15th

15TH FEASAC PROCEEDINGS

The Fifteenth FEA Student and Alumni Conference
AMERICAN UNIVERSITY OF BEIRUT
April 27–28, 2016
Beirut, Lebanon

ORGANIZING COMMITTEE 15TH FEASAC

FACULTY MEMBERS

Suidan Makram
Moukalled, Fadl [Chair]
Aboughali, Kamel
Ahmad, Mohammad N.
Al-Hindi, Mahmoud
Chabaan, Farid
Freije Makdessi, Mayda [Design]
Hamzeh, Farook
Mabsout, Mounir
Maddah, Bacel
Mahmoud, Samir
Moacdieh, Nadine Marie
Musfy, Leila
Mustapha, Samir
Nasser, Youssef [Editor]

STUDENT

Ajaeiya, Georgi [Website]
Tufenkji, Georges

FEA STAFF MEMBERS

Chamoun, Zeina
Itani, Maher [Web]
Khoury, Ibrahim
Kobeissi, Suzanne
Moufarrej, Nadia [Career Development Center]
Valerik, Vassili

GRAPHIC DESIGN

Deghayli, Karima [Website Design, Posters]
Karam, Sami [Logo, Flyer]
Kassarjian, Chantal [Banners]
Matar, Anastasia [Book]

WELCOME TO THE 15TH FEASAC

I would like to welcome you to the fifteenth annual FEA Student -and Alumni Conference. The format of this conference has been constantly evolving ever since Dean Ibrahim Hajj started this activity 15 years ago. This year is a special year for this great faculty and University. It is AUB's 150 birthday and because of this momentous occasion, we have expanded the conference and offered participation to faculty and students from universities other than AUB. The conference has eight half-day sessions each two running in parallel. In the sessions we have student speaker, both undergraduate and graduate students, and some of the students are from universities other than AUB. We have a distinguished alumnus and a distinguished young alumnus from the five departments that have been in existence for a while to have alumni. I hope that Chemical and Petroleum Engineering will start selecting their distinguished alumni soon. We also have a faculty member from each department making a presentation that covers their experience in academia, and for the first time, we have invited faculty members from outside AUB to give addresses about their experiences. We hope that with this format we have evolved this conference from an FEA Conference to one for all Engineers in Lebanon. ■ There will be a reception at the FEA Terrace on Thursday at 6:30 p.m. where monitory awards will be

given to the best student presentation and the best poster in each session. ■ This past Fall Semester FEA had 2,245 registered undergraduate students, 263 masters students, and 68 Ph.D. students. This year we established two new departments, the Department of Chemical and Petroleum Engineering, and the Department of Industrial Engineering and Management. We have also started a graduate only program in Biomedical Engineering, which admitted its first students this Spring Semester. We are awaiting approval from New York State of our new Program in Petroleum Engineering. ■ Now allow me to extend many thanks to the FEASAC organizing committee, the FEA students, the FEA faculty members, and the FEA staff members, for their dedicated work throughout the year to make this conference possible.

Makram T. Suidan, Dean

CONTENT

AUB FACULTY

DEPARTMENTS

FACULTY OF ENGINEERING AND ARCHITECTURE
LIFETIME ACHIEVEMENT AWARD

**BELOVED PROFESSOR, MENTOR AND FRIEND
WHO EXCELLED AS AN EDUCATOR AND A
PRACTITIONER**
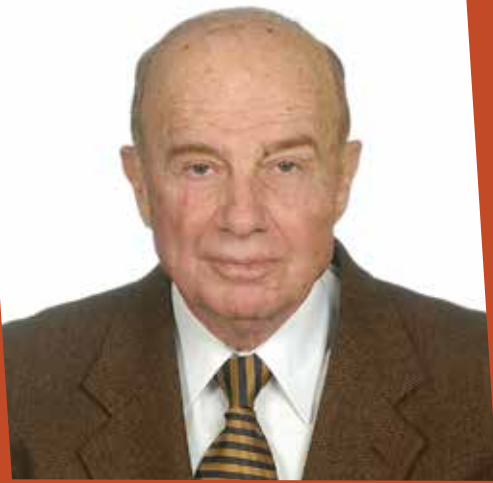
FATEH SAKKAL

BIOGRAPHY

- Born in Aleppo in 1926.
- Graduated from the University of Cairo with a B.Sc. degree in Mechanical Engineering (1949).
- Joined the Syrian University as a lecturer (1950-1954)
- Awarded a scholarship to do graduate work at the University of Manchester (UMIST) in 1954.
- Studied radiation measurement from curved surfaces at UMIST where he invented a special instrument for his research topic.
- Graduated from UMIST with a Ph.D. in Mechanical engineering (1956)
- Joined the G.N.Haden co. in London, working in the HVAC field (1956-1957) a while simultaneously attending the School of Jet Propulsion and Gas Turbine in Farnborough.
- Appointed as associate professor at the Syrian University (1957-1959)
- Joined the American University of Beirut's school of engineering as an assistant professor (1959).
- Promoted to full professor and chairman of the mechanical engineering department at A.U.B.
- Went on sabbatical as a visiting professor at the University of Minnesota (1968-1969) where he conducted research work on luminous flame measurement.
- Went on sabbatical as visiting professor at the University Of California, Berkley (1978-1979) where he studied the effect of cooling on sky radiation at the Lawrence Berkley Lab.
- Professional experience included HVAC, solar energy, and power diesel generating sets.
- Co-founded the Lebanese Society of Solar Energy and established a solar and weather station at A.U.B. that collected climatological data.
- Upon retirement in 1992, dedicated himself to charity and endowment work including establishing an award for the best renewable energy thesis in mechanical engineering at A.U.B.

RIAD AL SADIK

ABSTRACT

MEN TO MATCH MY MOUNTAINS, HOW TO MAKE DREAMS COME TRUE

Mr. Sadik's talk will address his journey through the various walks of life covering the pre-AUB/Faculty of Engineering days, and the early exodus of his family, then his AUB experience at the Faculty of Engineering, which he considers to be his "formative years" and thereafter, his journey in life following graduation from his beloved alma mater to complete his theoretical education with lessons from the "University of Life." ■ His address will cover the lessons learned along the way in the rich and complex walks of life. An expert in the world of construction, Mr. Sadik emphasizes the need to "construct" the human being and forge one's character as a well-rounded individual first and foremost. A life-long learner, who cherishes the importance of knowledge and culture in all their facets, Mr. Sadik ties the AUB Motto "That they may have life and have it more abundantly" to the opportunities opened to young graduates to develop a strong work ethic, and give back to the community.

BIOGRAPHY

Riad Tawfiq Al Sadik is a distinguished name in the Middle Eastern construction industry. A graduate of the American University of Beirut, Mr. Sadik, together with Mr. Khalaf Habtoor, founded Al Habtoor Engineering Enterprises in 1970. ■ From a workforce of less than 100 and two projects worth AED120,000, the company has grown to become one of the leading construction companies in the MENA region with a workforce in excess of 25,000 and with a turnover in billions of dollars. ■ The youngest son in a family of 4 brothers and 3 sisters, Mr. Sadik is widely traveled. Born in Palestine, his initial years were spent in Lebanon. After graduating he trained in the United Kingdom and then spent three years working, the last year of which he was Head of Engineering for the Eastern Region of Saudi Arabia. ■ In 1967 he moved to Dubai as the UAE became a key market in the Middle East. Mr. Sadik is well known for his philanthropic activities, and he holds a wider interest in humanities. He is an avid reader of contemporary literature and an associate of several literary associations and community service organisations.

■ Currently, Riad Sadik is the:
• Chairman of Habtoor Leighton Group
• Owner / Chairman of Construction Machinery Center (a co. specialized in Construction Machinery).
• Owner / Chairman of Dar Al Rokham (a co. specialized in Stone works).
• Owner of Grand Millennium Dubai, a five star hotel.
• Owner of Rosegrade Ltd. (a UK Services Co.).
• Owner of Gulf American, Los Angeles, U.S.A (Currently a Development Co.)
•   Owner of Business Bay Hotel.
•   Chairman of Kazax Minerals

Riad Sadik also holds membership of various boards including the Board of Trustees of the American University of Sharjah, the Board of Trustees of Birzeit University Palestine, the Board of Trustees of Bethlehem Development Board, and the Board of Trustees of the Institute for Palestine Studies. He is also Vice Chairman of the Board of Governor's of the Pearl Initiative.

BIOGRAPHY

Mr. Abbany is currently Director of Transportation Department at Dar Al-Handasah Consultants (Shair & Partner). He graduated from AUB in 1999 with a BE degree in Civil Engineering, and then earned, in 2002, his ME degree in Transportation, also from AUB. ■ Mr. Abbany joined Dar Al-Handasah immediately after his gradu-ation, and has since grown from junior engineer in his early years, all the way up the ranks until his election as partner and director of the firm in 2012. In his 16 years with the firm, he has moved between the various design offices in Beirut, Cairo, Riyadh and Pune. ■ During his career, Mr. Abbany has been involved in several major projects in the Middle East and Africa, mainly in the field of transport planning and the design of roads, airports and transit systems. He has lead the works of multi-disciplinary teams, and successfully managed many large-scale and fast-track projects such as the Riyadh Metro Project (Saudi Arabia), the Planning of Al Maktoum International Airport (UAE), the Automated People Mover System at Princess Noura Bint Abdulrahman University for Women (Saudi Arabia), and the Autoroute Est-Ouest - East Lot Project (Algeria).

ALI ABBANY

MANAGING LARGE-SCALE PROJECTS:
A JOURNEY OF SUCCESS

ABSTRACT

I joined Dar Al-Handasah Consultants (Shair & Partners) in July of 1999, just 3 days after my graduation ceremony at AUB, striving to build a long-term and stable career in Lebanon. Now after almost 17 years, I can confidently say that this "right" decision actually marked the beginning of a successful journey in DAR, which took me around 4 different offices and led me to the top ranks in the firm. ■ This journey was not only characterized by a lot of dedication, smart-hard work, and continuous improvement; it was also full of challenges and excitement, which are typically encountered in the engineering field, especially when working on large-scale and fast-track projects in the region. ■ Today in my presentation, I will briefly take you through my journey, giving insight on a number of selected noteworthy highway, airport and railway projects that I have managed, the challenges faced, and the way my involvement in these projects shaped my career path and helped me achieve my Vision of Success.

BIOGRAPHY

A graduate of AUB's Civil Engineering and Engineering Management Programs, as well as Harvard Business School and INSEAD managerial development programs, Mr. Marwan Abou Zeid has over 15 years of advisory and operational management experience in the Middle East. ■ Mr. Abou-Zeid is currently the Vice President of Strategy & Corporate Development at Alsaad Group, a Saudi-based conglomerate active in the contracting, trading, and manufacturing fields in the GCC and the Levant. As member of the group's senior executive team, he leads a wide number of initiatives of various types (strategic, commercial, operational, financial, and people-related) aimed at optimizing the performance of the group's operating entities. ■ Prior to joining Alsaad Group, Mr. Abou-Zeid was Vice President at CDS, a regional management consulting firm advising private sector organizations in the Arab World on strategy, organizational transformation, operational improvement, and human capital enhancement. He led teams on a multitude of consultancy assignments in the insurance, contracting, manufacturing, retail and distribution, and banking fields among others. Mr. Abou-Zeid resides in Lebanon with his wife and two sons.

MARWAN ABOU-ZEID

A CAREER IN EXTERNAL AND IN-HOUSE CONSULTING: LESSONS LEARNED

ABSTRACT

After obtaining my MEM from AUB, I started my career as a management consultant with CDS, a regional management advisory firm based in Beirut. Throughout the next 14 years, my engagements took me over 10 different countries in the MENA and GCC. While each engagement had its unique objectives and challenges, what was in common was how to deliver tangible and practical recommendations to client organizations, allowing them to implement the recommendations and build in-house capability to sustain them. Progressing from Business Analyst to Vice President, my role evolved from engagement delivery in early years, to leading consulting teams, and eventually to develop the firm's business and institutionalize its delivery mechanisms. ■ In October 2014, I made the transition from management consulting into a corporate role at Alsaad, a regional group with interests in contracting and related industries. As the group head of strategy group, I lead a variety of interactions with the management teams of all companies, with a primary objective to enhance the group's long term performance and reduce risk. A key element of my job is to ensure the companies have competent leadership and that they have the right culture and incentives in place. ■ My AUB education provided me a solid foundation for my career, with such skills as: problem solving, working in teams with people from different backgrounds and diverse views, business writing, delivering effective presentations, analysis and number crunching, etc. Most importantly perhaps was the good work ethic that was drilled into us by AUB professors.

BIOGRAPHY

Hadi El-Amine is a Ph.D candidate in Operations Research at Virginia Tech working with Drs. Ebru Bish and Douglas Bish. Before coming to the US he received his BE in Electrical and Computer Engineering (in 2009) and his ME in Engineering Management (in 2012) from the American University of Beirut. Hadi's research interest is the application of operations research techniques for improving healthcare operations, assisting the medical decision making process, and impacting policy. His current focus is on the formulation of robust blood screening strategies under limited information. Hadi is the recipient of the 2015 INFORMS Bonder Scholarship for Applied Operations Research in Health services.

HADI EL-AMINE

OPTIMAL POST-DONATION BLOOD SCREENING STRATEGIES UNDER UNCERTAINTY

ABSTRACT

Blood products are essential components of any healthcare system, and their safety, in terms of being free of transfusion-transmittable infections, is crucial. While the Food and Drug Administration (FDA) in the United States requires all blood donations to be tested for a set of infections, it does not dictate which particular tests should be used by Blood Collection Centers. Multiple FDA-licensed blood screening tests are available for each infection, but all screening tests are imperfectly reliable and have different costs. In addition, infection prevalence rates and other donor characteristics within the donor population are uncertain, while surveillance methods are highly resource- and time-intensive. Therefore, only limited information is available to Blood Collection Centers on infection prevalence rates and other donor characteristics. In this setting, the budget-constrained Blood Collection Center needs to devise a post-donation blood screening scheme so as to minimize the risk of an infectious donation being released into the blood supply. The focus is on "robust" screening schemes under limited information. Toward this goal, various objectives are considered, and structural properties of the optimal solutions under each objective are characterized. This allows to gain insight and to develop efficient, exact algorithms. My research shows that using the proposed optimization-based approaches provides robust solutions with significantly lower expected infection risk compared to other testing schemes that satisfy the FDA requirements. This has important public policy implications.

BIOGRAPHY

Nasser (M) Ghandour is Chairman and Chief Executive Officer of METS Holding, comprised of mechanical engineering; generator set assembly – supply – and maintenance, manufacturing, and a portfolio in real estate. ■ Born on June 19, 1965, Ghandour grew up with a remarkable interest for mechanics. During his childhood and at the early stages of his teenage years, his interest in fixing and assembling mechanical games grew into a passion. His determination to take part of the mechanical world led him as a young entrepreneur to establish his first company, METS (Mechanical Engineers, Technicians and Suppliers) during his days at the American University of Beirut. ■ Ghandour always believed that the dollar value of the project is temporary and what is the most valuable is a satisfied customer. With this vision, he was able to fill Baghdad's streets with generator sets that caught the eyes of the first Telecom network in central Iraq. This first stepping stone has resulted over the next five years in the distribution of the products in thirty- two countries, forty today. ■ During his EMBA studies at AUB (2009), Nasser launched the PowerMet network, METS Energy's own distribution and services network. Today PowerMet is operating in 13 countries with more than 500 employees in the Middle Eastern, Africa, and Asia. Ghandour was married to Suzanne Abou Zahr. The couple has three children. Ghandour is an avid reader, enjoys fishing, and listening to music.

NASSER GHANDOUR

A FLASH FORWARD TO WHAT LIFE AFTER AUB MIGHT BE:

ABSTRACT

What to expect after graduation?

Who are your future teachers?

What should be your guidelines?

Challenges and obstacles

Is it luck or opportunity?

How to keep growing?

Is AUB that good?

When can we rest?

BIOGRAPHY

Zahi Azouri is the Head of Corporate Development at FAMA Holdings (2012- Present), a Saudi Investment Company based in Riyadh, Saudi Arabia. He currently manages the monitoring of FAMA's six businesses, which are completely diversified across industries, geographies, value chain play and stage of development ([Technology, Hospitality, Steel Production, and Fruit and Vegetable farming], [Saudi Arabia, Bahrain and the United States], [Software Development, Farming, Production / Processing, Packaging, and Sales], [Greenfield, Developing and Mature Businesses and Organizations]). He also leads FAMA group's corporate initiatives (e.g., organizational restructuring, institutionalization of processes and governance) and assessment and execution of new investment opportunities. ■ Zahi is also a member of SAAC's Executive Committee, SAAC being one of FAMA's Green Field Projects he has lead from early feasibility stages to deal structuring, negotiation and closure, including EPC project oversight. (SAAC stands for Saudi Advanced Agriculture Company, founded in 2013, using advanced Hydroponics farming techniques). ■ Zahi's previous experience includes working as a Refrigeration Design & Contracts Engineer at YORK International in Dubai, UAE (2003-2005), as a Business Development Manager GCC & Egypt at Hansgrohe ME&A Ltd. based in Nicosia, Cyprus (2005-2008), and as a Lead Associate at Strategy & (Formerly Booz and Company) in the Financial Services practice based in Beirut, Lebanon (2010-2012). ■ Zahi holds an MBA from INSEAD (2009) and a Bachelor of Mechanical Engineering from the American University of Beirut (2003), where he received the Dean's Award for Creative Achievement on his final year project.

ZAHI AZOURI

ADDRESSED TO THE ENGINEERING STUDENTS

ABSTRACT

Dear Students,
I am looking forward to sharing with you some of the key learnings I have gathered during my twelve years of professional experience; I am sure they will answer many of the questions you would have. ■ First and foremost, you need to be relieved of the peer and societal pressure, by understanding that there is no perfect job or career, and that the ultimate goal, although difficult to reach, is happiness. In order to reach this happiness, you should manage your career; ensuring a proper education is a must but is only the start… the bulk of career management should take place at work. ■ At work, several challenges will have to be addressed. At junior levels, it is more about the job challenges themselves. The more senior you will get, managing people and their expectations becomes the key challenge, and this requires soft skills. ■ There are several ways to get to the senior management roles in any organization; we will touch upon the various possible paths for Engineers to get up there, and highlight the importance of an MBA as a complement to your academic acumen. You have to know that you are lucky to be an AUB student, because AUB will provide you with the needed so that you could navigate through the challenges and uncertainties of the professional world. ■ I will make sure to conclude my short presentation by demonstrating to you how the above mentioned insights were applicable to my own career progression to date.

BIOGRAPHY

Georges Makhoul is the CEO of Constellation Holdings, a Dubai-based private investment firm with assets in Energy, Hospitality, Aviation, Real Estate, and Financials. He is also a Partner in Alcazar Capital, a Dubai-based private equity firm with investments in Telecom, Renewable Energy, and Real Estate. He is a Director of multiple private and publicly-listed companies. ■ Until his Retirement from Morgan Stanley in January 2010, Mr.Makhoul was the President of the Firm in the broader Middle East and North Africa. He managed Morgan Stanley's regional headquarter in Dubai, and the other offices in Doha, Riyadh, Jeddah and Khobar. Mr. Makhoul spearheaded the firm's business in the region since its inception in 2005. He is an acknowledged franchise and business leader who counts among his advisory clients most of the Gulf governments, Sovereign Wealth Funds, Large MNCs and major private family investors. ■ Until 2005, Mr. Makhoul had been the Investment Banking Head for the European Consumer and Retail Industry, based in London. He had spent seven years in Tokyo leading the Telecom and Media Investment Banking group in Japan. Prior to Morgan Stanley Mr. Makhoul was the Managing Partner leading the Info-Coms and Entertainment consulting practice of PwC Japan. He had joined Price Waterhouse in the strategy consulting practice in New York in 1994. ■ Before PW, he led a National Science Foundation Research Center at Columbia University in NYC. He started his career with Strategic Operations and Management at Bellcore. ■ Mr. Makhoul holds a Ph.D. and M.S. degrees in Electrical Engineering from Syracuse University, and a B.E. in Electrical Engineering from the American University of Beirut. He is fluent in English, French and Arabic with capability of Japanese. He received his undergraduate and graduate education on full merit scholarships. Since 2005, he resides in Dubai with his family.

GEORGES A. MAKHOUL

SUCCESS AS OUTCOME OF NECESSITY

ABSTRACT

The speaker will share with the audience his experience at AUB and how the AUB experience helped him navigate through foundation-building life stages. He will describe how a career that is rooted in analytical training, transitioned to bear fruits in the world of business and finance. He will also argue for the necessity of increased emphasis on soft and diversified skills, in addition to rigorous analytical training.

BIOGRAPHY

Fabien Chraim earned a Bachelor's degree with honors in Electrical and Computer Engineering from the American University of Beirut in 2009. He obtained an M.S. in Civil Systems Engineering in 2010, then a PhD in Electrical Engineering and Computer Sciences in 2014 from the University of California, Berkeley. He was student researcher working with Prof. Kristofer S.J. Pister at the Berkeley Sensor and Actuator Center. He is the CTO and co-founder of Civil Maps. The company provides on-demand perception of the world, with a focus on automated driving applications.

FABIEN CHRAIM

ON-DEMAND PERCEPTION OF THE WORLD

ABSTRACT

For riders,manufacturers and infrastructure operators alike, the next leap in mobility is centered around the autonomous vehicle. The desire to automate transportation to this new level stems from a need to increase safety and efficiency, while reducing pollution and traffic. For this to become a reality, vehicles need to perceive the world at levels exceeding human capabilities. This talk explores novel methods for providing 3D semantic maps to autonomous vehicles at scale and with high accuracy. Centimeter-level localization of vehicles is also discussed. Challenges around data collection, processing and validation are presented.

IN RECOGNITION OF YOUR
PROFESSIONAL CONTRIBUTION TO THE
ADVANCEMENT OF KNOWLEDGE IN THE
FIELD OF CHEMICAL ENGINEERING AND
YOUR MENTORSHIP FOR WOMEN IN
ENGINEERING

BIOGRAPHY

Professor Khraisheh has a long record in environmental related research. She published upwards of 70 publications in this vital research areas. Since joining Qatar University, Dr Majeda has successfully demonstrated her ability to attract research funding from variety of sources and led a number of inter-nationally collaborative research projects related to gas hydrates control, CO_2, water management and ionic liquids. Professor Khraisheh is a renowned teacher winning the Qatar University outstanding Teaching award in addition to the highest university award (Merit Award) for excellence in research, service and teaching. She also won the Mentorship award for the leading Women in Science and Engineering in the middle east in 2013. She currently serves as the Director of the Qatar University Honors Program.

MAJEDA KHRAISHEH

HIGH PRESSURE CO2/N2 AND CO2/CH4 SEPARATION USING DENSE POLYSULFONE SUPPORTED IONIC LIQUID MEMBRANES (DPSILMS)

ABSTRACT

The main aim of this paper is to investigate the potential use of the synthesized DPSILMs in the industrial gas processing applications for high-pressure CO2 separation from N2 and CH4 streams with less or no loss of ILs. The synthesized DPSILMs were analysed using FTIR and SEM and showed a clear chemical and physical change in the structure PSF and well distribution of ILs in PSF. Binary mixtures of CO2/N2 and CO2/CH4 (5 mol% CO2) were used in the study. Selectivity values for the prepared DPSILMs were obtained using a high-pressure membrane unit obtained from Rubotherm Präzisionsmesstechnik GmbH apparatus (System 2). The highest CO2/N2 selectivity values were 36 for both PSF-0.5 wt% [DIP-C4mim][NTf2], PSF-25 wt% [N4441][formate], 29 and 21 for PSF-0.5 wt% [C4mim][NTf2] and PSF-50 wt% [P4441][formate] respectively. Whereas the highest CO2/CH4 selectivity results were 70, 63, 47, and 32 for PSF-2.5 wt% [C4mim][NTf2], PSF-2.5 wt% [DIP-C4mim][NTf2], PSF-0.5 wt% [N4441] [formate], and PSF-5 wt% [P4441][formate] respectively. Another system was used to measure the permeability of each gas to be plotted then on Robeson's upper bound (2008) with other PSF blends in the literature. The plot showed that the synthesized DPSILMs gave satisfying results and behave as well or better than different types of reported PSF blends. Stability measurements of DPSILMs were conducted regarding ILs loss and CO2/CH4 separation efficiency. DPSILMs with 5 wt% [P4441][formate] and [N4441][formate] showed about 30% and 20% ILs loss respectively at 10 bar after 12 hours with small reduction in CO2/CH4 selectivity; while no loss of [DIP-C4mim][NTf2] and [C4mim][NTf2] was observed.

BIOGRAPHY

Marcus Huebel supports as a Senior Advisor the top management of Chevron Phillips Chemical's business for polyolefins in Europe, Africa & Middle East. Focusing on commercial excellence and geographical growth, he is based out of Brussels, Belgium. Furthermore, he serves as Non-Executive Director at Kemgo, a buy sell platform for petrochemicals as well as on selected Advisory Boards, like that of the German-Arab Friendship Association. ■ Marcus Huebel joined Chevron Phillips Chemical from the management board of Muntajat, the Qatari petrochemicals marketing and distribution company, where he was in charge of corporate transformations, strategy and corporate communications. In his earlier career, he held senior responsibilities in global strategy development at SABIC in Riyadh, Saudi-Arabia, and led the implementation of transformational initiatives in their Saudi, Middle East and Africa Region. ■ Dr. Huebel gained a wealth of functional and geographical experience in the global chemical industry through various leadership roles in management consulting, initially with Arthur D. Little and later with Accenture, as a Partner, being based out of Germany and Shanghai, China. He started his career as a scientist with BASF in Germany. Dr. Huebel earned a Ph.D. in Chemistry, after studying at the universities in Bochum and Marburg, and at the Max Planck Institute for Coal Research, Germany. He conducted postdoctoral studies in Polymer Science & Technology at Durham University, UK. Dr. Huebel holds a Professional Director Certification from Hawkamah Institute of Directors in Dubai, UAE, accredited by the World Bank Group.

MARCUS HUEBEL

GLOBAL PETROCHEMICAL INDUSTRY AT CROSS ROADS

ABSTRACT

Moving into the second half of this decade, a host of trends materialize in shaping the global petrochemicals arena. On the feed stock side, there is next to the advent of shale gas primarily in the US, the drop in crude oil prices, lower and longer than anticipated. In the Middle East, a prime feed stock advantaged growth area in recent decades, buildup of capacity has slowed significantly, to a large extend driven by lack of additional gas as well as the petrochemical industry moves behind satisfying increased domestic energy needs. The US Gulf Coast witnesses a wave of capacity coming on stream these years, while Iran likely finds back into international business. In Saudi Arabia the feedstock pricing mechanisms has changed. All this has unfolded a significantly changed playing field, both from a geographical as well as a technological perspective. Flattening and shifting of the industry cost curve in polyethylene is a prime case in point. These shifts also affect competitiveness of aspired coal based petrochemicals in China, a country that has emerged as one of the most powerful demand centers, for polymers in particular. In times of a global economic slowdown, China's shift towards a more domestic demand driven economy goes in parallel with significant demographic changes – there, in neighboring economies, and in developed societies of North America and Western Europe. Disruptions of gradually shifting supply and demand balances by technical issues and Force Majeure, as they affected European polyolefins last year, are likely to remain a force.

BIOGRAPHY

Mona El Hallak Ghaibeh is a Lebanese architect and heritage preservation activist. She received her B. Arch from AUB in 1990 and her Master of Architecture from Syracuse University-Florence Program in 1994. She was an Associate at Rais and Jamal -Architects and Engineers until 2000 when she started her own practice and has many residential projects built and under construction in Beirut. ■ She is a member of APSAD "Association pour le Sauvegarde des Sites et Anciens Demeures au Liban", a founding member of "IRAB", an N.G.O. working on the preservation of the musical heritage of the Arab world, and a founding member of "ZAKIRA", an N.G.O. for promoting photography and its role in documenting and preserving Memory. ■ She led several heritage preservation campaigns and succeeded in the preservation of the "Barakat Building" in Sodeco. After 15 years of lobbying, the building was expropriated by the Municipality of Beirut to be rehabilitated into Beit Beirut : a museum of memory and a cultural urban center for Beirut. She is a member of the "Scientific Committee of Beit Beirut". ■ She is currently engaged in the "Civil Campaign to protect the Dalieh of Raouche" as well as many other campaigns to protect architectural and natural heritage sites around Lebanon threatened by real estate development, in an effort to keep them accessible to all as open shared public spaces. ■ In 2013, she was given the "Ordre National du Merite au grade de Chevalier" from the President of the French Republic in recognition of her work and achievements in preserving cultural heritage.

MONA EL HALLAK

AN ARCHITECT'S FIGHT FOR BEIRUT'S HERITAGE AND IDENTITY AGAINST AMNESIA

ABSTRACT

26 years ago when I graduated from AUB, I was dreaming to be an architect who builds projects in the city I loved. It was 1990: the war had ended and Beirut had started its reconstruction process in a frenzy of real estate speculation and sweeping post war amnesia. When history, identity and collective memory were weighed against money, money won. Buildings of unique architectural character were destroyed; intact heritage clusters documenting different periods in the city's history were fragmented by new developments; heritage buildings were dwarfed by modern towers and isolated from their traditional context: Beirut was irreversibly losing its identity. I found myself compelled to be also the architect who saves architectural and cultural urban heritage from greedy developers. ■ Of course, I managed to build many projects in Beirut along the way, got married and had a boy who loves to chase cats in AUB. ■ Hoping that it will be contagious, I will share with you my long fight starting with my success in saving the Barakat building, now becoming Beit Beirut, and ending with our campaign to save the Dalieh of Raouche: a fight that is still going strong to preserve the last remaining natural coastal site of the Ras Beirut headland, where the morphology is still almost intact in a city invaded by urbanization.

BIOGRAPHY

Khaled Malas is an architect from Damascus. He is also a doctoral candidate in Islamic Art and Architecture History at the Institute of Fine Arts, New York University. His current research interests center around representations of magic and other knowledge-based practices and technologies, whether in the medieval or modern Arab world. A firm believer in collaborative potential, most of his current non-academic work is conceived of and produced with a team that inevitably includes Salim al-Kadi (BARCH 2004), Alfred Tarazi (BGD 2004) and Jana Traboulsi (BGD 2000). ■ Khaled is a graduate of the American University of Beirut and Cornell University. He is a member of various associations including the Arab Image Foundation, the Historians of Islamic Art Association, and the London Institute of 'Pataphysics.

KHALED MALAS

SYRIAN MONUMENTS OF THE EVERYDAY

ABSTRACT

This presentation centers on two recent projects that examine alternative narratives of modernities and its creative resistances. These projects stem from a firm conviction that architecture can perform a transformative role that it is capable of transcending the oft-unquestioned distinctions between building and monument. The first project, 'excavating the sky', was commissioned by and exhibited at the 2014 Venice Architecture Biennale. 'excavating the sky' examined the role of heavy-mechanical flight, whether through bombing or observation, in producing the landscapes of Syria. Embedded into the earth, and defiantly challenging the violence that falls from above onto the surface, this research was coupled with the construction of two wells built with an activist group in rebel-controlled Deraa, Southern Syria. More than two years later, these wells still provide water for a community of 27,000 people. ■ The second project, 'current power in syria', is concerned with the role of electricity as an invisible force that is capable of producing new social relations, transforming the populace from subjects to citizens. These relations are manifested physically in space through infrastructures, appliances, and practitioners. In the besieged Eastern Ghouta region of Damascus, and in collaboration with a photojournalist and a blacksmith, a windmill to generate power for a public program was erected in June 2015. Representations of this project, including a brief alternative history of electricity in Syria, will be shown at the 6th Marrakech Biennale from February 24th-May 8th, 2016. ■ In Syria, where the land is the stake and the site of multiple cycles of excessive violence, the comprehension of past struggles in/of space and its meanings becomes essential. As such, these projects attempt to forge new ways of building in accordance with the circumstances of the Syrian context today.

MAZEN TABBARA [LAU]

BIOGRAPHY

Mazen Tabbara is currently an associate professor in the civil engineering department at the Lebanese American University (LAU). His three degrees are in civil engineering: BE at AUB (1981), MS at Stanford (1982), and PhD at Northwestern (1990). During his employment at LAU (starting in 1998) he served as chair of the department (2002-2004) and assistant dean of the school (2006-2011). ■ Employment before LAU included: Dar Al-Handasah Consultants (1982-1985), postdoctoral fellow at Northwestern (1991-1994), and senior member of technical staff at Sandia National Laboratories (Albuquerque, New Mexico, 1994-1998). ■ His research background is in computer modeling and simulation of physical engineering problems and he has worked on the following topics: quasi-static and dynamic fracture, strain localization, projectile penetration of targets, concrete columns wrapped with fiber reinforced plastic, structural analysis and design, material modeling of actively confined concrete under compression, post- earthquake survival of archeological monuments, flow of water over spillways, stability of earth slopes, and transport of air pollutants.

REFLECTIONS ON THE ROLE OF FACULTY @UNIVERSITY

ABSTRACT

As we move fast forward, we look back to the distance past and recall the slideRule, a landmark mechanical analog computer; a tool in the hands of the few abiding by its rigid rules to make a primitive computation. We look around us today and marvel at the iTool; a home button, a swipe and a tap and the world is at your finger tips in a flash. A tool in the hands of the many performing wonders, no rigid rules, no learning curves, and endless possibilities. A machine that equates between the ages, a machine that gives the younger generation the power to imagine the unknown. A machine that gives the young the opportunity to teach the old. An opportunity for the student to teach the teacher. ■ As we move fast forward, we ask: What has changed on the faculty front since the distance past? The 3Rule still dominates the lives of faculty since the slideRule. The old and famous triad of Teaching, Service, and Research. Teaching being an art as much as it is knowledge, Service being a calling more than a job, and Research being research in Lebanon. Yet, excellence in each dimension is a must! Academic Excellence comes to mind, counting the uncountable also comes to mind. How does Academic Excellence translate into an added value to the students? What are the indicators of success of our undergraduate students? The paradigm of thought has shifted and we as faculty have not (or have we?).

HADI ABOU CHAKRA [BAU]

BIOGRAPHY

Dr. Hadi Abou-Chakra has completed his PhD in Process Engineering in 1998 and postdoctoral studies from the University of Surrey, UK. Currently, he is the Chairman of the Department of Industrial Engineering and Engineering Management, and acting Chairman of Department of Petroleum Engineering at Beirut Arab University, Lebanon. ■ He has an international reputation for research projects through grants secured with a number of Lebanese and European organizations. He led several research projects in collaboration with different industries. Currently, he is managing two research groups at his department. These research groups are conducting various research projects in collaboration with different sectors of the Lebanese industries in the field of manufacturing processes and engineering management to design and implement processes that minimize numbers of unresolved engineering related problems. ■ He has published 43 research papers in reputed journals and conferences. He was selected as a member of the technical committee and paper reviewer of many conferences.

THE RESEARCH INTERESTS AT THE DEPARTMENT OF INDUSTRIAL ENGINEERING AND ENGINEERING MANAGEMENT IN THE FACULTY OF ENGINEERING AT BEIRUT ARAB UNIVERSITY

ABSTRACT

The IEEM department is involved in extending its research interests in manufacturing processes, material properties, project management advancement and supply chain management and design. Research in manufacturing processes involves the optimization of annual maintenance planning and scheduling, and the number of maintenance staff to be contracted in cement industry. As well as focusing on lean manufacturing in industry, machine condition monitoring and filtration systems. Current research in material properties concentrated on solutions for recycled municipal waste and fan palm natural fibres used in concrete industries, and on washing powders. Recent works in engineering management have investigated the practice of integrating PMI project management knowledge areas in the execution of projects in Lebanon. Furthermore, Research in supply chain management tackled the Bullwhip effect variance ratio to the market demand variance for aggregated orders.

BARBAR AKLE
[LAU]
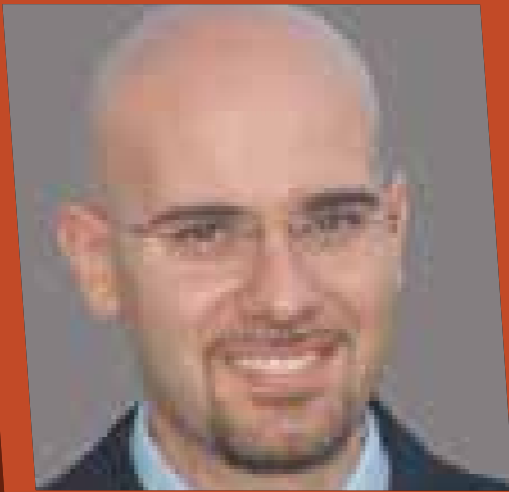
BIOGRAPHY

Barbar Akle holds a PhD in Mechanical Engineering from Virginia Tech and currently is an Associate Professor of Mechanical Engineering and the Assistant Dean of the School of Engineering at the Lebanese American University (LAU). His main research is focused on improving, modeling, and characterizing Electro-Active Polymers (EAP) that are also known as artificial muscles. His work led to better understanding of Ionic Polymer Metal Composites (IPMC) and to the development of applications such as robotic jellyfish, wall shear stress sensors, and currently working on developing Inner Ear Hair Cell sensors for the hearing disabled. Recently, he is leading a new research to develop a low cost and healthy exoskeleton system for the paralyzed. He has published more than 80 peer-reviewed international articles, and authored 5 US patents. Dr. Akle is an associate editor of the Journal of Intelligent Material Systems and Structures (JIMSS) and a member of the ASME Adaptive Structures and Material Systems Technical Committee.

OVERVIEW OF THE STATE OF UNDERSTANDING OF IONIC POLYMER METAL COMPOSITES AND ITS APPLICATION IN ROBOTIC JELLYFISH, BUOYANCY ENGINES, AND ARTIFICIAL EAR SENSORS

ABSTRACT

An Ionic Polymer Metal Composite (IPMC) is a smart material that bends due to the application of a small voltage (1V to 4V), and generates an electrical signal when it is mechanically deformed. IPMCs are also classified as Electro-Active Polymers (EAP) that are regarded as artificial muscles due to the similar force and displacement they generate when compared to human muscles. This seminar will discuss the state of understanding and modeling of the actuation and sensing mechanisms in IPMC. These efforts lead to significant improvements in the manufacturing and performance of this actuator. Finally, a snapshot of several applications of IPMCs will be presented; namely the biomemtic robotic jellyfish, the high efficiency buoyancy engine, and the undergoing work on hair cell like IPMC sensors that could be used in artificial inner ears.

RAYMOND
GHAJAR [LAU]

BIOGRAPHY

Dr. Raymond Ghajar is an Electrical engineer with over 30 years of experience in the area of power and energy systems; and energy economics. He spent the last 22 years at the Lebanese American University as a professor of electrical engineering, vice president for human resources and university services and now as associate dean of engineering. He worked for several electric utilities and consulting firms in Canada, United States and Lebanon. Dr. Ghajar published in internationally referred journals and conferences in the areas of power system reliability, electricity metering, renewable energy, and forecasting, and he organized several local and international events. Since 2008, Dr. Ghajar has also been working as senior energy advisor to the Minister of Energy and Water in Lebanon where he was instrumental in developing the current energy policy. Dr. Ghajar has been retained as an independent energy consultant for renewable energy projects in the Middle East and North Africa and he is on the Roster of Energy Experts of the World Bank.

BUILDING A POWER PLANT... WHAT DOES IT TAKE?

ABSTRACT

Building, owning and operating a power plant is a business that offers good returns on investment; however the financial investment is large and can be billions of dollars. It takes several years between the concept and the commencement of operation of the power plant with many pitfalls along the way – many projects fail at the concept stage after spending millions of dollars. In this presentation we will review the processes and decisions required to develop a power plant. One of the key decisions to make is which generation technology and which fuel are most appropriate for each situation. The presentation will address the key features of both thermal and renewable technologies, indicating the relative costs of each technology. ■ Understanding the project structure in terms of ownership, contracting and financing; the presentation will walk through the possible project structures and how a particular project structure is selected. The key components of the project cost structure and their contribution to the levelized cost will be presented. The power business is an international business and understanding the global power market is essential. The presentation will give an overview of the technology and fuel options for Lebanon and how these options are impacted by the changing oil prices.

AHMAD B. ALBADARIN

[LIMERICK]

BIOGRAPHY

Dr. Ahmad B. Albadarin is a Senior Research Fellow at the Solid State Pharmaceutical Centre (SSPC) at the University of Limerick. He is also a Visiting Researcher of Chemical Engineering at Queen's University Belfast (UK). ■ He has published over 45 international journal papers in the last 5 years (>425 citations, h-index 11). His research interests include particle/powder technology, environmental science and materials engineering. Papers are published on various aspects of chemical engineering and environmental science. ■ Dr Albadarin published the third most cited paper in Chemical Engineering Journal for the last three years (2014 AIChE Annual Meeting recognition), has a final stage under review project for the SFI Starting Investigator Research Grant (SIRG) Programme 2015 with total potential funding €400,000 and has been shortlisted for the Irish Laboratory Scientist of the year award 2016.

PHARMACEUTICAL PROCESS ENGINEERING

ABSTRACT

The Synthesis and Solid State Pharmaceutical Centre (SSPC), a Global Hub of Pharmaceutical Process Innovation and Advanced Manufacturing, funded by Science Foundation Ireland and industry, is a unique collaboration between 22 industry partners, 9 research performing organisations and 12 international academic collaborators. ■ The SSPC transcends company and academic boundaries and is the largest research collaboration in Ireland, and one of the largest globally, within the pharmaceutical area. The role of the SSPC is to link experienced scientists and engineers in academia and the pharmaceutical industry, to address critical research challenges. The SSPC leads the way for next generation drug manufacture and spans the entire pharmaceutical production chain from synthesis of the molecule, to the isolation of the material, and the formulation of the medicine.

ELIE HADDAD
[LAU]

BIOGRAPHY

Elie G. Haddad is a Professor of Architecture at the Lebanese American University, where he has been teaching since 1994. He was appointed Dean of the School of Architecture and Design in September 2012. ■ Dean Haddad holds a Ph.D. in Architecture from the University of Pennsylvania. He has taught several courses on the history and theory of modern architecture, which constitute the main areas of his research, and in which he has contributed several articles in international journals. He received two fellowships, in 2009 and 2012 to conduct research on modern architecture in Germany, specifically on the projects of social housing in Berlin and Frankfurt. ■ Haddad has co-edited a survey on architecture titled A Critical History of Contemporary Architecture 1960-2010, published by Ashgate in 2014. He also published a translation of Aldo Rossi's Scientific Autobiography, by Dar al Farabi (2011), and a collection of his articles on architecture, also by Dar Al Farabi (2014).

MOVING ACROSS CULTURES

ABSTRACT

In this presentation, I will trace my personal academic itinerary, highlighting the main points of an intellectual journey that I titled: Moving Across Cultures. ■ As implied in the title, I will emphasize the positive aspects of my cross-cultural education, which in the field of architecture, as well as in other fields, has become a necessity in an age of globalization. While this latter term carries some negative connotations worldwide, it is imperative to be aware and cognizant of other systems of exchange as the basis for establishing, or re-affirming, one's own traditions. As expected, this presentation will cover the various areas of research that I have embarked on over the past 20 years, culminating in my most recent interests.

SALAH SADEK

BIOGRAPHY

Dr. Sadek received his Bachelor's Degree in Civil and Environmental Engineering from the American University of Beirut, his MSc. in Civil/Geotechnical Engineering from Colorado State University and his PhD from the University of California, Berkeley. ■ Dr. Salah Sadek joined AUB in September 1993 as a full time faculty member. He served as Chairperson of the CEE Department from Oct 2008 to Sept 2011. ■ Dr. Sadek has accumulated over 20 years' experience in academia and associated teaching, research and administrative activities. He has acquired a wide and diverse experience as a consultant on a number of local, regional and international projects for private developers and governmental and international agencies. He is a founding member and current president of the Lebanese Geotechnical Engineering Society. ■ Dr. Sadek's on-going research interests are in the areas of Ground Improvement, Alternative Foundation Systems, Geotechnical Earthquake engineering, GIS Geo-Engineering applications (hazard assessment, decision-aid frameworks), Offshore Geo-Engineering and Energy Geotechnics.

**IN RECOGNITION OF YOUR DEVOTED SERVICES TO THE DEPARTMENT
OF CIVIL AND ENVIRONMENTAL ENGINEERING AND YOUR SCHOLARLY
CONTRIBUTION IN RESEARCH AND THE ADVANCEMENT OF KNOWLEDGE**

GEOTECHNICAL ENGINEERING... THE FOUNDATION OF ALL THINGS "CIVIL"

ABSTRACT

Geotechnical Engineering is arguably one of the earliest incarnations of what we now know as Civil Engineering. ■ Long before ours became a formal profession, "Engineers" have dealt with and tamed forces of Nature using available and native geo-materials, simple tools and skills. This knowledge honed and perfected over millennia is still relevant today and at the core of the field.■ In this talk we will briefly explore the "beginnings" of Geo-Engineering up to current practice and look ahead at future challenges and opportunities. A personal reflection on the research work conducted at AUB over the past twenty years is included in the context, along with a look at where we are headed from here and our role in contributing to the advance of the field and to local and regional practice.

BACEL MADDAH

Bacel Maddah is an associate professor in the Industrial Engineering and Management (IEM) Department at the American University of Beirut. He is also the chairperson of the IEM Department. Bacel holds a Ph.D. in Industrial and Systems Engineering from Virginia Tech, with a concentration in Operations Research. Bacel's research interests are in retailing, supply chain management, and stochastic processes. His recent interest is Financial Engineering that he teaches at AUB. His research was published in top-tier journals such as Management Science, Naval Research Logistics, IIE Transactions, and European Journal of Operational Research. He published 25 journal articles to date, along with several book chapters and conference proceedings. Bacel also has practical industry experience. He has worked as a senior business analyst with United Airlines and Hannaford Bros. He also provided limited consultation and training to large companies in the Middle East in areas closely-related to his interests.

**IN RECOGNITION OF YOUR SCHOLARLY CONTRIBUTION TO THE
DEPARTMENT OF INDUSTRIAL ENGINEERING AND MANAGEMENT
IN RESEARCH AND THE ADVANCEMENT OF KNOWLEDGE**

ROCKET SCIENCE RETAILING IS HERE TOO

ABSTRACT

The retail industry is witnessing fast growth. Concurrently, a large body of literature grounded in mathematical and statistical methods is produced to enable retail analytics. Some researchers call this "rocket science retailing". I describe my contributions to this area. ■ My early work is centered on developing analytical models that aid retailers in critical tactical decisions, on assortment, pricing, and inventory levels. One distinctive feature of this work is the demand function, which is consumer-centered, and based on the classical utility maximization principle. This captures realistic factors related to the dependency among products and categories. Part of my work considers optimizing retailing decision of a category of substitutable products, having stochastic demand, and sold over a short selling season, with a single opportunity to stock-up. This applies to fashion goods. My contributions here are mathematical results on the structure of optimal assortment and pricing, and on the interaction of pricing and inventory decisions. I also work on integrated inventory management and pricing schemes of "complementary" products. My recent work is along similar lines. However, I bring in new realistic ingredients and decisions. Specifically, I develop mathematical programming-based models that optimize large categories having hundreds of products with high computational efficiency. I also consider a different class of products, fast-moving-consumer-goods (e.g. grocery items), that are sold over long periods with several opportunities for replenishing inventory, and, price adjustment. ■ Finally, I analyze the space allocation decision by adopting a new holistic store-wide approach.

ISSAM LAKKIS

BIOGRAPHY

I graduated from AUB with a BE and ME in mechanical Engineering in 1991 and 1993 respectively. I joined the reacting gas dynamics lab at MIT in 1994 and earned my Ph.D. degree in mechanical engineering in 2000. From 2000 till 2002, I worked as a corporate engineer in Coventor (formerly Microcosm, developer of a CoventorWare for MEMS design and analysis). In 2003, I joined AUB as a visiting assistant professor. I was promoted to an associate professor in 2010. My research interests span multi-scale engineering applications in MEMS (RF Switches, Micro-mixers, Microfluidic Transistors), reduced-order modeling of microchannel flows, modeling of air and gases transport in the pulmonary system, grid-free computational methods for continuum (vortex methods) and non-continuum flows (DSMC), and more recently pollution transport and weather forecast.

**IN RECOGNITION OF YOUR SCHOLARLY CONTRIBUTION TO THE
DEPARTMENT OF MECHANICAL ENGINEERING IN RESEARCH AND
THE ADVANCEMENT OF KNOWLEDGE**

THE AUB PULMONARY MODEL - A TALE OF TWO FREQUENCIES

A B S T R A C T

Lung diseases (IRDS, lung obstructive diseases, apnea, lung interstitial emphysema, etc.) are a serious threat to human beings of all ages and typically lead to malfunctioning of the respiratory system that could cause death. The physical mechanisms behind clinical therapies that rely on various modalities of lung ventilation such as high frequency oscillatory ventilation (HFOV), are not yet fully understood. This is, in part, due to the limitations associated with in vivo measurements. ■ In this talk, I will introduce the AUB pulmonary model; a physically-based lung model that allows us to investigate the effect of lung abnormalities and treatment scenarios on exchange of gases with the blood. The model couples lung mechanics, gases transport in the airways and the alveoli, gases exchange with the blood by diffusion across the alveolar membrane, in addition to the the blood dissociation curves that characterize the O_2 and CO_2 binding properties of hemoglobin. The model also incorporates many previous measurements and correlations that allow estimation of lung parameters over wide ranges of height, age, gender, and body weight; rendering it applicable to a wide variety of individuals. ■ Unlike lumped models in literature, the model takes into account the dynamics of the spatial distribution of respiratory gases in the airways and within individual alveoli, which leads to more accurate prediction of gases transport. The model will be showcased for HFOV treatment of a preterm infant with IRDS. We will show that the model uncovers a characteristic frequency that, in many cases, limits the gas exchange, even when operating at the resonance frequency of the LC circuit modeling the air flow dependence on the pressure.

RABIH JABR

BIOGRAPHY

Rabih Jabr received the B.E. degree in electrical engineering (with high distinction) from the American University of Beirut in 1997, and the Ph.D. degree in electrical engineering from Imperial College London in 2000. Currently, he is a Professor in the Department of Electrical and Computer Engineering at the American University of Beirut. His research interests are in power system analysis and optimization. Rabih Jabr is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE), and he serves as an editor for the IEEE Transactions on Power Systems and the IEEE Transactions on Sustainable Energy.

IN RECOGNITION OF YOUR EXEMPLARY PROFESSIONALISM AND YOUR OUTSTANDING SCHOLARLY CONTRIBUTION TO THE DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

OPTIMIZATION IN POWER NETWORKS

ABSTRACT

The optimal power flow (OPF) is an optimization problem that seeks to minimize an objective function while satisfying physical and technical constraints on the power network; it is non-convex and generally hard to solve. This talk discusses a sufficient condition under which a semi-definite programming (SDP) solution can be used to recover the global optimum of the OPF problem; it then introduces the practical application of SDP on large-scale OPF problems by exploiting the sparsity of data matrices through matrix completion. A relation is established between the SDP solution and the second-order cone programming (SOCP) relaxation of acyclic distribution networks, and the use of SOCP for finding a global solution of the optimal feeder reconfiguration problem is presented.

MOHAMMAD AHMAD

BIOGRAPHY

Dr. Mohammad Ahmad is a Professor in Chemical Engineering and the Chairperson of the Department of Chemical and Petroleum Engineering at The American University Beirut (AUB). He is also the Chair of the Research Committee at FEA. Dr. Ahmad's research areas include: solid waste management; biofuels production from solid waste, production of activated carbons from solid waste; Heterogeneous catalysis and simulation and modeling. Dr. Ahmad supervised more than 25 PhD students and over 60 MSc and MPhil research projects. He has also been external and internal examiner for a large number of PhD students. He has given plenary talks and presentations at International Conferences Worldwide. Dr. Ahmad is acting as referee for a number of high impact International journals in areas related to his research. These include Industrial Engineering Chemistry Research, Biofuels, Renewable Energy, Bioresource Technology, Chemical Engineering journal Biotechnology Progress. He has also been acting as a referee to international research funding bodies in Canada, Saudi Arabia, Oman, Qatar and the UK. Dr. Ahmad has published over 60 papers in high impact journals.

IN RECOGNITION OF YOUR SCHOLARLY CONTRIBUTION TO THE DEPARTMENT OF CHEMICAL AND PETROLEUM ENGINEERING IN RESEARCH AND THE ADVANCEMENT OF KNOWLEDGE

BIOMASS TO BIOFUELS: THE DIFFERENT OPTIONS

ABSTRACT

In this presentation the different options of converting municipal solid waste (MSW) into biofuels and high value chemicals will be highlighted. The concept of the biorefinery will be considered in terms of pre-treatment, hydrolysis, and pyrolysis, aqueous phase reforming and dry reforming using heterogeneous catalysis. The yields of products as a function operating conditions such as temperature, catalysts loading and pretreatment will be determined and optimized. The results will also be modelled through lumping of the parameters in the liquid and gas phase. It is envisaged that this work will help in identifying solutions to the ongoing waste problem in the country.

KARIM NAJJAR

Associate Professor and Coordinator, Architecture Program, Department of Architecture and Design. ■ Dr. Najjar's research interest is in high-performing, responsive structures, beginning in the field of maritime and naval architecture, and currently shifting into climate-responsive design strategies. A large part of his research is dedicated to Responsive (Kinetic) Architecture. His research perspective is to continue to build inter-disciplinary collaborations with other engineering disciplines and with various industries. He synergistically combines research with design practice and teaching. His experimental design and professional work have achieved international recognition. ■ Najjar received his Diplom-Ingenieur (Master of Architecture) from the Vienna University of Technology.

FOR HIS INNOVATIVE AND CREATIVE INPUT, IN CHANGING AND RETHINKING ARCHITECTURE

RESPONSIVE STRUCTURES

ABSTRACT

Since the early 20th century artists and architects have introduced concepts of open and alterable artworks responding to the emergence of the machine age. In the first part, the presentation discusses reasons for this aspiration, through precedence. The second part introduces the concept towards a "dynamic architecture".

DEPARTMENT OF ARCHITECTURE AND DESIGN

CONTENT

A Synesthetic Graphic Design
exhibition design on the senses

Lama el Charif

Architecture and Design Department

American University of Beirut

Bliss Street

lge06@mail.aub.edu

Abstract

Whether it is the scientific revolution, or the Age of Enlightenment, or both that encouraged a rational interaction with the world, the world has faced periods showcasing a dismissal of the senses. Clear and functional thought of reason was contrasted with the unreliability of the body and the senses as sources of knowledge promoting the rationalization of nature. In the early half of the nineteenth century, the Romantics started the Counter Enlightnment that was interested in imagination and perceptual modes of engaging and experiencing the world. While advancement in technology after the industrial revolution and computer age brought along a sort of alienation of humans with nature, many thinkers including artists, philosophers, authors, poets, etc. encouraged a renewed interest in the senses. The establishment of the school of phenomenology, that re-established the primacy of perceptual experience as a reliable source of knowledge. In his book "The Eyes of the Skin", Juhani Pallasmaa writes about the weakness of taking into account the sense of vision alone, or even to depend on the mind and the neglect of the body. Accepting Pallasmaa's theory and debates of Maurice Merleau-Ponty and Martin Heidegger, I celebrate the cooperation of vision and each other sense and suggest an approach called "Synesthetic Design".

Through examining design works of James Goggin, Yves Kleine, and spacial works of Olafur Eliasson, etc, I suggest a definition of a design that refuses the hegemony of vision, a design that sprouts from synesthesia. The question I raise is one of long contemporary debates; how to not become estranged with our own body translated in a spacial design for an exhibitve context, and in a two dimentional visual.

Introduction

"Our body is both an object among objects and that which sees and touches them. There is an osmotic relation between the self and the world – they interpenetrate and mutually define each other." - Maurice Merleau-Ponty

My research identifies the human body as a unit that perceives. Processed information is conceived as a whole into this unit, not in parts of which some are received by the eyes, some by the nose. Through this research I aspire to uncover a design language that challenges the body not to perceive a visual element as pure visual, but as an experiential situation that communicates to its other senses. A language that although understands itself in the visual field, but refuses ocularcentrism. I have to elaborate on these theories in a context of exhibitions also tackling the basic senses. Through a visual journey, an exhibitory space could complete the task of challenging the human perceptual capacities to converge in moments of synesthesia. Through this project, I present a bodily experience for the audience of a communicative design instead of an encounter or understanding of visuals. I introduce "synesthetic graphic design".

Subjects will be invited to exist through their bodies and its involvement in the design for a series of experiential exhibitions and the space itself, and not as analytical minds with historical knowledge or analytical thoughts. The visitors are invited to thus not be estranged with their own bodies, a debate that has been raised between contemporary philosophers, and often in contemporary works. Hence I suggest a graphic design that joins on this philosophical debate.

Many visual artists have taken the position of translating moments of synesthesia, and many of suggesting a creation of it. The research discusses their works and other relevant cases. The question is how to design the space and build a graphic language for the body, and not for the eyes, by introducing a synesthetic graphic design.

I. THE SENSE & THE SHIFTS OF ITS IMPORTANCE

In the following chapter we follow the development of the study of the senses and how the different senses were represented in visual art through history, keeping in mind the growth of the dominance of vision over the other senses, and the rivaling of such growth.

A. History of the senses in visual art

Perceptual capacities have been the core of many theoretical, scientific, and metaphysical theories and discussions spread over centuries and cultures, concretely since Aristotle. Delving in an investigation of what the nature of the human soul is, Aristotle wrote "De Anima", a study of living things around 335 to 323 BC. Through his text, he suggested a more defined study of human perception, and its narrowing to five basic senses: sight, tactility, olfaction, gustation, and audio.

Employing an empirical tone and trusting that all knowledge comes from perception, he shaped an understanding that reigned over the many following centuries. (Fowler, 1995)

a. The Middle Ages

The theme of the senses in art, although recently in a boom of experimentation, was not born recently. Until the Romanesque art it had many scattered and shy occurrences. During the thirteenth century, in the early Middle Ages, senses in art started to make a more noticeable appearance however in sorts of uniform images. Sight and sound though stood in superior communal appreciation. Two ways of representations were developed to depict the senses in art. One depiction was based on a belief suggested by the early Gothic encyclopaedists that senses are more developed in certain animals than in mankind. It was only logical for artists to then depict a certain animal to
indicate each specific sense.

The other, based on Aristotle's "Parva Naturalia", was more considerate towards the imagery of Man in depicting his senses. Aristotle suggested thus to depict each sense as a human figure along a certain prop; mirror for sight, a musical instrument for sound, a flower for smell, a fruit for taste and a harp for touch since the French verb of playing of this instrument is "toucher". This iconographic depiction of the senses was only used in secular art, very popularly in fact. However ecclesiastical art, which was way more abundant, almost never adopted the theme of the five senses, although the theme reoccurred frequently in commentaries on the Bible, or sermons, or Christian educational scenarios, etc. (Nordenfalk, 1985)

b. Renaissance

The Renaissance, although preserved the ways, objectives and relations of the Middle Ages when depicting the senses, held changes in details. Sight was soloed in a hegemonic importance, abandoning sound in a less considered ground. Imagery of the Five Senses continued even then to have a sort of connection with love, similarly to medieval tradition for instance. Also they continued to be considered sometimes instruments of sins, particularly lust. This attitude was possibly why they were not referred to in religious art since the Middle Ages. Since the origin of the five senses was written in Latin, and the Latin nouns for each of these senses is masculine, it was almost inevitable to originally depict these senses in masculine bodies. However, in the sixteenth century, sensuality was appropriated to women figures. It was then that a sudden switch of the sex of the depiction took place in artistic depictions, and preserved for hundreds of years after.

Also with the invention of printing in the Renaissance, like woodcut printing representations of the theme differed from those of the Middle Ages in technique. The representation however stayed of a metaphoric visual symbol to hint to a sense. (Nordenfalk, 1985). In the early half of the 17th century, Descartian philosophy discussed the relativity of perception. Descartes did not believe that truth is revealed through human perception, on the contrary, he mistrusts what the senses transfer to the brain as information. Even visuals can be remembered or imagined, stimulated purely by the mind. He concludes that one has no absolute certainty but in one's own existence. Perceptions of the world, as nourishing as they are, could be deceiving. The warmth of a fire could be encountered in a dream of illusion, in a state of hallucination or unconsciousness. Descartes encourages doubt and thinking above all, "I think, therefore I am". However, he did consider perception part of the process of thinking, and he argues that it relies on the mind, and not on the body or physical organs. Although his questioning of alleged certainties, he emphasizes through his theories a strong network between the sciences, reason and philosophy in understanding the world. Perception is thus a mean for the mind, but also a trick of it. *(Marsh, 1988)*

c. Age of Enlightenment, Scientific Revolution

In the later half of the 17th and the 18th century, and in a time when the power in Europe belonged to the Catholic Church, an opposition rose against it. This opposition urged to counter every ideology, logic or process of thinking that the church enforced in its reign. As a contradiction to those, both the scientific revolution and the Age of Enlightenment, which interconnected, rose to dismiss subjectivity in obtaining truth, and accentuated a logic that covers objective physical scientific facts exclusively. Sound was nothing but the receiving of vibrations propagated through a body of either water or air, transferred as mechanical waves of displacement and pressure. Vision was nothing but reflected color rays, etc. This mindset carried the dismissal of the senses, and the encouragement of scientific studies. The beauty of nature is its components that consist of organisms of cells, elements of atoms, etc. It is what we have proven about it through case studies and analysis. Appreciation of beauty came from the understanding of the physical components and chemical reactions revolving around the subject.

However, art did have a very essential role in the movement, and so did the senses. To Enlightenment theorists, the celebration of life revolved around appreciating the rational order of nature. In the early Enlightenment, philosophers, especially French contributors to the movement, stressed on differentiating the objective realization of the latter, and a subjective sensual beauty. "The beautiful is the true". Art shall thus imitate nature in its order and rationality to achieve beauty. However, later in the movement, German theorists suggested that the artist himself had

a more elaborated task. Since the rational order of nature was only conceived through human's perception, an affinity to the senses materialized. German philosopher Christian Wolff discusses this relationship as a systematic beauty. He then founds a new term for this "science", and derives it from the Greek word for "senses": Aesthetics. For him, Aesthetics, or the science of the beautiful is the science of what is or can be perceived, naturally through the senses. We are only exposed to aesthetic pleasure, which is a certain national order, through human perception. Thus the senses are central in this age of Enlightenment, however not to be mixed up with subjectivity. Aesthetic pleasure was still a child of an objective realization of harmony and similar objective features of the world that are perceived even implicitly.

Artists of the century had a responsibility to propagate truth towards the people. Denis Diderot wrote "to make virtue attractive, vice odious, ridicule forceful; that is the aim of every honest man who takes up the pen, the brush or the chisel". An admiration for this truth and the beauty of it birthed an admiration for depictions of new machinery, scenes of experiments, or phenomena related to a technological or scientific advancement became more popular, a theme initiated by Joseph Wright of Derby. Just so, a rise of an appeal to allegories, historical paintings and art replaced the studies of representation of subjective sensual experiences. Wright in particular was fascinated with light contrast, the strength of shadows and darkness. *(Khan Academy)*

d. Romanticism

The Romantics rose in the early half of the 19th century and refused this objectivity towards and rationalization of nature. They countered the theories of Enlightenment with literature, art, and intellect emphasizing intense emotions, and deep appreciation of the beauty of nature, making way for its mystery and wonder. Nature is not the cells, not the atoms. Nature is the thunder that cries or the storm that threatens, it is the soil that springs and feeds. One does not understand nature counting its leafs and measuring its growth, however by vanishing the distance between one and nature and finding a sense of oneself in nature and a sense of nature in oneself. The romantics appreciated what is seen, what is heard, what is smelled, what is tasted and what is felt. Being a part of the mystery of the world, humans not only take in from nature or give to it, but also exist within it. They founded, or revived in a way, the concept of "genius"; the artist who creates beauty out of nothingness. The strong stances of the romantics were a universal wake up call about the accreditation of the body in contrast of the scientific mind. *(Bristow, 2010)*

The twentieth century brought along a considerable number of chronological art and design movements, and a general awakening about social, humanistic and even cultural and artistic unanswered questions lead to many investigations. Between them is the shift of the hegemony of vision and giving power to the body of the artist and his audience. In the 19th and early 20th century, blurring the separation between music and other forms of art became a "widespread obsession".

e. Postimpressionism

In the late 19th century, artists began to retranslate their interest in what they "saw". Postimpressionism starts to grow an interest in an unfaithful visual description of the object of the painting. However it achieves a faithfulness to the personal perceptive image that the object in subject draws in the artist. The eyes are no longer the windows for the perfect nature that an artist mimics and adores, but a part of an organism, himself being part of a natural phenomenon, embracing the imprint of nature on nature. Still in an era of a focus on sight, the art of the impressionist world starts to acknowledge the contribution of touch, smell, sound and taste in our encounters with the world. The human body stops being

a transparent tool for a mind that strives to understand nature, and becomes a rather complex part of nature that interprets and misinterprets, combines and divides its perception. Paul Cézanne conducted a considerable amount of research on the evolution of his own perception of the same scenery with different time lapses. He believed that "to paint is to register these color sensations". Both Cézanne and Pierre-Auguste Renoir were diagnosed to be nearsighted (myopes), the latter being known for stepping away from his own painting to experience it in his blurriness, allowing his eyes to exercise their personalization. Even when Cézanne was offered spectacles he is said to have screamed in rage: "take away those vulgar things!". Lucky for him, he also got other retinal disturbances with time causing color altering, which also he celebrated in series of paintings. According to Merleau-Ponty, Cézanne's paintings makes "visible how the world touches us". Postimpressionists not only noticed but also rejoiced the faultiness of the eye, opening the stage for later movements to engage in subjective bodily impressions. *(Crowther, 1933)*

f. Exposition Universelle in 1900 or World Fair

The 1900 introduced a brick stone in the evolvement of the organization of international exhibitions. The idea aimed at transforming exhibitions as specialist fairs that present industrial innovations before a professional audience into recreational spaces for the amusement of much wider sectors of the population. The exhibition was made as a celebration of what has been achieved in the art world, and an anticipation to what will be. Being in the heart of an industrial revolution, the public was communally acknowledged as being interested of technological advancements. The exhibition was recognized to have introduced the concept of "clou"; attractive looking objects that capture the imagination of the public and appeal to the human's soul. An impressive section of the fair used this craze into its tackling of human perception. Technologies were introduced as "clou" that would stimulate sensorial fields in what Vanessa R. Schwartz called a true fin-de-siècle.

One of the technologies was introduced by painter Hudo d'Alésie as a revival of an idea written by the artist Robert Barker in 1787. The Maréorama simulated a sea journey through the Mediteranian Sea, activating all five senses of the audience suitably. A canvas screen set up in a perimeter around the audience incorporated motion, light and visuals, sounds, certain smells and an adjacent restaurant to engage the audience in a full perceptual manipulated experience. D'Alésie's objective was to "make the spectators feel the same sensations that they would feel in the real place". "It is a holistic work of art" he said, indicating to how it tackles the human body as a whole. (Barbosa, 2015)

This approach of incorporating the idea of recognizing the body as a whole perceptual unit, in a celebration of the future and technological advancements, comprehends the beginning of the 1900th century thinking as with a revolutionary interest in the senses.

g. Cubism

The break free of the impressionist movement gave the floor for more independence for the artists of the following decades, who normally took advantage. Pablo Picasso was an essential figure for the cubist movement. The styles of painting he proposed issued many speculations on the reasoning behind his imagery. A theory suggested that Picasso would gradually abandon painting what the eye sees, to progress to paint what his hands touched. A cubist painting was interpreted as both a unification and a division of the perception of a three dimensional figure. After perceiving different angles of the subject, each detail was painted from a different angle to the taste of the painter. Painting was no longer of what is known, but what is seen, touched and selected.

h. Futurism

Futurism, started by Italian writer Filippo Tommaso Marinetti around 1909 as part of the avant-garde movement, was fascinated by the technology, the machinery and the urban life the world had come to. Futurist artists had an urge to propagate and advocate for the power and the beauty of speed, loudness, and quick change in the urban city. Paintings of the movement aimed for thus evoking all sensual experience with the visual not being the highest emphasis. This thrill about this perfect city forming drew them to represent its smells, heat, fast pace, loudness, etc, through the use of onomatopea, a hopeful imagination, a complete freedom in their compositions, and through creating a certain non rhythmic music to mimic the art of noises. The overwhelming power of the new machinery body was a celebratory consequence, especially the change of the urban city, etc. *(The Art Story)*

i. Op Art

Optical Art suggested a challenge for the equation of the eye and the mind coming to conclusions. Instead, experimentations of how an eye that abandons the mind would arrive to a much more desirable conclusion than if in collaboration with a reasoning. Bridget Riley's first op art work was "Kiss" in 1961. An experimentation for a communication with the human body, the painting presents two black masses on the very edge of touching, suggesting a result that is only completed in the viewer's mind. The fact of the touch being foreseen instead of visually presented, engages the audience's perception in creating the image. "The eye is the organ of distance and separation, whereas touch is the sense of nearness, intimacy and affection. The eye surveys, controls and investigates, whereas touch approaches and caresses," writes Juhanni Pallasma in his book "Eyes of the Skin". In her painting, Bridget Riley unveils an intimacy through two non-figurative elements that anticipate a caress.

Op art also simulates movement in a still. "The painting presents our eyes with contradictory data as we read part of the field in terms of diagonals and other parts in terms of horizontals and verticals. The painting practically forces us to move backwards and forwards and as we do so, the field appears to move, expanding, contracting and undulating," says Victor Vasarely, father of Op art. It is the eye only that then creates the perception of a movement in the mind of the viewer. "Movement does not rely on composition nor a specific subject, but on the apprehension of the act of looking which by itself is considered as the only creator", he continues. To create this movement, several techniques were studied. Op artists retrieved common characters of vision and movement like repetition, alternation and, to borrow from the auditory vocab, rhythm.

In 1967, Bridget Riley painted her first colored composition, Cataract 3. "The music of colors is what I want", she said. Already having caught the eye, Riley painted for the eyes to hear. The moving calm composition was dipped in a mixture of cold colors with a warm tint. The impact sang rhythm and harmony.

Optical art transitioned thus from being a manipulator of the eye, to an engagement with movement, to a visually narrated melody.

j. Contemporary

The next period until contemporary (new art*) records a huge variety and experimentations with the proliferation of the subject of the senses in art and design. Therefore, the documentation of the articulation of sensorial impressions stops here.

*New art, indicating to art in the current and a few decades back – the 20th century and back), where no specific school of art to constrict or directs artists.

II. A GRAPHIC DESIGN THAT FIGHTS OCULARCENTRISM

Centuries engaged in continual discussions around the perceptual capacities of a human being. The tools of perception were studied

in fields of science, and its effect on the person in the field of philosophy. This segregation gave a stereotype of dreaminess to philosophers, and one of detachment to scientists, two positions that were very often occupied by the same people in the renaissance and way before it. Where knowledge of the body occurred, knowledge of the soul proliferated. The Pythagoreans believed that "The eyes are made for astronomy, the ears for harmony and these are sister sciences".

A. Ocularcentrism

The 20th century brought along movements of rejection the hegemony of vision that had been surrendered to and admired since the renaissance. Rene Descartes often acknowledged the nobility of vision and its objectivity, and he equated it to tactility, the latter being "more certain and less vulnerable to error than vision". A theory started to form realizing the resistance of the vision's supremacy of the movements of its time, and put forth and theorized the contradiction, sometimes even aggressively. Whether objective or not, the process of internalizing the world is not reduced to a reception of light and creating mental images. If the whole body is involved then there should be a more diversified impression. Friedrich Nietzsche shamed this ocularcentrist attitude to be "treacherous and blind hostility towards the senses". Max Scheler blames it to be riveted by a "hatred of the body". This refusal propagated especially in French intellectual tradition like Maurice Merleau-Ponty, Rolland Barthes, Jacques Derrida, etc, according to Martin Jay's book "Downcast Eyes – The Denigration of Vision in Twentieth Century French Thought". Pallasmaa writes in the introduction of "Eyes of the Skin" "Significant architecture makes us experience ourselves as complete embodied and spiritual beings. In fact, this is the great function of all meaningful art."
History does not surrender as a whole to the supremacy of vision. Up until the Middle Ages sound was up to par with vision. The visual field however has retained its rank as opposed to sound. With this evolution, a collective development of the sense has occurred. Vision, because of its explicit usage, has had the most explicit understanding between the senses. However, the body acts on an experience as a whole, vision being only inclusive to the other senses. What is suggested is that then, this understanding be pushed to the viewer, for the latter to reach this relationship through the vocabulary of the suggested exhibitions of visual art. The twentieth century sprouted with epiphanies on the importance of celebrating the human as a perceptive unit, not as a mechanism. Since then, a human's sensorial capacities have been in constant questioning and under experimentations in philosophy, and in art. My project seeks to bring this debate to design.
The research seeks a stance of opposition to ocularcentrism through a proposition of visual and spacial argumentation rather than a debate. Although through a display of visual elements, the purpose is to direct the body to engage as a whole to have a full experience. It is to challenge the capacity of visual stimulators to intrigue the network of the other senses. It is to stimulate the body to hear, touch, smell and taste as the eye sees. *(Pallasmaa, 1996)*

B. Refusal of Occulacentrism

a. Through Phenomenology

The early 20th century brought a discipline that develops and embraces these theories regarding perception; phenomenology. The latter was founded by Edmund Husserl and involved parts of many disciplines in philosophy like logic, ontology and ethics as well as sciences such as neurology. It naturally dealt therefore with the structures of consciousness and the conditions of being. Under the new umbrella, philosophers like Martin Heidegger, Maurice Merleau-Ponty, Jean-Paul Sartre, Max Scheler carried through, discussed and initiated new ideas and studies.
In the wave of technological and scientific discoveries, philosophy revolving around the human soul, grew untrusted unless it covered a scientific dimension, hence the birth of phenomenology. It is in fact this discipline that gave this phenomenon the name of "Ocularcentrism". Juhanni Pallasmaa, Finnish architect, holds on to the philosophy of beauty. In "The Eyes of the Skin", he analyses the engagement of the human body, mainly through tactility, with the different objects that it encounters, realizing what draws a conclusion of beauty. His expertise in architecture and architectural theory directs him to analyze raw materials; wood, stone in contrast to steel, spatial characteristics, and human reactions upon encounters of these spaces. He links throughout his discussions these components to suggestions of an objective, in contrast to subjective, sense of beauty. The concept of relativity of beauty, feelings, and sensations has been established for a too long time. Phenomenology stands to wonder with our more or less similar organisms, how relative a feeling can be. Except in special circumstances, humans will always be intimate by touch, nostalgic by smells, or even angered by a storm. Neurology of the 20th and 21st centuries joins civilizations in history on a continuous pursuit of beauty, which evolved from perfection, proportions and emotions. It is relying on its concepts that this research develops its aim. A journey where the mind travels finds questions and answers. A journey where the body travels finds pleasures and relationships. The research longs for a journey that bridges both, finding the mind where the body touches, hears, tastes and smells what it can and cannot see. It urges for a journey that does not exclude the mind of the body and the body of the mind. The research understands thus that in this sort of inclusion, the body cannot be considered a partitioned entity with dissectible sensation. It is almost as if it is only through being one with the mind, that the body itself is a unit. Merleau-Ponty says "My perception is [therefore] not a sum of visual, tactile and auditory givens: I perceive in a total way with my whole being: I grasp a unique structure of the thing, a unique way of being, which speaks to all my senses at once". Pallasmaa revises the multisensory perception of encounters that might be hastily described as visual or tactile. Focusing on tactility the most, he reveals chapter by chapter the relevance of a human touch and perception in understanding the world. He aims then "to create a conceptual short circuit between the dominant sense of vision and the suppressed sense modality of touch". (Pallasma, 1996)

b. Through the study of the frame & peripheral vision

Renaissance artists defined their paintings to be a window to another world, a world of mystery, beauty, power, or emotions, etc. The paintings consequently create a space for the viewer to look at, projecting a space. Windows do not exist however without a distance and a separating wall. The frame of the paintings frames thus a world that we can only observe from behind our walls, and frames. The wall created has the tendency to project the viewer then to his/her own space. If the canvas is the presence of a desirable world, then what creates this wall is the absence of the world around the frame, thus the absence of peripheral canvas. Perception is hence interrupted, and a separation of the space created in the painting and the space occupied by the viewer is highlighted. The body thus understands the space as scenery to look at, and not be projected in, and have a full perceptual engagement. Science has adopted the importance of peripheral vision as a research topic. "The preconscious perceptual realm, which is experienced outside the sphere of focused vision, seems to be just as important existentially as the focused image. In fact, there is medical evidence that peripheral vision has a higher priority in our perceptual and mental system", tells Pallasma. "Photographed architectural images are centralized images of focused gestalt; yet the quality of an architectural reality seems to depend fundamentally on the nature of peripheral vision, which enfolds the subject in space".

Some paintings have overcome this obstacle by spreading on a large scale, in this way the composition would itself extend to a peripheral vision.

The question then seems to form around how to preserve or create the desired sensorial experience acknowledging the restrictions of the space through the subject's peripheral surroundings.

C. Introducing Synesthetic Design

In normal situations, signals from the eyes are interpreted in the brain as light, the ones from nose as smell, tongue as taste, etc. However, at least four out of a hundred people Synesthesia is a neurological condition where the brain interprets signals with the "mixing of the senses". The subject would then experience colors as smell, or sound as colors, or any combination of the sort.

Many artists like Vincent Van Gogh, Joan Michel and Wassily Kandinsky were said to be or even diagnosed with synesthesia. Kandinsky had experienced it the first time during a performance of composer Wagner of "Lehengrin". He had an unusually visual response to the performance that he started to study painting. *(Miller, 2014)*

Terms like the "haptic gaze" or "visual noise" come to life only when two organs interrelate; the eyes and skin, the eyes and ears. Such terms are so often used in a design context, describing nothing but the visual presented. Experiential art's objective lies mainly in bodily perceptual engagements that we experience what we could call a kind of synesthesia. Even architecture was described by the German romantic Goethe as "frozen music". This layer of synesthesia added, the visual language is enriched with the triggering of the body, the memory, and the mind. Designers have picked up on these links and have tackled them in different ways in experimentations throughout history, not necessarily causing them as much as translating them. Whether behind a screen or in hands, graphic design is a process of creating a visual. The visual language of the theme suggested will aim to carry the refusal of communicating to vision alone as a mode of perception. Synesthetic design is suggested as a design that understands itself as part of the visual field, but stands against ocularcentrism.

A series of brief studies will suggest some cases:

Sound

Music has been a source of inspiration for numerous artists and designers. This interest between the two arts has developed a rather very rich visual language to communicate sound in general. The basic concept of sound waves and artists and designer's experimentations in its visualizations, the need to advertise for a band or for a kind of music through posters, LP/cassettes/CD covers, musical notes and the musical scale, Visual Jockeys (VJ) designing light compositions as DJs play music, etc. all contributed in a very developed communal understanding of sound visuals. Of course there were other contributions to the visualization of sounds, like noises, silence, or simply certain sounds like screams, or familiar sounds such as the waves of the sea…

"Sight isolates, whereas sound incorporates; vision is directional, whereas sound is omni-directional. The sense of sight implies exteriority, but sound creates an experience of interiority […] The eye reaches, but the ear receives." – Juhanni Pallasmaa

Pallasmaa suggests that sound articulates and gives an understanding of a space. The echos of an occupied room are completely different than that of an empty room. Hearing thus not only draws a space, occupies it, but also puts a visitor in interaction with it. The walls of a visited room reverberate every syllable to the center, and reply to each intensity. Visual representations of sound often create or stimulate this certain feel of the space or this flexibility in echo. (Pallasmaa, 1996)

In 1913, Kandinsky painted this composition called Composition VII, one of many paintings that hold a similar name in reaction to his experience while listening to Wagner's piece. The dynamism and the colors of the painting mirror his synesthetic experience. He continues these experimentations as he tries to systematically translate each note in the symphony into a shape and color. (Image 1)

Sight

Descartes found the eye a rather easily erred mechanism. It is often affiliated with the incorporation of memory and dreams, preferential selection, highly dependent on angle of sight, on state of the eye, etc. However the visual world of the 20th century until now did not find a lack in these phenomena, but richness. Sight no longer consisted of a reflection of sceneries, but a personal internalization of images.

At the TATE Modern, the art is not revealed until after the first night of the exhibition in subject. This created a challenge for James Goggin as he was designing the poster of an Olafur Eliasson exhibition in London.

Goggin was only shown a scaled model of the main installation, and was told about the minimalistic and atmospheric approach that Elliasson followed.

In this case the designer had to communicate a sight in an image without transferring it. The poster being a promise to the visitors of what is in the exhibition, he decided to transfer an aftersight. The installation revolves around a big source of yellow light, blinding as the sun. Goggin condensed the concept into a typographic and colored background composition. He transferred the disorientation and the charge of the optic nerves after looking at the sun in a flat yellow background, where you cannot see anything other than the color, and a pointed typography translating the back flare as a result of the blinding effect. (Image 2)

Smell

Sharing the same organ as breathing, smells have been present everyday of our lives. We very often realize a smell and involuntarily experience a kind of nostalgia. That is because smell is the sense that is most related to memory. "The most persistant memory of any space is often its smell. […] A particular smell makes us unknowingly re-enter a space completely forgotten by the retinal memory," tells Pallasma. The moment that the memory is recorded, a smell, an atmospheric image and a feeling are interconnected. The smell is then preserved in our memory by its spirit; a succession of colors and forms. When experiencing the smell again, these images are triggered and if the subject is conscious about it, he realizes the "visual smell". (Arts Santa Monica Departament de Cultura, 2011)

This is why color is a very important in the representation of this sense, as colors themselves have an atmospheric symbolism, and are part of the image memory.

Taste

Taste is a very intimate sense, so intimate that its stimulation is often a multisensory process. For instance, it has a very strong connection to smell. For a very long time, they were even considered undistinguishable. Also, there is a strong connection between the taste and tactile experiences. Some colors are even very strong simulators of oral sensations. "Our sensory experience of the world originates in the interior sensation of the mouth, and the world tends to return to its oral origins. The most archaic origin of architectural space is in the cavity of the mouth".

Therefore the communication of taste, especially in food photography, often involves the suggestion of texture on the tongue, and a strong smell. *(Pallasmaa, 1996)*

Food photography, especially in advertising, holds infinite experimentation and studies on how to appeal to taste. In this case, we view examples of photographer and designer Michelle Min. In the set of photographs, Min transfers the texture of three different flavors of different densities, using more of less the

same color palette. The graininess through the photograph but the dissolving from the faded but existent typography of the salt, the creaminess and sliminess, but the delicateness of the taste of the butter, and the powder very light particles that fly like dust of flour, all documentation of their characteristics. The photograph of the yolk of the egg transfers its concentration and sliminess. The last photograph showcases a mixture of a texture and a pinch. (Image 3)

Touch

Touch is no doubt the sense that is most related to intimacy. In fact, a big lexical field involves both concepts which makes it tend to have this affiliation: skin, feel, touch, smooth, etc. Having this character of intimacy, hands have history. Cultures have made themselves acquainted with the offering of the intimate touch of the hand in a handshake. The dress of Botticelli's Venus has made itself intimate with the emphasis of a graceful touch of the fabric on the skin. "The skin reads texture, weight, density and temperature of a matter". The skin also can read time, within wrinkles, dust, countless waves on a shore that erase the rough texture from a smoothened pebble, etc. Spread on the biggest organ of the body, the skin, and the modern world involving all bodily reception of information as sense, tactile sensations are of very big variety. Even the widely used categorization

of warm and cold colors demonstrates an intimacy that one exercises with a visual. Also pain, discomfort like ich, temperature will all be considered within this section. *(Pallasmsa, 1996)*

Yves Kleine looked for the art of the after touch. He looked to visualize the imprint of a body. This could stand as the aftertouch, or the memory of the touch. The nudity and feminine character of the created silhouettes portray the intimacy of the sense of touch.

III. SPACE

This chapter will discuss a spatial element, and will tackle questions regarding this third dimension. It will follow the development of the definition and characteristics of an "exhibition" with commentary to put it in the sensorial framework. It also introduces the graphic designer in the context.

A. Ahistorical Exhibition

An "ahistorical exhibition" is a term born to describe a rejection of the tradition of displaying a chronology of art, a perception derived from the "mixed galleries" of the seventeenth century. Another way of classification was encouraged, one that involved "correspondences" between diverse works of different eras and cultures instead of time based or style based segregation. A correspondence is a link developed by the exhibition's curator to group a number of diverse works under a notion of his proposition. This link would therefore be far from this chronological or stylistic categorization and more challenging, the concept of style being debated to be a creation of art historians that belong to a later period looking back at art and systemizing it. However curators of this kind of classification called "thematic affinity" are criticized as "arbiters of taste". However they defend themselves as "intuitive classifiers", where the correspondences or sought "essential links between the different arts" that they suggest "transcend art-historical classifications in terms of period and style" and "communicate a message". This approach links an exhibition in a theoretical or historical framework, instead of a pure exhibit of visual forms. The introduction of these links starts to give the visitor more power of personal interaction or relationship with the work.

The later half of the twentieth century faced a collision and debates also around theories that try to define what a museum's role is. The "museum is an institution which plays a decisive part in determining the significance of works of art" mentions Meijers. The 1900 Expo Universelle showed a start of attempts of targeting sensorial stimulation in expositions. The postmodern era grows

interested in the display of diverse artists and characters across the exhibition, shyly existent in the classical Academy. Museums were sequenced in a bolder manner, like in the Romantic period where originality of an artistic voice was held above all. The sequence did not try to recall a traditional historical form or story. Chronology of development was no longer interesting, nor was it convincing. "Modern artists" did not learn from "Past Masters", but they created their own character, introducing philosophical concepts to the initiative of curating an exhibition, like human cognition and perception.

This research abandons notions of time and space in the study of the works of art in subject, baring them to the pure question of the bodily stimulation of the audience. *(Meijiers, 1992)*

B. Space and circulation in an exhibition – research study

Planning an exhibition cannot exclude the study of spaces. The exhibition does not exclusively communicate the works of art it showcases, but also sets the motion and the direction of the viewer or suggests it. Time and space are both dimensions an exhibit event tackles.

The exhibitions suggested is an access to a selection of works under a suggested narrative accentuated by surrounding visual, physical elements and the architectural space. As a more converged definition, the narrative suggested is the bodily engagement of the viewer, more precisely through a converging of his/her perceptual capacities. Also in this research, the term "exhibited material" will refer to all viewed material in the suggested space, thus will not exclude exhibition signage, display elements, and all intentional interventions on the space.

A walker in "Writing of History" is not a follower of a suggested plan. French philosopher Michel de Certeau suggests in his text that governmental and authoritative institutional maps indeed divide the cities and design them, but have no proper mimic when it comes to the walker in the city. These maps comprehend the city as one whole, an unrealistic expectation according to Michel de Certeau. The walker defines their own divisions according to preferences, comfort, knowledge of the roads, shortcuts and even memories. The map is thus broken into disproportional usage. In a study called "Walking, looking and longing: on the perceptual engagement with art exhibitions" by the TATE, this explanation is compared to the experience of exhibition spaces. Just like in the city, the visitor is a unit in the laid out space. The visitor does not experience the exhibition space as a unified whole, or as it is suggested by the designer, however by their own preference, perspective, memories, etc. An exhibition designer either directs, suggests, or even loses this human unit in the space composition, through the sequence of the works exhibited, the methods of display, the theme of the exhibition, the circulation or the divisions of space, etc. The designer does not have the power of imposing an experience of the space on the viewer. "A large part of the artworks they might never see, and there will be areas of the exhibition that they might completely skip. But their senses are continuously in play, and they form connections with the artworks and each other." In this research, the layers of an exhibition tackling the movement in space or the encounter of materials are acknowledged. The sensorial human walking experience has been a subject of study for British anthropologist Tim Ingold. He describes the movement as "making lines in space". "To correspond with the world, in short, is not to describe it, or to represent it, but to answer to it". He defines then the process of answering to the world. "The human being emerges as a center of awareness and agency whose processes resonate with those of its environment." In the course of this movement, the human then becomes an attentive intelligence, making correspondences and links through his experience. Ingold's theories could be seen in the light of the parallelism of the walking

experience of an exhibition and the encountering of materials. Other theories revolve around the relationship between the human unit with history, emotions, perspective and the "thing", in an exhibition context, the exhibited material. Jane Benett's "Vibrant Matter: a political ecology of things" discusses the coexistence of humans with other things in one space, or "matter" and relationship to matter. Benett also acknowledges the human as a perceptual capacity with an ongoing creative analysis that makes links and connections. Petra Tjitske Kalshoven approaches the subject from a musicological perspective. She mentions the curiosity of the viewers, and the romantic notion of magic of the artifacts when speaking of making relationships while walking in a museum. Kalshoven hence makes the position of the object and that of the human even more intertwined and flexible, and more emotionally invested, mentioning the desire and the curiosity behind this relationship.

C.   Questions regarding the space

The research suggests a space of perceptual harmony. While understanding that it is impossible to force a defined experience and a unified path on visitors, my role will include investigating how to suggest the objective theory keeping in mind this subjectivity of the viewer. The roles I investigate in the project are together the choice of the space itself, the curator of the visuals, the designer of the route of the viewer, and the graphic communicator of the five exhibitions.

a.   Regarding the neutrality of space towards the suggested experience

The experience sought is that of full perceptual engagement with the space. This experience is enhanced with an attempt at an elimination of diversions whether it is the history, time, the invasion of the daily life, etc. In the 1920s, with the rise of popularity of the ahistorical exhibition, an effort to abide to a neutral character in space and in time that metamorphoses with every correspondence it houses was done. A "sacred" museum room that held the title "The White Cube" was constructed. The room welcomed exhibitions with clashing visuals and diverse voices, and sought to preserve visual differences without manipulate their perception through an unintended atmosphere. Whether through the minimalistic methods of enclosure of the space, or the simplified character, the art was no longer classified in time or space, but offered a neutral ground to the subject. Another way of finding a neutral ground in this structure is through its color scheme. The abundance of white on contemporary exhibition walls are actually revivals of this neutrality of the cube, and a very typical type of museums of the fifties. The lighting of the museum also has a huge role in creating the atmosphere aimed for. The white walls and the strong and homogeneous distribution of light in the room was also of a certain neutrality towards the art. Harald Szeeman designed an ahistorical exhibition at the Boijmans van Beuningen Museum in Rotterdam in 1988 that recalled concepts of the juxtaposition of an artistic diversity the classical academy entitled "Ahistoric Sounds", in hopes of seeking "the essence of the work of art". For his spiritual idea, he chose to light the spacious exhibition strongly and almost equally to create a sort of meditative aura. These examples, however, are only taken as a case study, but not as an ideal solution for the problematic of unwanted characters being imposed on a space. *(TATE Modern research Center)*

b.   Regarding the sequence and content of the exhibition

Neither a historical chronology nor a special or cultural or stylistic regard will be emphasized in the choice or sequence of the works of art. The progressive position of a human of the twenty-first century giving a liberated but cultured overview over the art of the previous centuries will be taken advantage of. A series of ahistorical exhibitions that steer away from debates by shedding a light on their collections outside of categories and time constrictions is suggested. They instead depend on categories of my proposal of the relevance of the topic. In the exhibitions, the sequence or dispositions of the art in subject although will credit the artist but will not seek to inform of neither the historical nor the stylistic context in which the piece was created. It will seek to increase the visitor's engagement with his own body rather than his mind. The experiential graphic design incorporated – signage - in the space will take part in realizing the objective as much as the artwork itself, with the role of directing the audience towards a more intensified experience. *(Meijiers, 1992)*

d.   olafur eliasson – case study

Olafur Eliasson is an Icelandic Danish artist known for his perceptually challenging installation art and exhibitions. Having lived in Iceland, Eliasson has a very developed sense of appreciation and attachment to nature, and its expansion of the human perceptual capacities. This attitude is voiced in every work he presents through his sensitivity to the world of perception, the physical and the immaterial. In his work, he converts the passive visitors of an exhibition into "inhabitants" of the art, where the art is not only what we see but how we see, or in his words: "seeing yourself seeing". Eliasson's main challenge when designing his art is the understanding of the space of the exhibition, the art and what it stimulates in human perception and the integration of a nature that is emphasized, hinted at or missed. "I strongly believe that it isn't necessary to polarize the fields of art and architecture (of the exhibition). Rather, it's about trying to transgress the traditional boundaries to create a space of inclusion and hospitality, where differences of opinion are not only tolerated but encouraged. When you enter my exhibition, you do not step out of the city of Stockholm and into the protected world of art, but continue the processes of negotiation and co-production that characterize our shared reality," he says. I study therefore his different approaches to the challenging of the human perceptual realm.

"Riverberated" is an installation that simulates a rocky landscape crossed by a river, occupying the whole exhibition space of the Danish Louisiana Museum of Modern Art. The project expresses Eliasson's attack to a separation of the natural and the built, and nostalgia to a perceptually pleasing landscape absent from the urban city. "When I walk or drive through the Icelandic landscape, I sense the surroundings and sense myself searching for sense. This vast landscape is like a test site that nurtures ideas and helps me process them into felt feelings –maybe even into art" he says. Focusing on the progress of the walker in the exhibition, the route of sensorial stimulation is a reinterpretation of the core concept of: "the gallery is again primarily a place where we walk" as Marie Laurberg defines it.

The approach of covering the whole exhibitory area with elements from one nature strongly suggests a uniformity of space. The experience of the walk in the exhibition however, although embraces this homogeneity, but simultaneously highlights the transition between the different rooms of the exhibition space by tightening the orifice of connection between the rooms. This would enforce on the walk a certain movement that would introduce the body to a different spatial context. The walker flows in the simulated harmonious landscape, taking part of the installation.

His bodily movement takes an equal part in the exhibition as the rocks and the water, his walk being as much of the art as the physical intervention of the artist. The walk emphasizes a kind of experience rather than a distant visit to a cultural zone. The installation showcases an invasion of landscape on the walk, pushing the physicality of the architecture onto a perceptually engaging space. The experience of the exhibition abandons the walls and the intellectual analysis or memory to delve into the

body's intelligence. (Image 4)

Presenting a truth, and an element of esteem, the artwork in usual contemporary exhibitions looks untouched on a white setting that suggests a pure inhumane dimension, keeping away any human intervention with glass protection and synthetic material. Olafur Eliasson's installation rivals this attitude in engaging the body of the visitor in the art work. Map of Unthought Thoughts occupies the full volume of the room extending light forms from a central installation of a metal structure. The lit installation casts its presence on all the walls of the room, welcoming the visitor to join its presence. The experience is a witnessing to a space that is occupied with a perceptually tacit presence not with the occurrence of physical elements the room. When inside, the body takes part in that presence when its shadow is perceived interconnected with that of the central installation. The shadows casted intersect in a way one's perception of their own movement is flattened onto the perception of another object, the object that busies the room. The movement is paralleled instead of mirrored, as one's perception of oneself, which would be their shadow, grows when approached and shrinks when distanced instead of casting only a flat motor movement. This gives a power to the visitor that he or she lacks in reality, suggesting a non-rational situation. In this installation, Eliasson experimented with the occupation of space in volume vs mass, and the integration of the body of the visitor with the work of art. The presence of the visitor on the walls of the exhibition did create an important articulation of it, and was encouraged instead of disliked. (Image 5)

Eliasson also worked on intending to the question of the existence of the body in a space vs the existence of the body in a space with other bodies. To simulate solitude, or the alone presence of the body, he fabricated a room with a concentrated fog that creates a dense separation between different bodies that conceal them from each other. In this experiment, the space between the wall and the body is contracted to flatness, making the exhibition focused on within the body instead of the outer experience of space. Walking in the space does not suggest suggestions of different visuals, but a liberated choice on what the body interacts to. (Image 6)

Color studies were a big part of the experimentations of Oliafur Eliasson. A developed experimentation in regards to this question was the Feelings are Facts exhibition. Merge of strong colors and flatness are two characters that belong to only the lexical field of a painting, but not of a space. Eliasson suggests a dimension where this is reversed. What if the viewer existed in the flatness of the painting? The exhibition had an opening on a solid wall that introduced to the viewer a segment of what the exhibition room is. The visitor would enter through the opening into a dense fog of bright unnatural colors like magenta, emerald, yellow and electric blue. Standing in the fog, the visitor's sense of orientation, sense of depth and sense of reality are distorted. The circulation in the space is purely sensorial and an experience of feelings. In this work Eliasson intermingles the components of an exhibition, being visitor and work of art, to create pure sensory experience. (Image 7)

Olafur Eliasson deals with art, mechanisms, light, design, and curation. His work reshuffles the components or characters of the exhibition, to suggest a challenge to what is the habit or what is the more convenient. Creating a dialogue between the body, the space, and sensorial elements, he evokes a certain body consciousness, and often nostalgia to nature. Tackling with the dimensions of an exhibition, his work enriches my archive and understanding of the experience of a human circulation in one.

D. Graphic Design and Exhibition

It was until the second half of the 1900s that the Museum got rid of its reputation of being exclusive for the elites. Curators had found it more interesting to collect the rarest pieces and conduct their own lines of research rather than to present their findings to the biggest audience. The 1960s introduced however the start of the commercial museum that needs thus a visual identity and a graphic artist who would provide it. The interest of forming a visual identity for museums started to rise along the peak of advertising in the 1970s.

One of the first appearance of the phenomenon was by Willem Sandberg, museum director of the Stedelijk Museum in Amsterdam who started to design posters and catalogues for the exhibitions held in the museum between 1945 and 1963. In 1966 Dutch designer Wim Crowel and Total Design studio joined forces to further develop the visual identity of the museum. A defined grid, which was preserved from Sandberg's designs, was the base of the consistency of this identity, which eventually made the latter recognizable.

Jean Leering, director of the Van Abbemuseum in Eindhoven chose designer Jan van Toorn to work on posters and catalogues for the museum from 1965 until around 1973. Sharing Leering's provocative ideas about the absence of neutrality of an exhibition space, and rejection of a dependence on one housestyle followed through in the visuals of the museum, van Toorn designed with the idea that "stimulating the outrageous in order to awaken consciousness" is more interesting.

The 1980s held a very rapid boom in regards to museums, urging thus for the need of graphic artists to ensure an appealing communication, resulting in more variety of approaches and characters. Visual identities started being carried through more and more consistently, including identities for museums. Inevitably, different cultures did forecast a different spirit and sensitivities in their visuals, France with more image oriented graphics, Germany with a more direct and functional attitude, Netherlands with a non-decorative abstraction, Switzerland with a high respect for clear grids, organization and legibility, and the United States with a more commercial attitude. However, the Museum did not reach all countries of the world at the same time. Until the 1990s, Japanese collections of art were private collections of religious groups or elites of the society, never displayed to audiences except on occasions of festivals, anniversaries or ceremonies of any kind. Professor of the University of Tokyo Shuji Takashina described museums in 1992 as: "an institution not yet a century old which has been imported from the West".

a. The whitney museum of american art – case study

I also investigate about the visual communicative design entailed in an exhibition. Impressed with many approaches and projects of museum/exhibition design, I choose to present the recent design of a group of contemporary leading designers based in the Netherlands. I present and relate to the identity formed for the Whitney Museum of American Art by Experimental Jetset studio. Experimental Jetset were asked to redesign the identity of the Whitney after its relocation, and submit a "toolbox" that the in-house design would use for the future communication of the museum. The character and ideology of the Whitney was put under a scope. The Whitney presents, according to Experimental Jetset, a more complicated history of art. This generated a parallelism between a straight line of history, and one with more fractures. Taking into account the industrial architecture of the new host building, and the feel wanted from the prints in context of the city, a zigzag line was used in the logo as a metaphor. Linking one segment of the frame of the format to another, the zigzag line represents the cultural heartbeat of the city, the waves of sound or vision. This simple expression of the identity, of course with the strong convenience that the line mirrors the initial of the Whitney, gives a big dynamic range of usage. As long as a black

line divides the space into 4 equal areas, and starts with an upper left to lower right diagonal and ends with a lower left to upper right diagonal, the logo will be perceived and understood by the audience. (Image 8)

This approach to a visual identity is very liberating in the need to design for different characters and different concepts of the future coming exhibitions. An identity of simplicity, but also modernistic dynamism will also be imposed on the design by the feature of the logo. Thus the character of the museum will be present with the simple existence of the logo, without having to impose a style that will then affect the communication of the varied themes of the coming exhibitions.

This zigzag line also articulates a perception of a three dimensional space within the poster. Its variations almost symbolize the different angles the now called "Responsive W" is viewed.

Although the studio wanted to go with a typographic approach for the housestyle of the posters, the designers of the Whitney Museum were keen on going for a single image approach. This would require choosing one work of art from the subject exhibition and juxtaposing it with the logo now carrying the word Whitney. Experimental Jetset dislikes this approach as "reproducing this work of art especially for commercial purposes diminishes it". However with the request of the client, they proceeded with the project on this basis. The posters of always the same size will include works of different sizes. This forces a space that would almost always appear around the artwork. The background being white, this space is convenient to represent the space around the work, so in reality, the exhibition itself. "The new graphic identity needs to be adaptable to the Whitney's commitment to the primacy of the artwork. Art comes first, the institute follows."

The exhibition signage system was decided to be applied directly on walls instead of on a plaque or board. This approach suggests more freshness and causality during the circulation in the exhibition. To add perceptual depth or a sensorial impression, the mounted signage was enlarged to an extent that it would suggest a certain 3D, like supergraphics.

Experimental Jetset chose to use the font neue Haas Grotesk by swiss designer Max Miedinger and Edward Hofman, later redrawn by Christian Schwartz in New York 2010. This font is meant to be expressive or appropriate of New York, especially because of its similarity to Akzidenz Grotesk, according to Experimental Jetset. This case study is a more concrete study of one example of a designer of an exhibition space, and it mentions the dimension of the research regarding the design process of print design and signage system, etc. The interest in it comes from its answer to the question: How can we express a uniform identity while keeping a huge freedom to how it is used. The very dynamic and flexible line has such a very strong character that it reads in different perceptual languages in different posters without imposing a strict feel. However, this case study is also mentioned to express that I will not be using the single image approach. In fact, the visual communicative language of my project will not borrow at all any artwork of a different author than me. It will express my own design of the synesthetic experience.

Conclusion

In the research, I propose challenges on how to design a space that exhibits works of artists in a way that answers to the theoretic context suggested; how to design a space that presents art that simulates sensorial impressions in a way that enhances a perceptual experience. The language suggested for the communication material is also based on the definition presented of Synesthetic Design. The graphic language of the exhibition series employs the same task as the exhibition itself. It attempts to challenge the audience's perceptual linkage, triggering the body instead of the information

processing. However, the challenge could also complete this task in the different places and usages of the items. A poster is placed within an environmental surrounding, the brochure as of the entrance to the exhibition, etc. The emphasis is on approaching the body and not the mind. The research proposes and develops a definition proposed of "Synesthetic Design" in space and in print.

References

Although all links were part of the research, but not all are quoted directly in the research paper. I involve them all in my bibliography.
Of Philosophies on Senses:
Books:
(1) Carr, Donald Eaton. The Forgotten Senses, 1972.
(2) Crowther, Paul. Art and Embodiment: From Aesthetics to Self-consciousness, 1933.
(3) Ehrenzweig, Anton. The Psychoanalysis of Artistic Vision and Hearing: An Introduction to a Theory of Unconscious Perception, 1975.
(4) Macpherson, Fiona. The senses: Classic and Contemporary Philosophical Perspectives, 2011.
(5) Marsh, James L.. Post-Cartesian Meditations: An Essay in Dialectical Phenomenology, 1988.
(6) Merleau-Ponty, Maurice. Phenomenology of Perception, 1945.
(7) Pallasmaa, Juhani. The Eyes of the Skin, 1996.
(8) Pallasmaa, Juhani. The Thinking Hand, 2009.
(9) Reid , Thomas. An Inquiry into the Human Mind on the Principles of Common Sense, 1810.
(10) Robinson, Sarah and Juhani Pallasmaa. Mind in Architecture: Neuroscience, Embodiment, and the Future of Design, 2015.
Articles:
(11) Fowler, Thomas A.. More than Meets the Eye: Aristotle on Sense Perception, 1995
(12) Pink, Sarah and David Howes. The Future of Sensory Anthropology / The Anthropology of The Senses, 2010.
On Art on Senses:
Books:
(13) Kandinsky, Wassily. Concerning the Spiritual in Art, 1912.
Articles:
(14) Bristow, William. Enlightenment, The Stanford Encyclopedia of Philosophy, 2010.
(15) Hernandez Barbosa, Sonsoles. The 1900 World's Fair or the Attraction of the Senses, The Senses and Society, 2015.
(16) Khan Academy. A beginner's guide to the Age of Enlightenment.
(17) Marshall, Anne. Longthorpe Tower, Peterborough, Cambridgeshire c.1310: The East Wall, 2012.
(18) Miller, Renée B.. Wassily Kandinsky's Symphony of Colors, 2014.
(19) Nordenfalk, Carl. The Five Senses in Late Medieval and Renaissance Art, Journal of the Warburg and Courtaild Intitutes, Vol. 48, 1985.
(20) The Art Story Contributors, Futurism, 2015.
Of Science on Senses:
Articles:
(21) Carruthers, Peter. Animal Minds are Real, (Distinctively) Human Minds are Not, 2013.
(22) Hiskey, Daven. Humans Have a Lot More Than Five Senses, 2010.
Of Design of Exhibitions / Museums / Identities:
Books:
(23) Greenberg, Reesa, Bruce W. Ferguson and Sandy Nairne. Thinking about Exhibitions, 1996.
(24) Meijiers, Debora J.. The Museum and the "Ahistorical" Exhibition, 1992.
(25) Skolos, Nancy and Thomas Wedell. Graphic Design Process, From Problem to Solution 20 Case Studies, 2012.

Image 1

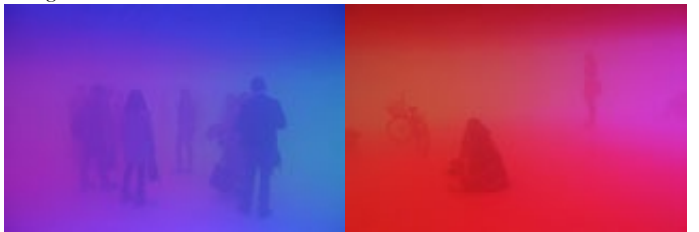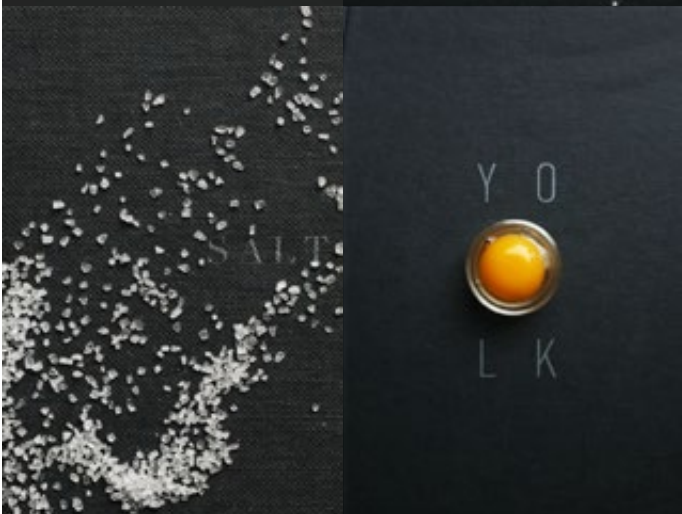Image 4

Image 5

Image 2

Image 6

Image 3

Image 7

Image 8

Adaptive Landforms

Ismail Hutet and Mariam Maria Yassin
Department of Architecture and Design
American University of Beirut
Riad El-Solh / Beirut 1107 2020 Lebanon
iah09@mail.aub.edu
mhy13@mail.aub.edu

Abstract- **In an attempt to rediscover the mountainous area of Cappadocia in Turkey as a habitable space for social and cultural opportunities, the design s**tudio entitled "The Fairy Pits and Towers" is an exploration of spatial possibilities and a weave of **social narratives through the study and research of landscape morphologies and geological layers of the area in question.**

"Adaptive Landforms" is a design **proposal based upon a** rigorous research on the "Fairy Chimneys" rock formation in **Cappadocia. Firstly investigating the volcanic area at a macro-scale, in an attempt to understand these extremely interesting natural formations, and secondly zooming-in at a micro-scale to the area of Zelveh, which is currently maintained as an open-air museum.**

In order to understand the nature of these formations, our research overlaps diverse aspects of the site; namely climate conditions and rock formation, natural formation vs. man-made interventions, materiality and atmosphere, land and leisure economy, and lastly culture and the human.

"Adaptive landforms" culminates in an architecture **intervention that moves from research and narratives into composed spaces through mapping of processes and connections, experimentation in large scale models, and exploration in form finding methods. The proposed programmatic scenarios carefully** occupy the "fairy chimney" landscape, combining space -**forming approaches with imagined narratives finally resulting in a mature urban scale architectural project.**

The project was presented at Studio X in Turkey (founded as an initiative of Columbia University Graduate School of Architecture, Planning and Preservation (GSAPP). Studio-X Istanbul is an urban laboratory that aims to identify the current and future issues facing the city and seeks to generate innovative forms of thinking for their solutions. Finally, the project main objective is to vitalize the area to its local inhabitant in an attempt against the gentrification of Cappadocia by creating a cultural center that understands the essence of the region.

I. A GEOLOGICAL STUDY

A. An Action

During the Tertiary period, frequent and continual eruptions and earthquakes occurred for thousands of years. The three volcanoes: Mount Erciyes, Mount Hasan and a smaller mountain named Güllüdag were then active.
These volcanos are located on the edges of Cappadocia. Their frequent eruptions created deposits of volcanoes ash, lava and basalt which laid the foundations for today's landscape. (Cappadocia, Ferika Özer Sari and Malike Özsoy)

The layers thus formed are now known as tuff, which is a soft porous stone (Cookson, 2007). The succession of the volcanic eruptions led to the layering of the Cappadocian ignimbrite. During the volcanic activities, the layer of tuff was covered by layers of lava and basalt. The thickness of these layers was not equal. Nowadays, the Cappadocian ignimbrite is divided into three sections which correspond to the different time periods of the volcanic eruptions. The section of figure1 is derived from a thesis entitled: "Fairy Chimney Development in Cappadocian Ignimbrites" by M. Naci Sayin, 2008. In this section we demonstrate these layers from which the different types of fairy chimneys take their names and characteristics.
The volcanic material slowly ran towards the depressed areas and covered the previously formed hills and valleys. This geological activity changed the general landscape of the region, giving it the appearance of a *plateau*.

B. A Reaction

The Cappadocian tuff is highly subjected to weathering and erosion. The climate of Cappadocia thus plays a major role in the formation of its landscape. It is characterized by heavy rains, snowy winters, and strong winds. The climate of Cappadocia is a continental climate in which the difference in temperature is high (Murakami, 2008). Thus, mechanical weathering plays a role in the fragmentation of the rocks which expand when heated and break up as they cool. However, the most important sources of erosion are rain and river flows. Heavy rainfall transformed the smooth surface of the plateau into a complex pattern of gullies. The erosion followed preexisting fissures in the rocks. These cracks were formed during the layering stages. As the lava layers covered the tuff, it cooled down. This change in the lava's temperature caused it to crack. The layers were then further covered with ashes and basalt, but the water still found its way through the existing cracks. As the soft layers of the ignimbrite eroded, the water reached the basalt layer, which is much harder. These layers later formed what is known as the fairy chimney cap. The sequence of the layers of the ignimbrite defined the type of the

Figure 1. Section through each type of fairy chimney

A further study of the precipitation pattern yielded to significant understanding of the formation of these different types of fairy chimneys. Figure 2 taken from "Fairy Chimney Development in Cappadocian Ignimbrites" by M. Naci Sayin, 2008, p:44, shows the location of the different types of fairy chimneys. The overlay of this map with the annual precipitation map (fig. 3) reveals an intriguing relationship between the two since each type falls into a region of different precipitation count.
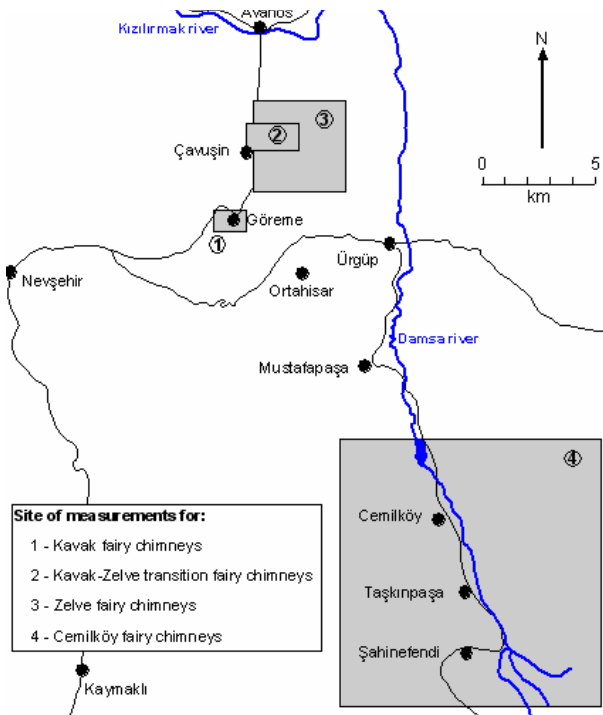

Figure 2. Location of the different types of fairy chimneys
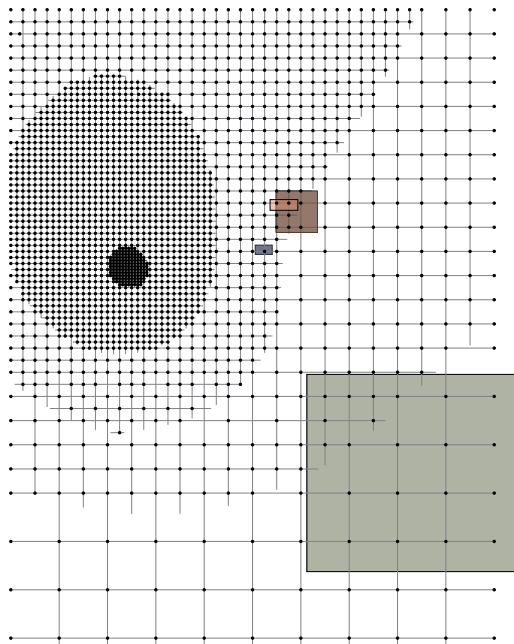

Figure 3. Precipitation map of Cappadocia

Through our research we could relate these types to different erosion patterns of each area. We could also identify through these types the layers of the ignimbrite which still existed on each site. We can anticipate that further erosion of the fairy chimneys would eventually lead to the formation of new types of fairy chimneys as different layers of ignimbrite are affected by erosion.

The different fairy chimneys are defined according to the layers at which they occur. The major difference is the sequence and the thickness of the tuff, lava, basalt and the deposits layers. This further defines the structural characteristics of these formations (fig. 4).
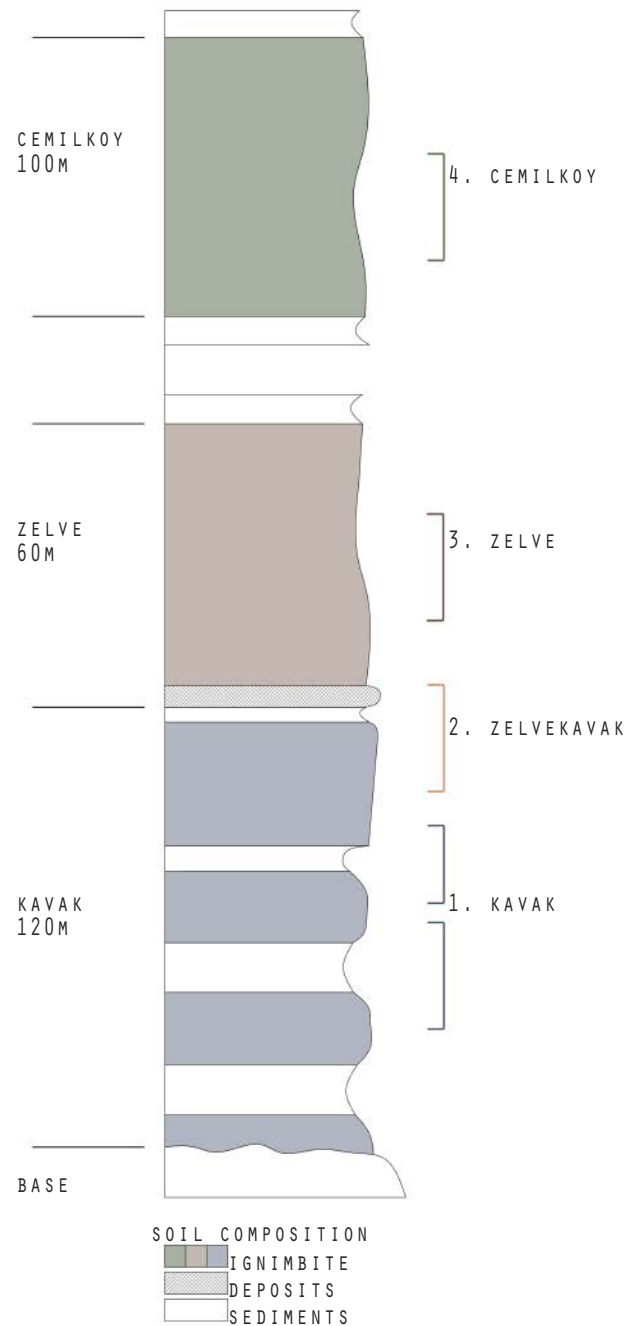

Figure 4. Section through soil composition of types of fairy chimneys

KAVAK AREA   ZELVE AREA   CEMILKOY AREA

Figure 5. Mapping of fairy chimneys in each location

Through further mapping of the fairy chimneys in each studied location, we were able to unravel a new relationship. Each type of fairy chimneys had a specific range for a diameter, and the distance between the chimneys was also related to the type. This pattern was most likely caused by the erosion pattern which was in turn defined by the cracks in specific layers. From this mapping (fig. 5) we were able to obtain various grids, each specific to a type of fairy chimneys. We mapped the distances between these formations. Comparing the grids allowed us to obtain a ratio of 1:2:8 to compare the different types.

C. A Human Intervention

We then zoomed into *Zelve* region where there's a clear shift of precipitation count creating a border of formation and an interesting region for our study. The *Zelve* ignimbrite is the middle layer, thus allowing further anticipation of future formation of new fairy chimneys in the Kavak layer and the reminiscence of the Cemilkoy layer (fig. 6).

Through a careful analysis of *Zelve,* which is maintained as an open air museum, we decided on a program that would further revive the site, yet relating to its unique morphology. The section in figure 7 shows the interaction between the soft nature of the soil in *Zelve* and the human. What is now the open-air museum of *Zelve* once was an agglomeration of dwellings, rock churches and a mosque. The abandoned settlement consisting of formerly occupied cave houses spreads over three small valleys. In the 50s, all inhabitants left the village because of the increasing risk of a collapse of the rock houses due to ongoing erosion.
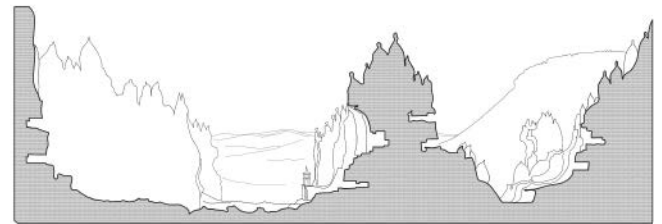

Figure 7. Section through Zelve open-air museum

II.   A NARRATIVE

A. Interpretation through Model, Recreation through Design

Basing on the geological study of *Zelve,* we built a study model in an attempt to understand the formation of this specific type of fairy chimneys in this area (type 2 and 3). Through the structural understanding of soil composition in terms of layering and in terms of the interaction of the external and internal forces, we created a structural unit. This unit can be manipulated in terms of scale, connection, joinery, repetition, pattern, and expansion. From our understanding of the composition of the soil, the fairy chimney is created due to different effects of erosion on the 2 existing soil types, the soft and the hard. This fact is translated in our structural unit in the presence of the soft point (the joint) and the hard point (the center). The connections of these units have a similar structural behavior of that of the soil composition. We stimulated the external forces of erosion through forces of compression (which forms the cap of the fairy chimney) and tension (which erodes the surrounding of it; the soft soil). The unique structure of the unit enables it to expand three-dimensionally, shift its center, push and pull the surrounding units, and ultimately interact with the other layers of soils that it is sandwiched between (figure 8). This exposed the forces in play between the

Figure 6. Vision

Figure 8. Apparatus model

between (figure 8). This exposed the forces in play between the different layers of soil. It also acted as the basis of the reconstruction of the eroded layers. Thus, when recreating Zelve's formation of fairy chimneys through the model, we mapped the present condition of the chimneys under three categories: intact chimneys, completely eroded chimneys, and altered chimneys. This layer, Zelve, is to react with the eroded/rebuilt layer of *Cemilkoy* and the anticipated layer of *Kavak.* This interaction will reveal two categories of fairy chimneys: the previously deformed chimneys and the now altered chimneys (figure 9). This experiment guided us to recreate/rebuild the eroded layer of *Cemilkoy* as an open platform for recreational purposes.

B. Evolution through Time and Space – Vision and Spatial Qualities

The platform morphing with *Zelve,* through specific integration with the site and with the fairy chimneys of the area (fig. 10). The structure is to expand horizontally between, above, and under the fairy chimneys of *Zelve.* The shifting and the intersection of the platform with the slowly eroding soil generate connections between the present layer, the eroded one, and the one to be discovered. This connection evolves through time and space with a structural element becoming an anchor for the present to discover the past and the future.

Two morphological principals of creation/erosion might be claimed to exist in the land of fairies pits and towers from our research and mappings: a verticality to that is directed towards isolation, versus a horizontality that is directed towards connectivity and filtration of the human through the site.

The sectional vision (fig. 10) evolves through time. The attempt to fuse the discovery of the site with its verticality and horizontality led to the reconstruction of the eroded layer with interconnected structures and platforms. Mimicking the fairy chimneys rock formation into functional spaces that erect vertically, these platforms are disrupted horizontally but further connected vertically.
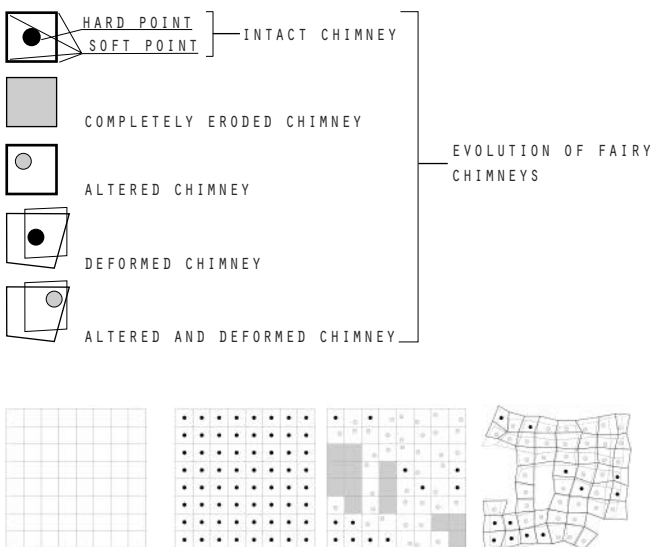

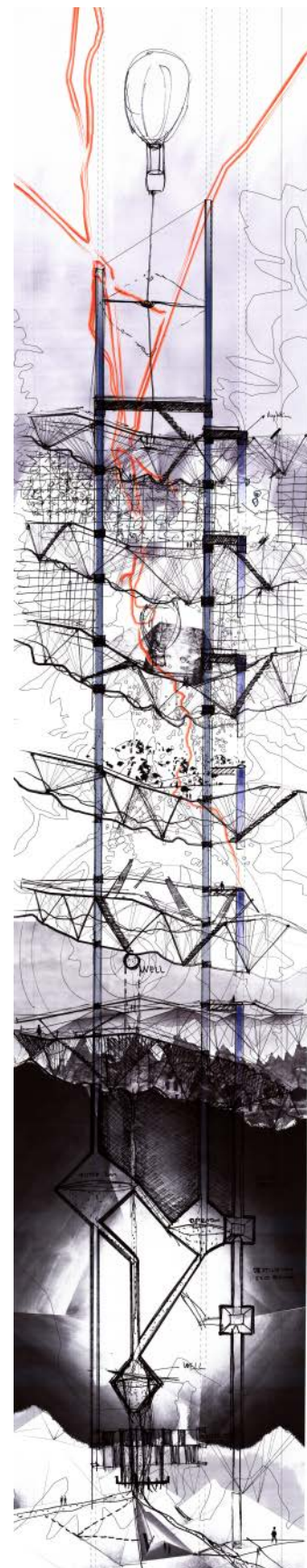Figure 9. Fairies chimneys' grid and pattern
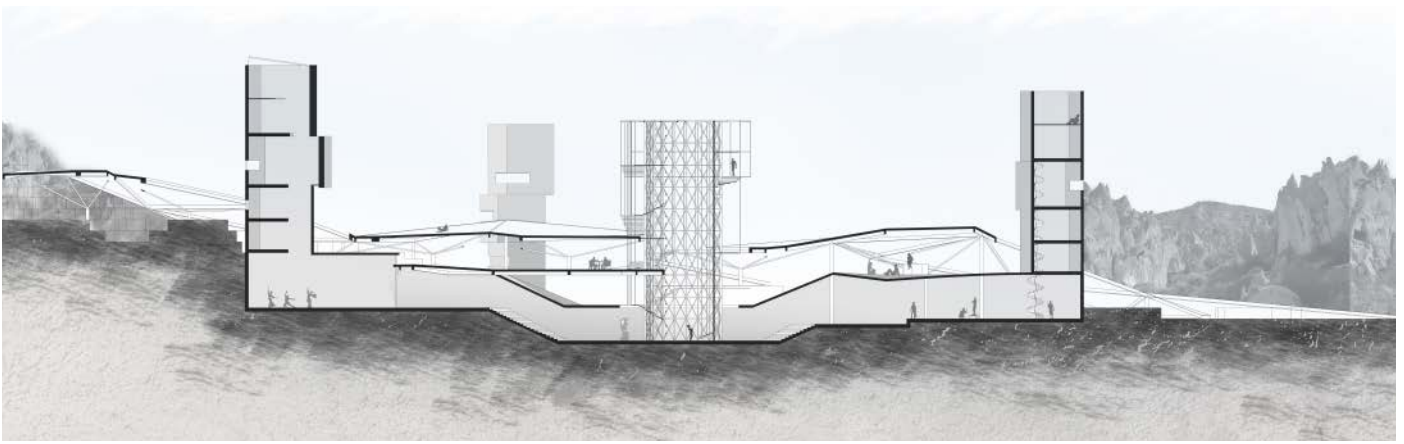

Figure 10. Sectional narrative

Figure 11. Section through Adaptive Landforms

III. DESIGN DEVELOPMENT

"Adaptive Landforms" is a proposal for spatial opportunities, social interaction, and art exploration that will connect the users to nature and guide them to discover the history of its making. The proposed program encloses a museum, galleries, workshops, artists' residences, and an open-air platform for recreational purposes such as cafes, amphitheaters, promenade paths connecting to different levels of the site, and finally a central skeleton connecting all platforms with the underground spaces and the shifting volumes and allowing the visitors to experience the site from above (fig 11).

The main function of the platforms is to provide accessibility to the site on different levels through the main central skeleton (fig. 12). The platforms are able to expand horizontally on the site in a radial manner which projects the potential of the structure and analysis at hand, allowing for potential future expansion. Whereas, the vertical elements encourages vertical exploration and wonder that will inspire the artist to create.
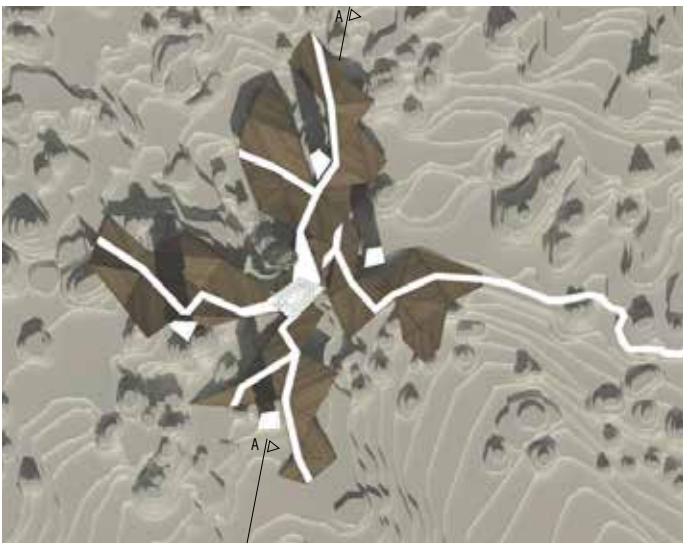
exploration and wonder that will inspire the artist to create. Each vertical element is a private residence for artists who have different emphasis/approach in their artistic exploration. Relating to our research, these towers are always located on the harder and more stable points on site. These towers further follow the morphology around them, creating an adapted shift with each unit. The towers extend underground to enclose the workshop spaces. A diagonal connection from these workshop spaces intersect in the center of the project enclosing the private galleries for each workshop space, and finally the museum in the center (fig. 13) From this focal point the skeleton erects to provide connection to each level of the platforms, ending with an observatory deck at the highest level portraying a deeper sense of wonder and speculation (Appendix A fig. 14).
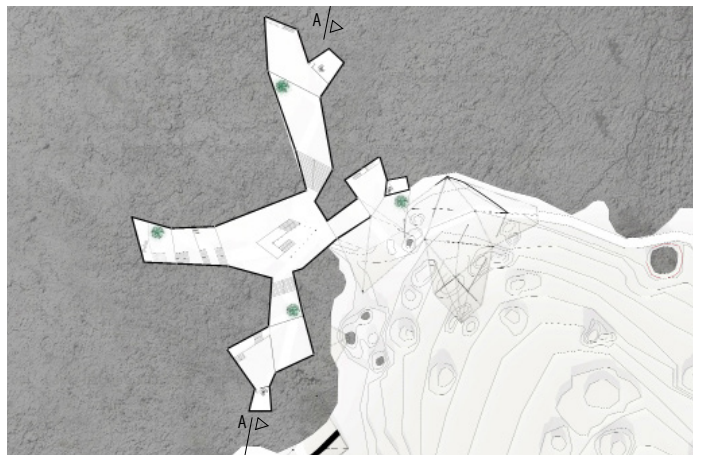
Figure 13. Underground Plan

Figure 12. Mass plan of Adaptive Landforms

REFERENCES

[1]   "Fairy Chimney Development in Cappadocian Ignimbrites" by M. Naci Sayin, 2008.
[2]   Cookson L., (28 July 2007). 'Away with the   fairies: Cappadocia is full of magical natural attractions' Newspaper
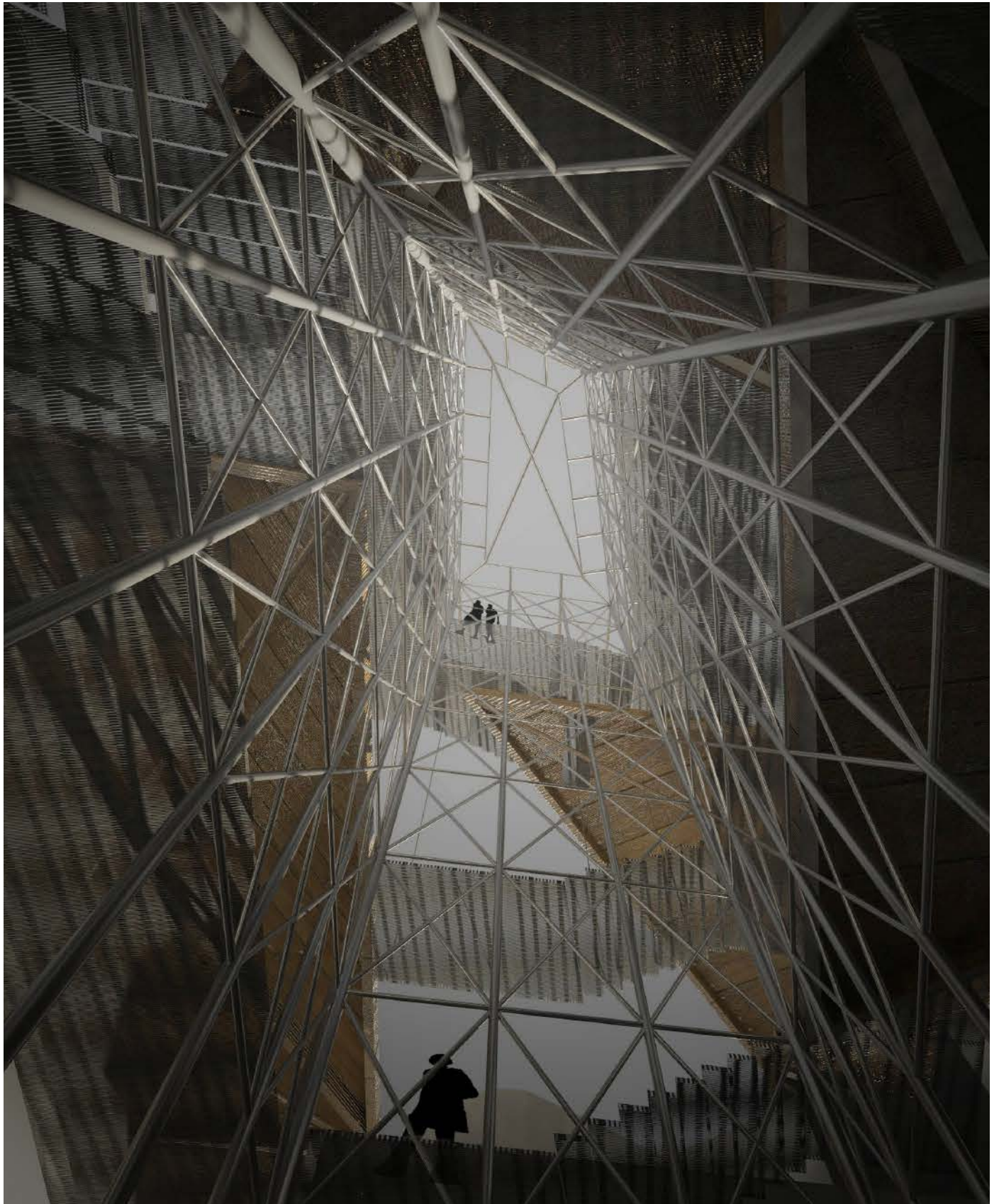
Figure 14. Perspective view from connecting skeleton

Archupation: *Arch*itecture as the Tool of Land Occ*upation*

Magdaleen H Bahour
Department of Architecture and Design
American University of Beirut
Beirut, Lebanon
mhb17@mail.aub.edu

Abstract- "Architecture is employed as a conceptual wall of understanding political issues as constructed realities". This has led to the development of the term archupation.

This research was carried out to identify how architecture was utilized as "one of the direct instruments of occupation" in the Israeli-Palestinian conflict, by referring to Eyal Weziman's Hollow Land: Israel's Architecture of Occupation as a main reference.

The main concept behind the architectural occupation used by Israel against the Palestinians is the enforcement of power embodied through architecture and spatial elements, utilizing it to inflict oppression and expropriate land. It is manifested in four main structures. First, the Separation/Apartheid Wall, which represents territorial claim and separation of land; it acts as a solid barrier between the Occupied Territories and the West Bank and Gaza Strip. It also creates a fragmentation on the landscape with its intense solid presence. Second, the watchtowers, which act as a surveillance element. It allows the Israeli entity to watch over the Palestinian lands and civilian movement. They are usually found dispersed on the roads, near military zones, and/or attached to the Separation Wall. The watchtower has become a prototype that easily duplicates on the landscape. Third, the checkpoints, which have more spatial qualities than the watchtower and the wall. They act as a border or a 'terminal', to enforce control on civilians transporting from an area to another. Fourth, the settlements, which are the most imposing on the landscape. Not only do they claim large areas of land, but also their presence on hilltops overlooking Palestinian territories has permitted them to become civilian surveillance mechanisms.

Therefore, the research carried out tackles each of the mentioned elements of archupation and analyzes its impact on the landscape, the people and the land, illustrated and investigated.

INTRODUCTION

"Architecture formed a visual language that was used to blur the facts of occupation and sustain territorial claims of expansion." [1]

1948: A year that marks Arab world history, only a couple of years after the end of World War II in 1945, and when the rest of the world was healing from all the losses and damages of the war, Palestine was going under its own political conflict, not knowing that until today in **2016**, its land is being scarred day after day.

In the midst of the political encounter for the right of land in the Israeli-Palestinian war, architecture became the main tool used in the conflict to construct political principles into tangible structures on the landscape. [2]

Elements of architecture and construction became the language forced on the landscape to expropriate land and manifest power and control through the building industry. These elements are defined as elements of '*archupation*', putting together the words *arch*itecture and occ*upation* [3], and they are as follows: the Separation Wall, hilltop settlements, watchtowers, and checkpoints, infrastructural blockades such as road barriers and closures, military outposts, and special security zones.

The main concentration in this paper will be to identify the four most imposing elements (the wall, watchtower, checkpoint, and hilltop settlements), their implications, the concepts that were infused in their construction in order for them to act as tools of power and occupation against the Palestinians, and how their presence on the land has given these mundane elements of construction the supremacy of control.

I. THE ELEMENTS OF ARCHUPATION: IDENTIFIED

The elements of architectural occupation referred to are all constructed within the confinement of the West Bank Palestinian Territories, where "*space becomes the material embodiment of a matrix of forces, manifested across the landscape in the construction of roads, hilltop settlements, development towns and garden suburbs*" [4]

A. THE SEPARATION WALL

The wall in this context is a representation of territorial claim and separation of land within the land. It acts as a solid barrier between the Occupied Territories by Israel and the West Bank and Gaza Strip. The Wall reaches 8 meters in height, and by the end of its construction it will cover a span of 723 kilometers.
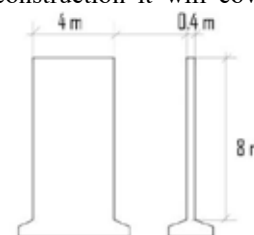

Figure 1. The dimensions of the unit used in the construction of the wall

Israel has regularly confiscated large plots of Palestinian land in order to build the wall. When the barrier is complete, most of it would have been built inside the occupied West Bank. The construction of the separation barrier began in 2002, and the route of it is around the West Bank where the 'Green Line' of 1948 has been drawn. Shown below is a comparison between the Berlin Wall, famously known for being a separating barrier in between East Berlin and West Berlin, and the Apartheid/Separation wall in Palestine.
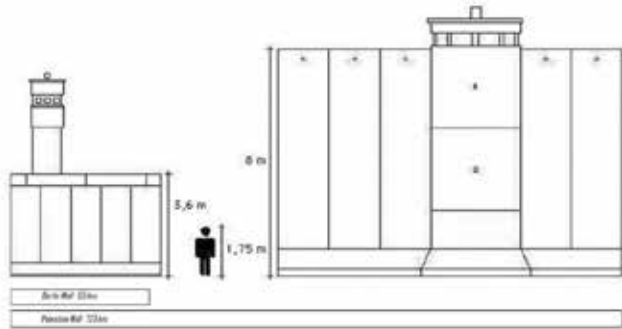
Figure 2. A comparison between the Berlin Wall and the Israeli Apartheid Wall

Israel's Separation Wall became a new global icon for oppression. Its trajectory, confining and oppressing the Palestinians within the West Bank, has led to restraining the people within their own rightful land.

Figure 3. Highlighted in red is the trajectory of the Separation wall around the West Bank Territories. Built and under-construction

Not only did the wall become an element of oppression, power and control, but also it has lost its traditional concept and material fixity when it comes to the contact between soldiers and the walls. The wall has been transformed into: (i)

a **transparent medium** through which soldiers can now see through which they can now shoot; (ii) **permeable elements** in which resistance and security forces travel through; and (iii) **flexible entities** responsive to changing political and security environments. [5]

transparent mediums   permeable elements   flexible entities

The following set of images show the different instances at which the wall is imposed on the landscape, expropriating land, separating neighborhoods and fencing around hilltop settlements as a 'measure of security'.

B. Military Zones: of which are Watchtowers

The watchtower acts as a surveillance element. It allows the Israeli entity to watch over the Palestinian lands and civilian movement. They are usually found dispersed on the roads near military zones, and/or attached to segments of the Separation

Wall. The watchtower has become a prototype that easily duplicates on the landscape.

The watchtower is thought to be the origin, the prototype, the model, and the mold of Israeli architecture. [6]

The watchtower is the ultimate machine of invasion. It becomes easily and freely multiplied and placed on the landscape, where it becomes an instrument rather than a space [7]. It always displays an image of a 'work in progress' allowing it to become a translation of a political agenda into the act of constructed realities.

Surveillance of 360 degrees   Creates a sterile radius
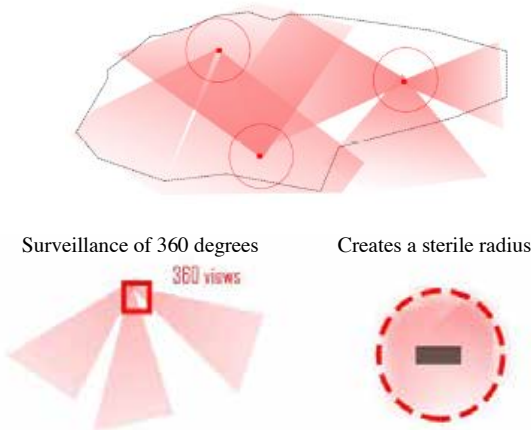
Figure 4. Conceptualizing the watchtower, where it multiplies on the landscape to become a 360 degree surveillance mechanism, which creating a sterile area around it, therefore extending the radius of expropriated land on which the watchtower sits.

There are three main distinct types of the watchtower as follows:

Type One: road-side watchtower
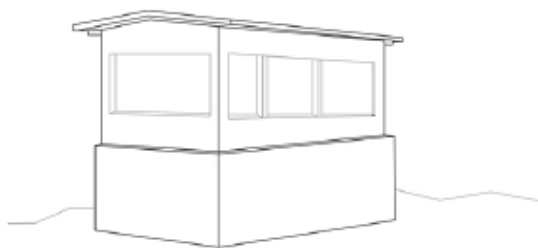
This type of the watchtower is usually found on road side locations, accompanied by soldiers on foot. It is sometimes present on hilltops along with a military outpost. This typology is not raised from the ground and is constructed out of concrete in most cases and is complemented by road blockades and barriers.

Type Two: raised watchtower
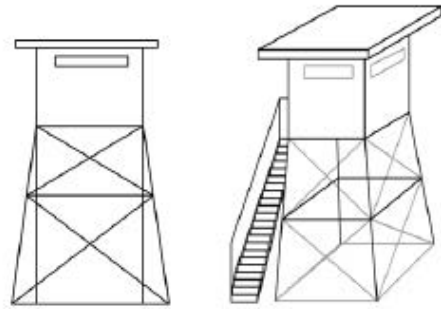
This typology of the watchtower is more common. It is constructed out of light-weight structures using steel and/or wood. It is raised above the ground and the soldiers do not need to be on the ground/road, as they remain inside the compartments positioned on the top of the structure, accessible through the staircase built with the tower. The compartment could be cylindrical or rectangular.

Type Three: cylindrical watchtower

The third type of the watchtower is usually found along the separation wall and/or stand-alone on the road. It is constructed out of concrete and the final finish of it is left bare, with no ornamentation. Although the watch compartment is situated on the top most area of the watchtower, similar to type two, however the vertical circulation is within the built space, not on the exterior as in type two.

Figure 5. Excerpt of images from Taysir Batinji's 2008 Project Watchtower

C. INFRASTRUCTURE AND CHECKPOINTS

Checkpoints have more spatial qualities than the watchtower and the wall. They act as a border or a 'terminal' in which civilians are forced to cross through to travel from point A to reach point B. The main function of the checkpoint is to enforce control on Palestinians transporting from an area to another, and to impose surveillance.

The checkpoint is also viewed as a *barrier* in the circulation and infrastructural road networks. It represents a turning point in the public space, where it becomes a space of its own, expanding continuously and affecting areas on its margins.

Translating from the narratives of Azmi Bshara in his book '*Al-Hajez*': "*The checkpoint spaces inhabit functions that expand outside the borders of 'security', such as food/drink stalls that service the 'users' passing through the area*" [7]

Checkpoints have been modified to function as a form of control, controlling Palestinian civilians and 'rupturing the social and economic fabric in which they live' [8]
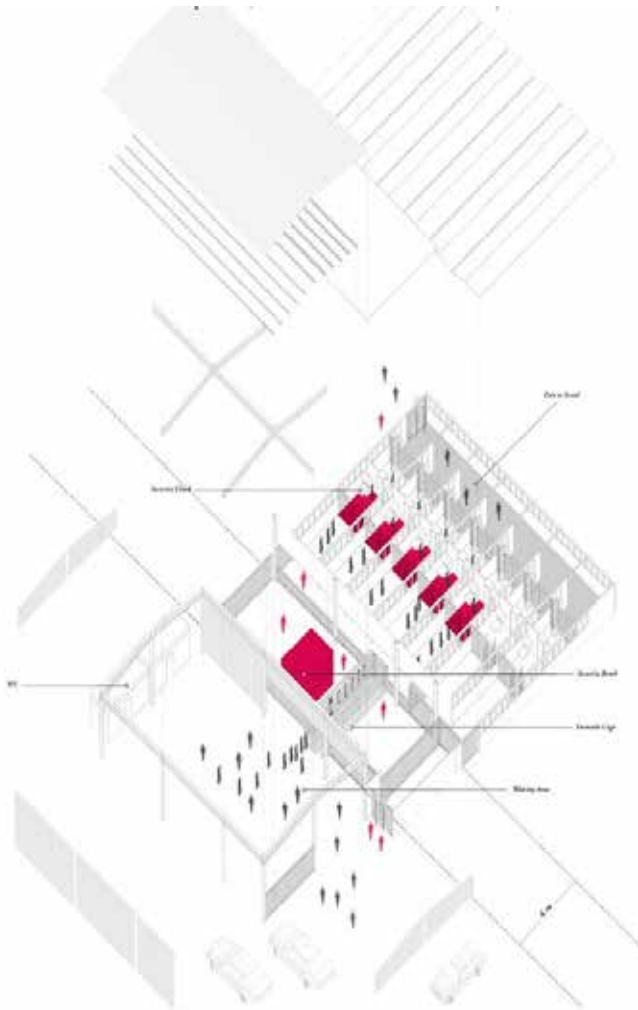

Figure 6. The Qalandia Checkpoint – exploded axonometric to show the different screening in which civilians have to go through to cross from the West Bank to Jerusalem District
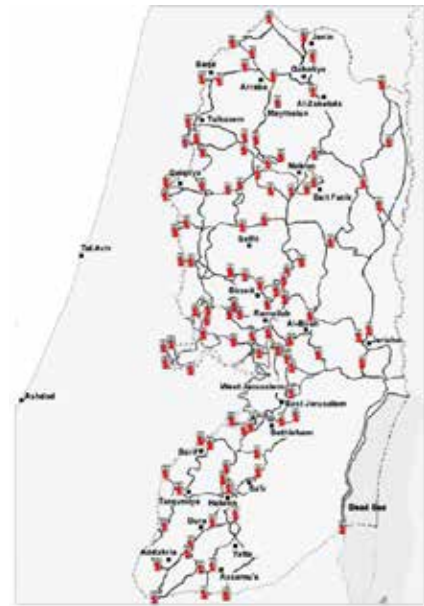
Figure 7. Shown is part of the map of Palestine, with the border of the West Bank Territory dotted. The red markers represent the checkpoints within the WB and on the borders to cross to the Occupied Territories

In addition to the disruption in the circulation due to the presence of checkpoints, the road networks have also been subjected to segregation. This restriction in accessibility has led to the creation of 'enclaves' that geographically separate Palestinian communities. Hence, checkpoints are mostly situated at the entrance and exit to these generated enclaves.[9]

The scheme of borders, enclosures, fences, checkpoints and controlled roads within the West Bank have developed a series of delay and interruption in the trajectory of travel of civilians. As an example, the exclusion and isolation of specific road networks and the dispersal of checkpoints cause a Palestinian civilian to stop and change mode of transportation multiple times to reach their destination, causing them time loss, whereas for an Israeli civilian, they just by-pass all these checkpoints and access more roads for a quicker and smoother circulation trajectory. For the same distance, from point A to point B, the Palestinian would spend almost 4 times the travel period of that of the Israeli. [10]

Completely segregated road   Blockades and Checkpoints
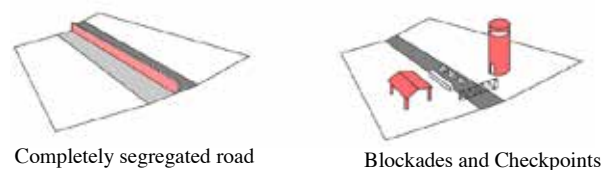
Figure 8. Diagrammatic representation of the obstacles in the road network

D. SETTLEMENTS

The settlements are the most imposing architecturally on the landscape. Not only do they claim large areas of land, but also their presence on hilltops overlooking Palestinian built areas has permitted them to become civilian surveillance mechanisms. They are characterized with modular and rigid construction, accompanied by the red-pitched roofs, which contrast the flat roofs of the Palestinians.

The strategic concepts of building these settlements can be explained and represented as follows:

See and be seen – *"settlements are locted on strategic summits, thereby allowing them to function as observation points: maintaining visual connection with each other and their surroundings, main traffic arteries, strategic road junctions, and Palestinian cities, towns and villages."* [11]
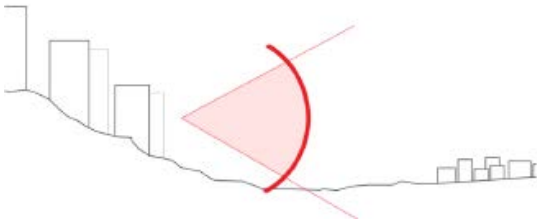
Claiming the landscape – *"Move, run, and grab as many hilltops as you can, everything we take now will stay ours. Everything we don't grab will go to them."* [Ariel Sharon, 1998]
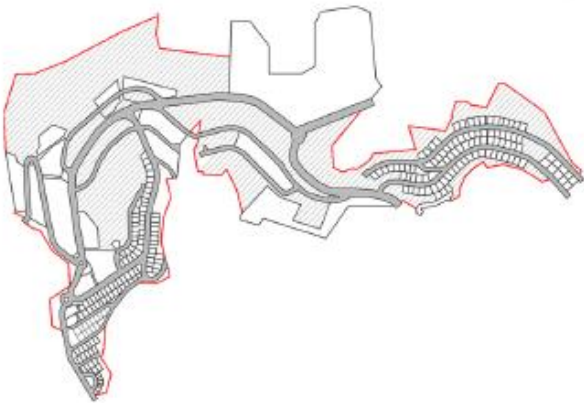
Planning and design – the modular, rigid construction of the settlements versus the organic models of the Palestinian built cities, towns and villages.
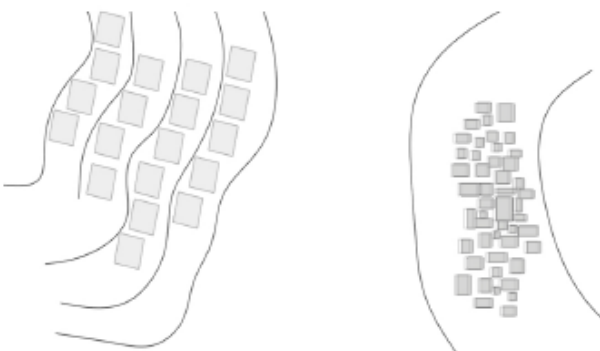
Figure 9. Israeli Settlements within the West Bank Territory highlighted as red dots.

In addition to these four main elements of *archupation*, the concept of 'Politics of Verticality and the Six Dimensions' is essential to the architectural occupation scheme.

This concept aims at finding a solution to the problem of connecting two entities separated by land and how to resolve the circulation issue.

This introduces the 3 Palestinian dimensions and the 3 Israeli dimensions, which are completely distinct, with no contact between the two, resulting in 6 dimensions. The scheme created is similar to the inbound/outbound terminals at an airport. They are represented diagrammatically as follows:
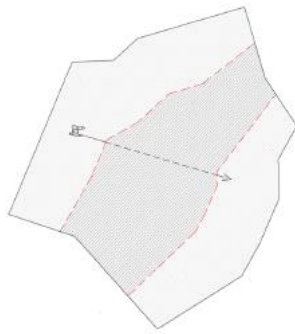
Diagrammatic plan
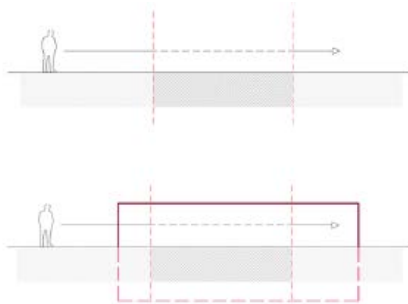
Diagrammatic section

II. THE ELEMENTS OF ARCHUPATION: RESULT?

These elements of architectural occupation have led to the disintegration of the spaces on the landscape. It created a segregated nation, with all the borders, walls, checkpoints, horizontal and vertical circulation routes, and utilizing construction and planning to zone and expropriate lands.

The impact? *A Palestinian archipelago.*

III. ARCHUPATION: WHAT NOW?

Understanding the approaches and techniques of the architecture of occupation has allowed me to build a set of design strategies, and to use these elements of archupation as design tools in my final thesis. Their significance in my research was to recognize the concepts behind each element and how it was utilized, in order to reverse and intervene on each in a hypothetical situation.

The culmination of this research lies in my concept of how to challenge this architecture and how to *'deny'* it its essence, stripping down its functions of power and control, and designing an architecture that reverses all the models of archupation.

ACKNOWLEDGMENT

REFERENCES

[1]   Weizman, Eyal (2007). *Hollow Land: Israel's Architecture of Occupation*. Page 26
[2]   Weizman, Eyal (2007). *Hollow Land: Israel's Architecture of Occupation*. Page 6
[3]   Weizman, Eyal (2007)
[4]   Weizman, Eyal and Segal, Rafi (2003). *Civilian Occupation*. Page 19
[5]   Weizman, Eyal (2005). *Walking Through Walls*
[6]   Rotbard (2003). *Civilian Occupation*. Page 46-51
[7]   Bshara, Azmi (2004). *Al Hajez*. Page 20
[8]   Bshara, Azmi (2004). *Al Hajez*. Page 27
[9]   UN Office for the Coordination of Humanitarian Affairs, CAP 2008
[10]   Boeri, Stefano et al (2003). *The Road Map, a project of Municipality*.
[11]   Weizman, Eyal (2007). *Hollow Land: Israel's Architecture of Occupation*. Page 81

FIGURES

Fig 1. The dimensions of the unit of the wall
Fig 2. A comparison between the Berlin Wall and the Israeli Apartheid Wall (2012). Retrieved from occupiedpalestine.wordpress.com
Fig 3. West Bank Wall – Map 2006. Published Dec 2011 by Council for European Palestinian Relations
Fig 4. Conceptualizing the watchtower – personal interpretation
Fig 5. Excerpt of images from Taysir Batinji's 2008 Project: *Watchtowers*
Fig 6. Axonometric of Qalandia Checkpoint. Retrieved from Kyle Bryer Waterloo Architecture Thesis
Fig 7. Israeli Checkpoints in WB. Published by Palestinian Economic Council for Development and Reconstruction
Fig 8. Diagrammatic representation of the obstacles in the road network
Fig 9. Map of Israeli Settlements in the WB. Retrieved from BBC Key Maps of Israel and the Palestinians.

Can Arabic Calligraphy Move Us?

Riham El Ghosseini
Department of Architecture and Design
American University of Beirut
rihamgh@gmail.com

Introduction

Calligraphy, the art of writing the Arabic script has played an important role in the Islamic culture. Through the Arabic script, a sacred and timeless message could be carried, that is the divine word of God. This paper will suggest an investigation of another relevance of Arabic calligraphy, one that allows calligraphy to be experienced emotionally rather than understood intellectually.

The term emotion implied here does not represent a conscious awareness by the person of their feelings, but rather it is an effortless experience that occurs and resonates naturally within human beings, it is a state of well-being.

This paper will investigate existing projects under visual and theoretical frameworks adopted from Islamic, Chinese and Japanese philosophies. While looking at these case studies, the paper will break down Arabic calligraphy to and at some points challenge its main constituents: the medium, the calligrapher, and the forms produced as well as the process of receiving it by an observer from the other end. As such, I will study: The Ink, The Hand, The Sign, and The Eye.

I The Ink

Calligraphy Extends through Time and Place Immortally

Intangible and infinite ideas can be expressed and preserved through a finite medium, that is calligraphy. Writers and receivers of Arabic calligraphy can connect with the past, the present and the future; The past, by engaging in a virtual conversation with the inventors of the script or with the reciters who relied on oral practices to transmit the Qur'an [1], the present, by experiencing an unfolding of the secrets of life as it occurs now, and the future, by the ink that can be reached by those who will come after. All are in sync with the infinite creator, that, similar to calligraphy, extends through time and place.

Rather than considering the ink as the calligrapher's tool, this paper aims to suggest that the calligrapher becomes the tool of which the ink manifests itself. Thus, the ink has the power to give eternity to calligraphers who got immersed in the act of repeating the same forms over and over again in order to master their skill of calligraphy, making it is a durable medium that can outlive the mortal lives of its calligraphers. Through this ink, they can become timeless and everlasting, and will live through their written words that remain alive as long as life exists.

Now what happens when one challenges this immortal quality of the ink while producing calligraphy? What occurs experientially when ephemeral artifacts are employed instead? Buddhists created sand mandalas with the intention of dismantling them. What does it take out and what does it add? Examples of ephemera include snow, ice and even fire in sculpture. For the sake of this study, we will investigate two examples of ephemeral calligraphy: water and light.

Ephemera and Immortality (Water as Ink)

In Chinese cosmogony, the square or 'dì' represents the earth and 'shū' signifies writing. The expression 'dìshū' translates to 'earth calligraphy'. Dìshū (地書) is a Chinese practice of ephemeral calligraphy using street pavement as its canvas and clear water as ink [2]. Anonymous calligraphers use hand-made brushes of sticks and sponges to write fluid lines of ancient Chinese characters that slowly disappear one by one as the water evaporates (Figure 1). Displaying literature, poetry or aphorisms, these monumental letterings experience infinite renewals.

The appreciation of the impermanent and the incomplete are two main concepts adopted in the Japanese philosophy, Wabi Sabi. Wabi Sabi is based on the observation of and learning from nature. It is the appreciation of the minor and the hidden, the tentative and the ephemeral. Wabi Sabi acknowledges that everything eventually fades into oblivion and nonexistence: the planets and stars (even in digital form) [3]. Thus, the spiritual values of Wabi Sabi challenge the view of ink as immortal, by arguing that ink presents an illusion of permanence while in fact everything wears down and comes to an end. In Wabi Sabi, the closer things get to nonexistence, the more exquisite and evocative they become and their effect becomes more potent and profound.

Dìshū brings together an element from nature (water) and artistic creativity (calligraphy) allowing the calligrapher and the observer to experience nature and art working as one. By embracing the water performing in its natural behavior, the calligrapher can contemplate on his own state of being and accept his inevitable mortality. This intellectual way of thinking

celebrates that he is nothing more or less than the water he is using. This is how the practice of Dishu has the potential in becoming a meditative, self-realizing act that can modify one's perception of the world.

Ephemera and Immortality (Light as Ink)

Using another element in nature, light, the French artist, Julien Breton carefully executes rehearsed motions that translate through photography to Arabic calligraphic strokes. The photos translate his existence by the light strokes that imply the presence of traces of his body movements. Presented to the audience as a still image, the photograph documents a busy and rich process. This process is a combination of many moments that existed and will never exist in the same way again. Even though the light strokes are ephemeral in nature, the immortality of the work and the artist is preserved through the technology of a camera. In its spatial illusionism and meticulous detail, the photograph inevitably points to a world outside of itself, a world of visual forms and sensations, always reminding us of its origin in it.

• • •

In such ways, artists can directly experience a relationship between themselves, their creations and the passage of time. Each medium calls attention to its relationship with the concept of time, whether through immortality or temporality. There is beauty and value in both. The appreciation of the idea of an 'immortal ink' reflects the human perseverance to outlive and resist death by creating calligraphy that can extend through time. My project will adopt the second relationship, by humbly accepting the impermanent and the natural. It will focus on the appreciation of an inevitable moment that will occur once and will never exist the same way again.

Figure 1. street Chinese calligraphy performance in Beijing using clear water as ink

Ibrahim ash-Shaybani (d.805) stated that calligraphy is 'the language of the hand". According to Schimmel, in order to successfully translate this language, a good calligrapher requires five things: the necessary utensils, a good hand, an understanding of calligraphy, a fine temperament, and an endurance of pain [1].

Significance of the Purity of the Calligrapher

To write the Qur'an in a worthy style was always the highest goal for calligraphers. The sacredness of the Qur'anic texts demanded the calligrapher 'not be unclean for a single hour' for 'purity of writing is purity of the soul'. Twenty-five hundred years ago Hippocrates acknowledged in his treatise Airs, Waters, Places that inner conditions require attention to the outer [4]. Inner and outer purity are interrelated, and are both equally as important in the field of calligraphy. It is of no surprise that calligraphers practiced ablution (typically performed before a prayer) prior to beginning their work [1]. Mystics considered letters as a veil between themselves and the immediate experience of the Divine, for which the mind and the heart have to be like a blank page, which is why they loved to dwell on the illiteracy of Prophet Mohammad who is called in the Qur'an 'ummi' or the one who needs no learning [1].

Immersion of the Calligrapher is a Spiritual Experience

Mir Ali stated,"Forty years of my life, were spent in calligraphy... If one sits leisurely for a moment without practicing, calligraphy goes from his hand like the color of henna" [1]. Mastering calligraphy requires a period of persistent practicing. Schimmel mentions the story of a calligrapher in Tabriz who did not notice an earthquake because he was so engrossed in producing a flawless 'waw'. What occurred experientially with this calligrapher? The focus needed while practicing calligraphy is similar to the focus needed on the breath during traditional meditation practices. Can creating things thus be a meditative act? Modern science shows us that such activities offer cues to the body-mind, lowering brain-wave frequencies, can resulting in absorption and meditation [4].

Calligraphy is an activity where mind and body work harmoniously. It demands focus on the letters formed by the hand as well as those seen in the mind. The Turkish calligrapher, Shaykh Hamdullah (d. 1519), described his practice as a mystical exercise that strengthened both his eye and mind [1]. Practicing calligraphy can be considered a form of creative meditation because it demands an effort from the right hemisphere (creativity, imagination) and the left one (mathematical study, language) which restores an inner balance in both brain hemispheres [4].

The Active Role of the 'Thinking Hand' in the Crafts

A 'craft' is skillful activity that involves making something by using one's hands. In the craft of sculpture, the hand acts as the sculptor's eyes as well as organ for thought. The active role of the hand in the crafts is emphasized by several philosophers such as the German philosopher, Martin Heidegger (d. 1976) in his writings in "What Is Called Thinking?". The human hand for Heidegger is more than an organ that can grasp, 'every motion of the hand in every one of its works carries itself through the elements of thinking' (Heidegger). The significance of the hand lies within its view as an organ with its own intentionality, and not as a passive tool to execute the intentions of the brain. The hand has knowledge and skills and is able to perform actively and make creative decisions. Defined as the 'Thinking Hand' by Pallasmaa, the 'thinking' process of the hand differs from that of the brain. The hand thinks through 'the realm of tacit, unconscious, internalized and embodied knowledge rather than abstract, conceptual thought processes' [5]. Note Abu Hayyan at-Tawhidi's (d.1023) remark, "Ibn Muqla is a prophet in the field of handwriting; it was poured upon his hand" [1]. Ibn Muqla had reached mastery in his field until it became natural to his hand to perform calligraphy. It is of no irony that Ibn Muqla's enemies had cut off his right hand.

Challenging the 'Thinking Hand' with Technology

In the context of his book, Eyes of the Skin, Juhani Pallasmaa expresses a concern towards computer imaging and its ability to flatten our "magnificent, multi-sensory, simultaneous and synchronic capacities of imagination by turning the design process into a passive visual manipulation, a retinal journey" [5]. Hand crafting can put the designer into a haptic contact with the object or space. In our imagination, the object is simultaneously held in the hand and inside the head, and the imagined and projected physical image is modelled by our bodies. One must note that an absolute rejection of technology in the field of design is not a practical decision. Rather, an awareness of the limitations that come with the machine must be taken into consideration while acknowledging the computer's potential to facilitate efficient design production.

The Body is the Tool

The technique and scale of Breton dynamic light compositions imply that the calligrapher was engaged in full-body gestural movements, which he turned into strokes (Figure 2). His bodily involvement (the movements of his fingertips, wrists...) becomes an active tool of a spacial and haptic exercise, one which fuses the external reality of space and matter with the internal reality of perception and thought. The art becomes unconsciously internalized in the artist's body. If one is to trace the body movements of the calligrapher while performing the practice, one can see them reminiscent of a form of martial arts. Martial arts are considered 'meditation in motion' which promote serenity and inner peace [6]. "Creative arts and traditional

Figure 2. Kalaam: Light Calli-graffitti by French artist Julien Bretong

meditative practices both demand focus, and they rely upon repeated physical actions" [4]. This intimacy with the art, allows the calligrapher to experience an internal state of awareness close to a meditative one.

A live performance of Shodo, or the art of Japanese Calligraphy, is an example of the involvement of the hand as it extends with the brush from the surface to depart onto another location. Here, the act of writing becomes a performative action, where the ink strokes are traces of moments of the calligrapher's bodily involvement (Figure 3). There is no hesitation, nor hierarchy. The calligrapher seems to naturally perform the calligraphy in sync with live music in the background, as he unites with the brush, the ink and the art piece. The subject becomes the object as boundaries between the inner and the outer are dissoluted. The calligrapher ends his performance by painting his face with the same brush strokes that he used on the art piece, in order to sign it with traces of his own face. This is another example of how the body becomes the media, the tool and the art.

●●●

The discussed examples highlight how the human body is suited to transcribing an emotional impression of the written language. The concept of the 'thinking hand' allows an appreciation of the human body as a tool with authority. Through bodily identification, empathy and compassion, the artist develops an awareness not only of the creative process but also of processes happening within his own body. In the case of calligraphy, through such an awareness, the calligrapher becomes his art. Through my project, I aim to allow my audience to act as active participants in the practice of calligraphy rather than passive receivers of it in a way that demands the involvement of their whole body.

Figure 3: snapshot from Shodo - Calligraphy live painting performance.

III The Sign

After exploring the ink and the human body as valuable tools in calligraphy, we will now explore what is born out of the two: the letters and their meaning. Through the study of signs, or semiology, this chapter will explore the different frames of meaning that can emerge through calligraphy. Is the letter an absolute reality or an interpretive relational concept?

Semiology of Linguistics (What is Arabic calligraphy?)

1. Forms as signifiers for letters: The sign in the pre-modern calligraphy was based on a relationship between a form that acts as a 'signifier' to a 'signified' letter, sound and meaning.

2. Letters as signifiers for faith: The letter "waw" is valuable for being a recurrent conjunction representative of faith in Islam, namely the "belief in God, AND his Books, AND his Messengers..." So, the sequence of "waws" can be read as symbolizing the connotations.

3. Letters as signifiers for imagery: Poetry has been used throughout centuries as an expressive form of literature. The aesthetic nature of the Arabic letters acts as a signifier to rationally unrelated concepts. Praise of beautiful script is common in the sources of the 9th and 10th centuries. When Ismail al-Katib saw good calligraphy, he expressed:

> "If it were a plant, it would be a rose
> if it were metal, it would be pure gold
> if it were something to taste, it would be sweet
> and if it were wine, it would be very pure."

[1]

5. Letters as signs of divine activity: Hurufism of the 14th - 15th centuries, was a Sufi doctrine that believed God is incarnated in every atom in the universe, which is manifested in sounds. The perfect form of uttering a sound is the word. Thus, Hurufists announced the letters as the foundations of all things, and in them saw the embodiment of the divine, love and beauty.

4. Letters as means to blessings: It was believed that with every stroke the calligraphers drew, 'baraka' came upon them. As the Hadith states: "He who writes the basmala beautifully obtains innumerable blessings". By copying the Mathnawi the calligrapher felt that he was receiving spiritual medicine for his heart's pain [1]. This is how the Arabic letters became considered vessels for the revelation and forms for a meditation for the divine.

5. Forms as pure traces: A modern Hurufi Art movement, starting from the fifties onwards, is an art movement that deployed Arabic letters as a source of inspiration. One can suggest that Hurufi movement is a transition from the traditional to the modern. The letters functionality as signifiers of literal meanings or of divine activity was dropped. At first glance, these recognizable forms seem to be signifiers towards the Arabic culture or the Islamic religion due to the nature of their strokes (curves, direction, dots) but they do not spell out anything specific nor do they aim to deliver a literal message. They are purely signs, traces, forms... Writing is mainly means for comprehension, clarification, transmission and delivery. In Hurufism, Arabic calligraphy focuses on the aesthetics of the form through a state of contemplation upon its inherent qualities. Form is a self-sufficient content that overrides text and its meaning. It initiates a dialogue of a language without a single letter, but rather one that is derived from the "alphabet of graphic forms" as described beautifully by Samir Mahmoud. This dialogue emerges from the line qualities, whether curved or straight, extended or cut, connected or detached, and from its playfulness as it comes closer to one another, intersects, or interchange.

Life through an Abstraction of the Letters

Samir Sayegh (1945) is an artist, critic, poet and historian. He is considered as an Avant-Garde Arab master calligrapher and as a pioneer of modernism in the Arab world. Sayegh, referred to by many of his students as the first to stir up attention towards the Arabic calligraphy in our region, considers the historical and cultural significance of Arabic calligraphy as important as calligraphy in Japan/China, sculpting in Greece and Renaissance paintings in Italy. Sayegh described that an issue with Arab civilization began with the introduction of the printing press in the 18th century by the Ottoman Empire, and the domination of typography by the Europeans. The master calligrapher who used to be admired as a man close to the divine, became considered merely a teacher of good handwriting.

His Arabic letters break the formal aesthetics of Arabic calligraphy as they present themselves as universal forms (Figure 4). They should be read through the lines in their nudity and their simplicity, in the straight and cursive lines, in the vertical and the horizontal, in the long and the short, in the thick and the thin, in the sharp and the relaxed, in the attached and the detached, in the closed and the open, in the ascending and the descending, and in the reclined and the flat. A universal visual

Figure 4. Works by Samir Sayegh

language is employed, one that emerges from the relationship between the lines, speaking to all observers similarly, regardless of their backgrounds, origins or knowledge. It is a language that talks to the eye, the body, and to the feeling heart rather than the thinking mind. For Sayegh, "A contemplation of the letter alif for instance unveils the stunning multiplicity of the possibilities derived from the straight line, especially when this straightness goes astray! With some imagination, this alif starts to move and dance, paving the way to various interpretations, as many as the eye can foresee. No wonder one can perceive the alif as a falling rain, a staggering stallion, a fruit-bearing branch or a razor-sharp sword. For Sayegh, the abstraction of the black and white squares in the quadrangular Kufic script represents the script's polarities and dualities: fullness and void, visible and invisible, vertical and horizontal shadow and light, body and soul, day and night, absence and presence and endless other possibilities through their language of universality [7].

● ● ●

My project will focus on looking at abstracted calligraphy letter forms through to an active intellectual process that can be inherently appreciated by the heart rather than understood by the mind.

IV The Eye

There are multiple ways of experiencing the world, and a lot of emphasis has been placed upon the five human senses in perception, especially that of seeing. This chapter will investigate the moment the eyes witness the calligraphy, by developing a better understanding of what occurs experientially while the brain and body are processsng it. It will raise the question: Can Arabic calligraphy move us?

How Can a Line Affect the Beholder's Sense of Well-Being?

Vischer (d.1933) suggests that we can react emotionally to perceived lines. He speculated that a horizontal line might be pleasing because it conforms to the structure of our visual apparatus, whereas a diagonal line is less so because it requires an awkward movement of the eye. A line with a gentle arc is more pleasing than a jagged one because of the 'congenial' nerve movements it induces, whereas a form displaying regularity is a happy one because it mimics our own corporeal regularity. Vischer also developed a notion of 'empathy': certain proportions can be harmonious because they are innately sympathetic to how the visual cortex, processes the perceived images [8]. Our artistic enjoyment is in large part conditioned by the complexity of our sensory, emotional, and intellectual (neural) patterns while viewing the art. William Hogarth stated about the 'line of beauty': "This curvilinear abstract line has an implicit movement that elicits and engages a beholder's creative acts of perception and thus his imagination. The 'serpentine line' comes to life waving and winding, flirting with the imagination, calling on it pleasurable even as the line itself becomes invisible. The space thereby created by the movement of the line, haptic space, is a space that does not invite a distant contemplation but intimate involvement; the eyes re-invited to move over the surface as though physically touching it through a haptic gaze" [9]. In a 19th century physiological sense, Vischer noted that 'the whole body is involved; the entire physical being is moved. Thus each empathetic sensation leads to a strengthening or a weakening of the general vital sensation.' The vital sensation here is our sense of life. Whereas, in a very modern understanding of physiology - certain sensations have an inhibitory effect on our nerves and muscles, while other enhance our vital sense of well-being [8]. Forms are pleasing when they intensify our biological life, by mirroring the perciever's own neurological properties.

Embodied Simulation of a Perceived Line

Lipps observed that when we look at a line, we 'create' the line with our bodies by following it from point to point with our eyes. It is more than just a process of contemplation but a feeling of being identical with the line by performing it and contributing to its existence. The beholder's eyes not only capture relevant data about shape, direction, and texture of the cut strokes, but by means of embodied simulation, they emulate the actual motor expression the artist used when creating these strokes. "The sensory-motor component of our image perception, together with the jointly evoked emotional reaction, allows the beholders to feel the artwork in an embodied manner." [5] . Neuroscience has been confirming many of the findings of the 19th century theorists of empathy. The movement or image seen is unconsciously simulated by the viewer (EEG study). Mirror Neurons provide new evidence of inter-subjectivity. It is triggered during spatial experiences and during the contemplation of objects [5] .

Expansion of the Self through an Abstraction of "HIMMA"

Joumana Medlej based her hand-crafted piece on Henry Corbin's view of 'himma'. Henry Corbin [1903 - 1978] was a philosopher, theologian and a professor of Islamic Studies in Paris. He argued that the imagination is the primary means to engage with creation and prayer is the supreme act of the

Figure 5. Himma by Joumana Medlej

creative imagination. For Corbin, himma is the power of the heart, spiritual aspiration and the creative energy it generates by its natural act of meditating, conceiving, imagining, projecting, and desiring. The heart creates by "causing to appear" what already exists. Similarly, Medlej is not the "author" of Himma, but a "preserver" of which already exists.

The artist adopted characteristics from nature to express the word "Himma". Interestingly, the piece does not resemble a specific object found in nature, but rather resembles the way nature exists and/or grows. The layering is reminiscent of the overlapping feathers of the kingfisher bird, the skin of a tree, or the petals of a flower. The proportions and the symmetry are other examples of existing inspirations from nature. The artists stretched the limits of legibility in her piece and transformed the word into a written and repeated module. An ornament based on repetition is produced, one that is absolute rather than patterned and that is self-regenerating. The repetition and consistency of the interlacing letter forms invite the viewer to experience the rhythmic flow and regularity of the whole composition. The harmony found in multiplicity of these abstract letters reflects the idea of Divine Unity in all things. The centre can be representative of the starting point of creation or the infinite unexpressed cosmos. It signifies unity, the origin. In meditations, concentrating on a point as it draws a circle around itself and expands brings the mind towards abstract concentration which is used as means to self realisation [4]. The viewer projects him/herself into the piece as it elicits his empathetic gaze. The observer can then complete the expansion of the form as it expands through his/her body similarly to the infinite expansion of the universe. Contemplation can unravel the Divine Beauty etched in the immersant's soul. Himma begins in the heart as a desire for God, and as it grows in intensity and concentration, it becomes a creative force capable of manifesting what is usually not visible on this plane. This piece is an example of how abstraction is able to signify what cannot be represented.

Audience Becomes the Performer

The Japanese calligrapher, Kotaro Hachinohe's attached a CCD (charge-coupled device) camera to his brush during a live calligraphy performance and projected it on a large screen. Similar to watching a dance performance, his audience watched the unbroken line of the brush as its danced around the surface. They witnessed its fluctuations, smooth curves, arcs, dramatic turns, and edgy stops as it gently or heavily touched the paper.

The viewers do not witness the art piece as a whole, but rather are immersed in the dynamic and changing brush movements moment per moment. It is about an appreciation of the moment the brush touches the surface of the paper and slides off rather than the final result. Through the haptic gaze, and embodied simulation, the observers are able to touch the surface with their eyes and mimic the line rhythms with their own muscles. They become absorbed with the brush as it floats or flies off the plane and lands in another location. Each beholder experiences it through his/her own body and senses, each experiencing a solitary immersive experience. Boundaries between inner, outer, mind, body, space and time begin to dissolve. Once they realize they cannot control the movements of the brush, an acceptance can occur through a shift of awareness as they let go of their rational urge to control. Thus it is an active but passive relationship (duality again): active through the of the body of the lines perceived, but also passive by the acceptance and awareness of what is inevitably occurring at that moment.

• • •

The theories and examples discussed in this chapter show us how calligraphy can shift the observer's perception through an embodied experience that resonates deeper within his/her physical body. Through bodily identification, the perceiver can see himself in the line gestures, where the subject and object become interdependent. My project will employ a similar approach to the calligraphic letters treatment (repetition, proportions, abstraction) as seen in "Himma" and the creation of an experiential atmosphere as in Hachinohe's live performance, in order to appeal to the observer's bodily intelligence, and dissolve the perceiver/perceived dichotomy.

Figure 5. Himma by Joumana Medlej

Conclusion

Many emotional experiences in life are very difficult to describe, define or express using language. The best we can do is contemplate upon and describe what was felt. Gaston Bachelard states in his book, The Poetics of Space "The poet speaks at the threshold of being" which allows us to suggest that poetry is the closest method of expressing one's awareness of his own existence. Similarly, the Wabi-Sabi state of mind is often communicated and contemplated through poetry, for it allows bringing together rationally unrelated images through larger frames than their literal verbal ones.

If one wants to express strong emotional states visually, poetry can be translated to a language of the abstract. The abstract and non-figurative art forms allow for self-expression because they can be felt intuitively by the heart rather than understood logically by the mind. They open an invitation to imagine by shifting the typical mode of perception to one that resonates deeper within the physical and emotional body in such a way that the immersant and the visible reciprocate one another.

After establishing that emotion could be expressed verbally through poetry and visually through abstraction, what about the dynamics of existence that cannot be written, seen or said? These subtle realms are beyond what our ordinary senses can perceive. As much as one attempts to describe them by words or forms, the only way one can truly understand such a phenomenon is by experiencing it.

Through my project, I aim to communicate intangible states of being by touching upon the moment the body experiences calligraphy through a haptic and embodied simulation. My project can thus bring closer the duality between logic and intuition, mind and emotion, language and experience.

References

[1] Schimmel, Annemarie. Calligraphy and Islamic Culture. New York: New York UP, 1984. Print.

[2] Chastanet, Francois. Dishu: Ground Calligraphy in China. Årsta: Dokument, 2013. Print.

[3] Downing, Jenny. "Wabi-Sabi: For Artists, Designers, Poets & Philosophers." Mindful. N.p., 30 Aug. 2010. Web. 10 Nov. 2015.

[4] Monaghan, Patricia, and Eleanor Viereck. Meditation--the Complete Guide: Techniques from East and West to Calm the Mind, Heal the Body, and Enrich the Spirit. Novato, CA: New World Library, 2011. Print.

[5] Juhani Pallasmaa, and Robinson, Sarah. Mind in Architecture: Neuroscience, Embodiment, and the Future of Design. N.p.: MIT, 2015. Print.

[6] Farrer, Douglas, and John Whalen-Bridge. Martial Arts as Embodied Knowledge: Asian Traditions in a Transnational World. Albany: State U of New York, 2011. Print.

[7] "Encountering Samir Sayegh." Interview by Saleh Barakat. MURAQQA. N.p., Oct. 2009. Web. 10 Oct. 2015. <http://www.muraqqa.com/samir-sayegh.html>.

[8] Mallgrave, Harry Francis. The Architect's Brain: Neuroscience, Creativity, and Architecture. Chichester, West Sussex, U.K.: Wiley-Blackwell, 2010. Print.

[9] Hogarth, William. The Analysis of Beauty: (London 1753). New York: Garland, 1973. Print.

Other Resources

AbiFarès, Huda Smitshuijzen. Arabic Typography: A Comprehensive Sourcebook. London: Saqi, 2001. Print.

"Ancient Chinese Water Calligraphy Gaining Popularity in Beijing." AWAZ.TV. N.p., 11 Oct. 2011. Web. 10 Dec. 2015.

Athr Art. Samir Sayegh Biography. N.p., 2015. Web. 04 Oct. 2015.

Bachelard, Gaston, M. Jolas, and John R. Stilgoe. The Poetics of Space. Beacon, 1994. Print.

Barjeel Art Foundation. Barjeel Art Foundation. United Arab Emirates, n.d. Web. 2 Oct. 2015. <http://www.barjeelartfoundation.org/>.

Barlow, Deborah. "Juhani Pallasmaa Archives." Slow Muse., 6 Sept. 2010. Web. 16 Nov. 2015.

Berlekamp, Persis. Wonder, Image, and Cosmos in Medieval Islam. New Haven: Yale UP, 2011. Print.

Botton, Alain, and John Armstrong. Art as Therapy. Phaeton, 2013. Print.

Dagher, Charbel, Nja Mahdaoui, The Traveler with Pen in Hand: The Passion of the Alif, translated by Samir Mahmoud

Ekhtiar, Maryam D., and Claire Moore. "Art of the Islamic World: A Resource for Educators." Met Museum. N.p., n.d. Web. 03 Oct. 2015.

"EXHIBITION | Writing Non-Writing: Wang Dongling." Ink Studio. N.p., 17 June 2015. Web. 10 Dec. 2015.

Fink, Michael. "Dishu: Chinese Water Calligraphy." The Art of Visual Thinking. N.p., 23 Nov. 2012. Web. 08 Dec. 2015.

Fouladvand, Hengameh. "Mohammed Ehsai's Modernist Explorations in Calligraphic Form and Content." ArteEast. N.p., 2008. Web. 2 Oct. 2015.

Gonzalez, Valerie. Beauty and Islam: Aesthetics in Islamic Art and Architecture. London: L.B. Tauris, 2001.

Haden, Peter. Elementary Knowledge: A Story of the Arabic Alphabet. Geneva: Galerie Patrick Cramer, 1982. Print.

"Interview with Samir Sayegh." Personal interview. 14 Oct. 2015.

Issa, Cynthia. "Prestigious Art Prizes Awarded to Three Lebanese Women." Editorial. The Daily Star [Beirut] 17 Nov. 2015: 16. The Daily Star Newspaper. Web. 3 Oct. 2015.

Kandinsky, Wassily. Concerning the Spiritual in Art. N.p.: n.p., 1912. Dover Publications, 1912. Web. 05 Oct. 2015.

Liberman, Mark. "Writ in Water." Language Log. N.p., 4 July 2014. Web. 10 Dec. 2015.

Moustafa, Ahmed, and Stefan Sperl. The Cosmic Script: Sacred Geometry and the Science of Arabic Penmanship. Print.

Pallasmaa, Juhani. "Body, Mind, and Imagination: The Mental Essence of Architecture."

Pallasmaa, Juhani. The Eyes of the Skin: Architecture and the Senses. Chichester: Wiley-Academy, 2005. Print.

Pallasmaa, Juhani. The Thinking Hand: Existential and Embodied Wisdom in Architecture. Chichester, U.K.: Wiley, 2010. Print.

Saleh, Zaki. Arabic Script.Egypt: Egyptian Public Form for Writing, 1983

The Daily Star (14 June 2013): 2. Ohrstrom, Lysandra. 'Samir Sayegh's Typography Revolution'. The Daily Star Newspaper. Web. 1 Oct. 2015.

"The Legacy of Henry Corbin." The Legacy of Henry Corbin. Web. 1 Aug. 2015.

Welch, Anthony. Calligraphy in the Arts of the Muslim World. US: University of Texas Press 1979.

شاهين ,محمود .الحروفية العربية الهواجس والاشكالات . الهيئة العامةN.p.:
Print. ٢١٠٢ ,للكتاب .

Earth: A House Reborn

Nadine Abdulsalam
Department of Architecture and Design
American University of Beirut
Ras Beirut, Lebanon
nha38@mail.aub.edu

Abstract- **In the light of architectural innovation and cultural exploration, this dissertation aims to investigate how the malleable, yet refined potential of building with earth reflected in the mature works of Egyptian architect Hassan Fathy has survived throughout time to incarnate in the designs of today and the architects of tomorrow. With the use of indigenous materials to develop traditional construction techniques, Fathy experiments with the lessons cultivated from the so-called man-made Gourna "failure" to generate universally-structural applications that consolidate the people's link with their past and traditions to induce a sense of mobilization and ownership. In order to reestablish a connection between the existing earth construction techniques and the demand for tomorrow's local people to build their own shelters, Fathy's architectural concepts begin to be reinterpreted in later works of Pakistani Architect Yasmeen Lari, who aims at redeveloping cities and communities, as a replicable source, in the aftermath of natural and man-made catastrophes. Highlighting precise moments in the practice fields of both architects Hassan Fathy and Yasmeen Lari, this dissertation adds on to illustrate the architecture of emergency and conflict, as a means of reimagining sustainable communities, in a 'traditional future'.**

I. INTRODUCTION

"We have radiations coming from the cosmos, from outside the Universe, and radiations coming from earth. Some are beneficial and some are bad according to the configuration of the earth. Man and trees were created in the very beginning on the surface of the earth.

Trees do not grow downwards. They have roots, I know, but the tree itself grows up into the atmosphere and we are all surrounded by this atmosphere; man, trees, and animals.

The tree is wiser than man, because the tree fans out from its roots to its trunk through its branches to the leaves which spiral out at the correct angle to take the right amount of sunlight and humidity.

Yes, the tree is wiser than man [1]. *"*

- Hassan Fathy, Abiquiu, New Mexico, (1984)

After his return from Greece in 1962. Fathy disrupts the anticipations of many, and sources the presumed Gourna "failure" to venture into the most prolific phase of his career, where he develops a contemporary reinterpretation of the limited traditional forms and materials used in the village, in response to the rootless import of new technologies from the West. Parallel to how a tree grows to absorb "the right amount of sunlight and humidity" without forgetting its roots, Fathy aims at promoting regional customs and styles while adapting to new demographic structures and climatological needs.

In reference to the Model House at El-Dareeya, Riyadh, and the New Bariz Village in Egypt, this dissertation aims at tackling how the guidelines adopted initially by Fathy in sustainable design advance to construct new codes and combinations. In 'An Architecture for People' by James Steele, Fathy demonstrates how natural cooling strategies and the use of local materials can not only abide by the pre-established standards of climate and structure, but also progress to challenge a change in the art community.

II. FATHY: THE REINTERPRETATION OF THE TRADITIONAL

A. Model House at El – Dareeya, Riyadh

In 1963, Hassan Fathy, delegated as the advisor to the Government of the Kingdom of Saudi Arabia, constructs a 'valid prototype', a Model House with a formal configuration of its architectural elements reminiscent of a typical existing house of El-Dareeya, a village near Riyadh, yet with a symbolic significance to the ruling family [2].

From the lightweight economical 'barasti' roof truss that endures the hot, humid climate of Saudi Arabia, to the upgraded baffled 'malkaf' that boosts the convective action of the Venturi effect, Fathy demonstrates how the structures of basic architectural materials and compositions adjust to different settings. Quoting his words in *'Architecture for the poor'*, Fathy describes the development of the 'barasti' as follows:

"When I was working in Greece I began to think about what other materials might be available in hot-arid and hot-humid climates where they don't have mud-bricks like we do. It was interesting because I like the challenge of change and don't like the repetition of anything in art. They have reeds and other plants, so I thought of using reeds for roofing, giving the resistance of geometric form to the reeds [1]."
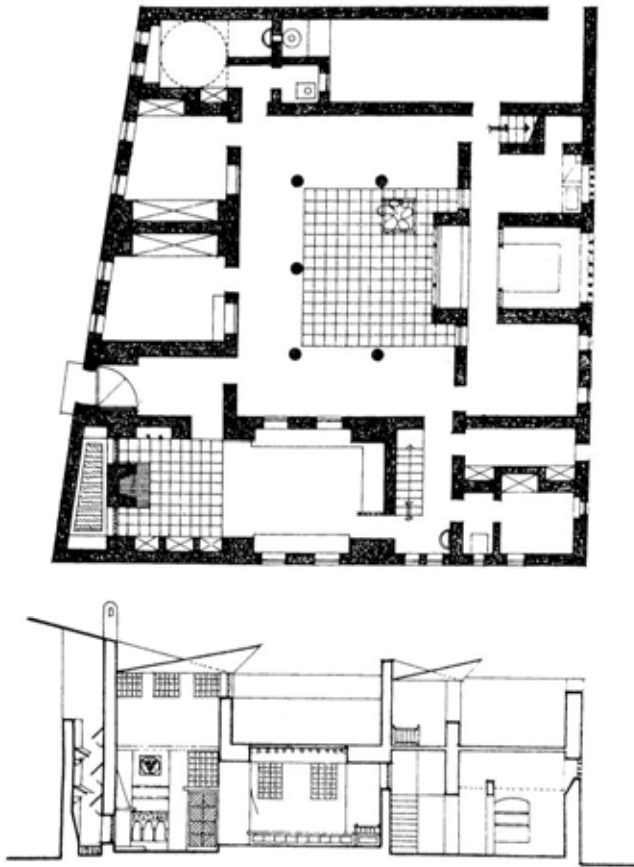
– Hassan Fathy

Fig. 1. Model House at El-Dareeya, 1975, Plan and Section [2]

Fathy confirms that the folded slab, with a ratio of 1:2000 between its thickness and its span, compiles in a thickness of 4mm only to roof a room with a 3-metre span. With a loading per square metre of one ton, the 'barasti' allows for cheap pre-fabrication, quick construction, and high performance with natural ventilation [1]. Along Fathy's terms as well, discrepancies in high and low roof levels highlighted more than a random notion and a formal expression; Fathy carefully mitigated the solar gain absorbed by a roof in mid-mornings by having high roofs cast a shadow on low ranges. A signature of the traditional architecture of this region, high and decorative parapets that extended from the plan and flat roof were also made possible with the use of wood, local stone and mud-bricks, known as the available local materials of the region.

With his passion for soft technology, that aims to adapt while it preserves fragile heritage, Hassan Fathy's proposals never escaped resistance from the people of Saudi Arabia, who despite their extensive contact with their natural surroundings, "did not want to hear about houses without air-conditioning [1]".

B. New Bariz Village, Kharga Oasis

Similar to the experience in New Gourna, yet twenty years later, the planning of New Bariz in 1963 served a challenge to Hassan Fathy; a testimony that could drive out the stereotyped image of Fathy - the anachronistic, the romantic, and the nostalgic. Sixty kilometres south of the Kharga Oasis, Bariz, a regional centre meant to promote extensive agricultural development, was to accommodate two hundred and fifty families, of which one hundred and fifty were farmers and the rest were service personnel [1]. Unlike New Gourna, where he had a chance to interview the 'clients', realize their preferences, and study their existing houses, Fathy here was allotted a limited given and had to focus on the climatological aspects of the area, along with the indigenous architecture of the surroundings, which included the ruins of Bagawat, a cemetery built by an ancient Christian community who has fled Roman persecution in the fourth century A.D. [3].

"Bariz was an interesting problem in which I was to create all the parts of a community, to bring together in the best manner possible people which I did not know. All that I had at my disposal were demographic, geographic and climatic surveys. I had to provide the aesthetics, the sense of man in a space constructed by man [1]."

Hassan Fathy

Fig. 2. New Bariz, Egypt, 1967, Plan of Village Centre [2]

From the introduction of the 'Moorish Café', an institution relevant to the Bariz population, to the discovery of the sand-brick, a means of resistance after the construction of the Aswan Dam, Fathy challenges existing circumstances to navigate unexplored cultural idioms and opportunities. With his belief in the physical material at hand, Fathy succeeds at defying the edict issued by the Egyptian Government that prohibited the stripping of the top soil and restores the deep-rooted practice of brick-making, with sand standing as an alternative to mud [1]. With this material on one hand, and the harsh environmental conditions on another, interior courtyards within each building replaced the vast open squares that baked in the sun (fig. 2). Narrow winding streets that ran in the north-south direction between the buildings allowed for shading throughout most of the day.
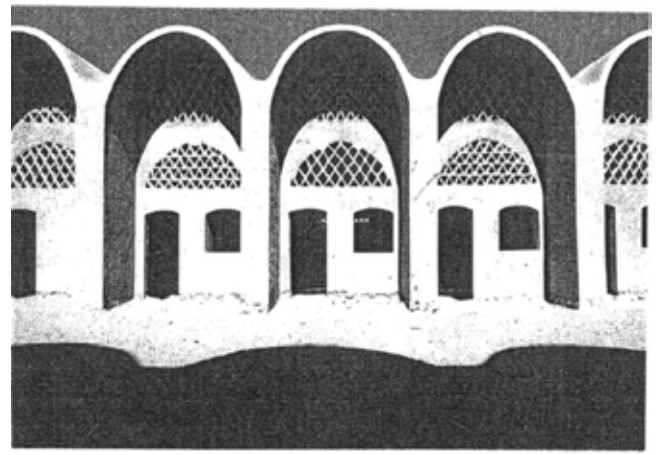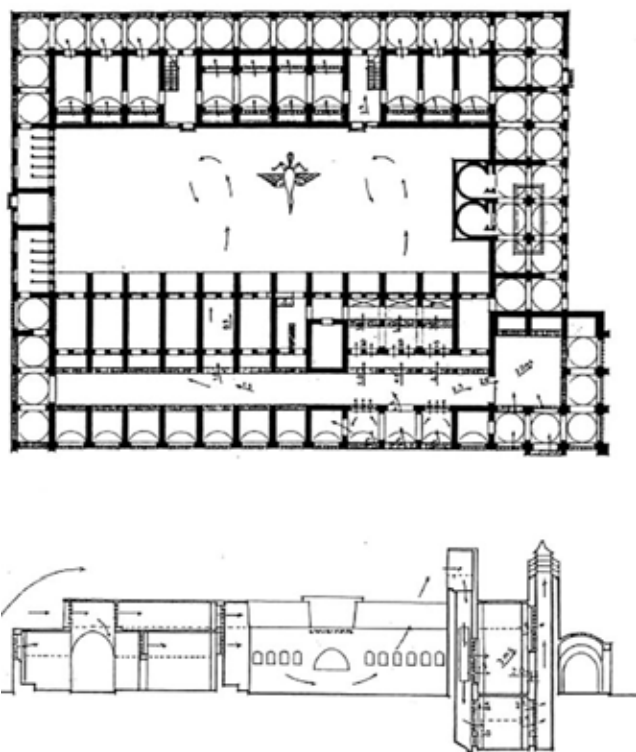

Fig. 3. New Bariz Village, Market Plan and Cross-Section [2]

The new village of Bariz consisted of a mosque, a souk, a hospital, administration offices, and the recently-introduced 'Moorish Café'. In order to sustain the perishable grains, fruits, and vegetables in this agricultural enterprise, Hassan Fathy managed to construct a cool storage area on the lower level of the souk or market-place, that beat the heat by an upgraded 'medieval-Cairo'-inspired malkaf or windcatch (fig. 3). Fathy arranged for a temperature reduction difference of 15°C by mounting a second chimney capped with inclined metal louvres to control and exaggerate the prevailing breeze [4].

Fig. 4. New Bariz Village, Market Courtyard [4]

Through solving the psychological and physiological needs of the people, Hassan Fathy becomes a living proof that contemporaneity lies in the understanding and application of definite physical principles and natural methods, that can be refined to potentially sensible forms and functions. Through the analytic approach of the village of New Bariz and the Model House at El-Dareeya, "Gourna-accused" Hassan Fathy represents the responsibility of today's architect to detect, test and experiment with what is abundant at hand, and to develop universal structural applications that could grow into resistance methods against today's proliferating architectural colonialism.

With the mentor's skills nearing the verge of expiration, a novel connection had to be made between the existing mud-brick construction and the demand for Egypt's poor to be given the adequate education for building a shelter for themselves again. Hassan Fathy, the winner of the first Chairman's Award in 1980, demonstrates how several of his architectural concepts, intended to "train local inhabitants to make their own materials and build their own buildings [2]", have survived throughout time to incarnate in today's modern designs and tomorrow's prospective architects.

III.   LARI: THE REJUVENATION OF THE TRADITIONAL

Thirty years after Hassan Fathy's recognition, Yasmeen Lari, denoted as the first woman architect of Pakistan, develops the objective of using "traditional" architectural methods to teach people not only to build their houses, but to rebuild their homes in the aftermath of man-made and natural disasters [5]. Acknowledged by London's Royal Institute of British Architects (RIBA), Yasmeen Lari's Green Women Centre in Darya Khan initiates a new dimension to the rehabilitation of reimagined cities and communities; one that uses 'Creation from Catastrophe' as a replicable source, in a 'traditional future'[6].

"You cannot come out of poverty if you keep re-building every year. The only way forward is to build the technical capability of communities and improve their earning capacity so that they can fend for themselves when disaster strikes [6]."

- Yasmeen Lari, 2014

A. The Green KaravanGhar: From Swat to Sindh

After the devastating flood of August 2010 submerged one-fifth of Pakistan's total area, Lari, director of the 'Heritage Foundation', along with her team of architects and engineers, translated their focus on cultural conservation to implicate post-epidemic relief and reconstruction, and thus initiated the Green KaravanGhar (GKG) or the 'green' module, in Swat, to construct 270 transient shelters, known as nucleus housing units, that comprised of a room, a verandah, a kitchenette, a W.C. and bath (fig. 5), all made with local materials such as mud, lime, bamboo, and stone [5]. With the 'kaccha' regions still prone to flooding, the team was assigned to design the Green KaravanGhar of Darya Khan Sheikh, a village along the banks of the Indus River, Khairpur, in the province of Sindh.

Fig. 6. Split bamboo planking used for roofing [7]

Stressing on the emerging threat of climatic change on Pakistan, along with the estimated rise in the frequency of floods on Darya Khan, Lari and her team's experiments with new materials suitable for disaster relief introduced a new costless design; a 'clay fortress' involving a bamboo cross-braced structure (known as *dhijji*) supporting stone-infill walls (fig. 6), completed with insulating mud-lime plaster, with split bamboo planking of ring beams and joists for the roof [7]. With almost the same materials used in cold and mountainous Swat after the devastating earthquake of 2005, the newly-introduced sustainable structures of hot and flat Sindh not only served as sturdy, strong, and relieving shelters, but as training workshops for prospective disaster risk mitigation management that aimed at community mobilization and empowerment [5].

"We were pleasantly surprised by the insulation provided by mud and lime," says Lari. "By using indigenous material "the locals develop an instant comfort level as they know how to use it even after you leave [6]".

- Yasmeen Lari

Fig. 5. Green KaravanGhar Module, Section and Plan [5]

With the abandonment of galvanized iron sheets and heavy concrete blocks used in former modular structures, the new housing units devised zero-carbon footprint, and utilized locally sustainable materials that were not only abundant, fast-growing and strong, as in the case of bamboo, but were also easily adjusted into climate-responsive windcatchers that could be easily prepared and repaired by the local household. Stones that could be broken by the local user replaced the use of high-cost and energy-consuming brick, with foundations made of compacted pure sand instead of soil, and thus decreasing wall masonry depth from 5-6 feet to around 2-3 feet [7].

As for the roofs and walls, standard split bamboo planking was replaced by strengthened interwoven local bamboo matting that could be sold in the form of long-lasting pre-fabricated panels in a nearby village, and could be easily transported and quickly installed by the people of Darya Khan in the form of doors and windows. These panels, in addition to being an adequate source of income for the nearby village, were exchanged by the people of Darya Khan for colorful date palm screens that operated on bathroom walls [7].

To keep in line with the local practices of Pakistan, the thick mud-lime layers within the walls were developed to sandwich a waterproof polythene sheet layer edged in grass that allowed for a temperature difference of 10 degrees between inside and outside [6]. With a low cost of $5oo and a completion period of three days only, the advanced GKG module proved to be a 'versatile emergency shelter', capable of customizing its capacity based on the diverse temperatures and climatic regions of Pakistan.

For a village whose people relied on animal husbandry, families lacking space inside the village were introduced with a small farm holding where the livestock was corralled and the fodder was grown, thus avoiding the dung to accumulate near their homes. The Swat Artisan Team, a 12-member group of workers that included carpenters and masons trained by the HF, who traveled from Swat to teach the people of Darya Khan new and sustainable techniques of construction, carried out training workshops in bamboo cutting, bolting, and anchoring, where student volunteers participated from various parts of the country to target education and hygiene in the community [5].

B. The Green Women's Centre, Darya Khan, Sindh

"The idea is that the women can meet there, they have their own social space, there is space for children to play. There is also, in the event of a flood, a first floor where waters won't reach [5]."

- Yasmine Lari

Concerned with the needs of the local village women of Darya Khan, Lari expanded the range of the green low-cost modular unit to design an elevated circular structure (fig. 7), that responded to the interests of the overlooked 'providers of food and carers of children'[6]. The resulting Green Women's Centre on stilts, marked a social meeting space for the women on the first floor, who could either cook on customized fuel-efficient twin burner stoves or sew handcrafted *'ralli'* and *'charpoy'* and sell them in their spare time [7]. The covered ground floor, on the other hand, being the coolest place in the village, served as a school for the children (fig. 8).

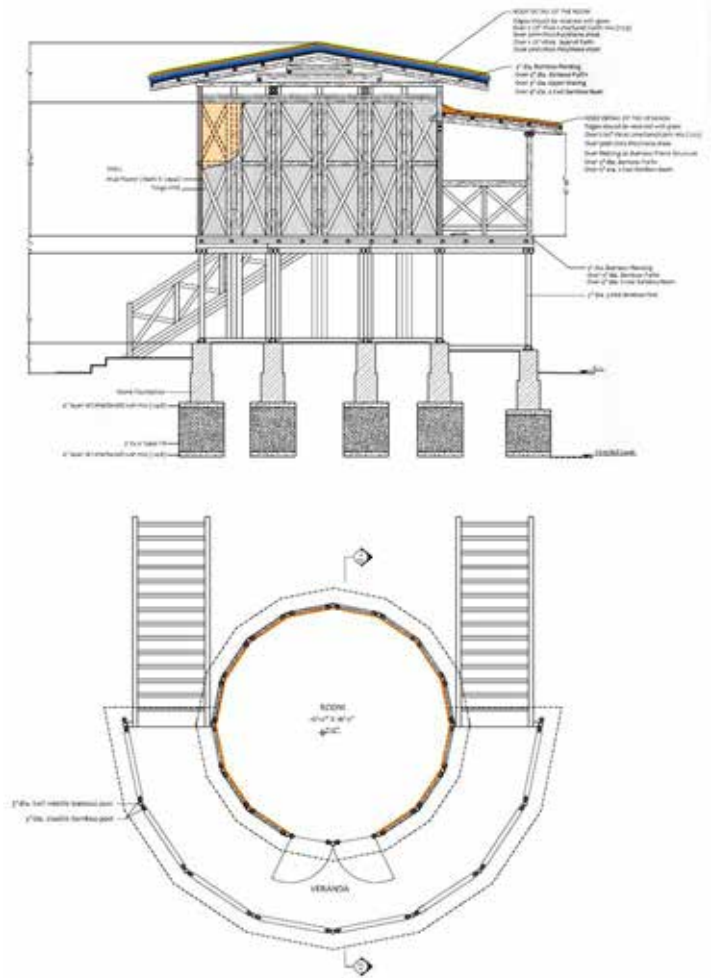Fig. 7. Green Women's Centre, Darya Khan Shaikh, Section and Plan [5]

With the designed unit advancing the technology of mud-plaster construction into a traditionally-based technique resistant to any prospective floods and earthquakes, the skilled women of Darya Khan found a winning opportunity in laying down the mud-plaster, thereby adopting enhanced skills and making money in the process [8].

Fig. 8. Green Women's Centre on stilts, Darya Khan Shaikh [7]

IV. CONCLUSION

With the culmination of significant moments from the 'architectural' practice of Hassan Fathy, man is demonstrated as having the exceptional impact of "the humble, the modest, and even the discreet". With a resolution towards the cultural, political, and socio-economical, contemporary architecture proves to be that in harmony with the traditional, critical of the 'international'. Failure, whether man-made or natural, no longer stands as a question waiting for an answer; but as an answer to the question. With the understanding of the natural, true innovation founds in believing, detecting, and testing with what is at hand, compiling it into the substantial, that could be taught to the people of the present, and moved on to the people of the future. It is only when the local meets the local, where the 'architect' vanishes into thin air.

REFERENCES

[1] James Steele, An Architecture for People: *The Complete Works of Hassan Fathy*. New York: Whitney Library of Design. 1997. pp. 87-95.
[2] Hassan Fathy, *Chorale: Man, Society and Technology*, Architecture for the Poor: An Experiment in Rural Egypt. Chicago: University of Chicago Press. 1973. pp. 41-127.
[3] James Steele, *Return to Egypt 1962-1978*, The Hassan Fathy Collection: A Catalogue of Visual Documents at the Aga Khan Award for Architecture, pp. 28-29.
[4] Abdel-Moniem M. El-Shorbagy, "The Architecture of Hassan Fathy: Between Western and Eastern Perspectives". Vol. 1. University of Canterbury. 2001. pp. 46
[5] Heritage Foundation Pakistan, *Green KaravanGhar and Beyond for Disaster Risk Reduction*. Karachi. September 2011. pp. 2-20.
[6] Zofeen T. Ebrahim, "Pakistani architect pioneers fresh approach to disaster relief". August 2014.
[7] Adele Cong, ed., *In the Face of Disaster*, The Heritage Foundation Pakistan: Green KaravanGhar Initiative. Architecture Asia. Issue 1. January-March 2011. pp. 21-23
[8] Yasmine Lari, "Threats to heritage sites in low-income countries: The application of low-cost emergency preventive interventions", Reflections *on ten years of cultural emergency response*. Heritage Foundation of Pakistan. Amsterdam, Netherlands: Pakhuis de Zwijger. November 2013. pp. 19-22.

Failure, A Part of Success:
Defining Failure in the Work of Hassan Fathy

Samih Zawaydeh
Department of Architecture and Design
AUB
Bliss Street, Beirut, Lebanon
saz09@mail.aub.edu

Abstract- **Hassn Fathy always had a deep connection to the countryside since he was a little kid. Such passion made him believe that there is something that could be done to restore the Egyptian countryside and improve the lives of the fellaheen. Therefore, he took it upon himself to find a solution, for he was aware that such a transformation should have been a result of firmly believing in the cause itself.**

When planning Gourna, Fathy understood that modernism ideologies would not be applicable in this case. He grasped the setbacks found in such strict beliefs and morphed them into ones that aided his design. A design that consults tradition, social values and needs, while still springing from the roots of the Goruin culture.

Gourna was a prototype that needed to break free from sanctified ideologies of modernism, and instead emphasis the ideologies of traditions and social values. However the latter was hard to implement because the Gourins were not willing to cooperate since they were being moved against their will, thus they rejected the project as a solution to their problem.

Fathy's work in Gourna was highly criticized to the point where it was labeled as a failure of a project. Therefore, in this paper I am aiming towards developing a hypothetical concept of failure in the architecture of Hassan Fathy.

However, framing the overall project under the title of failure would be obsolete, as the project needs to be dismantled into different chapters in the efforts of trying to pinpoint what actually went wrong. The latter would be plausible by tracing Goruna as a process and not an outcome.

I. INTRODUCTION

"I always say that we benefit more form failures than from successes. Failures makes us consider every aspect of a situation, but success has certain hidden dangers which when repeated on a large scale become a disaster. Only by knowing the mistakes in an idea can you find a solution to it and save it from the beginning. **Hassan Fathy**

Hassan Fathy has always had a deep connection to the country and the idea of it. Such association was supplemented by two contrasting images, one from his mother who spent most of her childhood in the country and the other from his

father who avoided the country altogether. From the two contrasting images, Fathy was able to develop in his imagination an amalgamation of both, which in its turn formulated a vision of the country as a paradise that is darkened from above by clouds of infested flies [2].

Such imagery haunted Fahty from a young age. However, he continually believed that there is something that could be done to restore the Egyptian countryside and improve the lives of the fellaheen. Moreover, his numerous visits to places like Talkha disturbed him, to a point where he could think of nothing but the hopeless resignation of the peasants to their cramped and undeveloped view of life.

Due to the above, Fathy believed that the solution was in his own hands, mostly because he knew that such transformation should have been the result of love and not by the dictate issued from office desks in Cairo [2]. Accordingly, one can dedicate a great deal of emotional attachment to the issue, since one need to love the "fellah" enough to build them homes in the country and devote their lives to work towards improving the rural life. The latter is highlighted by the fact that Fathy himself knew that the peasants could not afford the decent houses they need, especially because they were too sunk in their misery to initiate a change, thus accepting the horrible situation in which they were forced to put up with [4].

Consequently, all his efforts in trying to restore the Egyptian countryside and provide a decent home for the poor made Fathy a highly debated architect. His work in Gourna was weighed against the ideologies of modern architecture and was vastly criticized by many, to the point of labeling the project as a complete failure. However, for Fathy himself, Gourna was an innovative prototype that required breaking such ideologies to reach its full development. Henceforward, it was not the outcome that had the greater significance to him, but the great learning process it would provide.

II. THE EVOLUTION OF THE DESIGN PROCESS

A. Breaking Away From Modernism Ideologies

While tackling a project like Gourna, I personally believe that Hassan Fathy was aware that the rules taught in schools of architecture would have been irrelevant in this case, if not even destruction when put into practice. Being a social experiment of rehabilitating the Egyptian countryside, Gourna was a prototype that needed to break free from sanctified ideologies

of modernism, and instead emphasis the ideologies of traditions and social values [1].

Trusting that he could learn from what modernism lacked in terms of architectural growth through social interactions and traditions. Fathy focused on rethinking some of the main concepts of modernism, and morph them in a way suitable to that of Gourna.

One of the main letdowns in modern architecture in my opinion is the ideology that architecture has to deal with how people should ideally live rather than how they actually do live. Such notion was evident in the earlier phases of the project when Gourna was picked as the site of intervention. In the beginning, Fathy found the project to be rather daunting, since he was given fifty acres of virgin land to create a new life for the seven-thousand Gournis leaving their old town behind [2]. Here Fathy rejected the ideology of modern architecture that tends to redefine social and aesthetical values into rigid set of rules. Instead, he understood that Gournis have a complex web of ties made up of there own customs and habits that should be carefully dismantled and put together again in Gourna [1].

Evidence confirms that the ideologies adopted by modern architecture fail whenever the architect neglects the social and aesthetic values of the user [1]. Accordingly, Hassan Fathy wanted to fulfill his social requirement to the Gournis by allowing the design to spring from the roots of their own culture. He understood that he should come up with a conscious decision that is reached by consulting tradition rather than fixed modernism ideologies [2]. Therefore, Fathy was able to successfully challenge modern architecture principles by embracing cultural differences and taking into consideration his user-client needs. What made him successful in this sense is the fact that he did not base his choices on his own familiarities, and did not assume that his user-client would get habituated to living the way he expects them to live [1].

B. Establishing His Trinity in Design

Being granted the access to design whatever he wants in Gourna, Hassan Fathy faced the dilemma of how he could possibly grasp all the necessary details of the domestic life of a Gourni. He was aware that such information is vital to his design since it allows him to understand precisely what the Gournis want in their houses.

Fathy did not want to impose his values on a public that might not necessary share the same beliefs as him. He was concerned about maintaining the individuality of the village so that the people would feel at ease once they move in [4]. Consequently, Fathy did not neglect the social and ethical values of the Gournis by creating a project that ruptures the traditions they have, since he was aware of the negative public reaction towards modern buildings that tries to be "different" rather than fit in [1].

For Fathy the best way to plan Gourna did not revolve around making an odd or original design, as the public does not always accept such change since it disfigures the aesthetics they are used to embrace. Fathy believed that God does not have to alter the design's concept in order to fabricate individuality in humans, but can span the entire scale of beauty, by simply adjusting the position or the size of the elements in the face [2].

Likewise, Fathy also believed that like in nature were no two men are alike, the architecture of the houses must follow such concept and emerge from the needs of the inhabitants, hence making each house unique and fit to the family it hosts. Yet, he argued that when an architect is faced with a job that requires him designing thousand houses at once, rather than dream for the thousand individuals he design one typical house and places three zeros to its right [2].

The later notion discussed is directly related in my opinion to one of the modernism major flaws, that is standardization. Through the adaptation of standardization and mass-production techniques an impersonalized monotonous design is usually introduced, one that lacks diversity and actively discourage the reflection of human cultural differences and individual needs through design [1].

Fathy argues that such approach denies creativity to the architect and humanity to man. Fittingly, he made sure to integrate the people's needs and made them help in the design process, especially that he was aware that one architect could not hold thousands in his head [4]. Accordingly, Fathy was able to reestablish the trinity between the owner, architect, and craftsman as one unit that works towards one common goal. At Gourna the people were their own designers, contractors, supervisors making them as conversant with the process of construction as the architect himself. Such notion helped in maintaining the individuality of the village, which was one of Fathy's main concerns when he was given project [1].

III. Mud-brick as the Sole Hope for Rural Reconstruction

Rejecting traditional architectural forms and techniques was implicit by modernist architectures [1]. However, Fathy described tradition as not necessarily being old-fashioned and is not synonymous with stagnation. He believed that it was his duty as an architect to keep the traditions of the Egyptian countryside going, yet he should add his own invention and insight to give them that additional momentum they needed to transform themselves, thus saving them from coming to a standstill [3].

When it came to Gounra, Fathy's interest in the traditional techniques of building used by the peasant aided in the development of his project as a whole. In this case, Fathy avoided the attitude that is often adopted by architects when confronted with a peasant-community, the attitude that such community has nothing worthy to offer [2]. On the contrary, Hassan Fathy was aware such traditions go back to the beginning of human society, and yet they are still alive and will exist perhaps as long as human society does [4]. Therefore, the solitary solution for rural reconstruction would have been mud-brick, a material that is directly linked to the traditional techniques of building in the Egyptian countryside.

Fathy was aware that when working with a peasant-community, money is the main constrain. Therefore mud-brick was a heavenly sent material that should be used in the construction process of Gourna. It was economically sufficient

to be offered to the peasants by the government for free. Also, cheap enough for the peasants to obtain it themselves without the help of the government incase larger amounts were needed in later phases [2].

Furthermore, Mud-brick is a common local material that does not need skilled labor beyond what the peasants themselves can engage in, therefore most of the work can be done by unsupervised labor. The later aided Fathy in implementing the notion of aided self-help when it came to the construction of Gourna, especially that the material of choice was common to the peasant's found in the Egyptian countryside [2].

IV. THE GOURNIS BEING THE VITAL SETBACK

It is known that no peasant can ever dream of employing an architect, and no architect ever dreams of working with the miserable resources of the peasants. Knowing that the architect is an expensive luxury and is found where money is, the relationship between Fathy and the fellaheen was crucial in this experimental project. Especially that Fathy was emotionally attached to the project and wanted to build something that is rightful to the fellaheen as patrons of Gourna.

Fathy was aware that he was building a new site for a population that was being moved against its will from its accustomed home, therefore he was aiming towards creating houses that reflect the fellaheen's social values and needs so that they would be able easily adjust to such move. However, in order to successfully grasp their collective lifestyle and what they exactly want from the project, he was required to directly consult the Gournis and understand from them what they are hoping to achieve with such rehabilitation project [2].

However, the fellaheen were quite incapable of expressing themselves and talking about the style or beauty of the houses they aspire to have, or what they really wanted from the project. Moreover, when Fathy asked to consult the women, for the house is considered to be the province of the woman, his query was rejected because the women were kept jealously out of the way [2].

Another drawback that Fathy encountered was the difficulty Gourins faced when reading the plans. Even though he had built some twenty houses to show the fellaheen the kind of architecture he was offering, they were still unable to grasp neither the plans nor the notion of the project proposed [2].

From the mentioned above it is notable that the connection between Fathy and the fellaheen was dimensioning as the project proceeded, which in its turn resulted in a problem of communication that hindered the relationship between Fathy and the fellaheen to a point where they scarcely discussed the buildings with him [2].

However, despite all of that the major problem Fathy faced was the Gourins being against the project as a whole, due to the fact that they were being moved against their will. Such oppression caused the Gournis to despise the project and reject it as the solution to their problem [2].

In the perfect scenario, Gourna was an example of adopting the approach of the village's traditional cooperative building customs and metamorphosing them into large-scale project that has to deal with building a complete village [2]. Yet, due to the rejection the project has faced from the Gournis, the notions of aided self-help and cooperative building customs had to be dropped and replaced by hired laborers that are to be paid, which in its turn stopped the project from being really cheap.

V. DEFINING FAILURE IN GOURNA

The key question that remains unsolved here is one related to finding an accurate definition to the terms of failure found in a project like Gourna. According to Fathy himself, the notion of failure is nothing but an educational process that one learns and benefits from. He personally believed that failure makes one consider different aspects of the situation, that otherwise could be gone unseen by success. Therefore, through such a process one can learn the setbacks of a certain idea and work towards improving and saving it from the beginning.

What Fathy did in Gourna was an experiment that aimed towards improving the Egyptian countryside. Although the project did not serve its initial purpose of rehousing the fellaheen, yet it is still considered an innovative prototype that bridged the gulf that separates folk architecture from architect's architecture, which is a notion that Fathy strongly believed in.

Moreover, the work of Fathy has opened the way to many interpretations on how one view the life in the Egyptian countryside as a whole. The latter is seen through the alternative model Fathy has created in Goruna, where he adopted a sustainability-oriented architecture that responds to the social issues and needs of the community by using vernacular styles, materials and techniques.

Yet, the fact that the Gournis were never interested in the project as a whole made it impossible for Fathy to be able to succeed. Consequently, the notion of failure here could be directly linked to Fathy's lonely struggle to persuade his noble ideas and beliefs, which were against the prevailing culture during his lifetime. Hence, I believe that the rejection Fathy faced and was unable to counteract is the main reason of the so-called failure of Gourna.

VI. CONCLUSION

For Fathy himself Gounra was an experiment and example on how he would like to rebuild the whole Egyptian countryside. An aspiration that I believe is linked back to his dream as a youngster, where he wanted to create a village that is made from the earth and is tied up through the bonds of its inhabitants. Therefore, Gourna was not an end in itself but an initial tentative step on route of regenerating the Egyptian countryside. A social experiment that does not only aids Fathy in examining all his ideas, but a prototype for other larger villages to come.

Here, it is noteworthy to mention that modern architecture intentionally defies its older neighbors rather than standing beside them in peace. Its ideologies are systemic and rigid thus do not allow space for the architect to grow [1]. However, Fathy was able to avoid such notion when planning Gourna by respecting the traditional social and aesthetical values. Such sensitivity led to an outcome that is the result of user-

participation and flexibility, one that was not shaped by specialists, but by the spontaneous action of the people that share a common heritage and working towards creating a community of their own.

The course of building such a project was an educational process for Fathy himself, as he was learning, modifying, and improving the design as he goes to fabricate one final product that fits perfectly the needs of the fellaheen. His aim was to generate a land that grows according to the fellaheen needs and desire, a land that the fellaheen would call their own.

Accordingly and despite all the criticize it received, I do not personally believe that Goruna was a stillborn experience, for in Gounra a new concept of rural-housing was tried-out and proved feasible. Hence, it has paved the way to a plausible national rehousing policy that would have provided numerous housing opportunities to millions of Egyptians at a price they could afford.

In a nutshell, I do not believe that one can describe a project like Gourna as a failure. However, I can denote that what really led to the failure of the initial concept was the lack of contribution the user-client made to the design. Such a suggestion is highlighted by Fathy when mentioning that no matter how ignorant or even suspicious the client contribution is, it is something that the architect cannot do without [2].

From that I can safely say that Fathy took what occurred in Gourna as a learning experience that not only benefited him as an individual but the whole architectural society. Gourna was and will always be admired by numerous influential individuals who are perceptive enough to see the possibilities it offers. An example of the latter would be the report generated by UNESCO, in cooperation with the Egyptian Ministry of Culture and the governorate of Luxor, listing Gourna for conservation and safeguarding.

In conclusion, the definition of failure in such project does not hold. The latter is backed up by the fact that despite all the troubles Fathy faced he was determined to build a town specialized to its inhabitants. He understood his duty as an architect towards the fellaheen and was very passionate about the project and the enhancement in lifestyle it had to offer the Gournis. Also, one can notice that Gourna in itself presented many architectural possibilities and aided in emerging new principles of town building. Therefore, one cannot consider it as a failure but a testing ground and a prime example of all the innovative ideas Fathy had.

REFERENCES

[1] Brolin, B. C., "*The Failure of Modern Architecture*". Van Nostrand Reinhold. New York, 1976.
[2] Fathy , H., "*Architecture for the Poor: An Experiment in Rural Egypt*". The University of Chicago Press. Chicago, 1973. Pages 01-148
[3]Shearer, W., & Sultan, A.-e.-r. A., "*Natural Energy and Vernacular Architecture*". The University of Chicago. Chicago, 1986. Pages 03-28
[4]Steele, J., "*Hassan Fathy*". Academy Editions. London, 1988. Pages 07-85
[5]Steele, J. "*The Hassan Fathy Collection*". The Aga Khan Trust for Culture.

"From earth to earth… and beyond!"

Christine Ghossoub
Department of Architecture and Design
American University of Beirut
Riad El Solh, Beirut 1107 2020/ PO Box: 11-0236
cgg06@mail.edu.lb

Abstract- Solid waste management in Lebanon is a rising issue that remains unsolved until this day. Waste is rapidly accumulating in the streets endangering the environment and public health. I chose to conduct a research about trash management in Lebanon and abroad for my thesis and I am currently designing a composting facility and buildings with related functions in the Nahr Al Mott quarry site as my final year architecture project.

The question that my thesis answers is:

"How can trash be reused and transformed into building materials and landscaping elements to revive and reshape the wounded mountain site in Nahr Al-Mott?"

The project and paper focus on the benefits of trash in the fields of architecture, landscaping, building in nature, and design, and expand on how waste can be an opportunity to transform a 'dead' place into a hub of visitors and public activities.

Figure 1- Municipal Solid Waste Composition in Lebanon

Figure 2- Municipal Solid Waste Treatment in Lebanon

I. INTRODUCTION

Lebanon produces around two million tons of Municipal Solid Waste (MSW) per year. Figure 1 shows that more than half of the trash is composed of organic waste. Following are mostly recyclable materials: paper/cardboard, plastics, metal, and glass. These materials are generated by several light manufacturing industries: food and beverage, fabricated metal, non-metallic mineral, furniture, clothes and dyeing fur, wood products, leather products, and textiles. Figure 2 shoes that Municipal Solid Waste is treated in different ways. It is either landfilled (48%), dumped (29%), composted (15%), or recycled (8%) [1]. Thus, most of the trash is being landfilled or dumped (in sanitary landfills in the best cases) rather than composted and recycled. Moreover, municipal solid waste in Lebanon can be divided into two categories: hazardous (3000 tons/year) and non-hazardous (185000 tons/year). Both categories are currently being sent to landfills or dumped [2]. Figure 3 shows the distribution of sorting plants, landfills, composting plants, and anaerobic digesters in Lebanon, as well as the population per district. The city of Beirut is the most condensed relatively to its size and produces more than 600 tons of MSW per day.
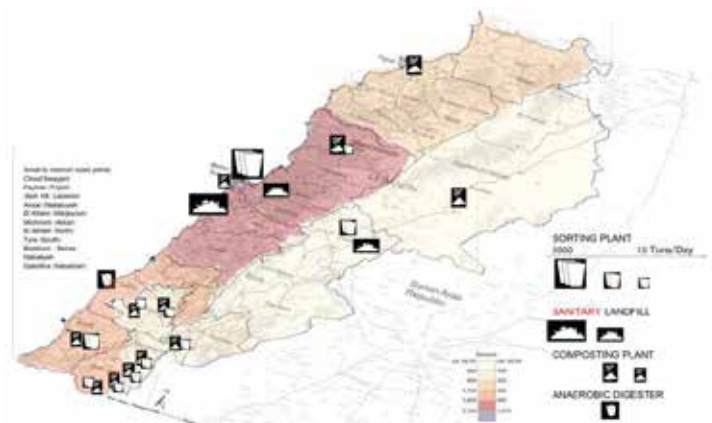

Figure 3- Plant distribution in Lebanon and density of population

Composting and recycling trash, after the sorting process, have far more advantages than landfilling, dumping, burning, or shipping it. Therefore, the general design plan of intervention includes recycling facilities for every type of waste (a paper recycling facility, a glass recycling facility, a metal recycling facility, and a plastics recycling facility), and a composting facility with all its related functions. Phases of the general recycling processes are shown in Figure 4. The project will focus more on the composting process that will be explained in detail in this paper. The proposal also includes a visitors' center that will receive students from schools, environmental specialists, and any citizen interested to learn about the methods and the benefits of recycling and reusing trash. Moreover, visitors of the site will be able to purchase recycled materials, artistic crafts made of recycled materials, plants, fruits and vegetables, and fertilizers, and other products from public markets built on site. The design proposal is a chance to treat the six hundred tons of waste generated daily in the city of Beirut by transforming the organic waste into fertilizers and recycling the other types of trash to create new materials (**APPENDIX A**), thus introducing new job opportunities (workers in the facilities). Furthermore, treated waste will be used in landscaping new terraces that will reshape a quarry site and as construction materials for facades of the various facilities and buildings on site. The different methods used to achieve this are described in the following paragraphs.

Figure 4- Recycling processes diagram

II. THE GOVERNMENT PLAN

The main reasons behind the trash problem in Lebanon are the increase in urbanization rate, the increased population due to the increased number of refugees in the past few years, the "poor enforcement of the law", the lack of environment policies, the bad habits, the "increase in income per capita", and the "social keenness to use new materials instead of used ones". According to the government and general studies, waste treatment should follow priorities as shown in Figure 4. The plan emphasizes the need for sorting the trash from the source. After that, the trash should be sorted in a Material Recovery Facility (MRF), then recycled or composted, and finally reused. The suggested design proposal will follow the government plan (Figure 5). Municipal solid waste will be collected from the city of Beirut by trucks and transported to the MRF where the sorting process will happen. The trucks are weighed before entering the facility to determine the quantity of trash they are bringing. Then, the sorted trash will be sent to the corresponding recycling or composting facilities depending on its nature (organic or non-organic) for treatment. Resulting elements will be used and seen by the public. The study of waste management in Lebanon was followed by a research about the financing of and cost of waste management. Figure 6 illustrates the average cost in $/ton of every step in the process of waste management leading to a total of 130$/ton. According to the government plan, the main sources for financing the project of waste treatment are the "independent municipal fund", international loans and grants, municipalities, and the government budget [3].
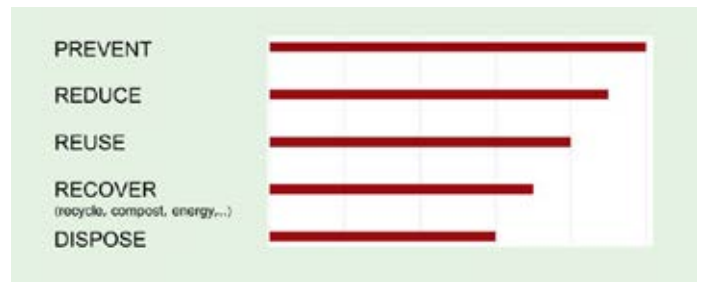

Figure 4- Treatment Priorities

SORTING FROM THE SOURCE - SWEEPING - COLECTING - TRANSPORTING

Figure 5 - Government plan for municipalities
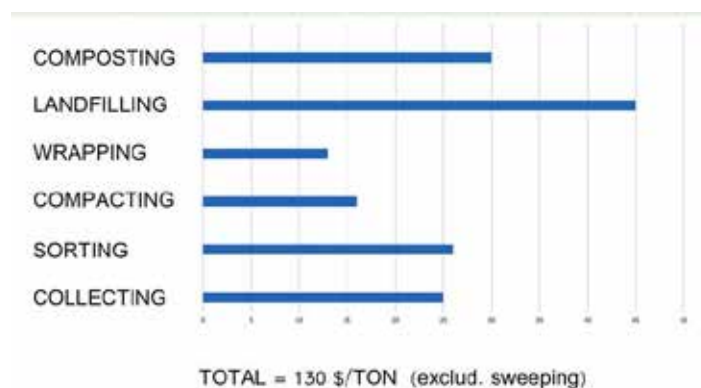

TOTAL = 130 $/TON (exclud. sweeping)

Figure 6- Cost of waste treatment

III.   CASE STUDIES

Lebanon is not the only country that suffers from trash management problems. Large cities all around the world have faced similar issues and some have already built facilities to treat their municipal solid waste and recycle it. One example is the Sunset Park Material Recovery Facility in Brooklyn, NY, by Selldorf Architects. The 13000 sqm. project includes a tipping building where the metal, glass, and plastic waste is dropped off, a processing building where the waste is recycled, a storage building where the compacted recycled materials are kept for later distribution, and an educational facility containing classrooms, exhibitions, and display areas for students and the public. The site of the project has outdoor green spaces for the visitors to enjoy. Additionally, an outdoor bio swale where runoff water is filtered from pollutants and silt is landscaped near the project. The filtered water is used by the facility [4]. General plans, sections, and photos of the project are shown in figure 7.

Figure 7- Aerial view, general project plan, and plan of the educational facility of the Sunset Park MRF

In other examples, architects and landscaping architects have found ways of using recycled trash in construction and building, especially in building façade systems. The Can Cube by Archi Union Architects is an innovative residential and office building of 1000 sqm., built in 2010 in Shanghai, China, where drink cans were used to form a layer of the building façade. The façade system consists of "aluminium carbonated drink cans which are enclosed in an aluminium frame". In order to save energy, the cans are not recycled. They are cleaned and kept in their original form. The aluminum frame keeps the façade light and easy to adjust. Openable sash windows provide the interior space with light and ventilation [5]. The system is shown in more detail in the picture and exploded diagrams of Figure 8.
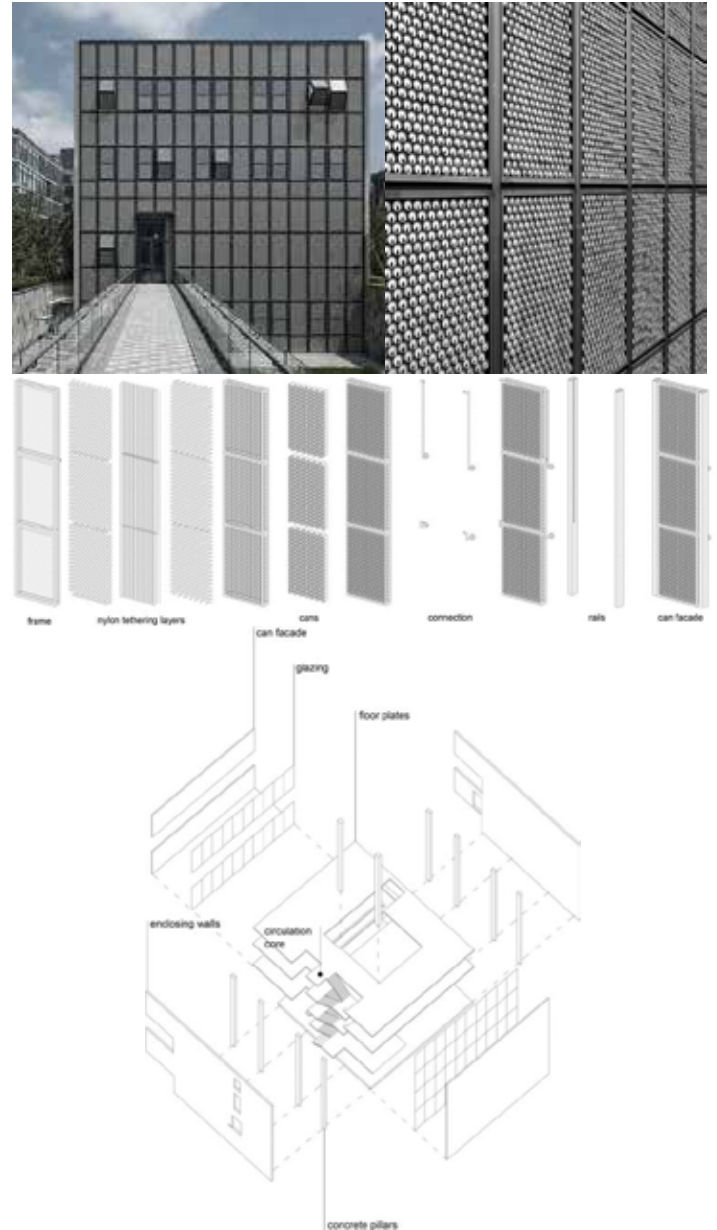

Figure 8- Detailed façade system of the Can Cube

Moreover, in 2012, Dratz&Dratz Architects built a small temporary workspace made of 550 compacted bales of paper in Essen, Germany. In this case, papers were compressed and stacked in the form of cubes to hold the building better. The system was "surprisingly well-insulated and was also quick-drying in spite of incessant rain". The façade has a double function. It reduces on the cost of using new materials for construction by reusing paper and cardboard and it also gives the building an innovative and very colorful creative look [6]. The bales of paper and cardboard, decaying over time, might be replaced by other similar ones, giving the facade an always changing vibrant aspect (Figure 9).
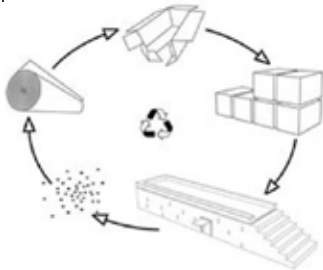
Figure 9- Paper/cardboard as building material

Furthermore, it is crucial to mention that local efforts have been made by several organizations and companies in Lebanon. For instance, CEDAR Environmental led by Ziad Abi Chaker [7] have already started using and implementing recycled materials in designed objects and construction. They were able to build a Material Recovery Facility using the Dynamic Composting technology that treats around 96 tons of waste daily to serve around 200000 people. They also built a prefabricated house made from recycled plastic bags and other furniture elements using the Ecoboard material that they fabricated, as well as artistic objects from recycled glass and others. Their work is shown in the following pictures.

The Material Recovery Facility

The prefabricated house

Furniture and objects made of recycled materials

Figure 10 – CEDAR Environmental projects

IV. THE DESIGN INTERVENTION

a) Site analysis:

Located in Roumieh, the site of intervention is a natural mountainous area that has been excavated forming a quarry. Located at around 5.5 km from the capital Beirut, it falls in an industrial zone, far from residential neighborhoods, and it is directly accessible by the Main Metn Highway. The latter stands above the old river of Nahr Al-Mott which today appears to have dried out. An old water canal, currently dumped, was located between the two mountain cliffs of the site and used to feed the river. One of the site's main advantages is that it is at an altitude of 105 meters, therefore, far enough from the groundwater level. This means that if trash is introduced on this site, it would not harm the underground layers of earth or water (given that insulating membranes and leachate tubes are placed under the trash). The maximal total built area would be 12000 sqm. out of a total site area of around 40300 sqm. The quarry is large enough to fit the program planned to be introduced. On the other side of the Main Metn Highway, a larger quarry site (~71500 sqm.) in Biakout faces the chosen intervention site and currently houses a cement factory. There, some buildings located on the top of the cliff were abandoned due to dangerous conditions and previous accidents. Both excavated sites are strategic and offer various advantages for the design project dealing with trash. They save on the cost of excavating a new land, they are close to Beirut, the major city generating trash, and they are accessible by the main highway by both trucks and regular cars. Figure 11 and **Appendix B** highlight major features of the Roumieh quarry that will be emphasized in the design proposal.

Figure 11- General view of the Roumieh quarry

b) Composting process and related functions

The design proposal focuses more on the composting process that can be explained as following. The organic type of waste consists of biodegradable waste such as plants and debris of fruits, vegetables, animals, and human waste. To begin with, it is transported to a tipping facility where it is mixed with yard waste (1/3 organic to 2/3 yard). Then, the mixture is shredded and grinded in a machine into smaller particles. The final mix is laid outdoors in the form of long 'windrows' (more than 50m long and 4.5m wide). This process consists of three phases where the compost mix is transformed into fertilizers after eight weeks in total: four weeks for phase 1, two weeks for phase 2, and two weeks for phase 3. In phase 1, the mix is covered with a special membrane (i.e. Gore Cover) protecting the compost material from weather (Figure 12). This cover is equipped with two sensors that allow measuring and controlling the temperature and the oxygen amounts in the mix for the best outcome [8]. All leaky substances from the mix are collected in a water trap and a leachate system is installed to prevent any damage of the earth below. The compost material is uncovered then turned to accelerate the reaction before it moves to phase 2 where the mixture is covered again. In phase 3, the compost material is uncovered to dry out. Once done, it goes into a screening machine that picks out any remaining plastic bags or metal pieces from the dry mix. The latter is grinded again into a finer product. The end result is a clean soil-like mixture and manure for agriculture. Before applying it to plants, the finished compost (or fertilizer), is sent to labs for testing then to a packaging facility. Resulting bags of fertilizers are sent to a public market where the visitors can purchase them for their private gardening activities. The project proposal will include all the buildings necessary for every step of the composting process (i.e. tipping facility, outdoor areas for laying the mixture, labs, packaging facility, and public market) along with designed paths connecting all the functions in order. Additionally, unpackaged fertilizers will be used onsite to grow plants, trees, fruits and vegetables, flowers, and others.
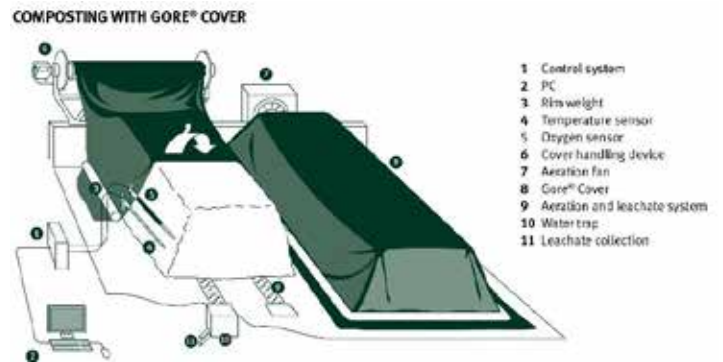
Figure 12- Gore cover for composting

c) Plan for sorting and recycling the trash

Prior to composting the organic type of waste, trash will be sorted and recycled in the Biakout quarry as follows. As mentioned before, trash bags will be collected from the city of Beirut and transported in trucks to a Material Recovery Facility. There, trash will be sorted and thoroughly divided into distinct materials: glass, plastics, paper/cardboard, metal, and organics (Electronic waste is transported and treated abroad). After the sorting process, every type of material is compacted into bales in the form of cubes. Part of these cubes will be used as construction units for the rest of the facilities and buildings on site (Figure 13).

Figure 13- Bales of trash as construction units

Additionally, bales made of recyclable materials will be sent to their corresponding recycling facilities. Bales made from materials that can't be recycled, combined with a steel structure, will be used for landscaping the Roumieh quarry and creating higher terraces for agriculture. The Roumieh quarry will house a large market for selling the materials recycled on the opposing quarry site.

d) Program for visitors

The Roumieh quarry will include a program for the local visitors of all ages, especially for students and academics willing to learn about the process of composting and agriculture in general. One of the buildings will include an amphitheater that will be used for conferences and teaching purposes, and workshops for both artisans and the general public. Artisans will have the chance to assemble new items from recycled materials and sell them in the public market. Moreover, visitors will have the opportunity to learn about the art of fabrication of these objects and try it using their own hands. Not to mention the guided tour passing through all the operational facilities. Outdoor terraces will also serve as display areas for these crafts made from recycled trash. Furthermore, the whole site will be landscaped and filled out in a way that will make it closer to its original mountainous form and condition. Superposed layers of planted terraces will reshape the lost layers of excavated rocks and re-green the site with trees and plants. Several public spaces such as restaurants, exhibition areas, and outdoor "promenades" overlooking certain features of the natural site will form the new "microclimates" created for the people. Not to mention the water which will be collected in a rainwater collection pond that will serve for keeping the composting mixtures moisturized. The water collected and filtered would equally serve for irrigating the onsite planted terraces. Youngsters will be able to enjoy a hiking path and mountain climbing activities wrapping around the water.

e) Some design strategies and personal input:

To begin with, I started by drawing and reconstructing the mountain site as it was before excavation. This helped me get a feel of the initial natural shape of the area and the level of distortion it has faced. Next, I drew and modeled the exiting site plan to finally complete it with my ongoing design intervention (**APPENDIX C**). The new landscape will be close to how it was before excavation. Yet, new elements will be included: buildings, bales of trash, rainwater collection, and agriculture. To create the new levels of the landscape, I have joined, using the plan, the "lines" with the same altitude from the right side of the quarry with those of the left side to create continuous overlapping layers. Those layers, elevated from the ground and following the principle of landfill, would be filled with bales of unrecyclable trash and structural elements to hold the planted roofs. Then, volumes representing the composting facility and other functions would be inserted within these layers and connected to each other by paths. Bales of recyclable materials will be used in the façade systems of certain buildings according to their material qualities and the function of the building. For example, bales of paper or glass would create colorful shades of light in the interior space.

What's more, sustainable systems will be introduced in the project. For instance, methane gas can be extracted from the covered bales of trash to generate electricity that would be used by the onsite facilities. Photovoltaics and wind turbines would generate similar energy.

V. CONCLUSION

In one word, trash can be an opportunity rather than a threat. It can be used creatively and purposely in the fields of design, architecture, and landscaping. In the course of two semesters, I have learned about trash management in Lebanon and referred to case studies from abroad to come up with a "part-solution" for our local serious problem. Instead of randomly dumping solid waste in landfills, burning it, shipping it, and even throwing it into the sea, we could use it for the benefit of construction and design. The Roumieh quarry will not only act as a large "container" for the trash, but it will also be a hub of visitors, a site of attraction, and a place for learning.

VI. ACKNOWLEDGEMENT

VII. REFERENCES

[1] "Country Report on the Solid Waste Management in LEBANON" by SWEEPNET, April 2014
[2] FITCHNER, CDR, Lebanon. "Treatment and disposal of Municipal Solid Waste in Lebanon - Request for Proposals - Part IV - Project Information Memorandum". October 2004
[3] Conference at Issam Fares Institute for Public Policy and International Affairs on Sept. 22, 2015: "Household Solid Waste Treatment Plan, approved by the Lebanese Council of Ministers Resolution No.1 on 09/09/2015."
[4] http://www.selldorf.com/projects/sunset-park-material-recovery-facility
[5] http://www.archdaily.com/85278/can-cube-archi-union-architects-inc
[6] http://inhabitat.com/dratzdratz-unveil-workspace-made-from-recycled-paper-in-essen-germany/
[7] http://www.cedarenv.com/machines.php
[8] Flisram, G. (2010). Favorably disposed. Planning, 76(7), 30-34. Retrieved from http://search.proquest.com/docview/763224695?accountid=8555

Additional references
- Interview with Mr. Paul Abi Rached, head of T.E.R.R.E. Liban and with Mr. Antoine Abou Moussa, environmental consultant at T.E.R.R.E
- Interview with Mr. Ziad Abi Chaker, owner of C.E.D.A.R.
- Interview with Dr. Issam Srour about the reuse of concrete and building materials in new construction.

Appendix A
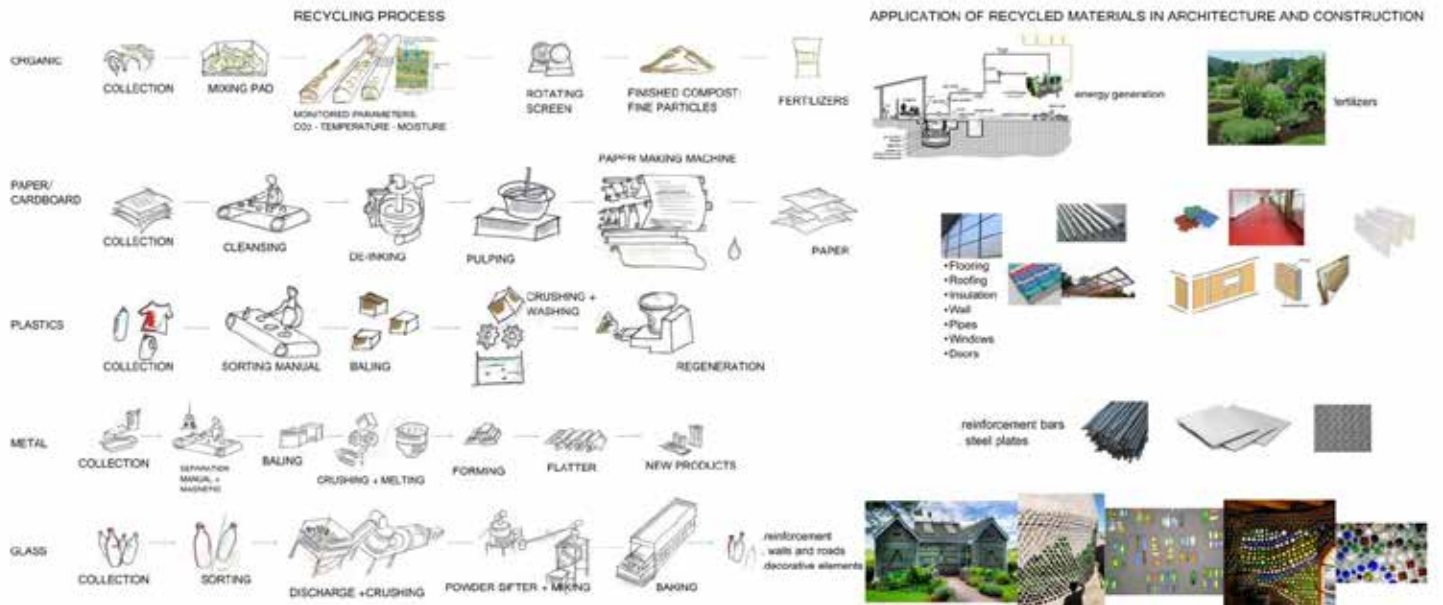
Appendix B
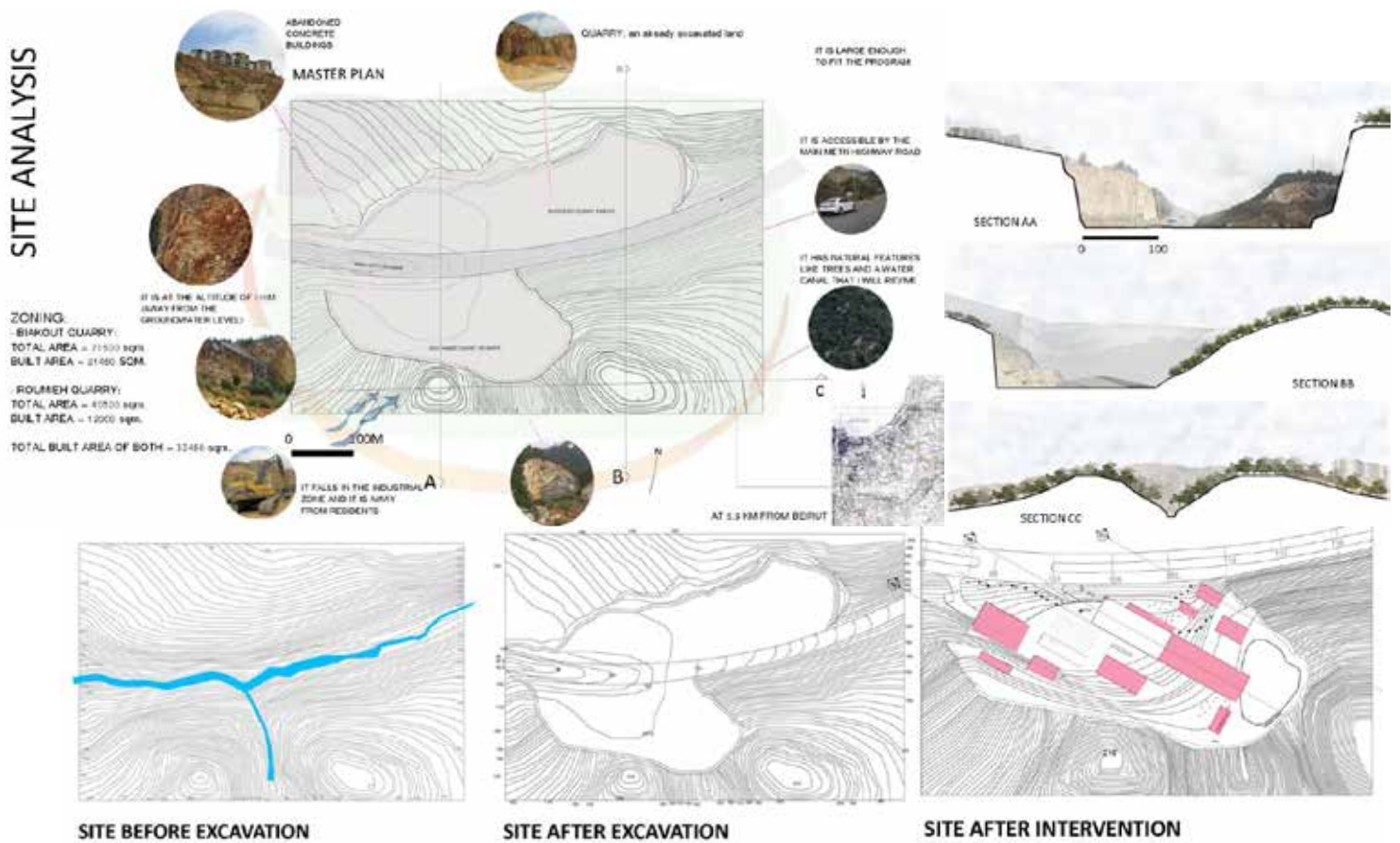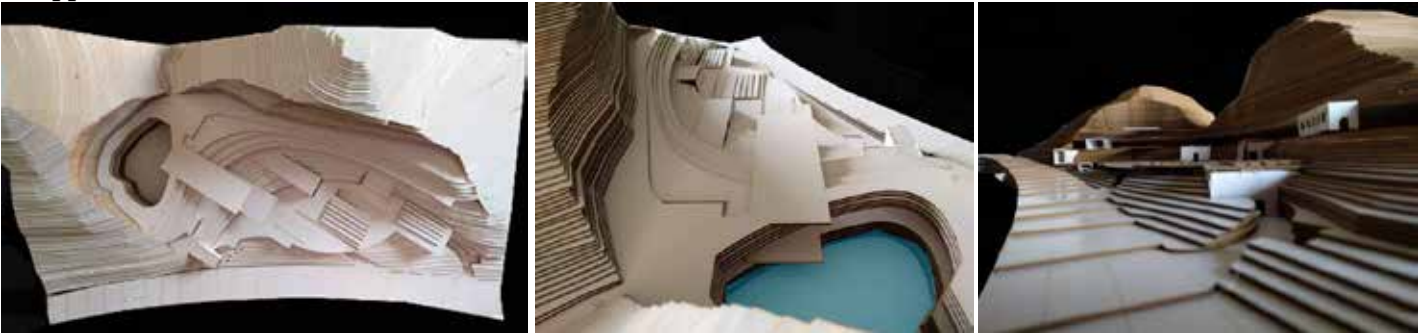
Current model showing the design proposal

SECTION BB

Section and elevation showing how bales of trash can be used in façade
systems and for landfilling, buildings with planted roofs, and the
rainwater collection

COMPLETE FACADE
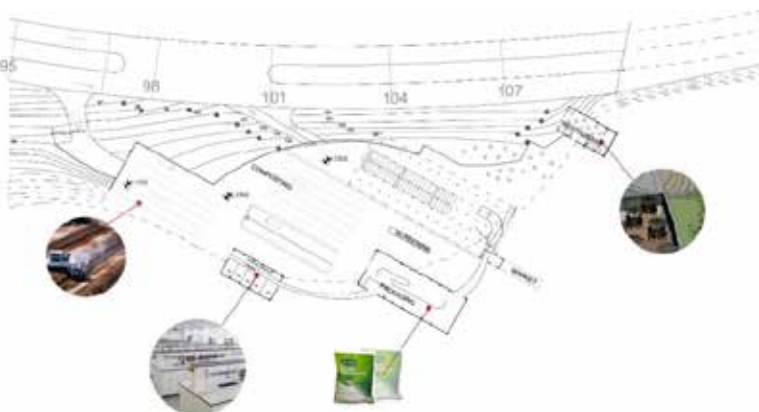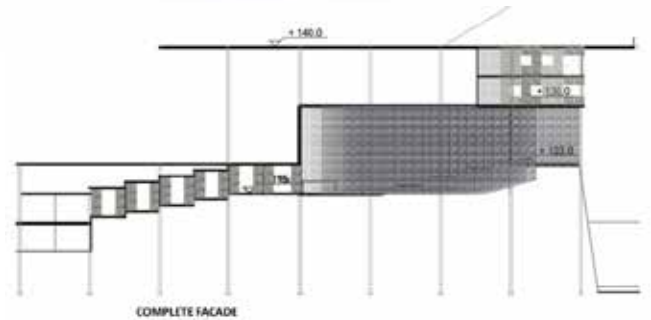
Plan showing functions on site

Tackling Waste Management
A Visual Proposal for a National Crisis

Sara Takkoush
Department of Architecture and Design
Faculty of Engineering and Architecture
American University of Beirut
Bliss Street Riyad El-Solh
Beirut 1107 2020 Lebanon
Sara.takkoush@gmail.com

Abstract - **A graphic design thesis that looks into the different aspects of the waste crisis as well as strategies and directions a local visual campaign on recycling/waste management could take.**

I. INTRODUCTION

Lebanon has always had infrastructure issues but nothing was quite like the garbage crisis that started in January 17, 2015. It all began when the Naameh landfill was shut down by activists and locals living in its proximity. The Naameh landfill is a piece of land just south of Beirut, used as a national dump for most of Beirut and its suburbs' trash. Opened in 1997, it had been proposed as a temporary solution (10 years) with a capacity of 2 million tons of garbage. 18 years after its inauguration it was harboring over 15 million tons of waste. Set to close in 2004, it witnessed a series of contract renewals that surpassed the landfill's capacity and slowly endangered the inhabitants around it. After January 2015, the deadline was extended twice for a period of three months each which caused massive stirs between activists leading to the final closure by protesters of the landfill. The inability to access the landfill by the collection company Sukleen caused the garbage in the streets to pile up and rot for weeks. This scenario, unfortunately timed with the expiration of Sukleen's contract led to our national garbage crisis. Although we're not legislators, graphic designers have the power to contribute in times of crisis. We are technically trained individuals capable of creating impact and simplifying communication into visuals. Responsible designers are those who feel a sense of urgency and responsibility towards their community and society, and aim to use their skills in the best interest and benefit of the population. As a graphic designer, I understood the role and power we have when it comes to communicating values and educating masses particularly in social matters that concern a big population. We have the tools to make technical and complicated issues seem simple and approachable and more importantly actionable. Today we are on the brink of an environmental crisis, and although it seems unresolvable, there is something the public can do about it.

The 17th of July, 2015 marked the expiration of the Sukleen contract and with it pickup services were set to a halt. Sukleen had been the private company in charge of garbage collection in greater Beirut and Mount Lebanon since 1994. It was also in charge of its own advertising and marketing, as well as garbage sorting and storage, landfilling and (in the past 10 years) recycling. After some negotiations, they resumed garbage pickup only in Beirut, and are storing waste or disposing of it in unsanitary landfills or arbitrary dumps around the city (the most famous being at the Monteverde area). Unsanitary landfills and dumps are often in sensitive areas above water aquifers where our rain water gets stored. This water is used for irrigation and residential service water. Soon enough all the leachate or garbage juices will seep into residential tap water and irrigation lines infesting the local agriculture and by extension the food we eat. The public doesn't seem to realize the gravity of the matter at hand. The current "solution" proposed and currently implemented by the government is not the most ideal or beneficial. But there is an alternative solution. There are ways to solve our crisis, by building on pre-existing infrastructure; making use of graphic design, we can launch a successful informative campaign to make a difference. Whether through awareness about how to dispose of waste through sorting and recycling or through actual design of a system and implementation of a program, there are many things that need addressing when it comes to the visual field.

Following the basic waste management pyramid we can begin with reducing, re-using and recycling our waste before we consider energy recovery and disposal. In Lebanon, we have a considerable amount of NGOs and associations currently taking on recycling and sorting initiatives, but we are in dire need of communication material that reaches the audience in ways that could alter behaviors, spread awareness and teach exact specifics about how to sort and manage waste. But this crisis is a national problem and shouldn't be only tackled by organizations, because it's first and foremost the government's responsibility. which is why, for my final year project, I would like to explore the voice of the government when addressing the general public in this situation which brings me to my problematic: how can I design and create a cohesive governmentally launched campaign that is capable of tackling the many branches of waste management in Lebanon, while addressing multiple publics (segments of the Lebanese population) and pertaining to the local and current context?

I will be looking at a variety of frameworks and cases throughout my paper, discussing and picking out what pertains to the points I make. We are looking at a broad spectrum of communication channels that will need to represent the voice of the government hence the need for a flexible or rich identity. My focus will be particularly about sorting waste rather than reducing and re-using it. Reduction of waste requires deep changes in practice, while re-use is easily a saturated practice

whether through the re-use of jars and bottles for storage or use of a variety of other materials for creation of artwork, it reaches eventually a kind of limit where you no longer need to store or need recycled artwork . My research will focus on national recycling campaigns such as Recycle Now in the UK, Keep America beautiful, and I want to be recycled in the US (and other American campaigns), local recycling and waste management initiatives and campaigns such as T.E.R.R.E Liban and Arc-en-ciel, but also the local national garbage pickup company Sukleen's communication. I will also investigate some local governmental campaigns, and communication strategies that will allow me to tackle the issue of recycling in Lebanon.

II. Basis of the Campaign

1-What is this about? The problem, in more tangible detail

The average Lebanese produces between 0.8 to 1.5 kilos of garbage per day. Approximately 50% of it is landfilled, 30% is openly dumped, 12% is composted, and only 8% is recycled. Our garbage is mostly made out of organic matter and food due to our Mediterranean lifestyle and service sector focused economy. The solutions proposed so far include new sanitary landfills, incineration, and deportation/export of waste. Now, while incineration has been a crowd favorite due to the small space required for its installation, it poses huge environmental problems whether due to the toxic gas that it produces into our air, the residual ashes produced from burning the garbage that is hard to control and difficult to dispose of, and due to the very nature of our garbage. Lebanese garbage is made of 50-60% organic material which contains high amounts of water requiring a larger amount of energy to burn making the incinerators very expensive and consequently not the best solution. Landfills are a problem due to their high maintenance and size which seems to be a very prominent issue in a small country with mountainous terrains like Lebanon. The latest proposals have been focused on deporting waste to foreign countries, but this will cost between $200 and $300 per unsorted ton of waste, in a country that produces 4000 tons a day, the numbers are staggering. Although I advocate none of these solutions, incineration and deportation could be made much more efficient if the garbage were sorted. Which is why regardless what the decisions are, it is imperative that we sort our garbage, and treat it as a resource rather than a nuisance.

Essentially, the problem is clear: we have a considerable amount of garbage that is treated like waste rather than a resource, and instead of capitalizing on what we produce, we're looking at ways to dump it. Granted, the solution is multidimensional, and ranges from infrastructural construction to reforming education curriculums in schools in order to educate the younger generation about the need to recycle. But let's say we were to start educating the public about this issue to get them involved; it would need a massive national initiative that targets everyone, only because everyone is directly concerned with the garbage problem.

2-So who's talking? The choice of a governmental over a non-governmental voice and direction

Lebanon has well over 100 NGOs, according to "daleel madani", concerned with all sorts of matters ranging from domestic violence to social inequality, children's rights all the way to environmental causes. The reason we have this many, is because of the array of issues that the country faces with no response or full participation from the government; The Environment is just one of our many causes. Non-governmental organizations have actually been very active in Lebanon, but sadly pass unnoticed due to their localization and size. T.E.R.R.E Liban was founded in 1995 just as Sukleen was launched, and remains one of the most active NGOs of the Lebanese eco movement today, partaking in school programs, installing 3 bin sorting systems and even offering paper pickup services. Many others like "Nidaa el Ard", "Beatuouna", "Green Line", and "Arc En Ciel" are also very active when it comes to spreading environmental behavior through working on different solutions; some of which were successful, but also many efforts that were rendered ineffective. The problem with NGOs is that they always go back to a group of people who are working within the organization, meaning that the public they address is also very targeted and local to their activities. "Nidaa' el Ard" has been one of the most successful organizations in the country, promoting village recycling and educating house-wives to sort their waste in their backyards and then assigning a kind of waste for different houses until they started receiving funding and donations from the UN, the Italian embassy and Germany, and began using that money to create small facilities where you see real waste treatment. But even the founder of this organization admits that she had to self-fund the work from her own money at some point, when there weren't enough volunteers or funds. They're also not very known in the country. Size is a limitation, locality is a limitation but also the resources of a non-governmental organization become an obstacle. At the end of the day an NGO runs on external funding, few employees and volunteers, and that can limit the room for outreach and efficacy. Not to mention that the essential message to be spread when it comes to green awareness is almost the same, aimed towards the same end result. The fact that it is being said by many groups in different locations with different tones and even procedures, confuses the audience and can cause unnecessary and in fact disruptive competition rather than uniting everyone over the same cause.

Which is why, I investigated the governmental voice. Today in Lebanon We have a set infrastructure and a number of private factories that rely solely on salvaged plastic or organic material in their industries. Sadly, the government currently does not promote them nor highlights their contribution to the waste sector. I believe it is the government's job to educate the concerned public about the existence of these factories, as well as to ensure transportation of recyclable materials to them, and benefit them in the market by offering them industrial benefits. The government is supposed to be a voice of authority in any country, capable of legislating laws, enforcing fines or creating urban modifications in service of citizens. This source of authority also has an allocated budget for any campaigns it chooses to endorse, and can mobilize many more people in contrast to the NGO scenario. There's also an advantage when it comes to spreading the message through varied communication channels, since the government has access to a wider spectrum of outreach.

Which is why I will be taking on the position of the government when addressing the general Lebanese public in this national campaign. This will allow me to take on a voice

of authority and to capitalize on the government's resources and ability to create policies and tangible change. I now have the opportunity to design a massive multi-channel multi audience campaign, allowing myself a reasonable access to a wider array of mediums. The government will have to stand as a uniform entity in front of the general public, hence the need of a very cohesive campaign brand that encompasses the content of my proposed campaign. The goal is to create a campaign identity that is cohesive enough but also versatile across mediums, capable of representing a voice of authority. I will be looking at international campaigns such as the national recycling campaign in the UK called Recycle Now, and the national American campaign "I want to be recycled" by Keep America beautiful.

3-To Whom? The audience, with an alternative take on traditional segmentation

A very essential part of any communication campaign is assessing the target audience. When the cause at hand is general and can target almost everyone, we find a need to split the message into parts in order to address the segments more effectively. A classical configuration is usually based on the following 4 types of segmentations: Geographic, Demographic, psychographic and behavioral segmentation. These are interchangeable and can overlap in service of the message. I needed a little more insight on the Lebanese population, which is why I conducted a survey with questions relevant to recycling in Lebanon. Around 100 people filled this survey; the majority of which were young adults (aged 18 to 25)

When asked about their knowledge of recycling in Lebanon:

5% said they knew nothing, 2% said they knew almost everything. The majority said they knew a little from word of mouth, friends, media and education but not enough to do anything tangible. This segment interests us because it shows us that a very big number of people has the potential to learn more.

When asked how often they recycle:

Only 10% said they always recycle , 20% said accessibility is the major factor in their decision to recycle, 50% said they recycle occasionally if pressured by peers or friends or surroundings, 15% answered rarely, only if the usual bin is absent and 5% never recycle.

When asked if they have ever recycled anything in their life, the majority answered yes (94%), showing that there is some level of acceptance of the issue.

But when you dig deeper to understand what stops the public from recycling, the answers fell under three categories:

Logistics: or accessibility to bins and drop off points

Knowledge: Including skepticism and distrust in the waste management system, lack of sorting knowledge of what goes where, and uncertainty regarding the impact and need to recycle.

And believing in lack of necessity: in fact no one who filled out the survey chose this as an answer, which means everyone who answered admits it is necessary, but we will take into account that methods of self-reporting are not always truthful, and some probably assume it isn't necessary. This question allowed me to pinpoint the important things that need to be addressed in the campaign.

The last thing I asked was what motivations could enhance their decision to recycle? And I collected and divided the results into 3 categories as well

Direct reward system such as reverse vending machines (20%)

More knowledge about how to sort and what happens after (30%)

More accessibility to bins and drop off locations (50%)

We see a pattern in the research that indicates that the lack of knowledge as well as the lack of accessibility and motivation to access bins is a major problem that needs to be tackled. This survey allowed me to not only understand my audience better, but also to understand how to tailor the message specifically so that the content of the campaign is relevant to the public.

One must not disregard classical means of segmentation because they do help when it comes to message delivery. A children's campaign designed for a school hallway billboard needs to have a certain aesthetic quality to it, using bright colors, more illustrations and engaging characters in order to reach out to the child looking at it. The message needs to be relevant to the age group while being engaging enough to encourage behavior change.

We see this in a local campaign by T.E.R.R.E Liban called Papivore Malin, an initiative created in order to collect and encourage paper recycling in schools and offices. This campaign makes use of a tree character always saying something and making eye contact with the observer. The decision to have a mascot was successful among children, because many years later, as soon as they saw the tree, they were capable of identifying the campaign.

Similarly, when designing for adults, one must also take some visual decisions that pertain to adults more, not only as an age group but also as a psychographic segment. You would address employees with a higher education degree differently than you would address possibly illiterate street vendors or taxi drivers. Just as you might focus on an informative or didactic tone with adults, over a friendly educational tone for kids. The communication method needs to be either universal enough to cut across different segments, or very targeted and well placed so that it maximizes effectiveness. Something universal would be like the Keep America beautiful campaign called " I want to be recycled", where one visual concept targets the entire population, In contrast with Recycle now's approach to separate the communication depending on the audience. The campaign contains a solution for the multi-audience issue, while fitting cohesively under its brand, and that's the "WRAP resource library" of campaign content. The patron company of Recycle now, WRAP, hosts a database of varied prints and publications that all fit under the brand, with the similar tone of

voice and images of the campaign. However, the content is classified in categories where you could select "office" or "school", and have access to a variety of posters and stickers that you could print or even customize for your use, and then print. Which audience targeting scheme is more effective, depends on too many factors to be able to measure. But if this informs us of anything, it's that we can stretch methods of communication across audience segments. We must always keep in mind the language specification of the Lebanese population, as well as the tone of voice for the campaign.

4-How? Methods of diffusion and communication strategies.

How do you address the Lebanese audience? And how do the results of the survey feed into my methods of addressing the public?

The classical methods and strategies one can resort to vary greatly. According to transparency.org, a successful campaign is tailored to specific audiences, generates a sense of communal responsibility, and uses clear images and prompts that increase human agency and makes use of successful strategies; According to Catherine Mann, at Transparency international, in order to have a successful campaign you need to tailor it to your audience following these 4 steps:

A- Use existing attitudes

It's very difficult to convince the public when the behavior required conflicts with their personal beliefs. 2- Make the issue publically accessible: the more technical and scientific, the less the public will relate or understand it. Recycle now released a series of videos that explain in very simple terms and visuals how a recycling factory operates, in order to teach the consumer how their waste is being recycled. 3-Make the issue culturally specific, and take into consideration that the more personal the campaign feels the better it resonates with the public. Kunhadi capitalizes on the association between the local traditional coffee cup, and the image of parents in their New Year anti drinking and driving campaign that went viral in 2012/13, and 4- Look at the issue from the target audience's point of view designing a targeted message that is relevant to a segment of the population. Another Kunhadi campaign targets the younger drivers, those who belong to a digital age of texting and driving. The billboard campaign simply had a local name (like karim and rana) followed by a "last seen at 18:45". It was launched around a time where social media and mobile communication tools like "Whatsapp", where you would normally read a last seen, were becoming very widespread. The campaign was tailored specifically to the age group that texts, and drives. The association we have of the content with the logo even without more explanation tells us that the driver was texting, then got in a car crash and therefore is now offline since the time of the car crash.

B- You generate a sense of communal responsibility by:

Making the issue socially unacceptable: If you show how individual behavior can bother others it becomes a community issue that you would more easily address because everyone is involved. In Canada, a measurable amount of people recycle only because not recycling harms their others. Highlighting the wider impact: a campaign that demonstrates the positive effect

of minor behavioral change on the community can become successful due to the sense of responsibility that comes with not practicing the behavior and thus depriving the community of the positive effects. Use of shaming: when putting mistakes in the spotlight you make use of 2 methods of getting your message across. By shaming the behavior you make sure the person responsible doesn't do it again, and the observers of the shaming don't make the same mistake. In a Lebanese city called Roumieh, sorting and recycling were being carried out by the municipality. If the residence did not sort their waste, they would get a big label pasted on their trash that said "This household does not recycle", and would have the bag left on the sidewalk in front of the building, to turn the spotlight on this house's negative behavior.

C- You increase the sense of agency by:

Developing a sense of self control as increasing people's sense of control can cause them to react against the issue which they at some point felt powerless against. So when you create a curbside recycling system, and you provide the public with a schedule, they feel in control of what they choose to do with their waste now that they have the options.

D- Offer alternative behavior

If there was ever a cause that needed our contribution, it's almost useless to advocate it if we can't create action to serve it. The Follow the frog campaign, by the rainforest alliance takes a massive global issue which is the destruction of rain forests and turns it into an actionable step that you can take by being a responsible consumer and buying certified rainforest alliance products rather than environmentally abusive products.

According to the Journal of Advertising Research, shock and fear strategies have been employed in social campaigns for a really long time, but there is no empirical evidence of their efficacy. As a matter of fact, a number of new evidence points out that fear based campaigns are effective in creating public attention but ineffective when it comes to behavior change, because individuals tend to dismiss situations as unlikely and distant. This kind of tactic would be effective when intending to hook someone onto a campaign or getting public attention. According to executive planning director at JWT, Angela Morris : "There are more varied approaches in the market now due to increasing depth of understanding of what it takes to motivate behavior change. The old assumption that a shock or fear based approach is always the way to go is no longer valid. For example, with young people, optimism bias sometimes means they self-exempt from the consequences since they don't think it will happen to them, it doesn't matter how shocking it is, and many older audiences have become desensitized to the shock over time", this statement goes along with most recycling campaigns that I will be looking into.

III. Components and Content - The Message

Looking at what's been done internationally:

Keep America beautiful is one of the oldest environmental organizations in America. Their recycling initiative is called "I want to be recycled". This award winning campaign launched in 2013 in partnership with ad council is a series of cute short

films, billboards, a public intervention, a website with infographics and pickup location search, and even a game and advocacy tools that allow anyone to advocate the cause. The campaign can be divided into 3 parts: motivation – manifested through the main billboards and videos, 2- education: manifested through the game, the infographics and information about the materials and the factories online, and 3- action: the advocacy page and recycling instructions to allow you to begin at home.

The billboards of this campaign as well as the videos take on a very optimistic and positive tone. It expands on the concept of trash that would like to be recycled in order to come back as something else. The colors dominating are a relaxing blue and a very fresh and clean green. The size of the typography is as dominant as the central element of the photography. The use of the link in the image is well integrated while being linguistically relevant www.iwanttoberecycled.com with even quotation marks that reference the recycling symbol. We see an integration of the message within the visual. We also see a personification of the garbage, giving it life to give it importance, which is a way to show empathy and relate-ability, something we must apply according to a TEDx talk by According to a TEDx talk by journalist Simran Sethi.

Another campaign that plays on personification is the Canadian campaign: Simply Recycle. They also capitalize on the need to relate garbage to the consumer and focus the tone of voice on making garbage speak, but this time to the recycling bin, creating a one way dialogue that involves the garbage flirting with the recycling bin. We also see a dominance of green and blue, clean garbage and speech bubbles. The campaign visuals also come with recycling tips on sticky notes that make the process seem familiar, homey and easy to do. This approach positions the campaign with the domestic campaigns where it addresses mainly residential areas and houses.

A national example with a wider target audience would be the national United Kingdom campaign "Recycle Now" supported and funded by the government managed by WRAP and used locally by over 90% of English authorities and municipalities. Launched in 2004, it offers a huge range of resources available for free to be used by consumers all over the UK. The campaign was designed by "how on earth" a strategic marketing and communications company that specializes in socially responsible causes. They realized that 70% of the population claimed that they recycled when in fact only 12% of the waste was being recycled. Which is why they created the recycle now brand. It comprises of an emblem: the swoosh – a dynamic logo that can handle city variations, a color scheme and illustration style, a voice and tone as well as photographic style that focused on portraying happy and positive messages when it comes to recycling. According to environmental behavior research by Richard Osbaldiston and John Paul Schott, showing positive attitude when initiating any behavior is more likely to encourage positive attitudes towards that behavior by the viewer. Recycle now pushes the photography medium to every location where recycling is possible. The campaign follows a segmented system where target not only every segment of the population differently but also prepare different visuals for different locations following geographic segmentation. The campaign includes videos and animations, infographics and posters, a variety of publication all available

to be printed, and one of the most important aspects of it, is the icon and color system used for bin signage. They also designed the actual waste following a very vibrant and colorful style. Due to the open nature of the campaign we see differences in the communication strategies. In some the authority behind it encourages recycling by telling us how limitless it could be, and in others the municipalities frown upon people's excuses. In some there's an informative approach to convince people why recycling is so important while in others there's focus on accessibility and commonality of the act of recycling and how ubiquitous it could be especially around the house (bathroom recycling bin campaign, kitchen recycling campaign etc...). If we dissect the components of this campaign we see a configuration of 1-encouragement and motivation through posters and positive messages, 2- informative material through videos and publications and finally 3- operational material including schedules for pickup and stickers for bin sorting.

Just to show how important it is to provide the public with information and knowledge, an American non-profit called "Recycle across America" advocates and believes that knowledge of proper sorting is enough to kick start better waste related habits. They launched a campaign called "let's recycle right" where they designed standardized labels that aim to teach using photography and color coding the importance of proper sorting. They resort to featuring celebrity endorsements and partnering up with concerned businesses in order to perfect the sorting system. Their input is relevant to the last step of the message: instruction. Teaching the public how to sort based on images on a trash can.

There are so many more campaigns out there, but maybe if we look at what the Lebanese market has to offer we could have a better understanding of the situation.

In Lebanon, waste has been managed by a private collection company called Sukleen. Funded by the government, it was in charge of garbage pickup, street sweeping and eventually recycling and composting. 10 years ago sukleen was using a 2 bin system for sorting. They had their material recovery facility in Karantina and were trying to take out as much recoverable and reusable material as possible after using compacting trucks that increase garbage contamination. They sold the recyclables to local factories that depend on them in their production. But it wasn't until April 2014 that they launched the RED and BLUE program. The latter has been the most successful local recycling campaign by comparison to others campaigns due to the most outreach it has received. The initial visuals used a simple and bright color system with basic clear bold typography in white over solid color. To launch their campaign they contacted businesses first and convinced them to partake in the initiative as part of their Corporate Social Responsibility programs. Slowly this initiative began growing on social media, and got broadcasted on Virgin radio as well as Facebook and Instagram. They then expanded into domestic models where they would distribute large bins for buildings and residential areas as long as a pickup date was agreed on by the building and an amount of houses pledged to sort before the bins were delivered.

But this campaign actually had a lot of problems regardless of its outreach. First, They were using a very artificial plastic blue for the paper and cardboard bin, and an apple red for the

plastics metal and glass bin which can be very counter intuitive, rather than sticking to white for paper and blue for plastic glass and metal for example. The reason behind this decision was the presence of other local campaigns before them (NISR by TERRE Liban) that used the color red for solid waste recycling, but the difference is that the other campaign used white and green for paper and organic recycling. So by comparison, the choice no longer made sense. People up till now still get confused every time they try to dispose of something due to the color connotation, having to stop and read to check every time is counterproductive and makes us no longer prepared to acquire the behavior. After asking some janitors and waste collectors about the content of the bins I received the same answer over and over: "AUB students treat these bins like normal trash cans dumping everything in them carelessly". This is a big problem. Some students have resorted to switching the lids of the boxes to make the colors more intuitive than what they are, but this becomes confusing due to the shape of the lid that is tailored to fit the material the bin is supposed to hold. (Rectangular slot for paper) In response to that however, Sukleen has revamped this campaign in October 2015, using a photographic approach rather than illustrative/icon based visual language due to the apparent confusion that the public was demonstrating when disposing of recyclables. They also minimized the use of the color on the bin sticker because 1- the red faded away in the sun with time, and 2- because it's redundant to have a red sticker on a red bin, and they needed the white background to show the details of the images on the bin. Now this approach should be a lot more effective because it minimizes room for error because you see exactly what goes where, answering the question of where do we put what.

Although Sukleen's work is seen on the streets and in institutions, it doesn't focus enough on awareness at a domestic level, which is why other NGOs have taken the reins on this one. "The three instead of one" campaign launched by T.E.R.R.E Liban is an old revamped initiative that encourages sorting in 3 bins. They highlight the use of an agenda for garbage pick-up per kind of waste. They advocate composting and sorting garbage to reduce landfills. Their main audience is found in the suburbs of Beirut particularly in Baabda and Hadath. Their actions are also quite centralized to those areas although they have been present for a long time. The problem with this campaign is that the visuals are quite primitive and incoherent. They make use of the Lebanese flag colors with a mix of photography, illustration iconography and almost randomly placed typography, with no clear sense of hierarchy and structure. They use their posters (in this case social media posts) as a platform for their demands and policies such as "no to incinerators no to landfills and no to sea landfilling" as well as an awareness tool with minor explanation of where to sort what. This message can be easily separated into the different messages it constitutes and reformulated in ways that address the public more efficiently. They use social media and protests as their main means of communication.

Another rather successful campaign that was launched in response to the Naameh landfill closing, "Sar lezem rassak yefroz" which translates to "It's about time you start sorting" by Cedars environmental. Designed by Intermarkets in December 2014, it was launched in January 2015 on social media, by email, and using door to door techniques. This campaign advocates what is known as single stream recycling,

by separating waste into only 2 bins. One for organic waste and the other for all recyclables combined. The benefits of this approach are: 1- narrowing the margin of error. 2- Occupy less space in the household, 3- minimize confusion when it comes to what goes where and 4- facilitating pick-up where we only need 2 separate trucks to 2 separate destinations. This plan has been successfully implemented in many municipalities with inhabitants of ten thousand or less. The organic waste black bag gets composted locally, at municipal spaces, and the blue bag gets transported to a material recovery facility (one can be easily provided to the municipality) where it gets sorted by professionals and then each kind of garbage gets treated accordingly and delivered to concerned parties. The entire campaign is summed up in one poster. It was eventually followed by a series of lectures and eventually a simple phone game where you are quizzed on which garbage item belongs in which bag. It's overly simplified and contains very minimal but crucial information on the kinds of things you can put in each bag. The poster uses flashy colors that in no way inspire or remind us of any other green movement that's been conducted, opening up the use of any color for a recycling campaign and not just blue and green. The key driver of this campaign was the research conducted on bag color usage and household habits. For a very concise campaign, it does a good job at delivering the message, however with the widespread of technology and social media, this campaign could have been better tailored to digital media. It could have also been followed up with more instruction and education. It combines all messages in one and compromises on the content of the message due to it serving the purpose in one visual. The format is very classical for posters while the majority of stakeholders saw it online. Targeting was not addressed here, but it spoke to everyone using the local colloquial Arabic dialect. And it does explain why sorting happens according to this system. But in no way does it feel like a complete system, because it stops after giving you the basic minimum information.

When it comes to communication language, If we look at Sukleen's material, we see that it's always bi-lingual (English/Arabic), while the eco movement's visuals use mostly modern standard Arabic (fus-ha) for their communication, in contrast with the sar lezem campaign which uses colloquial Arabic as it is spoken in Lebanon (pertaining to the local scene through language), and arc en ciel's campaign which is mostly in French and English. The choice of language is very important according to Antoine Abou Moussa, environmental consultant and project manager at TERRE Liban. Based on his field research, he documented that a majority of the people creating initiatives and advocating recycling in the country are using foreign languages as their primary communication tool, in the process losing a big segment of the population that would rather read in Arabic. Which gives me a direction of the kind of language that needs to be used and how. It becomes clear, that in order to successfully communicate I need to trigger the public or motivate them with something that will provoke them, inform using local information on factories and plants which make up the infrastructure for recycling and final instruct them on the exact details and process of recycling all while being pertinent and using a language and tone that speaks to them.

IV. Channels of Communication

1-Classical and alternative ways to address the public

If we were to think of the origin of print in communication media, we would find posters designed for streets and venues to be the first form of public communication. Up until today, we still see posters being designed and pinned all over the walls of the city whenever an event is coming up. Granted, a campaign doesn't usually advertise on a wall, but uses billboards and now animated led billboards to get its message across. Whenever you would develop an initiative, you would also follow it up with a publication or flyers or any kind of printed medium that allowed your message to be carried on without you physically being there. We're in 2016. The world has changed, and we can argue that people spend more time staring at screens rather than street walls. Which is why we cannot under-estimate the power of online presence as complimentary if not dominant channel for communication. Nearly everyone has access to the internet, and a big part of the public is very active with smartphone and internet penetration reaching 70% of the local market. Social media and online personas have changed the way we communicate and spread awareness. And although billboards and classic forms of communication are valid, they are sadly no longer enough. Which is why a very crucial aspect to consider for my campaign is the channels in which it spreads the message.

In July 2011, coca cola launched the Recycling king as part of their give it back campaign. They added over ten thousand recycling bins around Israel, and created Facebook locations for them. For every can or bottle you recycle you can document your encounter with the bins and check in online. The person who was most active, wins the title. By the time the campaign was over everyone knew how easy it was to recycle and access bins online. This kind of campaign starts with small steps but capitalizes on social media to spread awareness in somewhat non-traditional ways. No print advertising was created for this campaign.

Focusing more on locality, street posters are not nearly as effective as posters in an elevator or a restaurant for example, because when in those locations, one is circumstantially inclined to look at print. The better the placement the more effective the message delivery. In November 2015, Trident gum has launched an interesting campaign that includes very strategically placed ads in specific locations with specific taglines. The main concept of the campaign is that chewing gum keeps you focused; so they placed some ads in the entrances of parking lots with the tagline "someone's about to take your spot, chew and focus". Or at a bottleneck traffic zone with the message "don't lose your cool in traffic – pick up the refreshing habit". The campaign uses bold black type on flashy background colors and is very message oriented. What's interesting here is placement decision. From placement, we are capable to decide how to maximize the effect of a message just through studying the location. The children oriented campaign would be placed in and around schools, the business oriented campaigns in and around company head- quarters etc. The placement of the recycling campaign can make or break in message delivery. The message can even be placed on the recycling bin itself. Sukleen uses on bin information to indicate what each bin is for – arc en ciel do the same, while cedars

environmental use the color of the bag as the only indication, this is a form of information delivery. International "Recycle now" have their messages sometimes on the bin sometimes right above the garbage shoot, all decisions which maximize contact with the audience. A message placed above the garbage can allows me to spot it from a distance; this affects behavior and tackles bin visibility problems.

After investigating the different audiences and different ways we can segment the public into publics, it goes without saying that we can have a little more flexibility and innovation in the way we address them. Seeing that I can embody the role of the government, I can investigate in alternative communication channels such as "on product" recycling logic, where all the local Lebanese products would have a label that indicates where this item goes, with minor instructions as to carefully dispose of it. This idea has actually been implemented but slightly differently by OPRL in the UK, where a number of private companies volunteer to participate and have a little icon added onto their packaging indicating the specifics of how and in which bin should the item be recycled. This is especially helpful on slightly complex materials such as multi Package or multi material (composite) items like a milk carton with a plastic cover, or a tea box with plastic wrap and inside paper, or any kind of composite elements. The label is the "Recycle Now" swoosh framed by a black or green square. Green for recyclable, black for check local recycling services and black but crossed out for non-recyclable. I personally didn't notice this at first. Color usually is easily discernable and maybe they could have gone with a grey or white for "check local recycling". The consistent use of the patented Recycle Now swoosh makes OPRL fit very cohesively under the campaign. The choice of color, and informative tone of voice also falls well under the brand guidelines. This form of communication is revolutionary to say the least because it takes delivering the message to the absolute limit. You become aware of the product's recyclability while holding it for the first time. Other interesting channels would be the actual shape of the recycling bin, where we add a layer of meaning when designing the bin in order to emphasize or convey an idea through shape. This can be simplified in application by just altering the shape of the opening thus communicating through the design of the hole that takes in your trash. This also helps our reflexes realize if we're putting something in the right place. The sukleen office size bins follow this logic by having rectangular slot openings for paper bins and circle openings for other recyclables such as bottles cans and other materials. We can always analyze the shape of the opening and its relationship to the object it's meant to take in, but at this point, getting the message across is much more important.

Other innovative channels of communication include garbage dialogue, where the trash bin can carry a message that is meant to be read when disposing of garbage either in an interactive way (every time you dispose of a recyclable you get information or appreciation), or even a real size way finding system that leads the person to the concerned trash bin for the respective trash while using stickers or environmental interventions that encourage recycling. There are many interesting things that can be tested and tried and I believe this kind of innovation is what the sector needs in order to initiate a real national effort.

V. Conclusion

After surveying the Lebanese target audience, isolating the problem and investigating international solutions and communication strategies, I can wrap up what I believe would be the perfect solution for this problematic. I will need to create an entity that is present and prominent as part of the government - it will be given a brand – a logo - an emblem that will represent its authority and presence on all the media that it produces. Its responsibilities will be to first of all grab attention through a teaser campaign that can be but is not limited to a public intervention, or a visual social commentary. Then I will decide on all the material needed for the citizen to trust completely in the process of recycling, including company and factory names, factory owners and locations, Material recovery facility locations and numbers, dumps and landfills, as well as the prices and capacity to export well sorted trash. This information will then be disseminated in waves, either under sub-campaigns or activations of the major campaign. The goal is to teach at this point, to gain trust, to build a strong bond between the campaign and the public. I want to reach a stage where your sense of nationalism makes you want to recycle. The tone of voice will most probably be positive overall, but might contain comedy, sarcasm, or even dark humor in order to motivate the public. I would like to have a bin design system where the bin and its signage and instructions would be detailed and explained to the consumer. There will be a strong focus on channels and placement of the campaign, tailored to the governmentally accessible space in the country that allows the message in its many layers to be delivered. And hopefully this will allow the maximization of outreach and will produce national impact.

REFERENCES

Communication Strategies:

Andrews, Marc. Social Campaigns: Art of Visual Persuasion. Its psychology, its semiotics, its rhetoric. MA Thesis, 2008.

Floch, Jean-Marie, and Robin Orr-Bokdin. Semiotics, Marketing, and Communication: Beneath the Signs, the Strategies. Houndmills, Basingstoke, Hampshire: Palgrave, 2001. Print.

Andrews, Marc, Mattheis Lars Van Leeuwen, and Rickert Bart Van Baaren. Hidden Persuasion: 33 Psychological Influence Techniques in Advertising. N.p.: n.p., n.d. PDF.

Moore, Bridget, Designing for multicultural and international audiences: creating culturally-intelligent visual rhetoric and overcoming ethnocentrism, MA thesis, English Iowa State University 2010. PDF

Rose, Chris. 12 BASIC GUIDELINES FOR CAMPAIGN STRATEGY (2006): n. pag. Www.campaignstrategy.com. Web. PDF

Nation Branding: Concepts, Issues, Practice book - Dinnie, Keith. Nation Branding: Concepts, Issues, Practice. Oxford: Butterworth-Heinemann, 2008. Web.

Recycling and waste management awareness and behavior:

Vicente, Paula and Reis, Elizabeth Factors influencing households' participation in recycling. UNIDE, Department of Quantitative Methods, IBS - ISCTE Business School, Lisbon, Portugal.2008. PDF

Burn, M. Shawn, Social Psychology and the Stimulation of Recycling Behaviors: The Block Leader Approach. Journal of Applied Social Psychology. July 2016. PDF.

Republic of Lebanon ministry of Environment, Management of Recyclable Material for Lebanese Municipalities - Handbook for the proper preparation, handling, processing & selling of recyclable materials in Lebanon. 2011. PDF.

Werner, Carol M., and Eeva Makela. "Motivations and behaviors that support recycling". University of Utah, Salt Lake City, Utah 84112, U.S.A. University of Jyvaskyla, 1999.

Janz, N. K., and M. H. Becker. "The Health Belief Model: A Decade Later."Health Education & Behavior 11.1 (1984): 1-47

Fishbein, Martin, and Marco C. Yzer. Using Theory to Design Effective Health Behavior Interventions. N.p.: n.p., May 2013.

DARREN W. DAHL, KRISTINA D. FRANKENBERGER and RAJESH V. MANCHANDA (2003). Does It Pay to Shock? Reactions to Shocking and Nonshocking Advertising Content among University Students. Journal of Advertising Research, 43, pp 268-280

Osbaldiston, R., and J. P. Schott. "Environmental Sustainability and Behavioral Science: Meta-Analysis of Proenvironmental Behavior Experiments." Environment and Behavior 44.2 (2011): 257-99. Web

Ellis, Kail C. Lebanon's Second Republic: Prospects for the Twenty-first Century. Gainesville: U of Florida, 2002. Print

On Recycling and Waste:

http://www.getitsorted.org/ | www.recyclenow.com

www.simplyrecycle.ca | www.iwanttoberecycled.com

The Green MED initiative on waste in Lebanon: http://gmiproject.eu/?page_id=928

About nidaa' el Ard - http://www.yourmiddleeast.com/culture/a-ray-of-hope-in-lebanons-dark-age-of-garbage_34488

Marketing segmentation http://marketingflexibility.com/4-types-of-market-segmentation/

The City as a System

Rabab S. Hammoud
Department of Urban Design
American University of Beirut
Beirut
rms91@mail.aub.edu

ABSTRACT: Le Corbusier, an architect, urbanist and social reformer perceived cities as "dying" for their lack of geometrical order. His vision of the city in the modern age was based on strict geometrical grid. Le Corbusier's ideal city presented the _"perfect form"_ that lies in the juxtaposition of pure geometrical shapes. He believed that order defined in the urban form could bring order into society. This radical approach aimed for social equity and justice [1]. However, Le Corbusier's realization of the ideal city was mainly criticized for undermining the complexity of the real city life. Jacobs, the social activist, considered his vision as the application of "anti-city planning to existing cities" [2].

On the other hand, parametricism heads a new era after modernism. This new style is asserting a dominant style ingrained in advanced computational tools. While parametricism realized the simple order presented by Le Corbusier as limited, they valued the organized complexity, for it reveals the significance of life richness. Their research hoped to realize the hidden logic rooted in nature, its hidden regularity. Derived from parametric design systems, parametricism creates a sensitive design from multiple parameters. It creates a systematic, adaptive planning visions, that differentiates each site by its given parameters rather than presenting sterile fixed solutions. Parametricism is the journey of research and innovation that aims to present a deeper understanding of the complexity of the society [3].

Thus, I argue that parametricism presents a richer and _"evidence-based"_ approach to the city due to its flexibility and research oriented approach. It is derived from sophisticated programmatic tools that aim to articulate the city in computational models rather than presenting a one man's vision of the ideal city.

PREMISE: The purpose of this paper is to question the city as a system, comparing the vision of Le Corbusier (Modernism) and the vision of Parametricism, considering The Radiant City by Le Corbusier and The Kartal-Pendik Master plan by Zaha Hadid Architects as case studies.

I. MODERNISM AS THE GLOBAL STYLE OF THE 20TH CENTURY.

Le Corbusier was one of the major architects to lead what was then thought of as modern architecture. His architecture was based on his fascination with machines and how they work; he admired the concept to the extent that he turned his buildings to live-in machines. His work expressed this interest by implementing machine concept; a certain law is applied to achieve a desired outcome. The fact that he lived in the machine age no doubt was a factor of the dominance of the machine concept in his philosophy and his perception of its work to be the ultimate product and the ultimate producer. His buildings were an application of his thought; a product of art and technology; the two elements merged to create what he thought of as true pieces of genius. His avant-garde designs are mainly pure forms that were not possible before the new technology was introduced [4].

A. Defining the Style

The surface, volume and plan were Le Corbusier's perception of the design parameters that could be defined as follows. The volume is the first visual perception of the design model. The surface is the reflecting image of the building that holds a functional aspect by its defined openings. The grid-based division of the surface preserves the unity of the pure form. The structure is a concealed frame of order. The plan, lastly, maintains its own order, yet not constrained by the structural frame [4].

Le Corbusier perceived the grid as the main ordering system in the modern era. He regarded everything that wasn't formed on a grid as "accidental". The pure forms defined by "straight lines" and "right angles" are the main constituents of an ideal form [5]. He then rejected the curve as "ruinous, difficult and dangerous" [6]. Le Corbusier's realization was rational. He wanted to place order on plain sites replacing the non-geometrical present sites, thus extending his ideas from architectural morphologies to urban scale [5].

The perfect geometry Le Corbusier pursued extends from the building forms to all aspects of life in the city. He realized that all forms of production must fall into order. The order is then extensive, where the society, as a first step, required organizing [7]. This was realized in a syndical "pyramid of natural hierarchies" [8], where organization promotes a society that works together in harmony. Thus Le Corbusier went beyond the architectural form to be a sort of social reformer. His approach was truly revolutionary; he discounted any old notion of what a city might be. His theory was that a designer needed to start from scratch to achieve the ultimate result [7].

B. The Radiant City

Le Corbusier's vision was shaped in The Radiant City, a city that was first published in 1933 as the *city of tomorrow*. The plan directed the productive life in all its details. It forcibly administrates the market by professionals to match the society's demands to its production [7]. His aim was to improve the society rather its way of life. He created highly dense typologies ordered on a symmetrical grid [1].

The Radiant city presented "*the perfect form*". Le Corbusier perceived the present city as a "dying" city, given its planning was not based on geometrical order. His perception of geometrical order was the juxtaposition of simple geometrical forms. This clear definition was modeled in the Radiant City. The base idea of the city was the defined zoning that segregated administrative, commercial, and housing districts in the city. The city enclosed prefabricated, ordered and identical skyscrapers on vast green spaces. Two hundred meters skyscrapers were designated as the business center and housed eight hundred thousand people. In addition, a massive underground transportation system acted as the main link between the residential areas and the workplace [1].

As for the residential district, Le Corbusier designed the "Unite´s", prefabricated residences in the form of a vertical village. The fifty meters Unite´ houses would accommodate twenty seven hundred inhabitants. The apartments' areas are suited according to the number of the family members regardless of their grade in the industrial community. Le Corbusier's vision was mainly considered for respecting the scale of the human providing the needed not more or less [7].

Furthermore, the prefabricated units entail different recreational facilities easily accessed by the residents. The gymnasium and the swimming pools are provided on the roof of the Unite´. Other facilities such as a kindergarten and laundry are at different floors [1]. The uniqueness of the Unite´ lays in the collective services it provides rather than in being merely a residential building [7].

Le Corbusier's vision was radical to the existing city; he thought that a true realization of his vision could only be materialized after a revolution in the community. Le Corbusier believed in the pure architectural forms the advanced technology introduced. For him, it created the "*architecture of happiness*" [7].

Figure 1: The Radiant City (www.land8.com).

C. Brasilia

Brasilia, Brasil's capital, was built on a clear land offered by the president of Brasil. Even though Le Corbusier was not the urban planner of the city, it was built on modernism theories established by the Swiss-French architect. The city is seen as the largest implementation of Le Corbusier's theories. Lúcio Costa and Oscar Niemeyer, the urban planner and chief architect of the city built a *perfect* grid oriented plan administrated by the authority [1]. The main feature in the city was the total division between different districts. The administrative distract holding public buildings was aligned on a main axis. It shaped the city's civic sector with a monumental theme. Furthermore, the housing district housed five hundred thousand residents. It formed a neighborhood of collective uses. Massive units hosted recreational, commercial, medical and educational facilities [9]. By implementing Le Corbusier's principles, Costa and Niemeyer aimed to construct a city that portrayed "*equality and justice*" [1].

Figure 2: The monumental administration zones in Brasilia (Photo: Vesna Petrovic. http://www.getty.edu/)

D. Criticism

After many decades, it was clear that cities constructed on Le Corbusier's theories didn't live to their premise. Brasilia for instance, was considered a disappointment for not presenting people's lifestyles and aspirations and for not providing a place for city gathering [1].

Brasilia was more than a planned city. It aimed to create social order and revolutionize the community [10]. Nevertheless, the hopes of equity and justice were faced by a different reality. The city presented poorer situations for the workers than the previous cities did, and the classless society turned to be an entirety segregated community [11]. Brasilia, also, witnessed very little street activity. The dependence on vehicular movements produced modest interaction between the residents [10].

The Unité, or the superblocks residences, were built in many major cities. Located on the peripherals they were mainly houses of poverty and crime. Many have been destroyed or modified [1].

The Radiant City as defined by Le Corbusier is a theoretical plan that showed an "anti-city planning" to present cities, Jacobs underlined. The radical approach placed order not only on the physical setting but also on the whole society organization. The perfect form Le Corbusier envisioned undermined the richness of real life and its complexity. The

main evidence of the limitations of this planning approach, Jacobs concluded, was the fall down of these layouts [2].

In addition, the dependence on vast transportation lines to connect the highly segregated districts of the city split the city apart and destroyed the multi activity street and the dense living spaces that shape the urban life [12].

However, the idea of planning order in the city is still a temptation for urban planners now as when Le Corbusier proposed his vision in the early nineties. The ideal city providing a healthy environment free of present city problems such as congestion, pollution and lack of public spaces is still the main concern for designers nowadays [1].

II. PARAMETRICISM AS THE GLOBAL STYLE OF THE 21 CENTURY

Headed by Zaha Hadid and Patrik Schumacher, parametricism is being explored for the last fifteen years to head a new era after modernism. This new style is asserting a dominant style ingrained in advanced computational tools. Introducing itself as the new "*global style*", it states the death of other styles that were a reaction for modernism failure such as Deconstructivism and Postmodernism as Shumacher states. Parametricism, as modernism, is expressed in different scales from urban design to architecture to the smallest details in interior design. However, the complexity explored by the technological tools could be mostly expressed as the scale of the project increases. This large capacity to articulate an urban design has been researched for more than three years. Based on this style more than one urban proposal were designed by Zaha Hadid Architects. Many won the first prizes in different competitions [3].

Parametric design is the product of parametric advanced software. These techniques offer new form formulations. Algorithm techniques, on the other hand, are founded on codes. Going beyond the limitations of design softwares these techniques present the computer as a problem solver. It offers solutions that need excessive time to reach. Parametricism, as Shumacher explains, is the blend of these different design tools. While these tools open new way of thinking they are merely mechanical ones, where the designer can reach different aesthetic forms [13]. While modernists are using parametric tools to reproduce modernism, they leave behind the beautiful complexity these systems offers. Parametricism, however, seeks the uniqueness presented by design systems creating avant-garde models [3].

"*Aesthetically, it is the elegance of ordered complexity and the sense of seamless fluidity, akin to natural systems that constitute the hallmark of parametricism*" [3].

A. Formulating new styles

Parametricism is a style rooted in the exploration of advanced programs, thus it frames a new model of work and introduces new aspirations, techniques and standards. This opened the door for a new style era based on explorative design systems. It is a *radical* shift between styles liberated from previous design accumulations as modernism was. Parametricism could be summarized in series of creative work based on design tools where no definite end could be formulated [3]. These tools were merely part of science fiction in the early 90s, have become a reliant tool for creating unlimited complex forms [13].

B. Defining the Style

Parametricism rejected the ordered grid Le Corbusier believed in. The order that is based on pure forms and the repetition of isolated elements. On the other hand, parametricism believed that all shapes could be "*parametrically malleable*" and gradated in soft angles (Schumacher, 16). It considers the overall system of the model, rejecting the separation between the volume, the surface and the plan that was proposed by Le Corbusier. Its scheme shifts from one system ex. the surface to several sub-systems, "the envelope", "structure" , "internal subdivision" and "navigation void" [3]. Tools of parametric design offer design flexibility, where small adjustments in one part allow for the readjustment of the whole form. For example if one point on a curve is moved the whole curve readapt itself. Therefore, these systems can be helpful in modeling different scale project from single entity to an urban site design project [13].

Based on these arising technological tools, new themes arose. It is then required a move in the framework of design thinking, methods, and theoretical bases [14].

C. The spirit of order

Le Corbusier aimed for simple order presented in primitive geometrical forms. He perceived that while "*nature presents itself to us as a chaos ... the spirit which animates nature is a spirit of order*". Parametricism didn't reject order, however they rejected the limited order Le Corbusier proposed. They considered his definition to nature's order was restricted by the technology of his time. They are rather investigating the underlying order of those irregular patterns by "stimulating their material computation". In his writings Le Corbusier rejected the "patch donkey's path", a path that could be now explored and valued for its fundamental reasoning and intellectuality, thus discovering the hidden order and potential strength that lies behind it [3].

D. Understanding the Complexity of Urban Settlements

Unlike the limitations of the traditional urban design proposals, information systems provide various solutions for a definite urban site. Based on the input of information different diagrams can be derived. It is thus highly important for the research tools it provide [15].

Frei Otto, a pioneer visionary architect, researched the natural models. His main interest was settlement types studying two main processes the occupations and the forms of

connections it develops. The focus on these two forms was due to their existence in any urban settlement. In one of his experiments, he used water with floating magnets to stimulate far entity settlements. His experiments were based on the similarity between these settlement models and the existing urban models [3].

Otto, conceives material configurations that are able to arrange itself in optimal ways and compute the best network system between given ends. However, there is *no one best solution*, for each experiment give different solutions [3].

The multi-layered model promotes any modification in the parametric model to directly being adapted in the model as a whole. Modifying any of these quantifiable parameters often enhances qualitative features. This is where relationality could be best interpreted. The capacity of such models to arrange itself into various ordered systems is astonishing, thus showing the limited capacity of previous *ordered* models presented by Le Corbusier. The grid, Le Corbusier believed in, creates a uniform repetitive model that lacks the adaptive capacity parametricism offers in their models. The freedom Le Corbusier saw in his models could now be seen as definite frameworks where little could be modified [3].

Modernism and parametricism stand from different point of views. Modernism considers a "universal space" giving little to context and the uniqueness of the space; however, parametricism is based on fields where diverse parameters shape each model. Unlike modernism, parametricism establishes no discrete figures or zones with sharp segregation. It aims to introduce *new vision where axes are not found and there are no borderlines to cross* [3].

Parametricism launches new logics through advanced computational technology that highlights the deep relationality in the urban schemes. The parametric system holds the morphological urban form shaped on a network of streets and open spaces. Parametricism opens the door to unlimited number of parameters to shape the urban fabric, where the model could be adapted based on the orientation of sunlight, thus promoting better visual characteristics [3]. Different parameters creating virtual diagrams and proposing formal and functional guidelines that are highly adaptive, create a new structure to architectural and urban thinking [16].

E. Kartal-Pendik Masterplan

Kartal-Pendik, is an abandoned industrial site in Asian Istanbul. An international masterplan competition was held for one of the largest urban renewal schemes to redevelop the city. Designed by Zaha Hadid, the prize-winning architect, 555-hectare masterplan will be executed enclosing high skyscrapers, perimeter blocks and public spaces. The site will host 100,000 people to work and enjoy the city, in addition to visitors coming for shopping and different leisure activities [17].

Kartal-Pendik master plan is a mixed-use urban project setting all programmatic functions that defines a city space [3]. Activating the life in the city streets is the aim of the project. Hadid states that plan should "*animate the ground*".

She aspires to "*add or scoop out a civic space around every tower or building*" in the city [17]. The project objective is to project a new sub-center in Istanbul, thus decreasing the congestion in the city's historical side. Based on the themes of Parametricism, the plan didn't adapt repetitive geometrical units. It considered the street lines from the surrounding context as the main input generator of the urban network. Modeled by Otto's application, parametric generators could adapt the broadening and contracting of the streets based on urban climax. The main longitudinal line of circulation, placed at the center as the main network way, generates a series of smaller roads. This network shaped a hybrid-deformed grid. [3]. The newly defined grid is the ordering unit in the project. A grid that is extremely adaptive to different typological parameters in the urban site[18].

Figure 3: Kartal-Pendik, the path network generated minimised detour net [3].

Towers and perimeter blocks are the main two types brought in the plan to generate variety throughout the project. At the cross-points of the main circulation line, towers are placed to generate a sense of direction. The perimeter blocks are increasing in height from the adjacent context to the center of the project createing a soft transition between the old fabric and the new high-density urban fabric. This generated an overall sense of linkage despite the different types introduced. The model creates an interconnected network hosting multiple morphological fabrics and open spaces throughout the city [3].

Figure 4: Kartal-Pendik Perspective [3].

The masterplan is a combination of international and local parameters that merge to produce a coherent, easily differentiated skyline. Implementing the project will be likely based on definite guidelines and regulations to produce the aimed for model. The guidelines will retain the unique character of the project. "Ordered complexity" is positioned an alternative to the monotonic order planned by older developments and the urban chaos formulated in urban spaces [3].

However, Hadid's architects aspire to go further in the project to the architectural scale benefiting the most from the deep relationality and the adaptations modes the design systems offer [3].

F. Parametric Urbanism as Identity Generator

Globalization is the new dominant feature in the technological age. This led to uniformity and loss of individuality that each place offers in an increasingly connected world. However, Etriken presented globalization under two opposite resultants: "convergence" where identicalness is molding the diversity present in different environment, and "divergence" where environments preserve their spatial identity. This was summarized in the words of Dovey: Globalization "doesn't simply iron out differences between cities it also stimulate them" [19].

In reference to this definition, we can classify modernism and parametricism. Modernism realized pure forms present the deepest architectural meanings (Colquhoun, 17). They sought after the simple repetition of geometrically shaped units placed on a right angled grid. Regardless of the context these forms were placed producing a globalized uniform order. The universal theme they proposed destroyed the sense of identity the place offers [20].

In a globalized era, the design themes are much more homogenized, where information and techniques are widely shared and different architectural practices and teaching tools are being internationally offered [13]. This leaves little space for national or local identity to nurture and maintain its uniqueness. However, parametricism relies on the individual identity that could be generated from regional particularity such as the climate and the local site conditions. Places where activities are carried out in the open-air are easily differentiated from others having all their activities in closed spaces. The

climate have a direct relation in shaping a community, where each climate produce different settings whether hot, humid ,cold or even unstable climate. Parametricism uniqueness lay in its ability to identify adaptive designs to different climate conditions, thus creating different urban characters and identities to different environments. Other parameters could present the unique identities of space, cultural parameters is a one. The business district parameters are different from the parameters of a university campus or the ones of an industrial site [20]. Thus urban design thus is not perceived as a fixed unit but rather a combination of different parameters creating a dynamic complex design scheme. The form is perceived as the result of multiple layers and interactions. It is thus individual, distinctive and totally dependent on site details and history [16].

G. Parametric urbanism as Urbanity Generator

The urban realization of parametricism isn't fully recognized [3]. This broad field of research need to be seriously considered. The significance of this approach is in the generated models provided by such programs rather than the theme of style it presents. This framework presents a deeper understanding that goes beyond the shape to the parameters that could introduce a multilayered model. This advanced systems have influenced the design approach, it proved its capability to establish better design components and buildings that are more efficient [21].

Nevertheless, it is confirmed that, in spite of the vast capabilities presented by this new field to improve the quality of the urban life, parametric urbanism just applies limited parameters such as the "mix of uses" and "urban density strategies". The "formal, environmental and functional parameters" inspected don't assure a successful urban model. The parameters applied are vital however; they are not the only ones to improve urbanity. Other parameters explored by other researchers such as Frederico de Holanda are also vital. The percentage of open spaces over total study area is an example [21].

Models offered by parametric design systems are simplified versions of the real components of the urban space. They present new ways to explore the world, assume and design before implementing on the actual ground in an irreversible state. They are the promising link between hypothesis and real life. Even though these programs have limited capacities, they are reliable tools for generating data, given that they assist in testing hypotheses and comparing information [21].

However, the challenge is creating quantitative parameters that can be investigated to attain the highest capability of an urban space. Thus, urbanity is not a merely qualitative field, it is also a model that has the capability to be considered quantitatively, thus improving the design components of the urban spaces and promoting a better environment to live in [21].

"*Cities are complex, adaptive systems with their own characteristic dynamics, and—if they are going to perform well from a human point of view—they need to be dealt with as such. In that light we must re-assess our current systems of planning, building and managing cities—the laws, codes, standards, models, incentives, and disincentives that effectively make up the "operating system" for urban growth. To make better cities, we need to shift to an evidence-based approach, able to draw on the best lessons of science and history about the making of good cities, from a human point of view*" [22].

III. CONCLUSION

Parametric urbanism underlies the dream of Le Corbusier, "*the fusion of art and technology that was the basis of modern movement*" (Colquhoun, 17); however, Le Corbusier's implication was bounded by the limits of his time. Parametric urbanism implies a deeper realization of the complex order of the city. It aims to model qualitative measures into quantified ones enhanced by advanced parametric design tools, though it is still a matter of research that isn't fully mature. It can be developed by designing parametric design tools that would allow urbanity to be investigated in a more practical and interactive manner (Canuto &Amorim, 8109:16). Technology offers new ways to rethink the complexity of the city. This raises the question; can the city be fully planned by parametric design tools away from the direct contact with people in the city? Can we create an urbanity life by merely observations and research? What is the city that we aim for, or in other words, what is then a perfect city?

References

[1] Merin, Gili. "AD Classics: Ville Radieuse / Le Corbusier" 11 Aug 2013. ArchDaily. Accessed 14 Dec 2014. http://www.archdaily.com/?p=411878
[2] Wendt, Matthias. "The Importance of Death and Life of Great American Cities (1961) by Jane Jacobs to the Profession of Urban Planning." (2009): 1-24.
[3] Schumacher, Patrik. "Parametricism: A new global style for architecture and urban design." Architectural Design 79.4 (2009): 14-23.
[4] Colquhoun, Alan. "The Significance of Le Corbusier." Le Corbusier 23 (1989).
[5] Pedro Rebelo & Franziska Schroeder (2006) Performing the Order: The messiness of play, Performance Research: A Journal of the Performing Arts, 11:1, 3-8,DOI:10.1080/13528160600807372.
[6] Le Corbusier The City of Tomorrow and its Planning, trans. Frederick Etchells, New York: Dover Publications (1987).
[7] Fishman, Robert. Urban Utopias in the Twentieth Century: Ebenezer Howard, Frank Lloyd Wright, and Le Corbusier. MIT Press, 1982.

[8] Corbusier, Le. La ville radieuse. Vincent, Freal & Cie, 1964.
[9] *Macedo, Danilo Matoso, and Sylvia Ficher. "Brasilia: Preservation of a Modernist City" (article). The Getty Conservation Institute, 2009. Accessed,Web. 17 Dec. 2014.*
[10] *Holston, J. 1989. The modernist city: an anthropological critique of Brasilia. Chicago : University of Chicago Press.*
Jacobs, Jane. The Death and Life of Great American Cities. New York, NY: Random House, 1961.
[11] *Hall, P. 2002. Cities of tomorrow : an intellectual history of urban planning and design in the twentieth century. Oxford; Malden, MA: Blackwell Publishing.*
[12] *Ward, Stephen. "Jane Jacobs Critic of the Modernist Approach to Urban Planning Who Believed That Cities Were Places for People." The Independent. Independent Digital News and Media, 3 June 2006. Web. 17 Dec. 2014. <http://www.independent.co.uk/news/obituaries/jane-jacobs-480896.html>.*
[13] Leacha, Neil. "Parametrics Explained." (2014). *Web. 11 April 2016.* <http://isites.harvard.edu/fs/docs/icb.topic1412058.files/Week%209/Parametrics%20Explained.pdf>.
[14] Oxman, Rivka. "Digital architecture as a challenge for design pedagogy: theory, knowledge, models and medium." Design Studies 29.2 (2008): 99-120.
[15] Tang, Ming, and Jonathon Anderson. "Information Urbanism: Parametric urbanism in junction with GIS data processing & fabrication." ARCC Conference Repository. 2014.
[16] Stavric, Milena, and Ognen Marina. "Parametric modeling for advanced architecture." International journal of applied mathematics and informatics5.1 (2011): 9-16.
[17] *Marcus, J. S. "Designer Cities: The Development of the Superstar Urban Plan."New York: The Wall Street Journal (2008).*
[18] da Silva, Robson Canuto, and Luiz Manuel do Eirado Amorim. "Parametric urbanism: emergence, limits and perspectives of a new trend in urban design based on parametric design systems." Trans. Anja Pratschke. In V (2010).
[19] *Carmona, M., Tiesdell, S., Heath, T. and Oc, T. "The Perceptual Dimension", Public Places, Urban Spaces: The Dimensions of Urban Design. Routledge, 2010.*
[20] *Schumacher, Patrik. "The Impact of Parametricism on Architecture and Society." Interview by Angel Tenorio. N.p., Mar. 2014. Web. Accessed Dec. 2014. <http://www.patrikschumacher.com/>.*
[21] *Canuto, Robson, and Luiz Amorim. "Establishing parameters for urbanity."Eighth international space syntax symposium. 2012.*
[22] *Mehaffy, Michael. "5 Key Themes Emerging From the 'New Science of Cities'", City Lab. 5 Sep. 2014. Accessed,Web16 Dec. 2014. <http://www.citylab.com/design/2014/09/5-key-themes-emerging-from-the-new-science-of-cities/380233/>.*

The Mad and The Other
A Visual Exploration of Graphomania and Hypergraphia

Alessandro Karim Decaneva
Department of Architecture and Design
American University of Beirut
Riad El-Solh / Beirut 1107 20 20 - Lebanon
karimdecaneva@gmail.com

ABSTRACT

One of the many ways of referring to a madman in Italian language is "folle" from the Latin verb "FOLLERE", meaning to go around like the wind.[1] But it is only since the late 1970s that a gradual and very slow process of liberation of the mad from segregation in mental hospitals took place. A few decades after the movement of Democratic Psychiatry started sensitizing society about what the reality of mental hospitals was and about their function to maintain social control over the population, we still keep on escaping and segregating all those who cannot conform to society and who cannot communicate their individuality through a clear and rational speech. The difficulties in implementing this reintegration come from the resistance of society that is still bound to stereotypes, prejudgments and a general avoidance of the mad.

My interest in the theme of madness started because of many different personal experiences ranging from direct encounters with mad people in the public areas of my hometown Trieste, to indirect ones such as reading poems by Alda Merini about madness and about the decade she spent in an asylum of Milano. The research will develop from two premises. The first is the consideration that generally the tendency of avoiding the mad is still predominant in society and the second is that the most used media to know about the world and about other humans' lives is the book.

My research wants to explore the possibility of challenging the privileged sensation of safety that a book offers to its beholder in the understanding of a subject such as madness. The aim is to let the narrative structure and the visual content of the book create a dialogue between the mad and "The Other", and ultimately let "The Other" question his role and his function in relation to the mad. The reader will thus be forced to navigate throughout a journey from sanity to insanity, or vice versa. Can the immersion of "The Other" into the fictional subjectivity of the mad change the way we consider this emblematic figure in society?

SECTIONS

---

[1] L. Castiglioni and S. Mariotti, *IL Vocabolario della Lingua Latina*, Firenze,

THE MAD AS THE OTHER/THE MAD AND HIS OTHER
Madness as Limit/ Identity as Limit

"We could write a history of limits - of those obscure gestures, necessarily forgotten as soon as they are accomplished, through which a culture rejects something which for it will be the Exterior"[2]

French philosopher Michael Foucault makes it evident in the Preface to his 1961 work The History of Madness that wondering about the role, the function and the historical metamorphoses of madness in Western culture consists in an investigation into the *Limits* of that same culture. Establishing each time this boundary is in fact a constant ongoing process proper to all cultures at all historical times. It is only through an understanding of these relations and the acknowledgement that this process is usually a *Monologue of Reason* that we could eventually come to challenge the general perception of madness and of mad people. After all, "it is true that it is society that defines, in terms of its own interests, what must be regarded as crime: it is not therefore natural."[3]

More than defining its own identity, within a social context, it is the mad's other that defines what is and is not included in the ensemble of generally accepted behaviors. This process of defining the Limits is ongoing since the Classical Age and it was systematically accomplished by passing everything under the 'Gaze of Reason'. As Foucault exemplifies in his History of Madness, this gaze of Reason created a monologue according to which the mad's madness is self-evident:

"The madman is the other in relation to the others, the other, in the sense of an exception, amongst others, in the sense of the universal. All forms of interiority are therefore banished: the madman is self-evidently mad, but his madness stands out against the backdrop of the outside world, and the relation that defines him, exposes him wholly, through objective comparisons, to the gaze of reason."[4]

In what follows I will take a broad historical overview of the way in which the category of 'the mad' and 'the other' were constructed. Rather than being exhaustive, the overview is meant to demonstrate the existence of this process by looking at select historical moments.

The Classical World

In pre-Homeric Greece and up to the 5th century B.C.E, we cannot find an equivalent term for 'madness' in its contemporary multitude of meanings because in archaic Greek societies madness was not thought of as a phenomenon comprehending different behavioral manifestations. The experience of mania (μανία) was of multiple

---

[2] M. Foucault, *History of Madness*, London, Routledge, 2006, Preface XXIX

[3] M. Foucault, *Discipline & Punish*, New York, Random House, 1995, p.104

[4] M. Foucault, *History of Madness*, London, Routledge, 2006, p.181

natures, each consisting in a passion affecting the behavior of an individual or even of a group of people. The term for soul, psyche (ψυχή) meant only the vital breath, life itself. In Homer's poems there was still no concept of soul as a separate entity from the physical body. Hence madness was often seen as happening because of divine will. This is exemplified for instance in Plato's Phaedrus where Socrates states: "As the ancients testify, is madness superior to a sane mind (sophrosyne) for the one is only of human, but the other of divine origin."[5] Madness could also be self-induced as for instance in religious group rituals made in order to reach ecstatic states of mind as for instance in Dionysian practices. We can trace back a first division and so a first establishment of the limits between the rational and the irrational only when a moral dimension was attributed to the soul. This happened historically with the birth of Western philosophy. As Foucault points out, in Ancient Greece there was no real opposite of the logos before the advent of Socrates and Plato. It is only with another of Plato's dialogues that a first opposition between the sophrosyne (σωφροσύνη) and the hubrys (ὕβρις) emerged. These were soundness of mind and haughtiness respectively. The attribution of an ethical dimension to the soul was later transposed to Ancient Rome and included in the *mos maiorum*, the ensemble of Roman socially accepted behaviors. This unwritten code traced new limits and evolved incorporating Christian values after 313 B.C.E.

Middle Ages and Renaissance
"Since none who lives from fault is free,
We see ourselves in every man who'll
Say he's wise and not a fool."[6]

In the Middle Ages the concept of madness started being used as a term comprehending different realities. Madness and the figure of the mad started being connoted by disparate and ambiguous meanings and those are still present in our way of approaching the subject today. In the reinterpretation of the phenomenon through Christian values, madness started being associated with the notions of Sin, the Fall and the End of Time. As Foucault explains: "Madness and the figure of the madman take on a new importance for the ambiguousness of their role: they are both threat and derision, the vertiginous unreason of the world and the shallow ridiculousness of men." [7]

If on one side the figure of the mad is seen as a threat according to social norms and religious beliefs, on the other hand reflections about madness started haunting the imagination and the works of artists and literary writers from the Middle Ages onwards. It is in this period that madness started being an explicit subject of epic poems and satyrs such as Ariosto's "L'Orlando Furioso" or the literary *topos* of *the ship of fools*. It is in the latter that we can best acknowledge the ambiguity and double nature madness acquired in this period. This is extremely well expressed in the 1494 book *Ship of Fools (Das Narrenshiff)* by Sebastian Brant. The satyr had a great success in the years immediately following Gutenberg's invention of the printing press and it consisted in a critical and ironic narration about the ships that were carrying mad people and outsiders back and forth the Rein basin without scope or ultimate destination. This was according to Foucault one of the first forms of confinement the mad has historically been sent through, as he states by saying: "His exclusion was his confinement, and if he had no prison other than the threshold itself he was still detained at this place of passage. In a highly symbolic position he is placed on the inside of the outside, or vice versa."[8]

In the Middle Ages we also acknowledge the existence of a socially recognized figure closely related to madness: the Fool. This had sometimes both the duty of impersonating the madman and also of narrating about cosmic madness. It is common believe that the Fool had a privileged relationship with the truth. It was socially accepted for him to always say what he thought, often by mirroring the reality of mankind with prophetic tones. In his analysis of the figure of the Fool in early modernity, Midelfort highlights the following: "Court fools, we are told, were always granted a license to speak the unvarnished truth to those in power, as if they were humble prophets, permitted to say things no courtier would dare to say."[9]

Modern Period
"The instituting of the asylum set up a cordon sanitaire delineating the 'normal' from the 'mad', which underlined the Otherhood of the insane and carved out a managerial milieu in which that alienness could be handled"[10]

The association between madness and Satanism tightened more and more in the period following the Renaissance, between the Reform and the Counterreformation. Religious persecution was targeting mostly those who could not unequivocally make a statement of faith, which in turn is an irrational act. Foucault states regarding this historical period: "After defusing its violence, the Renaissance had liberated the voice of Madness. The age of reason, in a strange takeover, was then to reduce it to silence."[11]

The remedies sought to cure many forms of madness in early modernity were still related to the Four Humors and other religious rituals. Foucault insists that the disappearance of leprosy from Europe historically gave both the idea of creating a separate space (the exclusion in the leprosarium) for the mentally ill and a social predisposition (the fear of other epidemics) for the beginning of the process of seclusion that emerged since the Seventeenth Century. Asylums were thus created and became the place of madness all over Europe ever since early modernity.

During the Enlightenment madness began to be seen apart from Christian views as is evident in the words of Voltaire: "We call madness that disease of the organs of the brain that necessarily prevents a man from thinking and acting like others."[12] The process of medicalization of madness that started in Classical Greece with Hippocrates continued under the guidance of Reason along the Eighteenth Century. Madness stopped being considered of divine origin and scholars began studying the

[5] Plato, *Phaedrus*, The Internet Classics Archive, Web. 10 Dec. 2015.
[6] S. Brant, *Das Narrenschiff*, trans. E. H. Zaydel, New York, 1944, p.58
[7] M. Foucault, *History of Madness*, London, Routledge, 2006, p.13
[8] Ibidem, p.11
[9] E. H. C. Midlefort, *A History of Madness*, Stanford, California, Stanford University Press, 1999, p.231
[10] R. Porter, *Madness: A Brief History*, New York, Oxford University Press, 2002, p.122
[11] M. Foucault, *History of Madness*, London, Routledge, 2006, p.44
[12] Voltaire, *Philosophical Dictionary*, trans. Theodore Besterman, p.210

phenomenon in a close relationship with the body. Curing the mad in the asylums involved a number of inhuman practices that have been practiced until the birth of modern psychological schools of thought and in some cases even after. In the Eighteenth Century madness started being documented in the form of autobiographical accounts of many patients' experiences in the asylums. Some pioneer psychologists such as Philippe Pinel put forward since the end of the century a critique of the brutalities of some practices such as physical restraint and torture common in many of these institutions all over Europe. Indeed, "One of the characteristics of the early modern state, observable all over Europe, was an increasing concern about and even care for the troubled, dispossessed, and wretched."[13]

The rise of Capitalism and the Industrial Revolution had a huge effect on the quality of life in Western societies. Among the urbanized lower and middle classes people were hardly coping with the increasing working rhythms of an economy based on productivity and exploitation. This produced a vertiginous rise in the number of people interned in the asylums. Porter states: "Throughout Europe, it was the nineteenth century which brought a skyrocketing in the number and scale of mental hospitals. In England, patient numbers climbed from perhaps 10 000 in 1800 to ten times that number in 1900. […] Such increases are not hard to explain. Positivistic, bureaucratic, utilitarian, and professional mentalities vested great faith in institutional solutions in general -indeed quite literally in bricks and mortar."[14]

The Twentieth Century and The Present

In many regards, the history of madness collapses and merges in the Twentieth Century with the history of psychiatry. The asylums became ever since their establishment a great observatory point for pioneers of this field of study. This fostered an ever-growing interest in the analysis of madness from a scientific and positivistic perspective. Many scholars operating in asylums and higher education institutions attempted to make psychiatry receive a full recognition as all other sciences.[15] Porter also reports in his historical analysis of madness that "by 1900 Pinelian optimism had thus run into the sands: 'we know a lot and can do little, commented one German asylum doctor." The scarce results medical doctors have obtained in their attempt to cure madness lead to drastic conclusions during the period between the two Wars consisting for instance in Nazi Germany to consider the mad not worth living and being taken care of.[16]

Sigmund Freud's (1939) discovery of the unconscious opened up the path for the establishment and development of various schools of psychoanalytic thought regarding the different approaches that can be used to alleviate the distress we all suffer from in heavier or lighter forms. Freud's methods were based on speech but as we know, not all those considered mad are able to converse within a logical framework. Language came

to be defined as an insurmountable barrier by many, and possible cures for the mad started to take advantage of medical and technological improvements involving interventions on the patient's body. This resulted in the application of experimental medical practices such as lobotomy, the electroshock and after the 1940s of the prescription of chemical drugs that are still largely consumed at the present day.[17]

Scientific and social debates regarding possible approaches and actual practices related to madness culminated in the 1960s and 70s with the birth of anti-psychiatric movements that had a central role in the transformation and/or shutting down of asylums and a progressive sensitization of society and states towards tolerance and understanding. In 1974, psychiatrists Richard Hunter and Ida MacAlpine wrote:

"Today, it is assumed that mental pathology derives from normal psychology and can be understood in terms of faulty inter or intrapersonal relationships and corrected by re-education or psychoanalysis of where the patient's emotional development went wrong. Despite all efforts which have gone into this approach and all the reams devoted to it, results have been meager not to say inconclusive, and contrast sharply with what medicine has given to psychiatry and which is added to year by year. Patients are victims of their brain rather than their mind. To reap the rewards of this medical approach, however, means a reorientation of psychiatry, from listening to looking."[18]

Although the conditions of the mad have generally improved in many parts of the world in our closest past, debates about the social role and the clinical treatment of the mad are still ongoing at the present day. Visual culture appropriated especially in the Twentieth Century many of the concepts we relate to madness but the way we look at the mad is still socially biased and emotionally detached. The Mad lives and The Other, impassively, looks at him, avoids him, and fears him.

GRAPHOMANIA AND HYPERGRAPHIA

"I like thinking to have shown this possibility of graphic communication for the 'Outcasts' to become a subject of interest and discourse. Not a discourse revolving around the researcher but rather around the outcast. A discourse of the mentally ill rather than one about him"[19]

In the previous chapters I outlined the historical metamorphosis of the mad and of his other, I then used concepts of Foucault's philosophy in order to give a historical dimension to the research. Based on these principles we can move on to a critical analysis of the ways people labeled as mad visually express their individuality through graphic and artistic tools. The main focus of this chapter will be on the figure of the hypergraphic. I have chosen the figure of the hypergraphic because it is affected by the kind of madness that is most relevant to the project that is

---

[13] E. H. C. Midlefort, *A History of Madness*, Stanford, California, Stanford University Press, 1999, p.322

[14] R. Porter, *Madness: A Brief History*, New York, Oxford University Press, 2002, p.112

[15] R. Porter, *Madness: A Brief History*, New York, Oxford University Press, 2002, p.183

[16] Ibidem, p.186

[17] Ibidem, p.205

[18] Hunter, Richard, and Ida MacAlpine, *Three Hundred Years of Psychiatry, 1535-1860 a History Presented in Selected English Texts*, Open Library, Web, Accessed 10 Dec. 2015.

[19] V. Andreoli, *Il Linguaggio Grafico della Follia*, Milano, RCS Libri, 2009, p.233 [own translation]

going to be a publication.

Defining Hypergraphia and Graphomania Between Madness and Otherness

"There seems to be a continuum stretching from the general population who do not much enjoy writing, through creative writers, to hypergraphics"[20]

Within a scientific discourse hypergraphia is not considered as a disease of its own but rather a *side effect* of other conditions such as temporal lobe epilepsy, schizophrenia and manic depression. It can also be drug induced and it consists in a conscious and irrepressible urge for writing and for filling void spaces. Besides cases where hypergraphia has an observable biological base, other individuals can present the same behaviors without showing any consistent change in the brain cortex or in the nervous system. Hence we notice how hypergraphia is a behavioral condition that is at the present day not always medically classifiable based on a biological perspective.

Fluctuating between madness and otherness, hypergraphia acquired in different times and contexts multiple social attributes. It is interesting to note how in a non-medical context the term hypergraphia is rarely used and the term graphomania substitutes it in many regards. It is not certain why there is no consistency in the use of two different terms to indicate the same condition besides the fact that the term hypergraphia is scientifically used whereas graphomania is of popular use and does not often appear in scholarly or academic sources. The term graphomania is mainly used when it refers to a social phenomenon rather than a medical condition. This is exemplified for instance in its use by writer Milan Kundera in his work *The Book of Laughter and Forgetting* (1978). The author points out how graphomania is often associated with the desire of receiving acknowledgment for an extensive and prolific writing. It is in many cases also associated with the desire of seeing one's name being published. On this regard Kundera says:

"General isolation breeds graphomania, and generalized graphomania in turn intensifies and worsens isolation. The invention of printing formerly enabled people to understand one another. In the era of universal graphomania, the writing of books has an opposite meaning: everyone is surrounded by his own words as by a wall of mirrors, which allows no voice to filter through from outside."[21] The same author speaks about graphomania as a social phenomenon manifesting in extremely high rates within states where "The absence of dramatic social changes in the nation's internal life"[22] and "An elevated level of general well being, which allows people to devote themselves to useless activities"[23] are present.

Kundera's novel expresses a preoccupation about the increasing presence of graphomaniacs in developed countries' societies not because he believes they are mentally ill but because of the state of things in that particular society. It appears that hypergraphia and graphomania are respectively side effects of mental illness and of our way of living. Even provided that we accept them as separate terms, the first pertaining to the mad and the latter to the mad's other, there is undeniable evidence that these mental conditions manifest themselves in many regards with the same modes. We should then turn to an analysis of these modes in order to understand hypergraphia both from a scientific perspective and from an analysis of content and aesthetics in hypergraphic production.

Hypergraphia in Different Mental Conditions

Epilepsy Alice Flaherty describes epilepsy as resulting from chronic seizures happening because of a change in the wavelengths produced in the temporal lobe, a part of the brain cortex where a crucial interaction between the limbic system and our rational consciousness takes place. Flaherty exemplifies this complex mechanism saying: "It is largely the limbic system that gives us the motivation to do cortical tasks, and it is the cortex that allows us to have cognitively complicated desires"[24]. This is an area that is particularly susceptible to electric nervous stimuli and it is where the limbic system interacts with language. The seizures can amplify creative impulses and this is the origin of hypergraphic activity. Temporal lobe epilepsy is not the only cause for hypergraphic behavior but it is the best understood one. Hypergraphia is an interictal symptom [25], meaning it manifests between one seizure and another.

There seems to be evidence about the connection between temporal lobe epilepsy and what is called the Geschwind Syndrome. This is usually diagnosed to people showing five traits: hypergraphia, a deepened emotional life (hyper philosophical or hyper religious), emotional volatility, altered sexuality and over inclusiveness of details in descriptions and representations. The presence of this last symptom is the discriminatory in the diagnosis of the syndrome. Flaherty notes it is a behavior that "Comes from a strong, conscious, internal drive -say pleasure- rather than from an external influence."[26]

Seizures can gradually alter our consciousness because they increase in intensity until the distinction between the real and the unreal fades away. In these cases the effects produced by the seizures are called *experiential*. Objects can look as if they are changing dimensions. For instance it is believed that Lewis Carrol's *Alice in Wonderland* (1865) was inspired by epilepsy's experiential effects on the author.[27] Other experiential effects consist in feelings of déjà-vu, of jamais-vu (resulting in freshness of perception), and in feeling the presence of a double, an alter ego or a muse[28]. Right after each seizure communication with an epileptic becomes really hard. "The personality of the epileptic is adhesive to excesses in a way that makes him unable to track the mobility of the life-environment, unable to adapt to the demands of social life."[29]

[20] A. W. Flaherty, *The Midnight Disease*, New York, Houghton Mifflin Company, 2004, p.45

[21] M. Kundera, *The Book of Laughter and Forgetting*, trans. Aaron Asher, New York, HarperCollins Publishers, 1996, p.127-128

[22] Ibidem, p.127

[23] Ibidem, p.127

[24] A. W. Flaherty, *The Midnight Disease*, New York, Houghton Mifflin Company, 2004, p.23

[25] Ibidem, p.23

[26] Ibidem, p.24

[27] Ibidem, p.27

[28] Ibidem, p.28

[29] V. Andreoli, *Il Linguaggio Grafico della Follia*, Milano, RCS Libri, 2009, p.87 [own translation]

Francoise Minkowska studied the life and the works of Vincent Van Gogh and she believed he was an epileptic. Minkowska expresses how people suffering from epilepsy have the tendency of feeling rather than thinking[30]. She exemplifies this view by comparing Van Gogh's work to Seurat's. The first represented the dynamism of a reality that imposes itself on the object, the latter representing reality as the product of a rational thought where the object is considered solid, immobile and rigid. Besides the example of Van Gogh, it is generally very rare that an epileptic spontaneously starts drawing or painting. Andreoli reports that only 4% of epileptics paint while epileptic writers are extremely common.[31]

Schizophrenia If the epileptic condition results in a multiplication of the ties between the perceiving subject and reality in an attempt to grasp (specially in hypergraphic production) a vertiginous and over inclusive amount of data, the schizophrenic way is in many regards antithetical. "The epileptic way is significantly different from the schizophrenic way. The schizophrenic breaks his ties with the environment, the epileptic instead multiplies them."[32]

Eugen Bleuler, who first came up with the term schizophrenia insisted on the centrality of autism in the analysis of schizophrenia. "This detachment from reality with the relative and absolute predominance of the inner life, we term autism."[33] Another key concept to understand schizophrenia is provided by the studies of M. Mc Reynolds who described mental processes leading to schizophrenia in a direct relation with anxiety. Mc Reynolds' explanation starts with the definition of precepts. These consist in all the possible kinds of data we collect from our senses and from our feelings and that are then to be assimilated by our conscience. In schizophrenia new precepts cannot be assimilated harmonically with previously assimilated ones. This accumulation produces an unbearable degree of anxiety that in turn results in reactions of autism, dissociation and apathy.

"The schizophrenic has lost those mental schemes proper of 'normal people' because those schemes were not useful in arranging his precepts. In this way the schizophrenic is forced to go back to more extensive, primitive mental schemes through a system of 'regression'. This is a quest for schemes that could be able to rearrange reality, schemes typical of primitive mental structures."[34]

A first attempt of studying these primitive mental schemes was conducted by C.G. Jung (1961). A mental scheme that according to Jung is common to all ancient civilizations for a first understanding of reality is *quaternity*, an element that is also found in many of the graphic and artistic works of schizophrenics. In *Man and His Symbols* (1964) Jung and his associates analyzed the visual manifestations of *quaternity*.

Usually occurring in a circular form within which squared or triangular elements are inserted, they refer to these as *mandalas* (from Hindu language standing for *magic circle*). The mandala stands according to Jung as a symbol for the Self as a whole. One of Jung's associates reports that "The contemplation of a mandala is meant to bring an inner peace, a feeling that life has again found its meaning and order"[35].

Jung proved that the mandala is a symbol recurrent in creative individuals, shamans (religious symbols), and in schizophrenics, therefore we can say that these individuals exist on a continuum and that the boundaries between these three categories is blurred. But Jung was a psychologist and the aim and focus of his research was showing the humanity of the patient's (the Other's) condition not the humanity of the patient (the Other) himself. It is to Levinas that we need to turn to find a philosophy that humanizes the face-to-face encounter with the Other.

ETHICS, OTHERNESS AND HERMENEUTICS

Ethics and Otherness
"Every one is really responsible to all men for all men and for everything."[36]

We just saw in the previous chapter that graphic expression is proper and sometimes an inherent feature of many forms of madness. Therefore the aim of this chapter is to highlight a truly inclusive ethical approach in interpreting different forms of artistic expression regardless of one's label as being the mad or as the other.

We can say that in Western philosophy an increasing interest in the analysis and questioning of the relativity of social roles and cultural identities arouse after the end of World War II. Lithuanian-French philosopher Emmanuel Levinas (1995) is one of the philosophers who placed at the center of his philosophical system a mutual obligation between the self and the other. This was achieved by considering his ethics as the very metaphysics of his philosophical system. In one of his most influential works, *Totality and Infinity* (1961), Levinas questions the role of ethics in Western thought stating: "Western philosophy has most often been an ontology: a reduction of the Other to the Same by the interposition of a middle and neutral term that ensures the comprehension of being."[37]

In Western culture, such tendency is what Levinas came to identify as the Imperialism of the Same, which replaces the individual uniqueness of everyone with an abstract homogeneity. This in turn can only be avoided, according to Levinas, by giving greater emphasis to the face-to-face encounter as the only possible basis for a truly comprehensive ethic. "Only in the face to face encounter, insists Levinas, do we experience human others in their true exteriority, their absolute alterity. All other ways of encountering others reduces them to forms of being that are subject to appropriation and domination

[30] F. Minkowska, *Van Gogh, Sa Vie, Sa Maladie Et Son Oeuvre*, Paris, Presses Du Temps Present, 1963, p.26

[31] V. Andreoli, *Il Linguaggio Grafico della Follia*, Milano, RCS Libri, 2009, p.88 [own translation]

[32] Ibidem, p.87

[33] E. Bleuler, "Autistic Thinking", *American Journal of Insanity*, no. 873, 1913, p.69

[34] V. Andreoli, *Il Linguaggio Grafico della Follia*, Milano, RCS Libri, 2009, p.82 [own translation]

[35] Jung, Carl G., et al., *Man and His Symbols*, "The Process of Individuation", New York, Dell, 1968, p.213

[36] F. Dostoyevsky, *The Brothers Karamazov*, London, J.M. Dent, 1967, p.431

[37] E. Levinas, *Totality and Infinity*, trans. A. Lingis, Pittsburgh, Duquesne UP, 1969, p.43

by ontological categories."[38]

The glance in the face-to-face encounter becomes fundamental in Levinas for overcoming all social biases and for considering "the Other as not assimilable to the leveled down categories of history, sociology, psychology and of previous ethical theories themselves"[39]. Although reducing the essence of ethics to the short duration of a glance might seem trivial, it is indeed the only moment where our considerations about the other are still at what we might call a primordial and inexplicit stage. Levinas speaks about the face "as the scene of the saying before the said and this saying that is at one with the face and enters into a spiraling movement (*un movement en vrille*) with the said."[40]

Most of the schools of thought and perspectives in modern psychology focus on speech as the main media through which it is possible to investigate about the pathology a mentally ill person is suffering from. Levinas' ethic escapes the discriminating factor of speech and the social biases speech carries along. The psychological perspective that emphasizes the less on speech is the biological perspective. In this regard, Roy Porter comments:

"*You don't crack mental illness by decoding what the mad say: for, they held, mental disease had a biological base. Powerful psychiatric currents have furthered such tendencies to silence the insane, especially in institutional environments. [Ö] In any case, did not the methods of the natural sciences prescribe observation and objectivity, not interaction and interpretation?*"[41]

Whether there is always a biological cause for madness to manifest itself or not, our knowledge of the nervous system and of our cognitive processes across different cultures is at the present day insufficient not only to fully understand madness, but also to claim the right of establishing who is the mad and who is the other. In doing so, we would keep subjecting ourselves to the same socio-political self-referential values we are accustomed to.

If we consider glancing at any mad person apart from his condition of being mentally ill, at the stage of the glance he would be nothing else than a generic other to us. In case we notice in this other a condition of psychological or physical distress, then the glance would instantly make us feel sympathetic for that person and we would then start considering the hypothesis of trying to help the other. The creation of this obligation between the self and the other can easily become subject of misinterpretations if we try to integrate it with pre-existing ideologies or other forms of ethics.

Reducing the basis of ethics from a greater ensemble of postulates coming from natural right and democratic principles to the mere moment of a glance in a face-to-face encounter might seem drastic and not inclusive of any substantial content. But as Casey states, "Kierkegaard, Heidegger, Husserl and James conjoin in regarding the present moment as nonpunctiform, as open and extended, as capable of conveying content of undelimited expansiveness."[42]

Since the relation between the mad and the other cannot always develop based on logical speech and since in many of the cases of emotional and psychological distress a total and sympathetic identification between the mad and his other is impossible, the only plausible ethical system that can apply to all apart from social biases seems to be the one indicated by Levinas. However, a question presents itself here: How is any of this relevant to art?

Art, Ethics and Hermeneutics

"Art is art and as such it cannot be pathological, whether produced by someone who is sick or by someone who is not sick at all. In the assessment of the work, this factor does not intervene at all."[43]

Levinas ethics can provide us with an interesting framework for understanding artifacts regardless of the author's label of being the mad or the mad's other. In two of his essays the philosopher relates his ethics to some aesthetic considerations. In *Reality and its Shadows* (1948) and *The Transcendence of Words* (1949) he directly addresses art. However, one can argue against this use of Levinas and insist that in Levinasian ethics art is placed at a secondary level of significance compared to the face-to-face encounter with the Other because of the fact that art can be subject to different kinds of interpretations. These interpretations are in most cases biased by socially constructed *prejudices*.

This is true to a certain extent but there are, I think, two convincing responses to this criticism. First, there is a convincing argument from the works of Levinas that the encounter with a work of art can be like a face-to-face encounter. In a recent reflection on Levinas' ethics and art Benda Hofmeyr, transposes the notion of the face-to-face encounter and its ethical implications to the encounter with the work of art. Basing her work on Levinas, Hofmeyr argues convincingly that art is an activity capable of "giving things a face" and by doing so it can carry along ethical implications:

"Like the face, I believe that art has the power to address us in a way that stops our indifference and inertia and enables us to act effectively. Moreover, it is the ethico-political responsibility of artistic and cultural producers to come up with strategies to engender others with such em-powering paralysis ñ a passivity that provokes a pre-conscious awareness of the power we have to act in a time in which Western civilization's indifference towards the needs of others is rapidly becoming a global phenomenon."[44]

The way the work of art holds this power on us comes from the mimetic qualities present in many forms of artistic expression. If for instance we think about a representation of an individuated moment in time, where representation

---

[38] E. Casey, "The Ethics of the Face to Face Encounter: Schroeder, Levinas, and the Glance", *The Pluralist*, vol.1, no. 1, 2006, p. 79

[39] Ibidem, p.80

[40] E. Levinas, *Otherwise Than Being*, trans. A. Lingis, The Hague, Nijhof, 1981, p.44

[41] R. Porter, *Madness: A Brief History*, New York, Oxford University Press, 2002, pp.157-158

[42] E. Casey, "The Ethics of the Face to Face Encounter: Schroeder, Levinas, and the Glance", *The Pluralist*, vol.1, no. 1, 2006, p. 77

[43] E. Minkowsky, *Preface a Van Gogh: sa vie, sa maladie et son oeuvre* de F. Minkowska, op. cit. p.15

[44] B. Hofmeyr, "Isn't Art an Activity that Gives Things a Face? Levinas on the Power of Art", *Image [&] Narrative* [e-journal], no.18, 2007, http://www.imageandnarrative.be

paradoxically highlights the absence of its referent (reality), we are captured by an experience similar to that of the face-to-face encounter where the shadow of being reveals being itself. At this regard Hofmeyr refers to Levinas' considerations stating: "Face-to-face with a work of art we are therefore not free but immobilized -held fast by something addressing us. It is not about comprehending what is revealed but about being captivated by that which is precisely incomprehensible."[45]

Second, in regards to the criticism that interpretations of works of art are merely prejudiced and prevent a true face-to-face encounter, we can refer to the works of German philosopher Hans-Georg Gadamer, who in his major work *Truth and Method* (1960), analyzed and indicated the role of prejudice as fundamental in hermeneutics. For Gadamer the notion of prejudice does not necessarily carry a negative connotation but it is instead the essential starting point of each interpretation. The more accurate is the perception of us and of our socio-political situatedness, the more our interpretation of any work of art can extend and keep into consideration different influencing factors.

A comprehensive interpretation is in Gadamer's hermeneutics only possible by acknowledging the specificity of our way of perceiving things, the specificity of our prejudice in building our perspective in relation to the subject, to an other, or to a work of art. Besides the role of prejudice in our assessment of the work of art, Gadamer seems to confirm the fact that "the work of art, is not an object that stands over against a subject for itself. Instead the work of art has its true being in the fact that it becomes an experience changing the person experiencing it."[46]

<div align="center">STUDIES ON HYPERGRAPHIC PRODUCTION</div>

"Graphic language is therefore a parameter of the mentally ill, it is almost a mirror which can seize his features. But it is necessary not to expect seeing too much in it."[47]

As Andreoli highlights in his work, an analysis of the visual manifestations of these conditions can provide us with an interesting insight of the inner world of hypergraphics. The framework of this general overview will move in parallel for both epilepsy and schizophrenia and will include considerations about content, aesthetics and dynamics of their graphic production.

In Temporal Lobe Epilepsy We learn from Francoise Minkowska's studies on epilepsy that although in this condition writing is more common than drawing and painting, representation is not at all involved in a precise reproduction of shapes. We have seen how the epileptic is someone who is very bound to reality and who is guided by feelings more than by rationality. It is not surprising that the content of graphic production in epilepsy does never involve abstract images.[48] Typical of the epileptic's depiction of images is according to Andreoli the absence of perspective, of horizontal movement (resulting in an ascending and descending dynamism) and of a

high degree of mannerism, ornamentation and complexity.[49]

On the other hand, writing in epileptic hypergraphia is accessible in terms of meaning although it presents complex and intricate narrative threads. Narration does not flow in a chaotic way; we know instead that the typical process of narration in hypergraphia consists in what psychiatry named *flight of ideas*[50]. This is the correspondent of what in literature is called *free association*, a narration scheme used for instance in the Ulysses by James Joyce and in the novels of Marcel Proust.
Alice Flaherty describes the aesthetics of this condition as follows: "Hypergraphic handwriting tends to have distinctive physical characteristics. Hypergraphics often use highly elaborate or stylized scripts, even mirror writing like that used by Leonardo da Vinci. For emphasis, they frequently write in all capitals or in colored inks. They may not confine themselves to the main text, but may add exuberant annotations, drawings in the margins, and illuminated initials."[51]

Darin T. Okuda and Steve S. Chang have conducted a research entitled *Hypergraphia in the Age of Computers* in 2004 where they describe the development of hypergraphia in a temporal lobe epilepsy patient who used (a part from hand writing) computer technology to enhance his writings. This involved the use of many different fonts, dingbats, color-coding, graphs, charts, and animated clips. The patient wrote incessantly about his condition to let the doctors know about his symptoms in great detail.[52]

We have seen how in temporal lobe epilepsy hypergraphia is identified as an interictal symptom between one seizure and another. This creates an explosive dynamic ranging from being barely able to communicate (or write), where the handwriting looks trembling, to moments of sumptuous and frantic production.

In Schizophrenia In terms of meaning, graphic production in schizophrenia is far less accessible than in epilepsy, presenting a high degree of symbolism and distorted metaphors. The tone is characterized by emotional distance as for instance in the case of mathematician Theodore Kaczynski's *Unabomber Manifesto*. For these reasons it is less common that a schizophrenics' works are published and in many cases schizophrenics are not really able to write due to the degenerative nature of schizophrenia.[53] In many instances the writings of hypergraphics look childish and nonsensical because, as Arieti mentions, "Those we might consider manifestations of irrationality are instead archaic forms of rationality."[54]

Andreoli points out how particularly in schizophrenia nosology goes hand in hand with its graphic translation if we conduct a static analysis of the condition (as opposed to a

[45] Ibidem

[46] H.G. Gadamer, Truth and Method, New York, Seabury Press, 1975, p.92

[47] V. Andreoli, *Il Linguaggio Grafico della Follia*, Milano, RCS Libri, 2009, p.235 [own translation]

[48] F. Minkowska, *Van Gogh, Sa Vie, Sa Maladie Et Son Oeuvre*, Paris, Presses Du Temps Present, 1963, p.26

[49] V. Andreoli, *Il Linguaggio Grafico della Follia*, Milano, RCS Libri, 2009, p.88 [own translation]

[50] A. W. Flaherty, *The Midnight Disease*, New York, Houghton Mifflin Company, 2004, p.38

[51] Ibidem, p.26

[52] D.T. Okuda, and S. S. Chung, "Hypergraphia in the Age of Computers", *Neurology 62.12*, 2004, p.2329-330. Available from ProQuest, (accessed 10 December 2015).

[53] A. W. Flaherty, *The Midnight Disease*, New York, Houghton Mifflin Company, 2004, p.41

[54] S. Arieti, *Interpretazione della Schizofrenia*, Roma, L'Asino d'Oro Edizioni, 2014, p.179 [own translation]

dynamic analysis). Typical of schizophrenia are according to Andreoli [55] horror vacui, stereotypy, graphic dissociation, symbolic language. "In mental illness symbolic language is predominant, a language that has more magical than semantic functions, corresponding more to a kind of magical thinking than to a rational one. It is more mystical than historical." [56] Many schizophrenic patients show a stylistic evolution, as Andreoli exemplifies in his studies by means of comparison of several artworks done at different times by the same patient.

CONCLUSION

After having analyzed the general characteristics of madness, some of its visual manifestations and having highlighted the importance of the mad's graphic language as a media able to reintroduce an ethical dimension into the relation with the mad, I have clear guidelines for proceeding to the executive part of the project.

Researching provided me with a theoretical framework for creating a publication in which the mental sanity of the author/protagonist is not certain. In all the cases I took into consideration during the research I never came across a case in which the subject fluctuating between madness and otherness is a graphic designer.

Thus I decided to take advantage of this opportunity to create a book about the activity of a hypergraphic/graphomaniac designer.

I hope the publication could let one wonder about the binary sanity/insanity and at the same time could highlight how much graphic design practice can reflect but also challenge the dominant ideologies in society, with an emphasis on the relationship between the mad and the other.

REFERENCES

[1] A.W. Flaherty, *The Midnight Disease*, New York, Houghton Mifflin Company, 2004.
[2] B. Hofmeyr, "Isn't Art an Activity that Gives Things a Face? Levinas on the Power of Art", *Image [&] Narrative* [e-journal], no.18, 2007, http://www.imageandnarrative.be
[3] E. Bleuler, "Autistic Thinking", *American Journal of Insanity*, 873, 1913.
[4] E. Casey, "The Ethics of the Face to Face Encounter: Schroeder, Levinas, and the Glance", *The Pluralist*, Vol.1, No. 1, (Accessed Spring 2006), pp.74-97
[5] E. H. C. Midlefort, *A History of Madness*, Stanford, California, Stanford University Press, 1999.
[6] E. Levinas, *Otherwise Than Being*, trans. A. Lingis, The Hague, Nijhof, 1981.
[7] E. Levinas, *Totality and Infinity*, trans. A. Lingis, Pittsburgh, Duquesne UP, 1969
[8] F. Dostoyevsky, *The Brothers Karamazov*, London, J.M. Dent, 1967
[9] F. Minkowska, *Van Gogh, Sa Vie, Sa Maladie Et Son Oeuvre*, Paris, Presses Du Temps Present, 1963, p.26
[10] Jung, Carl G., et al., *Man and His Symbols*, "The Process of Individuation", New York, Dell, 1968.
[11] K. A. Harmon, *You Are Here: Personal Geographies and Other Maps of the Imagination*, New York, Princeton Architectural, 2004.
[12] L. Castiglioni, and S. Mariotti, *IL Dizionario Della Lingua Latina*, Loescher, Firenze, 2006.
[13] L. J. Davis, *Obsession: A History*, Chicago: University of Chicago Press, 2008.
[14] M. Foucault, *Discipline & Punish*, New York, Random House, 1995.
[15] M. Foucault, *History of Madness*, London, Routledge, 2006.
[16] M. Kundera, *The Book of Laughter and Forgetting*, trans. Aaron Asher, New York, HarperCollins Publishers, 1996.
[17] Plato, *Phaedrus*, Indianapolis, Bobbs Merrill, 1952
[18] R. Calasso, *La follia che viene dalle Ninfe*, Milano, Adelphi, 2005.
[19] R. Porter, *Madness: A Brief History*, New York, Oxford University Press, 2002.
[20] S. Arieti, *Interpretazione Della Schizofrenia*, Milano, Feltrinelli, 1963.
[21] S. Brant, *The ship of fools*, trans. by Alexander Barclay, Edinburgh, Scotland, W. Paterson, 1966.
[22] V. Andreoli, *Il Linguaggio Grafico della Follia*, Milano, RCS Libri, 2009.
[23] Voltaire, *Philosophical Dictionary*, trans. Theodore Besterman, ed. Benda, vol. I, Oeuvre completes.

ARTWORKS AND ILLUSTRATIONS

Albrecht Durer, Woodcuts for Sebastian Brant's *The Ship of Fools*, 1494

Brueghel, The Feast Between Carnival and Lent, 1559

Photograph from a mental hospital's dayroom

Oleg, Security Deposits, Cities, 1999

Hypergraphic Spiral

---

[55] V. Andreoli, Il Linguaggio Grafico della Follia, Milano, RCS Libri, 2009, p.78 [own translation]

[56] *Ibidem*, p.234 [own translation]

The Social Hierarchy

Mazen Zein
Department of Architecture and Design
American University of Beirut
PO Box: 11-0236. Riad El Solh, Beirut 1107 2020. Lebanon.
miz01@mail.aub.edu

Nabaa, a complex mixture of social, economic and political characteristics that constitute a unique community located in the northern suburbs of Beirut. Nowadays it is considered as the home of different nationalities, incomes classes and sectarian groups. The name Nabaa signifies in Arabic a spring and it is was named as such because of the presence of a water source that allowed people migrating from the south to water their land. But we cannot mention Nabaa without referencing Bourj Hammoud, the district to which Nabaa is considered as the non-Armenian neighborhood. The population that is now considered as the current settlers of Nabaa is affected by several historical and political turnovers that led to the creation of Bourj Hammoud and the formation of Nabaa. This population is diverse and its distribution among the buildings gives a sense of an interesting yet almost imposed hierarchical order. The paper will follow the process of tackling such social organization with a general introduction of the street representing the variation in the population, classified in residential buildings and following a certain hierarchy. Then, an investigation was held of two residential buildings found expressive of such hierarchy. Afterwards, a study of the exterior treatment of the façades, signs of refugees, a study of the interior adaption of the space, depiction of the settlement, entrances of the apartments, were conducted to better explore the differentiation of this hierarchy. Moreover, an investigation on the relationship between the tenants and the current renters was executed, in terms of rules and regulations, services, prices and number of people staying in one apartment. Other aspects were also considered, such as political and religious ones, and if these notions played a role in the classification. All these studies were conducted through a double comparison, in order to understand better the depth of the hierarchy and how does it work between the locals and the outsiders.

I. HISTORICAL BACKGROUND

During the end of the nineteenth century, the Ottoman Empire was held responsible of several massacres against the Armenian people, leading to the flee of Armenian families to countries such as Lebanon and Syria. In the 1920's, Lebanon was under the French mandate, to deal with this overflow of refugees, the French authorities established a hospital that would accommodate the incoming refugees for several days and examine them for search of any diseases that they might have. This area was later called as Qarantina. To contain the flow of refugees, a temporary camp was set in the location, and with time, authorities considered of having temporary shelters.

But, since these shelters became vulnerable to climatic conditions, and the period of the migration was undefined, permanent settlements were later considered to house the Armenian population.

Later on, the newly formed Armenian quarter turned into an important industrial district that housed several utility workshops and industries. This spatial transformation was the result of the know-how abilities and low income labor that came with the Armenian refugees in their settlement period. During the 1940's, many Armenian families went back to their country, they ended up selling their houses to Palestinian refugees, fleeing the Israeli occupation, and Lebanese families coming from the south and Bekaa regions to Burj Hammoud searching for job opportunities. Hence, more settlements were needed to accommodate for the increasing population, leading to the growth of the Nabaa neighborhood.

In the period ranging from 1975 to 1990, the civil became a turnover for Burj Hammoud and Nabaa, in terms of population shift and delocalization. Families migrated from the south and Bekaa regions, mostly Muslim Shiite, and went to settle in the Southern suburbs of Beirut. During this period, factors such as the Israeli invasion, crossing of Syrian troops, creation of different political parties, became the main causes of the escape of these families. Moreover, Syrian workers started coming over to Nabaa, for better work opportunities and for its close location to the different industrial districts, settling inside the houses later abandoned by the Lebanese families.

After the war, a phase of reconstruction was undergoing in the country. Major Private Sectors were taking in charge of most of the reconstruction projects. This action bolstered more Syrian workers to cross over Lebanon and start working for these companies. As for settlement, Nabaa was their main target, because of the affordable accommodation, the ongoing shift in population of different nationalities and the presence of numerous market spaces. With the blast of Syrian war back in 2011, families joined their relatives in Lebanon and started crossing over for safety and protection.
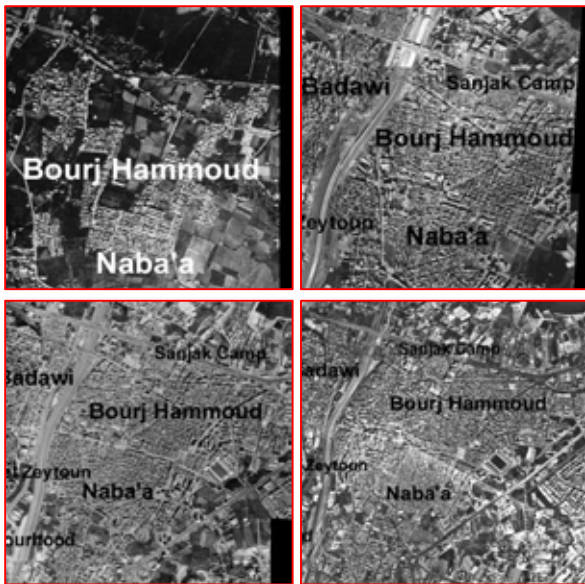
Figure 1. Maps showing the historical evolution of the Bourj Hammoud district with the Nabaa Neighborhood

II. SITE APPROACH

This brief historical background of Bourj Hammoud and Nabaa, along with the major political and economic phases, have helped in shaping Nabaa and its social structure. As a result, the people currently living there, can be classified into four main categories. One group composed of Lebanese mostly probable Muslim Shiite, another of Christian Armenians, a third group of Palestinians and the final group is made of Syrians. From the first site visit, while walking through the narrow streets and based on the initial observations of the neighborhood, one could sense that there might be a trace of a social classification in the buildings of Nabaa between the Lebanese and non-Lebanese residents.

To begin with, the street was located at a walking distance from the mosque of "Imam Ali Ben Moussa" and about eight blocks away from the main bridge (see Fig.2).
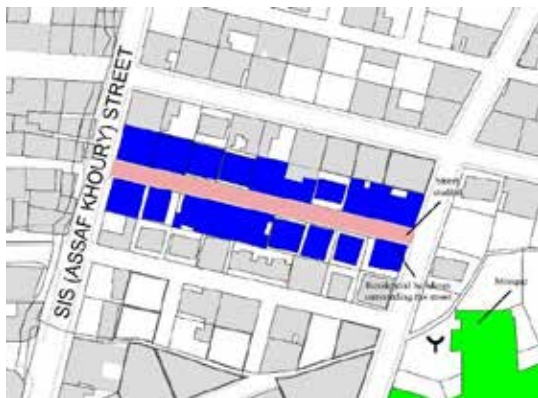
Figure 2. Maps showing street location in Nabaa quarter

At first, just by looking on the street level, we cannot but notice the kids playing around, of Syrian nationality based on their accent, in this narrow street. Their parents would be looking through the windows of their houses that were either located on the ground floor or first floor. The buildings of this street looked as if one was glued to the other and the balconies protruding from the buildings were so close that they were almost touching. From the balconies located on the upper floors, some locals started discussing between each other the purpose of the site visit and asking some questions. From this point on, the existence of a social pattern started to appear by locating the Lebanese residents on the upper floors and the Syrian refugees on the lower floors.

The lower floors of the buildings, because of their proximity, received little light, and were most of the time shaded. Even the entrances of the buildings with the stairs location, in the absence of artificial light, were barely visible. In terms of rain water, whenever there were showers along with the absence of an effective sewage system, water would reach high levels of the ground floor and cover large portions of the apartments. These conditions make it difficult for residents to live comfortably during all the seasons, and this is one reason for Lebanese people to stay on the upper floors and for the Syrian refugees to settle on the lower floors.

III. RESIDENTIAL CASE STUDIES

FIGURE 3. MAP SHOWING LOTS 1145 AND 1142 OF THE CASE STUDY BUILDINGS

Figure 4. Records of building lot 1142 showing existing apartments

After tackling the street and the distribution of residents in one building, two case studies were chosen for the subject, which were the two buildings located one next to each other that occupy lots 1142 and 1145 (see Fig.3). These buildings exist on one side of the street, almost the highest

buildings present, and they both contain Lebanese residents on the highest floors, and Syrian refugees, families and workers on the lower floors distributed among the rest of the apartments.

On lot 1142, the Em Fadi building exists, named as such because of the Em Fadi Market on the ground floor. Em Fadi was a resident of the building and an owner of the market, and now she runs the market with her son Fadi (see Fig.4).

The building is composed five floors with the stairs acting as the central core of the building. The ground floor is made of the market and two apartments that are currently housed by Syrian refugees. First there were Syrian workers living in these houses, and after the crisis in Syria, they brought their family so they could stay with them. The apartment is considered over populated compared to its area. Entering wasn't an option because of the presence of unveiled women in the house, but from the sight of the number of children, there were approximately seven people living in each apartment.

The first floor is composed of one single apartment, currently occupied by several Syrian workers, some of them have brought their families to live with them, and others are willing to. They are taking advantage of the space given, by increasing the members of the family living inside.

The second floor is made of two similar apartments, both occupied by Syrian refugees and workers. The number of residents is about six people, living in each apartment. When one worker came to Lebanon, he then brought his family with him, and later his other relatives followed to live with him in the same building.

The third floor is composed of two similar apartments that have small balconies looking towards the street, from the family of Zeinoun. One apartment is occupied by an old Lebanese lady living with her son, whom originally are from the south. The old lady was a bit resent, alone in the house, but she said that, they chose to settle on this floor for its spatial qualities and light conditions. And now since her husband lives in the southern suburbs of Beirut, her son is living with her temporarily.

On the last and fourth floor, there are two apartments. It looked as one on plan, but arrangements were made in order to make it two apartments to accommodate more people, in coordination with the owners of the building. In one apartment, a Lebanese couple is settled, from a long time. The man is called Hussein Issa, and he lives with his wife. They helped me in clarifying some elements about the strict rules applied by the owners over the Syrian tenants such as time limits, movement restrictions and contact with the other tenants. These directions, for them, help regulate the income flow of refugees in order to keep order and safety among the community, since they have been hearing a lot of crimes committed by Syrians coming to Lebanon in this period.

On lot 1145, the Em Hani building exists, named it after the Em Hani's house on the last and fourth floor. Em Hani is a Christian woman who married a Muslim, and settled in Nabaa. Couldn't get enough information because of their unavailability, but according to the Syrian residents in the building, they were the building representatives. Now the building is similar to Em Fadi's, with its central stair circulation, dividing each floor into two apartments, so in total, there are nine apartments, two on each floor, ranging from ground floor to the third floor, and the last floor is one apartment. The distribution of Syrian individuals in one apartment ranges from five to seven people. Whereas in the last floor, where the residents are Lebanese, the number is three.

According to these two buildings, the ratio of Lebanese residents to Syrians is about one to five. This unequal distribution in the buildings shows the continuous flow of Syrian refugees towards Nabaa. But what is more significant is that as we move towards the upper floors, the inhabitants of the existing apartments appear to be Lebanese and as we get closer to the ground floor, Syrians workers along with their families seem to reside there, showing this sense of dominance from the locals over the outsiders, from terms of choosing the best apartments to live in and the number of people living inside one space. The Lebanese chose the upper floors, maintaining a confrontable density of users leaving the Syrians on the lower floors, overcrowded and unable to live restfully. Hence, it was evident the sense of authority in the building exercised by the locals over the foreign tenants.

Moving on to the different exterior markers in the interior acting as signs of whether Syrian or Lebanese people were inhabiting the space.

IV.   ARCHITECTURAL ANALYSIS

On the architectural level, the windows on the lower floors were mostly closed and you rarely see anyone on the balconies, except for Lebanese citizens. In terms of how the apartment entrances were treated, on the lower floors, shoes and slippers rested outside Syrian apartments, which was a rare site on the entrances of Lebanese houses. Another marker relative to functionality, was the placement of wooden stops in front of the door steps in order to prevent the rats from entering the houses, which didn't exist on the upper floors.

But, the most important feature in the building was the central stair circulation that had a long landing acting as entrances towards the apartments (see Fig.5). The stair circulation acted as an additional feature towards the comfort for the apartments on the upper floors but not on the lower floors, since it tightened the space in front of the apartments. Moreover the circulation acted as a reason for locals to move and occupy the upper floors. First, in terms of daylight, the higher you are the better daylight you get. From the look of the ground floor,

how it was dark and gloomy, it was more reasonable to choose the upper floors. Daylight increased gradually while going up the stairs. And second, in terms of ventilation, the ground floor received little wind currents, whereas on the upper floors, one can feel the smooth breeze that ventilated the spaces.

Figure 5. Pictures showing the change in the level of daylight in the building
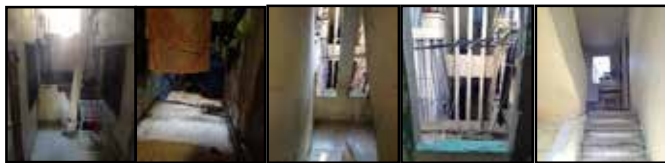

Figure 6. Pictures showing the different entrances to the apartments with the stair circulation

After discussing the exterior treatment of the building and the treatment of the spaces outside the entrances, there will be a shift towards the treatment inside the apartments, between families and workers, whether the space is considered as temporary shelter or a permanent house.

When entering the house of Hussein and his wife, the couple living on the upper floor of the building of Em Fadi, it felt as visiting a close relative. Hussein had been living in Nabaa with his wife for many years now, their apartment is about fifty meters squares. Their house was composed of a living room, a bedroom, a kitchen and a bathroom, not spacious, but enough for two people. They didn't seem to be bothered and they were managing themselves. They considered the apartment as their home, with their pictures decorating the walls and the tables, with coaches and carpets filling up the spaces. They like the neighborhood, but they started to feel a bit alone, after the flow of Syrian refugees. Some details in their houses such a vase of plastic flowers and a picture of a religious shrine hanged on the wall were enough proof that they consider this apartment as a permanent house.

Moving on to an apartment of a Syrian worker who brought his family recently. Mahmoud had spent almost a decade in Lebanon because of the work opportunities present. But when the crisis worsened in Syria, he asked for his family to join him. When visiting them, they were drinking tea on the floor, their house wasn't too much different of the Lebanese one. It was maintained and taken care of. Mahmoud bought most of his furniture from here. And according to him, he was picky a

bit in choosing the furniture, which was a sign of attachment to the apartment slowly becoming his home. What was striking, was that he even thought of buying pictures of flower vases, in order to add some decoration and finesse to the walls (see Fig.7). This delicate touch was a true indicator of considering the apartment as a house. And according to Mahmoud, is that as long as there is wok in Lebanon, he will stay in the country and think of making the apartment a permanent household.

Figure 7. Pictures showing the interior of a Syrian worker apartment

As for the Syrian workers, based on the short visit inside the apartment, they didn't seem to care about the place they were staying in. They had their matts packed on the corner and the rest of the furniture was chosen randomly. For them it was all about the work, where they spent most of their time. They didn't bother to think of how the inside would look like, with respect to their limited source of income. They thought of it as a temporary shelter to spend the nights in, a small requirement, easily replaceable, compared to their job duty.

Regarding the control and the relationship between the residents inside the building, through conducting different interviews between the residents and the tenants, and based on the little feedback received, there was this sense of inequality. The Syrian population was almost five times the Lebanese population in these buildings. They had lacked the freedom movement, under a direct pressure from the owners whom they were placed under their responsibility. This means that whenever the owners wanted to rent a Syrian migrant an apartment, they had to give all the necessary information to the municipality, with all the details as precautions and they would become the acting agents.

V. POPULATION DISTRIBUTION

Based on the historical events that were described at the beginning of the research, according to the testimonies provided by the current Lebanese residents and the thorough analysis on the social and architectural levels, the scenario goes as follow. Many families who migrated from the south and the Bekaa, and settled in Nabaa, left their homes during

the 1970's. As a result, many apartments were left empty. Some were located on the upper floors of the building, so the current Lebanese residents thought it would be a good idea to shift their houses upwards and become settlers of these floors. And, since there were Syrian workers coming in that period, they seized the opportunity of the available spaces in the buildings of Nabaa and settled in the houses that were evacuated in the lower parts of the building (see Fig.8).
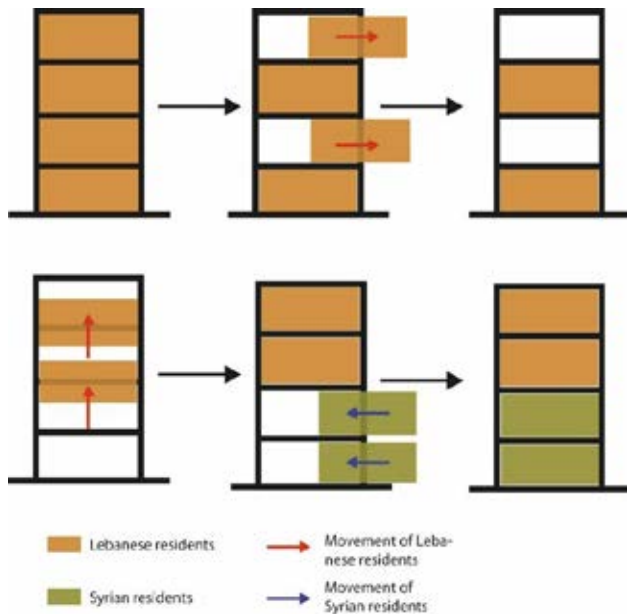

Figure 8. Diagram shows the movement of the Lebanese residents, when some of them leave their houses, then the rest move upwards and later the Syrian residents would occupy the lowest floors

VI. CONCLUSION

To sum up, Nabaa is a case of a mixed neighborhood, a testimony of a differentiated community living together peacefully. This structure is rarely to be found in other regions of Beirut, where people from different religious, political and economic backgrounds meet in one place. Some buildings act as an evidence to such way of living, representative of the social structure of the district. Nevertheless, even in such areas, especially in the case of residential buildings, one could find a certain social hierarchy constructed by the different nationalities present. A specific classification that relates to the social context of the residents, their income and nationality. Historical turnovers, major political events, signs and indicators never fail to create such social structure. With the presence of social hierarchy, even in communities such as Nabaa, the once mixed social gathering could turn into a classified, forcibly oriented social fabric, guided by rules and limitations replacing what was once a liberated quarter.

REFERENCES

[1] Building permits of lots 1142 and 1145, Municipality of Burj Hammoud, Retrieved October 2013

[2] Khayat, T. "Bourj-Hammud and the Armenian community in Lebanon: 70 years of housing policies", Escwa Beirut November 2002

[3] Khayat, T. "Population Shifts and Housing Problems in Nabaa", Escwa Beirut November 2002

[4] Maps of Bourj Hammoud and Nabaa, Department of Architecture and Design, Retrieved November 2013

DEPARTMENT OF CHEMICAL AND PETROLEUM ENGINEERING

CONTENT

Biomass and Sludge Gasification for Syngas Synthesis and CHP

Jad Halawi, Nour Sehnaoui, Ali Al Hady Hatoum, and Marwan Fakhr

Department of Chemical and Petroleum Engineering

American University of Beirut

Jch01@mail.aub.edu, nys04@mail.aub.edu, Anh15@mail.aub.edu, and mkf06@mail.aub.edu

Abstract-**The purpose of this project is to produce and recover the maximum amount of energy, from municipal solid waste and waste water treatment sludge, by syngas production via gasification and the combined heat and power cogeneration.**

I. INTRODUCTION

Lebanon is currently facing a serious energy problem in addition to a solid waste management crisis. The current demand for electricity in the Lebanon is around 2,300 MW [1]. Electricité Du Liban, EDL, is struggling to ensure the continuous delivery of power with its maximum production in the range of 1,500 MW [2]. The Lebanese energy sector suffers from crippling inefficiencies in its infrastructures including power plants, transmission and distribution networks[2]. In spite of the efforts, the electricity network does not represent a reliable source for the industrial and commercial sectors, where most beneficiaries have been compelled to install and use their own private generators to overcome the electricity shortages and ensure the continuity of their operations. To add to that Lebanon has recently been facing a dire solid waste crisis, due to poor governance of the sector and to the shutdown of the Nemeh landfill. Based on the state of the energy and waste management sectors in Lebanon, this project opts to solve a part of the problem. Gasification of biomass accompanied with a combined heat and power cogeneration cycle (CHP) would tackle the two problems at hand simultaneously. Gasification is a technique used for reducing the volume of the organic portion of solid waste in a cost efficient manner whilst producing synthesis gas or syngas to power gas turbines to fuel a CHP cycle.

This paper presents a complete design of gasification plant developed to be implemented in Lebanon.

II. GASIFICATION

Gasification is the process of partially combusting a carbonaceous fuel in the presence of a hypo-stoichiometric amount of oxygen at an elevated temperature. [3]Due to this partial combustion a gas stream that is rich in H_2 and CO is generated, i.e. syngas or synthesis gas. Thus this makes gasification an energy-efficient technique for reducing the volume of MSW and the energy recovery in the form of hydrogen fuel.

However, the gasification step itself is only the last step in a gasification system. In a typical gasification system the raw material i.e. Organic-MSW (O-MSW) and Waste Water Treatment-sludge (WWT-sludge) undergo the following steps: dehydration, pyrolysis, and gasification.

1) In Dehydration, whatever amount of moisture left in the feed or any other volatile liquid would evaporate due to the high temperature. Commonly this step occurs at a temperature range between 100ºC and 160ºC [3].

2) Pyrolysis or De-volatilization or thermal cracking is the thermal break down of bonds to produce char, tars, CO, CO_2, H_2O, NO_x, SO_x, H_2S, and NH_3 alongside a multitude of other products. Char is the carbon rich solid that remains after pyrolysis and is an integral part of the gasification process. The typical range of temperature that pyrolysis takes place at is between 550ºC and 700ºC. The operating temperature generally depends on the Ash melting temperature [3].

3) Char gasification occurs at high temperatures exceeding 950ºC and could reach up to 1200ºC [3]. Moreover, char gasification could be summarized in nine chemical reactions six of which are exothermic and the rest are endothermic reactions. This mix of endothermic and exothermic reactions is crucial for the gasification system, as we would dedicate a portion of the produced fuels and char to completely combust thus releasing heat to sustain the high temperature required for pyrolysis, i.e. designing an auto-thermal reaction system. This controlled combustion is achieved by varying the O_2 or air flow into the reactor, and maintaining oxygen at hypo-stoichiometric levels.

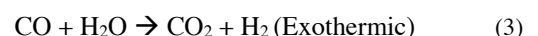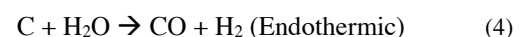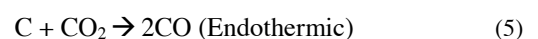The reactions of char gasification are:

- Combustion reaction:

$$C + O_2 \; \rightarrow \; CO_2 \;\; \text{(Exothermic)} \tag{1}$$
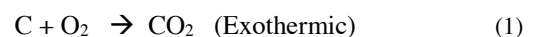
- Methanation reaction:

$$C + 2\,H_2 \rightarrow CH_4 \;\text{(Exothermic)} \tag{2}$$

- Shift conversion reaction:

$$CO + H_2O \rightarrow CO_2 + H_2 \;\text{(Exothermic)} \tag{3}$$

- Water-gas reaction:

$$C + H_2O \rightarrow CO + H_2 \;\text{(Endothermic)} \tag{4}$$

- Boudouard reaction:

$$C + CO_2 \rightarrow 2CO \;\text{(Endothermic)} \tag{5}$$

- Combustion reaction of CO:

$$CO + \frac{1}{2}O_2 \rightarrow CO_2 \text{ (Exothermic)} \quad (6)$$

- Combustion reaction of H_2 :

$$H_2 + \frac{1}{2}O_2 \rightarrow H_2O \text{ (Exothermic)} \quad (7)$$

- Combustion reaction of CH_4 :

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O \text{ (Exothermic)} \quad (8)$$

- Steam methane reforming reaction:

$$CH_4 + H_2O \rightarrow CO + 3H_2 \text{ (Endothermic)} \quad (9)$$
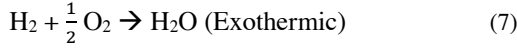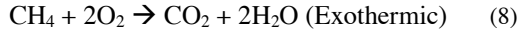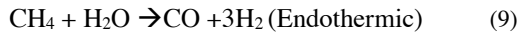
III. GASIFICATION SYSTEM SELECTION

There are multiple types of gasifier systems that are currently employed in the industry. In this project the type of gasifier that was deemed most suitable and efficient for the parameters of MSW in Lebanon, i.e. high moisture content [4], is the bubbling fluidized bed gasifier. In a bubbling fluidized bed gasifier the oxidant, i.e. air or oxygen, also plays the role of the fluidizing agent and is fed to the bed at the bottom of the gasification unit through a distributor plate. The bottom of the unit contains packing material usually silica sand particles and the waste is fed into the bottom of the bed of the unit to be mixed into the inert silica sand. The fluidizing agent should maintain a specific superficial velocity greater than the fluidization velocity of the bed to keep the bed in a fluidized state, which enhances reaction kinetics by increasing surface area of contact as well as heat and mass transfer parameters [5]. The optimal gasifying agents were found to be steam and air simultaneously. The optimal ratios of air to biomass and steam to biomass are 1:1 and 1.2:1, and the optimal temperature is 800ºC. This high temperature will not consume energy to achieve as the reactor will be operating under auto thermal conditions as a fraction of the fed reactants into the gasifier will undergo complete combustion to liberate heat [6]. The gasification process was simulated using the ASPEN PLUS® simulation environment, the kinetic reaction rates presented in Table 1 and the feed conditions presented in Table 2. With the aforementioned conditions the process was capable of generating 1224.87 kmol/hr of H_2 gas alongside an array of gases such as H_2S, SO_2, NO_2, CH_4, CO, and CO_2. The ASPEN PLUS® simulation of the process is presented in Figure1.

Table 1 Reaction Rate Kinetics

Reaction name	Reaction rate (s⁻¹)
Boudouard	$r_1 = \dfrac{k_{21}p_{CO_2}}{1 + k_{22}p_{CO_2} + k_{23}p_{CO}}$
Water-gas	$r_2 = \dfrac{k_{31}p_{H_2O}}{1 + k_{32}p_{H_2O} + k_{33}p_{H_2}}$
Methanation	$r_3 = \dfrac{k_{41}p_{H_2}^2}{1 + k_{42}p_{H_2}}$
Combustion (CO)	$r_4 = k_5 C_{CO}^{n1} C_{O_2}^{m1}$
Combustion (H_2)	$r_5 = k_6 C_{H_2}^{n2} C_{O_2}^{m2}$
Combustion (CH_4)	$r_6 = k_6 C_{CH_4}^{n3} C_{O_2}^{m3}$
Shift conversion	$r_7 = k_{81} C_{CO}^{n4} C_{H_2O}^{m4} - k_{82} C_{CO_2}^{n5} C_{H_2}^{m5}$
Steam methane reforming	$r_8 = k_9 (P_{CH_4} - \dfrac{P_{CO} P_{H_2}^3}{k_{eq2} P_{H_2O}})$

Table 2 Reactor Parameters

Parameter	
Temperature (ºC)	800 ºC
Pressure (kPa)	101.325 kPa
Mass flow (kg/hr)	19.96 kg/hr

IV. SYNGAS PURIFICATION

The syngas obtained from the gasification process should be scrubbed and purified. Acid gas removal is an integral part of our process to minimize hazardous emissions and prevent equipment corrosion and erosion. The major corrosive and acidic constituents of the syngas obtained are mainly hydrogen sulfide, carbon dioxide, sulfur dioxide, nitrogen dioxide, and ammonia.

The main reason behind the scrubbing of acidic gas is to minimize the environmental impact of the process. Moreover, the syngas is scrubbed to avoid undesired incidents such as corrosion of the pipelines and units downstream. In fact, CO_2 corrosion is also a major issue in the oil and gas industry because it is the main cause of equipment failure. The basic CO_2 corrosion reaction starts by the dissolution and hydration to form carbonic acid as seen in the reaction equations (10) and (11) [7].
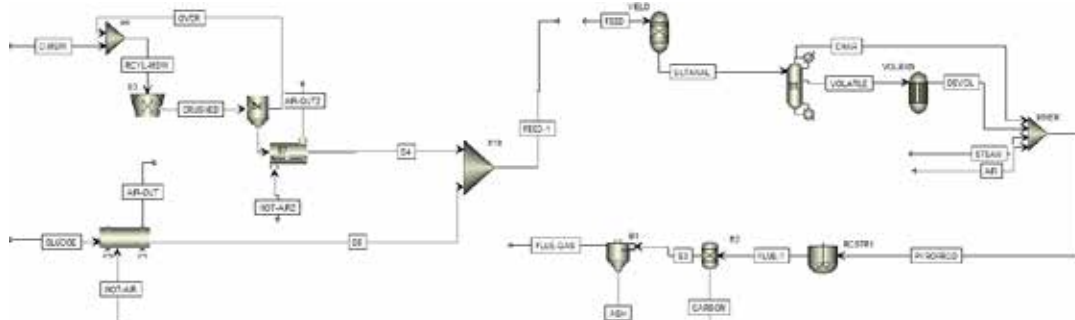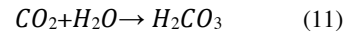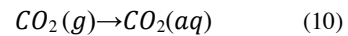
$$CO_2(g) \rightarrow CO_2(aq) \quad (10)$$

$$CO_2 + H_2O \rightarrow H_2CO_3 \quad (11)$$

Figure 1 ASPEN PLUS Simulation of Gasification

The CO_2 corrosion will lead to the production of $FeCO_3$ that will have different behaviors depending on its environment.

H_2S corrosion is also problematic. A proposed mechanism for H_2S corrosion will lead to two different paths and end with the formation of FeS. Furthermore, the corrosion rates for H_2S and CO_2 will vary as a function of the pH and the temperature of the system.

SO_2 and NO_2 corrosion are also important issues [8]. SO_2 plays a very important role in atmospheric corrosion in urban and industrial atmospheres; also it is mainly corrosive in the presence of moisture, as it will lead to the formation of H_2SO_4, a very hazardous component that can have severe effects on the health [9][10] . When H_2SO_4 comes in contact with carbon steel, it will cause the formation of ferrous sulfate [11].

$$H_2SO_4 + Fe \rightarrow FeSO_4 + H_2 \quad (12)$$

However the corrosion rate of SO_2 decreases when it comes in contact with stainless steel due to the anti-corrosive properties of stainless steel.

After surveying multiple techniques for the removal of acid gas, adsorption through an activated carbon column was found to be the most efficient in order to minimize energy consumption of the process.

The design of the adsorption column depends on many factors including the densities of the fluidizing agent and the adsorbent material, the void fractions, adsorption isotherms, and pressure drop through the bed. Moreover, the particle size distribution is used to specify the particles uniformity. For the rate of adsorption will increase as the particle size decreases.

The activated carbon adsorption process primarily relies on the diffusion and adsorption kinetics of chemical compounds through and on the activated carbon; and thus modeling this process mathematically should be done based on a mass balance that factors in the rates of mass diffusion, convection, and adsorption in the column. After performing a mass balance on the system the resultant equation is presented hereafter [9],

$$\frac{dC}{dt} = D_e \left(\frac{C}{C_0}\right)^{n-1} \frac{1}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial C}{\partial r}\right) \quad (13)$$

The above equation is a second order partial differential equation of Concentration C with respect to time (t) and radius (r) of the activated carbon bed [9].

Where,

$$De = \frac{\varepsilon_P D_P}{(1-\varepsilon_P)\rho_P nk} C_0^{1-n} \quad (14)$$

Where,

ε_P : Porosity of the adsorbent bed
ρ_P: Density of adsorbent (kg/m^3)
n: Fruendlich isotherm constant
k: Fruendlich isotherm constant
D_P: Diffusivity (m^2/s)

The Fruendlich isotherm is a mathematical relationship of the maximum concentration of a particular contaminant that could be adsorbed; it assumes that the adsorbent has a heterogeneous surface composed of adsorption sites with different adsorption potentials and based on empirical data [7].

$$\frac{x}{m} = KC^{\frac{1}{n}} \quad (15)$$

Where,

x: Amount of solute adsorbed (mg)
m: Mass of adsorbent (mg)

The next step in modeling the physical adsorption process of an activated carbon bed, is to solve the above partial differential equation (PDE) for each compound based on its given constants. To solve the aforementioned PDE the mathematical modeling software MATLAB was employed and the PDEPE function was used. The results of the MATLAB simulation are represented in Figure 2. Moreover the Freundlich isotherm constants employed are presented in the table below.

Table 3 Freundlich Isotherm Constants

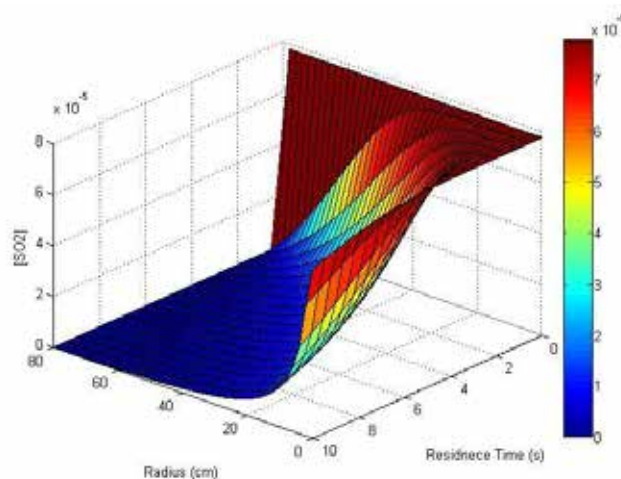|	H_2S	SO_2	NO_2	CH_4	CO_2	CO	H_2
D_P cm^2/s	33000	94	106	0.188	0.106	0.22	1.6
K	0.012	0.759	0.025	0.006	0.067	0.189	0.011
n	0.81	0.93	1.46	1.48	2.05	1.34	0.9

Figure 2 Change in Concentration of SO2 through the activated carbon bed

V. COMBINED HEAT AND POWER COGENERATION

Implementing a combined heat and power cogeneration cycle (CHP) in this process requires the presence of several elements. Those elements are a Gas turbine, a Heat Recovery System and a Steam turbine. After purification, the gas produced due to gasification would mainly contain hydrogen. Hydrogen would be sent to be combusted inside the gas turbine in order to generate electricity. The gases coming out of the gas turbine aren't combustible but are at a high temperature of around 1000 ºC. Thus to make use of the heat, a heat recovery system should be implemented. Hence the high temperature of the flue gases will be used to produce steam using a heat recovery steam generation unit (HRSG), and thus generating more electrical power via a steam turbine. The steam coming out of the turbine, will be condensed back to water, recycled and used again. Hence a closed steam cycle would be formed. [16]

Various types of HRSG are present. For this process it was
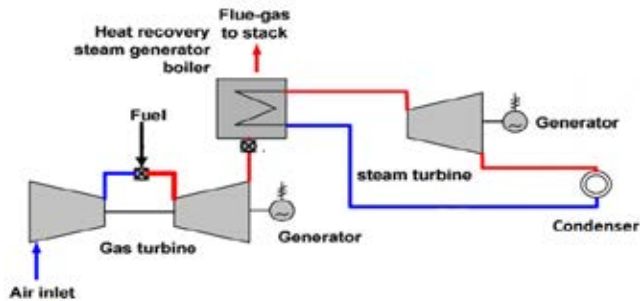
Figure 3 Schematic of an HRSG

determined that a once through steam generation unit (OTSG) would be optimal because it's easy to operate and maintain and has a lower cost than traditional HRSG units. OTSG is a unit with a bundle of tubes. Water passes through the tubes and gets transformed into steam after being exposed to flue gases that enter the OTSG unit from below and exit as stack into open air from above [17].

Using Aspen HYSYS®, the temperature, pressure and flow rates of entering and exiting fluids are determined. Hence a heat and mass balance on the OTSG gives the size of the proposed unit. The unit consists of 10 entry passes, where the inlet water is distributed equally. Steam exits from the lower part of the OTSG unit and is sent to a steam turbine as mentioned previously [17].

The complete results of the complete design process of the OSTG unit is presented in Table 4.

This project shows a promising solution to the two problems Lebanon has been facing for the last 40 years. The project aims at generating electricity using waste. Although one CHP power plant can't cover Lebanon's need, but constructing several CHP power plants in several locations would help close the gap between supply and demand of electricity.

Table 4 Specifications of OSTG unit

Specifications	Value
Number of tubes	1130
Tubes/row	10
Number rows	113
Pipe inner diameter	5.4 cm
Pipe thickness	0.5 cm
Pipe Length	8 m
Transverse pitch	3 cm
Longitudinal pitch	3 cm
OTSG height	12 m
OTSG length	10m
OTSG width	3 m

The net power generation by the implemented CHP cycle is presented in Table 5. Assuming that the plant operates for 24 hours a day, then the produced energy over a year (330 days) is 88,625 MWh/year. According to statistics conducted by the World Bank, the average consumption of electricity in Lebanon per capita is 3500 kWh per capita in one year [1]. Hence from this designed plant some 25,321 people would be provided with a 3500 kWh over a year. Assuming an average household contains four people, then process could provide electricity for up to 6300 households.

Table 5 Summary of Power Production

Unit	Produced (MW)	Consumed (MW)	
Pumps		0.5	
Compressors		14.87	
Gas turbine	20		
Steam Turbine	6.56		
Total	26.56	15.37	11.19 MW

VI. PROPOSED PLANT LOCATION

The plant location is a critical part of the process, there are many factors that have to be taken into account when choosing the preferable location. Although the plant should not be very far from Beirut for transportation of solid waste, it should be distant from residential areas as well as agricultural lands, urban areas and cities. Furthermore, the land should be in the coastal area in order to make sure that no major topographical limitations exist in that location such as mountains or sharp slopes and to be able use the sea water as a cooling fluid in the plant. These specifications were translated to the following four constraints:

1. Three kilometers away from the sea
2. Two hundred meters from agricultural land
3. Two hundred meters away from main roads(primary and secondary)
4. Plant area is greater than or equal to ten thousand square meters

In order to meet the constraint set, a program, geographic information system (GIS), was used. The constraints were

inputted into the software which allowed it to optimize the location and give the best fit locations.

The approximate location of one of the areas that was found by GIS to meet the applied constraints is in the village of Damour, which lies south of Beirut. Hence, and according to the calculations mentioned in the previous pertaining to the amount of power generated, the neighboring towns such as Saadiyat, Damour and Mechref could be provided with electrical power from the plant.

VII.   ECONOMIC FEASIBILITY

A rough economic analysis was performed on the entering and leaving materials of the process and yields the following. For every 1 ton of biomass processed and gasified this would provide 100 $ which is the price paid by municipalities for 1 ton after checking what other companies in Lebanon are charging for it [15]. 27.944 (wet) tons/hr of biomass and sludge are taken by the process. Hence a profit of $22,131,648 per year is obtained.

In addition a profit is gained from electricity which is being sold to neighboring areas. 1kWh of electricity costs 5 cents [15]. The process provides a net of 84,189.6 MWh per year. Hence a profit of around 4,000,000 $ is made from the electricity each year. Thus, this amounts to a gross revenue of $ 26M per year.

VIII.   CONCLUSION

The gasification of municipal solid wastes and waste water treatment sludge is a green process that not only produces syngas and power however, also aids in minimizing solid waste volume which is a troubling aspect of waste management. Furthermore, at the beginning of this project we set forth to recover as much power as possible from otherwise useless waste.

Our plant turned out to be producing around 11 MW of net electric power, which could power around 6000 households in Lebanon, and is comparable to a small scale electric power plant.

IX.   REFERENCES

[1] World Bank, *Energy efficiency study in Lebanon* (2009)
[2] V. Ganapathy, ABCO Industries Heat Recovery steam generation: understand the basics (1996)
[3] Nikoo, M., & Mahinpey, N. (2008). Simulation of biomass gasification in fluidized bed reactor using ASPEN PLUS. Biomass and Bioenergy, 1245-1254.
[4] Merhebi, F., Sabbagh, B., Moussallem, M., Rizkallah, M., Khalil, S., Seoud, J., . . . Khayat, Z. (2014). *Lebanon Environmental Assessment of the Syrian Conflict & Priority Interventions*. Beirut: Lebanese Ministry of Environment.
[5] Mumbai CII Naoroji Godrej Centre of Excellence. (2013, February). *A Practical Course on Industrial Mixing Technology & Equipment*. Retrieved from Mixing Experts: http://www.slideshare.net/mixing-expert2014/mixing-of-liquids-solids-and-high-viscosity-materials?qid=fa129bc1-782f-4d43-94ce-feb4f569d4f8&v=default&b=&from_search=1
[6] Hamad, M., Radwan, A., Heggo, D., & Moustafa, T. (2012). Hydrogen rich gas production from catalytic gasification of biomass. *Renewable Energy,* 1290-1300. Retrieved December 6, 2015.

[7] Koteeswaran, M. (2010, June 1). CO2 & H2S Corrosion in Oil Pipelines. Retrieved November 30, 2015, from https://www.scribd.com/doc/262135811/CO2-H2S-Corrosion-in-Oil-Pipelines
[8] Sulfur Dioxide - Acid Rain. (n.d). Retrieved December 1, 2015, from http://corrosion-doctors.org/AtmCorros/sulfur-diox.htm
 [9] Material Safety Data Sheet. (1999). Retrieved December 1, 2015, from http://avogadro.chem.iastate.edu/MSDS/H2SO4.htm
 [10] BRITISH STAINLESS STEEL ASSOCIATION Making the Most of Stainless Steel. (2015). Retrieved December 1, 2015, from http://www.bssa.org.uk/topics.php?article=37
[11] Panossian, Z., Lira de Almeida, N., Ferreira de Soussa, R., De Souza Pimenta, G., & Schmidt Marques, L. (2012, May 1). Corrosion of carbon steel pipes and tanks by concentrated sulfuric acid: A review. Retrieved December 1, 2015, from http://www.sciencedirect.com/science/article/pii/S0010938X12000595
 [12] Kalusche, H. (2015). Hydrogen sulphide safety and health issues. Retrieved November 30, 2015, from http://el.erdc.usace.army.mil/workshops/04jun-wots/kaluschue.pdf
[13] Hendriks, C. (1994). Carbon Dioxide Removal from Coal-Fired Power Plants. Retrieved December 1, 2015, from https://books.google.com.lb/books?id=8xJJCAAAQBAJ&pg=PT27&lpg=PT27&dq=NO2 and SO2 removal by amine solvent&source=bl&ots=ezK0RF39au&sig=K_9H0SB7UiLdr76rdEie5oVUpnQ&hl=en&sa=X&redir_esc=y#v=onepage&q=SO2&f=false
 [14] Matt Nash,waste [mis] management (September 2015) http://www.executive-magazine.com/special-report/wastemismanagement

[15] Liban, E. D. (n.d.). *Lebanon Electricity Tarrifs*. Beirut: Dynamic Eneregy & Water Solutions.
[16] V. Ganapathy, (2003) Technical Documentation: Integrated Gasification Combined Cycle Systems (IGCC) with Carbon Capture and Storage (CCS).

[17] Emad Hamid , Mike Newby,   Pericles Pilidis, (2011) THE PERFORMANCE MODELLING OF A SINGLE AND DUAL PRESSURE UNFIRED ONCE THROUGH STEAM GENERATOR.

Design of a Biomass to Gasoline Conversion Plant in Lebanon

Hanadi Abi Ghanem, Yara Beaini, Mohammad Itani, Tarek Lahoud, Malak Wehbe
Department of Chemical and Petroleum Engineering
American University of Beirut
Beirut, Lebanon
hka21@mail.aub.edu, yhb03@mail.aub.edu, mmi20@mail.aub.edu, tfl00@mail.aub.edu, mkw03@mail.aub.edu

Abstract - **With the recent increased interest in waste treatment in Lebanon, the process of production of gasoline from biomass was studied for the country. Simulations were used to assess the feasibility and operation of this process. Design wise, the process proved to be possible since methods were found to perform all the different stages of the process. The inlet feed was determined based on the biomass production in Lebanon, with a value of 150 tons/hour. Based on this biomass feed, the process yielded a gasoline production of 4710 barrels/day. The stages involved in this process are: the conversion of biomass to syngas through gasification, syngas cleaning and conditioning, conversion of syngas to methanol (MeOH), and finally the use of the MTG process to convert MeOH to gasoline using DME as an intermediate. The byproducts generated include sour gas, steam, waste water and ash. The design of the units employed in the process was also carried out and this paper includes the dimensions of a PBR and a distillation column. As for the plant location, the area of Sibline in South Lebanon was shown to be a suitable site for the construction of the plant. Safety and environmental considerations of gasoline were screened, displaying the complexities of using this fuel. At this phase of the project, the study is focused on the design considerations of the plant; only a preliminary economic analysis was performed. It will be further studied later along with energy consumption, after including optimization measures and heat integration.**

I. INTRODUCTION

Lebanon currently produces around two million tons of waste per year, half of which is organic. This quantity is expected to reach 2.22 million tons by 2020 [1]. Lebanon's waste management program relies on the collection and dumping of waste in landfills. Ideally, landfills should be confined spaces with a sealed liner to prevent seepage and leaching of contaminants to the ground. However, due to the lack of regulation in Lebanon, the integrity of the current landfills, and the new ones to be built, is questionable [2]. A very minor share of the country's waste is recycled or composted while a significant one ends up in uncontrolled dumpsites all over the country, most of which are located in close proximity of rivers, springs and streams and practice open burning [3].

Better waste management policies can be very beneficial for Lebanon. Recycling, construction of sanitary landfills and energy generation from waste are all viable options. The production of gasoline from biomass via gasification is one way of generating energy from waste. In addition to treating waste, this process could provide Lebanon with gasoline. The country is a heavy consumer of this fuel, all of which it imports, with a consumption of 140,000 MT in 2008 that has been increasing ever since [4] [5].

II. LITERATURE REVIEW

The process of going from biomass to gasoline consists of four major steps: (a) biomass conversion to syngas, (b) cleaning and conditioning of the syngas, (c) syngas conversion to MeOH and, (d) MeOH conversion to gasoline. The intermediate formed compound, syngas, is a mixture of carbon monoxide, carbon dioxide and hydrogen with a desired ratio of H_2/CO of more than 2 and a desired ratio of CO_2/CO of 0.6 [6].

A. Biomass to Syngas

Biomass can be converted to syngas either through biological or thermal pathways [7]. Biological processes, like anaerobic digestion or biophotolysis, involve the action of biological organisms. These processes are extensively time consuming and do not produce syngas at once, they produce biogas which then has to turned into syngas, which increases capital cost [8].

As for thermal processes, the main process that produces syngas is gasification, which takes place at very high temperatures reaching 1,000°C. Gasification uses a gasifying agent which can be oxygen, air or steam, and is chosen based on the required composition of the final product. It consists of four steps: drying, pyrolysis, gasification and combustion, occurring in series and breaking down biomass to eventually yield syngas [9].

B. Cleaning and Conditioning

The syngas produced contains various contaminants like particulate matter, tar, sulfur, alkali metals and hydrogen chloride. Before being further processed into MeOH, the contaminants' levels of syngas need to be as shown in Table 1 [9].

TABLE 1: CONTAMINANT CONCENTRATIONS IN SYNGAS

Contaminants	Concentration (mg m^{-3})
Particulate Matter (Char, Ash)	< 0.02
Tar	< 0.1
Sulfur (H_2S, COS)	< 1
Nitrogen (NH_3, HCN)	< 0.1
Alkali Halides (Primarily HCl)	< 0.1

The methods of cleaning of syngas differ for cold and hot gas streams and are chosen according to the syngas production method [9]. Syngas should also be conditioned to the conditions of the downstream processes using heaters and compressors.
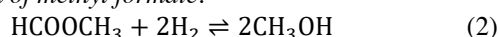
C. Syngas to Methanol

The production of MeOH from syngas can be done either via the formation of methyl formate or using steam. The first method involves the following reactions:

Carbonylation of MeOH:
$$CO + CH_3OH \rightleftharpoons HCOOCH_3 \qquad (1)$$

Hydrogenation of methyl formate:
$$HCOOCH_3 + 2H_2 \rightleftharpoons 2CH_3OH \qquad (2)$$

These reactions are performed on an alkoxide catalyst (RONa), which is poisoned by CO_2 and H_2O, abundantly present in the feed. Removing these compounds involves many technical and economic difficulties [10].

The formation of MeOH from syngas using steam, on the other hand, involves two different reactions:

MeOH steam reforming:
$$CO_2 + 3H_2 \rightleftharpoons CH_3OH + H_2O \qquad (3)$$
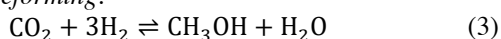
Water-gas shift:
$$CO + H_2O \rightleftharpoons CO_2 + H_2 \qquad (4)$$
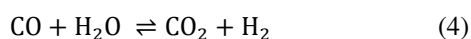
This process yields more MeOH than the first one. However, it needs high operating temperature and pressure and produces incompressible gases, which have to be purged downstream [11][12].

D. Methanol to Gasoline

This step can be mainly performed through either the MTG process or STG+ process. The MTG (MeOH to Gasoline) process consists of two stages. The first stage is converting MeOH to dimethyl ether (DME) and the second one is converting DME to gasoline using a catalyst [14]. This process has a high yield of fuel, reaching 80%, can reach complete conversion, is energy efficient and produces gasoline of a good quality with a low concentration of contaminants. Its drawbacks, on the other hand, are induced by its use of a catalyst: catalyst aging and coking are inevitable and have to be dealt with.

The STG+ process is also very close to the MTG one but involves more treatment of the produced gasoline to enhance its quality. This process also has a high yield of fuel and is able to convert syngas to gasoline without producing intermediate liquids. However, it is still in its early development stages with many aspects still unclear [14].

III. STAGES OF PRODUCTION

The simulation of the four stages involved in gasoline production from biomass used the Peng Robinson fluid package. The feed was considered as the organic waste of the whole country, which is approximately 150 tons/hour. The block flow diagram of the plant is shown in the appendix.

A. Biomass to Syngas

The thermal process of gasification was chosen over the biological ones. Biological processes are slow and may take up to 20 days to complete the reaction; hence, going with gasification would save us time and expenses. The gasification

TABLE 2 - BIOMASS COMPOSITION

Element	C	O	H	N	S	ASH
Percentage	48.038	38.393	5.899	1.124	0.178	6.368

plant includes the pretreatment of biomass using magnetic separators and shredders to reduce the particle size to be in the range of 0.25 mm and 0.55 mm to be used in the plant. The biomass feed used in this process was made up of the elemental composition shown in Table 2 [11].

Gasification takes place in a single unit in practice. This unit is divided into four main chambers, as shown in Fig. 1, each with its own operating temperatures and reactions. The pressure throughout the unit is maintained at about 300 kPa. For simulation purposes, each of these chambers was modelled separately using Aspen PLUS.

Drying is necessary to reduce moisture to acceptable levels of about 10%-20%. This is done by adding heat from downstream streams to heat the feed to around 100°C which removes water irreversibly. The process continues until a temperature of 200°C is reached. Biomass then undergoes slow pyrolysis. The residence time in this part of the reactor is between 5 to 30 minutes, with a low heating rate and at a temperature of 600°C. Pyrolysis breaks down biomass into non-condensable gases (such as CO, CO_2, H_2 and CH_4), bio-oil and char [9]. These products enter the gasification section where char is the major component that reacts.

Gasification requires a gasifying agent to allow the reaction to take place in the form of a fluidized bed. Steam was chosen as the gasifying agent for the purposes of this simulation. The operating temperature is in the range of 800°C and 1000°C [9]. The reactions taking place in the gasification chamber are shown below in (5) to (14) [12].
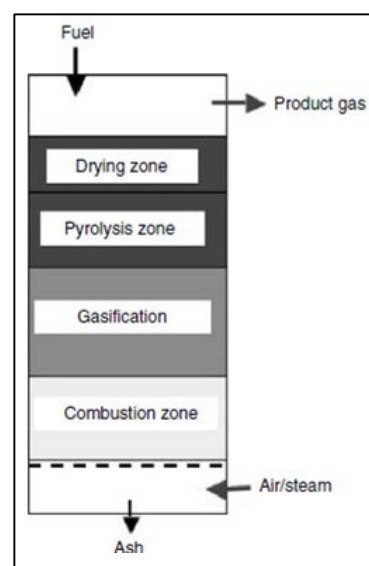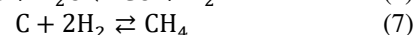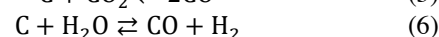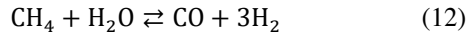
$$C + CO_2 \rightleftharpoons 2CO \qquad (5)$$
$$C + H_2O \rightleftharpoons CO + H_2 \qquad (6)$$
$$C + 2H_2 \rightleftharpoons CH_4 \qquad (7)$$

Figure 1 - Combustion Unit

$$2CO + O_2 \rightleftarrows 2CO_2 \qquad (8)$$
$$2H_2 + O_2 \rightleftarrows 2H_2O \qquad (9)$$
$$CH_4 + 2O_2 \rightleftarrows CO + 2H_2O \qquad (10)$$
$$CO + H_2O \rightleftarrows H_2 + CO \qquad (11)$$
$$CH_4 + H_2O \rightleftarrows CO + 3H_2 \qquad (12)$$
$$H_2 + S \rightleftarrows H_2S \qquad (13)$$
$$0.5\,N_2 + 1.5\,H_2 \rightleftarrows NH_3 \qquad (14)$$
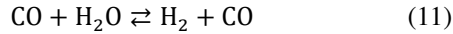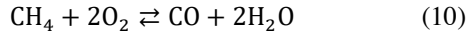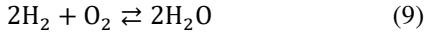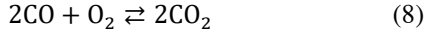
Finally, the combustion stage is necessary to provide some of the energy required by the previous stages. Since combustion is the only exothermic reaction of all four stages, the hot gases may be used in drying, pyrolysis and gasification reactions.

B. Syngas Conditioning and Cleaning

Raw syngas from gasification is produced at 900ºC and 303.4 kPa. This stream is imported into Aspen HYSYS with the same composition and conditions. However, before downstream processing of this gas, it should be conditioned and cleaned from contaminants such as hydrogen sulfide (H_2S), and NH_3 through scrubbing. The feed to the scrubbing section has to be at 60ºC and 7300 kPa [13]. The change from the initial conditions is achieved by using a series of heat exchangers for cooling as well as compressors for raising the pressure to the required value. Flash drums were used where necessary to remove any liquid water formed from cooling and compressing the stream.

The scrubbing of H_2S is carried out in an absorption column, in which chemical absorption takes place using methyl diethanolamine (MDEA). The streams leaving the absorption column are: sweet syngas (containing only 7.26 ppm H_2S) and rich amine. The rich amine stream is fed into a regenerator to recover MDEA and recycle it back to the system. It is important to regenerate the amine and keep it in a closed cycle as it is toxic, hazardous, as well as an expensive specialty chemical. The sour gas leaving the top of the regenerator, is mainly composed of CO_2 and H_2S and is a by-product of the process.

For downstream processing, the H_2:CO and CO_2:CO ratios in the scrubbed syngas must be adjusted. These ratios can be explained by studying the reactions involved in MeOH synthesis shown later. To attain these ratios the clean syngas undergoes a water-gas shift reaction shown in (4). This is done by passing the syngas stream over a bed of iron oxide-chromium oxide catalyst in a packed bed reactor (PBR). The syngas stream is introduced into the PBR at at 317.2ºC and 9500 kPa. A ratio of 5:1 of steam to syngas was found to yield a ratio of 2.877 for H_2:CO and 0.6002 for CO_2:CO, which satisfies the requirements for MeOH synthesis.

C. Syngas to Methanol

This step is also modelled on Aspen HYSYS. The reactions involved are shown in (3) and (4). These take place over Cu/Zn/Al oxide catalyst in a PBR. This step operates under isothermal conditions for which the allowable temperature range is between 180ºC and 280ºC according to Bussche and Froment [14]. The inlet pressure used is in the range of 75 bars to 82 bars. The conditions that were used for design purposes were 245ºC and 7600 kPa. At such conditions, the reaction occurs in the vapor phase and therefore, the effluent is composed of a mixture of MeOH, H_2O and unreacted gases.

The effluent stream is cooled down to 60ºC, allowing the separation of MeOH and H_2O from the unreacted gases via a distillation column. The liquid mixture of MeOH and H_2O is then introduced to another distillation column, where MeOH is recovered as a liquid stream from the top with 95% purity due to the presence of some CO_2 in the system, while H_2O leaves as the bottoms at 99.9% purity. The column is operated with a pressure gradient of 1255 to 1290 kPa with 23 trays.

D. Methanol to Gasoline (MTG)

This procedure is developed by the international company Exxon Mobil. The main idea is to polymerize MeOH into alkanes with the help of a catalyst. Typically, zeolite catalysts are employed, which are adsorbents and catalysts found in around 40 different forms. They would be used at the second stage of the process. The first stage involves the conversion of MeOH to DME through dehydration as shown in (15).

$$2CH_3OH \leftrightharpoons (CH_3)_2O + H_2O \qquad (15)$$
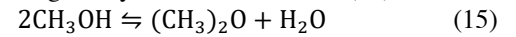
The reaction is equilibrium limited and has a conversion of around 84% at optimal conditions [15]. A minimum operation pressure is 1500 kPa. The reactor is operated adiabatically. It may be operated isothermally but a medium would be required to remove the heat generated. The reaction occurs over high-purity γ-alumina catalyst, and in order to prevent catalyst deactivation and degradation, a maximum temperature of 400ºC is to be used throughout the whole reactor. The catalyst pellets are 3mm in diameter and have a bulk density of 940 kg/m^3. The DME produced is then depressurized to 5000 kPa before being fed to the next stage going from DME to gasoline.

In the second stage of the process, DME is reacted over ZSM-5 zeolite catalyst, on which chain growth of molecules is sterically hindered, hence only allowing for the production of gasoline and lighter molecules. The kinetic models proposed can be grouped into lumped models which are a compromise between simplicity for modelling purposes and representation of the reality of the process. Detailed reactions of the transformation of MeOH to gasoline involve a large number of equations which include the intermediate carbene; therefore, it is very difficult to estimate the rate constants for each reaction [16]. For design purposes, a molecular model involving lumped parameters, shown in (16) to (19), is used. The rate constants of these reactions and the associated parameters are determined experimentally by fitting the results of mass fractions of various lumps to the corresponding mass conservation equations.

$$A \rightarrow C \qquad (16)$$
$$2C \rightarrow D_1 \qquad (17)$$
$$A + D_1 \rightarrow D_2 \qquad (18)$$
$$C + D_1 \rightarrow D_3 \qquad (19)$$
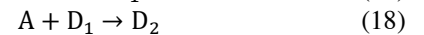
The lumped model can be described as having MeOH-DME mixture, also known as oxygenates, as a single species (A). The oxygenates, mainly DME, react to form light olefins, ethylene and propylene (C) which oligomerize to form products in the gasoline range (D1). These products can further react with

MeOH/DME or light olefins to produce heavier products of gasoline, D2 and D3, respectively [17]. For simulation purposes, certain assumptions had to be made to apply these kinetics for gasoline production:

1. Catalyst deactivation was not considered in this model.
2. The reactor was modeled as isothermal and a fixed bed.
3. Oxygenates (A) were modeled as DME due to its higher composition in the inlet stream.
4. Water was used to balance reactions where necessary.
5. Lighter hydrocarbons are assumed to be produced in (17), followed by heavier ones in the following two reactions due to greater branching.
6. Light olefins (C) were modeled as propene, the lighter hydrocarbons in the gasoline range (D1) are represented as cyclohexane, the middle hydrocarbons in gasoline (D2) are modeled as cycloheptane and the heaviest hydrocarbons in gasoline (D3) are modeled as cyclodecane.

The effluent stream is passed through a 3-phase separator followed by a series of distillation columns to get the desired composition and grade of gasoline. 4710 barrels/day of gasoline are produced in this plant, which can be sold in the market after adding certain agents to ensure that the gasoline abides by the EPA standards.

IV. DESIGN CONSIDERATIONS

Mass and energy balances were performed on each of the major unit operations in order to design them, including the PBR and the distillation column. The general mass and energy equations at steady state are given in (20) and (21).

$$0 = \dot{m}_{in} - \dot{m}_{out} + \dot{m}_{generated} - \dot{m}_{consumed} \quad (20)$$

Where, $\frac{dm}{dt}$ is the mass accumulation (kg/s), $\dot{m}_{in}$ and $\dot{m}_{out}$ are the mass flows going into and out of the unit (kg/s) respectively, and $\dot{m}_{generated}$ is the rate of mass generation (kg/s).

$$0 = \dot{E}_{in} - \dot{E}_{out} + \sum \dot{Q}_j + \dot{W}_s \quad (21)$$

Where, $\frac{dE}{dt}$ is the energy accumulation term (W), $\dot{E}_{in}$ and $\dot{E}_{out}$ are the energy input and output of the unit (W) respectively, $\sum \dot{Q}_j$ is the net energy produced by the unit (W) and $\dot{W}_s$ is the shaft work produced by the unit (W).

A. Packed Bed Reactor

The design of a PBR requires using the Ergun equation along with the overall component mole balance equation in (22).

$$\frac{dF_A}{dW} = r'_A \quad (22)$$

Where, F_A is the flowrate of component A (kmol/s), W is the catalyst weight (kg) and r'_A is the specific reaction rate (kmol/kg catalyst.s). These equations were applied to the MeOH synthesis reactor and yielded the parameters shown in Table 3.

B. Distillation Column

The conventional mass and energy balances along with the Erbar-Maddox correlation relating the minimum and actual reflux ratios to the minimum and actual number of stages of the column were used for the design of the distillation column. The

TABLE 3 – DIMENSIONS OF MEOH-SYNTHESIS PBR

|	
Tube outer diameter (d_0)	0.0381 m
Tube inner diameter (d_i)	0.0313 m
Total number of tubes per vessel (N_t)	2500
Tube volume	0.0146 m^3
Tube length (L)	13 m
L/d_0	341.2
Surface area of tubes in a vessel (A_o)	3889 m^3
Shell diameter (D_s)	2.73 m
L/D_s	~5

dimensions shown in Table 4 were obtained for the column separating water from MeOH.

V. OTHER CONSIDERATIONS

After modelling the plant, certain aspects have to be taken into account to assess the consequences of the implementation of such a plant. These include: by-products of the processes, plant location, safety considerations and environmental impact.

A. By-Products

There are several by-products of the previously mentioned processes that require further treatment. The first one is the sour gas produced from the regeneration of the rich amine in the scrubbing section of the plant. This stream is mainly composed of CO_2 and H_2S, with 86 wt% and 12 wt%, respectively. It is sent for further waste processing in a Claus process. This process converts H_2S into elemental sulfur in the presence of air as shown in (23) and (24). Fig. 2 shows the block flow diagram of the Claus process.

$$H_2S + 1.5 O_2 \rightarrow SO_2 + H_2O \quad (23)$$
$$2H_2S + O_2 \rightarrow 2S + 2H_2O \quad (24)$$
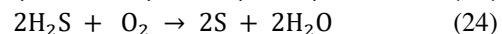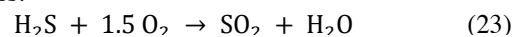
The second by-product is unreacted DME from the MTG. This can be obtained from the 3-phase separator following the second PBR in the process. DME is then separated from the mixture of gases and can be used as a fuel after undergoing necessary treatment.

Ash is another by-product of the gasification process. It can be used in the manufacture of cement concrete, roofing material and bricks. EPA supports the use of ash in an encapsulated form due to the various benefits this has, such as lowering the greenhouse gas emissions, reducing the cost of ash disposal and improving the strength and durability of material [18].

TABLE 4 – DIMENSIONS OF MEOH-WATER COLUMN

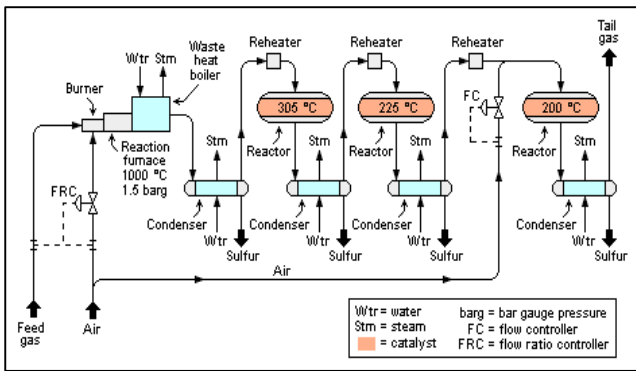|	
Number of stages (N)	52
Reflux Ratio (R)	0.52
Column Diameter (D_c)	4.8 m
Column height (H)	52 m
Pressure drop (ΔP)	330 kPa
Number of holes per tray (N_h)	17,472
Residence time (τ)	9.6 h

Figure 2 - Claus process schematic diagram

Any steam leaving the plant can be used as a heat source for heat exchangers elsewhere in the plant. This is done by carrying out heat integration calculations on the whole plant.

B. Scale of Biomass to Gasoline Plants

There is no current plant on a commercial scale converting biomass to gasoline, neither is there one converting biomass to methanol. However, several projects have demonstrated the biomass to methanol conversion on a pilot scale such as the *Hynol* project in the US, the *Bio-Meet* and *Bio-Fuels* projects in Sweden and the *BGMSS* project in Japan. These showed many problems associated with that conversion including its high energy consumption [19].

Each of the stages of production is available individually on a commercial scale. Gasification was used to produce combustible gases from organic feeds using blast furnaces since 180 years. Gasifiers are also used in Swedish strategic emergency plants. At present, there are about 64 gasification plants around the world operating on coal [20]. As for scrubbing, it is commonly used in the cleaning process of sour gas in the oil and gas industry. *Dow Chemical* and *Union Carbide* developed a CO_2 scrubbing process using MEA in the 1970s and 1980s. Since then, 21 commercial plants have been built, including 10 large plants and 11 skid-mounted ones [21]. Syngas to methanol plants also do exist on the commercial scale, but not all of them necessarily use the same process as the one described in this paper. One of these plants is located in Trinidad and Tobago, operational since 2004, and producing 1.7 million tons/year [22]. As for MTG process plants, they are as well available on the commercial scale. The *JAMG* plant in Shanxi, China, is an example of such plants, operating since 2009 and producing 2,500 barrels/day of gasoline [23].

C. Plant Location

Three potential locations of the gasoline production plant were screened using a selection matrix. The screened locations were the vicinities of Akkar in North Lebanon, Zahle, in Bekaa, and Sibline in South Lebanon, all of which are close to continuous sources of water due to the plant's constant need for H_2O. The selection matrix included ten factors, with a weight of 10% each: marketing area, raw materials, transport, availability of labor, utilities, environment, local community, land, climate and

Figure 3 - Hazards of Gasoline

politics. The selection matrix showed that Sibline, in South Lebanon, would be the best match for the plant construction.

D. Safety Considerations

Gasoline is a highly flammable fuel that requires extensive precautions to deal with to isolate it from any ignition sources. Its Threshold Limit Value (TLV) is 300 ppm as an eight-hour time weighted average (TWA). The NFPA diamond of gasoline is shown in Fig. 3 [24].

E. Environmental Related Issues

Gasoline has serious environmental impacts on its surrounding due to the extensive list of pollutants it emits. Three of these pollutants are the weightiest due to their impact on the environment: CO, hydrocarbons and NO_x. CO is the product of incomplete combustion of gasoline and can be life threatening at very high levels. Hydrocarbons, under sunlight, are precursors to the formation of ozone, which is detrimental to health at ground levels. As for NO_x, which are produced by the combustion of gasoline, they contribute to the formation of acid rain as well as ground-level ozone. Other pollutants also include CO_2, SO_2, PM and VOCs [24].

VI. ECONOMIC ANALYSIS

Since the aim of the paper is primarily the modeling and design of the biomass to gasoline process, a mere preliminary economic analysis on the operating costs was conducted. Initially, this analysis resulted is a loss of $300,000/day. However, with governmental support eliminating the tariff of sea water usage, profit will be generated, as shown by Table 5.

Optimization, heat integration and control of the units are not performed yet. However, when performed, they will help in further reducing the operating costs of the plant.

TABLE 5 – PRELIMANARY ECONOMIC ANALYSIS

Utility	Cost ($/day)
Water	27,700
Steam	278,000
Sea water	0
Power	362,000
Profit from MSW	654,000
Profit from gasoline	365,000
Total	351,000

VII. CONCLUSION AND RECOMMENDATIONS

In conclusion, it seems like, design wise, the process of converting biomass to gasoline through the thermal pathway of gasification is feasible. The units involved and their dimensions are achievable. The by-products of the process are also manageable and a suitable location to operate the plant is available. Although the process gets rid of Lebanon's waste, its product has emissions with a serious environmental impact. Safety wise, gasoline can also sensitive to deal with. At the current stage, governmental support is needed for the plant to be able to generate profit and break even.

It is recommended that the plant would be constructed since it can efficiently deal with more than half of Lebanon's waste while at the same time, provide it with its main transportation fuel. Stringent regulations have to be put on the plant's process to insure the production of a high quality gasoline with the least environmental impact. It would also be strategic for the public sector to operate the plant since this would imply a lower cost of utilities as well as cheaper gasoline for citizens.

REFERENCES

[1] M. Ei-Fadel and H. Sbayti, "Economics of mitigating greenhouse gas emissions from solid waste in Lebanon," Waste Management & Research, vol. 18, no. 4, pp. 329–340, Jan. 2000.

[2] M. N. Rahal, "State and Trends of the Lebanese Environment," 2011. [Online]. Available at: http://www.undp.org/content/dam/lebanon/docs/energy and environment/publications/soer_en.pdf.

[3] J. Ahlbäck, "Green Jobs Assessment in Lebanon - Prelimenary Assessment Waste Management," *International Labor Force*, 2012. [Online]. Available at: http://www.ilo.org/wcmsp5/groups/public/---arabstates/---ro-beirut/documents/genericdocument/wcms_210689.pdf.

[4] J. Al Assad, "Status of Lebanese Energy Statistics," *Escwa*, Apr-2012. [Online]. Available at: http://css.escwa.org.lb/sd/1810/lebanon.pdf. *Chicago | GAS TECHNOLOGY INSTITUTE*. [Online]. Available at: http://www.gastechnology.org/.

[5] "TRADING ECONOMICS | 300.00 INDICATORS | 196 COUNTRIES," *TRADING ECONOMICS*, 2016. [Online]. Available at: http://www.tradingeconomics.com/lebanon/fossil-fuel-energyconsumption-percent-of-total-wb-data.html. .

[6] A. M. Al-Jarallah, U. A. El-Nafaty, and M. M. Abdillahi, "Effects of metal impregnation on the activity, selectivity and deactivation of a high silica MFI zeolite when converting methanol to light alkenes," *Applied Catalysis A: General*, vol. 154, no. 1-2, pp. 117–127, 1997.

[7] P. Ana, H. L. Betania, and M. F. Rubens, "Application of Biomass to Hydrogen and Syngas Production," *The Italian Association of Chemical Engineering*, vol. 32, 2013.

[8] "A Two-Stage Anaerobic Digestion Process for Converting Waste to Energy," *Energy Research & Development, Training and Consulting -*

[9] P. Basu, *Biomass gasification, pyrolysis, and torrefaction: practical design and theory*, 2nd ed. San Diego: Elsevier Inc., 2013.

[10] P. Reubroycharoen, T. Yamagami, T. Vitidsant, Y. Yoneyama, M. Ito, and N. Tsubaki, "Continuous Low-Temperature Methanol Synthesis from Syngas Using Alcohol Promoters," *Energy & Fuels Energy Fuels*, vol. 17, no. 4, pp. 817–821, 2003.

[11] S. V. Vassilev, D. Baxter, L. K. Anderson and C. G. Vassilev, "An Overview of the Chemical Composition of Biomass," *Fuel*, vol.89, pp. 913-933, 2010.

[12] L. Zhong, W. Mei and Z. Hong, "Kinetic Model Establishment and Verification of the Biomass Gasification of Fluidized Bed," *IEEE*, vol.09, pp. 2112-2117, July 2009.

[13] Z. Aliabad, H. and S. Mirzaei, "Removal of CO2 and H2S using Aqueous Alkanolamine Solutions," *International Journal of Chemical, Molecular, Nuclear, Materials and Metallurgical Engineering*, vol. 03, pp. 50-59, 2009.

[14] D. Perko and J. Levec, "Kinetic Study of Methanol Synthesis over CuO/ZnO/Al2O3/V2O3 Catalyst Deposited on a Stainless Steel Surface," *American Chemical Society*, Slovenia, vol.51, pp. 710-718, 2012.

[15] A. Khaleel, "Methanol Dehydration to Dimethyl Ether over Highly Porous Xerogel Alumina Catalyst: Flow Rate Effect," *Fuel Processing Technology*, Al-Ain, vol.91, pp.1505-1509, 2010.

[16] S. Joseph and Y. T. Shah, "Methanol to Gasoline Process," *University of Pittsburgh*, Pittsburg, 15261.

[17] J. M. Ortega, A. G. Gayubo, A. T. Aguayo, M. Olazar and J. Bilbao, "MTG Process in a Fluidized Bed with Catalyst Circulation: Operation and Simulation of an Experiment Unit," *Industrial Engineering and Chemistry Research*, Bilbao, vol.37, pp. 4222-4230, 1998.

[18] "Coal Ash Reuse," *EPA*, 08-Jul-2015. [Online]. Available at: https://www.epa.gov/coalash/coal-ash-reuse [Accessed: 18-Mar-2016].

[19] H. Li, H. Hong, H. Jin, and R. Cai, "Analysis of a feasible polygeneration system for power and methanol production taking natural gas and biomass as materials," *Applied Energy*, vol. 87, no. 9, pp. 2846–2853, Sep. 2010.

[20] A. K. Rajvanshi, "Biomass Gasification," in Alternative Energy in Agriculture, vol. 2, CRC Press, 1986, pp. 83–102.

[21] D. G. Chapel and C. L. Mariz, "Recovery of CO2 from Flue Gases: Commercial Trends," Oct. 1999.

[22] "Atlas Methanol Production Plant, Point Lisas," Chemicals Technology. [Online]. Available at: http://www.chemicals-technology.com/projects/atlas_methanol/. [Accessed: 12-Apr-2016].

[23] T. Helton and M. H., "Methanol to Gasoline Technology An Alternative for Liquid Fuel Production ," exxonmobil, 30-Jul-2014. [Online]. Available at: http://cdn.exxonmobil.com/~/media/global/files/catalyst-and-licensing/2014-1551-mtg-gtl.pdf .

[24] "Gasoline Emissions," *Gasoline Emissions*. [Online]. Available at: http://www.ecooptimized.com/index.php/gasoline/76gasoline-emissions-.html . [Accessed: 19-Mar-2016].

VIII.   APPENDIX

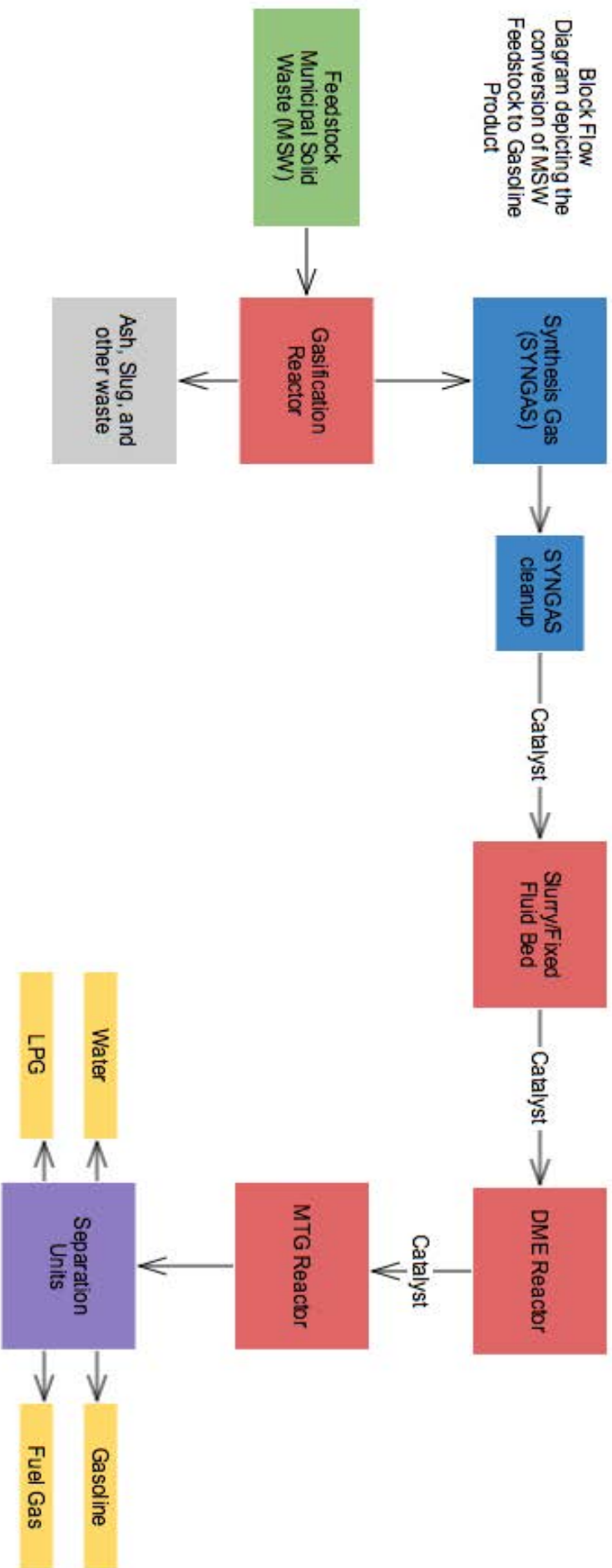Block Flow Diagram depicting the conversion of MSW Feedstock to Gasoline Product

Figure 4 - Plant Block Flow Diagram

151

Design of a Biomass to Gasoline Production Plant in Lebanon

Rasha Atwi, Chana Chammoun, Serine Haidar Ahmad, Caroline Rajeh & Lynn Salamoun
Department of Chemical Engineering
American University of Beirut
Riad El-Solh / Beirut 1107 2020, Lebanon
P.O.Box 11-0236
rma123@mail.aub.edu/ cfc02@mail.aub.edu/ shh44@mail.aub.edu
cnr04@mail.aub.edu/ lgs09@mail.aub.edu

Abstract - **The project at hand deals with the conversion of a biomass and sludge feed to gasoline via a series of reactions and processes. First, the biomass is pretreated, then it undergoes anaerobic digestion whereby different types of bacteria degrade it to produce biogas, which is a mixture of methane (60%), carbon dioxide (40%) and traces of other compounds such as hydrogen sulfide and ammonia. These sulfur and nitrogen compounds cause catalyst poisoning and corrosion, thus a desulfurization process is designed via the use of Fe(III)EDTA solution. Next, the biogas is converted to syngas, which is a mixture of carbon monoxide and hydrogen, through the process of reforming. Thus, upon the arrival to the reformer units, the carbon dioxide is removed via the first dry reforming reaction with a conversion of 21% and with a H_2/CO ratio of 1. The latter stream enters the steam reformer and achieves a conversion of methane equal to 46% and a H_2/CO ratio of 2:1. This stream is then taken to the MDEA unit to clean the syngas from any remaining carbon dioxide. Then, the syngas is converted to syncrude via Fischer Tropsch reaction. This step is followed by a hydrocracking process whereby fuel gas, LPG, gasoline, diesel and waxes are produced in the first stage, and whereby diesel and waxes are hydrocracked into gasoline in the second stage. Hydrocracking achieves a 92.21% conversion of waxes into 92% pure gasoline. The end result is the production of 137.1 kg/hr of 99% pure gasoline.**

I. INTRODUCTION

Energy requirements are increasing with the decrease in the abundance of reserves and an increase in the prices of energy sources and raw materials. Fossil fuels serve as a major source of energy in Lebanon whereby 29% of this energy is consumed by the electricity sector and 36% is consumed by the transportation sector [2]. Since Lebanon is not an oil producing country, it imports about 96% of its primary energy from abroad. Gasoline is among the most demanded because it is used for fueling internal combustion engines for cars, motorbikes, trucks and other transport vehicles. The fact that at least 75% of households own a car in Lebanon justifies the large demand for gasoline [2]. Aside from the deficit of energy sources in Lebanon, the country has been facing a garbage crisis due to the accumulation of 10,000,000 tons in a sanitary landfill near the town of Naameh [3]. Up until July 17, 2015, the landfill was receiving 45% of Lebanon garbage which is estimated to be 4,000 Ton/day, equivalent to 1.5 million tons/year, including domestic, industrial and medical waste [4].

We, as engineers, are expected to find practical and efficient solutions to these problems. One way to do this is to use the excessive waste that is accumulating in Lebanon and transform it through a series of treatments and reactions beginning with Anaerobic Digestion to produce biogas, a mixture of methane and carbon dioxide, then transform the biogas to synthesis gas, referred to as syngas which is a mixture of hydrogen gas and carbon monoxide, via a Reforming process, and finally the syngas is converted to gasoline by the Fischer Tropsch technology. A further Hydrocracking step is needed to achieve better yield of gasoline.

II. METHODS

A. Biomass Pretreatment

The chosen feed constituted of 75% slaughterhouse wastes, 15% food wastes and 10% cow manure which resulted in an overall waste composition of 23.5% carbohydrates, 12.18% proteins, 60% fats and the rest is ashes [7]. The loading rate for the solid biomass was chosen to be 10000 kg/hr, similar to the loading rate used in the A.D. plant in Sidon, Lebanon. The loading rate of the mixed sludge stream was chosen to be 14400 L/hr, equivalent to the amount of sewage produced in Lebanon [8].

Biomass underwent a pretreatment process before entering the Anaerobic Digestion (A.D.) plant to ensure uniform particle size and to aid the hydrolysis step of A.D. in particular. It was mechanically treated to obtain smaller biomass substrate shards that result in easier flow and better enzymatic activity in the A.D. process due to the increase in the substrate surface area. The substrate was also heated up to 125℃.

B. Biomass to Biogas

Biomass can be converted to a gas mixture mainly composed to carbon dioxide and hydrogen via different pathways, the most important of which are Anaerobic Digestion and Gasification.

Gasification is a process that converts biomass to producer gas in the presence of an oxidizing agent, such as steam, air or oxygen, at high temperatures (>700ºC). Producer gas is a mixture of carbon dioxide, hydrogen carbon monoxide, nitrogen,

sulfur, char and alkali compounds. It requires an extensive cleaning process to prevent plugging of downstream equipment due to the deposition of char, corrosion as a result of alkali compounds sticking on metal surfaces, the formation of ammonia from nitrogen which when combusted results in nitrogen oxide which is polluting and the poisoning of the catalyst due to presence of sulfur.

Anaerobic Digestion is the biological process that converts organic matter into biogas under the action of "close-knit" communities of bacteria in the absence of oxygen [5]. AD pathways involve complex biological steps: hydrolysis, acidogenesis, acetogenesis and methanogenesis. The temperature conditions for A.D. can either be chosen as mesophilic conditions, i.e. a temperature range of 20-37°C, at which the residence time is 15-30 days, or they can be chosen as thermophilic conditions, i.e. a temperature range of 50-65°C, at which the residence time is 12-14 days.

A.D. rather than Gasification was selected as the method of choice for the conversion of biomass to biogas. A.D. is less energy intensive than gasification since it requires significantly lower temperature conditions: the operating temperature for A.D. range from 20-65°C whereas the operating temperature for Gasification needs to be greater than 700°C. Moreover, A.D. does not result in ash and char formation as opposed to Gasification. This implies that a less expensive cleaning process for the gas produced in A.D. is required.

The kinetic constants corresponding to A.D. were obtained from experimental studies that rely on calculator blocks. The latter have FORTRAN program which calculates the products released at each step [7]. These calculator blocks take different variables, such as flow rate, volatile fatty acid and ammonia concentrations and temperature. Moreover, this study provides the reaction rates by using the power law, which includes specific growth rate of microorganisms and ammonia inhibitions [7].

A.D. was modeled according to wet digestion because the feed chosen contained high moisture content. Second, A.D. reactions were carried in three stage Continuously Stirred Tank Reactors (CSTRs) operating at thermophilic conditions (55℃) with a total retention time of 12 days to optimize conditions for each step of the A.D. In CSTR 1, hydrolysis reactions occur. In CSTR 2, acidogenesis and acetogenesis took place, and in CSTR 3 acetogenesis and methanogenesis to benefit from the synergetic relationship between hydrogen producing bacteria, (acetogens) and hydrogen scavenging bacteria (methanogens). The A.D. process was modeled on *Aspen Plus* with the Non-Random Two Liquid Model (NRTL) chosen as the fluid package.

A nutrient-rich digestate, typically used as a bio-fertilizer, resulted from the A.D. process, and it was subjected to dewatering and water treatment.

C. Biogas Treatment: Desulfurization

Hydrogen sulfide (H_2S) formed during A.D. process should be removed from the biogas mixture to prevent catalyst poisoning in the upcoming reforming process and to reduce risk of corrosion. Chemical oxidation using iron-chelated solution was chosen as the desulfurization method. Iron (III) held in Ethylene Diaminetetraacetic Acid (EDTA), a chelating solution, was capable of converting H_2S into elemental sulfur in a packed column [20]. The reaction that takes place between Fe (III)EDTA and H_2S is:

$$4Fe(III)EDTA + H_2S \rightarrow 4Fe(II)EDTA + 4H^+ + S$$
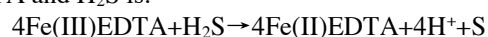
Using Iron (III) solution has various advantages. First, Iron (III) solution has a removal efficiency of 97%. Second, Iron (III) solution does not react with carbon dioxide and methane which are the primary components of biogas, so there is no loss of CO_2 and CH_4 that are used to produce syngas. Third, Iron (III) can be regenerated. Finally, the method is a low cost H_2S removal system.

The kinetic framework for the design of the proposed desulfurization method relies on absorption- reaction-transport model. This means that mass transfer is modeled in parallel to the chemical reaction [5]. This process was applied in an industry sized column packed with bio-balls, with diameter of 5 cm, and with a diameter of 0.5 m and a height of 0.8 m. The absorption column was considered to be a packed bed reactor since absorption and chemical reactions were taking place in series [5]. The rate constant of this process includes the rate of mass transfer of H_2S from the bulk gas to the gas-liquid interface and from the gas-liquid interface to the bulk liquid, and the reaction rate. It was found to be 5.7 seconds according to the following equation:

The chelated iron process is applied in an industry sized packed column with a diameter of 0.5 m and with a height of 0.8 m. The column is packed with bio-balls, with diameter (dp) of 5 cm, for removing entrained droplets from the gas stream. The absorption column is considered to be a packed bed reactor since absorption and the reaction are taking place in series [9].

The stream containing Fe(II)EDTA solution and sulfur was sent to a sedimentation tank. The sulfur settled at the bottom of the tank, and the Fe(II)EDTA solution was taken to another unit to regenerate Fe(III)EDTA by oxidizing it with air according to the following reaction:

$$4Fe(II)EDTA + O_2 + 4H_2O \rightarrow 4Fe(III)EDTA + 4OH^-$$
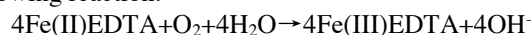
The reaction follows second order kinetics with respect to Fe(II)EDTA and first order kinetics with respect to oxygen in the air [20]. This results in a k value of 5.56×10^{-3} $m^6/mol^2.s$ and activation energy equal to 27.2 kJ/mole.

Desulfurization is modeled on *Aspen Plus* using the NRTL fluid package.

D. Biogas to Syngas

To convert the resulting biogas to syngas, a reforming process is implemented to produce gases with high hydrogen yield for fuel cells and gas engines and to obtain a pre-determined H_2/CO ratio for further industrial processes, like gasoline production in Fischer Tropsch reactors. There are three types of reforming processes: dry reforming, steam reforming and autothermal

reforming. Dry reforming is the most commonly used process in the industry which results in the almost 100% conversion of carbon dioxide, a greenhouse gas. However, coke formation that results in catalyst poisoning is an issue in dry reforming. Steam reforming requires low operation temperatures, and it gives the best H_2/CO ratio for fuel production. However, it results in high percentages of air emissions, and it initially requires high capital costs. Autothermal reforming requires an oxygen production plant which is expensive. Moreover, in the case of autothermal reforming commercial and scientific experience is limited.

In this project, dry and steam reforming are coupled in series to achieve a desirable H_2/CO ratio of 2 suitable for of two for gasoline production out of Fischer Tropsch technology. The types of reforming processes were chosen based on the energy requirements, the capital and operational costs and the environmental impact.

Dry reforming occurs in a homogeneous fixed bed according to the following heterogeneous catalytic reaction:

$$CH_4 + CO_2 \rightarrow 2H_2 + 2CO \qquad \Delta H_{298}^0 = +247 \text{ KJ/mol}$$
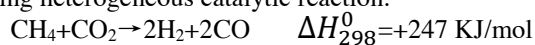
The reaction is highly endothermic; it requires an excessively high temperature to reach a high conversion. Coke formation inhibits the activity of the catalyst particles by filling their active sites at relatively high temperatures, between 500°C and 600°C. This results in a pressure increase inside the reactor's tubes. Coke formation and consumption are described by the following reactions:

$$C + H_2O \leftrightarrow CO + H_2$$
$$C + CO_2 \leftrightarrow 2CO$$
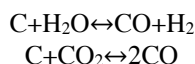
Nevertheless, one advantage of resorting to this method is the low H_2/CO ratio obtained, ranging between 0.8 and 1.2. This allows the production of liquid oxygenated products such as aldehydes and alcohols from biogas streams. Moreover, water gas shift (WGS) and reverse WGS reactions usually accompany all types of reforming reactions, but at different rates. These reactions are described below:

$$CO + H_2O \leftrightarrow CO_2 + H_2$$
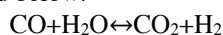
Steam reforming is a heterogeneous catalytic reaction that takes place in a packed bed tubular reactor (PBR). Coke formation also occurs but at higher temperature and remains less drastic than in the case of dry reforming. WGS reactions take place, leading to the following overall exothermic reaction:

$$CH_4 + 2H_2O \leftrightarrow CO_2 + 4H_2$$
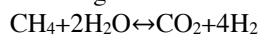
The choice of the catalyst influences the rate of the reaction, the level of coke accumulation as well as the operating conditions such as temperature and pressure. Nickel based catalysts were chosen for both dry and steam reforming because they are used extensively in industry due to their low cost (100 to 150 times less expensive than the Ruthenium based catalysts). However, the temperature and pressure conditions have to be accommodated in order to avoid carbonaceous accumulations and thus poisoning of the catalyst.

The reforming process is modeled on *Aspen Hysys* using Peng Robinson (PR) as the fluid package.

E. Treatment of Syngas

The syngas was treated to remove the carbon dioxide. Methyl Diethanolamine (MDEA) is a tertiary amine, used in industrial setups for its selectivity towards carbon dioxide. It was therefore used to treat acid gas and clean it with greater efficiency than pure water scrubbing. It is also resistant to degradation and is less corrosive than primary and secondary amines [4].
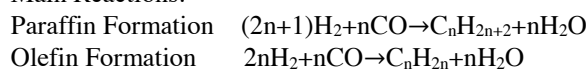
F. Syngas to Syncrude

To convert the treated syngas to the desired product which is gasoline, two types of processes can be used: Fischer Tropsch Synthesis (FTS) that produces syncrude which can be upgraded to gasoline and ExxonMobil process which converts syngas to methanol followed by the conversion of methanol to gasoline. In case of the ExxonMobil process, the thermodynamics are not in favor of methanol production, thus a high operating pressure (10 MPa) is required to drive the reaction. Whereas in the case of FTS, the process is almost free of thermodynamic limitations, thus low pressures (1 MPa) are used. In addition, FTS gives high conversion of syncrude to hydrocarbons without the need to pass through an intermediate as opposed to the ExxonMobil process which necessitates the passage of syncrude to an intermediate (methanol) to produce gasoline. In addition, the process of converting methanol to gasoline is a licensed technology by ExxonMobil, thus the access to scientific literature is limited. On the other hand, FTS is a well-established technology. However, the ExxonMobil process produces gasoline-range hydrocarbons rich in aromatics which meets the market requirements, whereas the FTS gasoline product has low quality for marketing and needs to be further refined to meet the market requirements. Thus, the resulting costs for gasoline production via ExxonMobil process are lower (0.82US$/L) compared to those needed for gasoline production via FTS (0.88 US$/L) [1]. Based on this comparative analysis between the two methods, FTS was selected as the method for the conversion of syngas to gasoline.
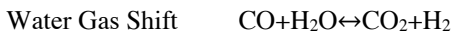
During FTS, syngas is converted into a mixture of hydrocarbons, oxygenates, water and carbon dioxide by a series of chemical reactions. The mixture of hydrocarbons and oxygenates is referred to as synthetic crude oil or syncrude, which like conventional crude oil requires refining to be used in the market as chemicals or motor fuels.

Reactions that take place during the conversion of syngas over FTS heterogeneous transition metal catalyst are divided into: main reactions, side reactions and reactions responsible for catalyst modification [1]:
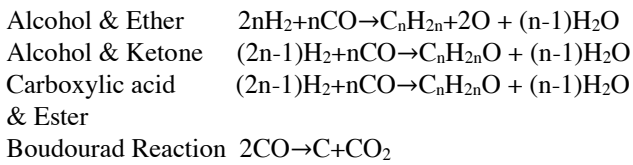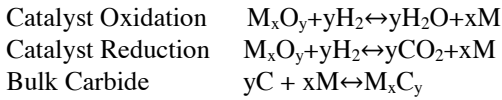
Main Reactions:
Paraffin Formation $\quad (2n+1)H_2 + nCO \rightarrow C_nH_{2n+2} + nH_2O$
Olefin Formation $\quad 2nH_2 + nCO \rightarrow C_nH_{2n} + nH_2O$

| Water Gas Shift | $CO+H_2O \leftrightarrow CO_2+H_2$ |

Side Reactions:

Alcohol & Ether	$2nH_2+nCO \rightarrow C_nH_{2n}+2O + (n-1)H_2O$
Alcohol & Ketone	$(2n-1)H_2+nCO \rightarrow C_nH_{2n}O + (n-1)H_2O$
Carboxylic acid & Ester	$(2n-1)H_2+nCO \rightarrow C_nH_{2n}O + (n-1)H_2O$
Boudourad Reaction	$2CO \rightarrow C+CO_2$

Catalyst Modification:

Catalyst Oxidation	$M_xO_y+yH_2 \leftrightarrow yH_2O+xM$
Catalyst Reduction	$M_xO_y+yH_2 \leftrightarrow yCO_2+xM$
Bulk Carbide Formation	$yC + xM \leftrightarrow M_xC_y$

Where n is the carbon number

FTS can be simplified by the combination of these reactions:

| FT | $CO + (1+m/2n)H_2 \rightarrow (1/n)C_nH_m+H_2O$ |
| WGS | $CO+H_2O \leftrightarrow CO_2+H_2$ |
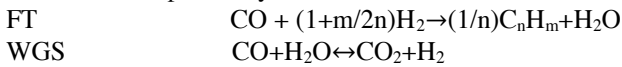
Where n is the average number of carbon atoms, and m is the average number of hydrogen atoms.

FTS produces a wide range of hydrocarbons. However, modeling all the resulting products is usually impractical and time consuming. Therefore, existing kinetic model based on Anderson-Schulz-Flory (ASF) distribution was used for modeling a realistic and consistent product distribution from FTS. The options chosen for the optimum production of gasoline are summarized in Table 1:

TABLE I
SUMMARY OF FTS CONDITIONS

| Factor	Choice
Catalyst	Iron
Reactor	Fluidized Bed Reactor
Temperature	315℃
Pressure	2.0 MPa
H_2/CO Ratio	2/1

G. Syncrude to Gasoline

Products obtained from FTS consisted mainly of paraffins and olefins with carbon numbers ranging from 1 to 130. Thus, the wide range of hydrocarbons required further treatment to maximize the production of gasoline which is the end product. Thus, hydrocracking unit is used to crack long hydrocarbon molecules into smaller ones that specifically fall in the range of gasoline formation. Gasoline cuts contain hydrocarbons ranging from C_5 to C_{12}, so the feedstock into the hydrocracking unit contains hydrocarbons ranging from C_{13} to C_{24}, which were cracked into smaller molecules and later on combined with the gasoline produced directly from FTS to get the final yield of gasoline out of the plant.

A catalyst consisting of 80 wt% zeolite and 20 wt% binder on Silica-Alumina with a density of 860 kg/m^3 was chosen. The use of maximum amounts of zeolite allows low operating temperatures resulting in low production of undesired gaseous products such as C_1 to C_3 hydrocarbons. Both FT and hydrocracking were modeled on *Aspen Hysys* using Soave-Redlich-Kwong (SRK).

H. Process Flow Diagram

The process can be summarized in the following block flow diagram.
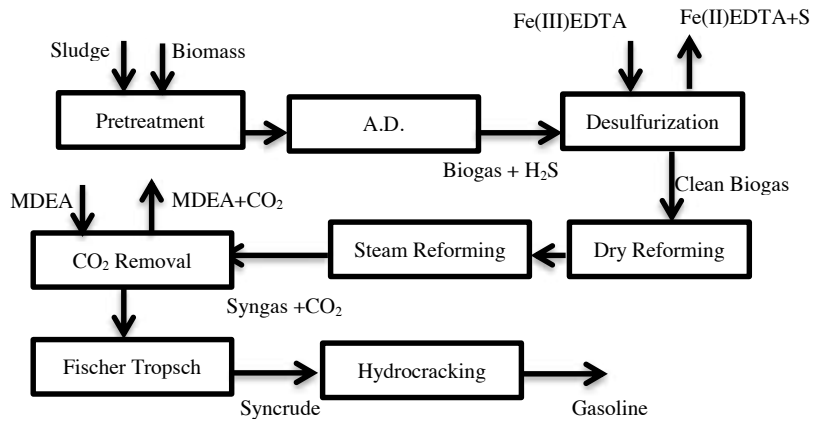
III. RESULTS

The plant was done on an industrial scale whereby the feed chosen was representative of the wastewater and the biomass amounts produced in Lebanon as a whole and the equipment sizes were scaled up.

A. Results of Anaerobic Digestion

During hydrolysis, complex organic compounds are biologically degraded to simpler intermediates by the addition of water and under the action of extra-cellular enzymes produced by enzyme secreting and fermentative microorganisms [6]. Hydrolysis is considered to be the rate limiting step [6]. Then, these intermediates undergo acidogenesis to obtain fatty acids: acetic acid, iso-butyric acid and iso-valeric acid, along with hydrogen gas, carbon dioxide [6]. Glucose, which was modeled as dextrose, is converted into acetate, butyrate and propionate, and glycerol is converted into propionate.

Finally, the biogas produced had a composition of 40wt% CO_2, 60 wt% CH_4, and traces of ammonia, hydrogen and hydrogen sulfide.

B. Results of Desulfurization

The Fe(III)EDTA solution had no effect on the other components of the biogas, as predicted, and was efficient in converting H_2S into sulfur. The mole flow of H_2S decreases from 12.2928 kmol/hr to 0.03 kmol/hr (99.7% removal).

C. Results of Reforming

The inlet stream to the reforming unit was heated to 500℃ using a furnace to avoid coke accumulation, then fed to a packed bed reactor (PBR) operating isothermally and at a pressure of 1 bar. The conversion was 21.55% for the dry reforming reaction

and 11.44% for the WGS reaction. This low conversion was expected due to the relatively low inlet temperature. The outlet stream entered the steam reforming PBR at a temperature of 750℃ and a pressure of 1 bar. A 46% conversion for the steam reforming reaction and a 7.974% conversion for the WGS reaction were obtained. The outlet H_2/CO ratio of the syngas reached a desirable value of 2.096 for the FT inlet and the stream was cooled to 15℃ to separate water form syngas.

D. Results of FT

The stream entering the FT reactor is free of carbon dioxide via MDEA cleaning. Results showed that an optimal temperature of 315℃ produced rapidly a great amount of hydrocarbons in the light gas and gasoline cuts, leaving little amount of heavy paraffins, at a pressure of 2 MPa. The hydrocracking unit resulted in a 44% increase in the production of gasoline, and the final product distribution after hydrocracking had between 58.8 and 60.6% of gasoline.

E. Results of Hydrocracking

The hydrocracking unit yielded an increase in the conversion in the gasoline cut up to 92.21% with traces of impurities mainly consisting of the unreacted hydrogen and heavier uncracked molecules. Different hydrocarbons with different carbon numbers yielded a different conversion percentage of their respective cracked molecules. The conversion ranges between 84.66% for C_{14} and 99.94% for C_{24}. The final aim was to send the end product to the export line and sell it as a transportation fuel. In order to do this, the temperature and pressure needed to be reduced to atmospheric conditions appropriate for transportation and storage. First, the hydrocracking products which came out of the hydrocracking unit at a very high pressure of 9.999 MPa were depressurized. The stream is liquid, therefore control valves were used in order to release some of the pressure and conduct separation. For safety reasons, the depressurization process needed to be done in four steps by using four control valves, each reducing the inlet pressure of the incoming stream by 1/3. The final stream had a pressure of 1 atm. This stream was combined with the stream of gasoline from the FTS section, and the outlet stream was cooled to room temperature (25℃).

IV. CONCLUSION

Although different operating conditions were tested during the implementation of the simulation, to try to optimize the results of yields and conversions mentioned above, the latter remain not optimal. Future work include the designing of the units, the optimization of the plant, and heat integration of the streams, the application of process control strategies to finalize a P&ID, the evaluation of a HAZOP study and finally, rendering the entire project an actually profitable one.

Some current recommendations can be suggested starting with the reforming process, where the dry reformer can be operated at a higher temperature, thus increasing the conversion.

Knowing that the latter will poison the catalyst due to the considerable rate of coke accumulation, a catalyst regeneration unit can be added to the plant. Moreover, one important product of the FT process is the LPG, which can be sold or burnt for internal energy requirements. The gasoline obtained can also be further treated or mixed with additives to yield a higher octane number especially that the price of gasoline is decreasing nowadays. Finally, a major quantity of water is obtained out of the plant's flash separators, and can be used either as a coolant in heat exchangers, or as a water makeup stream in the MDEA cleaning process, depending on its purity, i.e. the level of contaminants in it.

REFERENCES

[1] Wengenmayr, R., & Bührke, T. (Eds.). (2013). Renewable Energy: Sustainable EnergyConcepts for the Energy Change. John Wiley & Sons.
[2] Beheshti, H. (2010). Exploring Renewable Energy Policy in Lebanon: Feed-in Tariff as a Policy Tool in the Electricity Sector (Doctoral dissertation, American University of Beirut).
[3] (2015). Retrieved on December 1, 2015 from http://blog.etobb.com/en/lebanons-garbage-crisis-your-health-in-focus/.
[4] (2015). Retrieved on December 1, 2015 from http://eecoy.com/content/waste-problem.
[5] Arsova, L. (2010). Anaerobic digestion of food waste: Current status, problems and an alternative product. Columbia University.
[6] Wolfsberger, A. (2008). Modelling and control of the anaerobic digestion of energy crops. na.
[7] Rajendran, K., Kankanala, H. R., Lundin, M., & Taherzadeh, M. J. (2014). A novel process simulation model (PSM) for anaerobic digestion using Aspen Plus. Bioresource technology, 168, 7-13.
[8] El-Fadel, M., Zeinati, M., El-Jisr, K., & Jamali, D. (2001). Industrial-waste management in developing countries: The case of Lebanon. Journal of Environmental Management, 61(4), 281-300.
[9] Seibig, S., & van Eldik, R. (1997). Kinetics of [FeII (EDTA)] oxidation by molecular oxygen revisited. New evidence for a multistep mechanism. Inorganic Chemistry, 36(18), 4115-4120.

APPENDIX

Nomenclature Section:

- a_i: gas-liquid interfacial area per volume of bed (m²/m³)
- $C_{Fe(III)EDTA}$: concentration of Fe(III)EDTA in bulk liquid (mole/m³)
- E_{H2S}: the liquid film enhancement factor
- f_l: ratio of volume to liquid to the reactor volume.constant (m³/mol.s)
- $k_{1,d}$: rate constant for dry reforming reaction (mol/kgcat.s)
- $k_{2,d}$: rate constant for WGS reaction for dry reformining (mol /kgcat.s)
- $k_{1,s}$: rate constant for steam reforming reaction (mol/kgcat.s)
- $k_{2,s}$: rate constant for WGS reaction for steam reformining (mol/kgcat.s)
- $K_{1,s}$: equilibrium constant in $R_{1,s}$ (bar⁻¹)
- $K_{2,s}$: equilibrium constant in $R_{2,s}$ (bar⁰)
- $K_{CO2,1,d}$: adsorption coefficient of CO_2 in $R_{1,d}$ (bar⁻¹)
- $K_{CO2,2,d}$: adsorption coefficient of CO_2 in $R_{2,d}$ (bar⁻¹)
- $K_{CH4,1,d}$: adsorption coefficient of CH_4 in $R_{1,d}$ (bar⁻¹)
- $K_{H2,2,d}$: adsorption coefficient of H_2 in $R_{2,d}$ (bar⁻¹)
- $K_{CO,s}$: adsorption coefficient of CO in steam reforming reactions (bar⁻¹)
- $K_{H2,s}$: adsorption coefficient of H_2 in steam reforming reactions (bar⁻¹)
- $K_{CH4,s}$: adsorption coefficient of CH_4 in steam reforming reactions (bar⁻¹)
- $K_{H2O,s}$: adsorption coefficient of H_2O in steam reforming reactions (bar⁻²)
- K_{p1}: equilibrium constant in $R_{1,d}$ (bar²)
- K_{p2}: equilibrium constant in $R_{2,d}$ (bar²)
- k_{FT} is the Fischer Tropsch rate constant ($kmol.m^3.s.MPa$)
- k_g is the gas film mass transfer coefficient (m/s)
- k_L is the liquid film mass transfer coefficient (m/s)
- k_r is the reaction rate constant (m³/mol.s)
- P_{CH4} is the partial pressure of CH_4 (bar)
- P_{CO} is the partial pressure of CO (bar)
- P_{H2} is the partial pressure of H_2 (bar)
- P_{H2O} is the partial pressure of H_2O (bar)
- R_{des}: rate constant corresponding to the Desulfurization process (seconds)
- $R_{1,d}$: rate expression for dry reforming reaction (mol.s/kgcat)
- $R_{2,d}$: rate expression for WGS reaction for dry reformer (mol.s/kgcat)
- $R_{1,s}$: rate expression for steam reforming reaction (mol.s/kgcat)
- $R_{2,s}$: rate expression for WGS reaction for steam reformer (mol.s/kgcat)
- r_{FT} is the Fischer Tropsch rate constant (kmol/m³.s)
- T is the temperature (K)

The kinetic model for the Desulfurization process using Fe(III)EDTA solution is given as:

$$R_{des} = \left(\frac{1}{k_G \times a_i} + \frac{He}{k_L \times a_i \times E_{H2S}} + \frac{He}{k_r \times f_i \times C_{Fe(III)EDTA}} \right)^{-1} \tag{1}$$

The kinetic model corresponding to the Dry Reforming process is described using the following equations:

$$R_{1,d} = \frac{k_{1,d} K_{CO2,1,d} K_{CH4,1,d} P_{CO2} P_{CH4} \left(1 - \frac{P_{CO}^2 P_{H2}^2}{K_{P1,d} P_{CO2} P_{CH4}}\right)}{(DEN1_d)^2} \tag{2}$$

$$R_{2,d} = \frac{k_{2,d} K_{CO2,2,d} K_{H2,2,d} P_{CO2} P_{H2} \left(1 - \frac{P_{CO} P_{H2O}}{K_{P2,d} P_{CO2} P_{H2}}\right)}{(DEN2_d)^2} \tag{3}$$

$$DEN1_d = 1 + K_{CO2,1,d} P_{CO2} + K_{CH4,1,d} P_{CH4} \tag{4}$$

$$DEN2_d = 1 + K_{CO2,2,d} P_{CO2} + K_{H2,2,d} P_{H2} \tag{5}$$

$$K_{P1,d} = 6.78 * 1014 e^{-25966/RT} \tag{6}$$

$$K_{P2,d} = 56.4971 \, e^{-3658/RT} \tag{7}$$

$$k_{1,d} = 1.29 * 10^6 \, e^{-102065/RT} \tag{8}$$

$$k_{2,d} = 0.35 * 106 \, e^{-81030/RT} \tag{9}$$

$$K_{CO2,1,d} = 2.61 * 10^{-2} \, e^{37641/RT} \tag{10}$$

$$K_{CH4,1,d} = 2.60 * 10^{-2} \, e^{40684/RT} \tag{11}$$

$$K_{CO2,2,d} = 0.5771 \, e^{9262/RT} \tag{12}$$

$$K_{H2,2,d} = 1.494 \, e^{6025/RT} \tag{13}$$

The kinetic model for the Steam Reforming process is given using the following expressions:

$$R_{1,s} = (k_{1,s}/P_{H2}^{2.5}) (P_{CH4} P_{H2O} - (P_{H2}^3 P_{CO}/K_{1,S}))/(1/DEN_s^2) \tag{14}$$

$$R_{2,s} = (k_{2,s}/P_{H2}) (P_{CO} P_{H2O} - (P_{H2} P_{CO2}/K_{2,s})/(1/DEN_s^2) \tag{15}$$

$$DEN_s = 1 + K_{CO,s} P_{CO} + K_{H2,s} P_{H2} + K_{CH4,s} P_{CH4} + K_{H2O,s} P_{H2O} P_{H2} \tag{16}$$

$$K_{1,s} = e^{(-26830/T + 30.114)} \tag{17}$$

$$K_{2,s} = e^{(4400/T - 4.036)} \tag{18}$$

$$k_{1,s} = 1.17 * 10^{15} e^{-240100/RT} \tag{19}$$

$$k_{2,s} = 5.43 * 10^5 e^{-67130/RT} \tag{20}$$

$$K_{CH4,s} = 6.65 * 10^{-4} \, e^{(-38280)} \tag{21}$$

$$K_{H2O,s} = 1.77*10^5 \, e^{88680} \qquad (22)$$

$$K_{CO,s} = 8.23*10^{-5} \, e^{(-70650)} \qquad (23)$$

$$K_{H2,s} = 6.12*10^{-9} \, e^{(-82900)} \qquad (24)$$

The kinetic model for Fischer Tropsch Synthesis is described by the following equations:

$$r_{FT} = (k_{FT}P_{CO}P_{H2})/(P_{CO}+a_{FT}P_{H2O}) \qquad (25)$$

$$k_{FT} = 83127.305e^{-85000/RT} \qquad (26)$$

$$a_{FT} = 0.00463e^{8800/RT} \qquad (27)$$

7

Production of Diesel Fuel from Biomass and Wastewater Sludge

Abou Raad, S.; Azzam, M.; Saad, S.; Sebaaly, C.; Younes, G.

Department of Chemical and Petroleum Engineering

American University of Beirut

P.O.Box 11-0236 | Riad El-Solh / Beirut 1107 2020 Lebanon

sra27@mail.aub.edu; mwa23@mail.aub.edu; sms86@mail.aub.edu; chs04@mail.aub.edu; gry01@mail.aub.edu

Abstract – **This work describes the preliminary design of a plant used by a Lebanese industrial company to produce diesel fuel using biomass formed by municipal waste and wastewater sludge as a main feedstock. It mainly investigates the application of the anaerobic digestion treatment method in the process, as a possible pathway toward fuel from organic waste. Once methane is produced from the digester, the gas to liquid (GTL) process is implemented to get diesel as a final product. This process is simulated using Aspen HYSYS and Aspen PLUS software packages and is found to produce 3100 kg/h of diesel along with 581 kg/h of naphtha by treating about 472 tons/day of organic municipal solid waste and sludge. A detailed economic analysis shows that the total capital cost of the plant is $72.3 million with annual net earnings of $4 million. Therefore, the plant needs 18 years to return on its capital investment.**

I. INTRODUCTION

Diesel fuel in general is any liquid fuel used in diesel engines, whose fuel ignition takes place, without spark, as a result of the compression of the inlet air mixture and then injection of fuel. Synthetic diesel fuels can be made from any carbon-containing feedstocks, such as natural gas, coal and/or biomass through the Fischer-Tropsch (FT) process. This type of fuels is mainly composed of straight-chain alkanes ($C_9 – C_{20}$). While FT fuels made from natural gas bring no discernible greenhouse gas benefit relative to petro-diesel, the FT fuel made from biomass can provide a lifecycle carbon dioxide emission benefit. Moreover, the cetane numbers[1] of the FT synthetic fuels can reach up to 75 as compared to those of petro-diesel which range between 45 and 50 [1] [2]. This indicates that FT synthetic fuels has a much better combustion efficiency than petro-diesel fuels.

The raw materials used in this process consist of organic municipal solid waste (OMSW) and sewage sludge, both from Lebanese sources. To obtain a rough estimate of these materials, data is collected from a waste plant in Saida and a report submitted by the Lebanese Environment Ministry in 2010. The Saida Waste Plant data shows that organic waste is at a rate of 124.5 t/d. This number is assumed to represent 10% of the total Lebanese organic waste equivalent to the population percentage in Saida and its neighborhood. As for sewage sludge, available numbers are limited to 2010 at 310 t/d [3], so it is estimated that this waste generation is proportional to the population growth from 2010 to 2014 where production would be equal to 326.6 t/d.

The common first step to the different pathways to diesel is to produce methane and/or synthetic gas (syngas) from the biomass. The chosen process begins by treating the waste using an anaerobic digester which produces methane in majority, along with some impurities. The methane produced is sent to a purification unit to remove the carbon dioxide and the hydrogen sulfide. Next, the methane is transformed into syngas – a gas mixture of carbon monoxide and hydrogen – using an autothermal reformer. The syngas is converted into a wax of hydrocarbons using the Fischer-Tropsch process. The diesel and wax fractions are separated from the other hydrocarbons and finally sent to a hydrocracker to upgrade the fuel quality. The flow diagram of this process, drawn on VISIO®, is attached at the end of this paper.

II. BACKGROUND

A. Diesel

The FT reaction can be conducted at either high (HTFT) or low temperature (LTFT). These processes happening under different conditions result in diesel fuels having different properties. Table 1 provides a comparison between the two diesel fuels, and shows that LTFT fuels tend to have a lower density than HTFT-based ones because they contain less aromatics. Also, LTFT diesels have high cetane numbers (75-87) while HTFT diesels are more aromatic with lower cetane numbers (45-52) [2] [4]. In countries where the density of fuel is regulated, it might be difficult to market LTFT diesel as diesel fuel because of its low density. In Lebanon, such regulations are either weak or missing and are assumed to be inexistent.

---

[1] An important indicator of a fuel's combustion efficiency, showing the ignition properties of diesel fuel relative to cetane as a standard.

Properties	HTFT	LTFT
Density at 288 K (kg/L)	0.809	0.769
Sulfur (mass %)	<1	<5
Aromatics (mass %)	≈25	<1
Flash point (K)	351	331
Viscosity at 313 K (m²/s)	2.14×10^{-6}	2×10^{-6}

B. Different Options for the Process

Figure 1 summarizes the two investigated general processes to get to diesel from biomass using anaerobic digestion. Route (C) is the common process used in the industry to produce diesel fuel from natural gas. Route (B) has only been implemented on a small scale; and as the project seeks to produce diesel on an industrial scale, combining routes (A) and (C) seems more attractive. Moreover, the methanol produced in route (B) would not have been enough to produce diesel fuel in large quantities in the transesterification phase, and thus make-up methanol would have been needed – which makes this process less appealing. Therefore, the chosen process follows route (C) after producing the syngas using route (A).

III. METHODOLOGY

The process is designed and simulated using data and information found in the literature concerning each process step. The simulation is conducted using Aspen PLUS and Aspen HYSYS software packages developed by Aspen Tech®.

IV. PROCESS STEPS

A. Anaerobic Digestion and Biogas Purification

Rajendran et al. [5] proposed a process simulation model[2] for the anaerobic digestion (AD) of biomass using Aspen Plus. This model accounts for the composition of the biomass, from carbohydrates to lipids and proteins, and the inhibitions such as pH, ammonia, temperature and others. These inhibitions require FORTRAN or MATLAB calculators to be included in the Aspen Plus model. Since this is out of the scope of the project, these inhibitions are ignored and the reactions are input as the simplified first order kinetics rates as given by the authors.

The AD system is modeled on Aspen Plus in two different reactors representing the four major steps in an anaerobic digestion as follows:

- Hydrolysis step: modeled as a fractional conversion reaction in a stoichiometric reactor where the macromolecules are digested, in the presence of water, to micromolecules

- Acidogenesis, Acetogenesis and Methanogenesis steps: Modeled as three sets of kinetic reactions in a constantly-stirred tank reactor (CSTR)

When run, the model specifies a required volume of 13,000 m³ for the minimum necessary residence time of 8 days. The volume is considerably high and thus it was divided to ten reactors each with a volume of 1500 m³. The typical reactor is a big chamber 6.7 m high, 6.7 m wide, and 33.7 m long with concrete walls. The biomass is fed into the reactor chamber with a percolate spraying over it which is then drained and recirculated again. In order to control the temperature at 55°C, four heating pipes inside the chamber are advised.

The biogas resulting from this anaerobic digestion has a flow rate of 149 t/d and composed of 54% methane, 43% CO_2, 1% H_2S and 2% as other minorities. To help purify the biogas before entering the autothermal reactor, impurities are removed by amine absorption.

B. Autothermal Reformer and Air Separation Unit

Reforming of methane is undertaken to produce syngas suitable for the FT process. This part usually represents the bulk in terms of energy usage and total cost for Gas-to-Liquids (GTL) processes. Syngas is composed mainly of H_2 and CO, with the presence of CO_2 and CH_4 in low quantities. Autothermal Reforming (ATR) was chosen as the reforming method, as it can produce a syngas with a H_2/CO ratio close to 2.15 – a value proven to be ideal for FT synthesis [6]. ATR is a stand-alone process in which methane conversion is achieved in one reactor. This reactor consists first of a partial oxidation section where the methane-steam mixture is partially burned with pure oxygen which provides the heat required for the main endothermic
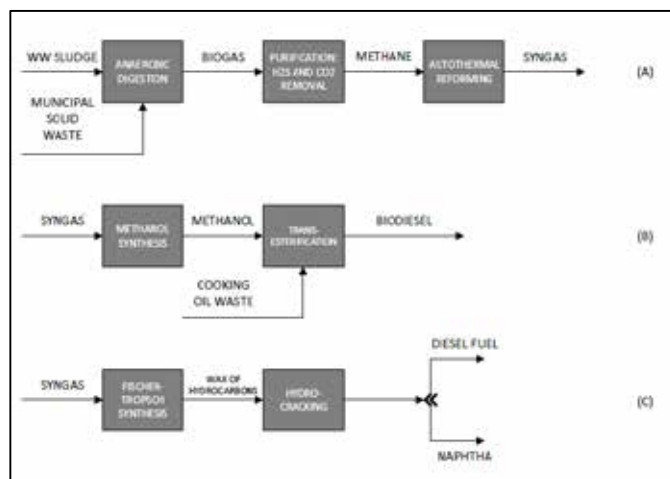

Figure 1 - Block-flow diagrams of the two main routes to diesel fuel.

---

[2] Model can be found online: http://hdl.handle.net/2320/12358

reforming reactions to occur in the subsequent catalyst bed section. The reactions that take place are listed in Table 2.

Many modelling attempts of the ATR reactor are reported in the literature. In most cases, it is deemed that reactions (5)-(9) have no significant rates and are ignored; as some of them are also nothing but combinations of the others. Successful models have been reported in the literature when approximating the entire reaction scheme with reactions (1)-(4) [7] [8]. Thus, in this work only reactions (1)-(4) are modeled.

Kinetics for such reactions are widely available. Xu and Froment (1998) established kinetics for the steam reforming reactions over a Ni-based catalyst that are used widely in the modeling of steam reforming synthesis. The kinetics of Trimm and Lam (1980) are considered to be the most rigorous in their description of methane combustion over Ni catalysts [9].

The most used catalyst in the ATR process is a Ni-based catalyst stabilized on an alumina carrier (Ni/Al_2O_3) [8]. Several companies produce such catalysts with a shape optimized for minimal pressure drop.

The process is simulated on Aspen HYSYS, whereby a syngas with an H_2/CO ratio of about 2.18, which is close to the ideal ratio, is produced. CO_2 constituted an acceptable fraction of 5%, while the water produced is eventually removed using a flash separator and recycled to the inlet of the ATR reactor in order to be mixed with the original methane stream.

ATR synthesis requires pure oxygen to be fed to the reactor. Air is usually 79% nitrogen by mole and providing air instead of pure oxygen would imply the presence of large amounts of nitrogen in the syngas product stream. Although nitrogen is an inert that can act as a heat sink, its presence is generally undesirable in the industry as it lowers the reactivity of syngas and imposes undue restrictions in the optimization of general process schemes. Pure oxygen is not usually available in Lebanon and its importation and transportation costs could be significant. Therefore, it is decided that an air separation unit (ASU) is to be run alongside the main process to provide pure oxygen. ASU is a relatively mature technology and many companies provide licenses for ASU plants. The Linde® process is the one used to recover oxygen with 99% purity. Although the process is relatively simple, it is energy extensive as it involves severe cooling and cryogenic distillation.

A design of the ATR reactor gives a volume of 50 m^3 with a cylindrical shell height and diameter of 5 m and 3.6 m respectively. The conical head and bottom angles are 45º and 60º respectively.

C. Fischer-Tropsch (FT) Reactors

The FT process transforms the syngas into hydrocarbons. This synthesis route has a large number of byproducts and intermediates including alkyls and alkenyls. It is highly exothermic and thus requires constant heat removal. Also, the set temperature, reactor type and H_2 to CO ratio of the syngas are important parameters to consider in order to be able to increase specific selectivity. The product required in this case is diesel (C_9-C_{20}).

There are two types of FT synthesis: Low Temperature Fischer Tropsch (LTFT) and High Temperature Fischer Tropsch (HTFT). Fe-catalyzed HTFT tends to give more light hydrocarbons (C_1-C_{10}) and aromatics (more than 70 wt%) while Fe- and Co-catalyzed LTFT tend to give more paraffins as distillates and waxes (more than 65 wt% for the Fe-catalyzed and more than 75 wt% for the Co-catalyzed). The most appropriate type for diesel production is therefore the LTFT [10].

The Co-catalyzed LTFT has a higher selectivity for diesel and waxes than Fe-catalyzed LTFT. Although cobalt is more expensive than iron, it is less prone to degradation. Co catalysts work at temperatures in the range of 225°C, while Fe catalysts require temperatures in the range of 340°C, thus making Co catalysts more suitable to LTFT synthesis. Finally, Fe-catalyzed LTFT has an activity towards a water gas shift (WGS) side reaction (reaction (4) in Table 2) that leads to inhibition of the synthesis whereas Co has a neglected activity towards this reaction [10].

Three types of reactors are possible: fluidized bed, fixed bed, and slurry reactors. Fluidized bed reactors cannot be used for this process because there would be a liquid phase production that can cause the agglomeration of the catalyst and by that the loss of fluidization. Slurry reactors are preferred over fixed bed reactors because temperature would be easier to control. Also, the catalyst can be fed and suspended inside the reactor which makes its regeneration process more practical. Moreover, the slurry reactors cost less than fixed bed ones [11].

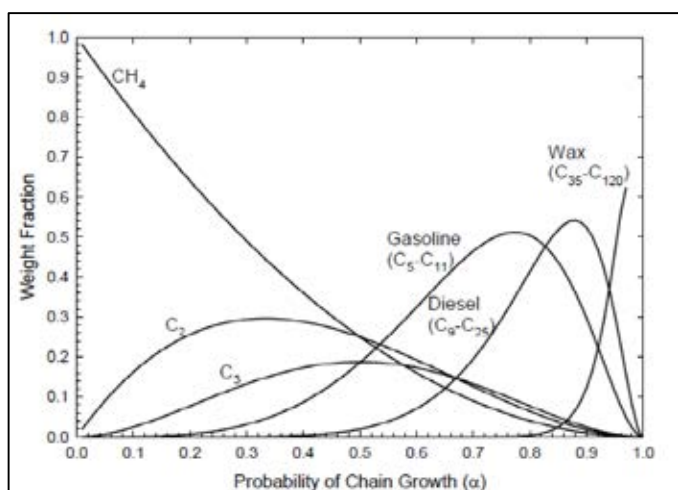To simulate the LTFT synthesis on ASPEN HYSYS, the following overall typical FT reaction (10) is used:

$$CO + 2.15H_2 \rightarrow Hydrocarbons + H_2O \qquad (10)$$

Table 2 - ATR reaction scheme.

Exothermic Combustion	$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$	(1)
Endothermic steam reforming reactions	$CH_4 + H_2O \Leftrightarrow CO + 3H_2$	(2)
	$CH_4 + 2H_2O \Leftrightarrow CO_2 + 4H_2$	(3)
Water-gas shift reaction	$CO + H_2O \Leftrightarrow CO_2 + H_2$	(4)
CO_2-reforming reaction	$CH_4 + 2CO_2 \Leftrightarrow 2CO + 2H_2$	(5)
Carbon deposition	$2CO \Leftrightarrow C + CO_2$	(6)
Methane cracking	$2CH_4 \Leftrightarrow C + 2H_2$	(7)
Carbon gasification	$C + H_2O \Leftrightarrow CO + H_2$	(8)
	$C + O_2 \Leftrightarrow CO_2$	(9)

Figure 2 - Weight factor as a function of the probability of chain growth

In parallel, the WGS reaction (4) can also occur. However, under normal FT operating conditions, cobalt-based catalysts have a low WGS activity and therefore only reaction (10) occurs. Reaction (10) shows that to ensure high conversions under Co-catalyzed FT, the ratio of H_2 to CO should be about 2.15 [2]. The overall rate of the reaction was obtained by Yates and Satterfield [12]. To implement the FT reaction in Aspen HYSYS, the Anderson-Schulz-Flory (ASF) distribution [11] is used:

$$W_n = n(1-\alpha)^2 \alpha^{n-1} \qquad (11)$$

Where W_n is the weight fraction of the hydrocarbon of n carbons, α is the chain growth probability and n the number of C atoms in a given hydrocarbon. When plotting the hydrocarbon weight distribution W_n as a function of the probability of chain growth α, Figure 2 is obtained.

Recent studies report a Co-catalyzed FT synthesis with an α between 0.85 and 0.95 [13]. An α value of around 0.9 should be assumed because at this chain growth probability both diesel and waxes are formed in large quantities.

Research has shown that Co-catalyzed LTFT is more likely to form paraffins without olefins. As a matter of fact, SASOL® reports a weight fraction of paraffins more than 90% in the product stream coming out of their Co-based slurry FT reactor [14]. Also, experiments showed that a ratio of H_2/CO equal to two or higher is responsible of reducing the amount of olefin products to approximately zero [15]. Since in this case this ratio is about 2.18, olefins in the products are neglected when implementing the FT reaction.

The ASF equation (11) is used to obtain an Excel® spreadsheet which assumes a certain mass flow rate of CO considered as the base reactant and predicts the mass flow rates of all the reactor products. The assumed mass flow rate is 100 kg/h. The available n-alkanes in the Aspen HYSYS database are the first 30. Using equation (11) with an α of 0.9, the weight percentages of the 30 n-alkanes in the product stream are

predicted with the requirement that 80% of CO must be converted [16]. After summing the mass flow rates of the n-alkanes, a conversion of about 78.15% is found which is close to the commercially-obtained conversion. Once the mass flow rates are predicted, the molar flow rates of the n-alkanes are calculated and the stoichiometric coefficients of the reactants and the products are found by applying atom balances on the carbon, hydrogen and oxygen atoms.

To simulate the slurry phase reactor in Aspen HYSYS, the CSTR model is used. When simulating the FT reaction and applying it to the CSTR, two product streams are obtained; one vapor containing n-alkanes in the diesel range and below and one liquid containing all the wax products. Two CSTRs in series are implemented and they ensure an overall CO conversion of 95.2% and the product stream composition is given in Table 3. After designing the two slurry phase reactors, their volumes are found to be 24.63 m^3 and 9.354 m^3 with lengths of 8.5 m and 6.1 m along with diameters of 1.9 m and 1.4 m respectively.

The products are then separated using a series of vapor-liquid-liquid (VLL) separators, flash separators and distillation columns. At the end, 581 kg/h of naphtha are produced in this process. Water accounts for 5146 kg/h with 99.6% purity. The diesel and wax streams are then mixed and sent to be upgraded in the hydrocracker.

D. Hydrocracker

This process upgrades the waxes into a commercial-grade product: diesel and light distillates. It produces high quality distillates and isomers of diesel. Even the n-alkanes in the diesel range produced by FT reaction have to go through the isomerization phase and their cracking can be limited according to the temperature and pressure conditions in the hydrocracker along with the chosen catalyst. Having more isomers enhances the combustion properties of the diesel and hence its cetane number.

The catalysts of this process are generally metals with a hydrogenation or dehydrogenation function (i.e. bifunctional metals). The LTFT waxes contain negligible traces of sulfur; hence, the use of catalysts formed of noble metals such as platinum and palladium is appropriate [17]. Hydrocracking is generally done in packed bed catalytic reactors, also called trickle bed reactors. This type of reactors consists of a liquid phase and a vapor phase flowing in a countercurrent motion through a fixed bed of catalyst [18].

Table 3 - FT product stream composition.

Alkanes Range	Mass Percentage
Liquid Petroleum Gas (LPG) (C_1-C_4)	10.89
Naphtha (C_5-C_8)	15.05
Diesel (C_9-C_{20})	48.5
Wax (C_{21}-C_{30})	25.56
Wax and Diesel Mixture	74.07

In order to simulate the hydrocracker in Aspen HYSYS, the kinetics found in reference (Pelligrini, Gamba, Calemma, & Bonomi, 2008) are used. These kinetic parameters and expressions are derived using the Soave-Redlich-Kwong (SRK) fluid package and working on the overall hydrocracking reaction which is given as follows:

$$n - C \Leftrightarrow iso - C \rightarrow Cracking\ Products \qquad (12)$$

The Aspen HYSYS database does not contain the alkane isomers in the C_{11}-C_{29} range. Therefore, it is assumed that reaction (12) can be simplified to the following reaction in the simulation:

$$n - C \rightarrow Cracking\ Products \qquad (13)$$
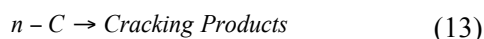
In the product stream, alkanes in the diesel range from C_9 to C_{20} are present along with the remaining H_2 that did not react. The final production rate of this process is about 3100 kg/h diesel including the n-alkanes and their isomers with 99.6% purity. The recovered H_2 is then recycled to the trickle bed reactor. In addition to diesel and naphtha, the process also produces liquefied petroleum gases (LPGs) with a rate of 500 kg/h.

V. HEAT INTEGRATION

Heat integration is implemented to make use of the streams that need to be cooled and this is by transferring their energy potential to the streams that need to be heated. As a result, six heat exchangers are needed in addition to one heating unit to supply the required extra heating of 1117 kW. The duty of the heat exchangers sums up to 4260 kW transferred from the hot streams to the cold ones. However, if no heat integration was implemented the amount of heat transferred should have been 5377 kW. This lead to a 79% saving in the total energy consumed by the process.

VI. PLANT LOCATION AND LOGISTICS

Several factors come into play when deciding on the location of an industrial plant. The decision could have great implications on costs associated with logistics, labor, maintenance, and licensing procedures. Seeing as most of the Lebanese terrain is mountainous; a property generally undesirable for a plant location as it presents transportation and climate-related complications, this lowers down the number of possible areas to two – the Bekaa Valley and the coast. These two areas present suitable geography for plant installation and material transportation. The coast represents the bulk in population density and thus the most availability of municipal waste which constitutes the main raw material for the plant. The presence of the mountains between Bekaa and the coast implies increased costs of transportation between the two areas. Thus, it seems that setting the site location on the coast would be the most beneficial as the main market and the source of raw materials exist in this area. The Zehrani area was finally chosen for the plant, and it is the easiest to access and is close to both the sources of raw materials and the main market for diesel fuel.

The location is also far enough from residential areas, and the costs of land acquisition for this area are reasonable.

Table 4 - The plant annual revenues.

Products	Production (kg/year)	Price ($/kg)	Revenue ($/year)
Diesel	24,448,000	0.679	$16,600,192
Naphtha	4,650,400	0.443	$2,060,127
LPG	4,000,000	0.422	$1,688,000
Nitrogen	4,950,000	0.45	$1,237,500
Raw Materials	**Usage (ton/year)**	**Price ($/ton)**	**Revenue ($/year)**
Waste	155,760	80	$2,227,500
Total Revenue	**$35,036,619**		

VII. ECONOMIC ANALYSIS

After calculating the preliminary design parameters of all the units present in the process, the total cost of equipment (abbreviated as E) including shipping is estimated to be 12.2 million dollars. This estimate helped in determining the total capital investment (TCI) of this chemical plant. The TCI required for a new design can be broken down into fixed capital investment (FCI) and working capital (WC), which can be approximated to be equal to 504% and 89% of E respectively [20]. These sum up to give a total capital investment of $72.3 million.

The operating costs of the plant are calculated on the assumed basis that the plant operates for 8,000 hours per year and are divided into direct (labor, utilities, maintenance…) and indirect (property taxes, insurance…) costs. They are found to be equal to $30.7 million per year. As for the revenues, they sum up to $35 million per year [13] and are summarized in Table 4 [21] [22].

The annual gross earnings (pre-tax) are calculated by taking the difference between the total annual revenue and the annual product cost (operating). For the plant at hand, the annual gross earnings are $4.3 million. In order to get the net annual profit, taxes should be accounted for. Seeing as this project is utilizing municipal waste and wastewater sludge, the Lebanese central bank offers the first ten years tax free and the following years 15% taxes. This results in $4 million as net earnings. The return on investment (ROI) is then calculated and is found to be equal to 0.055 which reflects that every year, 5.5% of the initial capital investment is compensated. Thus, in total, this plant needs approximately 18 years to return all of its capital investment.

Usually, a breakeven period larger than five years is unfavorable to investors and thus if looking only at the numbers, this project is infeasible. But, seeing as it is utilizing local waste and making a small profit on the side by selling fuel, it may be of interest to the government acting as its main investor. Therefore, this project can be considered feasible if its future investors favor sustainability over large profits.

VIII. Concluding Remarks

This work described a preliminary design of a process that transforms organic municipal solid waste and sludge into diesel fuel using anaerobic digestion and Fischer-Tropsch synthesis. About 3100 kg/h of diesel and 581 kg/h of naphtha are produced from 472 tons/day of waste, according to computer simulation data. Commonly, anaerobic digestion is used only as a waste treatment method with the main objective of producing organic fertilizers with the methane burned or sold. The work investigated the possibility of combining anaerobic digesters with a typical gas-to-liquid (GTL) plant, whereby the former would provide the necessary methane for the latter. Although the anaerobic digesters provided pure methane for the GTL process and in turn alleviated the need for secondary reforming – whereby the bulk of energy use is concentrated – the economics of the plant imply large capital investments with a large breakeven period of time. Thus, it would seem that the construction of such a plant would only be successful commercially if it was backed up by governmental support combined with a good market for diesel. A better way to produce fuels from waste would be the gasification process, whereby the gasification reactions allow for the elimination of the reforming section by directly transforming the waste into syngas. However, anaerobic digestion presents an advantage over gasification by the ability to produce high-quality organic fertilizers, which could be very useful for the Lebanese agriculture.

References

[1]   "Synthetic Diesel Fuel," 2002. [Online]. Available: https://www.dieselnet.com/tech/fuel_syn.php#f-t. [Accessed 19 October 2015].

[2]   M. E. Dry, "High quality diesel via the Fischer–Tropsch process: a review," *Journal of Chemical Technology and Biotechnology,* no. 77, pp. 43-50, 2001.

[3]   "Lebanon State of the Environment," Ministry of Environment/LEDO, Beirut, 2010.

[4]   P. M. Maitlis, Ed., Greener Fischer-Tropsch processes for fuels and feedstocks, Wiley-VCH, 2013, pp. 99-101.

[5]   K. Rajendran, H. Kankanala, M. Lundin and M. Taherzadeh, "A novel process simulation model (PSM) for anaerobic digestion using Aspen Plus," *Bioresource Technology,* pp. 7-13, 2014.

[6]   A. Steynberg and M. Dry, Fischer-Tropsch Technology, Amsterdam: Elsevier, 2004.

[7]   A. M. D. Groote and G. F. Froment, "Simulation of the catalytic partial oxidation of methane to synthesis gas," *Applied Catalysis,* no. 138, pp. 245-264, 1995.

[8]   V. Barrio, G. S. M. Rohde and S. Rabe, "Reactor modeling to simulate catalytic partial oxidation and steam reforming of methane. Comparison of temperature profiles and strategies for hot spot minimization," *Journal of Hydrogen Energy,* no. 32, pp. 1421-1428, 2006.

[9]   M. Halabi, M. d. Croon and J. v. d. Schaaf, "Modeling and analysis of autothermal reforming of methane to hydrogen in a fixed bed reformer," *Chemical Engineering Journal,* no. 137, pp. 568-578, 2007.

[10]   M. C. Valero and P. Raybaud, "Cobalt Catalyzed Fischer–Tropsch Synthesis: Perspectives Opened by First Principles Calculations," *Catalyst letters,* november 2012.

[11]   S. Saeidi, K. N. Maryam, A. Mirvakili, S. Bahrani, S. A. N. Aishah and R. R. Mohammad, "Recent advances in reactors for low-temperature Fischer-Tropsch synthesis: process intensification perspective," *Revue of chemical engineering,* vol. 3, no. 31, pp. 209-238, 2015.

[12]   I. C. Yates and C. N. Satterfield, "Intrinsic Kinetics of the Fischer Tropsch Synthesis on a Cobalt Catalyst," *Energy & Fuels,* pp. 168-173, 1991.

[13]   S. T. Sie, "Process Development and Scale-up: IVCase History of the developement of a Fischer Tropsch Synthesis Process," *Revue of Chemical Engineering,* pp. 109-157, 1998.

[14]   D. Leckel, "Upgrading of Fischer Tropsch: Products to Produce Diesel," Sasol Technology Research and Development, Munkerupgaard, 2010.

[15]   F. A. N. Fernandes, "Modeling and Product Grade Optimization of Fischer-Tropsch Synthesis in a Slurry Reactor," *Industrial & Engineering Chemistry Research,* no. 45, pp. 1047-1057, 2006.

[16]   R. Espinoza, A. Steynberg, B. Jager and A. Vosloo, "Low Temperature Fischer Tropsch Synthesis from a SASOL Perspective," *Applied Catalysis A,* pp. 13-26, 1999.

[17]   L. Pellegrini, S. Locatelli, S. Rasella, S. Bonomi and V. Calemma, "Modeling of Fischer Tropsch products hydrocracking," *Chemical Engineering Science,* pp. 4781-4787, 2004.

[18]   C. Bouchy, G. Hastoy, E. Martens and J. Guillon, "Fischer-Tropsch Waxes Upgrading via Hydrocracking and Selective Hydroisomerization," *Oil & Gas Science and Technology,* vol. 64, pp. 91-112, 2009.

[19]   L. A. Pelligrini, S. Gamba, V. Calemma and S. Bonomi, "Modelling of hydrocracking with vapour-liquid equilibrium," *Chemical Engineering Science,* pp. 4285-4291, 2008.

[20]   F. Azizi, "CHEN 470 Part 10 - Cost Analysis," American University of Beirut, Beirut, 2014.

[21]   OPIS, "OPIS Asia Naphtha and LPG Report," OPIS, 2015.

[22]   T. Economic, "Pump price for diesel fuel (US dollar per liter) in Middle East and North Africa," 6 12 2015. [Online]. Available: http://www.tradingeconomics.com/middle-east-and-north-africa/pump-price-for-diesel-fuel-us-dollar-per-liter-wb-data.html.

PRODUCTION OF DIESEL FUEL FROM BIOMASS AND WASTEWATER SLUDGE – PROCESS FLOW DIAGRAM

AREA 100: AD

AREA 200: ASU

AREA 300: ATR

AREA 400: FT

AREA 500: HC
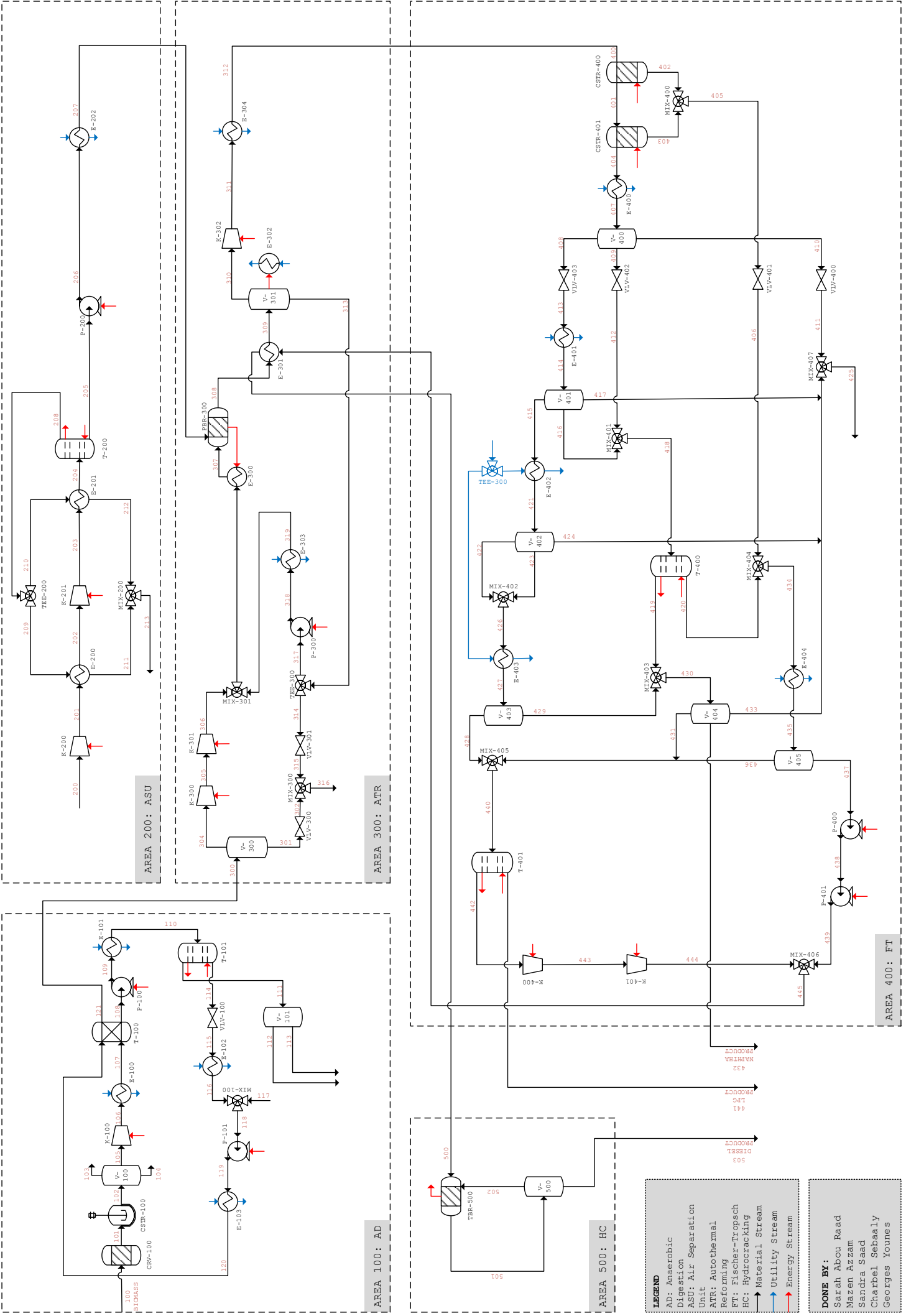
LEGEND
AD: Anaerobic Digestion
ASU: Air Separation Unit
ATR: Autothermal Reforming
FT: Fischer-Tropsch
HC: Hydrocracking
Material Stream
Utility Stream
Energy Stream

DONE BY:
Sarah Abou Raad
Mazen Azzam
Sandra Saad
Charbel Sebaaly
Georges Younes

Removal of pharmaceuticals from water using activated carbon

Reem Yahfoufi
Chemical Engineering department
American university of Beirut
Beirut, Lebanon
rty00@mail.aub.edu.lb

Abstract- **Pharmaceutically active compounds are considered to be unregulated trace level contaminants with potential chronic toxicity. Since the long term effects of these compounds on the human health and the environment are to date unknown, the removal of theses contaminants through conventional water and wastewater treatment is a crucial step. The large number of pharmaceuticals detected, their diversified properties, and their occurrence at trace-level have posed a unique challenge for the removal of these compounds from water. Several conventional and non-conventional "polishing" methods, such as filtration, advanced oxidation and membrane systems, have been used to reduce the concentrations of pharmaceuticals in water. Adsorption using activated carbon has proven to be a suitable treatment for the removal of trace-level organic compounds. The proposed research aims at assessing the effectiveness of activated carbon (derived from waste biomass) in removing trace pharmaceuticals from water. This will be achieved by conducting several bench scale adsorption experiments at different temperature and pH values, and for a number of common pharmaceuticals. The data generated will then be used to derive the appropriate kinetic and thermodynamic models.**

I. INTRODUCTION

A. Pharmaceutically active compounds

Pharmaceuticals have been identified as emerging contaminants in the aquatic environment. These organic compounds have been found to be present in surface and ground water bodies at low concentration levels (ng/L up to μg/L), hence their detection and quantification require the presence of highly complex analytical methods [1]. Some of these compounds are identified as endocrine disruptors even at low concentrations. In fact, the presence of pharmaceutically active compounds in water sources is of concern due to their adverse effects on human health upon consumption, such as antagonizing hormones, obstructing metabolic processes and causing problems in the reproductive system [2]. Some studies have investigated the possible effects on human health and aquatic organisms upon consumption of water containing low concentrations of pharmaceuticals. But these studies have not yet identified the risks associated with the persistent exposure to random combinations of these compounds and their toxicological significance remains, to date, an unanswered question. In addition, the increasing demand and ubiquitous consumption of pharmaceutically active compounds, has lead the level of exposure to increase, causing a larger threat to human health [2]. Consequently, their removal from drinking water and wastewater has

become crucial [3] [4]. However, the removal of pharmaceuticals from water is typically challenging for several reasons. First, a large number of compounds is detected due to relatively high consumption of pharmaceuticals. Secondly, the variety in their physical and chemical properties creates a unique challenge. Finally, these compounds are present at very low concentrations. However, activated carbon has proven to be suitable for the removal trace-level organic contaminants from water [5].

B. Activated Carbon Adsorption

Adsorption on activated carbon is generally used to remove natural and synthetic organic substances present in water and it is considered to be well suitable for the removal of low concentrations of these contaminants. Moreover, AC adsorption can eliminate a wide range of low to medium molecular mass compounds without generating by-products. Adsorption using AC is an effective process due to its high surface area and high pore volumes, in addition to the presence of varied functional groups [1].

The adsorption process has several advantages, such as: efficiency at low concentration of both organic and inorganic contaminants, possibility of regeneration and reuse, applicability both for continuous and batch processes. Due to its porous structure and surface chemistry properties, it is well suited to remove organic compounds. Moreover, activated carbon is a high binding adsorbent with relatively low cost, which let to it becoming a common adsorbent for the elimination of various organic contaminants in water [2].

Activated carbon can be produced at low capital cost from numerous carbon materials, such as wood [6], lignin, coal, coconut shells [7] [8], cork powder waste and polyethylene- terephthalate from used plastic bottles [8]. In fact, the latter carbonaceous material is an appealing and economic alternative for water treatment, waste disposal and recycling.

Wastewater treatment plants effluents are the main source of pharmaceuticals in surface waters. Consequently, several steps have been added to wastewater treatment to remove these compounds and avoid the contamination of the aquatic environment. In the wastewater treatment adsorption onto activated carbon is one of the most applicable technologies in terms of efficiency, ease in operation, low capital cost and energy requirements [3] [9]. The efficiency of activated carbon to re-

move organic compounds has been investigated in several water treatment plants and it has been shown to be dependent on the target compound and the frequency of carbon regeneration or replacement [10].

C. Removal of pharmaceuticals using activated carbon

In most of the studies that consider the adsorption of pharmaceuticals, a broad range of compounds were individually tested using different AC dosages in bench scale experiments [11] or in full-scale facilities [12] [8] [13]

Reference [13] studied the removal of β-blockers (propranolol and metoprolol), from hospital wastewater in a pilot scale treatment plant. The results showed that powdered AC could achieve high efficiency removal of the latter micro-pollutants (up to 100 % removal). A recent study was carried out on the adsorption of several pharmaceuticals (Carbamazepine, Ciproflocaxin, Citalopram, Metoprolol, Oxazepam, Sertralin) by GAC in water originating from sewage plants [9]. The authors indicated that despite the high solubility of these compounds and their dissimilar chemical structures, GAC provided high percentage removal of the investigated compounds (>90 %). Hence, activated carbon was shown to be a cost effective system and a competitive alternative to treatment with ozone. Reference [7] evaluated the adsorption of two pharmaceuticals, (Naproxen and Carbamazepine) onto two common activated carbon made from different material (Coal and coconut shell). The authors confirmed that both types of activated carbon were capable of almost completely removing both compounds (>90 % removal).

The purpose of this study is to comprehensively evaluate the potential of application of biomass waste-derived activated carbon for the removal of pharmaceuticals from water. The low cost high value carbonaceous adsorbents are prepared from olive stones and date pits. This strategy enables us to assess the potential possibilities for waste recycling by its application for the purification water with low contaminant concentration. Three pharmaceuticals from three different functional classes were chosen as target compounds: carbamazepine (anti-epileptic agent), ampicillin (antibiotic) and diltiazem hydrochloride (calcium channel blocker). The three targeted compounds are used worldwide and have been recognized as important organic pollutants increasingly found in wastewater from urban areas. Their persistence through wastewater treatment processes would lead to their continual introduction to the environment. Also, these compounds have noticeably different physiochemical properties in terms of log K_{ow}, pK_a and water solubility. Hence, they are considered representative of most pharmaceutically active compounds in terms of their adsorption characteristics.

II. MATERIAL AND METHODS

A. Material

The chemicals were used in this study as received and without further purification. Carbamazepine ampicillin, trimethoprim-13C3 and atrazine-13C3 were purchased from Fluka Analytical. Diltiazem hydrochloride was purchased from United States Pharmacopeia (USP). High performance liquid chromatography grade methanol was purchased from Fluka Analytical and Oasis hydrophilic-lipophilic (HLB) cartridges were purchased from Supelco. The two types of activated carbon were prepared from date pits and olive stones obtained from a local farm. Also activated charcoal obtained from Sigma was used in this study to highlight the adsorption potential of the activated carbon prepared in our laboratory. The contaminated water studied was prepared by spiking distilled water samples with a known concentration of pharmaceutical.

B. Preparation of adsorbent by H_3PO_4 activation

The carbon precursor was washed with distilled water repetitively and dried in an air oven at 60°C for 24 hours. The precursor was mixed with H_3PO_4 (85%) at an impregnation ratio of 3 to 1. The mixing was stirred for 24 hours using a magnetic paddle in order to ensure complete incorporation of H_3PO_4 into the precursor. After complete mixing, the impregnated material was dried overnight at 120 °C. The carbonization of the dried material was done at 550 °C under nitrogen flow (150 cm^3/ min). The carbonized product was then cooled at the room temperature and washed with distilled water to remove phosphate ions. Finally the washed product was dried at 110 °C for 24 hours.

C. Characterization of the adsorbent

Nitrogen adsorption desorption technique is performed using a Nova 2200e high speed surface area analyzer (Quantachrome instruments) to measure the specific surface area (SSA) according to the BET theory while the pore size and volume are calculated by the BJH method. The structural characterization of the adsorbents is carried out using a Thermo Nicolet 4700 Fourier Transform Infrared Spectrometer equipped with a Class 1 Laser; also, powder X-ray diffraction (XRD) spectra are recorded on a Bruker D8 Dicover X- ray Diffractometer. Scanning electron microscopy is carried out on a Tescan Scanning Electron Microscope at high voltage of 30 kV after coating the samples with a thin layer of gold.

D. Contaminated sample preparation

Stock solution of pharmaceutical was prepared by dissolving 2.5 ppm of standards pharmaceutical high-performance liquid chromatography grade methanol. Dissolved samples were stored at -20°C and used within months of preparation. The sample solutions for adsorption tests were prepared by adding the stock solution to ultrapure water to achieve the desired concentration.

E. Batch adsorption experiment

All adsorption experiments are performed in autoclavable bottles under shaking thermostatic water bath operating at 120 rpm in dark. An initial amount of adsorbent is added to the contaminated solution prepared in the laboratory, at set pH and temperature. When equilibrium is reached, samples are taken

out and filtered at using 0.45 μm filter. A filter control test was performed to assess potential losses to analytes during filtration. Recoveries close to 100% were observed. The amount of pharmaceutical compounds adsorbed at equilibrium, q_e (mg/g), is calculated by

$$q_e = \frac{(C_0 - C_e)V}{W}$$

where C_0 and C_e are the initial and equilibrium concentrations (mg/L) respectively, V the volume of the solution (L) and W the weight of activated carbon (mg).

The effect on adsorption of temperature and adsorbent dosage are studied. Also, the effect of the contaminated solution pH is investigated after adjustment using dilute HCl and NaOH solutions and a portable pH meter (LoviBond SensoDirect pH110).

F. Analytical technique

Samples are extracted using a solid phase extraction (SPE) with Oasis hydrophilic-lipophilic (HLB) cartridges, according to the EPA method 1694 for the analysis of pharmaceuticals and personal care products in environmental samples. Namely, the cartridge is conditioned using methanol, distilled water and acid water; then the sample is spiked with the surrogate standard trimethoprim-13C3. Later, the spiked sample is loaded onto the cartridge at a rate of 5-10 mL/min using a multi-position extraction manifold. Analytes are then eluted with 13 mL of methanol. The concentration of the extract is done to near dryness under a gentle stream of nitrogen in a water bath held at 50 ± 5 °C followed by the addition of the internal standard atrazine-13C3 and 3 mL of methanol.

Pharmaceuticals concentration will be determined using liquid chromatography with tandem mass spectrometry (LC/MS/MS). Previously determined chromatographic gradients and LC conditions will be set according to the polarity of the different pharmaceuticals as defined in the standard EPA method 1694 for pharmaceuticals and personal care products analysis in environmental samples with high performance liquid chromatography combined with tandem mass spectrometry (HPLC/MS/MS).

III. RESULTS AND DISCUSSION

A. Adsorbent characterization

Scanning electron microscopy (SEM) images reveal the porous structure of the adsorbent precursor. SEM images obtained are shown in figure 1. The examination of raw date stones shows interconnected porosity and formation of aggregates. After chemical activation, the porosity is more developed and elliptical macropores are formed. The images indicate that the chemical activation of date pits creates a regular macroporosity and a relatively homogeneous surface. Physical sorption of pharmaceuticals is more likely to occur onto highly porous activated carbon. Hence, the adsorption capacity of activated date pits is expected to be higher than that of natural date pits. Accordingly, preparation of highly porous activated carbon from low cost agricultural bio- mass presents an economically and environmentally viable application for water purification.
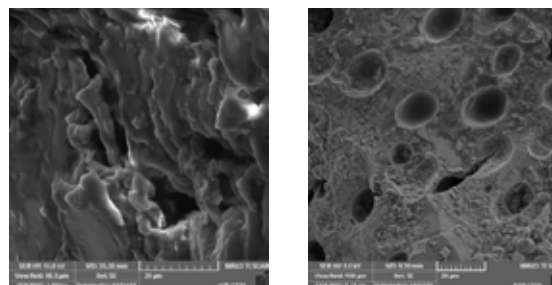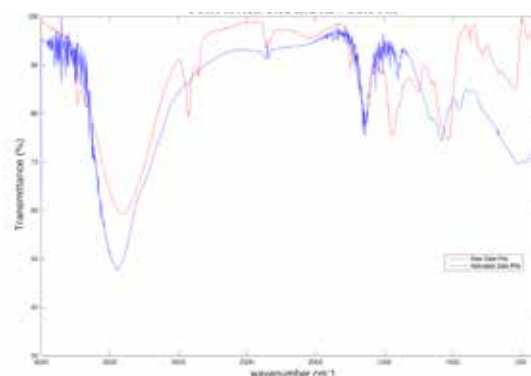

Figure 1. SEM image of raw and activated date pits.

Figure 2. FT-IR image of raw and activated date pits.

The FT-IR spectrum of the adsorbent precursor shows the presence of different functional groups namely alkene, aromatic, alcohol, ester, ether and hydroxyl. The sample shows IR bands at 3400- 3600 cm^{-1} which are attributed to -OH stretching of hydroxyl groups. The sample also shows peaks at 2926 at 2870 which correspond to C-H stretching in alkyl groups. The bands appearing at 1400-1600 cm^{-1} are ascribed to C=C stretching in aromatic rings, at the ones at 1000-1200 cm^{-1} are attributed to C-O stretching in ether groups. Finally, the bands appearing at 1600-1700 cm^{-1} correspond to the carbonyl group. IR spectrum of the activated carbon exhibits the same shape, which means that both materials have the same functional groups. The peaks observed at 2926 and 2870 cm^{-1} in the raw precursor disappear once the material is impregnated with H_3PO_4. Also, the bands between 500 and 1000 cm^{-1} attributed to aromatic C-H bending and O-H group stretching show a higher intensity. Hence, the aromaticity of the material increases during activation.

B. Adsorption of carbamazepine

A preliminary adsorption experiment was conducted in order to test the adsorptive capacity of the chemically activated adsorbent. The effectiveness adsorption effectiveness of the prepared activated carbon was compared with the commercially available activated carbon (activated charcoal purchased from Sigma). The contaminated sample was spiked with an initial carbamazepine concentration of 100 μg/L, and the adsorbent was added. The prepared and the commercial activated carbon achieved 99.6 % and 99.8 % removal, respectively. These results reveal that the adsorbent prepared from date pits can be a promising alternative to the highly expensive activated carbon commercially available.

C. Adsorption isotherm

Collected adsorption data at various adsorbent dosage ranging between 1 to 50 mg/L are fitted with two most commonly used adsorption models, namely Langmuir and Freundlich [14].

The Langmuir isotherm equation is represented by:

$$q_e = \frac{q_{max} K_L C_e}{1 + K_L C_e} \qquad (1)$$

where q_{max} is the monolayer adsorption capacity of the adsorbent (mg/g) and K_L is the Langmuir equilibrium constant (L/mg).

The Freundlich isotherm is represented by the equation:

$$q_e = K_F C_e^{1/n} \qquad (2)$$

where n is the heterogeneity factor and K_F is the Freundlich constant (L/g).

D. Adsorption kinetics

The adsorption kinetics are studied at 25 °C. The adsorption kinetics are monitored over 96 hours and the experimental results are fitted to four kinetic models, first, second, pseudo-first, and pseudo- second orders [15].

The first order equation is:

$$C_t = C_i e^{-Kt} \qquad (3)$$

where C_t is pharmaceutical concentration (μg/L) at time t (min), C_i the initial pharmaceutical concentration ((μg/L) and k_1 the first order rate constant (min^{-1}).

The second order equation is:

$$\frac{1}{C_t} = \frac{1}{C_i} + K_2 t \qquad (4)$$

with k_2 being the second order rate constant (L mg^{-1} min^{-1}).

The pseudo-first order equation is expressed as:

$$q_t = q_e \left(1 - e^{-k_1' t}\right) \qquad (5)$$

where q_e is the adsorption capacity of the prepared activated carbon at equilibrium (mg g^{-1}), q_t is the amount of pharmaceutical adsorbed (μg/g) at time t (min) and k_1' is the pseudo-first order rate constant (min^{-1}).

The pseudo-second order equation can expressed as:

$$q_t = \frac{K_2' q_e^2 t}{1 + K_2' q_e t} \qquad (6)$$

where k_2' is the pseudo-second order rate constant (g mg^{-1} min^{-1}).

IV. CONCLUSION

Pharmaceutically active compounds are serious emerging contaminant which are continuously introduced to surface water at low concentrations due to their persistence through municipal water recycling processes. Accordingly, in order to minimize environmental and human exposure, advanced treatment is essential to provide further reduction of these emerging contaminants. The main goal of this study is to examine the effectiveness of adsorptive treatment with activated carbon as an advanced treatment for the removal of pharmaceuticals from water. The activated carbon tested is biomass-derived activated carbon prepared from olive stones and date pits. The low-cost high-value adsorbent plays an important role on the economic feasibility of the overall adsorption process and the development of a sustainable process for the removal of pharmaceuticals waste. Also, the utilization of biomass waste as precursors is a suitable strategy to deal with the problem of waste disposal and recycling.

REFERENCES

[1] Delgado, L. F., Charles, P., Glucina, K., & Morlay, C. (2014). Adsorption of ibuprofen and atenolol at trace concentration on activated carbon. *Separation Science and Technology*

[2] Jung, C., Park, J., Lim, K. H., Park, S., Heo, J., Her, N., . . . Yoon, Y. (2013). Adsorption of selected endocrine disrupting compounds and pharmaceuticals on activated biochars. *J Hazard Mater, 263 Pt 2*, 702-710.

[3]   Reungoat, J., Escher, B. I., Macova, M., & Keller, J. (2011). Biofiltration of wastewater treatment plant effluent: effective removal of pharmaceuticals and personal care products and reduction of toxicity. *Water Res, 45*(9), 2751-2762.

[4]   Yoon, Y., Westerhoff, P., Synder, S., Esparza, M. (2003). HPLC-Fluorescence Detection and Adsorption of bisphenol A, 17B-estradiol, and 17a-ethynyl estradiol on powdered activated carbon. *Water Research,* 37, 3530–3537.

[5]   Delgado, L. F., Charles, P., Glucina, K., & Morlay, C. (2012). The removal of endocrine disrupting compounds, pharmaceutically activated compounds and cyanobacterial toxins during drinking water preparation using activated carbon--a review. *Sci Total Environ, 435-436*, 509-525.

[6]   Mestre, A. S., Pires, J., Nogueira, J. M. F., & Carvalho, A. P. (2007). Activated carbons for the adsorption of ibuprofen. *Carbon, 45*(10), 1979-1988.

[7]   Yu, Z., Peldszus, S., & Huck, P. M. (2008). Adsorption characteristics of selected pharmaceuticals and an endocrine disrupting compound-Naproxen, carbamazepine and nonylphenol-on activated carbon. *Water Res, 42*(12), 2873-2882.

[8]   Rossner, A., Snyder, S. A., & Knappe, D. R. (2009). Removal of emerging contaminants of concern by alternative adsorbents. *Water Res, 43*(15)

[9]   Ek, M., Baresel, C., Magner, J., Bergstrom, R., & Harding, M. (2014). Activated carbon for the removal of pharmaceutical residues from treated wastewater. *Water Sci Technol, 69*(11), 2372-2380.

[10]   Margot, J., Kienle, C., Magnet, A., Weil, M., Rossi, L., Felipe de Alencastro, L., Abegglen, C., Thonney, D., Chevre, N., Scharer, M., Barry, D.A. (2013). Tretament of micropollutants in municipal wastewater: Ozone or powdered activated carbon. *Science of the total Environment, 461-461,* 480-498.

[11]   Mestre, A. S., Pires, R. A., Aroso, I., Fernandes, E. M., Pinto, M. L., Reis, R. L., . . . Carvalho, A. P. (2014). Activated carbons prepared from industrial pre-treated cork: Sustainable adsorbents for pharmaceutical compounds removal. *Chemical Engineering Journal, 253*, 408-417.

[12]   Kim, S. D., Cho, J., Kim, I. S., Vanderford, B. J., & Snyder, S. A. (2007). Occurrence and removal of pharmaceuticals and endocrine disruptors in South Korean surface, drinking, and waste waters. *Water Res, 41*(5), 1013-1021.

[13]   Kovalova, L., Siegrist, H., von Gunten, U., Eugster, J., Hagenbuch, M., Wittmer, A., et al. (2013). Elimination of Micropollutants During Post-Treatment of Hospital Wastewater with Powdered Activated Carbon, Ozone, and Uv. *Environ Sci Technol* 47.14, 7899-908.

[14]   Duong, D.D. (1998). Adsorption analysis: equilibria and kinetics, *Imperial College Press, London*.

[15]   Qiu, H., Lv, L., Pan, B.-c., Zhang, Q.-j., Zhang, W.-m., & Zhang, Q.-x. (2009). Critical review in adsorption kinetic models. *Journal of Zhejiang University SCIENCE A, 10*(5), 716-724.

Syngas Production from Biomass Gasification for Methanol and DME Synthesis

Ali Albawi, Hamzy Ismail, Lara Abdallah, Omar Daouk, Sary Fayyad

Department of Chemical and Petroleum Engirneering

American University of Beirut

Beirut, Lebanon

aaa141@mail.aub.edu, hys05@mail.aub.edu, laa53@mail.aub.edu, omd04@mail.aub.edu

Abstract- **A waste-to-energy bio-refinery was designed in order to treat the municipal solid waste surplus in Lebanon and its negative repercussions. The aim is to treat municipal solid waste by gasification to produce syngas and proceed further to produce methanol and Dimethyl Ether (DME). The operating plant will receive about 20% of the total municipal solid waste in Lebanon which is 720 tons/day. The syngas produced from gasification will then be cleaned and treated, and sent to methanol production in three plug flow reactors (PBRs). The produced methanol will further used to produce DME in another PBR. The conversion of methanol in the reactor is 97.99% yielding 19 ton/hr of DME. The whole design of the plant was simulated using Aspen Plus and Aspen HYSYS.** *The income of the plant is generated by processing the feed at a price of 150$/ton and selling DME for 850$/ton.* **The yearly net earnings are 54,633,956$ per year giving a modest return on investment of 18%.**

I. INTRODUCTION

The new environmental concerns of the emissions of the greenhouse gases have turned the interest into environmental friendly energy sources such as solar energy, wind energy, and biomass energy. Thus, the industry is trying to make the best out of the organic materials, considered as waste, by providing new processes to convert them into valuable chemicals and energy. Currently, biomass, which can vary from food leftovers to solid waste and forests waste, is one of the renewable energy methods used. Biomass can be converted to transportation fuel, gaseous fuel or oil to generate electricity, plus it can be converted into useful chemicals. Locally, in addition to the environmental impacts caused by the use of fossil fuels, we can identify another problem which is the lack of proper management of waste.

Solid waste is spreading all over Lebanon and is becoming a striking problem that poses a serious risk on the future of the country. Therefore, an urgent plan needs to be developed in order to treat the waste. In this project, a solution for treating the bio-waste and providing Lebanon with the much needed bio-fuel to provide electricity is presented. Municipal solid

waste will be treated by gasification to produce syngas and proceed further to produce methanol and Dimethyl Ether (DME). Both, Methanol and DME, are used as intermediates to manufacture other chemicals, however, Dimethyl Ether can be used as a clean fuel. Due to its common characteristics with gasoline and diesel, their production and emerging use as an alternative fuel have been increasing.

II. LITERATURE REVIEW

A. Feed

Lebanon generates 22.5 million tons of waste annually in which each person is expected to generate 1.18 kg/day of municipal solid waste in urban areas. Figure 1 shows the composition of the solid waste in Lebanon.
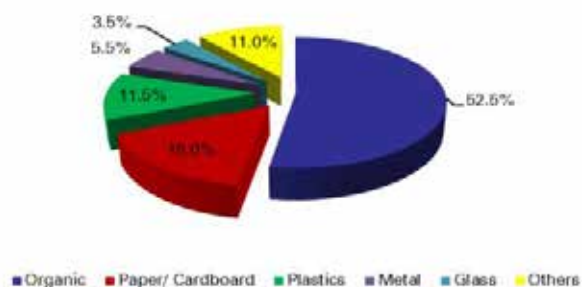

Figure 1 Composition of Solid Waste in Lebanon

The organic waste in Lebanon encompasses a big part of the generated waste due to the nature of the Lebanese cuisine. [5] This also contributes to the high moisture percentage of the waste reaching more than 60%. After treating the waste feed, a proximate and ultimate analysis is made as shown in table 1. A proximate analysis is physical analysis, while ultimate analysis is the chemical analysis of the feed.

TABLE 1 PROXIMATE AND ULTIMATE ANALYSIS OF FEED AFTER DRYING

Proximate Analysis		Ultimate Analysis	
Moisture	20 (wt %)	C (Carbon)	50.54 (wt %)
Volatile	63.7 (wt %)	H (Hydrogen)	7.08 (wt %)
Fixed Carbon	15.78 (wt %)	O (Oxygen)	41.11 (wt %)
Ash	0.505 (wt %)	N (Nitrogen)	0.15 (wt %)
		S (Sulfur)	0.57 (wt %)
		Ash	0.55 (wt %)

This plant will be operating at an industrial scale with a MSW feed of 30 ton/hr. this huge amount of feed enables us to classify this plant in the mid to high end industrial scale category.

B. Synthesis Gas

Synthesis gas is mainly known as syngas can be produced from a variety of different materials containing carbon. Its composition varies significantly depending on the feedstock and the gasification process involved with typical ranges of 30 to 60% carbon monoxide (CO), 25 to 30% hydrogen (H_2), 0 to 5% methane (CH_4), and 5 to 15% carbon dioxide (CO_2). Gasification is used to produce syngas which will be an intermediate product in the process of DME production.

C. Methanol

Methanol is the simplest alcohol with a chemical formula of CH_3OH. It is a light, volatile, colorless and flammable liquid with an alcoholic odor similar to but slightly sweeter than ethanol. Methanol will be produced by the catalytic synthesis from the syngas and will be also an intermediate product.

D. Dimethyl Ether

Dimethyl ether (DME) is a colorless, clean-burning, non-toxic gas with a chemical formula of CH_3OCH_3, simplified to C_2H_6O. DME is the final desired product of this process. It is used as an intermediate in the synthesis of important chemicals (such as dimethyl sulfate, lower olefins like propylene or higher value oxygenated compounds), as an electrolyte in high energy density batteries and as a solvent. Also, DME has been used gradually more as an aerosol propellant because of its environmental friendly properties. It will be produced by the dehydration of methanol.
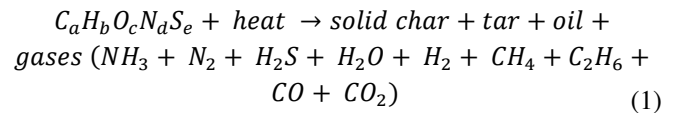
III. PROCESS BACKGROUND

A. Biomass Pre-Treatment

Before entering the gasifier, the feed is pretreated so that the gasifier works optimally and efficiently. Sorting system converts complex and diverse municipal waste using sorting streams into high quality and separated biomass feed. Then, mechanical treatment must be applied, so the waste gets shredded using an industrial shredder in order to homogenize the waste. Finally, the biomass waste is heated in a Direct-Fired Rotary Dryer in order to improve the combustion energy in the gasifier and reduce air emissions

B. Gasification

Gasification is the main step in the process in which the resulting producer gas consists of hydrogen, carbon monoxide, carbon dioxide, methane, water vapor, nitrogen, tar and dust. The producer gas should be cleaned form the impurities in order to obtain the syngas which consists of mainly H_2 and CO. In the bubbling fluidized bed gasifier used in this process (figure 2), the carbonaceous material undergoes several zones. The drying zone is the first, it requires substantial amounts of energy which can be provided by partial combustion. It is used to reduce the high moisture biomass feedstock to less than 5% in a temperature range of 100-200°C. Next, the pyrolysis zone breaks down the large molecules of biomass into smaller ones such as hydrogen, carbon monoxide, carbon dioxide, methane, ethane and ethylene. The resulting molecules can polymerize into bio-oil and tar. Pyrolysis occurs in the temperature range of 200-700°C in the absences of an oxidation agent. The overall reaction of pyrolysis is shown below:

$$C_aH_bO_cN_dS_e + heat \rightarrow solid\ char + tar + oil + gases\ (NH_3 + N_2 + H_2S + H_2O + H_2 + CH_4 + C_2H_6 + CO + CO_2) \quad (1)$$

In the partial combustion zone, the process occurs as the volatile products and some of the char react with the gasifying agent (air, oxygen or steam) to primarily form CO_2 and small amounts of CO. This provides heat for subsequent gasification reactions in temperature range of 1200°C to 1500°C. Letting **C** represent a carbon-containing organic compound,
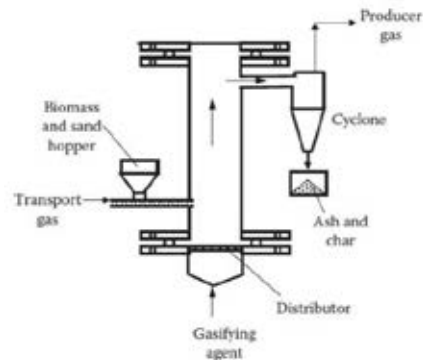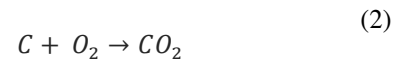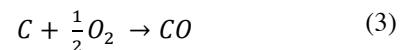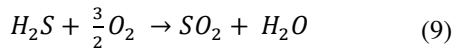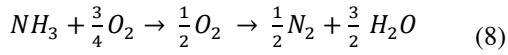

Figure 2 Bubbling Fluidized Bed

the basic reaction here is:

$$C + O_2 \rightarrow CO_2 \quad (2)$$
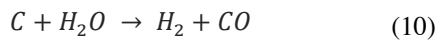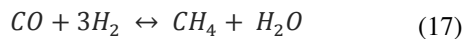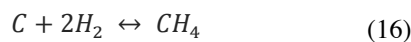
Below, some of the important reactions listed are:

$$C + \frac{1}{2}O_2 \rightarrow CO \quad (3)$$

$$C + O_2 \rightarrow CO_2 \tag{4}$$

$$CO + \frac{1}{2}O_2 \rightarrow CO_2 \tag{5}$$

$$H_2 + \frac{1}{2}O_2 \rightarrow H_2O \tag{6}$$

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O \tag{7}$$

$$NH_3 + \frac{3}{4}O_2 \rightarrow \frac{1}{2}O_2 \rightarrow \frac{1}{2}N_2 + \frac{3}{2}H_2O \tag{8}$$
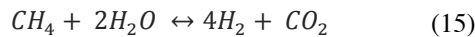
$$H_2S + \frac{3}{2}O_2 \rightarrow SO_2 + H_2O \tag{9}$$
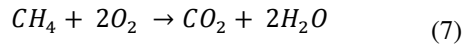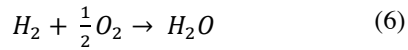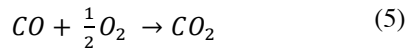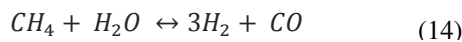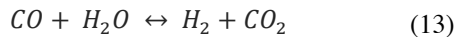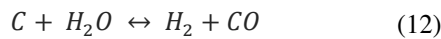
In the reduction zone, the process occurs as the char reacts with steam to produce carbon monoxide and hydrogen, via the reaction:

$$C + H_2O \rightarrow H_2 + CO \tag{10}$$

Basically, a limited amount of oxygen or air is introduced to the reactor to allow some of the organic material to burn producing carbon dioxide and energy, which drives a second reaction that converts further organic material to hydrogen and additional carbon dioxide. Further reactions occur when the formed carbon monoxide and residual water from the organic material react to form methane and excess carbon dioxide.
The main reactions are shown below:

$$C + CO_2 \leftrightarrow 2CO \tag{11}$$

$$C + H_2O \leftrightarrow H_2 + CO \tag{12}$$

$$CO + H_2O \leftrightarrow H_2 + CO_2 \tag{13}$$

$$CH_4 + H_2O \leftrightarrow 3H_2 + CO \tag{14}$$

$$CH_4 + 2H_2O \leftrightarrow 4H_2 + CO_2 \tag{15}$$

$$C + 2H_2 \leftrightarrow CH_4 \tag{16}$$

$$CO + 3H_2 \leftrightarrow CH_4 + H_2O \tag{17}$$

$$NH_3 \leftrightarrow \frac{1}{2}N_2 + \frac{3}{2}H_2 \tag{18}$$

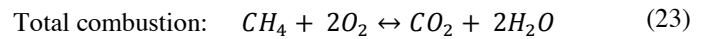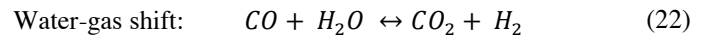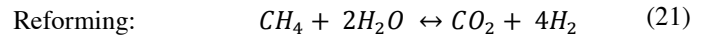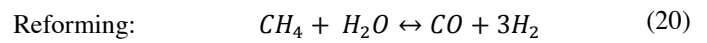$$SO_2 + 3H_2 \leftrightarrow H_2S + 2H_2O \tag{19}$$
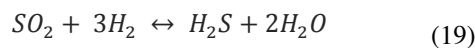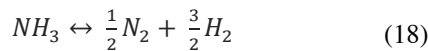
The first two reactions are the main char endothermic reduction reactions which reduces the gas temperature during the reaction for the production of H_2 and CO. H_2 is also produced by the water-gas shift reaction. Unreacted char and ash are considered solid residues and have to be removed continuously from the gasification reactor using a cyclone.

C. Producer Gas Cleaning

In order to obtain the syngas, it is necessary to clean the producer gas from impurities such as particulates, alkali metals, tar, nitrogen, and Sulphur compounds since they may cause clogging, catalyst poisoning, equipment fouling and blocking. First, the cyclone clean most of the large particulates greater than 50 μm in diameter. Then, scrubbers and filters are used to clean syngas from smaller particulates. So, a scrubber is used to remove mainly H_2S which is a must to be removed to avoid equipment damage, prevent catalyst poisoning, and meet the final product requirements.

D. Auto-Thermal Reforming

The cleaned syngas is sent to the auto-thermal section in order to increase the production of hydrogen and carbon monoxide. The catalyst for the reformer was selected to be 10 wt. % of Ni/Al_2O_3. In this section, the main reactions taking place are:

Reforming: $\quad CH_4 + H_2O \leftrightarrow CO + 3H_2 \tag{20}$

Reforming: $\quad CH_4 + 2H_2O \leftrightarrow CO_2 + 4H_2 \tag{21}$

Water-gas shift: $\quad CO + H_2O \leftrightarrow CO_2 + H_2 \tag{22}$

Total combustion: $\quad CH_4 + 2O_2 \leftrightarrow CO_2 + 2H_2O \tag{23}$

E. Methanol Production

Methanol synthesis reaction is exothermic, so the process is operated under low temperature and high pressure to obtain high methanol production. Even though CO is preferred to have higher concentration than CO_2, the concentration of CO_2 is important since it is found that the carbon in the methanol is obtained from CO_2. Hence, it is suggested to operate at a minimum concentration of 2 to 3 vol% CO_2, but not very high as to reduce the activity of the catalyst. It is found that the optimum conversion is obtained when CO_2 concentration is12 vol%. A Cu based catalyst was found to be most suitable for this reaction with temperature limitations; Cu based catalyst is found to be slightly active at temperatures below 230°C. The selected catalyst for methanol production is $Cu/ZnO/Al_2O_3$ Pressure is another important factor to consider in which increasing the operating pressure favors the conversion of CO and CO_2. However, increasing the pressure will increase the utility and operating cost. Thus, there should be a compromise between optimum pressure conditions and cost. The following two reactions take place in the methanol packed-bed reactor:

$$CO_2 + 3H_2 \leftrightarrow CH_3OH + H_2O \tag{24}$$

$$CO_2 + H_2 \leftrightarrow CO + H_2O \tag{25}$$
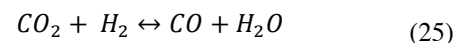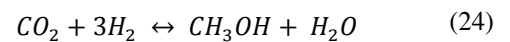
F. DME Production

After obtaining methanol from syngas, methanol undergoes a dehydration reaction, at a temperature range of 250°C to 400°C on a solid-acid catalyst which was selected to be HZSM-5, to produce dimethyl ether according to the following reaction:

$$2CH_3OH \leftrightarrow CH_3OCH_3 + H_2O \qquad (26)$$

IV.   ECONOMIC ANALYSIS

Dimethyl Ether Market is projected to reach USD 9.7 Million in terms of value by 2020, signifying firm annualized growth of 19.65% between 2015 and 2020.

The IEA states that approximately 27% of the global energy demand for transport could be met by biofuels in 2050; therefore, it is a lucrative and interesting investment. [4]

Nonetheless, the gasifier plant was designed so that it minimizes cost and increases profit while keeping in mind environmental considerations. The estimated capital cost of the equipment will be used from similar gasification projects. Since the capital cost is based on cost data for recent similar projects, the order-of-magnitude of the error could be estimated at >30 %.

TABLE 2 COST ANALYSIS FOR SIMILAR PROJECTS [6]

Cost Year	1991	2005	2000	2003
Plant Size (dry metric ton per day)	1,650	2,000	1,741	4,540
Feedstock	Generic biomass	Poplar	Poplar	Switchgrass
Fuel Output	Methanol	Ethanol	FT liquids	Diesel, gasoline
Feedstock Cost ($/dry short ton)	41	35	33	46
Capital Investment	N/A	191	387	541
Product Value ($/GJ)	15	12	16	15
Product Value ($/Gallon)	1.90	1.60	2.00	1.85

Using Tijmensen's plant cost and other similar plants as well as accounting for the location and inflation; the total equipment production cost was estimated to be 60,000,000 $.From this value the approximated Fixed Capital Investment was calculated. The Fixed Capital Investment (FCI) is divided into direct and indirect fixed capital costs. The direct costs are paid directly for the equipment and land preparation coupled with other considerations, while the indirect costs are paid for engineering designs, legal expenses, contingency, with other indirect charges.

In addition, there is working capital that consists of primary raw material stocks, salaries, taxes payable, etc...

The sum of the direct, indirect fixed capital investment, and working capital is defined as Total Capital Investment (TCI) which is the sum money that has to be paid at the beginning so that the plant can be built.

The operation costs should also be calculated in order to have a proper estimate for the cost. The operating costs are two types: direct operation cost including charges like power and raw material, and indirect costs including administration and insurance charges.

TABLE 3 OPERATING AND FIXED CAPITAL COST ESTIMATES (all values are in $)

Fixed Capital Investment	
Direct Costs	
Purchase equipment delivered	60,000,000
Equipment installation	28,200,000
Instrument and control (installed)	21,600,000
Piping (installed)	40,800,000
Electrical systems (installed)	6,600,000
Buildings (including services)	10,800,000
Site development	6,000,000
Service facilities (installed)	42,000,000
Total direct FCI costs	216,000,000
Indirect costs	
Engineering and supervision	19,800,000
Construction expenses	24,600,000
Legal expenses	2,400,000
Contractor's fee	13,200,000
Contingency	26,400,000
Total FCI indirect costs	86,400,000
Working Capital	50,400,000
Total FCI	302,400,000
TCI	352,800,000
Operating Costs	
Direct Operating Costs	
Raw material	5,676,480
Power	83,474,680
Labor (L)	8,409,600
Supervision (S)	1,261,440
Maintenance and Repair (MR)	15,120,000
Lab supplies	1,261,440
Operating Supplies	2,268,000
Direct Operating Costs	117,471,640
Indirect Operating Costs	
Property Taxes	4,320,000
Insurance	2,160,000
Overhead	14,874,624
Administration	252,288
Total indirect operating costs	21,606,912

Charging 150$/ton for the feed and 850$/ton for DME makes 181,713,060 $ in total sales per year. This number was not chosen at random as it was compared with the previous governmental price given to Sukleen at 147$/ton.[5] Subtracting the total operations cost and taking into consideration the government incentives as no taxes, net earnings will be 54,633,956$ per year. This gives a modest return on investment of (18%) and can be improved with 9further optimization and heat integration.

Although the price of gasoline has dropped in recent periods giving the plant a higher risk on investment, the DME forecasts remain positive giving the possibility of a hopeful DME price.

A depreciation rate of 5 %, an inflation rate of 7%, and a salvage of 5% gives the following monetary graph:
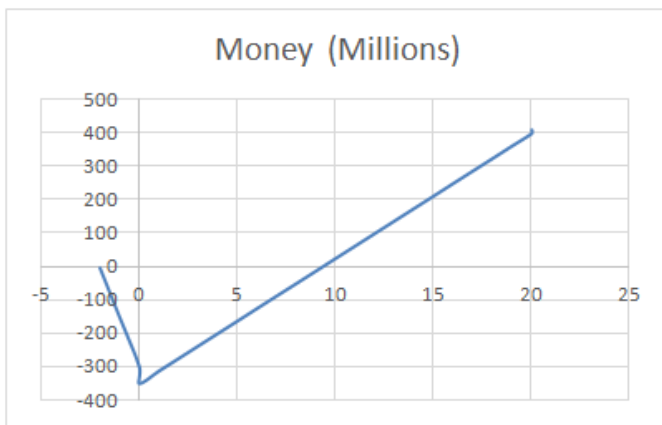

Figure 3 Monetary Quantity ($ million) vs. Time (years)

V. SITE SELECTION

The bio refineries' location plays a very important part in the design. The proper location of the plant can ensure effective, prosperous, and optimal functioning. However, an improper location can prove to be the biggest obstacle for plants and even be the cause of their failure. Therefore, extreme care and caution was taken in picking the optimal place for the plant. Since the solid waste management crisis is concentrated in the Beirut province and in order to decrease the cost of solid waste transportation, the plant should be placed somewhere close to Beirut. However, since this plant requires huge spaces in order to operate and needs a fresh stream of water, the site selected should be in a place where both of the requirements are abundantly available. After extensive research the suggested location of the plant was near the Damour River.

This choice was made due to several factors that include:

Transportation
Transportation costs play a major role in plant economy therefore it is imperative that transportation costs are reduced. Since the solid waste treated will be for Beirut, the plant will be located within a 19 min drive from Beirut and an 8 min drive from Naameh Landfill. This location thus can provide a constant feed of biomass for the plant at a stable and ensured rate.

Utilities
This plant location can provide plenty of utilities. The plant has access to river water that can be used for cooling and providing a fresh water stream for use. The Damour River has a yearly average discharge of 35 m^3/hr. In addition, the site has access to sea water which can also be used in big amounts and high pressures for cooling hot plant streams.

Employees
The plant location is ideal for many employees who can use shuttle buses and vans from two major cities in Lebanon, Beirut and Saida. The two cities are of close proximity and can provide the 240 staff of unskilled workers required.

Size
The 25,000 m^2 was reached after comparisons with other plants that gasify bio-mass, thus making the suggested land fit is consistent with the required size criteria.

Legal Permits
Since all sea viewed land were owned or are the property of the government, the government can provide an incentive in lowering the price of this land.

Skilled Labor
Skilled labor is available as Lebanon has the top universities in the region. These highly-ranked universities have the experience and reputation in providing Lebanon with a constant number of elite engineers and skilled workers.

Motorway
The motor way is available and in direct proximity to the land. This motorway is considered an international highway and thus can handle the traffic of several waste tanks per day and also at night.

Amenities
One of the things that make this plant location a decent choice is the possibility of making a DME export sea port. This port can be specialized for DME orders from turkey and closer European countries as they don't have enough DME plants. Further, this gasification plant can compete with the Chinese DME plants due to the proximity to possible markets and transportation cost reductions.

Safety
Safety is imperative and the location of the plant is in a considerable range away from communities. Thus, this location makes civilian interference and odor complaints obsolete.

VI. CONCLUSION

In brief, the plant transforms 720 ton/day of municipal solid waste, which makes 20% of the total municipal waste in Lebanon, into 19 ton/hr of DME. The conversion of methanol in the reactor is 97.99%. As for the economic analysis, the income of the plant is generated by processing the feed at a price of 150$/ton and selling DME for 850$/ton. The yearly net earnings are 54,633,956$ per year resulting in a ROI of 18%.

The Solid Waste Management crisis in Lebanon cannot be solved by a mere plant design. The problem is much deeper as it is rooted in the culture and in the minds of the Lebanese citizens. Governmental vacuum and political instability plays a major role in the deterioration of Lebanon.

REFERENCES

[1] Abdullah, C. "Provision of consultancy services for the preparation of a master plan for the closure and rehabilitation of uncontrolled dumps throughout the country of Lebanon," UNDP, pp.1-7, (May, 2015)

[2] Mikhael, M. "Solid Waste Management in Lebanon" Bloom bank, pp. 1-4, (Aug, 2015)

[3] Dimethyl Ether Market by Raw Materials (Coal, Methanol, Natural gas and Bio-based feedstock), by Applications (Aerosol Propellant, LPG Blending, Transportation fuel and others), and by Region - *Trends & forecasts to 2020*

[4] Dimethyl Ether Market Worth 9.7 Million USD by 2020. (n.d.). Retrieved March 21, 2016, from http://www.prnewswire.com/news-releases/dimethyl-ether-market-worth-97-million-usd-by-2020-523065471.html

[5] Mikhael, M. "Solid Waste Management in Lebanon" Bloom bank, pp. 1-4, (Aug, 2015).

[6] A. Kayode Coker (2007). Ludwig's Applied Process Design for Chemical and Petrochemical Plants, Volume 1 (4th ed.). Gulf Professional Publishing. pp. 732–737. ISBN 0-7506-7766-X.

DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING

CONTENT

A Sequential Data Assimilation Approach for In-Situ Characterization of Nonlinear Soil Behavior

Wael Slika, Cesar Yahia
Civil & Environmental Engineering Department
American University of Beirut
Riad El-Solh, Beirut, Lebanon
wgs02@mail.aub.edu, cny00@mail.aub.edu

Abstract- **Accurate local site response analysis allows for assessing the response of a soil column nearby an active seismic zone. However, the modeling of this behavior is accompanied with uncertainty that could lead to either an unsafe or conservative design. The main sources of uncertainty include the huge variability in the characteristics of the geologic strata, simplifications in mathematical models, errors associated with identifying properties via laboratory tests, and limited historical events. In this study, a probabilistic framework for estimating dynamic soil properties using a sequential data assimilation technique is developed. Particularly, real time acceleration data and simulations performed in OpenSees are coupled with an Ensemble Kalman filter framework (EnKF) to identify soil properties. Compared to a deterministic framework, the use of a probabilistic framework permits reliable prediction of model parameter amid uncertainty. A pressure dependent nonlinear soil model defined in OpenSees is used to characterize soil response of a multi degree of freedom model. This strategy would lead to an improved understanding of in-situ soil behavior, and it would enhance the analysis of soil-structure interaction models developed in OpenSees. Additionally, the calibration of a nonlinear hysteretic model allows prediction of soil behavior under various loading scenarios; this is not possible if the soil column idealized as a lumped mass model.**

I. INTRODUCTION

The response of a soil column subject to an earthquake ground motion depends on the properties of the soil, the stratigraphy, topography, extent of saturation, and several other factors [1]. Additionally, the dynamic behavior of a structure subject to an earthquake is dependent on the behavior of the soil column on which it is built. Therefore, it is required to model and analyze the behavior of the soil column in order to assess the soil-structure interaction. Evaluating the response of a soil column could be obtained using commercial software such as DEEPSOIL; however, predicting the response using DEEPSOIL or other similar tools requires the availability of a set of parameters that characterize the soil column [2]. The soil parameters that are used to calibrate soil models could be obtained by conducting laboratory tests. Nevertheless, laboratory experiments are associated with a high level of uncertainty, and they do not represent in-situ soil conditions accurately [1].

The development in sensing technologies permitted the emergence of system identification techniques that are used to calibrate soil models and estimate in-situ properties [3]. Particularly, downhole arrays are used to record the seismic response of soil columns at different depths [4]. Downhole vertical arrays are capable of measuring ground motions into and out of a layer as shown in Fig. 1; this allows the characterization of soil behavior, and estimation of soil properties in that layer using inverse theory (system identification) [5].
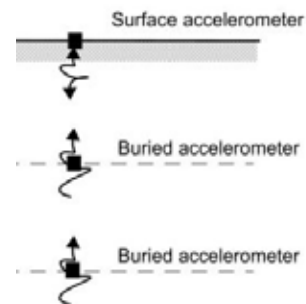
Figure 1 Vertical Array Setup for System Identification [5]

System identification (SI) techniques can be broadly classified into time domain or frequency domain methods [3]. Frequency domain system identification techniques are suitable for linear or equivalent linear systems. Time domain methods are required in order to capture the nonlinear and time varying behavior of soil columns subject to strong ground motions [3]. Application of SI methods that rely on time series include a framework proposed by Glaser and Baise; they estimated the nonlinear soil behavior using an equivalent linear idealization, a time series autoregressive-moving average model, and an extended Kalman filter algorithm [5]. Using the recursive method along with the segmentation scheme proposed, they were able to identify variations in estimated properties [5]. Similarly, Pavlenko and Irikura propose other idealizations of the soil column to estimate stress-strain relationships and study nonlinear behavior [6], [7]. However, time domain SI techniques that pertain to response characterization by using idealizations do not relate to the physical phenomena involved [1]. Furthermore, the estimated parameters from such idealizations could not be readily used for future site response in constitutive models that represent nonlinear soil behavior, further calibration of constitutive models is required [8].

Constitutive model SI methods are split into nonparametric and parametric techniques. Nonparametric SI techniques aim to reproduce observed earthquake motions without recourse to the physical phenomena involved [1]. An example of nonparametric SI methods is the use of sequential artificial neural network to predict stress-strain behavior of soils; the stress-strain curves were matched using observed behavior from triaxial tests [9]. Similarly, autoprogressive methods were used to train nested adaptive neural networks in order to idealize the mathematical formulation of a constitutive model [10]. Other than that, Tsai and Hashash extend the autoprogressive framework to develop a framework of self-learning simulations that aims at creating a neural network based constitutive model [11]. Nevertheless, hysteretic rules and soil constitutive models are required to represent 2-D site response analysis, or to analyze soil-structure interaction [8].

Determining the parameters of a constitutive model using parametric identification requires solving the equation of motion and relying on the constitutive relations [1]. The parameters obtained have physical significance in that they represent characteristics of the soil mass itself. This enables prediction of the soil behavior under different loading conditions [1]. A waveform inversion algorithm was proposed in [8] to estimate insitu soil behavior that is characterized by a formulated constitutive law. However, the robustness of the proposed inversion algorithm is significantly influenced by the number of parameters estimated [8]. Therefore, a different framework is required to estimate multiple parameters associated with a constitutive model.

Lin estimated the response of nondegrading structures by using an Extended Kalman filter iteratively in order to characterize soil behavior attributed to a Bouc-Wien model [12]. Similarly, the Extended Kalman filter had been used to predict properties of lumped parameter mechanical models that idealize nonlinear soil behavior [3] [5] [13]. However, this implies that the state dynamics are assumed to be linear for small vibrations while in reality strong nonlinear behavior exists [3]. This is due to the necessary linearization of model dynamics for propagation of error covariance matrix that is associated with the Extended Kalman filter [14].

Therefore, the aim of this work is to propose a novel system identification framework that is based on the Ensemble Kalman filter to identify nonlinear soil properties. The Ensemble Kalman filter permits the propagation of error covariance matrix based on nonlinear model dynamics [14]. The framework relies on a probabilistic data assimilation technique along with earthquake simulations generated via OpenSees in order to predict the parameters of nonlinear constitutive models. Particularly, downhole array sensor data could be used along with the proposed framework in order to estimate the aforementioned parameters. Thereafter, the

estimated parameters could be used for predicting the response of the soil mass under various loading conditions. The use of a probabilistic framework, as compared to deterministic approaches, permits reliable estimation amid uncertainties that are associated with geotechnical engineering applications. Furthermore, soil-structure interaction could be readily evaluated after using the proposed framework by integrating the calibrated soil models with superstructure models.

II. SEQUENTIAL DATA ASSIMILATION

Sequential data assimilation relies on a set of observational data and the underlying dynamical principles governing the system under observation for estimating unknown state variables and parameters. General schemes for data assimilation often relate to either estimation theory or control theory, but some approaches like direct minimization, stochastic and hybrid methods can be used in both frameworks.

In estimation theory, statistical approaches are used to estimate the state of a dynamical system by combining all available knowledge pertaining to the system including the measurements and the modeling theories. Of significant importance in the estimation process is the a priori hypotheses and melding criterion since they determine the influence of dynamics and data onto the state estimate. One of the most widely used tools in estimation theory is the Kalman Filter [15], which gives a sequential, unbiased, minimum error variance estimate based upon a linear combination of all past measurements and dynamics.

The Kalman Filter is an optimal sequential data assimilation method for linear dynamics and measurement processes with Gaussian error statistics. It provides a linear, unbiased, minimum variance algorithm to optimally estimate the state of the system from noisy measurements. The Kalman Filter loses its optimality for nonlinear dynamical systems. As such many extensions of the Kalman Filter have been developed to tackle the different challenges associated with the problem of sequential data assimilation. The efficiency of the use of Kalman Filtering techniques for system identification and model characterization problems has been demonstrated by Saad and Ghanem in different applications [16] [17]. In this study the Ensemble Kalman Filter will be used to calibrate the soil parameters so as to minimize the mismatch between the measured and predicted outputs (acceleration, velocity, and displacement).

III. ENSEMBLE KALMAN FILTER

The Ensemble Kalman Filter (EnKF) proposed by Evensen and clarified by Burgers et al. is based on forecasting the error statistics using Monte Carlo sampling [18] [19]. The EnKF

propagates an ensemble of state vectors forward in time and updates the state vectors as measurements become available. In this application, the model state vector consists of the dynamic state variables, the model parameters, and the model boundary conditions as shown in (1). V consists of the nodal displacements, velocities, and acceleration in the numerical solution scheme. P represents the model parameters that will be updated in the forward model. U is the matrix holding the ensemble members such that $U_i \in R^n$ as shown in (2). Where N is the number of ensemble members, and n is the size of the model state vector. The ensemble mean is given by (3). $1_N \in R^{N \times N}$ is a matrix with all its elements equal to $1/N$. The model error covariance matrix can be computed as shown in (4).

$$U = [V\ P]^T \qquad (1)$$

$$U = (U_1, U_2, \ldots, U_N) \in R^{n \times N} \qquad (2)$$

$$\bar{U} = U1_N \qquad (3)$$

$$P_e = \frac{(U - \bar{U})(U - \bar{U})^T}{N - 1} \qquad (4)$$

The initial ensemble is chosen so that it properly represents the error statistics of the initial guess of the model states. The initial ensemble is typically created by adding some kind of perturbations to a best-guess estimate, and then the ensemble is integrated over a time interval covering a few characteristic time scales of the dynamical system. The EnKF algorithm consists of two steps, a forecast step (U^f) and an update step (U^a). The forecast step is calculated by using the forward model discussed above to propagate the ensemble of state vectors from time step $(t - 1)$ to time step t as shown in (5).

$$U_{t,i}^f = f(U_{t-1,i}^a) \qquad i = 1, \ldots, N \qquad (5)$$

In the update step, the ensemble of forecasted state vectors U_t^f is updated to minimize the mismatch between the measurements and the corresponding predictions from the simulator. The ensemble of state vectors is related to the measurement variables via an observation matrix, H, that selects forecasted values corresponding to the measurement locations from the state vectors. A new ensemble of observations is generated at each update step. The ensemble of observations is generated by adding perturbations with zero mean and covariance equal to the measurement error covariance matrix to the true observation vector at time t [15]. Let $d \in R^m$, where m is the number of measurements, be the true observation vector at time,t. The ensemble of observations is shown in (6). The ensemble observations are obtained by

perturbing the measurement vector d as shown in (7). Thereafter, the corresponding measurement error covariance matrix is given by (8). $\gamma = (\epsilon_1, \epsilon_2, \ldots, \epsilon_N) \in R^{m \times N}$ is the ensemble of measurement perturbations. The updates states are computed as shown in (9). K_t is called the Kalman Gain matrix and is given by (10). Finally, once the probabilistic characteristics of the model parameters are calibrated, they are used to have a better seismic hazard analysis for the site and the surrounding area.

$$D = (d_1, d_2, \ldots, d_N) \in R^{m \times N} \qquad (6)$$

$$d_j = d + \epsilon_j, \qquad j = 1, \ldots, N \qquad (7)$$

$$R_e = \frac{\gamma \gamma^T}{N - 1} \qquad (8)$$

$$U_{t,i}^a = U_{t,i}^f + K_t(d_{t,i} - HU_{t,i}^f) \qquad (9)$$

$$K_t = P_e H^T (HP_e H^T + R_e)^{-1} \qquad (10)$$

IV. SOIL MODEL & ESTIMATION RESULTS

In order to estimate soil parameters using the aforementioned sequential data assimilation tool, a Tcl program that uses OpenSees framework was developed. Particularly, the program is capable of simulating the earthquake ground motion up to certain point in time. Afterwards, the user could impose displacement, velocity, and acceleration values which represent the system state at a specific time step. Then, the analysis and the ground motion proceed from the last time step while applying the imposed system states. This Tcl program was integrated with a Matlab script in order to use the EnkF tool for updating parameters in time domain.

A pressure dependent multi-yield material was used to represent the elastic-plastic soil behavior [20]. However, plasticity is only apparent in the stress-strain response of the soil element. The required parameters of the model include soil mass density, the reference shear modulus, the reference bulk modulus, the cohesion of the soil, and the shear strain at which maximum shear strength is reached. Afterwards, a single quadrilateral element of the pressure dependent material was coupled with a FluidSolidPorous material to represent a saturated soil mass [20]. Then, the 2 dimensional quadrilateral element was subjected to a gravity loads as well as a ground motion. The behavior was analyzed in time domain using a Newmark integrator.

For illustration purposes, the soil characteristics of a quadrilateral element were calibrated by subjecting it to a ground motion obtained from the Borrego Mountain, El Centro earthquake. The parameters to be calibrated were taken to be as the soil mass density and the reference shear modulus. Particularly, certain values of the parameters were assumed to be the true properties of the soil. Then, the aforementioned ground motion was used and the response at the surface of the element was taken to be the true response due to the subjected earthquake. In practice, the response at the top a soil column would be measured using sensors. The acceleration, displacements, and velocity are assumed to be obtained from the sensors. Afterwards, the initial values for the reference shear modulus and soil mass density were varied, and the framework was used to arrive at an estimate of the true values.

The same model that was used in the forecast step is adopted to generate the synthetic measurements. However, to represent a real scenario, the measurement errors was perturbed with an additive Gaussian white noise having a standard deviation equal to 0.5% of the actual data. The true values for maximum shear modulus and soil density were taken to be 32000KN per m^2 and 2.0ton per m^3 respectively. On the other hand, different initial sets estimates for maximum shear modulus and soil density were tested as seen in table 1, to asses the convergence of the presented scheme. Different source of uncertainty were added to the scheme; the initial Guess error was taken to be 5%, model error is 2%, parametric error is 1%, and sensors error is 1%. These uncertainties characterize the error in initial perturbation, the assumed sensors and the adopted model respectively. All these error are assumed as Gaussian white noises. Fig.2 and Fig.3 show the variation of the true and EnKF estimated values of displacement and acceleration respectively. Fig.4 and Fig.5 show the EnKF prediction results of the mass density and the reference shear modulus respectively. The graphs show a considerable improvement and accurate estimation for the dynamic behavior (displacement and acceleration) and the soil parameters. Moreover, these graphs reflect the robustness of the suggested scheme as it successfully converged to the true value independent of the initial guess estimate. Note that the preceding example is only for illustration purposes; the value of the EnKF framework lies in the capability of estimating multiple nonlinear model parameters.

Table 1nitial guess estimates

Initial Guess Set	Maximum Shear Modulus (KN/m^2)	Soil Density (Ton/m^3)
IG$_1$	34000	2.2
IG$_2$	28000	1.7
IG$_3$	36000	1.5
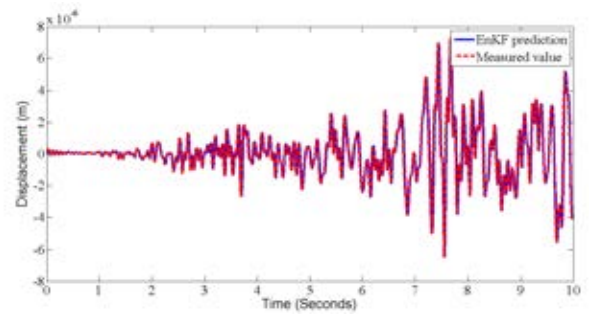
Figure 2 Estimated Displacement History and Actual Displacement History
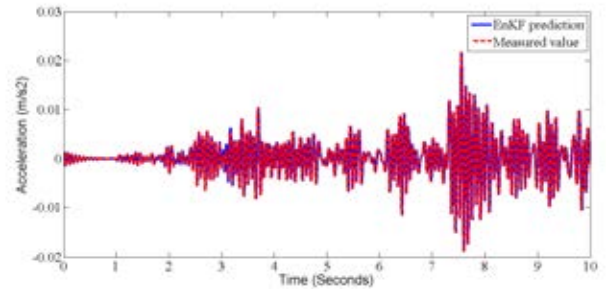
Figure 3 Estimated Acceleration History and Actual Acceleration History
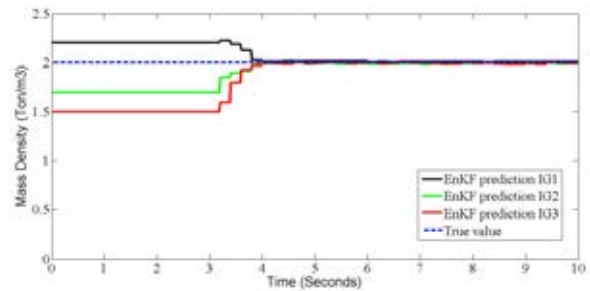
Figure 4 Variation of Mass Density Estimate
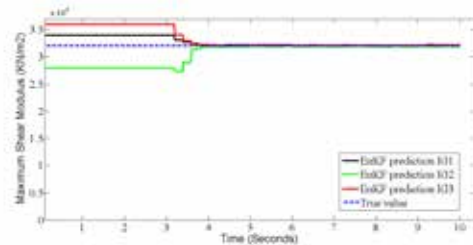
Figure 5 Variation of Maximum Shear Modulus Estimate

V.   DISCUSSION & CONCLUSION

The proposed framework relies on the Ensemble Kalman filter for estimating nonlinear properties of soils. Particularly, the proposed framework is used for calibrating soil constitutive models that are developed in OpenSees. The calibration of soil constitutive models allows predicting the behavior of the soil column under various loading conditions; this would not have been possible if the soil column was idealized by a lumped mass model. Furthermore, the use of OpenSees permits integrating the soil model with super structure models readily. Other than that, while the Extended Kalman filter requires

linearization of the error covariance matrix, the Ensemble Kalman filter permits analysis of full nonlinear models [14]. This enables accurate calibration of fully nonlinear constitutive models as opposed to the approximations that result from linearization.

The use of the proposed framework is illustrated via a quadrilateral element example. The results indicated that the soil properties and dynamic behavior of the soil element were accurately predicted. Therefore, the tool could be used to predict the properties of soils between sensors of downhole arrays as shown in Fig. 1. This could be used for modeling the behavior of structural as well as geotechnical systems with increased accuracy. The development in sensor technologies highlights the impact of predictive tools such as the EnKF on system identification. Other than that, the use of the Ensemble Kalman filter could be extended for navigation systems and other Electrical Engineering applications.

REFERENCES

[1] C. Oskay and M. Zeghal, "A survey of geotechnical system identification techniques," *Soil Dynamics and Earthquake Engineering,* vol. 31, no. 4, pp. 568-582, 2011.

[2] Y. Hashash, M. Musgrove, J. Harmon, D. Groholski, C. Phillips and D. Park, "DEEPSOIL 6.1, User Manual," Urbana, 2015.

[3] A. Mikami, T. Sawada and T. Ekawa, "Identification of non-linear and non-stationary soil properties during the 1995 Hyoguken-nanbu earthquake," *Soil Dynamics and Earthquake Engineering,* vol. 23, pp. 279-286, 2003.

[4] A. Elgamal, "Dynamic Soil Properties, Seismic Downhole Arrays and Applications in Practice," in *International Conferences on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics*, 2001.

[5] S. Glaser and L. Baise, "System identification estimation of soil properties at the Lotung site," *Soil Dynamics and Earthquake Engineering,* vol. 19, pp. 521-531, 2000.

[6] O. Pavlenko and K. Irikura, "Estimation of Nonlinear Time-dependent Soil Behavior in Strong Ground Motion Based on Vertical Array Data," *Pure and Applied Geophysics,* vol. 160, no. 12, pp. 2365-2379, 2003.

[7] O. V. Pavlenko and K. Irikura, "Identification of the non-linear behaviour of liquefied and non-liquefied soils during the 1995 Kobe earthquake," *Geophysical Journal International,* vol. 160, no. 2, pp. 539 - 553, 2005.

[8] D. Assimaki, W. Li and A. Kalos, "A Wavelet-based Seismogram Inversion Algorithm for the In Situ Characterization of Nonlinear Soil Behavior," *Pure and Applied Geophysics,* vol. 168, no. 10, pp. 1669-1691, 2011.

[9] C. Yao, D. Penumadu, R. Zhao and G. W. Ellis, "Stress-Strain Modeling of Sands Using Artificial Neural Networks," *Journal of Geotechnical Engineering,* vol. 121, no. 5, pp. 429-435, 1995.

[10] D. Sidarta and J. Ghaboussi, "Constitutive modeling of geomaterials from non-uniform material tests," *Computers and Geotechnics,* vol. 22, no. 1, pp. 53-71, 1998.

[11] C.-C. Tsai and Y. M. Hashash, "A novel framework integrating downhole array data and site response analysis to extract dynamic soil behavior," *Soil Dynamics and Earthquake Engineering,* vol. 28, no. 3, pp. 181 - 197, 2008.

[12] J.-S. Lin, "Extraction of Dynamic Soil Properties Using Extended Kalman Filter," *Journal of Geotechnical Engineering,* vol. 120, no. 12, pp. 2100 - 2117, 1994.

[13] Y. Zheng, W. Mansheng, H. Liu, Y. Yao and Z. Xiyuan, "Time-domain identification of dynamic properties of layered soil by using extended Kalman filter and recorded seismic data," *Earthquake Engineering and Engineering Vibration,* vol. 3, no. 2, pp. 237 - 247, 2004.

[14] H. Madsen and R. Cañizares, "Comparison of extended and ensemble Kalman filters for data assimilation in coastal area modelling," *International Journal for Numerical Methods in Fluids,* vol. 31, no. 6, pp. 961 - 981, 1999.

[15] R. Kalman and R. Bucy, "New results in linear filtering and prediction theory," *Journal of Basic Engineering,* pp. 95-108, 1961.

[16] G. Saad and R. Ghanem, "Characterization of reservoir simulation models using a polynomial chaos based ensemble Kalman Filter," *Water Resources Research,* vol. 45, 2009.

[17] G. Saad and R. Ghanem, "Robust Structural Health Monitoring Using a Polynomial Chaos based Sequential Data Assismilation Technique," *Computational Methods in Stochastic Dynamics,* vol. 2, pp. 203-213, 2013.

[18] E. G, "Sequential data assimilation with a nonlinear quasigeostrophic model using monte carlo methods to forecast error statistics," *Journal of Geophysical Research,* vol. 99, p. 10143–10162, 1994.

[19] G. Burgers, P. van Leeuwen and G. Evensen, "Analysis scheme in the ensemble kalman filter," *Monthly Weather Review,* vol. 126, p. 1719–1724, 1998.

[20] S. Mazzoni, F. McKenna, M. Scott and G. Fenves, "OpenSees Command Language Manual," 2006.

Adding Supplementary Floors and Basements to a Pre-Existing Building

Lara F. Gebara, Maya E. Habib, Ray T. Harran, Rita G. Mouawad and Adam D. Zarifeh

Department of Civil & Environmental Engineering

American University of Beirut

P.O. Box 11-0236, Riad El-Solh, Beirut 1107 2020, Lebanon

lfj01@mail.aub.edu, meh05@mail.aub.edu, rth07@mail.aub.edu, rgm17@mail.aub.edu, adz00@mail.aub.edu

Abstract- **"Between 1960 and 2012, Lebanon's resident population increased from approximately 1.9 million to 4.2 million" [1]. As a result, most of Lebanon's urban areas have become extremely congested, both with a lack of parking facilities and an increasing need for housing. This growing congestion inspires engineers to come up with innovative ideas to tackle this issue. Thus, the purpose of this paper is to address the possible implementation of one of the many solutions by accommodating for new parking spaces and extra floors in pre-existing buildings. This approach is justifiable by studying the nature of most structures built prior to the 1980s, which either have no basements or one underground story used for storage or shelter during the civil war. Thus, underpinning will provide an underground extension of the structure while a structural investigation will make sure that added floors do not threaten the safety and integrity of the building. This is comprised of the design of micropiles, foundations and a shoring system required for underpinning, as well as typical structural design including the possibility of jacketing. In addition, earthquake analysis and parking layouts are a necessity for the completion of the proposed solution. Finally, as part of the feasibility study of the project, environmental and economic aspects are tackled in order to examine whether or not such a radical undertaking would be beneficial to the Lebanese community.**

I. INTRODUCTION

This project was selected to tackle a problem that has been evident over the past several years, and is only growing in severity. The aim is to address both the technical aspects of the problem as well as the overall sustainability issues of implementing the plan. It consists of adding two supplementary floors as well as two underground basements to a pre-existing five floor residential building, in order to create more livable spaces whilst accommodating parking spaces, not only for the new potential residents, but also for the entire building. The pre-existing building that was selected stands in the Dekwaneh area of Beirut, in a relatively dense neighborhood surrounded by other residential buildings, offices, governmental establishments, and an arterial road that delivers cars to and from the highway.

Choosing this building was the first step into the project. It had to be done following a set of criteria that would make sure that this selected structure could be generalizable to similar buildings of the broader area of Beirut. All the related maps and drawings were obtained from the building's residents and reproduced on AutoCad to obtain workable maps of the structure. Next, the soil characteristics of the site were determined and structural and geotechnical analyses were conducted, including gravity and seismic load study, as well as underpinning and foundation design.

Finally, parking layouts in addition to the economic, social, environmental, and overall sustainable aspects of the project were rendered thus deeming the project feasible or not and to what extent.

II. LOCAL APPLICATIONS

In order to achieve the aforementioned objectives, numerous technical concepts were used in the project. Namely underpinning, which is the process of modifying or upgrading existing foundation. The most widely used type in the region requires the use of micropiles. These are typically reinforced concrete piles with small diameters that are drilled into the existing foundations in order to transfer the load of the structure to the underlying ground. They are designed to allow for excavations underneath the existing foundations while keeping the structure intact in terms of safety and serviceability [2]. The implementation of micropiles produces acceptable quantities of dust, noise and vibrations, and the adaptation of equipment used during their construction to work in confined spaces make micropiles most suitable for the project. Case studies procured by Bauer and International IGM, two geotechnical contractors, were proof that such underpinning projects that may seem too complex, were actually executed in several regions of Beirut. Although they remain uncommon, the option of preserving old buildings can be a very attractive concept. One of these cases is the underpinning of a building in Ashrafieh by International IGM. It is very similar to the proposed project in this study. The following photograph, in Figure 1, taken from this site during implementation shows a more concrete and illustrative representation of what has been so far only theoretical.

Figure. 1 Previous Underpinning Procedure in Ashrafieh, Lebanon

Underpinning of the building is the first step of the construction sequence. It is followed by the excavation down to the new foundation level leading to a bottom up construction. In order to link the new columns with existing ones, non-shrink grout is used to allow for a smooth load transfer procedure thus minimizing unnecessary displacements.

On the other hand, adding supplementary stories to the building might require jacketing for some of the existing columns. The objective of jacketing is to augment the structural member's strength and stiffness by increasing its section size. A steel reinforcement cage is thus fixed and cast to the existing members [3]. The concerned columns and their appropriate rehabilitation are determined following a structural and seismic analysis of the building. Note that the addition of shear walls might be necessary in certain cases.

III. RESULTS AND FINDINGS

A. Existing Structure

As a first step in the proposed project, choosing a building was the most critical part. It had to be representative of the buildings of its area in order for it to be generalizable once the whole analysis was complete. A key objective in the proposed study was to choose a building that satisfies a set of criteria that mainly includes: a concrete frame structure constructed between the 1950s and 1980s that does not exceed 6 stories in height or 1 basement in depth for the ease of the feasibility of the underpinning design and construction. The chosen building should also be located in a relatively congested area where land is expensive in order to justify the high costs emanating from implementing the proposed modifications. Several options in the areas of Verdun, Sin El Fil and Hamra, were examined but few were suitable. The best option was a building in Dekwaneh, constructed in the 1960s. It consists of five stories, a ground floor and a basement that only covers part of the footprint area.

B. Site Investigation

Before conducting any type of foundation or underpinning design, a geotechnical site investigation is crucial to obtain enough data that would contribute to the choice of the geotechnical design parameters. In the case of the proposed project, drilling boreholes and applying tests on soil samples was not possible. Thus, after contacting a few geotechnical firms, borehole logs obtained from different neighboring areas, were used to approximate the values of the geotechnical parameters of the site. Literature correlations as well as specialists' expertise contributed in the interpolation between the neighboring boreholes and the given site. Based on this analysis, the chosen soil stratigraphy and parameters, taking

Stiff Sandy Clay
$\gamma = 19 \; kN/m^2$
c = 100 kPa $\emptyset = 0°$
E = 30,000 kPa

9 m

Water table

Fractured Limestone
$\gamma = 21 \; kN/m^2$, f_s=250 kPa
E=250,000 kPa
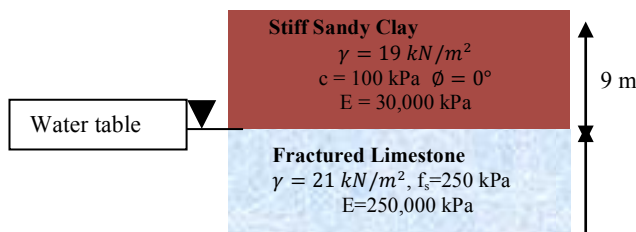
Figure. 2 Soil Stratigraphy

into consideration the worst-case scenarios were assumed to be as in Figure 2.

C. Jacketing and Seismic Design

In order to gain a better understanding of how stable the structure is, and in order to avoid any risk of failure, a structural assessment of the building was crucial to the project. Maps and drawings of the chosen building were available in hard copy. Using AutoCAD, a floor plan of the typical stories was drawn with column locations and dimensions. The hourdis slab thickness was taken to be 30 cm as per the instructions of the owner of the building and the clear height was measured to be 3 meters in typical floors and 2.5 meters in the existing basement. After importing the drawing into Etabs, the team modeled and analyzed the existing structure. The latter contained 39 columns of varying sections with an assumed 1% reinforcement ratio. Since the project will not be implemented, it was not possible to take core samples from the existing structure to determine its capacity, thus a concrete compressive strength of 20 MPa was assumed, given that the structural elements were in good shape. As per the advise of Dr. Bilal Hamad, Table 1 summarizes the super imposed dead load and live load assumptions that were used in the analysis.

After running Etabs, the team compared the load demands of each column to their capacities. The capacity of each column is calculated using (1), [3]:

$$P_u = 0.8 \times 0.7[.85 f_c'(A_{gross} - A_{steel}) + f_y A_{steel}. \tag{1}$$

It turned out that the structural concrete frame of the existing building was designed properly to sustain gravity loads at the time. Once the stability of the existing structure was insured, the team added the two extra floors and basements, and then checked the building's performance once seismic and gravity loads were applied. It is important to note, that Lebanon is located in a major seismic zone as it falls on major faults [4]. Unfortunately most of its buildings are not well designed to withstand such strong seismic activity. For this reason, the team decided to rehabilitate the structure to sustain both gravity and earthquake loading. Since Lebanon still lacks laws and regulations that dictate the seismic rehabilitation techniques, the ASCE 7-10 standard was used to design for earthquakes throughout this project [5]. The first step was to study the seismic behavior of the existing structure without any additional reinforcement to check whether or not it can withstand an earthquake occurrence. After all the necessary parameters, found in Table 2 were fed into Etabs, it turned out, as expected, that a large number of columns would not withstand the applied lateral load coupled with gravity loads, and thus would require some strengthening.

TABLE I
SDL AND LL ASSUMPTIONS

	Typical Floors	Ground Floor	Parking Lots	Roof
SDL	0.5 T/m²	0.3 T/m²	0.15 T/m²	0.3 T/m²
LL	0.2 T/m²	0.7 T/m²	0.35 T/m²	0.7 T/m²

TABLE II
SEISMIC DESIGN PARAMETERS

	Original Building	After addition of shear walls
Site Class	C	C
Period Parameters (Ct, x)	0.016, 0.9	0.016, 0.9
Response Modification Coefficient (R)	3	4.5
Overstrength Factor (Ω)	3	2.5
Deflection Amplification Factor (Cd)	2.5	4
Occupancy Importance (I)	1	1
Ss	1.2g	1.2g
S_1	0.4g	0.4g
Long-term transition period	8 seconds	8 seconds

As a result, the model was modified by jacketing some columns, adding shear walls and was then run again after adjusting some parameters. Many iterations were required in order to reach an efficient and safe solution.

Four shear walls of 25 cm thicknesses were added along the building facades and 15 of the interior columns of the structure were jacketed 5 cm from each side in both the ground floor and the first basement, thus increasing the reinforcements and their capacities. The reinforcements of the jacketed columns reached a 1.5% reinforcement ratio, where T25 and T32 bars were mainly used.

The design was implemented using ETABS and the final model after addition of shear walls is shown on Figure 3.

D. Shoring Design

Every type of soil has a tendency for landslides. Therefore, before proceeding to excavate any soil from a stable system, it is crucial to conduct a study on whether or not the cut soil would incur a landslide. In most cases, especially in the absence of rock, micropiles are needed to retain the soil. However, where the excavation is as deep as 9 or 10 m, the micropiles, as stiff as they could be, cannot be enough to hold the lateral earth pressure without failing. Thus, rows of pre-stressed anchors are used to hold the wall in place [6].

In the case of underpinning, it is not possible to start excavating soil under the building without making sure that the walls of the excavation would not collapse and slide into the open space. Thus, using the software Wallap, a shoring system was designed, as shown in Figure 4, while reducing micropile displacement to less than 1 cm and making sure the micropiles are stiff enough to withstand all shear and moment stresses (by varying their dimension and reinforcement). It consists of 25 cm diameter micropiles with 50 cm center-to-center spacing, going 5 m deep below the excavation into the soil; the micropile reinforcement consists of 17.78 cm outer diameter casing and 12.6 mm wall thickness. It is accompanied by 3 rows of anchors on each side of the excavation. The anchors are 15 cm in diameter, reinforced with grade 150 steel and their free length was estimated to be around 7 m, inclined at 15 degrees from the horizontal. The pre-stress load to be applied per strut varies between 400 and 500 kN according to its location on the micropiles. It is worth to be noted that special care needs to be taken in sites surrounded by adjacent buildings that have basements. The anchors would have to be inclined enough not to interfere with their stability. In the case of the studied site, the adjacent buildings had no basements and were judged to be far enough not to be affected by the designed shoring system. Also, municipal permits would have to be procured in order to be able to use penetrating anchors in the area around the site.
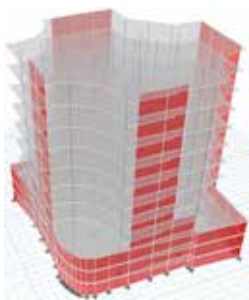
E. Micropiles Design

For the micropiles design, the Federal Highway Administration manual was used as a basis for the design procedure. The manual provides guidelines on determining the diameter, penetrating depth and steel reinforcements of the micropiles, depending on the service loads and soil properties of the chosen area. In terms of design, micropiles are governed by both their structural and geotechnical components. The geotechnical load capacity of the micropile is based on the fact that the contribution of end bearing is negligible, thus skin friction in the grout-to-ground interface only determines the load capacity of each micropile. This is because the cross sectional area of the micropile is negligible and thus barely provides any end bearing. It is also since the micropile will not settle enough to mobilize end bearing. On the other hand, the structural load capacity and stiffness of the micropile depend merely on the reinforcement percentage, dimensions and material of the piles.

According to the FHWA manual, a typical micropile consists of steel reinforcement and a minimum compressive strength of 35 MPa. Other requirements include a minimum spacing of 30 inches or 3 micropile diameters and typical lengths should be less than 90 meters. Due to the fact that limestone rock is present below the new foundation level, Type A piles (gravity grouted and fully bonded in rock or stiff clays) were used in the design. It was decided that 25 cm diameter micropiles were selected; as such dimensions are available

Figure. 3 Final Model on Etabs
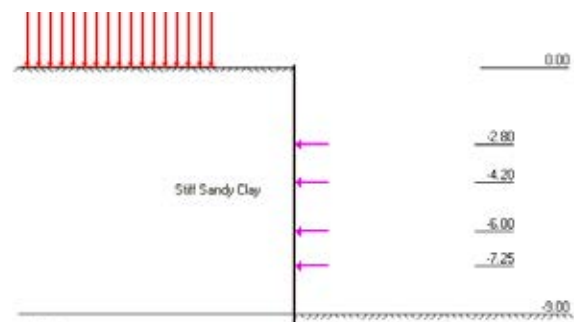
Figure. 4 Shoring System

locally. It was also assumed that the grade 60 steel should account for at least 1% of the area of the pile. The micropiles will be uncased since limestone is present. In addition, a factor of safety of 3 was used, as recommended by the non-seismic load groups. It is important to note that the micropiles will not be designed to resist seismic loads, as they are temporary elements [7].

Structural Design:

To evaluate the allowable structural capacity of the micropile, the Load Factor Design method was used. The allowable compression load is calculated using (2), [7]:

$$P_{c\text{-allow}} = \left[0.4 * f_c' * A_{gross} + 0.47 * f_{y\text{-steel}}\left(A_{bar} + A_{casing}\right)\right]. \quad (2)$$

The designed micropiles were found to have an ultimate capacity ranging between 72 and 88 Tons. Using this capacity, the number of micropiles was calculated using (3), [7]. It is important to note that an even number of micropiles symmetrically drilled was used in order to avoid the generation of moment because of slender columns.

$$\text{Number of micropiles} = \frac{\text{Maximum compression load}}{\text{Capacity}} \quad (3)$$

Geotechnical Design:

In order to evaluate the geotechnical capacity of the micropile, the length was first determined. The bond length of the micropile is the length of the portion of the micropile that will be embedded into the limestone rock after the excavation has been complete. The nominal strength that was used in the calculations was qs = 250 kPa. This value was obtained from consultations with design experts and contractors in the field. Once the number of micropiles had been determined, the next step was to determine their lengths by evaluating the geotechnical capacity of the micropile. Equation (4) was used to calculate the length [7]:

$$P_{G\text{-allowable}} = \frac{\alpha_{bond}}{FS} * \pi * D_{bond} * L_{bond}, \quad (4)$$

Where, α_{bond} is the grout to ground ultimate bond strength, D_b and L_b are the diameter of the drill hole and bond length. One needs to note that in cases where underpinning is used to add basements to an already existing building, the bond length is measured starting from the bottom of the new foundation level.

For cases where footings had edge columns, the procedure stated above could not be used as the minimum number of micropiles needed generated a large moment. Thus, Etabs was used to design the micropiles in a way to generate acceptable moments and take into account buckling purposes. Iterations were carried out until the micropiles provided a capacity that was larger than the design loads. PCA Column was used to validate that the moments generated in the micropiles with the given loads and steel reinforcements were acceptable.

Because of the present fractured limestone, which is relatively strong and stiff, it was easy to reduce settlement to a value less than one centimeter, thus considering it negligible. It is important to note that settlement of the building must constantly be monitored for serviceability and safety issues. Overall, the design consisted of 176 micropiles located in red as shown in Figure 5.

F. Foundation Design

As per the drawings obtained from the owner of the building, it is evident that isolated footings were not a feasible option for the design of the new foundation system. With the increase in axial loads due to the additional stories and basements, the new isolated footings would have to increase in dimensions. This option is not viable, as this will cause the footings to collide together because of the limited space and the difficult site environment. As such, a raft foundation was chosen as a feasible option. Columns were assumed to be axially loaded and a factor of safety of 3 was taken. The area of the raft foundation would have to cover the entire foot print area of the structure, which is around 795 square meters. The raft foundation will be assembled onto the top of the fractured limestone layer.

Based on Meyerhof's bearing capacity equation, which includes the shape and depth factors, the allowable bearing capacity was calculated to be 7648 kPa with a total axial service load of 6308.9 Tons. The depth of the raft foundation was chosen to be one meter based on the structural analysis of the raft on Safe. Due to the presence of fractured limestone, the settlement was reduced to less than one centimeter, which was considered acceptable.

G. Parking Design

As planned, the three basements of this building were to be turned into a parking facility to meet the needs of the residents for parking spaces. According to the Lebanese Law, parking stalls are to be 1.8 m wide and 5.2 m long, while providing for 5.5 m turning radii. In addition, the maximum allowed ramp slope is 20%, except for the first and last 5 m of exterior ramps where the allowed slope is 10% [8]. The design allowed for the creation of 57 parking spaces covering the need of all the building's residents. However, one major limitation is to be mentioned: as per the Lebanese Law, the access for more than 30 cars should be 5.5 m wide. However, in the case of this project's building, an access ramp limited to a width of 2.8 meters could only be provided since disturbing the ground floor's slab was not an option. Thus, in order to solve this
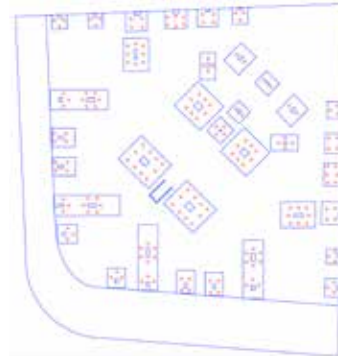

Figure. 5 Micropile Layout

issue and preserve the residents' safety, a number of mirrors as well as a sensor system will be installed notifying drivers of possible opposite traffic emerging. The plan of the parking was first drawn and then the design was extruded into a three dimensional model using AutoCAD as shown in Figure 6.

IV. FEASIBILITY STUDIES

A. Economic Feasibility

In order to come up with a decision regarding the feasibility of this proposed project, an environmental and economic study had to be conducted. With respect to the economical aspect, average prices were obtained from different local contractors in the field and a final bill of quantities was erected as shown in Table 3. It is important to mention that prices included in the BOQ comprise material, labor and equipment costs per unit.

After consulting with a surveyor who visited the site, one square meter of basement in the building's area was found to be worth approximately $1060 and one square meter of residential area was found to be around $2400. Considering that the new buyers will have to pay for their parking stalls of dimensions 5.2 m x 1.8 m yielding an area of 9.36 m^2 per lot, the total parking structure will then generate $158,746 in revenue. To calculate the balance of the implementation of the project, the total cost was subtracted from the selling price of eight apartments with their parking spots ($ 2,976,000), which yields a profit of $723 in addition to two parking stalls per existing apartment. However, it is important to note that the residents need to procure a capital of $156,702 per apartment in order for the project to initiate.

B. Socio-Environmental Feasibility

The aim of this study is to make the construction phase as socially and environmentally friendly as possible and to then consider the long-term benefits of the project. This generalizable case study can be imitated later on to yield more perceptible benefits on a community level.

During Construction

The most important aspects to be controlled are dust, noise and traffic. In addition, while aiming at a "zero waste" approach, the outcome of the demolished and excavated material should be considered.

For Dust Control, concerning the addition of extra floors to the existing building, the construction activities are almost identical to any typical construction site.

In order to insulate the surroundings, the building is to be covered with polymer sheets that can either be taken from previous projects, or purchased.
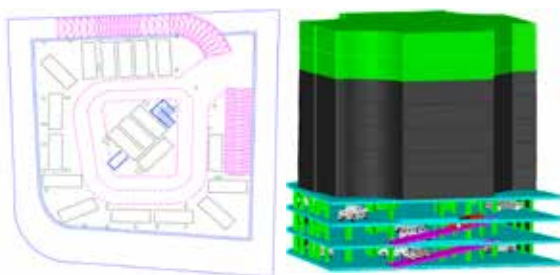
Figure. 6 Plan and 3D views of the parking facility

TABLE III
BILL OF QUANTITIES

	Quantity	Cost per unit ($)	Total Cost ($)
Underpinning Micropiles	176 with 12 m linear depth	80	168,960
Shoring Micropiles + Reinforcement	224 with 14 m linear depth	100	313,600
Excavation (fill material)	6956.25 m^3	10	69,565
Lump Sum Basement Concrete Works	1010.66 m^3	277	280,325
Raft	954 m^3	175	166,950
Jacketed Columns	30	500	15,000
Shear Walls	332.6 m^3	310	103,106
Additional Stories	1240 m^2	400	496,000
Lump sum basement finishes	2385 m^2	130	310,050
Lump sum floor finishes	1240 m^2	250	310,000
Parking control equipment	1 unit	600	600
Parking sensor system	2 units	500	1,000
Parking round 24'' mirrors	6 units	87	525
Metal staircase with rails	4 units	2000	8,000
Anchors	250 embedded 8 m	55	110,000
Epoxy coating: parking floor	2385 m^3	43	102,555
Demolition: basement floor	795 m^3	50	39,75
Demolition: existing footings	246 m^2	50	12,300

Sub-Total	$ 2,508,286
12% Overhead + Profit	$ 300,994
8% Design + Supervision	$ 224,742
Permits	$ 100,000
Total	$ 3,134,023

This also provides mild protection against falling debris. Whereas for the basements works, it would be a good idea to place barriers around the building according to dust control practices to prevent airborne dust at the street level and serve to limit the safety perimeter. Also, the trucks should be covered and possibly have their wheels sprayed to clean the dirt off. For Noise Control, no loud activities will be scheduled before 7:30 AM. In addition to this basic regulation, further restrictions can be applied when special facilities are at proximity, namely hospitals, school universities and/or religious buildings. For this project, there are no special buildings therefore basic noise control is adequate.

For Traffic Control, some potential traffic-congesting equipment will be used including excavators, trucks, and concrete pumps. However, the rate of turnover of trucks is expected to be lower than the rate for a site erecting a whole

new building. This will require a minor traffic routing in collaboration with the municipal police in order to momentarily route the concerned traffic of the two-way major artery of Dekwaneh. Furthermore, no road blocking activity will be scheduled during the morning and afternoon peak hours of the day, between 7 to 8 AM, 2 to 3 PM and 5 to 6 PM.

The Dekwaneh building will yield about 500 m^3 of demolished concrete and 48 Tons of steel reinforcement. This rubble will be generated from breaking down partition walls and other structural members that can serve as coarse aggregates for non-structural concrete or sub-base material for paving roads. The steel reinforcement can be separated from concrete and sold to recycling facilities. Also since the amount debris is not considerable, it is manageable to segregate the construction waste to ease later recycling.

For Excavated Material, depending on the number and area of basements needed, the depth and volume of the excavation material might be moderate or significant. The Dekwaneh case is expected to generate roughly 7,200 m^3 of fill. It is essential to consider proper usage of these byproducts before proceeding to wasting them as excavated material could be sought after in case there are land reclamation projects nearby or roads construction that might require some fill. Other more singular solutions may apply such as using the diggings for nearby parks and plantation curbsides or along road medians.

Post-Construction

Benefits are best divided along three vantage points: benefits for the residents of the building, the whole street and for sustainability.

Adding two additional stories, each accommodating 4 apartments yields a total of 8 new apartments that the residents can sell or rent for a relatively high value considering the crowded geographical location. Also, the existing residents now benefit from an added value to their homes by having 2 parking spots per apartment that also relieves them from the daily stress of finding a parking space.

Moreover, recall that the rehabilitated building falls at a right angle intersection connecting a minor street to the major artery of Dekwaneh. Forty cars that were previously being parked curbside or in paid parking areas are now well accommodated for, underneath the building. For the narrow streets of the network, the absence of curbside parking will make the car flow much smoother. If only a few other buildings join the model, this could result in improving the transportation infrastructures that could revive the commercial or economic aspect of the region, otherwise labeled as "continually jammed". More so, less traffic jams will result in a reduction of air pollution; particularly dangerous pollutants released by car engines such as Carbon Monoxide, Nitrogen Oxide, total hydrocarbons and volatile organic compounds [9].

Regarding the sustainability aspect, the biggest advantage is that the existing structure will not be completely demolished to build a whole new one. Keeping the structural skeleton of the building saves big amounts of resources and energy that would

have otherwise been consumed. Reducing the need for new concrete can be translated into a reduction in demand for aggregates (decreasing the activity of quarries), for cement (reducing the emission of Carbon dioxide). Also, reducing the construction period spares the environment from polluting activities including truck trips, electricity generation and products consumption (machines, energy, tools, paper work…).

V. CONCLUSION

The integration of several engineering disciplines allowed for the design of a system that can provide a pre-existing building with additional stories and basements. A structural analysis coupled with geotechnical design components were needed to come up with a stable transitional foundation consisting of 176 micropiles that would allow for a bottom up construction sequence. The final product generated 8 new apartments, as well as 57 new parking spaces accommodated in a 7-story high structure that agrees with the minimum safety and serviceability requirements. Taking this into consideration along with the data generated in this study, the benefits of such a project out way the costs of implementing it, only in cases where it is possible to add extra floors, thus deeming it feasible from both a financial and an environmental standpoint.

REFERENCES

[1] UN Sustainable Development. (2012, June). National Report to the United Nations Conference on Sustainable Development (RIO+20).

[2] Kordahi, R. Z. (2004). *Underpinning Strategies for Buildings with Deep Foundations*. Massachusetts Institute of Technology.

[3] ACI Committee 546. (1996). *Concrete Repair Guide* (ACI 546R-96).

[4] Huijer, et Al., 2015, Re-evaluation and updating of the seismic hazard of Lebanon

[5] ASCE/SEI 41-06, Seismic Rehabilitation of Existing Buildings 2007

[6] Najjar, S. (2016). Applied Foundations [Class Notes]

[7] Sabatini, P. J., Tanyu, B., Armour, T., Groneck, P., & Keeley, J. (2005). Micropile Design and Cosntruction (FHWA NHI-05-039). National Highway Institute, Federal Highway Administration, U.S Department of Transportation, Washington, D.C.

[8] Ministry of Public works and Transportation. Lebanese Building Decree (2004)

[9] United States Environmental Protection Agency. (2008, October). Idling Vehicle Emissions for Passenger Cars, Light-Duty Trucks, and Heavy-Duty Trucks.

Adding Two Basements for Parking Spaces in an Existing Building

Michel Lebbos, Nassib Sakr, Paul Maamari, Mohammad El Mikati, Robert Jamal

Department of Civil Engineering

American University of Beirut

Hamra-Lebanon

rgj06@mail.aub.edu, pnm02@mail.aub.edu, mke22@mail.aub.edu, mgl02@mail.aub.edu, nns20@mail.aub.edu

Abstract- **Due to the high density in Beirut, and the lack of construction planning in the area, it is very hard to find parking spaces. A proposed solution is to build a basement parking for an existing building that does not have a basement. This research will discuss the way this basement will be built, and a feasibility study will be implemented.**

I.   Introduction

Throughout the years, Lebanon has been suffering from parking problems due to many reasons. One of the main ones is the lack of proper legislations that enforce enough parking spaces to accommodate the traffic. Every time a plan is implemented, the problem is shifted somewhere else rather than being solved. In Beirut, most of the buildings were constructed back in the 1970s and 1980s, at a time where one can hardly find a family having more than one vehicle, hence there was no need for providing parking spaces. However, recently, the need for parking spaces increased dramatically as the population increased and the already limited amount of space decreased; therefore a quick and feasible solution is needed. One of the suggested solutions was to build a basement parking to an existing building that was originally built without a basement. Since this work has been extensively studied in the past couple of decades, the team worked with well-known and experienced engineers (ACC) in addition to the technical advisor of this project. In order to reach the objective of this case study, a sample building was to be chosen which fits the latter conditions. A seven floor building was taken as that sample; it was built around 30 years ago. The building's plan was drawn on AutoCAD, then modeled on ETABS producing the loads on each column, in addition to the manual calculations of the own weight of the building. Finally, a comparison was done for the results of both manual and software calculations. Moreover, a parking design was implemented for the basement of the building. The space available for the parking was around 200 m^2, and abiding by the Lebanese regulations for parking, the parking basement was able to accommodate seven cars only, hence the need for another parking basement since it is required to provide two parking spaces for each apartment. Furthermore, a car lift was chosen over a ramp for this project due to the relatively small area of the building. For the construction phase of this project, micro-piles were found to be the best solution available for underpinning. However, since the building consists of 17 columns, it is recommended to have transfer beams in order to provide space for the basement parking, and maximize the number of parking spaces. Based on the soil report, the type and bearing capacity of the soil underneath the building were found. Additionally, the cost of this whole project will be estimated. The construction of such an underground parking for this particular type of building, which is representative of around 60-70% of the buildings in Beirut, could help solve the parking problem. Residents of Beirut shall not use the parking spots in the streets of Beirut anymore, leaving them free for people that come to Beirut daily. This plan will offer more parking spaces, reduce traffic, and help people save time by parking safely in the confines of their own building.

II. Parking Design
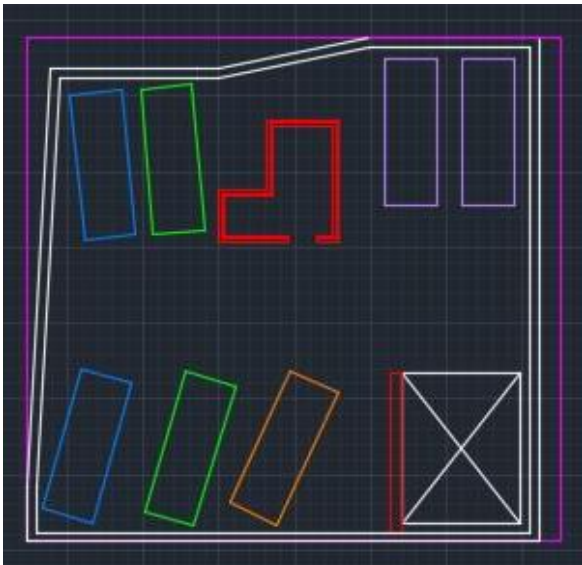
Figure 1: Parking Design Drawing

The building has 7 apartments, so 14 parking spaces are needed. Due to the limited area of the land, 2 basements should be constructed to accommodate for the needed parking spaces. Due to the limitations in the space, a ramp could not be constructed, so we opted for an electric car lift instead (5.5x3m). The lift will be at ground floor level down to the level of the 2nd basement. The Lebanese standard code stipulates that for a parking less than 800m^2, the dimensions of each car should be 4.8mx1.7m keeping a clear distance of 0.4m on the right and left side of each car, and 0.2m from the front. The parking was designed based on a 90º scheme, with an aisle width of at least 5m. The turning radius of the car, the freedom to maneuver, and the large area for the car to park were all accounted for in the design. The figure above shows the adopted parking design for the additional basements. The shaft shown in the 3D Revit model represents the car lift.

III. Site Investigation

In order to understand the project, a site investigation should be carried out. We visited the building under study to find out its actual conditions. Measurements of the building were taken, but due to the lack of structural maps, the plan of the floors of the building were measured by hand on the field, then modeled on AutoCAD in order for later use. Then the building was modeled on ETABS using standard loads from each slab to the columns. The load combinations used were:

$$\text{Design Load} = 1.2DL + 1.6LL \qquad (1)$$

$$\text{Service Load} = DL + LL \qquad (2)$$

With

DL= dead load

LL= live load

During the site investigation, we found out that the building is surrounded by three buildings, one on each side, and the fourth side faces the road. We asked about the depth of the surrounding buildings, and each one turned out to have basements with retaining walls.

For the geotechnical aspect of the project, soil characteristics and type are of great importance since a shoring design is needed during the excavation phase in addition to the micropiles used for the underpinning of the foundations. Since the building is located in Verdun, the soil report of ABC Verdun site were used since it is around 300-400m away from the building under study.

The soil profile consisted of the following:

- 0-20m: destructive drilling but it was found out to be composed of sand
- 20-30m: Yellowish Silty Sand
- 30-60m: Fractured Limestone

Since the borehole depth was 60m and no water table level was found, it was assumed the water table is at least 60m below GF level. It was also assumed that the first 20m of sand share the same characteristics as the lower sand layer with an elastic modulus of 40 MPa

The tables below shows the soil characteristics used in the shoring and micropiles design:

Table 1: Soil Characteristics

	Cohesion (kPa)	Unit weight (KN/m^3)	Angle of friction	Elastic Modulus (MPa)
Yellowish Silty Sand	0	18	30	80
Fractured Limestone	100	22	15	150

Table 2: Soil Characteristics (continued)

	Poisson's ratio	K_0	K_a	K_p
Yellowish Silty Sand	0.35	0.5	0.33	3
Fractured Limestone	0.24	-	-	-

IV. Shoring System

Since our building is surrounded by other buildings from three sides and the road from the fourth side, excavation will be done from the side facing the road. The adjacent buildings have retaining structures, so we need to provide a shoring system from the 4th side only. The shoring system method adopted consists of driving micro-piles along the length of the exposed wall on the side of the wall, and also installing anchors for lateral support of the wall at their designed location. As for the soil report, one borehole is enough since the length and the width of our area do not exceed 30m, which minimizes the probability of having different soil types in the area and increasing the reliability of the borehole. It was found that the first 20m consist of sand followed by 10m of yellowish silty sand and 30m of fractured limestone. According to the geotechnical report of adjacent buildings, the water table was found to be on a depth of 60m. The preliminary design methodology used for the shoring system of this project was modeled on WALLAP using micro-piles as a retaining structure, and anchors as bracing for the retaining structure. Micro-piles of diameter 25cm were used and the length of around 20 meters was determined in order to account for the buckling effect. These micropiles will serve as a supporting wall against the soil. As for the anchors used, the elements used have a cross sectional area of 0.85 in^2 with an ultimate strength of 127.5 kips. They will be installed on 3 levels with 1.5m spacing between them along the horizontal length of the wall at an inclination angle of 45 degrees to avoid underground utilities. The steps below describe the methodology for excavation and shoring system used in this project:

1. Excavation to the level of existing footings (-1.0m below GF)
2. Installation of micro-piles (25cm diameter) up to 20m below Ground Floor level
3. Excavation to a depth of -3m below Ground Floor Level
4. Installation of Anchors at depth of -2.5m below Ground Floor level
5. Excavation to a depth of -6m below Ground Floor level
6. Installation of Anchors at depth of -5.0m below Ground Level
7. Excavation to a depth of -9.2m below Ground Level
8. Installation of Anchors at depth of -7.5m below Ground Level

The figure below shows the shoring system as modelled on WALLAP:
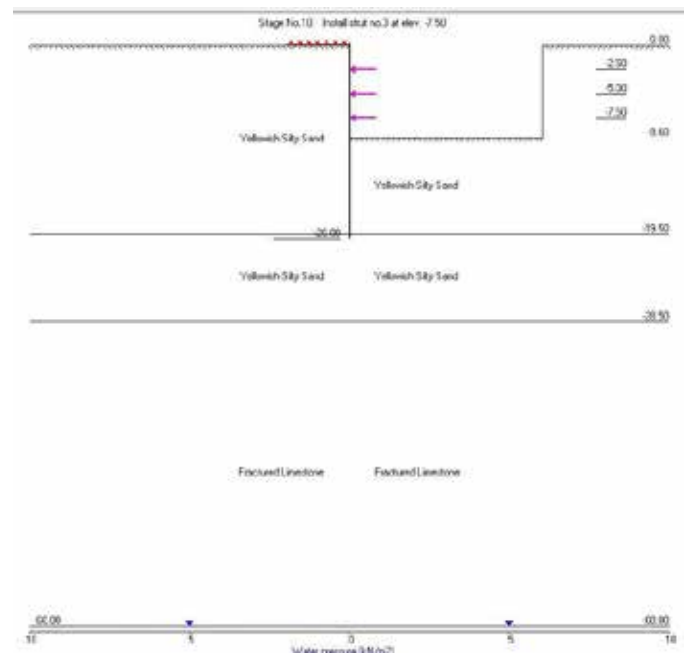
Figure 2: Shoring System Model

After inputting the model on WALLAP, the following results were found: 1cm maximum displacement at -9m elevation, -7 KN.m/m run maximum bending moment and 400 KN/m run shear force while the displacement at the top of the wall was minimal.

In order to obtain the results, it was assumed that a 20m x 6.2m truck surcharge is present on the side of wall for conservative reasons.

The following figure shows the obtained results.
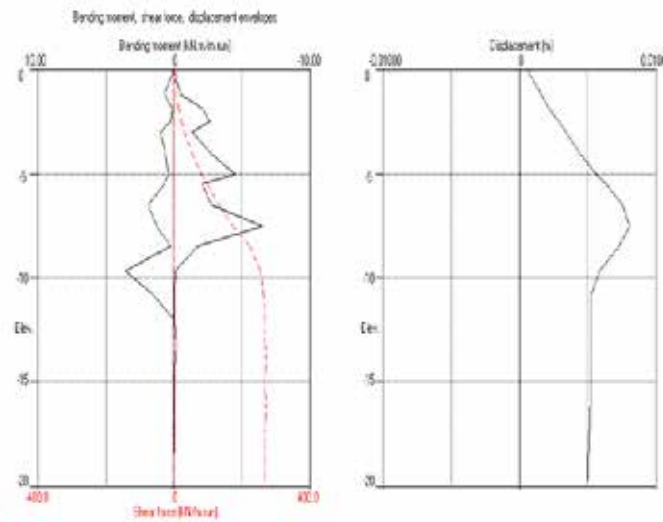
Figure 3: Shear, Moment and Displacement of Wall

V.   EXISTING FOOTINGS

The load combinations (1) and (2) were used to find the loads on the footings of the building, since the most critical part of our structure will be the foundations. The loads on the footings were calculated then proceeded in designing the footings, this is the case since there were no structural maps at hand. Moreover, there were two types of designed footings, combined and isolated, the choice was based on Engineering logic. The area of each footing was calculated using the equation:

$$Af = (P \times F.S.)/Q\_ult \quad (3)$$

Where:

A_f = Area of the footing

P= Service load on the footing

F.S. = Factor of safety

Q_{ult} =Ultimate bearing capacity of the soil

The figure shown in the next section above shows the existing footings dimensions based on the calculations.

VI.   Micropile Design

Micropiles will be constructed beneath some of the footings. Under those footings, four micropiles will be constructed, a pile on each corner. These micropiles will have a free length of 7to9m, which might cause some buckling effect. To overcome this effect, the micropiles will be tied by anchors on 3 levels (-2.5, -5 and -7.5m below ground floor level respectively). The micropiles should carry the load of the column above the footing, each micropile carrying Pc. The length needed will be calculated using the equation:

$$L = (Pc \times F.S.)/(\pi \times D \times \alpha) \quad (4)$$

Where:

L= Length of micropile (maximum 29m)

D= Diameter of micropile

α= Bond strength between soil and pile (100 KPa)

The excel sheet shown in the appendix was created in order to facilitate the design calculations for the micropiles supporting the structure. It shows the resulted dimensions of the micropiles based on the soil characteristics and load on each footing. The micropiles designed varied by length from 10m to 26m depending on the load supported. The micropiles will have steel casing of 1cm with 2cm reinforcement bars.

The functions of these micropiles is to support the building during the excavation and the construction of the proposed basements. Since these micropiles are temporary, they will be demolished after the construction of the basements.

VII. Raft Foundation, Walls, Columns and Transfer Beams

A raft foundation will be constructed at the base of the lowest basement. It will have an area of $252m^2$ with a thickness of 1m. The bearing capacity of the soil and settlement of the raft were calculated accordingly. All the vertical elements, slabs and beams were designed according to the ACI code. After the raft is finished, the columns and core walls will be constructed followed by the slab and transfer beams of the first basement. Similarly, the second basement will be constructed, reaching the ground floor. The design of basement walls is related to the adjacent soil. Using SAP 2000, the wall was modeled to obtain the moments and loads. The thickness of the basement wall to be used is 30cm. The loads on the basement walls are soil loads and water hydrostatic forces. As for the core wall, since it will be continuous throughout the building, we obtained the seismic loads using ETABS. Since the building is in Verdun, the parameters concerning the seismic design incorporates the use of normal shear wall frame in a normal structural environment. The transfer beams are mandatory for this project since they are supporting huge concentrated column loads which had to be removed in the basements to abide by the parking rules and account for the freedom maneuver of the cars. The transfer beams will be designed to carry the load of the hanging columns. Using ETABS, the loads on each column were determined, and using SAP 2000, moments and shear forces were obtained so that the transfer beams could be designed. After construction is complete, the micro-piles will be removed by cutting them. Basement columns were also designed in order to support the load of the structure above. Various excel sheets were developed in order to facilitate the design of the columns and beams. They are shown in the appendix, along with the mentioned designs of the basement as plan view.

VIII. Scheduling and Cost Analysis

The construction components were broken down into the most basic tasks that had to be done in order to accomplish such project, and a timespan was given for each task based on both the quantity demanded and the unite rate of each. First step was to manually compute the quantity needed (the demand) of the project, such as concrete (m^3), steel (kg), earthwork (m^3) and items. Secondly, using Engineering logic and with the advising of experienced Engineers, the construction sequence of the project was implemented (check the "Construction Sequence" table of the project with their relationships in the Appendix). After finding the unit rate (per hour)

of the activities, the duration (days) of each activity was computed. Since the relationships of the activities and their durations were found, the project can be implemented on PRIMAVERA. However, since the unit rates were not finalized so far, the data mentioned will be inputted on PRIMAVERA at a later stage, and a rough schedule will be planned. On the other hand, the cost of the project was subdivided in to three main components: reinforced concrete (RC), earthwork and items. The unite rates of the components were as follows: for the RC a $250\$/m^3$ was taken, $7\$/m^3$ for earthwork, and $5000 for the items (excluding both "braces for micro-piles bracing" and "anchors for shoring system"). Then after having both quantities and the mentioned unit rates, the total cost of the project was computed (check the table presenting the prices of the three parts along with the total cost of the project in the Appendix).

X. References

[1] Applied Foundation Courses Slides (CIVE 630)

[2] Design and Construction of Micropiles, Shong 2003

[3] Back to the Future: Renovation of a 1930s Apartment Building into an Art Gallery.Matt Thomas, P.E., S.E.1 ; Ken Maschke, P.E., S.E., M.ASCE1 ; and William D. Bast, P.E., S.E., SECB1

[4]Micropiles for Re-support and Lifting of Two Buildings under Construction. Brent Byford and Mark Hampton
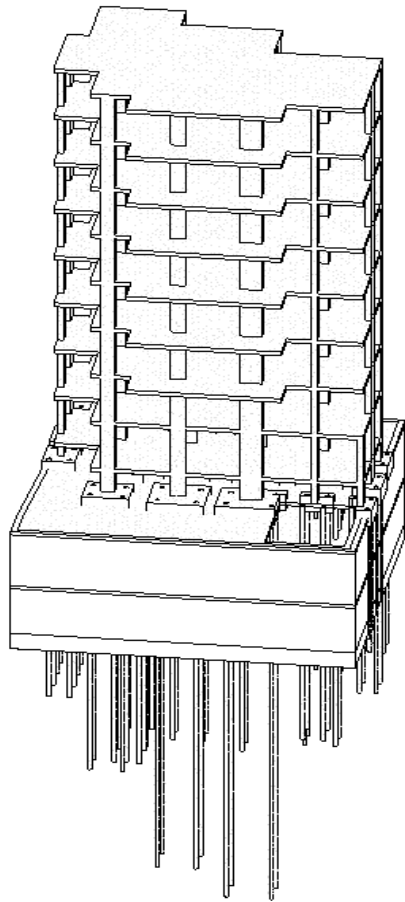
Figure 4: REVIT Model
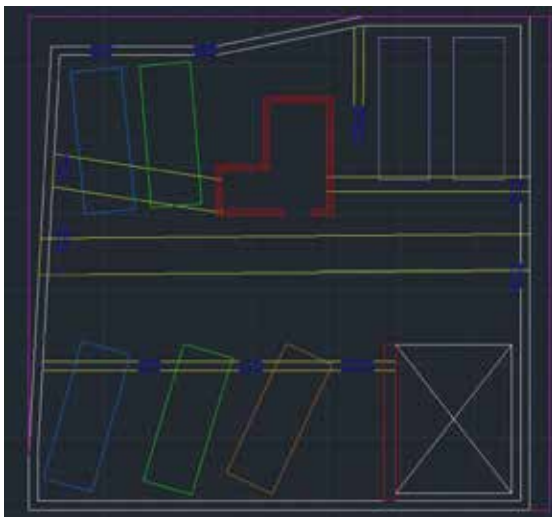
Figure 5: Plan View of the Designs Done in the Basement

Table 3: Total Length of Micropiles Arranged as per the Plan View

15.87	14.58				
			12.66		
15.90	19.70			14.49	12.94
16.09	19.67	18.33	18.83	18.15	14.10
	18.54	24.52	26.29	15.88	9.11

Table 4: Total Cost of the Project in US Dollars

Earthwork	Concrete	Items	Total Price of Project
11900	184445.4	5000	201345.4

Table 5: Construction Sequence

Activity ID	Activity	Relationships
1	Install a Transfer Beam on C(F4)	-
2	Demolish Existing Slab on GF	-
3	Install Micropiles on road side	-
4	Excavate till the bottom level of the footings	FS (2 & 3)
5	Install Micropiles in the footings	FS (4)
6	Excavate till first level of anchors	FS (5)
7	Install first layer of anchors	FS(6)
8	Brace Micropiles	FS(6)
9	Excavate till second level of anchors	FS(7 & 8)
10	Install second layer of anchors	FS(9)
11	Brace Micropiles	FS(9)
12	Excavate till third level of anchors	FS(10 & 11)
13	Install third layer of anchors	FS(12)
14	Brace Micropiles	FS(12)
15	Install Raft	FS (13 & 14)
16	Install Basement Walls	FS (15)
17	Install Columns	FS (16)
18	Install Core Wall	FS (16)
19	Install Beams	FS(17 & 18)
20	Install Slab	FS(19)
21	Install Basement Walls	FS(20)
22	Install Columns	FS(21)
23	Install Core Wall	FS(21)
24	Install Beams	FS(22 & 23)
25	Install Slab	FS(24)
26	Demolish the Micropiles	FS(25)

American University of Beirut Campus: Parking Management Technology

Charbel Abboud, Anas Abou Hawili, Mohamad Anas Bin Ahmad, Ghina El Sabbagh, Reina Tabbara
Department of Civil and Environmental Engineering
American University of Beirut
Bliss Street
P.O. Box 11-0236
Beirut, Lebanon
caa27@mail.aub.edu, aza13@mail.aub.edu, mfb19@mail.aub.edu, gme18@mail.aub.edu, rjt04@mail.aub.edu

Abstract- In the past, several attempts have been made to address the parking issue that the American University of Beirut (AUB) students face. Around 70% of the surveyed students spend more than 15 minutes looking for a parking space due to parking deficit in the area. The goal of this project is to find practical solutions that aim to reduce the parking problem. The proposed solutions focus on the management of parking in the upper and lower AUB campus. For the upper campus, the Municipality of Beirut has proposed to build a multi-story parking garage in the location of the current HSBC surface parking on the Abdul Aziz street (495 – Ras Beirut). An automated parking system has been proposed for the area, where such a system is considered more efficient than a traditional design from physical, behavioral, and financial aspects. For the lower campus, this paper proposes that the poorly utilized Biel/Solidaire parking lots be used as satellite parking for AUB students during the day. Finally, an integrated system between the upper and lower campus locations will be proposed through the development of a mobile application.

I. INTRODUCTION

Lebanon is suffering from an uprising parking problem especially in the capital Beirut due to the increase in the rate of car ownership. Given the location of the American University of Beirut in one of the most congested areas of Beirut city, students have to waste a significant amount of time on a daily basis to find an available parking spot. The problem is that there are not enough parking places to meet the demand, thus the importance of designing a new parking facility.

Preliminary analysis indicates that the Hamra area has a shortage of around 900 parking spots according to Salameh, Abou Zeid and Kaysi in their report "Estimation of Parking Deficit in the Neighborhood of the American University of Beirut" [12]. In response to such estimates of the deficit, the Municipality of Beirut has proposed to build a multi-story parking garage, with a capacity of 450 spaces, in the location of the current HSBC surface parking on the Abdul Aziz street (lot 495 – Ras Beirut) represented in Figure 1, of an area of 1,860 Km^2. Efficient solutions to cover the deficit include building an automated parking garage (upper campus), and designing a satellite parking system between AUB and Biel (lower campus). For this reason, this paper intends to investigate the elements, potential and feasibility of introducing Tower Parking System into the proposed parking garage and to study the design and feasibility of a satellite parking system in the Biel area.

II. SCOPE OF WORK

The goal of this study is to build a connected and efficient parking system for AUB students and to come up with innovative ideas to reduce the parking crisis.

This will be achieved by designing an Automated Tower Parking System for the parking lot on the Abdul Aziz Street in order to increase the supply of parking spaces for AUB students.

Complementary to the design of the Abdul Aziz parking lot, the main contribution of this project is to further increase the supply of parking spaces through a satellite parking system from AUB to Biel. Students can park their cars in the Biel/Solidaire area where a rotation of scheduled shuttle busses is going to be provided to transport students from the latter location to the lower campus. Moreover, a mobile application will be built to manage the upper and lower AUB campus parking systems. This application will provide various services such as tracking the location and arrival time of shuttle busses as well as notifying the students about the remaining available spots in the Tower Parking System built on the Abdul Aziz street.

Figure 1. Satellite Image of Lot 495

III. UPPER CAMPUS

A. Methodology

The demand for the new Abdul Aziz parking was considered to be large to justify building the facility since providing a parking facility attracts more people to park in addition to the existing congestion in the Hamra area.

In order to build an efficient parking system, AUB students were surveyed to know their tendency to park in an automated parking garage instead of a conventional one. Moreover, different types of automated parking designs implemented in the Middle East were studied and three main alternatives were considered: Rounded Automated Parking as shown in Figure 2.A in Appendix A, Robotic Automated Parking and the Lebanese version of the Automated Tower Parking System. The main target is to select the strategy that will accommodate the largest demand considering the small area of the designed parking lot.

A comparison of the construction cost and time, the number of provided parking spots and the vehicle emissions for each alternative in contrast to the conventional parking system was done to study the economic feasibility of the project. Moreover, in order to avoid congestion of traffic in the area from Abdul Aziz Street to Hamra or Makdisi Street, a queue analysis for each alternative was done in terms of computing the number of queuing cars and the average waiting time for each car outside the parking facility. A scoring analysis was done in order to choose between the three studied automated parking designs and the conventional one. After looking at different aspects, the Tower Parking System was chosen to be designed in Lot 495 in Abdul Aziz Street.

B. Feasibility Study

1) Alternatives' Scoring

As previously mentioned, the project included researching various types of automated parking systems that are available in the Middle East and comparing it to the conventional parking system.

First, the capacity of the conventional parking in the given lot was estimated to be 450 spaces. Thus, the capacity parameter was kept constant across the other systems, in order to vary the cost and area footprint parameters. According to a study conducted by EITO&GLOBAL INC. [1] that is specialized in Round Parking systems, the required diameter for each tower is 20m[1], which eliminates this option when considering area (1,116 Km^2, which is 60% of the total area according to the Lebanese construction regulations) and altitude limitations. However, the robotic and tower systems can accommodate the requested capacity and using 50% of the area required by the conventional according to an article published by Robotic Parking Systems, Inc. [2].

Further comparison was conducted across the alternatives in terms of construction cost per space, queue analysis and reduction in emissions relative to the conventional parking.

Construction costs of conventional systems were obtained after conducting an interview with Mr. Ramadan Harb, the head of the transportation department at Khatib&Alami [3]. As per the queue analysis, the system was assumed to be $M/D/c^2$ assuming a Poisson distribution for the arrival patterns with a maximum rate equals half the capacity (225 vehicles/hour), based on a revision of a study conducted on the previous HSBC parking and a deterministic uniform service process.

Finally, since the project is considered a green project, a study of the emitted gasses, mainly CO_x and NO_x, was conducted in order to understand the effect of each system on the environment. The percentages obtained were relative to conventional systems according to results included in the "Green Garages" article [4]. The results of the conducted study are summarized in Table I.A in the appendix A.

2) Detailed Feasibility

After analyzing all the listed aspects above, the Automated Tower Parking system received the highest score, where a large weight (70%) was assigned to the cost and traffic impact analysis, since they represent the major constraints, followed by the environmental aspect (30%). In addition, this system's software and material are manufactured in Lebanon by Smart Parking Co. Thus, the project would be 100% local and thus supporting the Lebanese economy.

After choosing the most feasible option, a detailed feasibility study was performed regarding all incurred costs and expected revenues in order to identify the ROR period of the project.

Two major players are considered in this project, the municipality and the contractor. As agreed with Dr. Bilal Hamad, the head of the municipality of Beirut, the municipality has purchased the land and will incur the construction fee to the contractor in order to carry out the project. As for operation and maintenance, after conducting an interview with Smart Parking Co.'s primary lawyer, Mr. Ali Jaber [5], the company consists of operating and maintenance departments since the used code was privately encoded and won't be released to the public. The operation and maintenance services consist of providing specialized employees to make sure the system is functioning properly in addition to an inspective every 3 months accompanied with structural maintenance. Mr. Ali Jaber [5] indicated the fees for the latter services are represented by 20% of the income. Therefore, the revenue streams were first studied.

In terms of revenues, different scenarios were considered in terms of turnover rates. The fees that will be enforced are similar to the market prices: 3,000 L.L for the first 3 hours, an additional 1,000 L.L for each additional hour and a maximum of 8,000 L.L per day. The parking offers an automated car washing service in the bottom right spot which charges 10,000 L.L per service.

The parking spots, as previously mentioned, will be divided into AUB reserved spots and spots for the public. The AUB reserved spots (225 spots) will be subsidized by AUB. The monthly subscription for a spot costs 100$, where AUB will cover 50% of that amount and the student will pay the rest. We

---

[1] Built up area = 1116 m^2 (after eliminating set offs). Area/car is estimated to be 24m^2 [1], so 10,800m^2 is required for the suggested capacity.

[2] c = 15, number of servers

have discussed the subsidy issue with various professors, where they all agree that AUB has to take responsibility for the parking crisis and that such a solution is one of the most suitable for the current situation. Thus, the turnover for those spaces is considered to be equal to 1. However, AUB spots won't be available for students during the summer vacation (60 days), where during this period, the spots are available to the public. A realistic scenario was derived for the behavior of the public, where it was related to the HSBC parking in terms of turnover and parking durations that was covered in a previous FYP report in 2014 [6], where the results were extrapolated to represent the bigger capacity of the proposed system. Detailed results are represented in Table II.A in Appendix A. The revenue streams are described in the Figure 2.
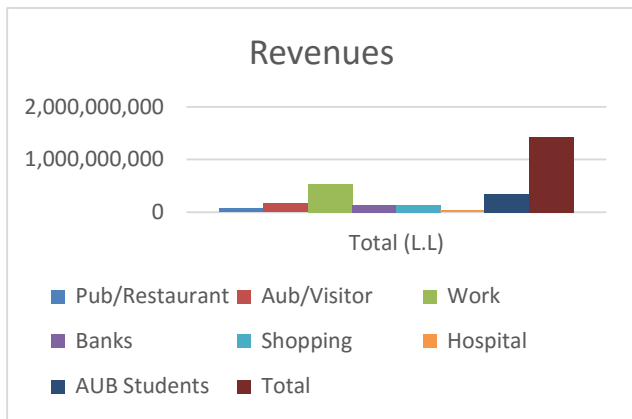

Figure 2. Revenues for Proposed Parking

The total obtained revenue is estimated to be 1,426,025,000L.L per year. In addition, on average, the washing service is estimated to generate 60,000,000 L.L/Year. Also, since the towers are going to be covered with shutters, the possibility of using the façade as a marketing platform was studied. Since the location of the parking lot is recognized as prime, advertisement billboards on the first row of towers facing the HSBC bank can generate around 45,000,000 L.L per year, according to PromoMedia, a local outdoors advertising specialist. Thus, the revenue is increased to $1,020,683/Year (using the current dollar-Lebanese pound exchange rate), which is divided 20% for the contractor and 80% for the municipality.

As per the costs, a cost-benefit analysis was conducted for both parties. The municipality has to incur the capital costs, cost of land and construction, license and permits, and security and maintenance that were obtained from the 2014 FYP report [6] as represented in Table I. The maintenance cost was obtained from Smart Parking Co. which charges 5,000$/tower for a structural inspection every 3months.

TABLE I
DIRECT COSTS TO MUNICIPALITY

Item	Unit Cost ($)	Total ($)
Cost of land	40,000,000	40,000,000
Construction Cost	13,000 (Per Space)	5,850,000
Total		45,850,000

O&M Costs		
License/Year		5,000
Security/Year		52,000
Maintenance	5,000	75,000

When looking at the project's feasibility from the municipality's aspect, the yearly profit is limited to 80% of the total revenue (816,546$). This results in a 50% return on capital investment after 40 year of operation, assuming 1.88% real interest rate (8% nominal interest rate; 6% inflation rate). The ROI period of the municipality is considered very long after finding the NPV of the revenue since the land was purchased as an asset. Regardless, the municipality target is focused on enhancing the transportation sector of Beirut and benefitting the public. It is worthy to note that after 40 years, the municipality will still own the land of an inflated estimated value of $130,480,000, assuming a property inflation rate of 2.7% according to Bank Audi records. As for the contractor, his revenue consists of the construction and annual maintenance fee; in addition to 20% of revenues for the operational services.

TABLE II
REVENUES AND COSTS FOR OPERATOR & MAINTENANCE

Revenue	Amount
Construction Revenue	$5,850,000.00
Construction Cost	$4,680,000.00
Construction Profit	$1,170,000.00
Operation Revenue/Year	$204,137.00
Operation Cost/Year	$80,000
Operation Profit/Year	$124,137
Maintenance Revenue	$75,000
Maintenance Cost	$45,000
Maintenance Profit	$30,000
Total Profit / Year	$154,137

The operational costs were assumed to be constant including the wages and benefits of 4 operational employees (2 employees/shift) and utilities that sum up to $50,000/Year as indicated by Mr.Ali Jaber. Therefore, the contractor obtains an amount of 154,137$ per year as profit from operational and maintenance services, which is a tolerable income as discussed with Mr. Ali Jaber.

C. Design
1) Material Used
The Tower Parking System is mainly made of steel. Steel columns and beams sections have a W shape and trusses have C shape. Galvanized steel was used in the upper floors to resist a wind speed of 137 km/hr. Bracings of equal angles are used to resist lateral load and galvanized bolts and anchor bolt were used at the base. Cover of the towers will consist of perforated steel or aluminum while the inside is copper.

2) Lot Physical Layout
The area of Lot 495 is 1,860 m² and is located in the area of Hamra, so in turn it is categorized as zone 3 by the Lebanese

Real Estate Law. No setbacks are required in the underground floors of the parking, and so the whole footprint area of the lot can be utilized. However, exploitation factors were taken into consideration in the upper and ground floors as per Lebanese regulations, thus the area in the latter floors used is 1,062 m^2 as mentioned in a previous FYP report in 2014 [6].

The setbacks were taken into consideration by designing circulation lanes within the lot. The route consists of two lanes, a continuous flow lane and a waiting lane. The continuous flow lane allows the vehicles to circulate inside the lot in order to arrive to the towers set at the back of the lot. In addition, the entry and exit of vehicles will as well be accommodated on this lane. Thus, all caused traffic will be restricted within the lot without spilling over to the adjacent street. As for the waiting lane, it allows customers to wait at the gate of a tower before being served, minimizing traffic inside the lot.

3) Parking System Design

The footprint of each tower is 50 m^2. The tower is composed of 2 parking areas separated by a vehicle lift. According to a review in 2011 [7], "Mixed Fortunes: Lebanon's automotive sector", the demand of SUVs in Lebanon adds up to 17%, and this percentage is expected to keep increasing due to the poor quality of roads across the country. Thus, all the lots of the tower system were designed to accommodate SUVs of an average weight of 2 tons. Each tower consists of 15 floors: 4 underground, 10 above ground and the ground floor. Taking into consideration the required setbacks of the lot, the land can accommodate 15 towers which adds up to 450 parking spaces. Daily and subscribed customers don't have a reserved lot; they can park in any available spot. As a first step, the model was built on ETABS, as represented in Figure 1.B in Appendix A, using the materials listed in the previous sections. Supports at the base of columns were assumed to be fixities while most of the rest of the connections were assumed as pins in order to withstand the lateral wind loads. After assigning all materials, the shear and moment diagrams were plotted indicating the stability of the structure.

After that, a model was built on AutoCAD in order to represent the parking's cross section and was later imported to ArchiCAD, where a 3D model of the tower with the circulation routes was rendered, as represented in Figure 3. Since the Lebanese law doesn't specify the features or required aspects of automated mechanical parking systems, the requirements of The National Parking Association's Guide to the Design & Operation of Automated Parking Facilities were implemented throughout the design.

Parking spaces (6) outside the towers, as represented in Figure 3 above will be provided inside the lot to accommodate vans and large vehicles that exceed the stall dimensions provided. In addition, green areas will be provided in order to maintain the environmental aspect of the garage.

Furthermore, the towers will be covered with solar panels on the sides that are subjected to sunlight whereas other faces of the tower will be covered with aluminum shutters. These faces will be used for advertisement billboards as previously discussed in order to increase revenue.
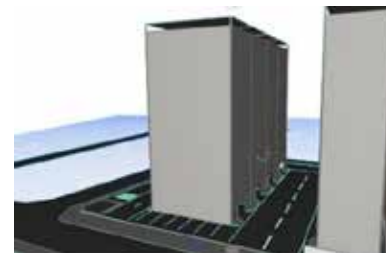
Figure 3. 3D ArchiCAD Rendered Model

IV. LOWER CAMPUS

A. Methodology

In order to implement a bus schedule for the Biel satellite parking to AUB new women's gate and vice versa, the demand for this parking was calculated, based on the following factors: (a) Biel choice (BC), which estimated the percentage of people willing to park in Biel [8]; (b) Car (C) and carpool (CP), which assessed the percentage of people who commute by driving or carpooling, entering Beirut to the Lower campus area [9]; (c) people switching (PS), which valued the percentage of people willing to switch from their previous parking spaces; (d) number of AUB students (NS). The demand is calculated based on this equation, which assumes that no demand will result from students who previously used non-car modes.

$$\text{Demand} = \% \text{ BC} \times (\% \text{ C} + \%\text{CP}) \times \% \text{ PS} \times \text{NS} \qquad (1)$$

In order to create a bus schedule, the following information was required: (a) average trip time, for peak and non-peak hours, i.e. the time needed for a bus to go from AUB new women's gate to Biel and vice versa, with consideration for bus loading and unloading time; (b) the daily partitioning of the students based on their class schedules, represented in the table below; (c) bus capacity.

The optimization of the schedule took into consideration the following inter-reliant factors: (a) headway (HW), associated with the average waiting time between the arrival of one bus to Biel and another, inversely proportional to student comfort; (b) driver comfort (DC), which estimated the hourly resting time allocated to the bus driver throughout the trips, calculated by averaging the difference in time between departure of a bus and its arrival to AUB, disregarding a 5-minute break after each trip; (c) number of two-way trips (NT); (d) number of buses (NB); (e) efficiency of operation (EO), related to the occupancy of the bus per trip, which is calculated by studying the ratio of bus occupancy to bus capacity throughout the trips per day. It should be maximal in accordance to the aforementioned factors.

Different schedule trials were prepared, but for the simplicity of this work, three trials were instigated, for a typical Monday: (a) in the 1st trial, a short bus headway was imposed, a total of two buses throughout the day were operated and driver's comfort was overlooked; (b) in the 2nd trial, a more flexible headway was allowed, two buses were operated till 10:00 a.m., and one bus till 7:50 p.m. and driver's comfort was still overlooked; (c) the 3rd trial was similar to the 2nd one; however, driver's comfort was imposed.

Figure 4. Demand factors results [1], [2]

B. Results

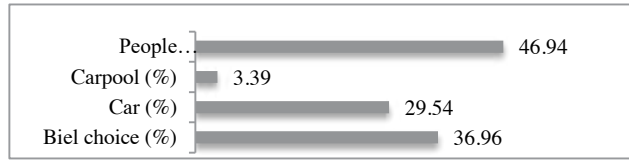The demand factors are represented in Figure 4. Using equation (1), with 8,300 AUB students [10], the daily demand was found to be 474 students. Note that this demand considers only students coming to Lower campus through Biel, and therefore isn't related to the demand of parking spaces for Upper campus area. According to Mr. Radwan Moukalled, the responsible of the parking lots in Biel, AUB students can access 230 spots out of 700 spaces. 3.39% of the students share their car (Fig. 4), and an average of 2.5 students in one car was taken [10]. Therefore, the daily actual number of students eligible for this satellite parking was 245, i.e. 51.68% of the daily demand.

For Monday, the partitioning of the students on an hourly basis was performed, from 8:00 a.m to 8:00 p.m., and shown in Table I in Appendix B. The results of the three trials aforesaid, in accordance with the schedule optimization factors are shown in Table III and Figure I, II in Appendix B. The number of buses before 10:00 pm, not shown in Table III, was two for the three trials, and cannot be lowered, since demand was very high: 24%, 25.8%, respectively, for the 8:00 and 9:00 am classes [10].

TABLE III
TRIAL 1,2,3 RESULTS

	NT	AUB: SC (min)	Biel: SC (min)	DC (per hour)	NB (10:00 am - 8:00 pm)
Trial 1	72	21	21	4	2
Trial 2	42	37	37	4	1
Trial 3	38	41	44	10	1

According to Table III, Trial 1's purpose was to show the negative impact of advantaging student comfort on other factors. Trial 2 and 3 showed that imposing driver comfort decreased the number of trips but decreased student comfort. Fig. I and II in Appendix B indicated that the efficiency of operation is the lowest for trial 1. From Biel to AUB (Fig. II), trial 2 and 3's efficiency of operation varied between 12:00 p.m. to 6:00 p.m., while from AUB to Biel (Fig. I), the efficiency of operation was the highest for trial 3.

Among the available options, trial 3 was the optimum solution proposed since it took into consideration all factors moderately,

and would provide the best financial outcome since it has the lowest number of trips, the least operated number of buses, advantaged driver comfort, while satisfying the hourly partitioned demand. However, student comfort was the least for this trial, since imposing a small headway would lead to tremendous increase in costs, which is perceived in trial 1.

C. Impact Analysis

The feasibility analysis of the satellite project was studied for two different scenarios. The first scenario entails AUB to buy a new bus, and the second scenario requires AUB to hire a bus operating company to operate the shuttle bus system. A cost benefit analysis for AUB was done to select the most profitable scenario, where AUB's benefits are limited to providing a better service for students.

1) Costs and benefits

The costs are different for the two scenarios. For the first scenario the main costs were the price of the bus, vehicle insurance, wage of the bus driver, workman and health insurance of the bus driver, and the vehicle operating costs (VOC). However, the cost to AUB incurred for renting a bus operating company, was only the fee set by the company for operating the shuttle system. The costs were discovered by several interviews done with respective people responsible for each cost such as insurance company, bus drivers, and others. The costs are summarized in Table I.C in Appendix C. However, the VOC was calculated by the following empirical equation:

$$VOC~(\$/100Km) = 192.702 + 490.389 \times e^{-0.086 \times speed~(kph)}~(2)$$

TABLE IV
COSTS SCENARIO 1 AND 2

Scenario	Cost	Value ($)
Scenario 1	Price of Bus per study period	65,000
	Vehicle Insurance per year	150
	Wage of Driver per month	450
	Insurance of Driver per year	840
	VOC per year	5,501
Scenario 2	Rent per year	27,000

There are four main benefits for this project: (a) travel time reduction*:* having an insured parking spot will reduce the parking search time. It is assumed that on average an AUB student spends between zero (already have a parking registration plan) to 30 (peak hour search time) minutes searching for an available parking spot, so a 15-minute parking search time was taken in this feasibility study. Also, it is assumed that the AUB student values an hour for 6.76$ [5]; (b) parking fee reduction: the students pay on average 4.7$ per day on parking. Since Biel is relatively far from AUB, the parking fee is 2$ per day, where the latter values were obtained upon interviewing cashiers at Biel parkings (c) reduction in emissions: according to the ministry of environment, the emissions by vehicles have two costs on the environment: the pollution cost, and the climate change cost shown in Table V [11]; (d) salvage value: according to the ministry of environment, Lebanon lacks exact standards to calculate the depreciation value of vehicles, and prices of used cars are randomly assigned so the American Standard will be used to estimate the Bus's salvage value [11]. The Modified Accelerated Cost Recovery System (MACRS) is shown in the

table found in appendix C, and the salvage value will be the difference between the initial price and the depreciation after 10 years.

TABLE V
CLIMATE CHANGE AND POLLUTION FOR DIFFERENT CAR SIZES

Renting a bus operating system instead of buying a bus will

	Small Car	Medium Car	Bus
Climate Change (c/veh.km)	2.3	2.5	2.8
Pollution (c/veh.km)	1.2	1.2	1.2

generate the same benefits but without the salvage value. The total benefits are summarized in Figure 5.

Figure 5. Total Benefits

2) Results

The results of the study show that it is more profitable for AUB to buy a new bus for it generates a higher net present value of $1,050,297 and more benefits, with a benefit to cost ratio of 7.1. The results are shown in table VII.

TABLE VII
SATELLITE PARKING FEASIBILITY RESULTS

Results	Scenario 1	Scenario 2
NPV Costs	$173,058	$225,722
NPV Benefits	$1,223,355	$1,222,154
Difference in NPV	$1,050,297	$996,431
Benefit to Cost Ratio	7.1	5.4

V. MOBILE APPLICATION

A mobile application restricted to AUB students was created. AUB students will have to register using their ID number and they will be provided with several services through this application. The scope of the application covers the upper and lower AUB campus. For the upper campus, the application mainly displays the operation hours for the automated parking system built on Abdul Aziz Street and notify the students about the remaining available spots in the parking. Moreover, students have the chance to reserve parking spots using the application and the number of available spots will be updated automatically. For the lower campus, the application mainly focuses on the schedule of the shuttle buses and tracks their location. This will notify the students about the remaining time for the bus to arrive to Biel or to AUB (New women's gate).

The program used to build the application is Android Studio and the programming language is Java and XML (eXtensible Markup Language). Due to space limitations, snapshots of the application will be later provided in the full report.

VI. CONCLUSION

In order to reduce the parking crisis AUB students face, two areas were approached: Upper campus area, and lower campus area. For the Upper campus area, three possible designs were studied for the "Abdul Aziz parking" in Hamra. The Lebanese tower automated parking was chosen, after a thorough feasibility study comparing the three alternatives. The chosen parking supplies 450 parking spots divided for AUB students and public use. As for the lower campus area, an optimized schedule of buses for a Monday, running from Biel to AUB new women's gate and vice versa, was proposed: two buses operating from 8:00 a.m. to 10:00 a.m., then one bus till 8:00 pm, with a total of 38 trips per day. A detailed economic impact was then performed comparing scenarios where AUB rents or purchases the bus. With a net present value of 1,050,297 $, AUB should buy the bus. For better management of both parking areas, a parking application "AUBSpot" was designed and created in order to facilitate for the students to track the location of buses and the availability of parking spots in "Abdul Aziz parking."

However, the study faces several limitations: Biel satellite parking solves part of the demand of parking spots for AUB students coming only towards Lower campus from the Biel area. It shifts the congestion out of our study area, but doesn't

REFERENCES

[1] EITO&GLOBAL INC. "Round Automated Parking System." EITO&GLOBAL INC.Web. 17 Apr. 2016.
[2] "When Is Robotic Parking Not Robotic Parking?" Robotic Parking. Web. 17 Apr. 2016.
[3] Harb, Ramadan. "Conventional Parking Costs." Personal interview. 13 Feb.2016
[4] "Terra Nova Ventures: Robotic Parking: Green Garage." Terra Nova Ventures: Robotic Parking: Green Garage. Web. 31 Mar. 2016. Green Garages
[5] Jaber, Ali. "Smart Parking In Beirut." Personal interview. 1 Mar. 2016
[6] Lot 495 Multimodal Parking Garage An Introduction of Sustainable Parking Supply. Rep. Beirut, 2014. Print. FYP Report
[7] Cochrane, Paul. "Mixed Fortunes: Lebanon's Automotive Sector." Back in Beirut, 2011. Web. 31 Mar. 2016
[8] Abboud, Charbel, Abu Hawili Anas, Mohamad Anas Bin Ahmad, Ghina Sabbagh, and Reina Tabbara.. Raw data. Lebanon, Beirut.
[9] M.Danaf, M. Abou-Zeid, and I. Kaysi, "Modeling Travel Choices of Students at a Private, Urban University: Insights and Policy Implications", Case studies on Transport Policy, Vol. 2, No.3, pp. 142-152 (2014)
[10] Naghi, Al, and Hani Abdul Rahman. "Evaluation framework for organization-based ridesharing: service design considerations and potential for AUB." (2014)
[11] Ministry of Environment, (2015). "Mobility Cost A Case Study for Lebanon".

Appendix A

TABLE I.A

FEASIBILITY STUDY ACROSS AUTOMATED ALTERNATIVES

Parking System	Cost of Construction/Space ($)	Queue Length (Vehicles)	Maneuvering Time (Mins)	Waiting Time To Enter Facility (Mins)	Percentage Reduction of Emissions (%)
Robotic	30,000 $	3.3	0.2	6	13.15
Tower	13,000 $	0.34	0.64	1.36	13.64
Conve ntional	25,000 $	1.04	3.3	0.3	0

Appendix B

TABLE I.B
HOURLY PARTITION OF THE DEMAND

Monday	Percentage (%)	Number of students to AUB	Percentages (%)	Number of students to Biel
8:00:00 AM	24	59	-	-
9:00:00 AM	25.8	63	0.5	1
10:00:00 AM	16.7	41	1.4	3
11:00:00 AM	10.8	26	3.7	9
12:00:00 PM	7.4	18	6.9	17
1:00:00 PM	3.7	9	8.3	20
2:00:00 PM	4.2	10	12.1	30
3:00:00 PM	2	5	15	37
4:00:00 PM	2.5	6	13.3	33
5:00:00 PM	2.5	12	13.9	34
6:00:00 PM	0.4	1	7.4	18
7:00:00 PM	0.2	0	11	27
8:00:00 PM	-	-	5.5	13
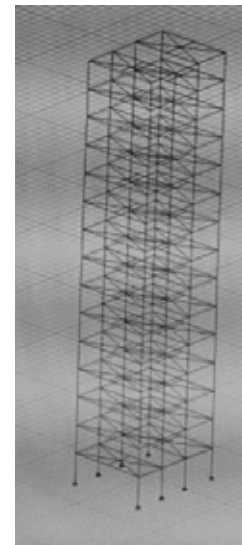
Figure 1.B ETABS MODEL

Efficiency of Operations
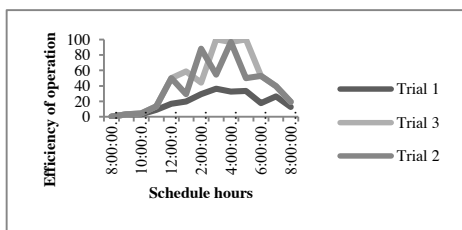
Figure I.B AUB to Biel efficiency of operation
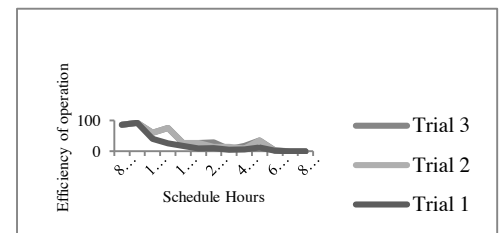
Figure II.B Biel to AUB efficiency of operation

Appendix C

TABLE I.C
SCHEDULE OF
BUSES

Schedule BIEL to New Women's gate		Schedule New Women's gate to Biel	
Departure time from BIEL	Arrival Time to AUB	Departure time from AUB	Arrival time to BIEL
7:30:00 AM	7:45:00 AM	-	-
7:35:00 AM	7:50:00 AM	-	-
8:30:00 AM	8:45:00 AM	-	-
8:35:00 AM	8:50:00 AM	9:10:00 AM	9:25:00 AM
9:30:00 AM	9:45:00 AM	9:50:00 AM	10:05:00 AM
9:35:00 AM	9:50:00 AM	10:10:00 AM	10:25:00 AM
10:30:00 AM	10:45:00 AM	10:50:00 AM	11:05:00 AM
11:10:00 AM	11:25:00 AM	11:30:00 AM	11:45:00 AM
11:50:00 AM	12:05:00 PM	12:10:00 PM	12:25:00 PM
12:30:00 PM	12:45:00 PM	1:10:00 PM	1:25:00 PM
1:30:00 PM	1:45:00 PM	1:50:00 PM	2:05:00 PM
2:10:00 PM	2:25:00 PM	2:30:00 PM	2:45:00 PM
2:50:00 PM	3:05:00 PM	3:10:00 PM	3:25:00 PM
3:30:00 PM	3:45:00 PM	4:10:00 PM	4:25:00 PM
4:50:00 PM	5:05:00 PM	5:10:00 PM	5:25:00 PM
5:30:00 PM	5:45:00 PM	6:10:00 PM	6:25:00 PM
6:30:00 PM	6:45:00 PM	6:50:00 PM	7:05:00 PM
7:10:00 PM	7:25:00 PM	7:30:00 PM	7:45:00 PM
7:50:00 PM	8:05:00 PM	8:10:00 PM	8:25:00 PM

Appendix D

TABLE I.D
PERCENTAGE DEPRECIATION OVER YEARS

Year	Percentage Depreciation
1	10
2	18
3	14.4
4	11.5
5	9.22
6	7.37
7	6.55
8	6.55
9	6.55
10	6.55

Analyzing Driver-Pedestrian Interaction in a Mixed-Street Environment

Hassan Obeid and Hoseb Abkarian
Department of Civil and Environmental Engineering
American University of Beirut
Hamra, Beirut
hassanobeid1994@gmail.com, hsa40@mail.aub.edu

Abstract - **This paper presents the design, analysis and results of a driving simulator experiment conducted to study the interaction between drivers and pedestrians in a mixed-street environment. 111 students of the American University of Beirut (AUB) from different majors and faculties participated in the experiment that took place in the Transportation and Infrastructure Laboratory in AUB. The study looked at the driver-pedestrian interaction from the driver's perspective, by quantifying the effects of different scenario variables on the driving behavior of the participants. One-way ANOVA shows that drivers' behavior in proximity of pedestrians tends to be significantly less aggressive when their approach velocity is lower, side street parking is not allowed, a crosswalk exists, and the number of pedestrians crossing the street is higher. The paper also derived a discrete choice model for the yielding behavior of the drivers as a function of different predictor variables. Five out of the six predictors considered (except for gender) had a significant effect on the yielding behavior, with high significance of side-street parking, number of pedestrians crossing, and approach velocity. This model can be used to predict the effect of different measures on the yielding rate of the drivers.**

The results of this study lead to a better understanding of the drivers' behavior and their interaction with pedestrians, and help planners propose and evaluate safety measures and traffic calming techniques to reduce the risks on the pedestrians.

I. INTRODUCTION

Pedestrians are one of the most unprotected road users. Hence, their safety should be addressed in all countries. WHO (2015) states that almost half of all traffic deaths involve pedestrians (22%), cyclists (4%) and motorcyclists (23%). In particular, pedestrian death constitutes 27% of all road traffic deaths in the Eastern Mediterranean part of the world. Moreover, NHTSA (2013) records a 73% pedestrian fatality occurring in urban areas compared to rural ones and a 69% occurring at non-intersections compared to intersections. These high percentages reflect the naturally unsafe behavior of drivers, and thus, confirm the need to study the different factors that contribute to a safer environment for pedestrians.

The WHO report on pedestrian safety (2013) examines some of the risk factors associated with pedestrian injury. The speeding behavior definitely affects the probability of a crash occurring and determines the accident's severity. Furthermore, the compliance of drivers with speed limits and the extent of enforcement of the laws regarding pedestrian crossing determine the type of driver-pedestrian interaction. Also, lack of pedestrian facilities and inadequate visibility due to obstructions increase the pedestrians' vulnerability to crashes.

This paper develops a methodology to study the driver-pedestrian interaction from the driver's point of view. Data was collected of 96 AUB students using DriveSafety's RS 600 driving simulator located in the Transportation and Infrastructure Laboratory of the American university of Beirut (AUB). These data points were used to study the effect of different experimental variables on the speeding behavior of drivers. Moreover, a binary logit model was constructed to capture the yielding behavior of drivers to pedestrians. Our aim of the model is to further understand the factors that affect the drivers' behavior and their interactions with pedestrians. This would help planners propose and evaluate safety measures to reduce risks on the pedestrians (crosswalks or other traffic calming techniques).

The remainder of the paper is organized as follows: Section 2 presents a review of the literature focusing on the driver's point of view in a driver-pedestrian interaction. Section 3 describes the methodology followed to collect the data and analyze it. Section 4 presents the results & analyzes them. Finally, section 5 concludes the paper.

II. LITERATURE REVIEW

This section summarizes some of the studies which are most relevant to this research. The studies presented focus on the drivers' speeding behaviors based on different driver-pedestrian interactions and on predictive models for driver yielding behavior at pedestrian crossings by using logistic regression.

A. Driver speeding behavior

Reviewing the pedestrian chapter in the *Highway Capacity Manual* (Transportation Research Board (2000)), the yielding compliance rate of drivers at crosswalks is expected to be 100%, since the pedestrians have the right-of-way on zebra crosswalks. However, many observational studies refute this assumption. Katz et al. (1975) studied factors that influence drivers' behavior in giving way to the pedestrians. Trained pedestrians performed the crossing of the street, while three observers recorded the variables of interest. Studying the data, crossing speed (average vehicle speed over the 20m stretch before the crosswalk) was significantly lower at marked

crosswalks compared to unmarked ones. Mean relative speed reduction was increased from 17% to 30% at one of the sites and 4% to 17% at the other. Crossing velocity was also significantly lower when the distance from the pedestrian to the vehicle was higher, when pedestrians were crossing in groups and generally lower when pedestrians were looking at the approaching driver. Furthermore, female drivers and older drivers slowed down more than other drivers.

Bertulis et al. (2013) studied the yielding behavior of drivers to pedestrians in marked crosswalks. The study was conducted in nine uncontrolled, marked crosswalks in two locations: Boston and Brookline, Massachusetts. Eight of the studied locations had similar geometry, lane configuration, parking and surrounding land use. Also, the experiment was staged so that all the drivers would face the same condition, where the trained pedestrian would only start crossing the street when the driver reached the stopping sight distance. The yielding rate ranged from 75% at 20 mph to 17% at 37 mph. Further analyzing the 85th percentile of approach speeds versus yielding to pedestrians, an inverse correlation was found, with an R^2 of 0.99, indicating a very strong correlation to low yield rates on high-speed roadways. This shows that drivers tend not to respect the pedestrians' right of way on marked crosswalks.

Bella et al. (2015) used a driving simulator to study the influence of several safety treatments at zebra crosswalks, combined with different driver-pedestrian interactions, on the speeding behavior of drivers and their yielding compliance. Curb extensions, parking restrictions, and advance yield markings were the safety measures studied. Moreover, driver-pedestrian interaction was categorized based on the arrival time of the vehicle to the zebra crosswalk at the moment the driver perceived the presence of a pedestrian. The results showed that curb extensions induced a significant change towards a safer speeding behavior; the distance from the crosswalk at which the driver reacted to the presence of a pedestrian, the minimum velocity during deceleration and the distance where the minimum velocity occurred were significantly higher for the curb extensions compared to the other safety treatments. Furthermore, the highest percentage of yielding was recorded for the curb extensions, and the lowest for parking restrictions.

B. Modeling driver yielding behavior
Himanen et al (1988) used multinomial logit modeling to study driver-pedestrian interaction on seven crosswalks at street junctions. The study was done in two different towns: Helsinki and Salo, Finland. Data was collected from videotapes recorded and variables of interest were extracted. Modeling the driving behavior, the authors found that the driver's probability of yielding to the pedestrian significantly increases as the number of pedestrians in the group and the distance of the pedestrian to the driver increase. On the other hand, the probability significantly decreases as the speed of the approaching vehicle (20 km away from the crosswalk), the number of cars in the platoon and the city size increase. More interestingly, the probability of yielding significantly increases

by 0.4 for pedestrians crossing the street on marked crosswalks rather than crossing the street elsewhere.

Sun et al. (2002) developed a binary logit model to capture the motorist yielding behavior at a two-way two-lane unsignalized midblock crosswalk, using a set of video cameras to collect the data. The studied crosswalk was on Springfield Avenue at the University of Illinois, Urbana Champaign, and data was collected during two different peak hours: the peak hour for vehicle traffic (16:30 – 18:00) and the peak hour for pedestrian traffic (11:30 – 13:00). The dependent variables taken into consideration to model yielding were the gender of the driver and his/her age, the type of the vehicle (car, SUV, truck or bus), the number of pedestrians waiting and the traffic flow condition on the opposite direction (stopped or not). The binary logit model predicted the highest probability of yielding for trucks and buses, since these vehicles have a longer braking distance, implying the cautious behavior of their drivers. Also, the model suggested doubling the probability of yielding when the number of pedestrians increased from three to six. Sun et al. intuitively explained this by fact that a bigger group quickly captures the attention of the driver. Moreover, older drivers showed a higher probability of yielding than younger ones. Finally, the opposite traffic condition had the highest coefficient in the utility function of yielding. The developed model was estimated using 900 data points and was used to predict the yielding behavior of other 354 data points. The model correctly predicted 87.1% (with 1% accuracy level).

Lastly, Schroeder et al. (2011) studied different factors affecting driving yielding at unsignalized, marked crosswalks and developed logit models from data collected at two sites in North Carolina. The two sites were chosen because of future crossing treatment installation plans, enabling the authors to look at the effect of the safety treatments on the yielding behavior of drivers. Pedestrian activated in-pavement flashing lights were installed at the first site while an in-road pedestrian warning sign was installed at the second site. Descriptive studies showed a significant increase in the yielding percentage at both locations. Then, binary logistic regression was used to model the yielding behavior of drivers. The developed model predicted an increase in the likelihood of yielding if the pedestrian was assertive in his/her approach to the crosswalk. Moreover, vehicle platooning decreased the likelihood of yielding, explaining the low overall yielding rate at the first site. Finally, the crossing treatments predicted a higher probability of yielding; however, the yielding probability was dependent on the activation of the pedestrian flashing lights at the second site.

III. METHODOLOGY

A. Data collection and procedure
The data collection phase for this project consisted of recruiting volunteers from the American University of Beirut to drive the experiment. Prior to starting the official data collection, a pilot was done with 7 participants to get initial feedback and unveil potential unexpected problems with the

scenarios. Participants were requested to give their overall feedback on the experiment, and were asked specific questions to see whether or not they observed the controlled scenario variables of interest. The experiment was then updated as necessary and the participants' feedbacks were addressed as required. The official data collection was then launched and 111 participants drove the experiment in total. Out of these data, 15 were removed later on during the data analysis phase for various reasons related to non-serious driving behavior, occurrence of accidents and crashes, and some participants failing to follow the direction instructions and reaching dead ends. The call for participants was mainly done by personal approach, through contacting friends and asking them to tell their friends. The rest of the participants were recruited through posting on university related social media groups and forums, and contacting students individually via email.

The experiment consisted of four phases: consent form and screening interview, test drive, actual experiment, and post-driving survey. In the first phase, participants were informed of the purpose of the experiment and were asked to sign a consent form confirming their willingness to participate. A screening interview was then conducted with the participant in order to determine whether or not he/she is eligible to participate based on several physical, mental, and health conditions. Afterwards, the participant was requested to drive the driving simulator for a 5-minute period in order to get familiar with it and make sure he/she does not suffer from any dizziness or discomfort while doing so. Then, participants were informed that the actual experiment will start shortly and that accordingly the data will be recorded. Participants were asked to drive as they would do normally under real life conditions. The driving experiment was in a suburban setting, where individuals drove on a sequence of streets before reaching the streets of interest: two streets, a university and a non-university setting, where scenario variables to be tested were controlled. Depending on the ID number which was assigned randomly, a participant will drive on these streets three times: twice on the university setting and once on the non-university setting, or vice versa. The scenario variables being controlled in the experiment are discussed in the subsequent section. After finishing the experiment, participants were asked to fill a survey which contained questions related to their general driving behavior, their attitudes to specific driving incidents, and their overall impression on the conducted experiment.
Table 1 summarizes the distribution of participants between genders and age groups.

B. Experimental design and data analysis
This section explains the methodology used in analyzing the data of the experiment. The dependent variables of interest are summarized in table 2. On the other hand, the independent

Table 1: Age and gender distributions

	Age	Gender
Average	20.271	79.12% Male
Standard Deviation	1.578	20.88% Female

variables controlled in the experiment are: presence or absence of crosswalks, street setting (university vs. non-university), number of pedestrians crossing, and the direction from which the pedestrians cross the street (which gives an indication of the effect of side street parking, and hence the visibility of the pedestrian, on the driver's behavior, since parking was only allowed on the left side). Another independent variable that was not controlled in the experiment but was taken into account is the Time To Zebra (TTZ), defined as the distance of the driver to the pedestrians when the pedestrians start to cross (65m, fixed) divided by the velocity of the driver at this distance. In other words, the TTZ represents the expected time for the driver to reach the crossing pedestrians if he/she continues driving at the same speed. The motivation behind adding this last variable is that the driving behavior of the driver and his/her interaction with the pedestrians will depend on his initial velocity prior to the interaction. Based on the pedestrians crossing speed controlled in the experiment, pedestrians need 5 seconds to reach the midpoint of the street. Consequently, a threshold TTZ value of 5 sec was chosen to indicate whether or not the driver needs to react to the pedestrians movement; if TTZ < 5 seconds, the driver (if he continues at the same speed) can cross the conflict point before the pedestrian reaches it, whereas if TTZ > 5 seconds, the pedestrian will reach the conflict point before the driver, and the driver will have to react by either braking to let the pedestrian cross, or accelerating to force his way before the pedestrian.

First, an exploratory descriptive analysis was performed on the data using one-way ANOVA to understand the effects of the different scenario variables on the dependent variables of interest. Confidence intervals were then plotted for the variables that have significant effects on the dependent variables to better understand their effects. Following the descriptive statistical analysis, a discrete choice model was developed to model the yielding behavior of the drivers. Since the decision making process of the driver has two possible outcomes (to yield or not to yield), it can be captured using a binary logit model. A further discussion of this model is presented in the next section.

Table 2: Dependent variables of interest

Dependent variable	Description
Distance at reaction (D_r)	The distance at which the first reaction occurs
Minimum velocity (V_{min})	The minimum velocity that occurs along the 65m stretch before the crosswalk
Average crossing velocity (V_{20})	The average velocity of the driver over the 20m stretch before the crosswalk
Deceleration rate (Dec)	The difference between V_r^2 and V_{min}^2 over the distance $2(D_r - D_{min})$
Yielding	The driver lets the pedestrian pass

IV. RESULTS

This section presents the results of the descriptive statistics as well as the development of the discrete choice model for the yielding behavior of the drivers.

A. Descriptive statistics

A one-way ANOVA was performed to study the effects of the independent variables on the metrics of interest. Table 3 summarizes the results of the test. To further interpret the ANOVA test results, 95% confidence intervals were plotted for the variables that had significant effects on the response metrics (see Appendix A). The plots show that the distance at which reaction occurred, either by pressing the brakes or releasing the gas pedal, was significantly higher in the scenarios that had installed crosswalks (p=0.000), had three pedestrians crossing rather than one (p=0.039) or had pedestrians crossing from the right where no visibility obstructions were present (p=0.001). This result presents the effectiveness of the crosswalk to decrease the reaction time of a driver. It is also a clear indication of how well the driver perceives the experimental situation and hence, the possibility of a conflict arising between him/her and the pedestrian(s). Hence, when the number of pedestrians was higher or no visibility obstructions were present, the driver started slowing down earlier. Also, the minimum velocity was significantly lower for the scenarios that had three pedestrians crossing the street (p=0.001) or were of a non-university setting (p=0.008).

This result presents the cautious behavior of the driver if he/she perceives the pedestrians earlier if a group of pedestrians were crossing, thus reducing the vehicle's velocity significantly until the pedestrians cross the street safely, then continuing his/her way. More interestingly, the driver tends to be more cautious in non-university street settings. Although not expected, one possible explanation can be the fact that drivers expect pedestrians themselves to be more cautious when they are crossing university-setting streets, leading to drivers' tendency to rely solely on the pedestrians' cautiousness. Furthermore, the average crossing velocity was significantly lower for the scenarios that had 3 pedestrians crossing (p=0.004); this result confirms that of Katz (1975). The average crossing velocity was also significantly lower were at the non-university street setting (p=0.020). These results have similar justifications with those of the minimum velocity recorded along the street. As for the deceleration rate (absolute value), it was significantly higher when drivers had TTZ values less than 5 seconds (p=0.000). This is expected, since drivers who had low TTZ values were associated with high initial speeds. Therefore, these drivers needed higher deceleration rates to yield. It is also important to mention that the TTZ value, which is an indication of approach speed, had a significant positive effect on all other response metrics, confirming the importance of following posted speed limits to insure the safety of pedestrians.

Table 3: ANOVA test results

p-values	Crosswalk	Number of Pedestrians	Street Setting	Direction of Crossing	TTZ
Distance at reaction (D_r)	0.000	0.039	0.349	0.001	0.004
Minimum velocity (V_{min})	0.321	0.001	0.008	0.058	0.000
Average crossing velocity (V_{20})	0.082	0.004	0.020	0.097	0.000
Deceleration rate (Dec)	0.293	0.576	0.678	0.939	0.000

B. Driver yield model: binary logit

A discrete choice model was developed to model the yielding behavior of the drivers as a function of several explanatory variables. As mentioned earlier, a binary logit model was chosen to capture the decision making process of the drivers, who are faced by two alternative decisions: to yield or not to yield. This type of models has been widely used by researchers to model behavior. For instance, the gap acceptance behavior of pedestrians, where the choice set also has two alternatives, was frequently modeled using binary logit models in the literature (Rao (1993), Pant et. al (1994), Brewer et. al (2005), Danaf et. al (2016)). In such models, the probability of choosing an alternative is dependent on its utility function, defined as follows:

$$U_i = \beta_{0i} + \beta . X + \epsilon, \qquad (1)$$

Where:
- U_i is the utility of alternative i
- β_{0i} is the alternative specific constant for alternative i
- X is a vector of independent variables (predictors)
- β is a vector of coefficients of the independent variables
- ϵ is a random error term

The probability of choosing alternative i, P_i, is thus obtained from the utility function defined above using the following formula:

$$P_i = \frac{1}{1+e^{-V_i}}, \qquad (2)$$

Where:
- $V_i = U_i - \epsilon$ is the systematic utility.

Since only two possible outcomes (i and j) exist in binary logit models, the probability of choosing alternative j, P_j, is given by:

$$P_j = 1 - P_i. \qquad (3)$$

In this experiment, the independent variables considered were the following:

- *Cross*: 1 *if there is a crosswalk and* 0 *if there is no crosswalks.*
- *Setting*: 1 *if university setting* 0 *otherwise*
- *TTZ*: 1 *if TTZ* > 5*s and* 0 *otherwise*
- *NumPed*: 1 *if three pedestrian crossing*, 0 *if one pedestrian is crossing*
- *Direction*: 1 *if pedestrians are crossing from Left to Left*, 0 *otherwise*
- *Gender*: 1 *if male*, 0 *if female*

Notice that all the variables above are categorical.

As such, the utility function of the yielding alternative has the following form:

$$V_{yield} = \beta_0 + \beta_1 Cross + \beta_2 NumPed + \beta_3 Direction + \beta_4 Setting + \beta_5 TTZ + \beta_6 Gender. \qquad (4)$$

Since each participant is observed over three streets, the data obtained contains multiple observations from the same person. This type of data, also called panel/agent data, can lead to unobserved correlation between the choices made by the same person due to some common unobserved individual trait. For instance, if a driver has a tendency to stop for pedestrians, his/her choice over the three streets will be dependent on this tendency and the resulting outcomes are thus not independent. One way to capture this dependency is to add to the utility function an individual specific random error term that is fixed for each individual across the three streets. This term is random and is assumed to have a normal distribution with a mean of 0 and a standard deviation (sigma) that will be estimated during the model calibration. Including this term and estimating its standard deviation will correct for the correlation of the choices made by the same person, if such correlation exists. If the estimated standard deviation was not significantly different from 0, the agent effect is considered insignificant and the error term can be removed. The model estimation was done in R, using the package "pglm" (panel generalized linear models), and the output is shown in the table below, with an explanation of the meaning of each estimate obtained and what does it reveal about the driving behavior.

Table 4: Binary logit model results

	Estimate	Std. error	t-value	p-value	Explanation of the estimate
Intercept	-2.08	1.12	-1.85	0.06	The negative value of the intercept means that, everything else being the same across alternatives, a driver is more likely not to yield than he is to yield. The p-value (0.06) of this alternative shows that this estimate is almost significant on a 5% significance level.
Cross	1.39	0.71	1.95	0.05	The positive value of the crosswalk coefficient indicates that a driver is more likely to yield when there is a crosswalk. This estimate is significant on a 5% significance level.
Setting	1.35	0.68	1.98	0.04	The positive value of the Setting coefficient indicates that driver is more likely to yield when he/she is driving on a non-university setting. This result, although unexpected, is in line with the results of the descriptive analysis which showed that the drivers tend to be less aggressive on the non-university setting. This estimate is significant on a 5% significance level.
TTZ	4.62	1.26	3.67	0.00	The positive value of the TTZ coefficient estimate shows that the driver is more likely to yield when his/her approach velocity is lower, and hence his/her TTZ is higher. This is because the driver entering the street with a lower velocity has more time to perceive the pedestrian and react by yielding, whereas a driver with a high approach velocity will not have as much time to make a stop and let the pedestrian cross. This estimate is significant on a 5% significance level.
NumPed	2.22	0.81	2.73	0.00	The positive value of the NumPed coefficient estimate shows that the driver is more likely to yield when more than one pedestrian is crossing the street which indicates that drivers tend to be more cautious when more pedestrians are involved. This estimate is significant on a 5% significance level.
Direction	3.99	1.06	3.75	0.00	The positive value of the Direction coefficient estimate shows that the driver is more likely to yield when pedestrians are crossing from the right where no parking is allowed in this experiment. This shows side street parking has a significant effect on increasing the risk of pedestrians crossing by limiting the visibility of the pedestrians trying to cross the street, making it less likely for the driver to yield. This estimate is significant on a 5% significance level.
Gender	-1.05	0.89	-1.18	0.23	The negative sign of the Gender coefficient estimate indicates that male drivers were less likely to yield than women drivers. This result, although not significant on a 5% significance level, can indicate that the driving behavior of men tends to be more aggressive than that of women.
Sigma	2.83	1.05	2.67	0.00	The value of the standard deviation Sigma of the individual specific error term has no important indication regarding the driving behavior. However, the fact that this term is significant indicates that the agent effect exists and that there is a correlation between the individual's choices across the streets. As such, this term cannot be removed from the model in order to capture this correlation and obtain accurate estimates of the βs.

The model derived above can be used to predict the probability of yielding of a certain pedestrian when faced with a specific scenario. For instance, a male driver driving on a university setting with no crosswalk and having an approach velocity of 40 km/h has a probability of yielding of 0.63 when faced with one pedestrian crossing from the side where street parking is not allowed.

V. CONCLUSION

This research paper aimed at studying the different variables that determine the driver-pedestrian interaction using a driving simulator and developing a discrete choice model to describe and predict the yielding behavior of drivers near pedestrians.

The use of a driving simulator in such experiments offers great advantages, specifically in allowing a high level of control over the variables being tested while causing no harm or threats on real pedestrians' safety.

The most interesting findings of this study were the statistically significant positive effects on the driving behavior (and hence pedestrians' safety) of prohibiting side street parking, adding crosswalks, and reducing the approach velocity of the drivers near pedestrian crossing locations. When these conditions are present, drivers were found to yield at higher rates and had lower minimum and crossing velocities near pedestrians. As such, planners are advised to direct traffic calming techniques so as to ensure these conditions are satisfied.

The binary logit model developed in this paper can be a very useful tool in evaluating these traffic calming techniques. The statistical significance of the estimated coefficients can produce reliable estimates of the yielding rates expected from different scenarios. It is advised, however, that the model is tested on a bigger data set prior to using it in evaluation.

REFERENCES

[1] World Health Organization, 2015. Global Status Report on Road Safety, Retrieved from www.who.int/violence_injury_prevention/road_safety_status/2015/G SRRS2015_Summary_EN_final.pdf on April 2016

[2] National Center for Statistics and Analysis. (2015, February). Pedestrians: 2013 data. (Traffic Safety Facts. Report No. DOT HS 812124). Washington, DC: National Highway Traffic Safety Administration.

[3] World Health Organization, 2013. Pedestrian Safety: A Road Safety Manual for Decision-Makers and Practitioners. Retrieved from http://apps.who.int/iris/bitsream/10665/79753/1/9789241505352_eng.pdf on April, 2016.

[4] Highway Capacity Manual (HCM) 2010, National Research Council, Transportation Research Board, Washington, DC, 2010.

[5] Katz, A., Zaidel, D., & Elgrishi, A. (1975). An Experimental Study of Driver and Pedestrian Interaction During the Crossing Conflict. *Human Factors: The Journal of the Human Factors Society, 17*(5), 514-527.

[6] Bertulis, T., & Dulaski, D. (2014). Driver Approach Speed and Its Impact on Driver Yielding to Pedestrian Behavior at Unsignalized Crosswalks. *Transportation Research Record: Journal of the Transportation Research Board, 2464*, 46-51. doi:10.3141/2464-06.

[7] Bella, F., & Silvestri, M. (2015). Effects of safety measures on driver's speed behavior at pedestrian crossings. *Accident Analysis & Prevention, 83*, 111-124. doi:10.1016/j.aap.2015.07.016

[8] Sun D., Ukkursi S., Benekohal R., Waller S. (2012). Modeling of Motorist-Pedestrian Interaction at Uncontrolled Mid-block Crosswalks, unpublished.

[9] Himanen, V., Kulmala, R. (1988). An Application of Logit Models in Analyzing the Behavior of Pedestrians and Car Drivers On Pedestrian Crossings. Accident Analysis and Prevention20, 187-197.

[10] Schroeder, B. J., & Rouphail, N. M. (2011). Event-Based Modeling of Driver Yielding Behavior at Unsignalized Crosswalks. *Journal of Transportation Engineering J. Transp. Eng., 137*(7), 455-465. doi:10.1061/(asce)te.1943-5436.0000225.

[11] Rao, S.B.S., (1993) "Sequential binary choice models for urban travel to work," *Hwy. Res. Bull.*, Vol. 48, Indian Roads Congress, Hwy. Res. Board, 91-107.

[12] Pant, P.D. and Balakrishnan, P., (1994) "Neural Network for Gap Acceptance at Stop-Controlled Intersections", *ASCE Journal of Transportation Engineering*, Vol. 120(3), pp.432-446.

[13] Brewer, M., Fitzpatrick, K., Whitacre, J., Lord, D., (2005). "Exploration of Pedestrian Gap Acceptance Behavior at Selected Locations". Submitted for Transportation Research Board, Washington, D.C.

[14] Danaf, M., Sabri, A., Abou Zeid, M., and Kaysi, I., (2016) "Pedestrian-vehicular interactions in a mixed street environment", Submitted for the World Conference on Transportation Research - WCTR 2016, Shanghai, China.

APPENDIX A
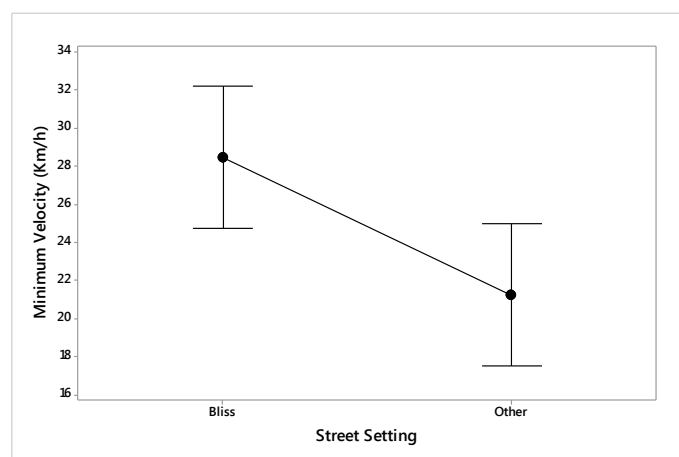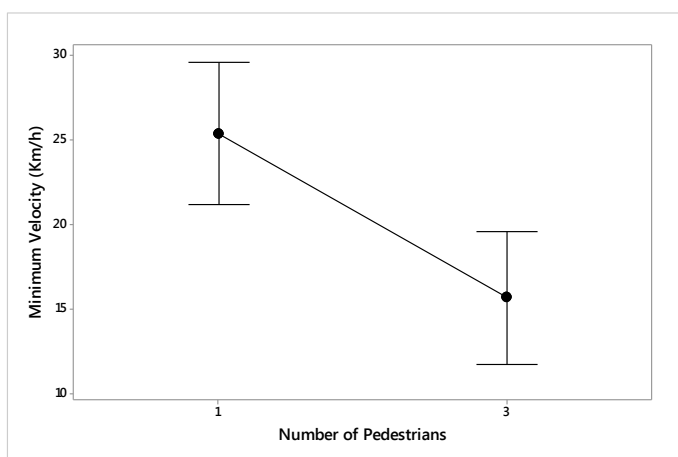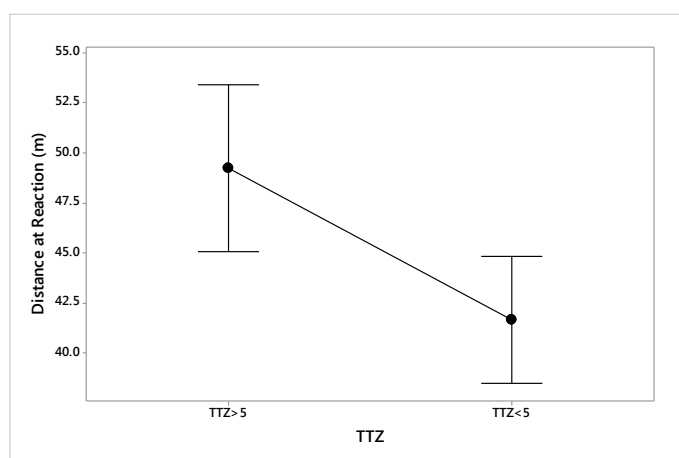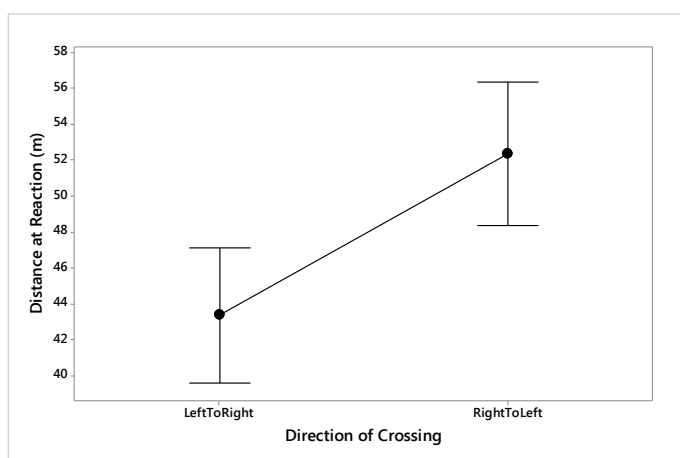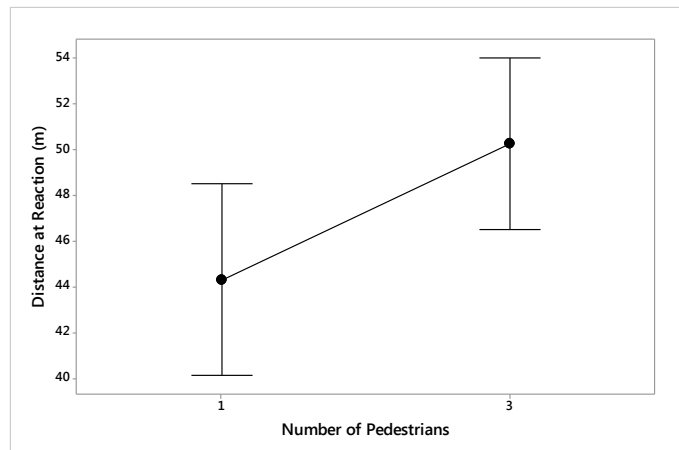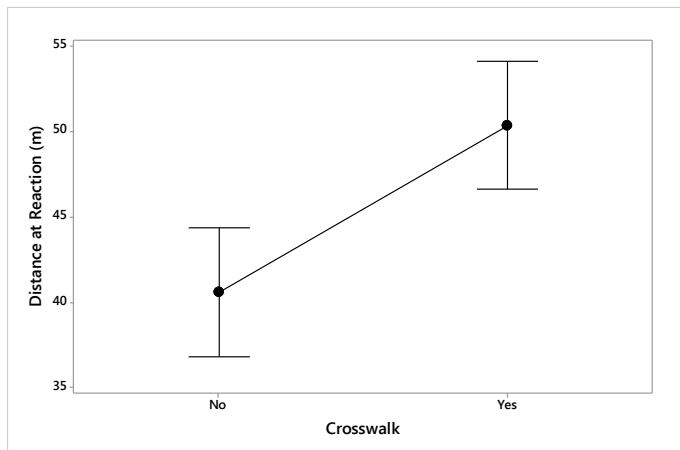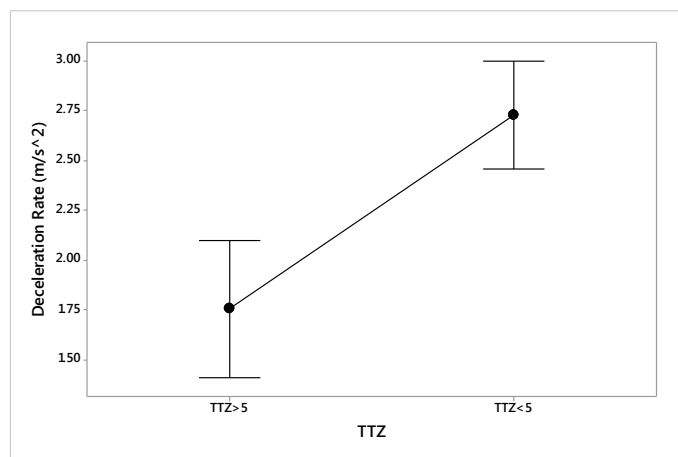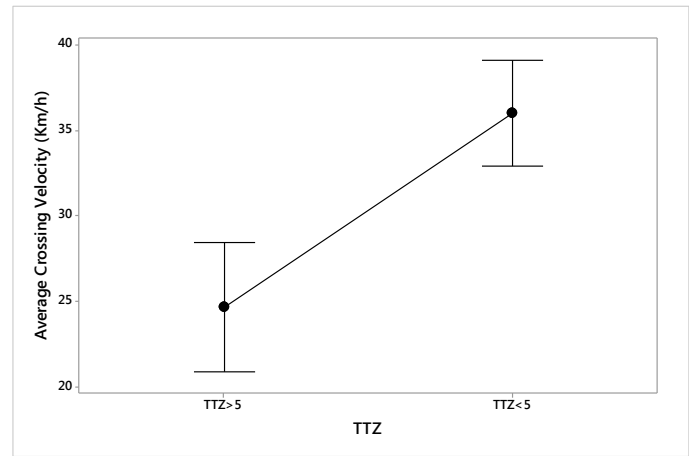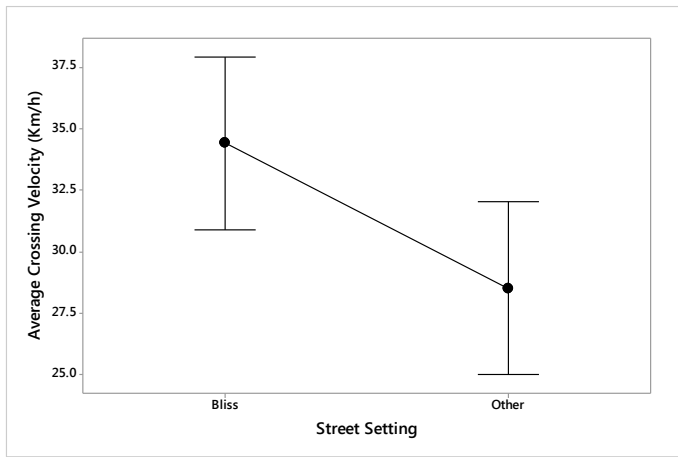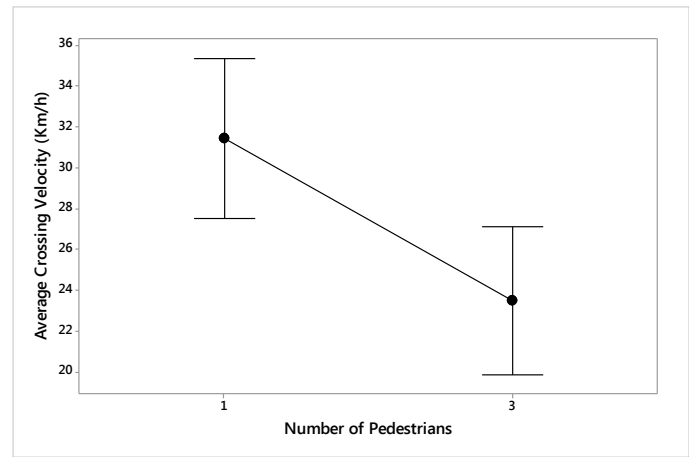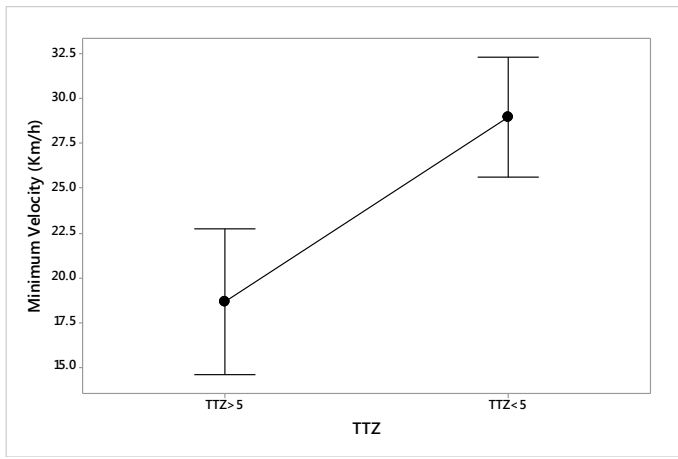
Assessment of the constructed wetland in Remhala and Feasibility of upscaling it

Assaad El Assaad[#1], Nour Chamseddine[#2], Raed Moustafa[#3], Nourhan Nassar[#4], Omar Tayyan[#5]

[#]Civil Engineering Department, American University of Beirut

Riad El-Solh - Beirut, Lebanon

[1]ase20@mail.aub.edu

[2]nwc06@mail.aub.edu

[3]rmm63@mail.aub.edu

[4]nmn22@mail.aub.edu

[5]oht01@mail.aub.edu

Abstract— **This paper presents the results of a study on a wastewater treatment plant which operates on the principle of wetland treatment using beds of reed, implemented in the village of Remhala, Lebanon. Furthermore, it displays the conclusions reached when considering the feasibility of implementing the same plant on a bigger scale. As a first step, water samples were taken from the plant in Remhala, at three stages: inflow, midsection and outflow. These samples were analyzed and tested in the laboratory and yielded an extremely high BOD, inconsistent with normal domestic wastewater, which hindered the proper functioning of the plant. Throughout this research, continuous sampling was done to further understand the obstacles faced by the plant, and to be better able to propose effective solutions to them. Preliminary results show some defects within the existing treatment plant, mainly due to evident anaerobic conditions in the sedimentation tanks as well as a diversion in the flow away from the reed bed. Moreover, the lack of management and maintenance by village officials may have further contributed to the poor conditions.**

Keywords—Wetland; Wastewater treatment plant; Reed beds; BOD; COD.

I. INTRODUCTION

Wastewater generated from domestic use in most Lebanese villages was habitually disposed of without any treatment, often into poorly designed septic tanks. However, after the end of the Lebanese war in the 1990s, the economic growth and progress in housing patterns led to a greater need for wastewater treatment systems. Indeed, the conventionally dispersed villages were renovated into denser modern-day communities and the wastewater discharge became much more significant. Therefore, proper sanitation and treatment were required to protect public health. Over the past sixteen years, remarkable progress has been made in building a link between the Lebanese population and wastewater collecting systems in all Regional Water Authorities (RWAs) in Lebanon. However, while evaluating the effectiveness of wastewater management along with these achievements, the Effective water management (EWM) indices performed poorly with a score range between 19.6 % (BML) and 27.6 % (North Lebanon) in 2014 [1].
In this paper, a promising environmental project is assessed which could prove a viable alternative to existing non-effective plants. The wastewater treatment plant in question is located in Remhala, a small Lebanese village. The plant encompasses is a constructed wetland system, which utilizes natural reed beds to treat domestic wastewater. If functioning properly, the wetland could offer an energy and almost cost free solution. The mechanism which the plant operates under will be further explained in the following section.

II. BACKGROUND AND LITERATURE REVIEW

A. Site study

The constructed wetland used in this research is in the small village of Remhala, located in Aley, which occupies an approximate area of 4.53 km2. This wetland was designed by engineer Rached Sarkis to treat domestic wastewater generated from the small population of the village (420-450 people). The constructed wetland in Remhala is very basic. Wastewater from the sewage network is first collected at an elevated storage tank and then flows to a sedimentation tank with a sand/gravel filter. After primary treatment, the water is discharged into the wetland cells where it is undergoes further treatment in the reed beds. Sewage pipelines are only connected to approximately 100 households with wastewater flowing through by the sole force of gravity; very minimal pumping is needed, since only about 10 households reside at a lower elevation than the plant.

B. Literature review

In comparison to neighboring countries, Lebanon has relatively abundant water resources. It is the only Arab country that doesn't contain a desert and does not suffer from extreme droughts. However, mismanagement often leads to a constant deficit in domestically usable water. Most resources go to waste or are inefficiently treated and hence deemed non-potable.

Contemporary literature offers examples of initiatives to install wastewater treatment plants in Lebanon. However, most of these have been carried out (or at least planned) in the "traditional" way, which has usually not been very effective.

This study is particularly interested in a specific, newer, type of wastewater treatment plants: constructed wetland. These are not widespread in Lebanon, to say the least, but a few examples were found.

One interesting project is the ''Litani River Constructed Treatment Wetland''. The project was conceived by the Litani River Basin Management System (LRBMS) program. Its main purpose was to treat polluted water from the Litani River, to which it is adjacent. It would also restore riparian (relating to wetlands near rivers) habitat and raise environmental awareness. The wetland was designed to be approximately 2.5 ha in size, receive 30 L/s flow in the dry season and 60 L/s during the rest of the year. Moreover, it was expected to remove between 30% to over 90% of pollutants coming into the river (the rate depends on the nature of the pollutant and the time of the year). The system is supposed to consist of a basin of deep and shallow zones; 2-3 m and 30-50 cm respectively, with appropriate vegetation for each zone. The design is such that minimal maintenance is required; this is an important trait in constructed wetlands, as it significantly reduces costs. The project was completed in April 2014.

Another completed project is a constructed wetland in Garzouz-Jbeil, funded in its early stage by the American University of Beirut and later on by the convent (where the plant was actually implemented) present in Garzouz and the University's environmental laboratory for water analysis. The project was initiated in 1992 and implemented in 1999, after undergoing several evolutionary steps. The project's research paper states that the possibility of implementing a similar project in the Cedars resort (North of Lebanon) and an educational pilot size system in Burj- al-Barajneh (south suburb of Beirut) was being investigated at the end of the Garzouz project. However, no evidence of a follow-up appears in the literature.

III. SCOPE AND METHODOLOGY

Figure 1. Schematic representation of a constructed wetland

The wastewater flows, via sewage pipes, to a storage tank where it is collected and then driven into a sedimentation tank. There, the water undergoes primary treatment to remove large solids, grease, and oils. The effluent is then passed

through a gravel/sand filter to trap any remaining solids before it reaches the reed bed. The main mechanism of the reed bed system operates on what is formally known as a "Gravel Bed Hydroponic System", which is comprised of a natural reed bed planted on top of a thick layer of gravel. "A reed bed is essentially a basin that is lined with an impermeable membrane, filled with gravel and planted with macrophytes such as reeds and rushes. Wastewater (black or grey) passes through the root zone of the reeds where it undergoes treatment via physical, chemical and biological interactions between the wastewater, plants, micro-organisms, gravel and atmosphere. The reed plants leak small amounts of oxygen out through their roots, creating small oxygenated sites within an otherwise anaerobic environment. This mix of aerobic and anaerobic conditions creates an ideal environment for the growth of microorganisms on the surface of the gravel and plant roots. These microorganisms are largely responsible for the pollutant removal that occurs in a reed bed, as they feed on and breakdown organic matter and nutrients, and compete against pathogenic organisms" [2]. Inlet and outlet pipes are usually positioned below the gravel surface, so that the water always remains below the surface, thus minimizing the risk of human exposure to the wastewater and foul odors. The system can keep working with a continuous flow and no artificial thermal insulation.
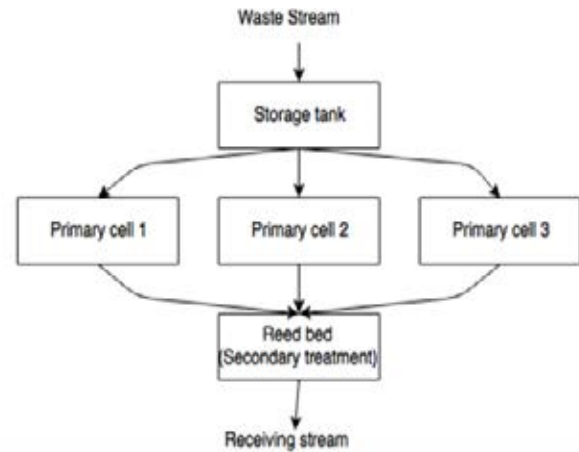
Figure 2. Diagram representing Remhala's wastewater treatment plant

Waste-stream characterization

Throughout this study, the main method used for collecting data was primary research. Several samples were collected from the treatment plant in Remhala. Four samples were obtained during each visit: from the inlet pipe, from the ponds, from the reed bed and after treatment. The sample from the inlet pipe was taken in order to test the water directly exiting the households. Moreover, water collected from the ponds appeared to be inconsistent with domestic wastewater and hence an issue with the functionality of the ponds arose.

Samples of influent and effluent water were taken with varying lag times during each visit in order to evaluate the wetland system's performance. With the help of the laboratory assistants at the American University of Beirut (AUB), conventional water quality parameters including BOD5, TSS, COD and TF were analyzed by Standard Methods [3], along with TDS, Nitrate, Nitrate-Nitrogen and Sulphate. The BOD test was based on a 5-day incubation, at a temperature of 20°C.

All samples were kept at a temperature of 4°C from time of retrieval until testing procedures. In addition, measurements of the ponds and their surroundings were taken. Subsequently, the plan of the wastewater treatment plant was drawn using AutoCAD.

IV. RESULTS AND DISCUSSIONS

Note: in the following tables, 1 stands for before arriving to the collecting pond, 2 stands for the first pond, 3 stands for samples from the reed bed and 4 stands for after treatment.

Tc	Too numerous to count	Too numerous to count	Too numerous to count	Too numerous to count
Fc	Too numerous to count	Too numerous to count	Too numerous to count	Too numerous to count

TABLE I: SAMPLES FROM FIELD VISIT 1

Samples Tests	1	2	3	4
COD (mg/L)	1230	2510	380	274
BOD	194	99	83	66
TSS(mg/L)	250	250	110	60
TDS(ppm)	959	516	500	495
NO3(mg/l)	110.6 25-N	92.6 20.9-N	80.9 18.3-N	48.6 11.0-N
SO4(mg/l)	33	34	54	53
T.C	Too numerous to count	Too numerous to count	Too numerous to count	Too numerous to count
F.C	Too numerous to count	Too numerous to count	Too numerous to count	Too numerous to count

TABLE 2: SAMPLES FROM FIELD VISIT 2

Samples Tests	1	2	3	4
COD (mg/L)	400	460	220	270
BOD	238	413	290	466
TSS(mg/L)	63	360	180	420
TDS(ppm)	510	560	534	294
NO3(mg/l)	47.8 8-N	24.7 5.6-N	30.0 6.8-N	8.7 2.0-N
SO4(mg/l)	68	0	52	14

TABLE 3: SAMPLES FROM FIELD VISIT 3

Samples Tests	1	2	3	4
COD (mg/L)	-	421	410	358
BOD (mg/L)	-	159	74	59
TSS(mg/L)	-	7525	540	700
TDS(ppm)	-	277	257	265
NO3(mg/l)	- -	111.8 25.2-N	43.2 9.8-N	8.1 1.8-N
SO4(mg/l)	-	11	45	50
Tc	-	Too numerous to count	Too numerous to count	Too numerous to count
Fc	-	Too numerous to count	Too numerous to count	Too numerous to count

TABLE 4: SAMPLES FROM FIELD VISIT 4

Samples Tests	1	2	3	4
COD (mg/L)		165	262	373
BOD (mg/L)		1194	148	56
TSS(mg/L)		440	60	13.3
TDS(ppm)		529	497	451
NO3(mg/l)		6.8 1.6-N	83.6 18.8	60.8 13.6-N
SO4(mg/l)		2	8	1
Tc		Too numerous to count	Too numerous to count	Too numerous to count
Fc		Too numerous to count	Too numerous to count	Too numerous to count

The tests performed give some preliminary results. Prior to laboratory testing, it was evident that the samples taken from the ponds were darker in color and more polluted than their counterparts at the inlet pipe.

The data obtained from the laboratory tests confirmed the initial observations. Samples taken from the pond displayed a very high BOD (of the order of 1000 mg/L) whereas "normal" domestic wastewater is expected to yield a BOD not exceeding 200 or 300 mg/L. Based on criteria for surface water

discharge, the secondary treatment standard for BOD has been set at 30 mg BOD/L (i.e. 30 mg of Oxygen are consumed per liter of water over 5 days to break down the waste). The treatment plant reduced the initial BOD values considerably, however this reduction in BOD was not enough to meet the standard for domestic non-palatable use or even for irrigation purposes. This is due to the existing anaerobic conditions in the septic tanks caused by the lack of aeration, as well as eutrophication (which leads to algal growth and further oxygen depletion). These anaerobic conditions caused a change in color (black) and a foul smell and finally raised the BOD.

Too much BOD can cause the excessive growth of bacteria in the ponds. If the BOD is higher than a certain threshold, all available oxygen will be consumed and the ponds can become anaerobic. This causes the desirable bacteria to die, resulting in a diminished treatment of the sewage. The low oxygen encourages the growth of anaerobic bacteria, that produce a mucilaginous coating which can quickly cause congestion. Thus, excessive BOD can cause a treatment plant to function poorly, and even fail prematurely.

High concentrations of pollutants may cause many complications in the wetland. A high presence of TSS may result in the clogging of irrigation systems as well as the formation of deposits in the wetland. Moreover, high BOD or COD concentrations can cause both microbial fouling and bacterial regrowth in distribution systems. High levels of total coliforms may also cause bio-fouling of sprinklers and nozzles in irrigation systems.

In order for water to be used for irrigational purposes, a maximum allowable concentration for different pollutants must be set. In Lebanon, the set of requirements is shown in the table below.
Note: 1 mg/l NO3-N = 1 kg N/1000 m^3 of water.

TSS	<50 mg/l
BOD	<30 mg/l
COD	<250 mg/l
Fecal coliforms	<1000 cfu/100ml
TDS	<500 mg/l

TABLE 5: MAXIMUM ALLOWABLE CONCENTRATIONS OF DIFFERENT POLLUTANTS

These values must not be exceeded in 80 percent of samples per month. Min. No of samples 5.

As for nitrate concentrations, the requirements are dependent on the types of crops planted; since sensitive crops may be affected by nitrogen concentrations above 5 mg/l.

While most other crops remain unaffected below a nitrogen levels of 30 mg/l.

Furthermore, it was seen that the effluent leaving the septic tank was too low, which ultimately affected the flow. The effluent was also observed deviating from its proper course and hence did not fully reach the reed bed for treatment. This phenomenon might be due to the unleveled terrain as well as the obstruction caused by dead and/or fallen plants.

V. LIMITATIONS

Some limitations to the methodology followed, and the topic itself, certainly exist. First, vehicular access to the plant was very limited and mostly had to be covered by foot. Moreover, as underlined in the literature review, prior research and experiments were limited and not readily available at our disposal. Access to some parts of the plant was not possible. For example, checking for seepage was not viable because it requires excavation and removal of existing reeds.

As for the sampling and testing, these also can sometimes be delicate: samples have to be kept at a low temperature until taken to the lab, containers should be filled as to avoid air occupying the empty space, etc. These precautions should be taken into consideration in order to preserve the integrity of the samples.

VI. CONCLUSIONS

A comparison of the resultant wastewater parameters was made with the international standards [5]. As previously mentioned, the tests conducted proved that the plant was not effectively clarifying the water and did not meet the required (allowable) standards. Field visits also showed that the flow of effluent recorded was below the expected, which might be due to seepage problems. This particular issue was not thoroughly studied due to the limited access available to the membrane layer.

The new constructed wetland design (5,000 inhabitants) focuses mainly on eliminating as much seepage as possible by using a suitable type of membrane for the area's conditions and characteristics. Also inlet pipes and flow equalization methods will be enhanced, along with the addition of vertical flow reed beds in order to improve the treatment. Up to date screening and filtration techniques will also be used in the new design. The only constraint that seems to stand in the way of such an upscaling is the need for empty surface area. This constructed wetland would indeed require a large space and therefore would not be suitable for implementation in a crowded city such as Beirut. However, further research will be conducted to decide whether such a project could be carried out in another Lebanese village.

As for the issues facing the plant in Remhala, as already stated, they seem to essentially originate from anaerobic conditions in the septic tanks. Poor maintenance and lack of management are also the reasons behind this poorly functioning plant. The clogging of the filters in the pond and

the seepage problems prove negligence on behalf of the municipality officials and all parties involved in the sustainability of this project.

REFERENCES

[1] A. Karnib, "A Methodological Approach for Quantitative Assessment of the Effective Wastewater Management: Lebanon as a Case Study", Environ. Process., vol. 1, no. 4, pp. 483-495, 2014.

[2] The Use of Reed Beds for the Treatment of Sewage & Wastewater from Domestic Households, 1st ed. Lismore: Southern Cross University, 2005.

[3] APHA. Standard Methods for the Examination of Water and Wastewater, 19th Ed.; American Public Health Association: Washington, D.C., 1995; 10–137.

[4] Engineer Rached Sarkis, designer of the plant.

[5] "RECOMMENDED STANDARDS for WASTEWATER FACILITIES", 2014. [Online]. Available: http://10statesstandards.com/wastewaterstandards.pdf. [Accessed: 17-Apr- 2016].

Appendix: CAD drawings

Sewer Network

Decanting basin

Inflow basin
2.00
4.00

400
100

Natural soil

vegetable land for Implantation 40cm

Gravel Surface 10cm

40.00

Decanting basin

Emptying

Outflow basin
2.00
4.00

Outflow basin
2.00
4.00

25.00

Superficial Screen (Fine gravel)
Isolate surface from air (5 to 10cm)

Fine aggregates
Intermediate aggregates 2,5-5cm
Gravel 5-10cm

e=5-10cm
e=20cm
e=30cm
e=40cm

Natural terrain (natural permeability)
Waterproofing Membrane

Basin for vertical flow

Fine aggregates
Intermediate aggregates 2,5-5cm
Gravel 5-10cm

e=10-20cm
e=25cm
e=25cm

Natural terrain
Waterproofing membrane

Basin for horizontal flow

219

Superficial Screen (Fine gravel)
Isolate surface from air (5 to 10cm)

Fine aggregates
Intermediate aggregates 2.5-5cm
Gravel 5-10cm

0cm
m
m
m

Natural terrain (natural permeability)
Waterproofing Membrane

Basin for vertical flow

Fine aggregates
Intermediate aggregates 2.5-5cm
Gravel 5-10cm

0-20cm
5cm
5cm

Natural terrain
Waterproofing membrane

Basin for horizontal flow

Beirut Rafic Hariri International Airport: A Step Towards Perfection

Dani El Kaissi, Bachir Lakkis, Mohamad Basma, Mariana Souki, Moustapha Awali

Civil and Environmental Engineering

American University of Beirut

Beirut, Lebanon

dik04@mail.aub.edu; mkb15@mail.aub.edu; bsl01@mail.aub.edu; mis16@mail.aub.edu; maa226@mail.aub.edu

Abstract- **The only operating airport in Lebanon, Beirut Rafic Hariri International Airport has been facing a challenge in the past couple of years. This challenge is represented by the increasing number of passengers that exceeded the airport's capacity.**

This Paper consists of a multidisciplinary study that aims at examining the existing airport conditions along with a procedural plan to overcome the prevailing challenge. The strategic plan of the project consists of preparing a new pavement design for the operating runways, adjusting the characteristics of the airfield of the airport, and studying an expansion design for the existing terminal building. These modifications shall enable the airport to cater for the Airbus A380 jet which is becoming a necessity for all international airports. In addition it will lead to the increase in the number of passengers.

I- Introduction

Description

Despite Lebanon's political instability and continuing security issues, it remains a prosperous touristic destination; thus, its current airport should provide a state-of-art service. During the year 2015, Beirut Rafic Hariri International Airport (BRHIA) has exceeded its capacity by 1.2 million passengers where its designed capacity is around 6 million annual passengers, in total, while the airport served more than 7.2 million passengers during that year. On the other hand, new types of aircrafts have emerged after the reconstruction of the airport was completed after the Lebanese Civil War. The Airbus A380 is a newly designed jumbo jet that began its operations in 2005. This aircraft type is joining the fleet of numerous airlines in the region, which forces BRHIA to modify its characteristics to become able to accommodate it. These accommodations are vital since the A380 belongs to Code F type of aircrafts which are not currently designed for in the airport.

BRHIA Passenger Traffic

BRHIA has shown some astounding figures with regards to the number of passengers commuting through this facility. In the year 1998, the first phase of the airport was completed with the capacity being around 6 million passengers while the actual number of passengers didn't exceed 2 million passengers. As the aviation industry evolved in the region and the number of expatriates increased, the number of passengers grew significantly. In the past 10 years, the records show a continuous increase in the commuters with the passenger growth rate peaking at 22.6% and averaging at 6.7%. This rate showed a decline only during one exceptional year which witnessed the Lebanese-Israeli War that caused the closure of the airport. Such positive numbers encourage the investment in this well performing facility. In addition, since BRHIA is the only airport in Lebanon, an expansion to the existing terminal building, leading to the increase in the capacity of the passengers, is the only solution for this phenomenon.

The figure below shows the variation in the passenger growth from the year 2005 till 2015
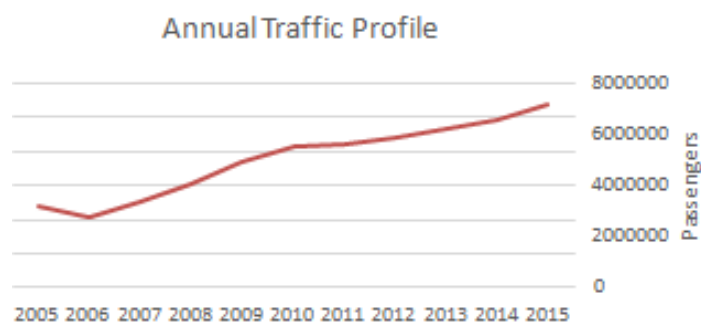

Figure 1: Annual Traffic Profile

II- Design
A- Pavement Design

Since the existing runways are formed of concrete, the Rigid Pavement Design will be considered in the study.

Two scenarios were performed on FAARFIELD which is a program used to reach the optimum thickness of the concrete layer on the surface layer. The first scenario includes the current aircraft mix that is utilizing the airport on a yearly basis. While the second scenario is formed of the expected aircraft mix that is expected to be present after the beginning of the Airbus A380 operations in Beirut. A forecast of 3 daily A380 flights is made and the annual growth rate is taken to be 6.7% according to that calculated from year 2005 till 2015. The design life for the new design is 20 years according to the FAA standards [2]. The table below represents the yearly number of departures of the different Aircraft types in the two cases.

Table 1: Aircraft distribution in the two cases

	A320	A321	A330	B777	B737	A380	Annual Growth
Case 1	12410	4015	2920	1095	2190	0	-
Case 2	11315	4015	2555	1095	2190	1095	6.7%

The different layers of the pavement design are as follows:

• PCC Surface: This surface is responsible to offer a texture of nonskid qualities, impermeable to water, and offers flexural strength able to cater for the aircraft mix expected to use the runway. The modulus was taken to be 4,000,000 psi and strength to be 700 psi, while the thickness is to be calculated by the program.

• P-306 Econocrete: This layer is formed of crushed concrete formed of the old existing concrete surface that was demolished. The thickness was taken to be 6 inches and modulus to be 700,000 psi.

• P-209 Cr Ag: This layer consists of a base course formed of crushed aggregates present over an underlying course. The thickness for this layer was taken to be 6 inches and modulus to be 52,802 psi .

• Subgrade: The stresses reaching the subgrade layer are lower than that reaching the upper layers. The controlling stress is that at the upper of the layer and as the depth increases the stress decreases. The modulus was taken to be 27,238 psi with an infinite thickness.

Results

After simulating both cases, two different thicknesses of PCC Surface were displayed. With regards to the existing airplane mix, a thickness of 15.46 inches was demonstrated. While for the thickness for the new mix that includes the Airbus A380 was 15.69 inches.

By further examining the effects of adding the A380 carrier to the aircrafts operating in the airport on the pavement, several figures arise from the simulation.

Figure 2: Pavement Layers

The following table summarizes the results of the impact of each airplane on the pavement of the designed runway.

Table 2: FAARFIELD Results

No.	Name	CDF Contribution	P/C Ratio
1	B777-200 ER	0.00	4.04
2	A330-300 std	0.05	1.88
3	A380-800	0.02	3.61
4	A320-200 Twin std	0.08	3.70
5	A321-200 std	0.79	3.46
6	B737-800	0.06	3.52

Where CDF (Cumulative Damage Factor) Contribution is the amount of structural fatigue life of a pavement that has been used up [2].

$$CDF = \frac{number\ of\ applied\ load\ repetitions}{number\ of\ applied\ repetitions}$$

While P/C Ratio is the required number of passes that lead to the formation of one full load application to a unit area of pavement. [2] For rigid pavements, the passes are determined by measuring the maximum stress occurring at the bottom of the PCC layer.

The results displayed show that the introduction of a new type of aircraft to the aircraft mix doesn't have major effects on the pavement of the runway by which the Airbus A380 doesn't play the role of the critical aircraft for the runway in terms of CDF contribution and P/C ratio since in both cases, the contribution of the A380 isn't the highest.

B- Structural Design

1. Architectural Layout

The designed extension of the terminal building follows the vertical distribution as the current BRHIA terminal building. In other words, the first level is assigned as the reception floor, along with its specific facilities like conveyor belts, security check points, immigration control stands, duty free…etc. The second level is specialized for the departing facilities, such as the holding rooms, extended duty free and cafes...etc. A plan view for the two levels is shown in the following figures.
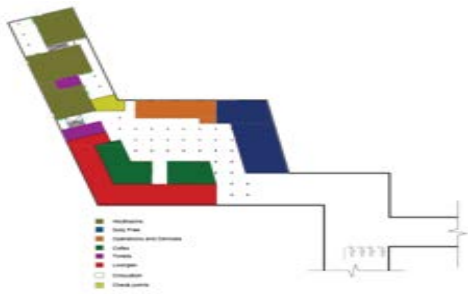
Figure 3: Second Floor
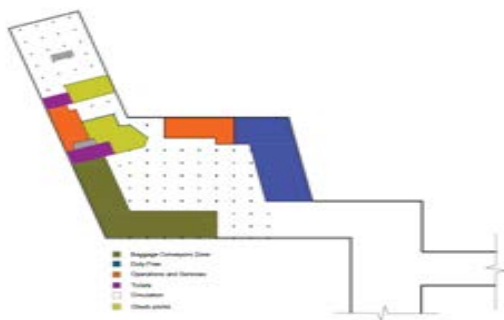
Figure 4: First Floor

The configuration of the extension was set after studying the aerodrome design considerations, guided by ICAO design standards for aerospace (i.e. separation distances between the terminal building and taxiways and/or runways), in addition to the fact of balancing the plan of the whole airport. In other words, the extended terminal conserves the horizontal symmetry of the new BRHIA plan. The total area of the building is around 10630 m2, including three holding rooms, as shown in the figure above, each consists of 650 m2.

Moreover, to conserve the homogeneity with the current terminal building, the clear spans between columns are 9 m and the height of each of the two floors is 4.4 meters.

2. Structural Specifications

The proposed new building consists of 117 (60 cm x 60 cm) columns on each floor, connected by rectangular beams that are 100 cm in width and 42.5 cm in height. The authors preferred a two-way joist system over other slab systems for the following reasons[1]:

1. Longer spans with lesser number of columns, thus architecturally appealing and practical.
2. Lower deflections due to large spans (9 m) subjected to heavy loads.
3. Saving on weight and materials.
4. No trenching is needed if it rains, thus higher productivity and less site work.

Most of the waffle pods dimensions were taken from manufacturer standards, while some other pods' dimensions were customized to fit the edges. Standard 77.5 cm x 77.5 cm x 32.5cm forms were used, separated by 12.5 cm thick ribs. A typical section of the slab, beams and ribs are shown in the figures below.
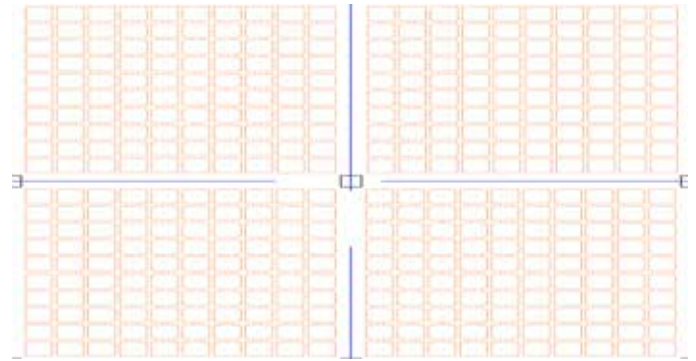
Figure 5: Slab Section

Figure 6: Typical Ribs Section
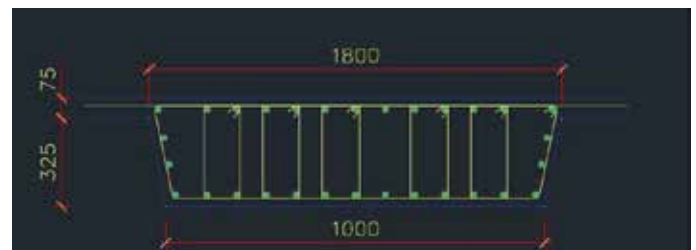
Figure 7: Typical Beams Section

Regarding the loads assigned to the floors, the new operating building can be compared to a library in term of live loads, so according to the ACI building code, it is recommended to take the live loads 500 Kg/m^2. And for the same reason, the super imposed dead loads are 400 Kg/m^2. Note that the roof will have a live load of 200 Kg/m^2.

3. Modeling

In order to assess the performance of the designed terminal building, a detailed model was done on ETABS, according to the architectural and structural specifications discussed above. Deformation, displacements and further design results will be presented in future paper. The following figure shows the proposed structural design of the terminal building.
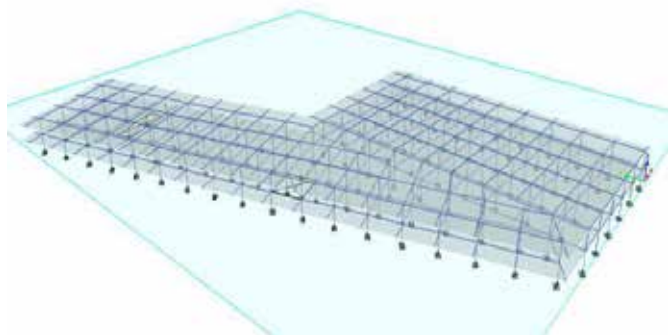
Figure 8: ETABS model

4. Geotechnical Investigation

The soil profile, especially the water table and the bearing capacity of the soil, is one of the constraints facing the study. The only thing known about this profile is that, the underground soil of the site is composed of red sand layers only (according to the geotechnical map of Beirut City).

.

C- Transportation

The characteristics of the existing airport ranging from runways, to aprons, to taxiways are set according to ICAO Code 4E standards. The airport consists of three concrete runways, rapid exit taxiways along with aprons. The two main operating runways are 16-34 and 03-21 which have an open V configuration.

Table 4 summarizes the runway and taxiway characteristics.

Table 3: Runway and Taxiway Characteristics of Existing Airport

Runways/Taxiways	Length	Width	Shoulders
Runway 03-21	3800m	45m	7.5m on each side
Runway 16-34	1900m out of 3400m extending into the sea.	45m	7.5m on each side
Taxiway system		25m	10.5m on each side

In order to cater for the A380 airbus, certain measurements are taken into account to upgrade the current airport conditions. The basic runway width shall be increased by 15m, while the taxiway width remains the same. In addition, the shoulders of the runways shall be increased by 7.5m on each side, while the taxiway shoulders include 7 m on both sides.

Table 4: Airport Airside Requirements

Items	ICAO Code E	ICAO Code F
RWY width + shoulder	45m+ (2x7.5m) = 60m	60m +(2x7.5m) = 75m
TWY width (+shoulder)	23m (+ 2x10.5m) = 44m	25m +(2x17.5m) = 60m
TWY width of taxiway/taxilane strip	47.5/42.5 m	57.5/50.5 m
RWY/TWY separation	182.5m	190m
OFZ width	120m	155m
TWY/TWY separation	80m	97.5m
TWY/Taxilane/Object separation	47.5m - 42.5m	57.5m - 50.5m
Taxiway bridge width	44m	60m
Clearance at gate	7.5m	7.5m

D- Sustainability

The sustainability of the airport facilities is a very attractive aspect that increased the interest in this project. To begin with, the new terminal extension has an important sustainable feature which is the waffle slab structure. Not only

does the two-way joist system reduce the ceiling materials used (i.e. concrete and steel) and the slab cost, but also decreases the weight on each of the footings, which in turns cuts from the foundation materials, therefore potentially means less foundation expenses. In addition to that, especially cost wise, this technique reduces the overall framework of a building charge due to the use of standard, modular, reusable formwork, and recyclable waffle pods material.

Moreover, the curtain walls covering most of the façade parameter allows more exposure of sunlight to the interior, which leads to less consumption of energy during the day.

Also, the use of crushed old runway concrete as a subgrade in the new design will be of great advantage to the project since no transportation and disposal costs are implied, thus less CO_2 emissions are produced. Consequently, such a step results in the decrease of space burdens for the wastes caused by the demolition of the old concrete layer. The fact that the apron is connected to the new building with passenger tubes, no busses are used, therefore no impact on the carbon footprint.

E- Project Planning and Cost

1. Planning

The project construction plan covers the airside and the landside of BRHIA: the terminal building extension, runways and taxiways widening leading to the pavement reconstruction. For the purpose of avoiding interruption in the airport traffic as well as using the remaining life span of the existing pavement (about 2 years), the terminal building construction is scheduled first on the calendar without interfering in the traffic movement. After finishing the construction, the pavement would be already exhausted, and thus the widening of taxiways and runways can be initiated along with the construction of a new pavement. The two operational runways will be reconstructed one after another in order not to obstruct the airport's operation. Thus, the reconstruction will take a place during the period of minimal traffic demand, which, according to the graph below, falls between February till the end of April.
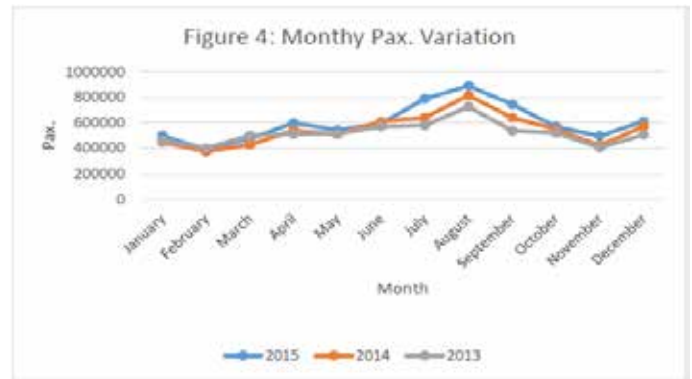
Figure 9: Traffic in the past three years

After the exhaustion of the pavement, it will be demolished and replaced by a new pavement. The existing pavement was estimated to be 40 cm thick. Areas(including runways, shoulders, and aprons), volumes, and production rates of demolition and reconstructing the pavement were used to give an estimate of the project duration. Below is a table that shows an overview of the scheduling of the demolition process:

Table 5: Factors Affecting Demolition Process

Area	586,693 m^2
Thickness	0.4 m
Volume	234,677 m^3
Demolition Rate (fast track)	2,250 m^3/day
Duration	104 days

The above production rate is derived for 6 working Hammers (1,700 kg) and 6 excavators (330C).

As was mentioned previously, the demolished concrete is then used as a 15 cm thick layer of base course for the new pavement (P-306 Econcrete). Spread and leveling of the crushed concrete will be applied using 4 dozers with a 1.2 m^3 bucket. The table below shows the production rate of spread and leveling:

Table 6: Factors Affecting Leveling of crushed concrete

Area	586,693 m^2
Thickness	0.15 m
Volume	88,000 m^3
Rate (Fast track)	2136 m^3/day
Duration	41 days

It must be noted that the above phase may not consume from the project's time because it can occur simultaneously along with the phase of pavement demolition.

The final phase of the pavement construction consists of the concrete work:

Table 7: Factors Affecting Pavement Construction

Area	586,693 m^2
Thickness	0.4 m
Production Rate (Fast track)	4180 m^2/day
Duration	140 days

Thus, the overall time duration for the pavement demolition and reconstruction is estimated to be around 250 days. This does not mean that the airport will be dysfunctional over this period of time, however, it will function continuously because only one runway at a time will be reconstructed.

2. Project Scheduling

The schedule prepared is divided into two phases with different wind orientations discussed in each case.

Phase 1: This phase includes the closure of Runway 16/34 in order to perform the widening and pavement works for the runway. On one hand if the wind was North –South, Runway 03/21 should be used for landing and Runway 35/17 should be used for take-off. On the other hand, if the wind was S-N, Runway 21/03 should be used for landing while Runway 17/35 should be used for take-off.

Phase 2: This phase includes the opening of Runway 16/34 and closure of Runway 03/21. If the wind was in the N-S direction, landing should occur on the Runway 35/17 while take-off on Runway 34/16. While if the wind was in the S-N direction, the landing should be on the Runway 16/34 and takeoff on Runway 17/35.

3. Project Cost

The estimated cost of a project consisting of an expansion of the existing airport and the modification of runways can be calculated from cost rates that are supplied from Dar al Handasah (Shair and Partners). The factors that played a role in forming the sum are the structural construction work of the terminal building, the technical and safety equipment for the terminal building and the runways, and the construction of the pavement.

Table 3 presents the rates and sums for each fixture and the estimated total cost for the project.

Table 8: Estimated Costs

Airfield	Rate ($/m^2)	Unit	Quantity (m^2)	Cost ($)
Airfield Rigid Pavement				
Runway Rigid Structural Pavement	205	m²	464,914	95,307,452
Runway Rigid Shoulder and Blast Pad Pavement	58	m²	115,779	6,715,165
Passenger Aprons Pavement	205	m²	6,000	1,230,000
Airfield Flexible Pavement				
Taxiway Shoulder Pavement	38	m²		
Buildings				
Passenger Terminal Building	4,000	m²	21,570	86,280,000
Construction Cost of Building (62%)	2,480	m²	21,570	(53,493,600)
Equipment & Systems(Provisional) (38%)	1,520	m²	21,570	(32,786,400)
Passenger Loading Bridge Movable	1,150,000	Nr.	3	3,450,000
			Total	$192,982,617

III- Conclusion

The multidisciplinary study prepared provides a well planned solution for the next several years. When the project is completed, the airport will be developed to become able to cater for aircrafts up to type F, while previously it was only able to accommodate aircrafts up to type E. However, further plans are important to follow such a step. As this vision is applied, a well designed airport will be ready to cater for the needs of the Lebanese citizens and tourists. In addition, our only operating international airport will gain a positive reputation and may play a role in competing with the neighboring airports.

References

1. AC 150/5320-6E. CHAPTER 3. PAVEMENT DESIGN FOR AIRPLANES WEIGHING MORE THAN 30,000 POUNDS. Retrieved from https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentID/99762

2. Al Ayyash, Z., & R., Yaghi. (2014). Airfield Pavement: A Case Study On BRHIA Runway 16/34.

3. Packard, R. G. (n.d.). *Design of Concrete Airport Pavement* (pp. 13-30).

Design and Assessment of a Dry Port in Lebanon

Al Hajj Hassan, Lama; Al Rayess, Jana; Damaj, Omar; Gemayel, Karen; Obeid, Hassan

Department of Civil and Environmental Engineering

American University of Beirut

Beirut, Lebanon

lah24@mail.aub.edu; jke04@mail.aub.edu; omd05@mail.aub.edu; kkg02@mail.aub.edu; hao11@mail.aub.edu

Abstract- **This paper presents the design of a dry port in Lebanon, its economic and environmental impacts, and financial feasibility. The site for this facility, designed to be connected to the Port of Beirut (PoB), is selected to be Rayak. Its main purpose is to relieve congestion at PoB. In terms of capacity, the dry port is designed to handle 202,000 TEU's of domestic and transit containers, as well as 250,000 tons of both dry bulk and general cargo, by 2030. The port's pavement was designed to prevent rutting and cracking. Preliminary studies show that this project is financially feasible. The estimated IRR for the operator proved to be attractive with values ranging from 17% to 24% based on different scenarios of the handling fees.**

INTRODUCTION

In Lebanon, for the past decade, the economic sector has been struggling. In spite of the stagnant economic conditions, Port of Beirut (PoB) has been flourishing. In 2014, the port received 1.2 mil containers, against a current capacity of 1.4 mil containers [1]. Aided by PoB's high operation efficiency, this activity is anticipated to keep increasing, especially when stability is restored in the region. At this rate, the facility is expected to exceed its capacity by 2018 [1]; in fact in 2014, PoB utilized 85% of its capacity which reflects a high congestion level [2]. Until now, efforts to relieve this congestion were directed at increasing PoB's capacity either by; expanding it toward Beirut River, filling the fourth basin, or decreasing container dwell time. However, all these measures have their limitations. The expansion, which was met by a storm of objections, can't be implemented without a governmental decree, and according to PoB officials interviewed in January, the fourth basin project was also terminated for the same reasons. Further, even though improving the level of operations can solve the issue in the short term, this measure will only increase handling capacity hence attracting more volumes to PoB and further contributing to the problem.

This project shifts PoB's off-dock functions and "extends" its gate to a dry port. A dry port is an inland extension of a sea port that acts as a terminal for dropping off, receiving, and clearing containers. In addition, it provides other services such as un/packaging cargo, container storage and maintenance, and warehousing. By assuming the sea port's off-dock functions, the dry port will achieve its main purpose of relieving congestion at PoB. The designed port will also facilitate transit, an alternative more economically-viable than sea shipment for most countries in Lebanon's hinterland. Furthermore, as Lebanon's economic indicators, such as GDP and unemployment rates, do not reflect any promise of future growth, a transportation facility like the dry port has the potential of boosting the country's economic welfare.

This paper is structured as follows: Determination of dry port's location; Design capacity and area; Port pavement design; Environmental and financial feasibility; Economic and social impacts.

DRY PORT LOCATION

The dry port's site was determined based on several criteria to facilitate the construction process, and insure easy access with the neighboring countries. These criteria included flat site slope, proximity to road networks, distance from densely populated areas, and distance from rivers and agricultural lands. Another criterion was the proximity to existing rail tracks to allow for future establishment of a series of rail connected dry ports across the Arab countries. These requirements were then incorporated in a GIS model where the input files were a digital elevation model (DEM), a land-use raster, and shapefiles containing information on road networks, river streams, and cities and villages with their respective population. The model's final output was a rated map of Lebanon where all lands were rated based on how much they satisfy the mentioned requirements. A sensitivity analysis was performed with different rating scales, and Rayak, Bekaa, scored the highest in all scenarios.

DEMAND FORECAST

According to PoB officials, dry ports are designed to handle container volumes predicted to arrive in 15 years. For demand forecast, the dry port main trade partners were identified based on their past and current share size of Lebanon's imports and exports, their accessibility by land, and their proximity to Lebanon. Based on Lebanon's profile in the Observatory of Economic Complexity (OEC), an economic online platform, the country's leading trade destinations which met the aforementioned criteria were as follows: Turkey, KSA, UAE, Jordan, Syria, Kuwait, and Iraq.

The dollar values for land freight, imports and exports, by country was calculated as the difference between the total trade monetary value and the sum of the dollar values of freight transported by air and sea. OEC website provides total trade monetary values between Lebanon and each of the stated countries for the period from 2002 till 2011, and the distribution of these imports and exports by mode (air and sea) were taken from Ministry of Finance (MoF) reports. Sea and land freight monetary values were then converted to equivalent number of 20ft containers (TEU) after dividing each route's transport cost($) by the product of ($/ton.km) rates and route lengths (km) taken respectively from [3], [4] for sea routes and [5] for inland routes. Finally, knowing that a 20ft standard container holds 24 tons [6], the number of equivalent containers was calculated.

Based on the resultant number of 2011 containers arriving by land and sea, the transport multiplier (TM) concept, which is used by PoB officials, was applied to predict the container demand at the dry port in 2030. This multiplier is the ratio of change in container traffic to change in the country's GDP which is assumed constant over an interval of five years. For each time interval up to 2030, the percent growth in container traffic was calculated based on TM estimates provided by PoB officials and GDP growth estimates taken from International Monetary Fund (IMF) website. Then container growth values were used to calculate expected traffic at the end of each interval.

Other than the sea ports, Lebanon has three main overland access points: Masnaa, Abboudie, and Kaa border points with Syria. The Masnaa point represents the dry port's "gate" and takes up 15% of exports and 18% of imports in Lebanon [7]. Therefore, annual TEU arriving at the dry port is the product of the calculated annual TEUs and these percentages. The forecasted container traffic was inflated up to 25% to account for demand growth due to the dry port's implementation and generated transit passing through it, and the traffic was deflated down to -25% to reflect worst case scenarios; this range is based on the growth pattern of the Riyadh dry port in KSA [8]. Then the number of containers was corrected for capacity considerations by dividing the annual TEUs by 0.8 knowing that a container yard reaches capacity at 80% utilization [9]. Finally, the annual number of containers expected to pass through the dry port in 2022 and 2030 are presented in Table I.

CONTAINER YARD AREA

The container yard area was calculated based on the maximum daily container inventory. The calculation accounted for container dwell time and the inventory consists of export, import, and empty containers. 2022 and 2030 are both non-leap years so the total annual container traffic was divided by 365 days to get the daily traffic of which 53% are exports and 47% are imports based on previous calculations. The dry port should also have a stock of empty containers needed for packing; this stock is the difference between imports and exports.

The dry port was designed in compliance with United Nations Conference on Trade and Development (UNCTAD) guidelines and assuming that truck scheduling software is used to designate for each truck an arrival time interval and corrects for delays.
Consequently, the standard UNCTAD dwell times (Exports: 3 days, Imports: 7 days, Empty: 15 days) were applied for inventory calculations as shown in (1), (2), and (3) [10].

$$Average\ Export\ Inventory=3\ days*Export(\tfrac{TEU}{day}) \qquad (1)$$

$$Average\ Import\ Inventory=7\ days*Import(\tfrac{TEU}{day}) \qquad (2)$$

$$Average\ Empty\ Inventory=15\ days*Empty(\tfrac{TEU}{day}) \qquad (3)$$

The total daily inventory is the sum of the results of the above three equations. PoB officials recommended using stacks of three containers for easy shuffling, and each stack has an area of $15m^2$, so the container yard area is equal to five times the daily inventory. Container yard areas are presented in TableII.

TABLE I
ANNUAL CONTAINER TRAFFIC IN DRY PORT (TEU)

Growth (%)	-25	-15	-5	5	10	15	20	25
2022	119,824	135,797	151,784	167,757	175,743	183,730	191,716	199,716
2030	131,973	149,581	167,176	184,770	193,568	202,365	211,162	219,959

TABLE II
CONTAINER YARD AREA (X100M²)

Growth (%)	-25	-15	-5	5	10	15	20	25
2022	9.5	10.9	12.1	13.5	14.0	14.7	15.3	15.9
2030	10.61	11.9	13.4	14.7	15.5	16.2	16.9	17.6

CONTAINER FREIGHT SECTION AREA (CFS)
This area is used for un/packing dry bulk (DB) and general cargo (GC) that are not delivered in containers. Since no data for DB and GC exists for Lebanon, PoB officials recommended that the CFS area be designed to handle 125,000 yearly tons of each for both 2022 and 2030. An area of 10,417 m^2 is required based on formula below [10]:

$$CFS\ area= \frac{TEUs\ per\ year\ \times 40 \times 25}{300}\ m^2 \qquad (4)$$

A 15% container growth rate was assumed since for both 2022 and 2030 choosing any growth rate value negligibly affects the storage area requirement. Thus, a 15% growth rate was chosen so that the dry port's capacity is not over-estimated. However, the total area taken is $60,000\ m^2$ because this area is owned by the government in Rayak according to PoB officials. This area includes the container yard, CFS, the area needed for parking, equipment storage, buildings, and internal roads, Fig.1.

RAIL- ROAD NETWORK
Usually, a railway connects a dry port with the sea port however in this case it was disregarded because a railway is not feasible financially and environmentally. According to Dar Al Handasah engineers, to reach the Bekaa valley, a tunnel would have to pass through Mount Lebanon at a cost of 15 mil $/km and the construction of a single track costs 25 mil $/km. They further added that railways are not done for distances shorter than 100km and for this case only a 50 km track is needed. Moreover, the tunnel excavation would impact the biodiversity and surrounding environment.

As for the roadway network, it was necessary to assess the roads connecting PoB and Rayak and then the network with the hinterland to ensure facilitated freight mobility. Mr. Ghaleb Naamani, a transport engineer at Dar Al Handasah, recommended that the trucking road between PoB and dry port

follows the already existing highway connecting Beirut and Jamhour, a plan that was approved by Ministry of Public Works and Transport. Then it was decided to connect Bhamdoun and Choueit by a bridge to avoid the Kahale road known for its high truck related accident rates [11]. Afterwards, trucks continue to Dahr Al Baydar along the Pan Arab highway to reach the dry port in Rayak. The total road route has a length of 61.51 km, Fig. 2. Also, to limit the bridge span, avoid entering the Kahale region, and limit truck time in residential areas, a 1.53km extension of the existing road is suggested. This extension is represented in blue in Fig. 3 and compared to the Kahale road (in yellow) which is much more difficult to navigate a heavy truck along because of the steep slopes and narrow curves. The bridge connects the extended road to Choueit over a length of 2.41 Km. Although the transit traffic is limited in the projected years; this road may be used by heavy and medium trucks transporting goods with in Lebanon and not just limited to the studied transit trucks. The additional truck traffic is expected to be roughly 20 trucks per round trip and thus will not have excessive effect on traffic. The number of extra trucks is based on the expected service and arrival rates calculated using the generated transit traffic between PoB and the dry port which accounts for 5% of total daily dry port TEUs and a one way trip time of 95 min assuming a 30 km/hr in Beirut and 45km/hr in Bekaa[12].

As for road network with hinterland countries, six routes represent the transit roads between the Arab countries. The Lebanon-Syria-Jordan route has an acceptable speed reaching 54 km/hr, but has a relatively high time wasted at the borders with an average of 50 hr/trip due to bureaucracy and poor custom practices which the dry port can help mitigate [13].

PORT PAVEMENT DESIGN

Three main aspects were considered in the pavement design. First, pavement should withstand the static load from the containers and the moving load from the different trucks and equipment. Second, the pavement should withstand severe temperatures as Rayak witnesses relatively extreme high and low temperatures shown in Fig.4. Finally, the pavement design should be economically and environmentally viable.

Therefore, the pavement was designed as a grid, where the vertical and horizontal lines that represent the movement of the trucks and RTG's within the port are paved with concrete, and the areas in-between are paved with asphalt that is because the container load is spread over a larger area compared to load over equipment wheels. Furthermore, this approach guarantees economic and environmental viability for it utilizes more asphalt than concrete. Even after the costs of wear and tear and maintenance are taken into account, asphalt's residual value is 203% more than that of concrete [14]. In terms of sustainability, asphalt pavements are 100% recyclable, they are not harmful to a water environment, and have a lower carbon

footprint than concrete since they require half the energy to construct when compared to concrete pavements [14].

The concrete beneath the trucks' tires and the RTG's was designed as standard PCC with maximum required thickness, 30.5 cm. This conservative thickness was chosen to guarantee effective load resistance no matter the load carried by the equipment.

The first step was to compare temperatures in Rayak to a city whose flexible pavement design is part of Mechanistic-Empirical Pavement Design Guide (MEPDG) software database. Using online weather database, accuweather.com, the daily temperatures for the year 2015 were obtained and inputted into a temperature graph. This revealed a very close pattern between Rayak and Dallas, Texas (USA), especially in the higher temperatures, Fig. 4. Therefore, the input file for Dallas was used to design the pavement knowing that the temperature correlation between Zahle, close to Rayak, and Dallas was verified in "unpublished" [15]. Other design inputs included, a 15 km/hr speed for all machinery within the dry port [16]. In the design, a 15% fatigue value for a service life of 20 years was targeted. The pavement was designed to handle 20,000 Average Annual Daily Traffic of heavy load trucks. Truck load was considered heavy to account for the load of the carried container.

For an asphalt thickness of 23.5cm overlaying a 38 cm layer of crushed stone, MEPDG predicted a 1.4cm rutting in the asphalt and 2cm of total rutting, as well as 15.25% fatigue at the end of the design life thus results were acceptable. Another mix was used to verify the MEPDG results and was used as input to design the flexible pavement. This mix was designed by Mr. Hussein Kassem, PhD candidate in Civil Engineering at AUB, to withstand relatively high loads, namely at airports.

ENVIRONMENTAL IMPACT ASSESSMENT

In general, ports have adverse effects on the environment including solid waste generation, water pollution from possible fuel spillage, damage to vegetation and wildlife, and noise pollution. Environmental factors were considered in the decision-making process and adequate preventative measures were suggested to minimize any possible negative impacts [18]. The main impacts considered include effects on site biodiversity, runoff, and CO_2 emissions.

The assessment of the dry port's site allows the identification of landscape features, existing ecosystem, vegetation areas, and existence of threatened species. However, the identified inland location of Rayak does not include any critical habitat.

In terms of post-development runoff, the dry port's impermeable surfaces will disrupt rainwater infiltration which leads to increased runoff and possible spread of runoff polluted by spillage of fuel, oil, and other contaminants. In terms of precipitation, the average rainfall intensity "i" of 685 mm/year,

and runoff coefficient "C" equal to 0.2 and 0.8 for predevelopment and post-development respectively [17] . Using the Rational Method and taking the 60,000 m^2 port area to be completely paved with asphalt, the runoff is Q = CiA = 22 m^3/day predevelopment, and Q = CiA = 90 m^3/day post-development. This increase in runoff volume shows that an adequate storm water network should be considered.

As for emissions, the amount of CO_2 released by any mode of transportation depends on the traveling and waiting processes. Equation 5 is used to estimate the amount of CO_2 produced by a certain mode of transportation [19].

$$TE = \sum_{m=0}^{m=n} E_{m,l,d} + \sum_{m=0}^{m=n} I_{m,d} \qquad (5)$$

Where TE is the total emissions, E is the emissions due to transportation and I is the emissions due to the waiting process, for different modes of transportation m. However, in the design of the dry port, the amount of CO_2 emitted due to waiting time was neglected as it is negligible compared to the traveling time. A comparison between ship and truck transportation allows for an adequate assessment of the change in CO_2 emissions incurred by the implementation of the dry port. According to the network of transportation and the environment, while carrying one ton of cargo for one kilometer, a truck emits 59 grams of CO2 and a ship 10 [18]. The impact on CO2 emissions by taking the land route over the sea route was investigated in Table III based on the emission factors just mentioned and transit route length adapted from [13] and sea route lengths from [4]. The results show that the extra CO_2 is minimal in most of the routes except the longest.

TABLE III
CO$_2$ EMISSION: LAND VS. SEA ROUTE

OD	Land	CO$_2$ (kg)	Sea	CO$_2$ (kg)	Difference
Lebanon-Jordan	352	20.768	705	7.05	13.718
Lebanon-Kuwait	1250	73.75	6608.7	66.087	7.663
Lebanon-KSA	550	32.45	3900	39	-6.55
Lebanon-UAE	2446	144.314	5822.5	58.225	86.089
Lebanon-Syria	70	4.13	200	2	2.13
Lebanon-Iraq	500	29.5	6640.32	66.4032	-36.9032

As a result, several mitigation measures, in addition to the storm water network, can be taken to alleviate the impacts of the dry port on the environment. For instance, unnecessary movement of trucks and equipment must be reduced to minimal, and this is achieved through system optimization. Cargo handling equipment that is older than ten years of age must be retired at the new facility. It is also preferable to use alternative fuel and eco-friendly equipment.

FINANCIAL FEASIBILITY

For this project, the net present value (NPV), internal rate of return (IRR) were calculated under different scenarios using a forecasted stream of costs and revenues. These economic indicators helped establish the most feasible and financially attractive alternative to carry out this project.

The main costs to be incurred in the project are as follows: land acquisition cost, infrastructure and building costs, equipment costs, IT cost, personnel costs, and maintenance costs, whereas the benefits are generated mainly from the handling fees of containers, general cargo and dry bulk. Given that the dry port is to be in Rayak, there will be no cost of land acquisition as most of the required area is owned by the government and is assumed to be provided at no extra cost to the dry port operating company as a way to encourage investors and to make the project more attractive. In return, the government will obtain a certain percentage of the revenues from the handling fees collected by the dry port operator. The other fixed costs are the infrastructure and building costs, equipment costs, and IT costs. These costs are summarized in Table IV.

TABLE IV
DRY PORT FIXED COSTS (x 1000,000 $) [20]

Year	Infrastructure	Equipment	IT	Total
2017	34	0	0	34
2018	17	0	0	17
2019	17	0	0	17
2020	17	30.65	0.8	48.45
Total	85	30.65	0.8	116.45

The infrastructure cost constitutes earthworks, paving, and facilities. In terms of earthworks, the land considered for the dry port is relatively flat, and thus a medium level of earthwork is required. At this stage of the feasibility analysis, the required earthworks include excavation, cutting and filling. As for surfacing and pavement works, the yard surface will be a grid with concrete pavement for the service roads and asphalt pavement for the storage areas. Building and warehousing costs were also included in the analysis such that the buildings include the headquarters, an operational and control building, the custom and inspection building, and the maintenance garage, and a large warehouse for un/packing GC/DB. In the feasibility analysis, the government is assumed to share infrastructure costs with the operator to make the project more attractive.

The number of equipment was obtained by benchmarking the dry port's expected demand with other dry ports around the world. The total equipment cost is estimated at 30.65 ml $ and was then obtained as per Table V.

The IT costs were approximated based on information obtained from an interview with PoB IT staff. It includes required software and hardware to insure digitization and optimum coordination of the dry port's data and synchronization with PoB. The costs were approximated at $800,000.

TABLE V
DRY PORT EQUIPMENT COSTS [20]

Equipment	Unit Cost (mil$)	Units needed	Total Cost (mil $)
RTGs	1.55	15	23.25
Empty Handler	0.299	5	1.495
RS	0.47	7	3.32
Forklift Trucks	0.97	10	0.97
Tractor Trucks	0.11505	14	1.61

As for the maintenance costs, they are split between infrastructure, equipment and IT maintenance. Infrastructure maintenance is assumed to be 1% of the buildings and infrastructure initial cost [20]. Equipment maintenance is dependent on the number of TEUs handled at the dry port, with a cost of $2.77 per TEU, and the IT maintenance is 10% of the IT cost and is required yearly for maintenance and updates of the software [20]. These costs are incurred starting 2021 and are subject to an inflation rate of 3% over the study period of 30 years (until 2050) [20].

Finally, the number of employees required at the dry port was obtained by comparing the equipment and tasks required at the dry port to those at the newly constructed container terminal at PoB. The number of employees was then obtained by interpolation. An estimate of the salaries was obtained by interviewing personnel at PoB, and these salaries are subject to an inflation of 6% per year [20].

As for the benefits these are generated from the container handling fees imposed on container traffic as well as GC/DB. As such, the benefits are dependent on the fees which should be carefully set as to attract shippers while maintaining an attractive margin of profits to make the project attractive to investors. The average handling charges for the different type of container traffic at the PoB: 245$/TEU, $31.5/ton of GC and $19.25/ton of DB [21]. It is important to note that shipping lines typically obtain lower fees in exchange for a guaranteed traffic volume.

Given the traffic volume expected in the dry port, the costs and benefits over the entire study period can be calculated for different values of the fees involved. To encourage shipping lines to use the dry port, it is important that dry port fees be lower than those at PoB. As such, several dry port fees taken as fractions of the PoB fees were applied to see the fraction below which the dry port is no longer financially feasible. Additionally, two scenarios were considered regarding the contribution of the government to the infrastructure costs and the fraction of the revenues it earns from the handling fees. The first scenario assumed that the government will contribute to 25% of the infrastructure cost and will obtain 30% of the generated revenues. The second alternative assumed that the government will contribute to 50% of the infrastructure cost, and will consequently obtain a larger fraction of the revenues,

assumed also at 50%. The NPV and IRR values accrued by both, the government and the operator, are presented in Table VI for two different fractions of the fees imposed at PoB and discounted at 12%, a typical value in the region [20].

TABLE VI
OPERATOR PERSPECTIVE

Case	% Gov. Contribution	% Fees for Gov.	Fraction of fees wrt PoB			
			0.5		0.75	
			IRR (%)	NPV (mil $)	IRR (%)	NPV (mil $)
1	25	30	17	48.34	24	113.85
2	50	50	16	27.61	22	74.4

TABLE VII
GOVERNMENT PERSPECTIVE

Case	% Gov. Contribution	% Fees for Gov.	Fraction of fees wrt PoB			
			0.5		0.75	
			IRR (%	NPV (mil $)	IRR (%)	NPV (mil $)
1	25	30	13	6.04	17	34.11
2	50	50	24	60.17	30	106.9

As shown in above, all scenarios are financially feasible. Comparing the IRRs obtained to those of other similar projects shows that the values obtained are reasonable. The Ismailia Dry Port in Egypt and the Rajiv Ghandi Dry Port in India both have comparable IRRs of 22.2% and 20.2%, respectively [22] and [23].

Intuitively, the project is more feasible for both the government and the operator if the fees imposed at the dry port are higher. However, the higher the fees, the less attractive the dry port is for shipping lines. As such, the sensitivity of the demand with respect to the handling fees should be determined to decide on the optimum handling charges. It is however expected that shippers will be in favor of the dry port even when the fees are 75% of those at PoB due to the reduction in dwell time and storage fees.

ECONOMIC ASSESSMENT

SWOT analysis is done to assess a project's viability, to highlight possible issues, and to come up with adequate preventative measures. This analysis consists of identifying the organization strength (S) and weaknesses (W) that give the organization advantages/disadvantages over others as well as other external opportunities (O) and threats (T) that affect the project outcomes.

In terms of strength, the dry port benefits from its proximity to the border Masnaa point, Zahle industrial city, and to 27 major exporting factories, Fig.5, in addition to the agricultural exports which the Bekaa valley is known for. Moreover, the project site is on a flat land requiring minimum earth work, and

it is connected to a well-established road network most importantly the Pan Arab highway built mainly to boost transit. The region itself provides a large labor force base with a population of 840,300 including refugees where 60% of this population is in the working age, 10% are deprived Lebanese, and the unemployment rate is 6% [24]. These advantages outweigh the potential weaknesses such as excessive border waiting times [13] and difficulty in installing a rail which was proven to be not feasible or justifiable. Finally, the weather conditions may be problematic specifically during snow storms, but the dry port can have its own snow plowing machine.

In terms of external factors, the dry port benefits from Lebanon's strategic position, closeness to main capitals, and to the old rail in Rayak which was connected to Syria and is still in an acceptable situation [25]. Furthermore, there is a profitable future market for construction material in Syria to feed the reconstruction process because the supply of products through PoB, which is closer to Damascus than Latika and Tartus ports, will reduce transport costs for the expected volumes considerably; that in addition to the fact that the national infrastructure network connecting Syrian cities requires major reconstruction which will hinder the transport of products [26]. Another market may open for recycled material and aggregates from demolished buildings in Syria whether in Lebanon or overseas so packaging and customs done in the dry port would facilitate the process and limit the possible congestion at PoB while generating revenues. Nevertheless, the dry port activity may be affected by the region's stability which is why the project can be phased and expanded appropriately. Also, government tax regulations should be coordinated with dry operators so as not to adversely impact container activity. Finally, the adoption of the TIR convention, by which a container is sealed at the origin and only opened at the destination port, by all hinterland countries is a necessity to increase the adoption of container use and facilitate trade.

Based on the above, the project is viable and the Keysian School further supports the need for government to subsidize such projects. This school advocates governmental spending, based on historical data that links improved economic growth to an increase in government expenditures. For example, the Czech Republic's anti-crisis plans in 2009 included increased expenditures on infrastructure projects especially transport ones [27]. Moreover, engaging in a resource intensive project like the dry port has numerous advantages. By engaging in a project of this scale several jobs are created, during construction and operation phases, and this will help develop the Bekaa region. Meanwhile the capacity problem in PoB is elevated allowing the authorities to offer competitive rates, and hence attract more investors. In addition, PoB authorities can focus on boosting the port's transshipment role after gaining

capacity from empty containers and stored containers that shifted to the dry port. Other possible advantages are available in Fig. 6 which is based on surveys conducted with dry port managers.

The construction of a dry port in Lebanon is expected to be provided by the government or a public private collaboration. The realization of the project is accompanied by many risks. These risks incorporate: saturating the market by logistic services, the risk of decreased tariffs for transportation operations, inflation due to injected funds. These risks are to be analyzed by a professional consultancy company once the project is accepted for development; the risk assessment is not expected to change the decision to proceed with the project.

Conclusion

In conclusion, the dry port was designed with a total surface area of 60,000 m^2 of which 16,200 m^2 is container yard area and 10,417 m^2 is CFS area. This facility will handle 212,800 TEU's of storage, including DB and GC. At an IRR of 17% the project will provide 100 direct jobs. The pavement was designed as a grid with 30.5 cm thick PCC strips along which the equipment move and 23.5cm thick asphalt layer covering container stack areas. Finally, the project was found to be environmentally, economically, and financially feasible, with minor mitigation measures. This facility gives rise to new projects, both regional and national. Such projects include a series of similar railway-connected dry ports in the Arab countries as well as developing software to optimize the operations and coordination between PoB and the dry port.

References
[1]. Demarque, G. (2015). Beirut Port: Holy waters. Executive Magazine. Retrieved 28 March 2016, from http://www.executive-magazine.com/economics-policy/beirut-port-holy-waters
[2]. Assaf, C. (2015). Alternatives for Developing Containers' Activity at the Port of Beirut. Beirut:BlomInvest Bank.Retrieved from http://blog.blominvestbank.com/wp-content/uploads/2015/07/Alternatives-for-Developing-Containers----Activity-at-the-Port-of-Beirut.pdf
[3]. Korinek, J., & Sourdin, P. (2009). Maritime transport costs and their impact on trade. Accessed March, 11, 2011.
[4]. SeaRate,. (2015). Routes Explorer. Retrieved 4 December 2015, from https://www.searates.com/services/routes-explorer/
[5]. JIB. (2011). jib.jo. Retrieved 27 March 2016, from http://www.jib.jo/Portals/0/library/PDF/634464695292832500.pdf

[6]. Sea Container Specifications. (2014). Retrieved from http://www.nrs-international.com/wp-content/uploads/2014/11/containers.pdf

[7]. Openknowledge.worldbank.org,. (2015).The Impact of the Syrian Conflict on Lebanese Trade. Retrieved 9 February 2016, from https://openknowledge.worldbank.org/bitstream/handle/10986/21914/The 0impact0of00ct0on0Lebanese0trade.pdf?sequence=1

[8]. Big.sa.com, (2016). Logistics & Port Services | Baas International Group. Retrieved 8 January 2016, from http://big.sa.com/logistics-port-services/

[9]. Bryan Salminen, J. (2013). Measuring the Capacity of a Port System: A Case Study on a Southeast Asian Port(Master of Engineering in Logistics).MIT

[10]. Handbook on the Management and Operation of Dry Ports. (1991). Geneva.

[11]. AssafirNewspaper. (2014).Lebanon: Death on the Kahale Road. Retrieved 2 March 2016, from http://assafir.com/Article/3/353087/MostRead

[12]. Climatechange.moe.gov.lb,. (2011). GREENHOUSE GAS MITIGATION ANALYSIS. Retrieved 14 February 2016, from http://climatechange.moe.gov.lb/viewfile.aspx?id=36

[13]. Meeuws, R. (2013). Developing a Trade and Road Transport Facilitation Strategy for the Arab World. Jordan: IDB-AULT-IRU. Retrieved from https://www.iru.org/cms-filesystem-action/mix-publications/IDB-booklet-Road_Transport-EN-web.pdf

[14]. ASPHALT FACTS. Alasphalt.com. Retrieved 17 April 2016, from http://www.alasphalt.com/asphalt%20facts.htm

[15]. G. Chehab, E. Touma, and E. Khawaja, "Implementation Initiatives of the Mechanistic-Empirical Pavement Design Guides in Lebanon," (unpublished), August 2010.

[16]. Cushman-Roisin, B. (2015). ENGS-44 Sustainable Design:Water. Lecture, Dartmouth University.

[17]. LITANI RIVER BASIN MANAGEMENT SUPPORT PROGRAM WATER BALANCE REPORT. (2011) (1st ed.). Beirut. Retrieved from http://www.databank.com.lb/docs/Water%20balance%20report-Dec%202011.pdf

[18]. Carbon Emissions | World Shipping Council. (2016). Worldshipping.org. Retrieved 17 April 2016, from http://www.worldshipping.org/industry-issues/environment/air-emissions/carbon-emissions

[19]. McKinnon, A. CO2 Emissions from Freight Transport: An Analysis of UK Data (1st ed.). Logistics Research Centre, Heriot-Watt University. Retrieved from http://www.greenlogistics.org/SiteResources/d82cc048-4b92-4c2a-a014-af1eea7d76d0_CO2%20Emissions%20from%20Freight%20Transport%20-%20An%20Analysis%20of%20UK%20Data.pdf

[20]. Kaysi, I., & Jamaleddine, K. (2016). Term Project (March 2016). Lecture, American University of Beirut.

[21]. Sky Handling Partners (2012). Environmental Impact Assessment Study and Mitigation Plan for Construction of a Cargo Storage Warehouse at https://www.miga.org/Documents/EIA_Study_Cargo_Storage_Facility

[22]. DRY PORTS IN INDIA –NEED AND CHALLENGES. (2014). www.unescap.org. Retrieved 17 April 2016, from http://www.unescap.org/sites/default/files/6.4.India__0.pdf

[23]. Ismailia Dry Port Project. (2015) (1st ed.). Cairo. Retrieved from http://bombaychamber.com/admin/uploaded/Download/Ismailia%20Dry%20port.pdf

[24]. Murphy, M. (2015). Lebanon: Bekaa Governorate Profile.Beirut: UNHCR. Retrieved from https://data.unhcr.org/syrianrefugees/download.php?id=9524

[25]. PCM Meeting Minutes. (2014). Pcm.gov.lb. Retrieved 29 March 2016, from http://pcm.gov.lb/arabic/subpg.aspx?pageid=5958

[26]. Demarque, G. (2015). Port of Beirut. Executive Magazine. Retrieved 31 March 2016, from http://www.executive-magazine.com/opinion/a-port-policy-for-all

[27]. A, K., & Khesin, E. (2013). EU member states economies after the introduction of the Euro: From Euphoria of 1999 to the debt crisis of the 2010s. 250-251.

[28]. Chang, Z., Notteboom, T., & Lu, J. (2015). A two-phase model for dry port location with an application to the port of Dalian in China. Transportation Planning and Technology, 38(4), 442-464

APPENDIX
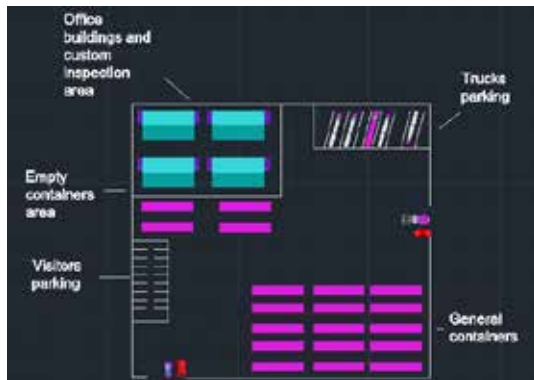
Figure 1 Dry port layout
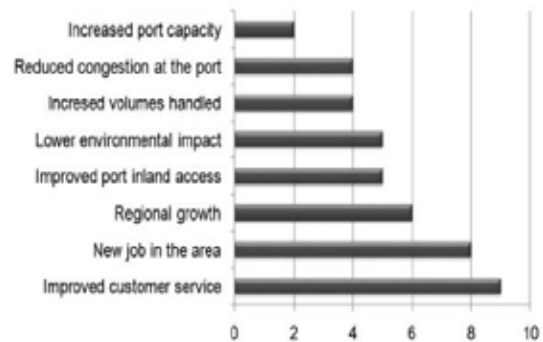
Figure 6 Benefits of the dry port Source: Chang, Notteboom & Lu, 2015 [28]
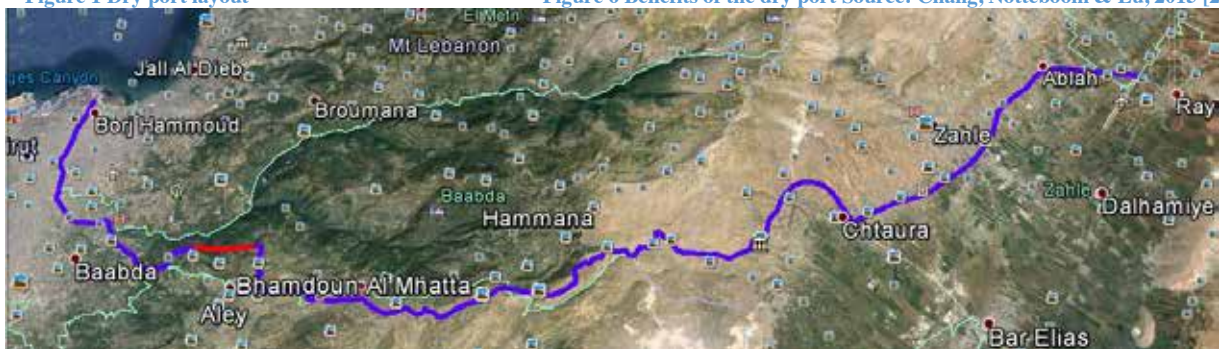
Figure 2 Truck PoB-dry port suggested route

Figure 3 Extension Road vs Kahale Road (Left), Bridge (Right)
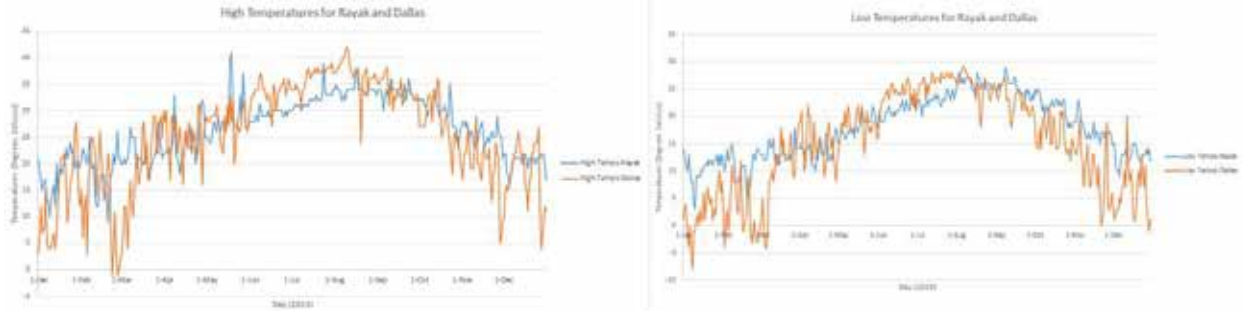
Figure 4 Temperature comparison for Dallas and Rayak

Figure 5 Major factories close to the dry port and listed in table below

TABLE VII

LIST OF FACTORIES IN FIGURE 6 BY INDICATED NUMBER

1.Liban Lait	5.Mimosa	21.Ethel Chocolate Factory	12.Taanayel Dairy Products
2.Al Dirani Factories	6.Tanmia	23.Dairy Khoury	13.Sicomo
3.Daher Foods	7.Somo Plast	25.Chateau Ksara	14. SLAC
4.Kaa El Rim Carton Company	8.Al Yaman	22.Domaine Wardy	18.Chateau Kanafar
9.Chateau Ksara	15.Cave Kouroum	24.Jaber Jaber	19.Wadi El Kheir
10.Center Jdita	16.Chateau Marsyas	26.Domanco	20.Al Manara Dairy Product
11.A-Z Manufacturing	17.Chateau Kefraya	27.Conserves Chtaura	

234

Design and Management Plan for a Refugee Camp in Lebanon

Joanna Chatila, Nayer Daher, Saeed Nehme, Helmi Oueidat and Karim Tabaja
Civil and Environmental Engineering Department
American University of Beirut
Beirut - Lebanon
jac11@mail.aub.edu, nmd25@mail.aub.edu, ssn21@mail.aub.edu, hwo02@mail.aub.edu and kat10@mail.aub.edu

Abstract- **This study presents a design of a sustainable refugee camp in Tall Znoub, Lebanon. The camp in question includes housing units and common areas, medical and educational facilities, and water and wastewater supply and collection networks. The scope of work is divided into site planning, structural design, environmental and societal impact assessment, and software modeling. The implementation of the scope is executed through multiple site visits to potential campsites, regular visits to the UNHCR, and to actual settlements in order to better approximate the refugees' current conditions and develop solutions to cater their basic needs. Findings of this study include the complete design of an official refugee camp in Lebanon, across which wooden shelters are distributed with Concrete Canvas shelters in order to accommodate the different needs of refugees. In addition, water and wastewater networks are designed to satisfy the refugees' sanitation requirements. These facilities contribute to positively impact the environment and society in which the camp will be implemented. Finally, rendered views of the refugee camp are developed using 3Ds MAX.**

I. INTRODUCTION

One major headline trending nowadays is the refugee crisis around the world. The situation has yet to improve as nearly 19.5 million refugees were registered under the United Nation High Commissioner for Refugee's (UNHCR) mandate in 2014. Besides issues faced by refugees themselves, the hosting nations also get their share of problems, ranging from religious, sectarian to denominational affiliations. For instance, Lebanon, due to its proximity to the major war zones in the region, provides a significant platform for refugees, mainly of Syrian, Palestinian, and Iraqi origins. Yet, Lebanon isn't fully capable of accommodating such a large number of people, hence the existence of unregulated informal settlements. This project proposes a design and management plan for an environmentally sustainable refugee camp in Lebanon.

The study focuses on five primary dimensions that need to be addressed for establishing the first official, legal and sustainable refugee camp in Lebanon: site planning, structural design, environmental and societal impact assessment, and software modeling. The project begins with providing the refugees with structurally stable shelters that withstand the weather conditions of Lebanon's mountainous regions distributed according to specific criteria across the selected

land in Tall Znoub. Following the site layout and structural design of shelter units, an adequate water network is designed in order to ensure continuous water supply to the camp. A reduction in the environmental implications is obtained by resorting to dry sanitation and finally an assessment of the social impact and future prospects of the energy efficient, sustainable refugee camp design is established using a detailed schedule and work breakdown structure of the activities required for the actual implementation of the layout design.

II. LITERATURE REVIEW

Refugee camps started developing in the middle of the twentieth century. Various conflicts across the years were responsible for the establishment of new and larger camps across the continents. The Palestinian conflict was one major event that led to the creation of multiple camps in Lebanon. Other refugee camps were created in order to shelter communities escaping the horrors of civil wars in their home countries – mostly in Africa [1]. Therefore, the need to shelter and protect all moving populations resulted in the creation of refugee camps.

The different amenities provided in camps are accommodation, hygiene facilities, clinics, hospitals, and educational centers. Therefore, the following section explores the services provided in different landmark refugee camps. The chosen camps are Dadaab in Kenya – the world's largest refugee camp, Al Zaatari in Jordan – the largest camp in the Middle East and Northern Africa region, and Ain El-Hilwe in Saida, Lebanon. These camps were selected in order to evaluate the differences in offerings in these camps based on the camp size and location.

A. Dadaab Camp

Located around 100 kilometers away from the Kenyan-Somalian border, the Dadaab camp is considered to be the world's largest refugee camp. The camp is composed of five different settlements that were established chronologically in order to accommodate for the increasing number of refugees in the region. The first settlements were built in the 1990s to provide shelter for Somalis escaping the civil war. More recently, the combination of regional conflicts and natural events were responsible in increasing the number of refugees

in the Dadaab camp to reach a number greater than 423,496 refugees [2].

Ever since its establishment back in 1992, Médecins Sans Frontières has been providing medical care in the different settlements. Regarding education, around 37% of the camp population is of school-going age; however, only 51% of these are attending schools as efforts of the UNHCR are being put into action in constructing new schools [3]. On the other hand, worker programs are offered in some settlements in order to help refugees contribute to the environment and society they are living in and help them aid their families, especially with a daily average of 1,300 refugees arriving to Dadaab. Finally, regarding water availability and accessibility, UNHCR's efforts to establish a 314 kilometers water pipeline network theoretically provides 23.5 liters of water per day for each refugee, which is definitely lower than the standard water requirements of individuals.

B. Al Zaatari Camp

Al Zaatari camp is the second largest refugee camp located in Jordan in "Mafraq" governorate. It was established in July 2012 by the Jordanian government and worldwide agencies such as UNHCR to accommodate Syrians that fled their country due to the ongoing Syrian civil war which broke out in 2011. Around 79,284 refugees currently reside in the camp [4].

In coordination with the Jordanian Ministry of Health, several facilities provide primary healthcare, including clinics administered by the United Nations Population Funds, Médecins Sans Frontières and others. However, the camp suffers from shortage of medical support utilities such as ambulances. Concerning water and sanitary services, the Ministry of International Cooperation and Development and UNHCR announced that work is in progress to produce a full-scale wastewater and drainage system which will be used for agricultural purposes in the camp. Additional works for the provision of sewage collection for each refugee household in the camp are underway. Then, another piping network will transport the sewage accumulated in communal tanks to the wastewater treatment plant [4]. Regarding education, the cooperation between the Ministry of Education and other relief agencies such as United Nations International Children's Emergency Fund (UNICEF) ensures that Syrian refugees are receiving proper education through newly established schools [5]. Despite the presence of such facilities, the Educational Sector Working Group in Jordan found that 48.4% of all school-aged children are out of school, 38.6% are not currently receiving any form of education whether formal or informal, and 28.3% have never attended any form of education [3].

C. Ain Al-Hilwe Camp

Taking a more local perspective, the investigation of the different offerings in local refugee camps in Lebanon is highlighted through the Ain Al-Hilwe Palestinian refugee camp in South Lebanon. Located 45 kilometers south of Beirut, the camp was created in 1948 in order to accommodate the Palestinian refugees escaping their country. Today, the camp hosts more than 140,000 refugees mainly Palestinians in addition to Syrians which have been fleeing the civil war since 2011. According to Médecins Sans Frontières, the living conditions in Ain Al-Hilwe are far from being exemplary [6].

In fact, most refugees are currently living in old overcrowded buildings that need rehabilitation while some of them are in tents. As security poses an important obstacle against conducting studies and censuses in the camp, no information regarding water supply is available. However, the United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA) confirms that water supply in the camp does not meet the demand. From an education point of view, the UNRWA is also responsible of providing schools for children in the camp despite the huge challenges faced by the organization. With the new arrival of Syrian refugees, the UNRWA has been keen on rehabilitating the camp's shelters and infrastructure and planning for several electricity and water supply projects.

From the different aspects of the refugee camps analyzed, focus will be on determining the optimal combination of housing, clinical and hygiene facilities in order to satisfy the refugees' needs. As shown in Table 1, tents are the most common facility used in order to shelter the migrant populations. However, the use of Concrete Canvas shelters within common areas of the camp is an appealing solution.

TABLE 1
CHARACTERISTICS OF THE CAMPS UNDER CONSIDERATION

	Dadaab	*Al Zaatari*	*Ain Al-Hilwe*
Housing	Tents & Prefabricated houses	Tents/Shelters	Tents & Buildings
Education	Classrooms	Classrooms	Classrooms
Water Supply	Low supply	Low supply	Low supply
Wastewater Treatment	N/A	Transport to wastewater treatment plant	N/A
Health	In-camp clinics	In-camp clinics	N/A

D. Concrete Canvas Material

Concrete Canvas (CC) – developed by a British company in 2005, is a flexible, fabric-rich concrete that hardens when hydrated to form a long lasting, water proof and fire resistant concrete layer [7]. CC permits concrete construction without the need for plant or mixing equipment.

This innovative product was first used for ditch lining and mining activities. CC's ease of installation and flexibility allows the opportunity to construct a wide range of custom ditches in different regions of the world, as well as temporary blast and vent structures in mines. More recently, CC has been used for erosion control and slope protection. The use of this revolutionary product as part of a refugee camp is yet to be adopted so its applicability will be considered in this study as part of a combination with the wooden shelters discussed in the following section.

A. Site Layout design

A Refugee camp has to be designed in accordance with the Lebanese Regulations and the UNHCR latest handbook[1]. The site that was found the most suitable to host the camp in is a 495,000 square meters of land located in Tall Znoub, West Bekaa which is a property of the Ministry of Information.

The site survey established by the authors of this paper concluded that the number of refugees currently residing in the area amounts to 3,200 residents. The camp is designed to host 5000 refugees assuming a 4% growth rate over the next decade. The area allocated for one person is 30 square meters according to the UNHCR handbook, yielding to a total area of 150,000 square meters of space required. In order to efficiently supply the refugees with provisions and services with the minimum time required, the camp would comprise of five main clusters linked by main access roads, each hosting a thousand refugees. Each cluster would contain a medical clinic, a food center and two classrooms along with the necessary latrines, shelters and showers which will be discussed in more details throughout the paper. The outer margins and inner separations comply with the firebreak requirements providing 75 meters of firebreak allocated space for every 300 meters of measured distance [8]. A total of five storage areas are integrated along with a recreational space designed for every cluster and one common cemetery. Fig. 1 presents a plan view of the camp layout identifying the five main divisions. Refer to Table 4 in appendix for comparison between Tal Znoub camp and camps mentioned in "Literature Review" section.
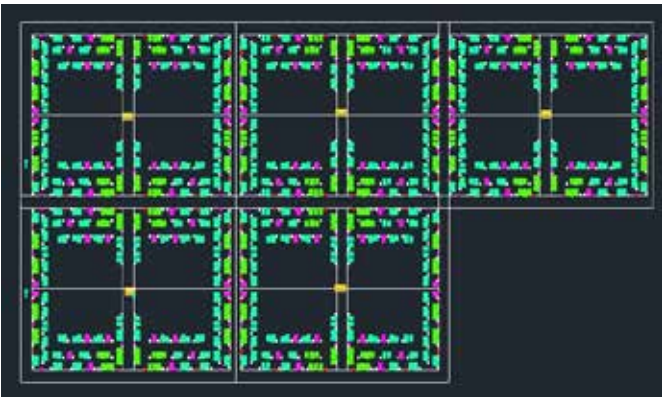

Figure 1. Plan view of the camp layout

B. Shelter Design

The primary structural element to be designed is the shelter unit that would host the refugees. According to the UNHCR latest handbook, the standard shelter space allocated per person is 3.5 square meters [8]. Two shelter sizes have been devised: the first is a small size shelter of 14 square meters for four residents while the second is a 21 square meter shelter fitting six refugees. Fig. 2 presents the two standard shelter sizes.
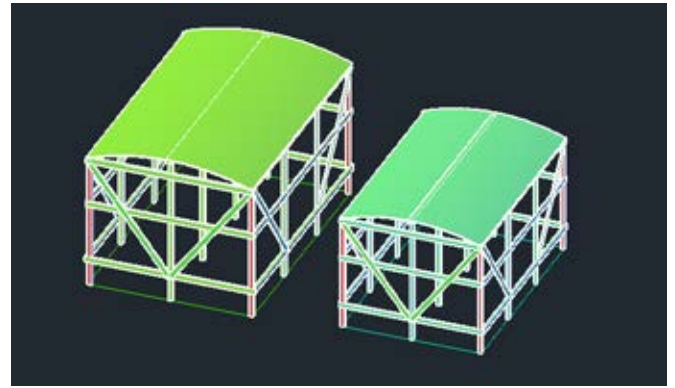

Figure 2. Standard shelter sizes

The refugee families the team has visited extend on average between six, ten and up to fourteen persons per household. Table 2 below presents data on the number of shelters provided per cluster (1000 people).

TABLE 2
DISTRIBUTION OF SHELTERS PER CLUSTER

Family Sizes	Customized Houses: Shelter Unit(s)/Household	Number of houses/cluster
Size 1: 6 people	1 big shelter unit (21 m^2)	20
Size 2: 10 people	1 big shelter unit (21 m^2) and 1 small shelter unit (14 m^2)[1]	60
Size 3: 14 people	1 big shelter unit (21 m^2) and 2 small shelter units (14 m^2)	20
Total	20x**6** + 60x**10** + 20x**14** = **1000 people per cluster**	

The shelters in order to comply with Lebanese Regulations must not be considered as long term settlements and therefore cannot be made from concrete. The material used for the shelter units is Douglas fir wood as it can be easily supplied in Lebanon. Moreover, the wooden structure can be assembled and dismantled by the refugees themselves allowing them to build their own shelters and reach self-reliance. The roof of each shelter is equipped by polystyrene foam (which is biodegradable by Pseudomonas putida) for thermal insulation [9].

To hold the shelter down, the authors initially thought of pouring a thin layer (about 7 cm) of blinding concrete directly onto natural ground, that acts both as a floor and a support system for the shelter. Considering the camp is a short-term project, the concrete option was replaced by a bed of hollow construction blocks to act as a floor. To help transfer the load to the ground, the main columns are extended about 50 cm into an isolated opening in the ground. The shelter's roof is curved, to help alleviate the snow and rain load (and for rainwater harvesting purposes). Horizontal and inclined wooden beams provide lateral bracing, and a wooden frame spanning the length of the shelter and extending to the ground supports the curved roof. Using SAP 2000, the shelter was tested to sustain

---

[1] Note that shelters must have a minimum distance of two meters separating them according to the UNHCR guidelines

the weather conditions in Bekaa rural area, the loads were computed for the lager shelter size by referring to Bekaa's average snowfall and wind speed provided by "Unifrom Building Code" [10]. The loads for which the shelter was tested are provided in Table 3.

TABLE 3
LOAD TEST VALUES

Loads	Magnitudes (kN/m²)
Dead	0.328
Live	0.294
Wind	0.511
Snow	0.981

Primary structural testing on SAP2000 shows that the shelter remains structurally sound despite the multitude of applied loads and load combinations that were taken according to ASCE Section 7.10.

C. Elevated Water Tank

To provide the tank with usable water, the authors sought out multiple options, mainly drilling wells, and buying water on demand. Yet the idea of designing and testing a water tank seemed both challenging and rewarding, prompting the team to proceed with it. The camp's daily demand was estimated to be about 100 m³/day, and this figure was offset by a factor of safety of 1.5 to account for leakages and uncertainties in the demand (the new demand figure is therefore 150 m³/day). The design of the tank was performed on SAP2000, including a reinforced concrete support system. Primary testing shows little to no issues concerning the structural integrity of the water tank.

D. Water Network

The total water demand required is 30 liters per person per day, this includes basic water demand for cooking, drinking, and basic hygiene practices (showers and latrines). In practice, water demand for similar camps ranges between 20-30 liters/capita/day; so extra amount was added to compensate for losses and firefighting demands.

A water network – depicted in Fig. 3, accompanied by the water tank was used as the main distributor for the entire camp. Since this kind of network will allow uneven water harvesting hence, a Venturi meter will solve practically the problem by allocating desired rate of water (m³/day) for each location. Polyvinyl Chloride (PVC) piping of 32 mm diameter size will
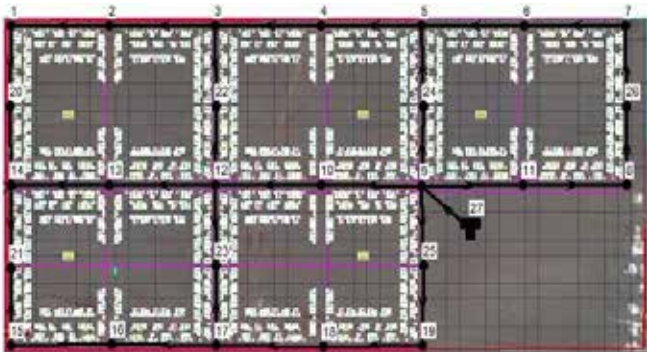

Figure 3. Plan view of the site layout along with the water network

be used. PVC piping systems are environmentally stable, have a long service life, are easily installed, are corrosion resistant, and are cost effective.

All demands were distributed on the different nodes of the network with proper elevations and the network was modeled on EPANET. Pressures on major nodes are acceptable and follow the distribution depicted in Fig. 4 where all resulting pressures are high enough to provide a head for the water to reach the top of the shelter or showers and hence insuring proper distribution of demands.
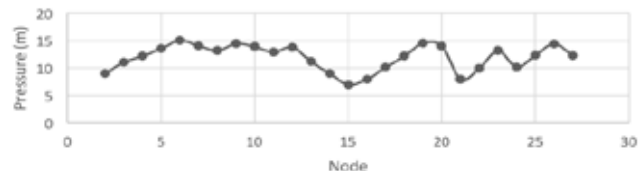
Figure 4. Nodal Pressure Distribution

E. Dry Sanitation, Latrine Design and Greywater System

One of the general considerations that need to be thought of during the site planning phase is the immediate services that ought to be provided for the refugees; these include proper provision of latrines and showers in order to ensure a sanitary environment across the refugee camp. With regards to the environmental implications these structures may subsequently generate if not adequately designed and constructed, an innovative and sustainable latrine design had to be implemented.

The latrine design makes use of the dry sanitation concept . Dry sanitation approaches usually require the separation of urine and feces. Urine, which generally poses little threat to human health, also contains nutrients (88% of the nitrogen, 67% of the phosphorus, and 71% of the potassium). Separation of urine allows it to be used safely as a fertilizer after minimal treatment. Similarly, feces that contain most of the pathogens also can be safely used as a fertilizer after storage at ambient temperatures for two years or composting at high-temperatures for six months.

According to the UNHCR, one latrine is to be allocated for every 20 people, thus a total of 50 latrines are provided for each cluster hosting a 1000 refugees. However, after discussing with the UNHCR and Dr. M. Abou Najm, ensuring privacy and security of the residents became one of the primary targets of the project, hence a latrine is provided next to each household – with a cross-section shown in Fig. 5 from AutoCAD.

Figure 5. Cross-section of the latrine

On the other hand, every shower installed accommodates a maximum of 50 people, which means that the minimum number of showers needed for 250 people is 5 showers yet a total of 8 has been chosen and evenly distributed across each block containing 250 people as to ensure availability, accessibility and mostly privacy and separation between men and women showers providing 4 showers for each gender.

Also, greywater generated from the camp – represented by water received from sinks and showers may contain traces of dirt, and certain household cleaning products. However, it is a safe and beneficial source of irrigation. Greywater release into rivers and lakes lead to the contamination of the mentioned media, but to plants, they are valuable fertilizer as the nitrogen and phosphorous in the greywater serve as necessary food source for the plants. Aside from the obvious benefits of saving water and money, reusing greywater keeps it out of the sewer or septic system, thereby reducing the chance of polluting local water bodies.

In the camp only "plant friendly" products which are salts-free detergents, boron, or chlorine bleach are to be used since the build-up of salts and boron in the soil can damage plants.

F. Service and Storage Facilities

Concerning the provision of educational and healthcare services in the camp, the classrooms and medical facilities will be Concrete Canvas shelters. These structural elements can be added as long term investments as they present future prospects for the land once the refugee crisis is resolved. The material in questions – the canvas, is flexible, and hardens when hydrated forming a long lasting water proof and fire resistant concrete layer. The canvas is laid out, inflated and cured for 24 hours. Every cluster would have three Concrete Canvases each of 50 square meters containing two classrooms and a medical clinic. Figure 6 below presents the Concrete Canvas model developed.
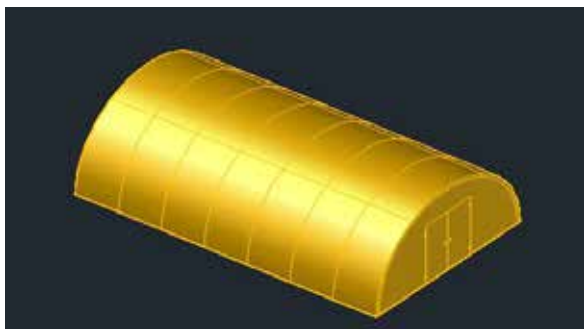

Figure 6. Concrete Canvas model

On the other hand, a storage area is allocated for every cluster. The storage facility is made of prefabricated steel; this material has been widely used for storage facilities as they are easily assembled and can be procured from multiple suppliers available in Lebanon such as Dalal Steel who was contacted for information about their products, standard sizes and delivery procedure as well as the durability, structural stability and impact resistance of their structures. The volume of the storage facility is taken as 180 cubic meters abiding by the UNHCR standards [8]. Similarly, two food distribution centers each of 100 m^2 area will be hosted in a prefabricated steel structure in each block. This center will host the food preparation activities from the collection of ingredients, to cooking and finally distributing the food to the refugees. Subsequently a total of 15 steel prefabricated structures are incorporated in the camp divided between storage and food distribution centers.

G. Other Facilities and Amenities

The camp offers several other facilities, which help improve the refugees' living conditions. First, playgrounds are provided for young refugees where they will be able to enjoy their time. Second, soccer fields are provided across the camp to allow all refugees to participate in sport activities. Finally, a cemetery is provided to burry deceased refugees close to their families.

H. Solar Panels

One aspect of the camp that still needs to be discussed is the generation of electricity from solar panels placed on the shelters. In fact, one main feature of this refugee camp is the fact that it is an environmentally renewable and sustainable camp.

Assuming that electricity will be used mostly for lighting and heating in addition to personal use from the socket, the estimated energy required per year totaled 212 kWh per shelter.

In order to satisfy this demand, a 200 Watt rated solar panel having a 15% solar panel yield and 1.34 m^2 area will be installed on the roof of each household which means 500 solar panels will be installed across the camp. The annual average solar radiation which depends on the latitude of the area and the angle of inclination of the panel, is estimated to be around 2000 kWh/m^2.year as well as an efficiency of 75% is considered. Hence the energy generated from a single panel [11] is 300 kWh, 88 kWh more than the required energy amount per shelter.

IV. CAMP VISUALIZATION AND RENDERING

In order to visualize the camp, 3Ds MAX was used. Objects of same layer enter as 1 object so materials from the "*iray*" and "*vray*" databases are assigned to each object – Wood for the wooden shelters, steel for the latrines, showers, storage, and food centers, concrete for the classrooms. Then the daylight systems alongside the camera are set up in order to complete the rendering process. A general rendered view of the camp is shown in Fig. 7.

Figure 7. Rendered view of the entire camp

V. Project Scheduling and Work Breakdown Structure

A. WBS: Work Breakdown Structure

The first step performed was to break down the project into activities. The WBS approach organizes and defines the scope of the project allowing a preliminary baseline schedule to be set. A total of 55 activities had been set out for the organization of the Refugee Camp project, starting from the meeting with stakeholders and suppliers, to the design phase and ultimately the construction phase.

B. Project Schedule

Critical Path Method approach (CPM) was used for scheduling the Refugee Camp project from its initiation and planning phase until its construction phase. Using Primavera P6, the construction logic was established. The software was chosen as it identifies critical activities, is self-adjusting when updated and is currently the most effective scheduling system implemented. After defining the predecessors, successors and durations of each activity, the total duration for the actual construction of the camp ideally amounts to a total of three months (Figure 8 in Appendix), yet a safe lag time of additional three months must be attributed in order to account for possible weather conditions that may hinder the construction process or the productivity of laborers and to account for delays in obtaining the necessary permits from regulatory agencies.

C. Project Cost

Once the Work breakdown structure was completed and the activities were scheduled using Primavera P6 software, the project cost was estimated by preparing an excel sheet simulating a bid tabulation activity (Figure 9 in Appendix). The cost of laborers was generated and the direct costs of the several activities in each of the divisions that need to be accorded to the construction of the refugee camp were processed. Such activities include the drainage pipes procured and installed, the trenches excavated for service and drain pipes, the woodwork performed for the shelter units, the main tank of a capacity exceeding 2000 gallons, the sanitary fixtures and installation, food and storage areas procured and health clinics along with their provisions. The costs were determined based on Lebanon's construction cost data collected; the unit prices included date back to 1996 and therefore the data was brought to present date by taking into account the inflation indices for consumer prices from the International Monetary Fund official website.

The implementation of the first officially legal and environmentally sustainable project in Tall Znoub, Lebanon is estimated at 2,774,133 US dollars. Given the number of estimated refugees that could ultimately be hosted along with the environmental and sustainable dimension allotted for the camp, the project proves to be cost efficient. A refugee camp that provides its residents with structurally stable shelters that are accommodated to the Bekaa's climatic conditions while providing them with all the necessary care (food provision, access to clean water, dry sanitation, health clinics) not only is desperately needed in these times in Lebanon, but also provides an interesting future outlook to the endless possibilities such a camp can provide for the actual residents of the Bekaa area once the refugee crisis is resolved.

References

[1] K. ELMadmad, "Mixed Flows And The Protection Of Migrants With Special Reference To Sub-Saharian Migrants", *ISPI*, vol. 26, 2008.

[2] Dadaab Refugee Camps, Retrieved from http://www.care.org/emergencies/dadaab-refugee-camp-kenya, 2008.

[3] S. Dryden-Peterson, "Refugee Education: A Global Review", *United Nations High Commissioner For Refugees Policy Development And Evaluation Service*, pp. 26–27, November 2011.

[4] Syrian Observatory For Human Rights, "UAE to fund sanitation project to help Syrian refugees in Zaatari camp". Retrieved 19 October 2015, from http://www.syriahr.com/en/2015/08/uae-to-fund-sanitation-project-to-help-syrian-refugees-in-zaatari-camp/

[5] Relief International. Retrieved 22 October 2015, from http://www.ri.org/newsroom/news-article.php?ID=83

[6] A. Lachant, "From Damascus to Ain el-Helweh: Palestinians in Syria Flee to Lebanon", *Medecins Sans Frontieres*

[7] Milliken Infrastructure, *Concrete Cloth GCCM Milliken Infrastructure*. Retrieved 4 December 2015, from http://infrastructure.milliken.com/concrete-cloth-gccm

[8] *Handbook For Emergencies*. 3rd ed. Genève: UNHCR, 2007. Web. 5 Jan. 2016.

[9] Ward, PG; Goff, M; Donner, M; Kaminsky, W; O'Connor, KE (2006). "A two step chemo-biotechnological conversion of polystyrene to a biodegradable thermoplastic".*Environmental Science and Technology* **40** (7): 2433–7. doi:10.1021/es0517668

[10] *Uniform Building Code*. Whittier, CA: International Conference of Building Officials, 1997. Print.

[11] "How To Calculate The Output Energy Or Power Of A Solarphotovoltaic System, Excel PV Calculator To Estimate Solar Electricity Output". *Photovoltaic-software.com*. N.p., 2013. Web. 9 Mar. 2016.

APPENDIX

Table 4. Comparison between Tal Znoub camp and other camps

	Dadaab	Al Zaatari	Ain Al-Hilwe	Tal Znoub
Housing	Tents/Prefabricated houses	Tents/Shelters	Buildings/Tents	Wooden Shelters
Education	Classrooms	Classrooms	Classrooms	Classrooms
Water Supply	Does not meet demand	Does not meet demand	Does not meet demand	~ 30 m^3/person/day
Wastewater Treatment	N/A	Transport of sewage to nearby wastewater treatment plant	N/A	Grey water reuse system
Health	On-site clinics	On-site clinics	N/A	On-site clinics
Electricity	Municipal supply	Municipal supply	Municipal supply	Solar Panels

Activity ID	Activity Name	Original Duration	Early Start	Early Finish	Late Start	Late Finish	Free Float	Total Float
FYP FYP: Refugee Camp		108	01-Jun-16	16-Sep-16	01-Jun-16	16-Sep-16	0	0
1	Start	0	01-Jun-16		01-Jun-16		0	0
2	Site Survey	2	19-Jun-16	19-Jun-16	30-Jun-16	30-Jun-16	0	11
3	Camp siting	2	19-Jun-16	19-Jun-16	30-Jun-16	30-Jun-16	0	11
4	Notice to proceed	10	09-Jun-16	18-Jun-16	20-Jun-16	29-Jun-16	0	11
5	Site Mobilization	2	20-Jun-16	21-Jun-16	01-Jul-16	02-Jul-16	0	11
6	Order of Prefab structures	7	04-Jun-16	10-Jun-16	07-Sep-16	13-Sep-16	12	95
7	Site Conditioning: Clearing and Grading	10	19-Jun-16	28-Jun-16	20-Aug-16	29-Aug-16	0	62
9	Fencing Design	1	01-Jun-16	01-Jun-16	18-Aug-16	18-Aug-16	21	78
10	Pouring Grout	2	22-Jun-16	22-Jun-16	03-Jul-16	03-Jul-16	0	11
11	Water piping installation	15	22-Jun-16	06-Jul-16	26-Aug-16	09-Sep-16	0	65
12	Drain Piping	15	29-Jun-16	13-Jul-16	30-Aug-16	13-Sep-16	0	62
13	Latrine Installation	7	07-Jul-16	13-Jul-16	10-Sep-16	16-Sep-16	65	65
14	Shower Installation	7	07-Jul-16	13-Jul-16	10-Sep-16	16-Sep-16	65	65
15	Shelter Installation	70	04-Jul-16	11-Sep-16	04-Jul-16	11-Sep-16	0	0
16	Ordering Concrete Canvas	60	04-Jun-16	02-Aug-16	07-Jun-16	05-Aug-16	0	3
17	Concrete Canvas Installation	5	03-Aug-16	07-Aug-16	06-Aug-16	10-Aug-16	0	3
18	Concrete Canvas Curing	2	08-Aug-16	09-Aug-16	11-Aug-16	12-Aug-16	0	3
19	Lighting	7	08-Jul-16	14-Jul-16	03-Sep-16	09-Sep-16	0	57
20	Placing Road Gravel	1	14-Jul-16	14-Jul-16	16-Sep-16	16-Sep-16	64	64
21	Paving	3	14-Jul-16	16-Jul-16	14-Sep-16	16-Sep-16	62	62
22	Landscaping trees	7	15-Jul-16	21-Jul-16	10-Sep-16	16-Sep-16	57	57
24	design of Water Tanks	1	01-Jun-16	01-Jun-16	12-Jun-16	12-Jun-16	0	11
25	Wastewater collection	1	01-Jun-16	01-Jun-16	16-Sep-16	16-Sep-16	107	107
26	Construction of a drainage ditch	7	22-Jun-16	28-Jun-16	10-Sep-16	16-Sep-16	80	80
27	Watertank Installation	1	23-Jun-16	23-Jun-16	16-Sep-16	16-Sep-16	85	85
29	Furnishing prefab structures	3	12-Sep-16	14-Sep-16	12-Sep-16	14-Sep-16	0	0
30	Furnishing common areas	3	09-Sep-16	11-Sep-16	12-Sep-16	14-Sep-16	3	3
32	Ordering Hospital Equipment	30	10-Aug-16	08-Sep-16	13-Aug-16	11-Sep-16	0	3
33	Preliminary waternetwork design on EPAN	5	03-Jun-16	07-Jun-16	14-Jun-16	18-Jun-16	0	11
34	Final design on EPANET	1	08-Jun-16	08-Jun-16	19-Jun-16	19-Jun-16	0	11
35	Meetings with shareholders	3	01-Jun-16	03-Jun-16	14-Sep-16	16-Sep-16	105	105
37	Site layout	1	19-Jun-16	19-Jun-16	30-Jun-16	30-Jun-16	0	11
38	Fencing Installation	15	23-Jun-16	07-Jul-16	19-Aug-16	02-Sep-16	0	57
39	Shower design	2	01-Jun-16	02-Jun-16	12-Jun-16	13-Jun-16	0	11
40	Latrine design	2	01-Jun-16	02-Jun-16	12-Jun-16	13-Jun-16	0	11
41	Storage design	2	01-Jun-16	02-Jun-16	01-Jun-16	02-Jun-16	0	0
42	Built water tanks	1	01-Jun-16	01-Jun-16	16-Sep-16	16-Sep-16	107	107
43	Installing storage areas	3	23-Jun-16	25-Jun-16	14-Sep-16	16-Sep-16	83	83
44	Order classroom equipment	15	10-Aug-16	24-Aug-16	28-Aug-16	11-Sep-16	15	18
45	Installing generators	2	15-Sep-16	16-Sep-16	15-Sep-16	16-Sep-16	0	0
46	Wind Load Testing	1	03-Jun-16	03-Jun-16	03-Jun-16	03-Jun-16	0	0
47	Snow load Testing	1	03-Jun-16	03-Jun-16	03-Jun-16	03-Jun-16	0	0
48	Live Load Testing	1	03-Jun-16	03-Jun-16	03-Jun-16	03-Jun-16	0	0
49	Dead Load Testing	1	03-Jun-16	03-Jun-16	03-Jun-16	03-Jun-16	0	0
50	Geotechnical testing for water tanks	1	02-Jun-16	02-Jun-16	13-Jun-16	13-Jun-16	0	11

Figure 8. Project Activities List

Activity	Detailed Description	Value	Unit	QTY	Total (Value x QTY)	Unit Price	TOTAL PRICE
				1- Structural			
Site Work and Excavation							
To receive foundations, pile caps and ground beams; depth not exceeding 1 m	500x14 m² shelters and 500x21 m² shelters, below each 1 m of excavation	17500	m³	1.00	17500.00	$ 9.71	$ 169,925.00
Trenches for service pipes, drain pipes, cables or the like including disposal and filling	Trenches to suit pipes not exceeding 200 mm diameter; avg depth 1 m	3000	length of main pipes (m)	3000.00	3000.00	$ 5.81	$ 17,430.00
Renting equipment	Hydraulic excavator (backhoe)	10 (2 per cluster, for 30 days)	day	300.00	300.00	$ 250.00	$ 75,000.00
Foundations							
Foundations, combined and isolated bases: 150 to 300 mm thick	14 m² area and 0.5 m thickness	7	m³	500.00	3500.00	$ 75.00	$ 262,500.00
Foundations, combined and isolated bases: 150 to 300 mm thick	21 m² area and 0.5 m thickness	11	m³	500.00	5500.00	$ 75.00	$ 412,500.00
Framing							
Douglas-fir white wood (ceiling joists, roofing and exterior walls)	14 m² shelter	3.06	m³	500.00	1528.13	$ 110.00	$ 168,094.16
Douglas-fir white wood (ceiling joists, roofing and exterior walls)	21 m² shelter	2.04	m³	500.00	1018.75	$ 110.00	$ 112,062.78
Roof plastic sheathing	assuming each roll is 1 m wide and 300 m long, the shelter needs three passes, and the roll covers up to 20 shelters	1	roll	50.00	50.00	$ 50.00	$ 2,500.00
Prefab Structures							
Storage center		60	m²	5	300	$ 50.00	$ 15,000.00
Food center		100	m²	10	1000	$ 50.00	$ 50,000.00
Concrete Canvas							
Imported sets from UK	Used as classrooms and clinics	1	piece	15.00	15	$ 20,000.00	$ 300,000.00
Sanitaryware, complete with fittings							
Showers	32 showers per cluster	32	pieces per cluster	5.00	160.00	$ 300.00	$ 48,000.00
Latrines	Sink	100	per cluster	5.00	500.00	$ 20.00	$ 10,000.00
	Dump composte tank under each seat	100	per cluster	5.00	500.00	$ 20.00	$ 10,000.00
	Toilet seat	100	per cluster	5.00	500.00	$ 10.00	$ 5,000.00
	Outer envelope of the toilet chamber (100 latrines per cluster)	100	pieces per cluster	5.00	500.00	$ 150.00	$ 75,000.00
Tanks							
Wall + Columns	for the 150 m³ tank	25	m²	1.00	25.00	$ 210.00	$ 5,250.00
Top	for the 150 m³ tank	20	m²	1.00	20.00	$ 210.00	$ 4,200.00
Bottom	for the 150 m³ tank	20	m²	1.00	20.00	$ 210.00	$ 4,200.00
Steel reinforcement	for the 150 m³ tank	9	ton	1.00	9.00	$ 480.00	$ 4,320.00
Plastic water storage tanks with lids	114 litres, one for each family	100	tank per cluster	5.00	500.00	$ 54.05	$ 27,025.00
				2-Electrical			
Solar panels	Capacity: 200 W			500.00		$ 450.00	$ 225,000.00
				3-Specialities			
Insulation	Styrofoam insulation sheets (5 cm thick)	4	sheet	500.00	2000.00	$ 23.00	$ 46,000.00
Clinics	Health center (50 m²/cluster)	50	m²	5.00	250.00	$ 1,612.00	$ 403,000.00
				4-Mechanical			
		Linear feet of piping distance	Unit	Material Cost per LF	Installation Cost per LF	Unit Price	Total Price
Piping - Installed	1" Gray PVC pipe	9843	LF	2.2	10.65	$ 12.85	$ 126,482.55

5-Security							
Security (Alarm System)	Detailed Description	Value	Unit	QTY	Total (Value x QTY)	Unit Price	TOTAL PRICE
	Burglar Alarm (for storage areas)	1	piece	5.00	5.00	$ 340.00	$ 1,700.00
	Smoke detectors (in food area and classrooms	1	piece	15.00	15.00	$ 51.00	$ 765.00
	Fencing (chain link, galvanized)	1	per LF	1.00	5905.00	$ 9.70	$ 57,278.50
	Fire Fighting equipment (13A, 9 litre extinguishers)	1	per 200 m²	88.00	88.00	$ 50.00	$ 4,400.00
6-Finishes							
Shelter Doors	Plain wood	1	piece	1000.00	1000.00	$ 50.00	$ 50,000.00
Swings	For recreational purposes	1	piece	50.00	50.00	$ 80.00	$ 4,000.00
7-Labor Ressources							
		Number of days		Qty of workers	Hours	Employers cost of labor per day	Total Cost
General Labourer	na	120	days	20.00	2400.00	$ 20.00	$ 48,000.00
Groundworks Labourers	na	30	days	20.00	600.00	$ 30.00	$ 18,000.00
Plumber	na	60	days	5.00	300.00	$ 30.00	$ 9,000.00
Electrician	na	20	days	5.00	100.00	$ 25.00	$ 2,500.00
						TOTAL	$2,774,132.99

Figure 9. Detailed Cost Estimation for Camp Implementation

Development and Implementation Plans of a Healthcare Center in an Underserved Area

Wassim Bou Khzam - Ali Doughan - Houssam El Rifai - Nader Hoballah - Julien Ragheb

Department of Civil & Environmental Engineering

American University of Beirut

Bliss Street, P.O. Box 11-0236, Beirut, Lebanon

wwb05@mail.aub.edu   ard08@mail.aub.edu   hhe15@mail.aub.edu   njh07@mail.aub.edu   jnr05@mail.aub.edu

Lebanon's advanced medical facilities are generally located in its urban centers, primarily in Beirut. These facilities are generally privatized, and many average Lebanese citizens cannot afford privatized treatment. Recent regional developments have resulted in a significant increase in population, resulting in existing medical facilities being stretched thin. Population increases have been concentrated in rural areas, such as the Bekaa region, resulting in already underserved regions being further strained. Although extensive efforts have been exerted by non-governmental organizations, the Lebanese government has proposed no concrete solution. The purpose of this Final Year Project is the proposed construction of a polyclinic in the underserved neighboring towns of Anjar and Majdel Anjar, located in the Bekaa region. The combined populations of the two towns approximately doubled in the past few years, while existing medical facilities have not expanded, justifying the need for a new facility to help meet a potentially doubled demand. After conducting a visit to the region, medical demands specific to underserved regions with similar socioeconomic conditions. Consultations were held with the Department of External Affairs at the American University of Beirut Medical Center to determine the design capacity of the facility. The proposal of a polyclinic includes architectural and structural blueprints, treatment of the existing water supply in the area, design of parking facilities for the polyclinic, and finally management of the clinic to ensure the financial feasibility of the project. Findings of the project have shown the area of the proposed facility to be approximately 1800m^2, split over two floors, with shallow foundations being suitable for the design loads. The complete Final Year Project, if launched as a real-life project in the region, will satisfy both current and long-term demands of Anjar and Majdel Anjar. Furthermore, the complete design could be replicated with minor adjustments to fit the needs of other underserved regions.

I.   INTRODUCTION

Lebanon's advanced medical facilities are generally located in its urban centers, primarily in Beirut. These facilities are generally privatized, and many average Lebanese citizens cannot afford privatized treatment. Recent regional developments have resulted in a significant increase in population; as of 2016, the Syrian civil war is entering its fifth year, resulting in the influx of over one million Syrian refugees into Lebanese territory, effectively increasing the country's population by nearly 25%.

The towns of Anjar and Majdel Anjar, located in the Zahle region of the Bekaa, comprise approximately 40,000 Lebanese inhabitants. Medical facilities in the region are relatively small and unsophisticated, comprised of small facilities and modestly funded polyclinics; the capacity of the existing facilities appears to have been designed based on the demand of the Lebanese population with little room for growth.

The population of Anjar and Majdel Anjar doubled to approximately 80,000 due to refugee influx, severely reducing the operational level of service of the existing medical facilities. The area surrounding the two towns, which was already at or above capacity prior to its sharp increase in population, became even more severely underserved. As a result, the region of Anjar and Majdel Anjar appeared to be an ideal location for the construction of a much-needed medical facility.

The purpose of this Final Year Project is to develop the design of an essential medical facility in an underserved region. The efforts of the facility will be focused on the demands specific to underserved regions, with minor adjustments specific to Anjar and Majdel Anjar, determined through a site visit to the region. The finalized facility follows a design based on two floors, with a designated parking lot above ground.

While wealthier members of the targeted community can seek healthcare in advanced medical facilities outside of the region, the majority of the population that composes the middle and lower economic classes generally depend on government assistance or subsidized services provided from polyclinics. In order to appeal to the economic constraints of the majority of the target population, the facility will provide subsidized care, and funding will likely be acquired through subsidies from the Lebanese Ministry of Health, rather than through the typical revenue of privatized healthcare. Since the goal of the facility is not profit maximization, institution construction and operating costs will need to be carefully considered. Thus, the facility will be developed using the most economically efficient engineering solutions.

The major civil engineering disciplines that will be covered in this project are:
- Management, in studying the costs, operation, and economic feasibility of the facility
- Structural, including the design and architectural layout of the facility

- Water treatment, in assessing water quality in the region and proposing an appropriate treatment method
- Transportation, in the design of the facility's designated parking lot

II.   LITERATURE REVIEW

With the start of Syria's violent conflict in 2011 and the rise of the Islamic State of Iraq and Syria, nearly half of the country's population has been displaced, resulting in nearly seven million displaced within Syria, and over four million refugees stranded abroad (Bixler - 2015).

The vast majority of refugees fled to neighboring countries, primarily Lebanon, Turkey, Jordan, Iraq, and Egypt. As of September 2015, Lebanon was recorded to host over one million refugees. While Iraq, Jordan, and Egypt are hosting a combined number of one million refugees (Bixler - 2015). Unsurprisingly, many of those fleeing from the conflict are in need of medical attention. Whether due to injuries sustained in the conflict, or preexisting medical conditions requiring continuous treatment or medication, many refugees are found without proper medical care.

Provision of Medical Facilities in Neighboring Countries

Various organizations such as Doctors Without Borders have taken steps to provide healthcare for refugees, both in neighboring countries and in Syria itself. A report by the World Health Organization revealed that only 49% of hospitals in Syria reported themselves to be fully functioning, with 21% reporting themselves not functioning and the remaining 30% not reported (WHO - 2013). However, neighboring countries have significantly less assistance from the organization; Iraq and Jordan have a combined seven health structures, and Lebanon has nine health structures, with only three receiving regular support for over 90,000 consultations (MSF - 2013).

While healthcare systems in countries such as Jordan and Turkey are already in place and functioning, they have been significantly strained by the influx of refugees depending on the state-provided healthcare (MSF - 2014). In Jordan, Doctors Without Borders collaborated with the Jordanian Ministry of Health to construct a mother and child hospital, providing maternal and pediatric consultations, among others. The hospital served two purposes: providing much-needed healthcare for the refugees, and also freeing up the Jordanian state healthcare resources by creating a healthcare facility specifically for refugees (MSF - 2014).

Doctors Without Borders is not the only organization constructing medical facilities in Jordan. In 2013, International Federation of the Red Cross and Red Crescent (IFRC) began construction of an Emergency Response Unit (ERU) hospital near a refugee camp approximately 100km from Amman. However, it is not a traditional emergency response unit facility; it is a hybrid between a permanent hospital and an emergency hospital.

Refugee Healthcare in Lebanon

As of 2015, nearly a quarter of Lebanon's population is made up of one-million-plus Syrian refugees. Lebanon has the highest number of refugees per capita in the world; an appeal by the United Nations for a $1.7B aid package to Lebanon had met less than 20% of the required funding in 2014 (Amnesty International - 2014). As a result of the lack of funding, combined with Lebanon's preexisting financial struggles, the Higher Relief Council temporarily halted Syrian refugee aid operations, including medical care, in 2012 (Amnesty International - 2014).

Findings of a 2012 report published by Doctors Without Borders on Lebanon determined there are significant gaps in healthcare provision; nearly 40% of interviewed refugees in need of treatment reported themselves unable to access a hospital. Nearly 90% of refugees in Tripoli reported receiving free healthcare, likely due to the relatively high number of hospitals in Tripoli. In less developed areas, such as the Bekaa and Wadi Khaled, refugees reporting free medical treatment composed less than 25%, indicating such regions may be ideal for the construction of a medical facility providing subsidized care (Amnesty International - 2014).

The specific types of care needed vary as well. Approximately 50% of the respondents of the survey distributed by Doctors Without Borders required medication for chronic illnesses, including high blood pressure, diabetes, arthritis and asthma, with respondents stating the availability of their necessary medications varies depending on the availability at health centers. Nearly a fourth mentioned the need for dentists, urologists, gynecologists, and financial assistance for surgeries. Nearly 80% of refugees expressed need for access to basic healthcare, with child refugees reporting a vaccination rate of only 60% (Amnesty International - 2014).

These statistics indicate the demand for specific health care treatments; while advanced surgical facilities may not be feasible for the healthcare facility, the provision medical

services such as dentistry, vaccinations, and consistent supply of chronic illness medications can meet the growing demand of refugees while still being feasible.

Ghata Initiative

In the absence of a government response plan with regards to the influx of refugees, the American University of Beirut has taken the initiative to assist the refugee population by dedicating its civic and community services to developing the Syria Relief Project (SRP). An interview with Mr. Ali Basma, a consultant at the Center of Civic Engagement and Community Service, presented the recent activities AUB has launched for providing the necessary educational, sanitary, and recreational services.

In February 2014, the Community Development and Projects Unit at the CCECS in AUB, cooperated with KAYANY Foundation for the project "*Ghata*: Bringing Education to Informal Tented Settlements." Through the use of portable classrooms within refugee settlements, this project was able to provide basic schooling for the children. Each *Ghata*, which translates to "Cover" in English, is constructed from materials locally available and easily accessible, such as wood panels and steel bars found in standard sizes. The shelter can endure severe weather conditions, while maintaining a design that is flexible enough for further modification and expansion. Most importantly, the structure is easily assembled and dissembled, and can be constructed by refugees themselves.

Such a setup is not adequate for the intended outpatient clinic in Bekaa. However, similar sustainability practices will be taken into consideration during the construction of the facility by mainly reducing energy and water consumption.

III. SCOPE AND METHODOLOGY

A. *Determination of Location and Size of Facility*

In order to determine the location of the facility, several factors were investigated. It was already well known that rural areas located far from Lebanon's main cities (Beirut, Tripoli, and Sidon, for example) were relatively underserved; population fluctuations, due to internal and regional factors, further influenced the decision on the facility's location.

The primary region of interest was the Bekaa Valley, due to its relatively rural environment and vulnerability to population fluctuations as a result of its proximity to Lebanon's borders. A site visit to the area was proposed, in order to determine the treatments most highly demanded in

underserved regions, as well as the approximate size of the target populations to be served.

To approximate the size and cost of the facility, various consultations would be held with doctors, engineers, and medical consultants at the American University of Beirut. A space program will be produced to determine the various rooms of the facility, as well as the total area of the facility.

B. *Completion of Architectural Drawings and Structural Details of the Facility*

The findings of location and size of facility discussed in the previous section will be used to produce the architectural drawings of the facility. An architect will be employed in order to produce a detailed architectural plan view of the facility, using previously mentioned space program as the primary reference. The architect will be informed of the necessary criteria and constraints present, such as the width of hallways, location of columns, and room sizes that would allow full functionality of the facility. To determine the structural integrity of the facility, ETABS and SAFE software simulations will be used to check the adequacy and design of vertical and horizontal elements, including column and slab dimensions, as well as their reinforcement.

C. *Assessment of the Water Quality in the Region and a Proposed Treatment Method*

The Litani River is one of the most important water sources in Lebanon, providing water supply for domestic, irrigation, and hydroelectric uses in the country. The Ghozayel River, one of the two main tributaries of the Litani, passes through the town of Majdel Anjar, the region of interest in this study (Litani River Authority, 2009). As a result, the scope of this study will include examining the water quality in the Litani region and determining whether it's suitable for use in the proposed medical facility.

Point source and non-point source pollutants seriously affect the Litani River Basin (U.S. Agency for International Development, 2011). Several contaminants have been discharged into both surface and groundwater, rendering the quality of water as inadequate in the region.

The main source of water supply to the medical center will rely on groundwater pumping from a legally drilled well. It is important to note that the water pumped will be subjected to proper treatment methods rendering the water suitable for medical use.

United States Agency for International Development (USAID) conducted a study on the water quality of the basin in 2010. This study tested for the presence of contaminants in the groundwater samples that were extracted from different sources of the Litani region, and consequently contaminant concentrations were determined. The findings complemented

those of the Litani Basin Management Advisory Services (BAMAS) study, which was conducted in 2005. The results of the analyzed groundwater samples confirmed the presence of zinc, mercury, iron, total coliform, E-coli, and organic matter as presented in Figure 5 of the Appendix (BAMAS, 2005). The contaminants mentioned above pose severe risks to both environmental and human health.

It is necessary to rely on a self-sufficient water treatment system to remove the contaminants previously mentioned from the groundwater. One of the most commonly adopted systems is the reverse osmosis that utilizes membranes for water purification. This process hinders the passage of suspended solids, bacteria, viruses, cysts, pesticides, multivalent ions, and monovalent ions, only allowing water H_2O to pass through. The RO system proves to be the best water treatment method, since it is compact, efficient, simply operated, and requires minimal labor.

In order to ensure the long-term performance of the RO treatment plant and achieve optimal cost-effectiveness by reducing maintenance costs, it is of primary importance to delay the fouling process by taking preventative measures through pretreatment.

Fouling is an inevitable phenomenon that will appear on the RO membrane. Failure in periodical cleaning of the system will result in severe damage that might require membrane replacement, significantly adding to the cost of operation. In order to ensure the long-term performance of the RO treatment plant and achieve optimal cost-effectiveness, it is of primary importance to take preventative measures through pretreatment. Some pretreatment processes include the use of a dual media filter (sand and anthracite), chlorination, adding anti-scalant, etc. Additionally, to ensure that the water produced by the reverse osmosis system is adequate for its intended use, post treatment is required. Methods of post treatment include adding anti-corrosion and an alkaline, or post chlorination.

D. Completion of Parking Lot Design

The parking lot for the facility will be designed using AutoCAD software. The parking facilities will be split into separate staff and patient parking lots, and will include entrance/exit points, aisles, islands, and adequate parking stalls for vehicles. Knowledge from previous transportation courses on parking requirements will be utilized in this process, abiding by design requirements and limitations

IV.   LIMITATIONS AND CONSTRAINTS

Financial Constraints
The financial constraints primarily deal with budget limitations and the allocation of funds in the development of the facility. The final product, performance, functionality, and quality of care are directly affected if the project is not well managed. Furthermore, government subsidies will be crucial in eventually recovering the initial investment in the facility's construction.

Environmental Constraints
Public concern and government regulations set specific constraints in order to protect the environment. Wastes generated by health care activities include a broad range of materials, including used syringes and needles, blood, pharmaceuticals, chemicals, medical devices and radioactive materials. Poor management of health care waste potentially exposes the community to the spread of bacteria, disease, toxic effects and injuries, as well as a risk of polluting the environment.

Political Constraints
Lebanon is currently experiencing a dire political climate, as a result of presidential vacuum and a nearly paralyzed government. Obtaining approvals and permits will not be straightforward, and registration processes will likely face delays. This will affect project proceedings, and such delays will likely result in increased costs.

Health and Safety Constraints
Due to the poor living conditions faced by the Syrian refugee population entering Lebanon, concerns are raised regarding the risk of transmission of contagious viral and bacterial infections into the country.

V.   SYSTEM DESIGN AND SIMULATION

A.   Determination of Facility Location and Size

In order to determine the location and approximate demand of the proposed medical facility, a site visit was made to a makeshift medical center designed specifically for the growing refugee population. During the visit, it was determined that the region encompassing the towns of Anjar and Majdel Anjar was severely underserved; previously established medical facilities barely managed to provide for the demand of the Lebanese population, and the refugee crisis produced a population spike of 40,000 in the past few years. As a result, the area encompassing the two towns was selected as the location of the medical facility.

In order to determine the size of the facility, it was first necessary to decide what services would be treated. During the site visit, it was made clear that a variety of clinical services were required, with a specific emphasis on gynecology, dentistry, and pediatrics. After consulting with Mr. Namir Kanaan, a mechanical engineer at the American University of Beirut Medical Center, it was determined that the hospital would ideally follow guidelines specified by the American Institute of Architects. However, it was decided that rather than attempting to replicate an American hospital in the Bekaa, a template for a hospital design in a similar environment would be more suitable.

A consultation was held with Mr. Imad Sadek, Managing Director of Administrative and Management Affairs of the American University of Beirut's Department of External Medical Affairs. Having worked on various hospital designs throughout Lebanon, Mr. Sadek was able to provide an area approximation for an advanced, full-fledged medical facility designed for construction in Erbil, Iraq. Many aspects of the hospital, such as advanced treatments beyond the scope of the proposed medical facility, were reduced or removed completely. However, the capacity of the facility for treating patients was crucial in finalizing the size of the facility.

The population of the target region was estimated to be 80,000; however, approximately half of the population was composed of non-permanent residents. In designing for future population growth, it was necessary to slightly increase the target population; however, given that half of the target population are not considered permanent residents, they were not factored in to the future growth out of concern for severe overestimation of the population's demand. After consulting with Mr. Sadek, it was determined that the facility would be designed for a population of 70,000 at an ideal level of service, while maintaining the possibility that the facility could later accommodate a larger population.

Mr. Sadek, having recently conducted a study with the Lebanese Ministry of Health, stated that on average, individuals visit medical clinics approximately four times per year. Thus, the total demand of the target population was estimated to be 280,000 clinical visits per year. Two nearby medical facilities previously existed in relatively close proximity; however, the design capacity of the proposed facility exceeds that of the existing facilities, and thus it was determined that approximately 50% of the population's demand, or 140,000 individuals would be served by the new medical facility. Mr. Sadek explained that approximately 45% of the target population would have medical coverage, while 55% would depend on government subsidies, reaching the conclusion that only about 70% of the 140,000 would seek care at the new facility, with the other 30% seeking care at more advanced facilities outside of the immediate area.

Having finalized the design capacity of the facility, adjustments were made to the area spreadsheet to produce a final total area of approximately 1900m². A large, empty plot of land was found nearly equidistant between Anjar and Majdel Anjar, with a total area of more than 60,000m². After consulting with the notary public of the region, it was determined that renting a sufficient part of the land would have a nearly negligible cost (less than $5,000 per year) compared to the total cost of the project. Although it was initially proposed for the hospital to span a single floor, it was eventually divided into two floors after considering the unnecessary added cost of a foundation for such a large area.

After consulting with Mr. Kanaan and Mr. Sadek, it was estimated that the cost per square meter of construction would be approximately $1,100. Factoring in equipment costs (estimated at approximately $1.75M USD by Mr. Sadek for the specific services provided) and design/supervision fees of approximately 10%, the total cost of the project was estimated at just over $4,000,000USD.

In determining revenue, it was necessary to determine the total number of staff, as well as the different types of staff that would be employed. Consultations with Mr. Sadek, as well as Dr. Jamal Hoballah, Chairman of the Department of Surgery at the American University of Beirut Medical Center, the distribution of number of patients per specialty was determined through the general demand of rural areas in Lebanon (i.e. significant focus on dentistry, pediatrics, gynecology, etc.). From the demand per specialty, the number of physicians per specialty was determined through a physician-population ratio document published by Practice Support Resources, Inc.

Using a general estimate of approximately $500/day of clinical services per physician, with a total of 18 physicians holding clinic between one and four days a week, total physician cost was estimated at approximately $89,000 per month. Factoring in other staff including nurses, lab technicians, janitors, and secretaries, total employee salaries reached $113,000 per month; number of staff and salaries can be seen in Tables 5 and 6 of the Appendix.

In determining revenue, base checkup costs were set at 20,000L.L. for all patients, with an additional government subsidy of 30,000L.L. for the 55% dependent on government assistance. Additional services, such as x-rays, blood tests, and dentistry produce additional revenue, with prices shown in Table 7 of the Appendix. Annual revenue totals were approximated at $1.872M USD, with annual total costs approximated at $1.406M USD.

B. Structural Design

The architectural drawings obtained, as presented in Figures 1 and 2 of the Appendix, included a general layout of the facility and a detailed plan for each floor. Sustainability aspects were implemented into our design by reducing energy consumption through the orientation of the building to maximize natural light, as well as the implementation of two glass domes to further enhance the effect. Hospital construction is generally limited to either steel or reinforced concrete; in order to build a structure resilient to changing weather conditions, fire hazards, as well as vibrations (for the sake of sensitive medical equipment), the "conventional" approach will be implemented, i.e. reinforced concrete will be used to construct the facility's skeleton. Concrete structures generally require less maintenance than steel structures. Furthermore, the majority of construction projects in Lebanon are completed through the use of reinforced concrete, for economic reasons as well as material availability (CRSI).

Specifications on material properties and construction methods were decided based on common practices in Lebanon. The concrete to be used will have a compressive strength of 30 MPa, and steel will have a yield strength of 45 MPa. The modulus of elasticity of concrete was calculated based on the ACI code. As for the construction process, isolated footings will be used for the foundations. Columns and walls will be used as vertical structural elements. As for the horizontal elements, a flat slab will be implemented in the construction of the facility. A flat slab is a two-way reinforced concrete slab that does not use beams or girders to transmit the loads from the slab to the columns. Instead, loads are directly transmitted to the columns. Using reviews from the literature and consulting with professionals in the design and construction industries, dead and live loads were taken to be 500 kg/m^2.

Manual calculations were performed in order to check the adequacy of the structure, specifically the performance of columns in compression. Assumptions made in this first iteration included taking the steel to concrete ratio equal to 1%. An ETABS simulation will determine whether or not this ratio is sufficient. The first step before starting the calculations included the determination of the tributary areas supported by each column. For every type of column, the column with the largest tributary area was taken. The ultimate load on the column, due to both dead and live loads, were calculated and compared against the axial capacity of the column. The columns that were taken into consideration were the most critical, because they were subjected to the largest loads. Initial hand calculations showed that all the columns under study are indeed adequate. This indicates columns of similar types subjected to smaller loads will also be adequate.

After importing the architectural drawing into AutoCAD, four main layers were kept so they can be imported into ETABS. They included Boundaries, walls, columns, and openings. Note that there is no layer for beams will be used. All other components in the original drawings were not imported into ETABS because they have no structural function. The imported layers were used to generate the model on ETABS. Load sets and combinations, material and section properties were then defined. In order to generate the models presented in Figures 3 and 4 of the Appendix, all the structural components of the facility were placed on top of the architectural floor plan and the load set was applied on the slab. After running the software, two major conclusions were directly made. The results showed that the deflections generated by the loads were within acceptable limits and that the columns were adequate with a steel-to-concrete ratio of 1%. This indicates that no serviceability issues will be encountered and that the structure is stable. The hand calculations verified the software run on ETABS, as all columns proved adequate and the steel ratio turned out to be the same in both trials.

The next step of structural analysis using ETABS will focus on system optimization through reducing the number and sizes of columns in order to ultimately reduce the cost of construction without risking structural integrity. Major constraints that should be avoided are one-way and two-way punching shear that might happen as a result of the reduction in contact area between columns and slabs, thus increasing the punching pressure. After changing the sizes of the columns, the new steel to concrete ratio will be calculated and rebar detailing will be presented. A similar methodology will be conducted for horizontal elements such as slabs using the SAFE software. Further findings and conclusions will be presented in the final report.

C. Design of Water Treatment System

Attempts to use various softwares for the design of the medical facility's reverse osmosis (RO) system yielded irrelevant results, due to the relatively low water demand of our facility (elaborated in the steps below). Indeed, the calculated value of the facility's daily demand suggests that a small RO unit would be sufficient in supplying the facility with treated water. However, the design for such small units required resorting to the manual system design of international RO manufacturing companies and applying it step by step mainly through tabulated data.

- Referring to Table 1 in the appendix and using the TDS concentration of the feed water (863mg/l), BWRO membrane (standard) type was selected.

- Referring to Table 2 in the Appendix and using brackish wells as feed water type, the average permeate flux was determined to be 29 L/m^2/hr.

- Based on an estimated daily user (outpatients and personnel) demand and total number of users, shown in Table 3 of the appendix, the computed daily water demand is 1.6 m^3/day.

- Using the values of the average permeate flux, the daily water demand and the membrane area of a small RO unit (8.4 m^2) obtained from Figure 6 of the appendix, we compute the number of RO elements using $N = \frac{Q}{f_{avg} \times A_m}$ which yields 1 element.

- Aiming for the highest recovery rate possible while taking into account the low TDS level (863mg/l) and the minimization of the resulting adverse affects (scale formation, osmotic pressure increase, etc.), the most suitable recovery rate is thus selected to be 75%.

- Based on the selected 75% recovery rate, tabulated correlations with number of system stages, as seen in Table 4 of the appendix, yield a 2-stage system. Moreover since the demand is relatively low, a single pressure vessel for each stage is sufficient.

Note that emergency preparedness requires the presence of a water storage tank with a calculated capacity of 3.5 m^3 to supply the facility in case of unexpected interruptions.

D. Design of Parking Lot

In designing the parking lots for the facility, various design considerations were taken into account. The entrance and exit points were located away from nearby intersections in order to prevent traffic interference. The semicircular design within the hospital entrance serves two purposes: the initial aesthetic benefit, as well as allowing incoming vehicles to circulate within the hospital grounds without exiting onto main roads. Furthermore, it prevents the queuing of vehicles on the street in order to prevent the slowing down of traffic on the main road.

For efficient use of the land, the aisle is parallel to the longest dimension of the lot with stalls on each side of the aisle. Although angled parking provides easier parking maneuvers, 90 degree stalls were implemented due to relatively low turnover rate i.e. incoming patients stay in the facility for at least half an hour. This configuration provides the maximum number of stalls and allows wider aisles and two-way circulation, resulting in shorter travel time and distance.

Abiding by the Lebanese code for parking lot design, the following restrictions were considered:
- Stall dimension: 5.2 X 1.8 m
- Aisle width= 6m to enable maneuvers into and out of stall
- Door clearance=40cm on each side of the car
- Bumper clearance= 20cm from wall

According to common practice, visitor and reserved parking spaces were separated into two different zones. Based on course materials provided from CIVE460: Highway Engineering with Dr. Fawwaz, 25 m^2 is allocated for each car in a surface parking lot for an efficient design. This figure takes into account entry, circulation, and parking spaces.

Parking lot A: 14 stalls
- Users: doctors and staff
- Area of parking= 350 m^2
- The medical facility employs 18 doctors in alternating shifts. When operating at full capacity, 11 doctors will be present in the facility. Assuming that all ten doctors have their vehicles, 3 stalls will act as a buffer.

Parking lot B: 32 stalls
- Users: patients
- Area of parking= 800m^2
- 210 patients / 8 hours = 26 patients every hour
- Assuming that every patient owns a vehicle, a minimum of 26 patient parking stalls should be provided, and the remaining 6 stalls will act as a buffer.

Another consideration taken throughout parking lot design is pedestrian access. To minimize pedestrian- vehicular conflict, raised sidewalks were provided. Although this increases cost and takes up space, it ensures pedestrian safety.

In order to provide adequate site distance and turning radii, islands at the parking access points were provided. This will control circulation, and increase safety aesthetics.

VI. CONCLUSION

In conclusion, the construction of a medical facility in an underserved region in Lebanon appears completely feasible. Due to the relatively modest size of the facility, the structural aspect of the project is traditional and straightforward. Furthermore, the water demand and purification requirements are reasonable. The economic feasibility of the project, based on preliminary investigation, suggests the entire capital investment of the project could be recovered in approximately nine years or less (in the case of increased demand). After the initial investment is recovered, the facility would likely produce annual profits of approximately $450,000.

REFERENCES

"Agonizing Choices: Syrian Refugees in Need of Health Care in Lebanon." *Amnesty International*. Amnesty International, May 2014. Web. 22 Oct. 2015.

"Availability of the Health Resources and Services at Public Hospitals in Syria." *World Health Organization*. WHO and Syrian Ministry of Health, Fall 2013. Web. 22 Oct. 2015.

Bixler, Mark, and Michael Martinez. "Migrant and Refugee Crisis: War Uproots 1 in 2 Syrians" *CNN*. Cable News Network, 11 Sept. 2015. Web. 22 Oct. 2015.

"Emergency Response Units (ERUs): Types." *International Federation of Red Cross and Red Crescent Societies*. IFRC, n.d. Web. 22 Oct. 2015.

"How Much Water Does the Average Person Use at Home per Day?" *USGS Water Science School*. United States Geological Survey, n.d. Web. 17 Apr. 2016.

"How Much Water Is Needed in Emergencies." (n.d.): n. pag. *World Health Organization*. Web. 17 Apr. 2016.

Litani Basin Management Advisory Services. *Litani Water Quality Management Report*. Rep. N.p.: U.S. Agency for International Development, 2005. Print.

"MSF Opens Mother and Child Hospital for Syrian Refugees in Jordan." *Médecins Sans Frontières Canada*. Médecins Sans Frontières, 17 Mar. 2014. Web. 22 Oct. 2015.

"Syria: Latest MSF Updates." *MSF USA*. Medecins Sans Frontières, 26 Dec. 2013. Web. 22 Oct. 201

APPENDIX

Figure 1 Medical Facility Layout

Figure 2 Medical Facility Floor Plans
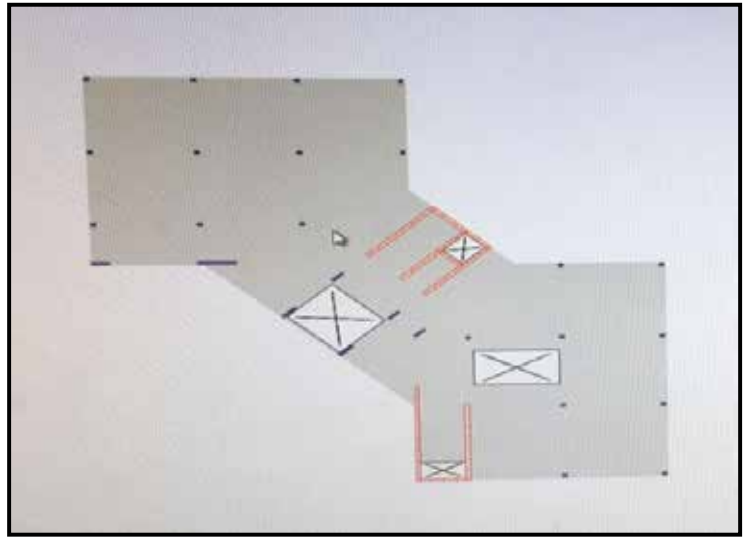
Figure 3 Structural Floor Plan on ETABS
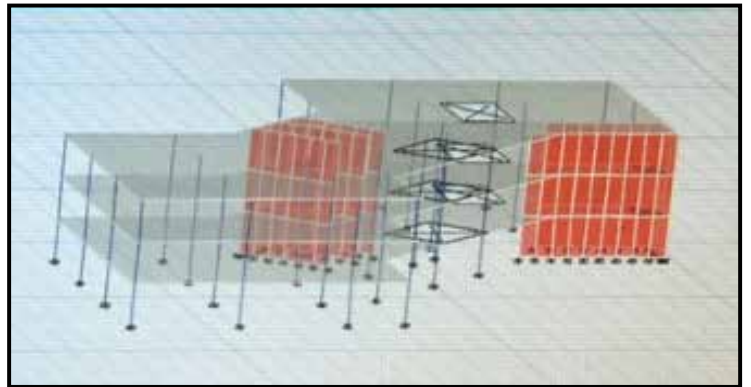
Figure 4 3D Model of Structure on ETABS

Figure 5 Groundwater Quality Profile

Figure 6 Lewabrane 4" Membrane Specifications

Table 1 Association of Membrane Type for Different Water Salinity

Low conc. Brackish water (up to 500mg/l) :	BWRO (Low energy)
Brackish water (up to 5,000 mg/l) :	BWRO (Standard)
Brackish water (more than 5,000 mg/l), Seawater :	SWRO

Table 2 Correlation Between Feed Water Type and Flow Parameters

Feed Water Type	Average Permeate Flux	Lead Element Permeate Flux	Concen-trate Flow Rate	Feed Flow Rate	Pressure Drop per Vessel	Element Recovery Rate	Flux Decline Rate	Salt Passage Increase
Municipal Supply	22 (19-25)	< 32	> 3.6	< 13	< 2.0	< 15	> 10	> 10
Brackish Wells	29 (25-33)	< 43	> 3.0	< 16	< 3.0	< 20	> 7	> 10
Surface Water Media Filtration	21 (18-24)	< 31	> 3.6	< 13	< 2.0	< 15	> 10	> 10
Surface Water MF/UF Filtration	27 (23-29)	< 39	> 3.0	< 15	< 3.0	< 17	> 7	> 10
Secondary Waste Media Filtration	11 (9-13)	< 19	> 3.6	< 12	< 2.0	< 12	> 15	> 15
Secondary Waste MF/UF Filtration	17 (13-20)	< 25	> 3.6	< 13	< 2.0	< 13	> 15	> 10
Seawater Intake Media Filtration	14 (12-16)	< 28	> 3.6	< 14	< 2.0	< 10	> 10	> 10
Seawater Intake MF/UF Filtration	17 (15-19)	< 32	> 3.6	< 14	< 3.0	< 12	> 10	> 10
Seawater Beach Wells	17 (15-19)	< 35	> 3.6	< 15	< 3.0	< 12	> 7	> 10
RO Permeate	35 (30-39)	< 48	> 2.4	< 17	< 3.0	< 30	> 3	> 5

Table 3 Computation of Daily Water Demand

Estimated daily demand per user:	Estimated daily number of users:	Volume of water per toilet flush	Estimated daily number of users using toilet flush	Total demand:
5L/day	200	6L	100	200x5 + 100*6 = 1600 l/day = 1.6 m3/day

Table 4 Percent Recovery versus Number of System Stages

1 stage system: < 50%
- Usual recovery SWRO (< 50%)

2 stage system: < 75-80%
- Usual recovery BWRO (< 80%)
- High recovery SWRO (< 60%)
- High recovery 2nd pass (< 90%)

3 stage system: < 85-90%
- High recovery BWRO (< 90%)
- High recovery 2nd pass (< 95%) (special case)

SWRO: seawater desalination, BWRO: Brackish water desalination

Table 5 Physician Salaries

Physician Type	Number	Daily Wage	No. of Days/Week	Total Cost
Urologist	1	$500	2	$4,000
Gastroenterologist	1	$500	4	$8,000
General Surgeon	1	$500	1	$2,000
Neurologist	1	$500	1	$2,000
OBGYN	2	$500	2	$8,000
Pediatrician	3	$500	1.5	$9,000
Dentist	2	$500	2	$8,000
Dermatologist	1	$500	2	$4,000
ENT	1	$500	3.5	$7,000
Ophthalmologist	1	$500	4	$8,000
Cardiologist	1	$500	2.5	$5,000
Internist	3	$400	5	$24,000

Table 6 Salaries of Non-Medical Staff

Non-Medical Staff	Number	Monthly Wage	Total Cost
Lab Technician	8	$1,000	$8,000
Janitor	6	$600	$3,600
Secretary	8	$800	$6,400
Nurse	6	$1,000	$6,000

Table 7 Sources of Revenue

Sources of Revenue	Number	Price	Total
Checkup	65333	$13.33	$870,893.33
X-Ray	9800	$20.00	$196,000.00
Subsidies	35933	$20.00	$718,666.67
Lab/Blood Tests	13067	$6.67	$87,154.67
Dentistry	3394	$10.00	$33,939.39

Elevated Light Rail System from Tabarja to Beirut

Lara El Baaklini, Michel Maroun, Jimmy Massaad, Joe Sassine, Danielle Stephan

Department of Civil and Environmental Engineering

American University of Beirut

Beirut, Lebanon

Lse09@mail.aub.edu, mcm08@mail.aub.edu, jjm17@mail.aub.edu, jrs15@mail.aub.edu, dys02@mail.aub.edu

Abstract- In cities all over the country, traffic congestion, which often leads to accidents, national economic losses, and increased vehicular emissions, can be found like clockwork. This paper aims to alleviate the traffic congestion found on the 24 km stretch from Tabarja, Lebanon to St. Michel, Beirut, Lebanon by studying and designing an elevated light rail transit (LRT) system. A mass plan including station locations and a horizontal alignment for the railway system are proposed. The factors governing the operational characteristics of the transit line and every relevant aspect are culminated in a comprehensive summary containing the data required to properly run the system. The gathered peak-hour demand data was also used and adjusted for non-peak and low demand periods in order to design the system using guidelines provided by the Transit Cooperative Research Program (TCRP) and Transit Capacity and Quality of Service (TCQS) Manual. This demand data were incorporated in a dynamic model that calculates the travel speed, frequency and station schedule of trains. The travel characteristics were computed to meet the current demand and accommodate the forecasted demand on the system until 2030. Lastly, a fixed block control system that divides the track into blocks of variable lengths has been investigated to ensure safe operations.

Environmental sustainability, structural functionality, and construction feasibility were the main foci of the structural design of the post-tensioned-box-girder-segmented-bridge system proposed for the LRT. The various sections and structural elements were designed in accordance with the guidelines found in the American Association of State Highway and Transportation Officials (AASHTO), Precast Concrete Institute (PCI), American Concrete Institute (ACI), American Society of Civil Engineers (ASCE), the Federal Transit Administration (FTA), and Federal Highway Administration (FHA) of the United States of America. A full structural analysis was conducted using SAP2000, Reinforced Concrete Members ACI-Builder, and SPColumn in order to analyze the associated dead load, moving load, and earthquake load combinations. Accordingly, complete post-tension and reinforced concrete designs/models of the superstructure and substructure elements are provided within their respective sections in the paper.

As the installation of a LRT system affects the environment, measures will be implemented in order to mitigate effects on the surrounding area. The critical aspects investigated will be related to emitted air and noise pollution. The train will be designed to minimize energy consumption by employing modern braking methods that save and reuse energy – a crucial element for the efficient use of natural resources. Also, the train crosses several densely-populated areas, such as Jounieh, Lebanon, in which stricter measures will be applied to reduce the costs for residents. Construction practicability is also taken into account according to project-specific costs to evaluate the advantages of the system and economic feasibility.

INTRODUCTION

Traffic has been a major concern for citizens commuting on the roads all over Lebanon. In cities all over the country, traffic congestion, which often leads to road rage and accidents, can be found like clockwork. One of the most severe areas of traffic is located on the 24km kilometer stretch of highway from Tabarja to Beirut. With one in three citizens owning cars and a projected 5 million daily-motorized trips within Beirut alone (Choueiri, 2010), it is no surprise driving in Lebanon leads to unhealthy competition for the same deteriorating stretch of asphalt and extra movement. With little to no room for expansion on much of the roadway, an elevated light rail system is proposed to help decrease the traffic and congestion on the road.

Many different alternatives were drawn up to reduce the problem of traffic from Tabarja to Beirut including expanding the highway, creating an underground railway system, creating an on-grade railway system or an elevated light rail system. Expanding the highway was quickly ruled out due to the little usable land to execute such a development. Similarly, an on-grade system is not suitable due to geographic and demographic restrictions and an underground railway system would be prohibitively expensive. Fortunately, an elevated light rail system gave the most promise to be both cost-effective and efficient at reducing traffic due to the fact that it requires less land and is easily adaptable to the dense demographic and geographic nature of the road.

The Elevated Light Rail System from Tabarja to Beirut will employ the use of an alternative mode of transportation not yet used in Lebanon, an elevated light rail. This railway system works similarly to a normal light rail system (using railcars to transport the urban population to their destination), but will be elevated on piers along the coast to allow for adequate space of the project. The light rail will be composed of multiple lanes of railcars, providing mass transportation to and from Beirut and Tabarja. The implemented design will be easily adaptable to the overall larger railway system which would ideally run to and from Tripoli and Sidon. The following sections explore the structural design, geotechnical design, transportation analysis, environmental impacts, and initial investment costs associated with the project.

STRUCTURAL DESIGN

A. Pre-design Planning

In order to develop the structural system for the 24km stretch from Tabarja to St. Michel, an innovative bridge

structure with longer spans and increased construction flexibility must be considered. A post-tensioned-box-girder-segmented-bridge was proposed for the elevated LRT system and developed in accordance with the proper structural design codes and specific load cases provided by AASHTO LRFD for strength and extreme limit states. High-strength, anti-corrosive concrete with a compressive strength of 40MPa was designated as the main structural material in order to facilitate construction through cast-in-place and precast methods, increase load resistance through earlier strength attainment, and resist the corrosive elements characteristic of coastal areas. A 40m simply-supported span supported by single piers was chosen, based on the 36m maximum single span of the Dubai Metro.

B. Load Considerations

The elevated LRT system consists of two steel track lanes for occupied train travel on a 5cm reinforced concrete (RC) wearing surface casted atop the bridge deck. The Bombardier MF2000 train model of the Paris Metro was chosen in coherence with a direct fixation track to define the exact dimensions and load values, as well as the top slab clearance widths required. The self-weight and capacity of the train were used in order to calculate the empty train and occupied train load values. The geometry of the train, consisting of five railway cars, was modeled into point loads simulating the contact points between the train wheels and tracks, as shown in the profile view in Figure 1. Impact factors including a dynamic load allowance due to surface imperfections and characteristics of the vehicle suspension, as well as a vehicular breaking force, were factored into the live load combinations throughout the analysis. The loads, 2D modeling application classifications, and service values are summarized in Table 1.
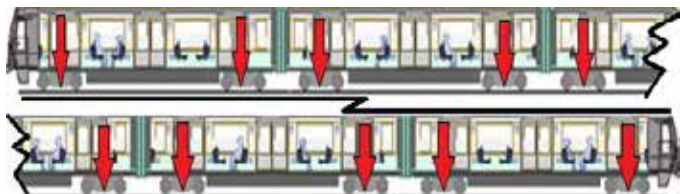

Figure 1
TRAIN PROFILE VIEW

Table 1
LOADS AND APPLICATIONS

Load Case	Finite Element Method Modeling	Class	Magnitude
One Empty Train	10 Point Loads	Live	12.4T Each
One Full Train	10 Point Loads	Live	16.3T Each
Railway Tracks	Distributed Load	Dead	0.2T/m
Superstructure Self Weight	Distributed Load	Dead	14.7T/m
RC Surface	Distributed Load	Dead	1.1T/m

C. Superstructure Design

a. Box Girder Geometry

The top slab width of the box girder was determined through dynamic train clearance values found in section 3.8 of the TCRP Report 155- Track Design Handbook for Light Rail Transit. The top slab width was found to be 8.9m, not

including the trapezoidal parapets with a 1.0m height, 0.6m base width, and 0.3m top width. The remaining cross section member sizes of the box girder were found using section 5.3.3 of the U.S.A. FHA Post-Tensioned Box Girder Design Manual. Figure 2 illustrates the final box girder geometry and Table A1 describes all cross-sectional characteristics. The dimensions were used to determine all necessary section properties and analysis inputs through Excel spreadsheet equations.

b. Shear Keys

Shear keys were added to the top slab, bottom slab, and webs of the box girder cross-section as seen in Figure 2. Shear keys were used to connect the units and provide stronger joints through a puzzled connection of indentations and keys. Research conducted by Dr. Rombach of the Technical University of Hamburg (2002) provides evidence for shear key resistance to shear forces, torsional forces, and earthquake effects, as well as geometric design schemes for key placement.

c. Longitudinal Load Analysis

A longitudinal load analysis was conducted along the 40m span to determine the critical bending moment and shear force. Impact factors were added to the live load combinations to take into account vertical oscillations in the system. Excel spreadsheet coding with embedded moment and shear influence lines propagated the factored series of point loads along the span and outputted the critical demand for bending moments (4432T-m) and shear (623T).

d. Transverse Load Analysis

A transverse load analysis to design the deck was conducted using SAP2000 software according to section 8.5 of the U.S.A. FHA Post-Tensioned Box Girder Design Manual. Figure A1 in the appendix illustrates the SAP2000 model of the cross-section and a typical bending moment diagram due to loading. The AASHTO LRFD Service 1 load combination was used in coherence with three factored live load cases: two full trains on the section, one full train alone on the section, and one full train with one empty train in order to analyze and redistribute loads on interior strips designated by AASHTO LRFD Article 4.6.2.1. Table A2 summarizes the analysis moments and classifies them as positive top slab moments, negative top slab moments, and negative cantilever moments.
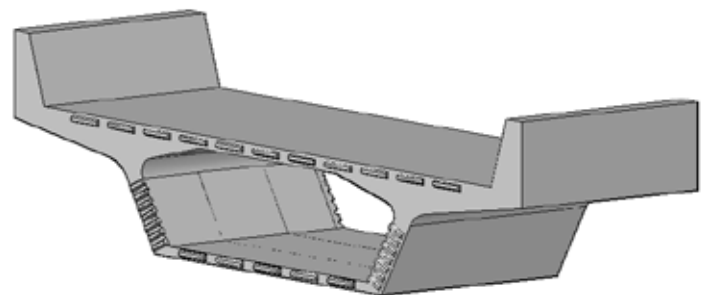

Figure 2
BOX GIRDER GEOMETRY

e. Longitudinal Post-Tension Design

An Excel spreadsheet with the post-tension equations needed to design the 40m span was used to evaluate post-tensioning at the transfer of force and at service in the future. The output consisted of using a 9440T force to pre-stress 546

15mm strands (1.4cm^2) placed within 14 circular ducts of 14cm diameter (136.5cm^2). Deviator ducts are used to guide the external tendons throughout the system and provide structural integrity to the span, as seen in Figure 3.

  f.   *Transverse Post-Tension Design*

The transverse post-tension design was performed in accordance with AASHTO LRFD 4.6.2.1.3 and the values found in Table A2. The required pre-stressing force per one meter run of interior strip is 98T for 4 13mm (1.3cm2) strands within oval ducts. Figure 4 illustrates the post-tensioning within the top slab and Figure A2 illustrates the application of a pre-stressing force on such oval ducts.

  g.   *Punching Shear and Web Shear/Torsion Analysis*

The top slab of the box girder was checked for adequacy with respect to punching shear in a pre-stressed slab. The top slab was proven to be adequate according to ACI 11.11.2.2 for the case of four wheels in the same region and one wheel alone. The webs of the box girder were designed according to AASHTO 5.8.6.3 and 5.8.6.5 for shear and torsion with the related reinforcement details, shown in Table A3.

D. Substructure Design

  a.   *Axial Pier Analysis*

The single piers used to support the superstructure were designed for the corresponding axial and maximum earthquake loads at the base. The piers span a 5m clear height followed by a pier cap and the overlying superstructure. The Strength Limit State 1 load case provided by AASHTO LRFD for the critical compression load on the piers was used to design for a 1520T load in compression with the output of 0.90m diameter piers and 38T20 reinforcement.

  b.   *Seismic Pier Analysis*

The single piers were then evaluated using the Uniform Building Code 97 (UBC97) earthquake guidelines and zone specific factors provided by Huijer et al. (2015) in order to calculate the maximum bending moment about the column base. A maximum bending moment of 1297T-m was calculated through an Excel spreadsheet analysis.
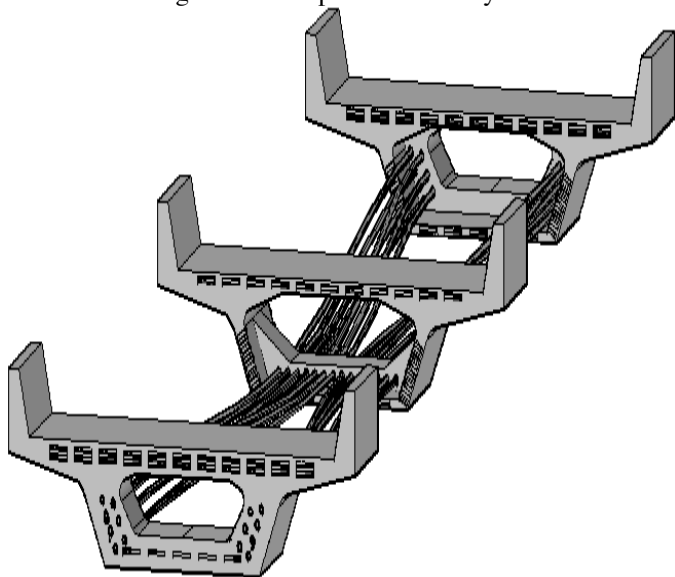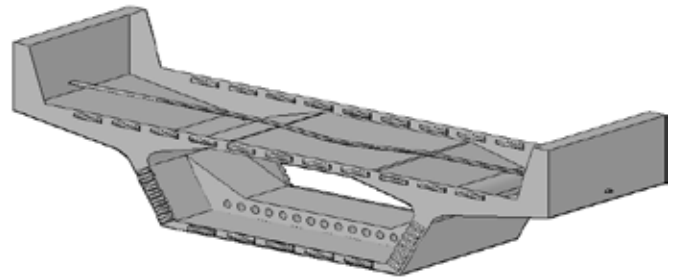

Figure 3
LONGITUDINAL POST-TENSION TENDONS

Figure 4
TRANSVERSE POST-TENSION TENDON

  c.   *Pier Final Design*

The final pier design was conducted using the AASHTO LRFD Extreme Event 1 Limit State with SPColumn software. The resultant pier is shown in Figure 5, a 1.6m diameter circular pier with 40T40 (2.50%) and spiral tie reinforcement, as seen in Figure A3. The resultant load-moment strength interaction diagram for the system is shown in Figure A4.

  d.   *Foundation Design*

A soil profile that represents poor soil conditions the system runs through was constructed through borehole logs retrieved for Antelias, Lebanon (Advanced Construction Technology Services [ACTS] SAL). The soil profile, found in Figure A5, was used to design the piles and pile cap to support the above-ground elements. Axial service loads and uplift/compression loads due to earthquake effects were taken into account with traditional geotechnical factors of safety. After integrating through the soil profile and employing spreadsheet analyses, six drilled shafts with a 1.6m diameter and 27m length were chosen. The pile cap was then designed with a length and width of 14.0m and 10.0m respectively to satisfy the minimum clearances between shafts and the shafts and edge of cap. The pile cap depth was designed through modeling on RC Members ACI-Builder in order to adequately resist bending of the cap and punching shear from the overlying pier, resulting in a 1.7m pile cap depth. The drilled shafts were checked for possible tension due to the earthquake uplift effect and designed with a minimum 1% steel reinforcement due to the lack of net tension. Steel reinforcement of the pile cap was also designed in order to resist the bending and punching forces, as seen in Figure A6. The overall structural design is illustrated in Figure 6.
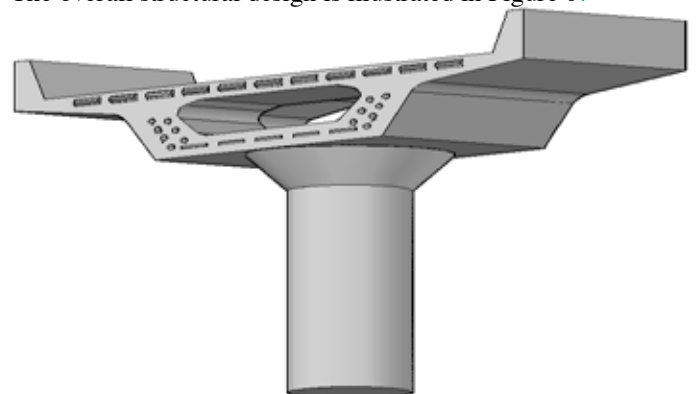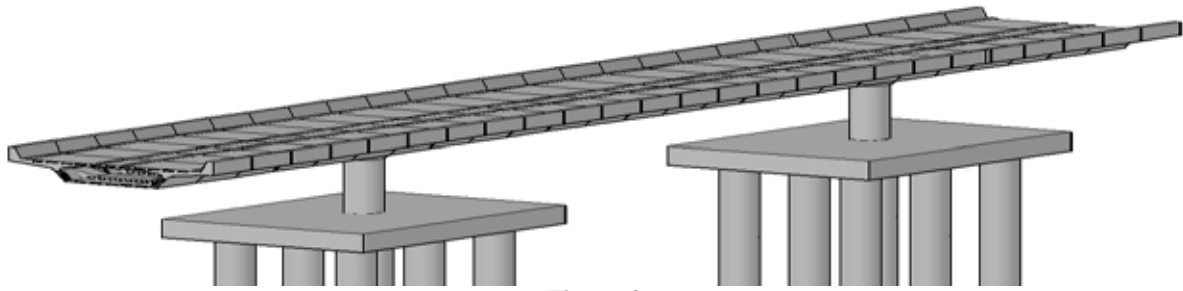

Figure 5
RESULTANT PIER

Figure 6
OVERALL STRUCTURAL SYSTEM

TRANSPORTATION ANALYSIS

I. *Data Collection and Adjustments*

In order to start the analysis, ridership data was obtained through a 2012 Egis-Group report commissioned by the Council for Development and Reconstruction (CDR). The report provided demand data for various stations along a coastal route in Lebanon for a "weekday peak hour" during the years 2020 and 2030. The obtained data was developed in order to account for different congestion hours throughout the week and more efficiently placed train stations by distributing demand through tributary areas of ridership, as seen in Figure 7. The fluctuating nature of transportation demand throughout the hours of a day and different days of the week are is taken into account through redistribution coefficients diving the day into peak, non-peak and low demand hours, as seen in Table A3. The week was divided into weekdays and weekends and Excel spreadsheets were used to develop the data into a more accurate representation of reality. Table A4 provides the coefficients to adjust demand relative to weekdays.

Figure 7
STATIONS ALONG COAST

II. *Operational Characteristics*

a. *Fixed Block Control System*

The third edition of the TCQS Manual was used to determine the operational characteristics of the system. Having analyzed the control modes available, a fixed block system (FBS) was chosen to ensure safe operations of the trains. The FBS divides the track into 600m long blocks/spans with incorporated traffic signals at the beginning of each block, ensuring a safe minimum separation of two blocks between trains. Figure 8 illustrates the system in motion. A green light indicates there are two or more blocks separating trains, a yellow light indicates there is only one block of separation and informs the conductor to decelerate to 30km/h from a maximum design speed of 70km/h, and the red light alerts the conductor of a train presence in the same block and urges the conductor to stop immediately, waiting for further directions through a yellow caution light.
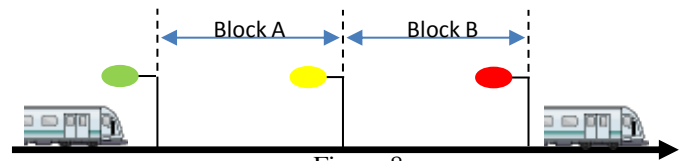

Figure 8
FIXED BLOCK SYSTEM

b. *Non-Interference Headway*

The non-interference headway is defined as the minimum safe separation between two adjacent trains, composed of the minimum headway, dwell time and operating margin. Equation A1 provided by section 8.5 of the TCQS Manual was iterated in order to calculate the minimum headway time between two trains of 43 seconds for an optimal station approach speed of 23 km/h (6.389 m/s). Figure A7 illustrates the iterated relation between the minimum headway and approach speed. Then, using Section 8.5 of the TCQS Manual, a conservative average dwell time of 45 seconds per station and operating margin of 25 seconds were determined. The total non-interference headway was calculated by summing up the three components to obtain 113 seconds.

c. *System Capacity*

The Bombardier MF2000 train consists of five railway cars with a comfortable capacity of 581 passengers and 1000 passengers at surcharge. Taking into account the non-interference headway and single train capacity, a LRT system with the maximum number of trains satisfying the headway would be able to service demand until the year 2052, accommodating 13,881 passengers per hour with a maximum frequency of 32 trains per hour. The capacity of the system increases to 23,891 passengers per hour when surcharge values are taken into account. The mentioned characteristics were

integrated into an Excel spreadsheet model linking forecasted demand values and system properties to compute the frequency of train movement and hourly passenger capacity according to Equation A2 found in section 8.5 of the TCQS Manual. The model also calculates the number of trains needed for any given year by projecting ridership according to a population growth factor derived from the forecasted demand data. Figure 9 shows the yearly ridership and number of trains needed between the years 2020 and 2040 as calculated through the Excel spreadsheet model.
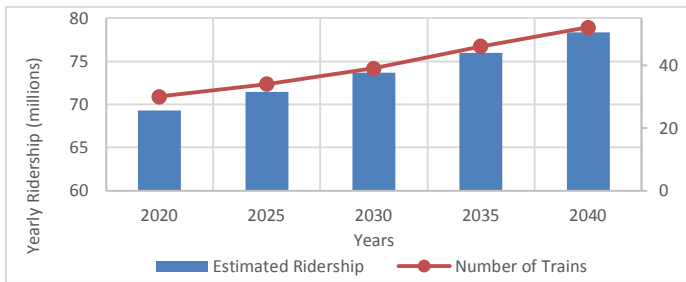

Figure 9
YEARLY RIDERSHIP AND NUMBER OF TRAINS

The frequency of the movement of trains and system capacity values are shown in Figures A8, A9, showing the frequency of trains during the different hours of the weekdays and the weekends respectively, and Table A5, showing the system's capacity for different hours from 2020 to 2040.

III.   Modes of Operation

Two modes of operations were analyzed in order to optimize the system characteristics: an alternating stopping pattern where the trains stop at every other station (used in times of peak demand for faster trips) and a pattern that stops at every station (used in times of non-peak and low demand). Using inter-station distances and the kinematic characteristics of the train due to acceleration and deceleration rates found on the Bombardier Technical Specification Sheets, the travel times required to pass from one station to the next were obtained and used to yield the total trip duration through an iterative process with the optimal train speeds. In order to find the optimal speed at which the travel time is a minimum, an optimization model was set up with the objective of minimizing total trip duration. The optimal travel speed is a function of the number of stops along the system, leading to the derivation of varying optimal speed values according to the modes of operation. Figure 10 shows the variation of trip duration for different travel speeds for both modes of operation.
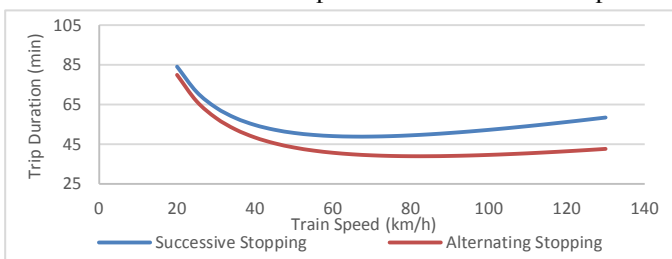

Figure 10
TRIP DURATIONS AND TRAVEL SPEEDS

For the successively stopping mode, the optimal speed was found to be 70km/h. The Bombardier MF2000 technical specifications sheet validated the optimal speed of 70km/h as the "recommended operating speed". The alternating-stopping mode's optimal speed was found to be 90km/h. In order to reduce the discomfort caused by high noise levels due to trains passing at high speeds in an urban context and to relax the design criteria of the structural elements, the maximum speed was chosen to be 70km/h for both the alternating and successive modes of operation. Also, the difference between operating the alternating train at 90km/h instead of 70km/h was found to improve the travel time by only 0.79% (from 39.31 minutes to 39 minutes). The alternating mode of operations, adopted in peak hours, sacrifices mobility: in case a passenger needs to commute between any two successive stations, they would be required to go to the nearest transfer station (Nahr El Kalb and Jal El Dib stations) and commute back to the station of choice. However, since demands during peak hours follow the trends of travelling southbound (into Beirut) during morning peak hours and northbound (out of Beirut) in the afternoon peak hours, the alternating mode of operation prioritizes the flow into and out of Beirut and therefore minimizes the time required to cross the entire system. The mode has reduced the travel time for a trip to and from the line's extremities by 10 minutes, allowing a trip from Tabarja to Beirut (and vice versa) to be completed in 39 minutes during peak hours, compared to the average 2 hour car ride. During non-peak and low demand hours, no clearly defined pattern of demand is visible and therefore accessibility and low waiting times are deemed more valuable than prioritizing the flow along the extremities. When operating under the successive stopping mode, a maximum waiting time of 5 minutes is maintained by the system. The time needed to cross the line from extremity to extremity is 49 minutes. Under this mode of operation, mobility is maximized such that passengers can board and un-board at any station.

In conclusion, the model developed takes the train and system characteristics as an input and generates the hourly ridership and operational characteristics consisting of frequency of trains, number of trains, system capacity and trip duration for any given hour and day of any year.

ENVIRONMENTAL ASPECT

The main source of air pollution in Lebanon is transportation, with an estimated 2 million cars (434 cars for every 1000 people[1]) amounting to 21% of greenhouse gas emissions in the country[2]. Implementing a LRT system will help to reduce air pollution, noise pollution, energy consumption, and vibration disturbances due to vehicles. Considering the aforementioned factors, the elevated LRT was designed in order to mitigate the

[1] *URBAN TRANSPORT DEVELOPMENT PROJECT*. (n.d.).
A Council for Development & Reconstruction Project
[2] *State of the Environment Report*. (2010).
Chapter 4

effect on the environment and decrease greenhouse gas emissions, while also being aesthetically appealing with regards to structural design and addition of greenery along the route.

I. Air Emissions

The air emissions are mainly composed of carbon dioxide (CO_2), along with particulate matter and other gases such as nitrogen oxides (NOx), sulfur dioxide (SO_2), methane (CH_4), nitrous oxide (N_2O), and carbon monoxide (CO). According to the Ministry of Environment, an average car travels 15,00km/year[3] and releases 260.4 g/km[4] of CO_2 emissions. With 2 million cars, the total amount of CO_2 released by passenger cars per year is 7.5 million tons.

The Bombardier MF2000 used in the system saves 30% more energy[5] when compared to other train models, totaling to a use of 24.5 KWh/km. In order to find the emissions produced by the LRT system, the energy needed to fuel the train is computed by setting the distance travelled by the train per day. Knowing that the system emits 0.638 kg of CO_2 per KWh[6], it produces 176 tons of CO_2 per weekday and 148 tons per day in the weekend.

In order to calculate the emissions saved by the train, the number of passengers on board per day between respective stations was found, considering peak, non-peak and low demand hours. The number of passengers was multiplied by the distance between stations. According to the relation that every 2.3 people share a car, the "equivalent" distance saved is computed. Finally, this distance was converted to carbon dioxide emissions using the respective factors. The net savings is equal to deducting the emissions produced by the system from the amount of "equivalent" emissions saved. The system saved 108 tons/weekday and 91 tons/weekend-day. This is an equivalent of approximately 37,500 tons/year. A mature tree absorbs 21.77 kg/year[7], making the emissions saved by the elevated LRT equivalent to the action of 1.7 million trees. As for the remaining air pollutants, savings of each chemical compounds were calculated for weekdays and weekends. Notable savings include 5,283kg of CO in weekdays and 4,432kg of CO in weekends.

II. Sound Pollution

The LRT traverses densely populated areas, which are affected by noise pollution and vibrations due to vehicular vertical oscillations. The desired sound level is 67dB in residential areas and adopting an LRT, in itself, mitigates the amount of noise and vibration to the surrounding areas due to the reduced vehicular traffic.

III. Energy Use

The energy required to power a Bombardier MF2000 is as low as 24.5 KWh/km due to regenerative braking, that saved 17% of energy use. Regenerative braking reuses the energy normally lost as heat due to friction. As the train decelerates, the energy saved could be used to fuel lights, reacceleration, or for other operational services. This technique is desirable for this system due to the frequent stops of a commuter train. Regenerative braking both reduces energy consumption and extends the mechanical lifespan of the braking system – which would reduce future maintenance and repair costs. Overall, the 5 KWh/km savings result in a decrease of CO_2 emissions of 35.75 tons/weekday and 30 tons/weekend-day.

INITIAL COST ESTIMATE

The initial investment cost of the system was calculated taking into account many project-specific costs. Figure 11 illustrates the cost breakdown of the $621,000,000 budget for the initial investment.
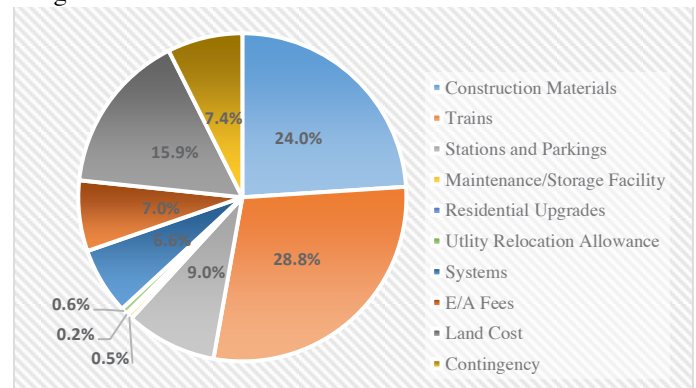

Figure 11
INITIAL INVESTMENT COST BREAKDOWN

The cost per km for the LRT system is $25,875,000 including land acquisition costs, compared to an approximate $39,000,000 cost per km for a two lane highway.

REFERENCES
[1] AASHTO LRFD Bridge Design Specifications, Customary U.S. Units. Washington, DC: American Association of State Highway and Transportation Officials, 2012. Web.
[2] CHAP. 16, DIV. III 1615 1621.3. "1997 UNIFORM BUILDING CODE." N.p., n.d. Web.
[3] Choueiry, Georges. "Analysis of Accident Patterns in Lebanon." Arab Studies Quarterly 4.4 (1982): 309-23.Http://www.4ishgd.valencia.upv.es/index_archivos/62.pdf. Web.
[4] EgisInternational. "Etude De Faisabilité D'un Système De Transport Public De Masse à Beyrouth.", Apr. 2012.
[5] Hujeir, C., M. Harajli, and S. Sadek. "UPGRADING THE SEISMIC HAZARD OF LEBANON IN LIGHT OF THE RECENT DISCOVERY OF THE OFFSHORE THRUST FAULT SYSTEM." U PGRADING THE S EISMIC H AZARD OF (n.d.): n. pag. Http://lsj.cnrs.edu.lb/wp-content/uploads/2015/12/harajli.pdf. Web.
[6] "Metro MF2000 – Paris, France - France - Bombardier." Metro MF2000 – Paris, France – France Bombardier. N.p., n.d. Web. 31 Mar. 2016.
[7] Rombach, G. "Precast segmental box girder bridges with external prestressing."(2002). http://www.ciccp.es/biblio_virtual/Precast%20segmental%20box%20girder%20bridges.pdf. Web
[8] "Specification for Steel Girder Bridges. 3B Stresses 4 Design and Construction."(n.d.):n.pag. Web.
[9] The Federal Transit Administration, Transportation Research Board. TCRP (n.d.): n. pag.Http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_rpt_165fm.pdf. Web.
[10]Track Design Handbook for Light Rail Transit. Washington, D.C.: National Academy, 2000. Web
[11]USA Department of Transportation Federal Highway Administration. "Bridges & Structures." -. N.p., Sept. 2015. Web. 31 Mar. 2016.

---

[3] M., & U. (2015). *MOBILITY COST: A Case Study for Lebanon.*

[4] *State of the Environment Report.* (2007). Chapter 5

[5] Transportation, Press release. (2007). *The RATP orders an additional batch of 49 MF 2000 Metro Trainsets for 189 Million Euros.*

[6] International Energy Agency. (2007).

[7] Discussions with Monty Maldonado, U.S. Forest Service, Forests Management, tree planting program, October 5, 2011

Appendix

Table A1
BOX GIRDER CROSS-SECTION PROPERTIES

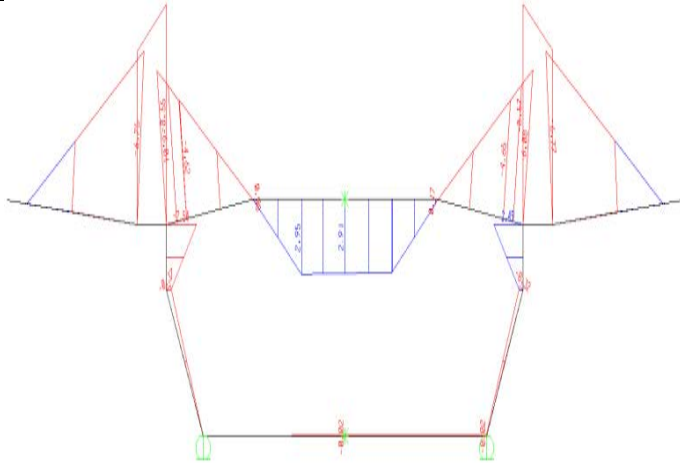| Property	Value
Cross-sectional Area	$5.88m^2$
Moment of Inertia (X)	$13.40m^4$
Moment of Inertia (Y)	$89.91m^4$
Centroid Location along X	2.25m
Centroid Location along Y	1.41m
Overall Depth	1.80m
Top Slab Thickness	0.20m
Top Slab Net Width	8.90m
Bottom Slab Thickness	0.135m
Bottom Slab Width (Top)	3.98m
Bottom Slab Width (Bottom)	4.50m
Web Thickness	0.30m
Wing Length	2.13m

Figure A1
SAP2000 TRANSVERSE MODELING

Table A2
TRANSVERSE MOMENT VALUES

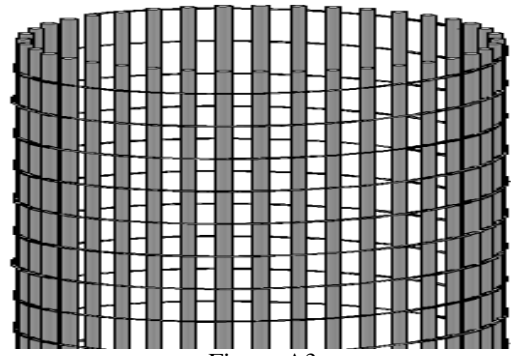| Moment Cause	Overhang (T-m)	Top Slab Edge (T-m)	Top Slab Center (T-m)
Parapet	-2.42	-0.46	-0.21
RC Surface	-0.36	-.22	0.16
Live Load	-3.04	-1.88	0.87
Metal Tracks	-.10	-0.06	0.04
Service Moment	-8.78	-3.90	1.55

Figure A2
PRESTRESSING OVAL TENDONS

Figure A3
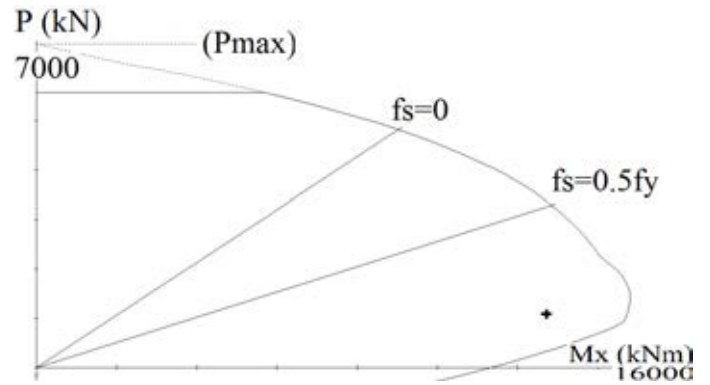PIER SPIRAL REINFORCEMENT

Figure A4
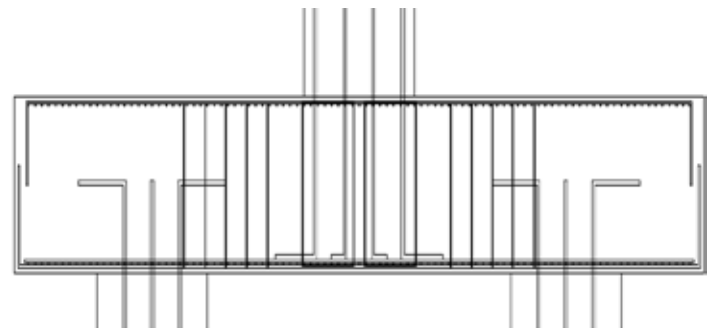P-M STRENGTH INTERACTION DIAGRAM

Figure A5
SIMPLIFIED SOIL PROFILE

Figure A6
PILE CAP REINFORCEMENT PROFILE VIEW

259

Table A3
WEEK AND DAY DISTRIBUTION OF DEMAND

Weekday	Peak	7:00 to 9:00 and 17:00 to 19:00
	Non-Peak	6:00 to 7:00, 9:00 to 17:00 and 19:00 to 21:00
	Low Demand	5:00 to 6:00 and 21:00 to 24:00
Weekend	Peak	7:00 to 9:00 and 14:00 to 18:00
	Non-Peak	6:00 to 7:00, 9:00 to 14:00 and 18:00 to 24:00
	Low Demand	N/A

Table A4
DEMAND COEFFICIENTS ACCORDING TO TIME

Time Period	Demand Coefficient
Weekdays	1
Weekends	0.75
Peak Hour	1
Non-Peak Hour	0.6
Low Demand	0.3

Equation A1
MINIMUM HEADWAY TIME

$$t_{cs} = \sqrt{\frac{2(L_t+d_{eb})}{a+a_gG_o}} + \frac{L_t}{v_a} + \left(\frac{1}{f_{br}}+b\right)\left(\frac{v_a}{2(d+a_gG_i)}\right) + \left(\frac{(a+a_gG_o)l_v^2t_{os}^2}{2(v_a)}\right)\left(1-\frac{v_a}{v_{max}}\right) + t_{os} + t_{jl} + t_{br}$$

Where:
- L_t is the longest train length
- l_v is the line voltage as percentage from specifications
- a_g is the acceleration due to gravity
- G_i is the grade into station
- G_o is the grade out of station
- d_{eb} is the distance from the front of the train to the start of the station exit block
- v_a is the station approach speed
- v_{max} is the maximum line speed
- f_{br} is the braking safety factor
- b is the separation safety factor
- t_{os} is the time for an overspeed governor to operate an automatic braking system
- t_{jl} is the time lost to braking jerk limitation
- t_{br} is the brake system reaction time
- a is the initial service acceleration
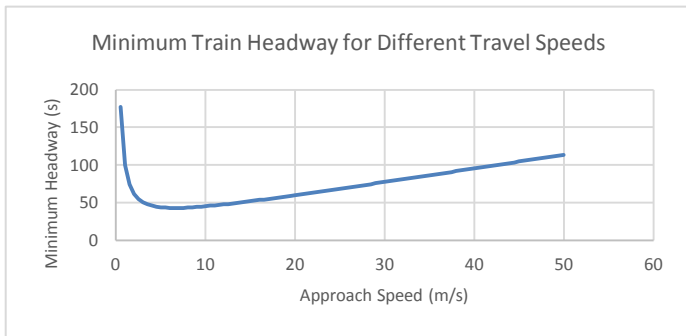- d is the initial service deceleration

Figure A7
MINIMUM TRAIN HEADWAY RELATION

Equation A2
RIDER CAPACITY

$$P = TN_C P_C (PHF)$$

Where:
- P is the design person capacity (passengers per hour)
- T is the line capacity (trains per hour)
- N_C is the number of cars per train (cars per train)
- P_C is the maximum schedule load per car (passengers per car)
- PHF is the peak hour factor
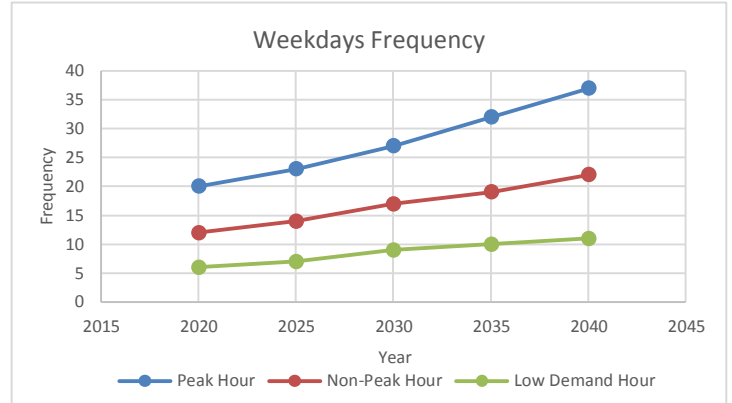
Figure A8
WEEKDAYS FREQUENCY
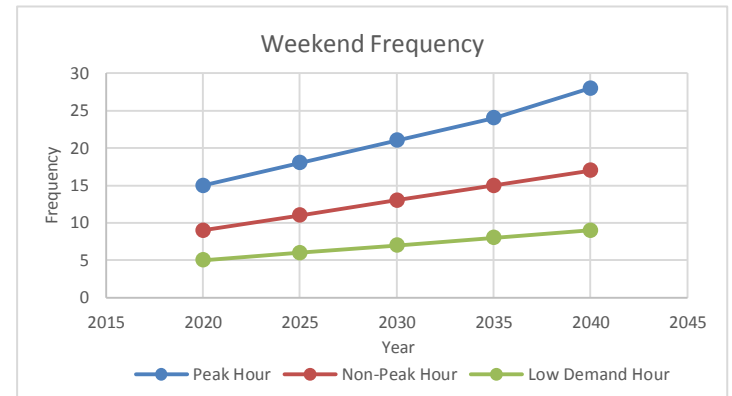
Figure A9
WEEKEND FREQUENCY

Table A5
WEEKEND (WE) AND WEEKDAY (WD) CAPACITIES

Year	WD Peak-Hour	WD Non-Peak-Hour	WD Low Demand	WE Peak-Hour	WE Non-Peak-Hour	WE Low Demand
2020	11,620	6,972	3,486	8,715	5,229	2,905
2025	13,363	8,134	4,067	10,458	6,391	3,486
2030	15,687	9,877	5,229	12,201	7,553	4,067
2035	18,592	11,039	5,810	13,944	8,715	4,648
2040	21,497	12,782	6,391	16,268	9,877	5,229

Feasibility of Establishing a Network for Air Ambulance Transfers in Lebanon

Kassar, Tania – Hakim-Rahme, Tarek – Mansour, Karim
Civil and Environmental Engineering Department
American University of Beirut
Bliss Street - Hamra, Beirut, Lebanon
tnk05@mail.aub.edu; tfh03@mail.aub.edu; kwm06@mail.aub.edu

Abstract - **With 13 traffic deaths per 100,000 capita per year, Lebanon ranks among the top-20 list of countries with the highest traffic deaths rate. It is therefore crucial to take action. Traffic Law enforcement was incapable of providing a remedy for this problem mainly because of the public unawareness. The bad cultural context regarding traffic laws, and the recurrent traffic jam problems are also to blame. Thus, a more elaborate action plan has to be developed. This project examines the feasibility of implementing the first medical aerial transport system in Lebanon. which is bound to save thousands of lives that are lost due to ambulance delays in traffic jams. The project handles two major issues: the first is the technical and structural execution of the helipads' and helicopter's network, whereas the second tackles the network operation and managerial aspect of the proposed project. This paper also illustrates a proposition for amending the Lebanese construction and urban laws in order to enforce the adoption of aerial transport through building helipads on each hospital and within each newly developed area.**

I. INTRODUCTION

Car crashes are leading causes for mortality, a study by the UNDP suggests. In fact, every year, 1.2 million people worldwide die on the road and 50 million are injured in traffic accidents [1]. A rough estimate of 15% of these deaths are either due to the delay of arrival of the ambulance to the site, or due to a delay during the transportation process of the patient to the nearest hospital. Taking a closer look at the case in Lebanon, one notices that the elite hospitals are clustered in the Greater Beirut Area (GBA) and expands gradually to Mount Lebanon district. On the other hand, a survey of car accidents per district reveals that most of the fatalities and injuries resulting from car accidents take place within Beirut and Mount Lebanon [2]. Figure 8 in Appendix A illustrates this statistical study.

In Lebanon, according to the Red Cross, 10.3% of people involved in car crashes have died in 2014 [3]. In fact, the delay due to traffic jams worsens the case of many patients that incur drastic negative health impacts Thus, implementing a system that decreases the transportation time of these victims is expected to significantly reduce the number of casualties. Accordingly, putting in place an air ambulance network consisting of helicopters, helipads and heli-stations will serve this purpose perfectly and reduce the number of casualties. This system has been implemented and tested in many developed countries and has proven to be extremely beneficial [4].

This project aims at examining the feasibility of creating a network of aerial transport that connects the different hospitals within the Greater Beirut Area, with main highways and major axis. The design stage of the medical aerial transport network includes the different physical design components such as choosing a design helicopter, designing the helipads and heli-

stations. Moreover, it studies at length the feasibility of implementing such a program, while examining its advantages and drawbacks, within the context of the different general and technical constraints.

II. LITERATURE REVIEW AND PROJECT MOTIVATION

Establishing an aerial network for emergency transportation is rather expensive and problematic in a country such as Lebanon. For that reason, a detailed analysis of the traffic accidents in this country and the design elements associated with the project is a must.

2.1 Traffic deaths in Lebanon

Statistical results of accurate quantification of the number of accidents, and the resulting casualties and injuries are substantially scarce. The problem with Lebanon is the lack of governmental care regarding such issues; however, this has been slightly modified during the last couple of years. Most of the data regarding accidents and casualties are acquired from NGO's and global traffic surveillance organizations. Table 1 below shows the number of traffic deaths per 100 000 capita, in Lebanon for the years 2008, 2009, and 2010 [2].

It it is estimated that by the year 2020 [5]:

- Road traffic deaths and injuries will experience a rise by 65% in middle and low-income countries.
- Road traffic deaths and injuries will experience a reduction of 28% in high-income countries.

Table 1 Deaths due to Road Traffic in Lebanon

Year	Deaths per 100 000 population
2008	11.2
2009	12.1
2010	12.9

Source: Section of Service and Operation, Internal Security Forces (ISF)

Before examining the reasons behind these differences across countries with varying income, it is crucial to examine the user categories of road casualties in Lebanon.

Figure 1 below illustrates the traffic deaths by road user category in Lebanon, 2010 [2].
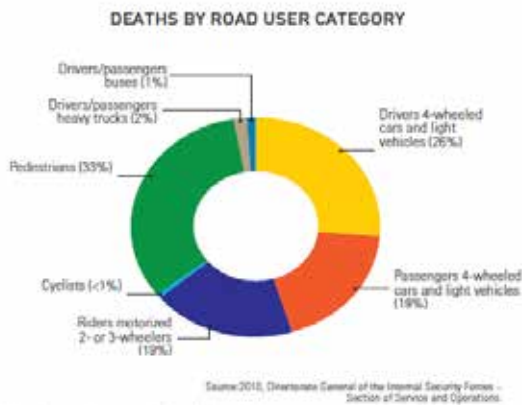
Figure 1 Death by Road User Category

It is clear that pedestrians have a substantially high share of the traffic deaths (33%), whereas buses and cyclists have negligible percentages.

2.2 Cost of Traffic Death

Traffic collisions are costing Lebanon 3.2-4.8% of its GDP, yearly.

According to the World Health Organization (WHO), the estimated GDP lost due to road traffic crashes amount to 1.5M − USD 2.2M yearly [5].

The latest report conducted by the WHO reported:

- Yearly traffic fatalities of 649 (77% males, 23% females), these are reported
- The estimation of the actual number is 1088, which is equivalent to 22.6 per 100,000 capita
- A yearly GDP loss of 3.2-4.8% which is equivalent to 1.5 to 2.5 M USD
- 43% of fatalities are pedestrians
- 1 in every 5 victims of traffic accidents dies inside the ambulance while reaching the hospital

2.3 Death Fatalities in Greater Beirut Area

Statistics showed that most of the road accidents, fatalities and injuries, occur within the greater Beirut area. Most of the hospitals located in this zone lack the existence of a helipad, and if present, it is inoperable, except for three hospitals: Rizk hospital, Hotel Dieu de France hospital and Saint George Hospital.

After analyzing data from numerous sources, and referring to the Kunhadi organization and the Lebanese ISF, a GIS mapping of the hottest accident zones within the GBA were highlighted on a map resented in Figure 2 below.

The zones Khalde, Airport road, Ouzaai, Mansourieh, Sin El Fil, Dbayeh, and Jal El Dib have the highest reported traffic accident rates per year. These zones share in common 2 features. First, they all house a major highway, on which most of these accidents occur. Second, they do not have any major hospital surrounding

them. This leads to the extension of transportation time of the victims to hospitals existing in other zones.

2.4 Relevant case studies

Two of the most important cases reflecting the success of medical air transport programs are the London and Canada programs.

2.4.1 London medical air transport program

The London Air Ambulance charity saved 1818 patients in 2013, 34% of which resulted from road traffic collisions. Research showed that the service reduced the chance of death of severely injured patients by 48% and that of serious trauma by 30-40%. (Annual Review– London's Air Ambulance, 2013-2014). The operation of this service costs $3.5million to run every year. It is partly funded by the government, but it relies mainly on corporate donors and charitable donations.

2.4.2 Canada medical air transport program

The "Canadian Journal of Surgery" published a 10 yearlong study in which it claimed that "Patients transported to hospital by helicopter have a better chance of surviving traumatic injuries than those transported by ground ambulance despite having more severe injuries and needing more surgical interventions" (Canadian Journal of Surgery, February 4, 2014). However, due to the relatively high cost of air transport, it is still limited to emergency cases that present high risks if transported using typical land means.

Figure 2- TAZ in the GBA Showing the Frequency of Reported Accidents/Year

2.5 Design helicopter

Based on literature and research, Schweizer helicopters presented themselves as the most eligible and credible models of medical helicopters. This type of helicopter comes in a wide variety of

sizes; Table 3 in the Appendix C illustrates the different models available alongside their attributes [6].

After thorough investigations and analysis, and taking into consideration the fact that this aerial transport network is new to the country, a relatively light helicopter was chosen, which has a low lease price and small dimensions that fits the existing helipads. The helicopter chosen is the Shweizer S-434, which has the following characteristics:

- Max takeoff weight: 2900 lbs
- Length: 31.2 ft
- Diameter: 27.5 ft
- Crew required for operation: 1
- Additional patients/passengers: 2-3

III.   METHODS AND OPERATION

After analyzing the issue related to traffic accidents in Lebanon, and after looking at several case studies of countries that once encountered similar problems, it became evident that action must be taken in order to save the lives of many citizens as well as reduce the costs of these accidents. Considering the fact that Lebanon is in debt, and the execution of large expensive projects is rather impossible, it was decided to limit the project to the Greater Beirut Area. In fact, the GBA houses more than 50% of the Lebanese population and scores the highest rate of reported accidents.

The solution proposed is to develop the first air ambulance transport network in Lebanon, allowing patients to be transferred directly to hospitals, bypassing the traffic of the GBA, and minimizing the transportation time of passengers that are in critical conditions.

Establishing such a network requires the following:

1. A map showing the hottest accident zones within the GBA, with the associated yearly number of fatalities and injuries.
2. A detailed course of action in cases where aerial transport is needed, and a categorization of the cases.
3. Detailed guidelines for the construction of helipads.
4. Mapping of the locations of major hospitals, and installed helipads and helistations.

3.1 Hot Accident Zone
The Lebanese Internal Security Forces (ISF), with the help of Kunhadi, recorded the location and severity of accidents across Lebanon for the last 5 years. This data was integrated in a Geographic Information System (GIS) in order to identify hot accident zones. Figure 9 in appendix B illustrates the mapping of these zones; the columns illustrate the total number of accidents per zone per year – ranging between 0 and 28, the yellow section represents the number of injuries, and the red section represents the number of fatalities (Refer to Map 2 in Appendix A).

3.2 Operation and Action in case of Emergency
In air ambulance transport, there are three relatively distinct operations:

1. Primary response: A case that is so urgent, which requires the transport of medical personnel and/or equipment to the site of the accident, as the state of the victim is highly unstable.
2. Secondary response: Direct delivery of a helicopter to the nearest helipad to the site of the crash. In this case, the victim is then transferred aerially to the nearest hospital or medical center.
3. Tertiary response: Planned urgent transfer of patients in between hospitals, due to the lack of equipment, space, or expertise at the medical center at which the patient is currently at.

The phone operator receiving the call for emergency assesses and categorizes the call. This person then calls the operation base of the air ambulance network. The base then answers the request according to the case at hand.

Operating such a network is faced by a series of constraints:

1. Helicopters are costly to procure and operate.
2. Weather conditions are a major concern, as they can ban the possibility of delivering a helicopter to the site.
3. Primary responses are limited to daytime, as the operation of a medical crew on-site is rather impossible during night-time.
4. Categorization of the case is a hard job, as people reporting accidents are often traumatized and might exaggerate in describing the on-site case.
5. Aerial transport is expensive, and if the government doesn't sponsor this service, then the patient being transferred will have a large bill on his account in order to be transferred.

Managing the aerial ambulance network requires high efficacy and good managerial role distribution, therefore, the following chart developed, dictates the chain of command required in order to operate the network. Note that this chart illustrates the base managerial hierarchy, neglecting the medical crew and pilots operating on site, as these depend on sponsorship level corresponding to the project.

Figure 3-Managerial Hierarchy at the Operation Base

3.3 Helipad construction guidelines

Such a network, even if managed perfectly, cannot operate with the current number and spread of helipads. A detailed scan of the numerous hospitals present within the GBA revealed that only three hospitals have a current fully operational helipad. Therefore, it is extremely important to impose the construction of a helipad associated with each existing hospital. Moreover, the project proposes the idea of establishing on-site helipads on several major highways and in hot accident zones, in order to answer urgent calls and bypass traffic jams.

The construction of a helipad is subject to many constraints and guidelines, regarding its dimensions and its accessories. These guidelines are illustrated in the following list, and derived from the guidelines of the ministry of health in Australia [7]:

1. The location of a helipad is very important, as it shall be placed in a high-rate accident zone, and shall be as far as possible from buildings and surrounding residential areas. Preferably on elevated area such as rooftops.
2. The minimum size of a helipad is 9x9 m, as a landing area, which is going to be set as 12x12 meters for our case, in accordance with the design helicopter chosen in the literature section.
3. A final approach and takeoff area (FATO) shall be provided around the helipad-landing center, which shall have a radius of 27.5 meters at least.
4. Safety shall be provided by ensuring that no obstacles stand in the path of the helicopter while landing or taking off. Moreover, safety nets shall be established around the helipad and approach segment if it is elevated.
5. Lighting equipment shall be provided heavily around the helipad in order to allow for night operations, and during fog and bad weather conditions.
6. Wind velocity information device shall be installed on each helipad, in order to assist the operation of the helicopter.
7. Approach guidance lights and land markings shall be provided as well in cases where the helipad is not clear enough from the sky.
8. Ground markings on the landing pad shall be large and clear enough for it to be distinguished from aerial elevation.
9. A pathway to the helipad shall be provided with a width of at least 1.2 m and a length exceeding 8 m.

Figure 10, in Appendix D shows the layout of a roof helipad [4].

3.4 Helistation Design

The helistation is a conceptual structure design, which aims at providing a helipad on major highways and locations of interest for emergency evacuations.

The helistation is a fully equipped structure that comprises of a series of 6 columns and 2 beams, spanning over the two sides of a highway. The columns and beams withstand a helipad on top, which is linked through ramps with two elevators, one on each side of the road. The elevators help carrying the ambulance up to the level of the helipad, and then the patient is transferred to the helicopter waiting on the helipad. The helistation was designed in a manner that doesn't interrupt moving traffic, and designed structurally in order to withhold the dynamic loads incurred by the landing and taking off of the aircraft.

Figure 4 below illustrates the dimensions and general layout of the helistation.

The clear span from the pavement ground level to the bottom of the drop beam is 10 meters, and the elevation of the ground of the helipad is 12 meters with respect to the PGL.

This structure was designed using ETABS, in order to withstand the live loads imposed by the helicopter's landing and takeoff movements.

Input data:

Table 2 below illustrates the basic data used to design the helistation [8].

Table 2. Input Data on ETABS for Helistations

Element	Characteristics
Placement	6-lane highway
Height of Station	12m
Slab Thickness	0.5m
Static Loads	1.6DL+SDL+1.2LL
➢ Dead Load (DL)	
➢ Superimposed Dead Load (SDL)	10 KN/m to be uniformly distributed
➢ Live Load (LL)	3 KN/m to be uniformly distributed on all floors.
Helipad Special Loading	
➢ Uniformly distributed.	1.92 KN/m2
➢ Load applied on 10x10 cm as to produce maximum effect.	13.35 KN
➢ Two concentrated loads	2.5 m apart of 0.75xMTW (Max Take-off Weight which is 2900lb)
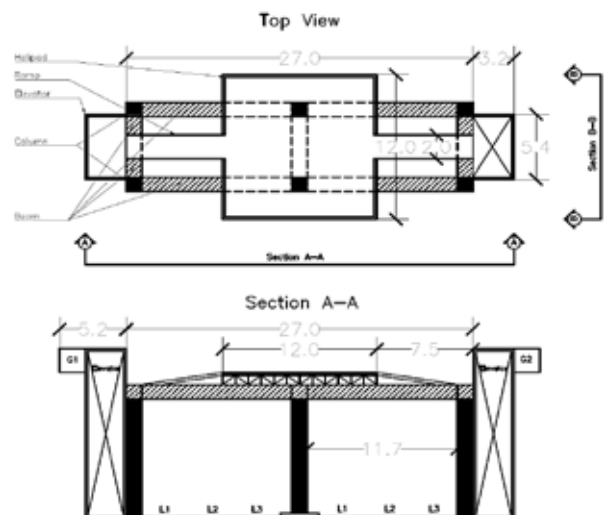
Figure 4-Dimensions and General Layout of a Helistation

Figure 12 in appendix F illustrates the 3D modeling of the heli-station, and the following list elaborates on the base data used in modeling the structure.

Design:

From the results we got, we noticed that the best outcome we will get is by having two types of columns (col1 at the extremities and col2 at the center beneath the helipad) and two types of beams (beam1 at the extremities and beam 2 at the center beneath the helipad).

Tables 5 and 6 in Appendix F illustrate the design recommendations for the beams and columns associated with the helistation.

Figure 13 in appendix F shows the deflected shape of the station. The maximum displacement occurs at mid-span on either way of the station, and has a value of 4 mm which is an acceptable value.

3.5 Aerial network

Based on the hot zones analysis, an aerial network is established as to cover as much areas as possible and insure proper functioning of this service, while minimizing the capital cost. This network comprises of two main components. The first one is the hospital helipads, which shall be installed on each hospital rooftop. The second component is the heli-stations installed on major arterials within hot traffic accident zones.
Trying to limit the imposed capital cost of the project, only four heli-stations are to be set up:

1. At Khalde, on the Beirut-south highway
2. At Jnah, on the Verdun-Ouzaai highway
3. At Mkalles, on the Siyyad-Mansourieh highway
4. At Jal El Dib, on the Beirut-North highway

The radius of influence of each helistation was studied and illustrated in Figure 5 below. Moreover, the relation between the heli-stations and the hospitals is studied and delineated in Figure 6 below.

Figure 5-Radius of Influence of each Helistation

Figure 6-Relationship between Helistations and Hospitals

IV.   STRUCTURAL DESIGN OF THE AUBMC HELIPAD AND HELISTATIONS

From a structural design perspective, two structures were analyzed, the first is the new AUBMC building, and the second is the designed heli-station.

4.1 Case Study: AUBMC

To check if the existing building can withstand the helipad and helicopters' accompanying forces, the AUBMC building was modeled using ETABS software as shown in Figure 11 in appendix E.

Input

Table 3 illustrates the data used and the underlying assumptions that were engaged in the structural analysis of the AUBMC building [8].

Table 3. Input Data on ETABS for AUBMC

Element	Characteristics
Columns	800x800 mm with 20 bars #6
Beams	400x200 mm
Static Loads:	Load combination is as follow: 1.6DL+SDL+1.2LL
➤ Dead Load (DL)	
➤ Superimposed Dead Load (SDL)	0.5KN/m to be uniformly distributed on all floors.
➤ Live Load (LL)	3 KN/m to be uniformly distributed on all floors.
Helipad Special Loading	
➤ Uniformly distributed.	1.92 KN/m2
➤ Load applied on 10x10 cm as to produce maximum effect.	13.35 KN
➤ Two concentrated loads	2.5 m apart of 0.75xMTW (Max Take-off Weight= 2900lb)

Output and Analysis

The deflected shape is shown in Figure 7 below. Note that the maximum deflection, which is highlighted in red below, has a maximum vertical displacement of 5mm which is an acceptable value.

Moreover, regarding columns and the moments applied on the slabs, it was noted that:
1. Two frames, column C16 and beam B17, are under high stresses and failure might occur. The absolute maximum moment in the column C16 is 52.55 KN.m, which is still in the acceptable range.
Also, for B17, we have:

- Absolute maximum shear= 26.23 KN (in the range)
- Absolute maximum moment= 31.856 KN.m (in the range)
- Maximum deflection= 1 mm (acceptable)

2. The maximum moments applied at story 4 (the most critical) for the slab are all within the acceptable range.

This structural analysis of the AUBMC building yielded the conclusion that it surely can withstand the additional weight of a helipad. Moreover, the dimensions of a helipad were integrated within the rooftop dimensions and it was found that there is enough space to integrate one. However, the AUBMC is located in the heart of Hamra, and the construction of a helipad might disrupt inhabitants of this area due to noise contours, and to safety measures associated with landing a helicopter in the midst of a building's cluster. This is problematic and could cause many inconveniences to the residents of the area.
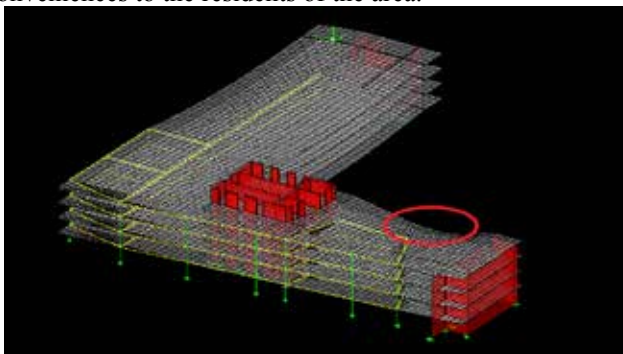

Figure 7 -The Deflected Shape of the Building

V.   FEASIBILITY STUDY

Referring to Kunhadi data, and some of the ISF data, an assessment of the annual cost of accidents in Lebanon is made. First of all, the total accident costs, including disabilities, and car repairs, are assessed at 1.8 Billion USD per year. However, isolating the mortalities costs only (as this project focuses on reducing these) they are estimated at 1.5 to 2.2 Million USD per year [9]; which is paid by the government to the families of the deceased as compensation (cost per life of a Lebanese citizen).

This number is compared with the cost of running the aerial network proposed. The cost of the system is split between the cost of building the helipads and helistation and the cost of operating the helicopters and base. Table 4 below illustrates the operation cost of each helicopter, and its estimated generated revenues. These numbers are based on the report of feasibility of helicopter emergency medical services in Humboldt County USA [10].

Table 4-Feasibility o operation per helicopter

Item	Unit cost	Cost$/year
Mid-size helicopter lease, fully equipped	55,000$/month	660000
Operating crew wages	8,000$/month	96000
Base operation costs	22,000$/month	264000
Yearly training programs for crew	10,000$/year	10000
Depreciation of equipment	15,000$/year	15000
Lighting and helipad maintenance	2000$/month	24000
Fuel cost	120$/hour	100000
Yearly cost of operation per helicopter		**1,170,000 $**
Fee per response	500$	
Estimated revenues, based on 2responses/day	360000	
Yearly net cost of operation		**810,000$**

Each helicopter will cost around 800,000$ to operate on a yearly basis. Considering the fact that the proposed network requires 4 helicopters, the gross cost is assessed at 2,400,000 USD yearly, which is relatively equivalent to the cost imposed by the traffic deaths yearly. Moreover, the capital cost of building helipads on each hospital is the responsibility of the hospital itself, and that of building the helistation is rather negligible on the long run.
Therefore, the proposed aerial network is feasible.

VI.   TRAFFIC LAW AMENDMENT

Traffic laws are poorly enforced in Lebanon, of which only five are currently applicable. These laws are the national speed law, the national drink-driving law, the national motorcycle helmet law, the national seat belt and phone use law, the national child restraint law. Also, for security reasons, according to the Lebanese regulation, only the Lebanese Air Force has the right to fly helicopters over the country. This project is highly related to the government, and it requires the amendment of this regulation and the execution of a new law, forcing all hospitals to acquire a helipad on its rooftop or surrounding environment, in accordance with global standards, that were previously listed in this paper. However, if that is not possible, this project is jeopardized and cannot be applied to Lebanon.

VII. SUMMARY

This paper analyzed the traffic deaths situation in Lebanon, and assessed the severity of the problem. Moreover, it tackled this problem by proposing a solution in the form of an aerial ambulance transport system, which covers the greater Beirut area. This system was found to be feasible on the long term. However, it faces a huge constraint, which is the governmental engagement in order to implement new laws related to this matter and making sure they are enforced.

ACKNOWLEDGMENTS

REFERENCES

[1] Misbah, M.(2010). Traffic accidents in Lebanon becoming an epidemic. Yasa web, Yasa. Retrieved from http://yasa.org/en/Sectiondet.aspx?id2=237&id=1, access date 2016.
[2] Lebanese statistical road data, 2010, <www.GPRS.com> , access date 2016.
[3] Private interview, Red cross staff.
[4] Air Ambulance Services in the US Market Research | IBISWorld. (n.d.).
[5] Choueiri M. Elias, Choueiri, M. George, Choueiri M. Bernard, Analysis of accident patterns in Lebanon, 4th international symposium on highway geometric design, June 2010.
[6]U.S Department of Transportation, Heliport design, AC No: 150/5390-2C, 201
[7] Alexander, R. (2008). Hospital Helipads [PDF document]. Retrieved from https://www.rotor.com/LinkClick.aspx?fileticket=xMhQLMWMc9Q%3D&tabid=179, access date 2016.
[8] Helicopter Hangars Design and Construction. (2014). Retrieved from http://www.reidsteel.aero/helicopter-hangars/, access date 2016.
[9] Traffic deaths data and statistics, 2016, www.Kunhadi.com, access date 2016.
[10] Statistics Archives - Air Ambulance Guides. (n.d.). Retrieved from http://www.airambulanceguides.com/category/air-ambulance-services/statistics/, access date 2016.
[11] Annual Review 2013/14 – London's Air Ambulance. (n.d.). Retrieved from https://londonsairambulance.co.uk/our-service/our-work/annual-review-2013-14, access date 2016.
[12] ArriveAlive - Helipads. (n.d.). Retrieved from http://www.childrenshospitals.org.uk/helipads.html, access date 2016.
[13] Canadian Journal of Surgery, (2014, February 4, "Helicopters save lives: Helicopter transport increases trauma survival over ground ambulance". *Science Daily*.
[14] Peltier, E. Hospital Heliport Inspection Basics [PDF document]. Retrieved from http://www.mhcea.org/Hospital%20Heliports.pdf

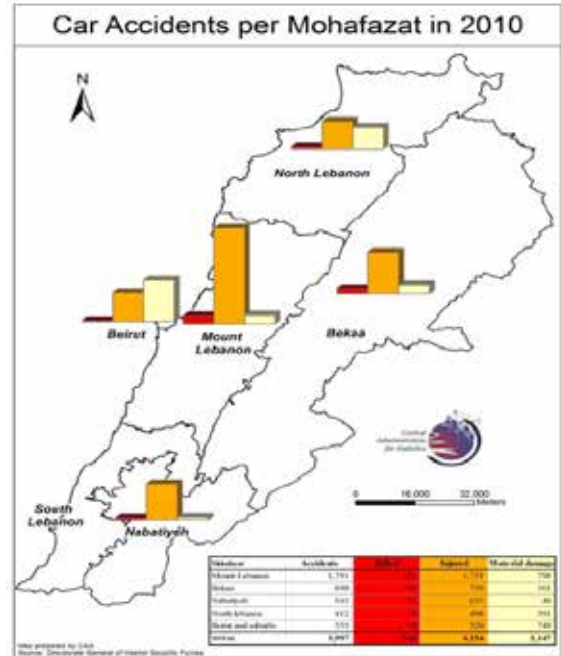## APPENDIX A CAR ACCIDENTS PER MOHAFAZAT

Figure 8 Distribution of Car Accidents in 2010

APPENDIX B: GIS MAPPING OF THE HOTTEST ACCIDENT ZONES WITHIN GBA

Figure 9- Mapping of the Hottest Accident Zones as Function of Mortalities and Injuries

Table 3-Shweizer helicopters catalog

APPENDIX D – ROOF HELIPAD GENERAL LAYOUT
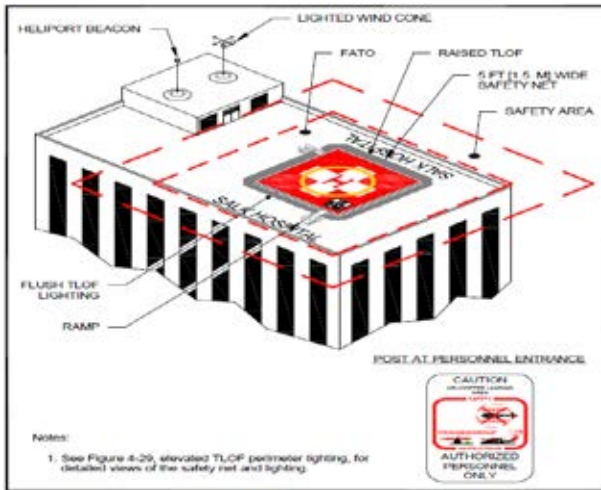
Figure 10-Helipad general layout

APPENDIX E – AUBMC ETABS MODEL
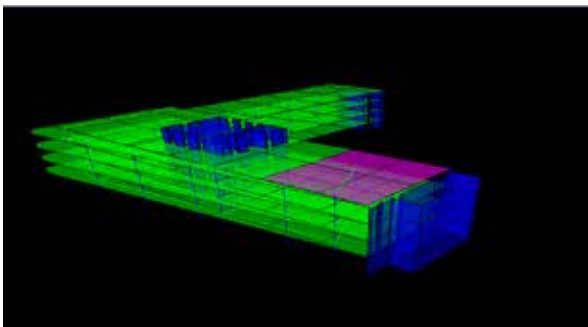
Figure 11 -the AUBMC building modeled on ETABS

APPENDIX F – HELISTATION STRUCTURAL ANALYSIS
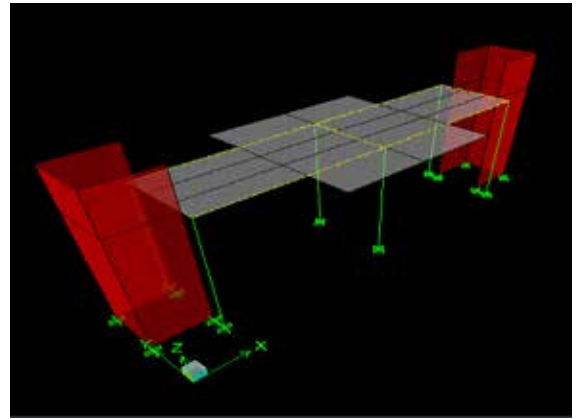
Figure 12-Helistation ETABS model

Table 5-Design requirements for columns

Design Input\Column	Type 1	Type 2
Cross-Section (m2)	1.2x1.2	1.2x1.2
Length (m)	12	12
Rebar (cm2)	140	140
Design Pult (KN)	463	2487
Design Mx (KN.m)	62	-55
Design My (KN.m)	379	-133

Table 6-Design requirements for beams

Design Input\Beam	Type 1	Type 2
Cross-section (m2)	1x1	1x1
Length (m)	13	6
Rebar Top(cm2)	30	10
Rebar Bottom(cm2)	20	5
Design +ve Moment (KN.m)	546	125
Design -ve Moment (KN.m)	-1092	-250
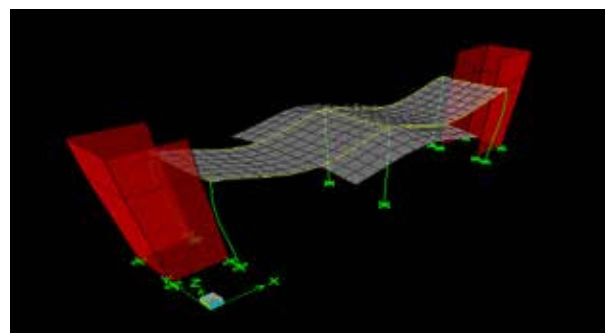
Figure 13- the deflected shape of the heli-station

Green Automated AUB Parking Facility

Mira Mohamadieh, Mohamad Awada, Hassan Fouani, Razan Naaman, Ali Jaber

Department of Civil and Environmental Engineering

American University of Beirut

Beirut, Lebanon

mym09@mail.aub.edu, maa222@mail.aub.edu, hsf08@mail.aub.edu, rwn03@mail.aub.edu, aaj20@mail.aub.edu

Abstract - **For years, AUB students have been suffering from the lack of a parking facility in the university which is solely dedicated to them. Thus, students are forced to search for a space to park their cars around the university's area which can be very time-consuming. The street parking spots and the small-sized parking lots nearby are not sufficient. Therefore, the main objective of this project is to design a Multi-Story Automated Parking Facility replacing the existing lot found in OSB, lower campus which will be dedicated to AUB students. It is of great importance to emphasize that the proposed parking facility is aiming to be environmentally friendly, where the design will intend to acquire the Green Garage Certificate. This certificate defines sustainable practices utilized in the structure, management, design and technology of the garage. Hence, the facility will be another AUB innovation that abides to international green standards. A design will be proposed and the feasibility of the project will be investigated.**

I. INTRODUCTION

A. General Overview

For AUB students that drive to campus, there has been a long-term struggle when it comes to finding a parking space for their car. In the early hours of the morning, it is almost impossible to find an empty parking spot since working individuals manage to occupy these spaces most of the time. Moreover, on Ein El Mraysse Road and Bliss Street lays a large number of coffee shops and restaurants that have reserved parking spaces for their customers. In addition, the narrow road on Bliss Street makes it unmanageable to park any car on street without causing traffic congestions.

Therefore, in a serious attempt to solve the parking crisis, this project presents a concept that will ensure the availability of parking spaces for AUB undergraduate students, while taking into account the status of AUB as a pioneer in incorporating sustainability and environmental technologies into its campus. The proposed project is constructing a Green Automated Multi-Story parking facility for AUB undergraduate students.

B. Project Description

The proposed facility is planned to be constructed on the existing OSB parking lot facing Irani Oxy Building, lower campus owned by AUB. **Fig. 1** below shows the exact location of this parking lot as shown on Google Earth. The final decision of the authors is to construct the Green Multi-Story Parking Facility which would be composed of ground floor parking spaces reserved for bicycles, motorcycles and handicapped spots, 4 upper floors and 3 underground floors that employ an innovative automated parking system.

Figure 1 - Google Earth image of the existing OSB lot.

It must be noted that the automated parking system utilizes computer-controlled machines that are similar to lifts, whereby a car that arrives on the lift's platform is transported to an available parking space, and then brought back to the passenger without any human intervention. A section of the facility entrance will

include aisles that have the base of the machine that will carry the parked car to an empty parking spot. To understand this concept, **Fig. 2** provided in the appendix can visually explain the concept. This system is designed to minimize the area used to park cars, eliminate most disturbances of conventional parking and withstand an immensely higher capacity of parked cars in comparison with a regular parking lot. Therefore, according to Automotionparking.com, the automated system is environmentally friendly and cost efficient, which are significant features that the authors want to attain in their proposed design. Further investigations will ensure the feasibility of such a structure.

II. SURVEY RESULTS

In this part of the study, the authors will present a survey to get the probable demand for parking spaces. The survey will also tackle the monthly parking fee to get an incentive on the pricing that will be adapted by the end of the project. The main objective of this survey is to get the peak number of vehicles arriving to the parking facility and the corresponding peak time. The document in the appendix shows the survey presented to AUB students. A brief summary of the obtained findings is found below:

A. Mode of Transportation

Students were asked about their mode of transportation upon arrival to and departure from AUB. The results acquired are presented in the **Fig. 3** below.
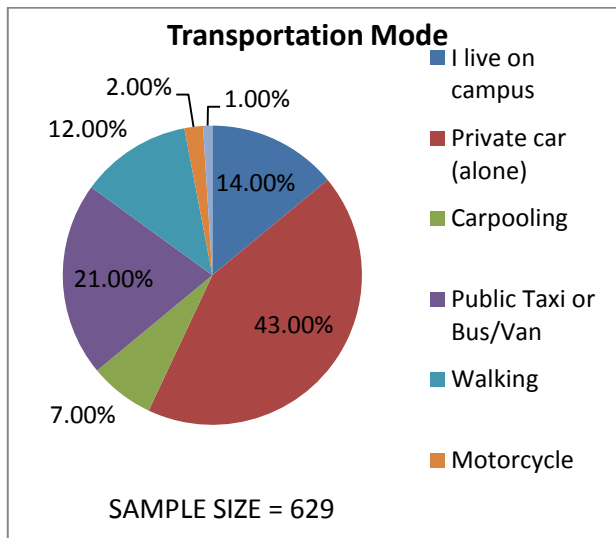
Figure 3 - Survey statistic of Current Student Transportation Mode.

The students who need a parking spot are those driving private car (alone). They represent 43% respectively of AUB students. Knowing that AUB has around 7500 students, according to AUB's registrar office, this means that 43% × 7,500 = 3225 students currently drive to AUB. These 3225 are most likely to use the new parking facility.

In addition, from **Fig. 3**, 28% currently use service or bus, or are carpool passengers in a car. (21%+7%) 28% of 7500 students is 2100 students, and of these 41% (obtained from question 6 in the survey) might be willing to change their commute mode to driving to AUB and thus must be considered as potential demand for our parking facility. 41% of 2100 students = 861 students who are most likely to use the new parking facility. Unfortunately, it is inevitable that a portion of students will convert from public to private transportation, but it is mainly due to the chaotic Lebanese public transportation means and that it is more comfortable for some students to drive to AUB.

B. Cost

Of the total (3225 + 861 = 4086) students under study, around 20.6% (obtained from survey question 7) will be willing to pay a monthly fee of the proposed 100$ according to willingness to pay information obtained from the same survey. This subscription rate was chosen on the basis that it is equal the average subscription charge in the Bliss area, and at the same time, high enough to control the demand on the parking facility generating enough income to eventually cover the costs of the construction of the facility in a reasonable amount of time.

C. Demand

The estimated parking demand is: 20.6% × 4086 students = 842 parking spaces.

In addition, this demand estimate can be further adjusted to more accurately model the actual demand. This is bearing in mind that not all the students considered for the parking facility will be on campus at the same time (around 79% are present at the same time), and that there are a number of students who have classes only 2 or 3 days of the week (90% have classes every day). Therefore, the adjusted daily demand is:

842 × 0.79 × 0.9 = 599 parking spaces. Therefore, our design demand will range between a minimum of 600 to a maximum of 700 spaces. Currently, the parking is designed for a capacity of 600 cars.

D. Peak Hour

From the survey, about 24% of students will arrive on 8:00 am approximately on both TR and MWF schedules, followed by a 21% on 9:00 am. This needs a queuing and operations system that can handle 24% of the 599 parking spots at a time. This 144 cars peek time will be distributed along 4 car lifts that move with 0.6 m/sec. This is achievable and does not cause congestion in the parking structure. However in terms of departure of students, peek time does not have a constraint on the operations of the structure since its widely distributed along all times of day with a maximum departure of 17% at 2:00 pm. The values obtained above are from survey questions concerning the schedule of students.

III. GEOTECHNICAL FINDINGS

In this section, the researchers will illustrate the geotechnical data corresponding to the site. Since conducting a full geotechnical investigation, with samples to be collected and tested, is costly and requires well-experienced professionals, the necessary data will be analyzed based on a past geotechnical investigation conducted for the construction of the Olayan School of Business (OSB) building which is directly adjacent to the proposed site. For this purpose, the authors consulted the Facilities Planning & Design Unit (FPDU) at AUB to collect the soil report prepared for the OSB building. The main soil parameters that will be taken into consideration to complete this study are: the allowable soil bearing capacity, the shear modulus, and the modulus of the sub grade reaction. These parameters will later be introduced to the SAFE program in order to be analyzed and eventually to determine the type of foundation that is adequate for this project.

A. Soil layers
The geotechnical layers of the ground levels are divided into two layers.

1. Topsoil layers
   - Medium firm to stiff sandy clay
     50% → 83% fines
   - Thickness varying from 1 → 3.5 m
2. Rock Mass Layer
   - Fractured hard strong Limestone with Chert or Marly Limestone
   - Calcium carbonate contents 88% → 94%
   - Thickness varying from 1 → 9.5 m

B. Water Table Levels

In order to find the water table levels, readings in the piezometers and in the open boreholes were taken at different dates, monitored at installation and daily after installation over a period of 11 days. The readings show water table levels varying between -0.20 meters below sea level < el. < +0.90 meters above sea level. It is safer to consider that the water table level is 0.90 meters above sea level. According to FPDU, the proposed land is 4.5 meters above sea level. Therefore, with 3 basements of 2.8 meters each, upon excavation, the water table will be encountered at a depth of 3.6 meters below the ground floor, upon the excavation of the second basement. Later, the uplift pressure must be calculated in order to find the weight of the building required to outweigh the uplift pressure, and thus, determine the duration of the dewatering process. So, the weight of water to be displaced is $(5.5 \times 59 \times 30)$ m^3 × 1.5 T/m^3 = 14,602.5 tons. The weight of the building is divided into: $(36×0.4×1×3=43$ m^3 for columns) + $(59×30×0.3=531$ m^3 for slabs) + $(90×0.3×0.29=78$ m^3 for walls) so total is 652 m^3 × 2.5 Tons/ m^3=1630 tons/floor, and the raft is 59×35×0.75×2.3Tons/m3 = 3622 tons. Therefore, after constructing the third floor (building weight = 15032 > uplift pressure = 14602), dewatering process can cease.

C. Concrete

The sulphate and chloride contents in the rock mass are expected to be low in the area of construction. Therefore, Ordinary Portland Cement Type I can be used in all concrete below ground level. The foundation concrete should be dense, with a minimum recommended cover of 40-mm.

The excavations for the 3 basements will reach a depth of about 8.2 meters below the ground, whereby the top soils are clayey and cannot be excavated without having support walls. However, the underlying Marly Limestone is hard and strong, and expected to have a good vertical stability upon excavation.

IV. FACILITY DESIGN

A. Zoning Regulations
With the help of the FPDU, the authors were able to get the total area of land at the existing parking lot next to OSB owned by AUB which is 2971 m^2. In addition, a-according to the zoning plans for the city of Beirut, the AUB lot is located in zone 5 and the general building

requirements applicable are according to the new decree 15874 dated 5/12/2005. The setbacks on public streets are 3 meters, the surface exploitation factor (SEF) is 40%, and the floor to area ratio (FAR) is 1.25. The following calculations were performed in order to develop an adequate facility design:

1. The building height should not exceed 25 meters according to the decree. Due to capacity and cost calculations that will be discussed later, the authors decided to have 4 floors above ground each of 3.2 meters height with the ground floor being 3.5 meters high, and basements of 2.6 meters, so the total height will be almost 20 meters including the roof parapet.

2. Building footprint = $2971 \times 0.4 = 1188.4$ m^2
   According to the decree, additions to the building footprint are: 20% for balcony areas, 20% extra area due to the fact that the land is at an intersection of two roads, 54 m^2 for external double walls, and 40 m^2 for stairs and elevators. Therefore,
   New building footprint = $(1188.4 \times 1.2) \times 1.2 + 54 + 40 = 1805$ m^2
   Upon these calculations and the land geometry found in **Fig.4**, the authors decided it is best to have the building geometry as such:
   Area = $59 \times 30 = 1770$ m^2

B. Floor Layouts & Operation

First, a detailed AutoCAD plan was established as shown in **Fig.4** in the appendix. This drawing shows the divisions of each floor. As previously mentioned, the parking facility is composed of 4 floors aboveground, 3 basements and the ground floor which does not utilize an automated system since it is reserved for handicapped, bikes and motorcycles. As seen in **Fig. 4**, each floor is composed of 4 zones. Each zone contains an elevator that will be used to lift/lower the vehicles.

1. The ground floor shows the entrance and exit of the cars in addition to the handicapped vehicles, bikes and motorcycles parking spots. The cars enter the facility, as the arrow indicates on the drawing, then goes into the available aisle leading to the designated podium. After parking the car, the driver moves out of the chamber and it closes up on the vehicle. The podium then rotates 90˚anti-clockwise to reach the shaft

location. The elevator then lifts the car to the designated floor where a parking spot is available. Note that the entering cars have a space to queue inside the facility with minimal effects on the main road. In addition, toilets, offices, stairs and elevators also exist in the space.

2. The design of the floors excluding the ground floor is similar, whereby each zone has a central aisle upon which the rail-machine moves back and forth on. When the car arrives to the floor, the rail-machine slides under the podium on the elevator then lifts the car onto the aisle. Then, the rail-machine places the car to the left or right of the aisle. Note that the basements have a larger area since they are not included in the setbacks and contain water tanks for water collection and building services, mechanical/electrical rooms and storage units.

3. The dimensions of the discussed elements are summarized in the following table.

Dimensions	Length (meters)	Width (meters)
Ground floor aisles	55	6
Pedestrian walkway	12	70
Podium	5.3	2.5
Handicapped space	5.5	3
Bikes/Motorcycle space	1.8	1.2
Central Aisles (railings)	5.3	2.5
Vehicle space	5.2	2

Priority for carpoolers will be given which will be controlled by security guards, whereby the last elevator will be dedicated to them for faster operation time.

It is also worth mentioning that after the end of the lifespan of the parking facility, AUB has the freedom to convert it to any type of building and still be covering the requirements of floor height and setbacks. This can be considered a privilege.

V.   STRUCTURAL ANALYSIS

The AutoCAD drawing was then exported to E-tabs to assess the validity of the proposed design. The

proposed parking facility is a reinforced concrete structure. The concrete compressive strength was assumed to be 35 MPa. The live load was calculated with the help of the ACI code, and the dead load was developed with the help of Mr. Ali Jaber, the head of the Smart Parking Co. in Lebanon. The values of the loads are represented below:

Loads on the Slab:

-Live Load = 2KN/m^2, which includes mainly the weight of cars (2.5 Tons), rails and elevators (12 Tons). This value was verified based on the guidelines of the ASCE 7-10 for Passenger Cars.

-Super Imposed Dead Load = 1KN/m^2. This value was assumed to be minimal because no important dead loads are usually assigned in a parking facility, but the assumption of 1KN/m^2 was taken for more safety.

Loads on the Roof:

-Live Load = 2KN/m^2, for any unexpected additional loads on the roof level of the facility.

-Super Imposed Dead Load = 3KN/m^2, which includes the weight of the water collection system, AC systems, etc.

Also, the loads of the rails and the elevator were calculated and found to be equal to 26KN per linear meter of slab along the rails.

In addition, the building is exposed to wind loads and seismic action, which are critical indicators to the level of serviceability of the structure. According to the ACI code, the maximum value of side sway was found to be less than H/500 meters (H: height of the building) and joint displacement was found to be less than 0.02 meters. Shear walls are included in the design in order to resist the horizontal movement of the machine and the seismic effect. After that, the ultimate load combination was also run through the program whereby the slabs were checked against punching shear and deflection; beam sections were checked against deflection, flexure, and shear; and column sections were checked mainly against buckling.

VI.   ENVIRONMENTAL BENEFITS

In this part of the study, the researchers investigate the environmental effects of implementing the automated parking technology. Green Parking Systems are considered environmentally friendly projects that conserve fuel and reduce pollution for a clean parking facility.

According to Apcpark.com, toxic emissions are considerably reduced upon the use of an automated parking system, whereby the following pollutants show the percentage by which they will be reduced in an automated parking system in comparison with a conventional parking garage:

- Carbon Dioxide (CO2) 83%

- Nitrogen Dioxide (NO2) 81%

- Carbon Monoxide (CO) 77%

- Volatile Organic Compounds (VOC) 68%

- Energy Consumption (Fuel Use) 83%

In addition, since there are really minimal lighting usages in automated parking systems, substantial energy consumptions are reduced.

VII.   CONVENTIONAL VS. AUTOMATED

There is a need to compare the automated parking facility in hand with a case study of a conventional facility on the same land in order to further highlight the advantages of constructing such a project in terms of cost, space exploitation, and capacity.

Based on the Lebanese decree, the design car size is 5.2 × 1.8 meters, distance between 2 cars and between a car and a side wall/column is 40 cm, and the distance from a back wall is 20 cm. The aisles should be 6 m wide for two-way access and other regulations exist for the turning radii which take up more space (up to 50 meters squared per turning radius). To move from one floor to another, the ramps are of length of 23 meters and width of 3 meters which results in a 300 m^2 loss of space per floor dedicated to ramps. Therefore, almost 50% of the total space is lost per floor due to these reasons in comparison with the automated parking facility that does not include these considerations. In addition, the capacity of a conventional parking is 36 cars / basement, 31/ground floor, 38/upper floor which results in a total of 276 cars. However, in the automated facility, the capacity is 98 cars / basement, 6/ground floor (for handicapped), 94/ upper floor which results in a total of 676 cars. Therefore, the automated parking facility can accommodate more than double the capacity of that of the conventional It is worth mentioning that the current parking lot has a capacity of 130 cars parked in a disordered manner.

Mandatory costs for both facilities are concrete and steel costs. According to Engineer Nada Kaddoura, the cost of concrete is (90$/m^3 material + 50$/m^3 labor = 140$/m^3) and of steel is (600$/ton material + 20$/ton labor = 620$/ton). On average, the average weight of steel needed in reinforced concrete column or beam is 150 kg/m^3. Electricity costs are 20% of the total cost and an additional cost for mechanical works is also 20% of the total cost. For the proposed design, the preliminary cost is 2,332,680 for concrete and 430,000$ for steel, so total is 2,762,680$. The conventional parking has an added cost for ramps which is almost 15,000$ versus the

500,000$ extra cost for the car lifts and other "green additions" in the automated parking. Although this seems like a huge difference in cost, in terms of the system cost and operation, but the profit will definitely be significantly larger due to the higher capacity not to mention the accommodation of the needed demand. (A detailed cost analysis was assessed but due to space limitations, it couldn't be added, so it will be included will in the full report)

VII. SUSTAINABILITY CRITERIA

In this section, the authors will address the sustainability aspect of the project. The project will try to achieve a *Green Parking Certification* bronze standing. A minimum of 110 points is required for this purpose. To ensure that the facility design will hit the margin of 110 points, the researchers are aiming to complete a design that targets 120 points to keep a certain margin of error. The authors will choose the different points that are applicable in their design and then implement them within the scope of the project.

The Green Garage Certification Standard provides new parking facilities with a path to diminish environmental impact and boost structural longevity, operational efficiency, revenue diversity and community relationships. The Standard is organized in the following three major categories, with an added area for innovations:

1. **Management:** Highlights ways in which garage operations can maximize the use of a parking asset while minimizing waste. Embracing these practices ensures facility staff utilizes resources to their full potential.

2. **Programming:** Guides garages to new revenue sources, greater customer satisfaction and stronger community relations. Green garage programs ensure effective vehicle ingress/egress, provide access to alternative mobility solutions, and leverage the garage's potential as a public space.

3. **Technology and Structure Design:** Outlines the physical attributes a garage can deploy to increase energy efficiency, lower waste and support customer mobility choice. (Council, 2014).

4. This project will integrate multiple elements from the certificate to ensure its sustainability and environmental standards. Mainly, the project includes: Indoor water efficiency, green roofing systems, renewable energy generation, ventilation, lighting control, storm water management and the use of low volatile organic compounds. The other themes found in this guideline are currently being studied to decide which aspects will be integrated within this project.

VIII. CONCLUSION

In conclusion, such an investment promises great profit for the investor in addition to the fact that it solves the AUB parking crisis. It is important to mention that upon receiving the Green Garage Certificate, this project would fit AUB's high environmental standards. It will also be a new innovation for AUB knowing that it is the pioneer in that field. As civil engineers, the ultimate goal is to design a structure that would fit the surrounding setting, solve the ongoing parking struggle and do that while adhering to high safety and sustainability standards. Upon completion, this project would be considered as an original multi-story structure that will serve AUB students for almost 80 years.

IX. ACKNOWLEDGMENT

REFERENCES

[1] Green Parking Council, 2013, greenparkingcouncil.org

[2] Abou Zeid M.,Department of Civil and Environmental Engineering, American University of Beirut.

[3] American Association of State Highway and Trasportation Officials, 2011, transportation.org

[4] Abou Ghali N., Assaad F., Ghizzawi F., Houssami L., Salameh S., HSBC Multi-Modal Parking Garage, 2014

[5] Automated Parking Corp. | Green benefits, apcpark.com/green

APPENDIX
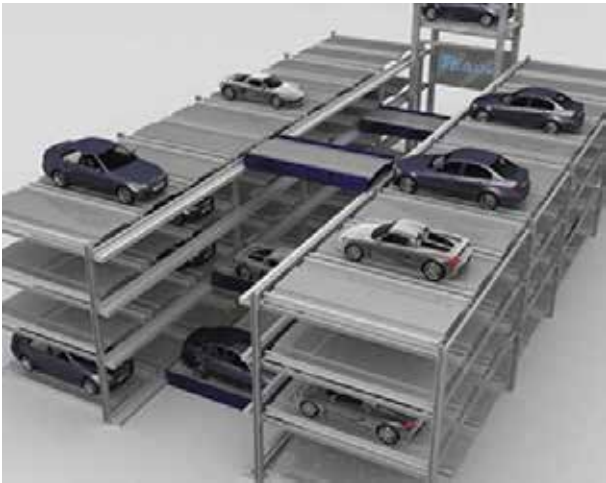
Figure 2 – The Proposed Automated Parking Demonstration
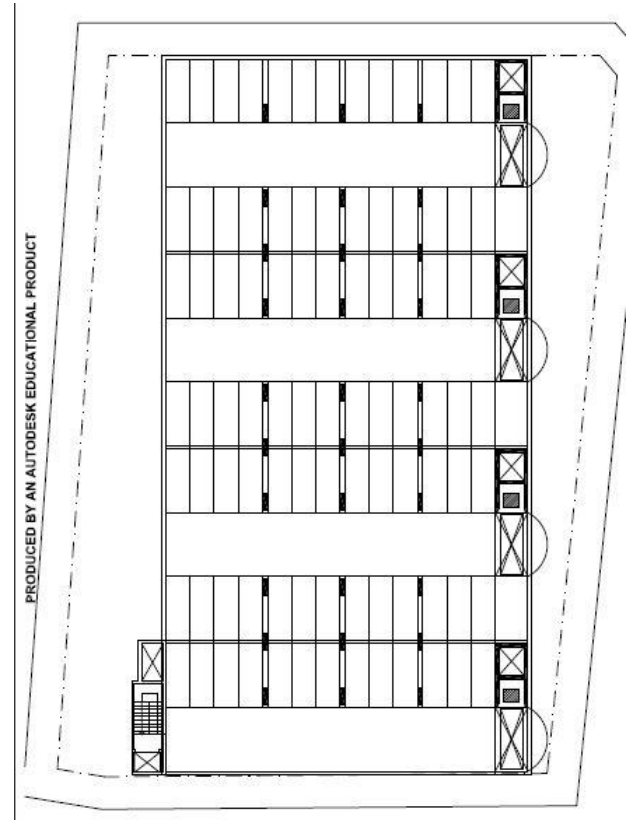
Upper floor design

Figure 4 – Detailed AutoCAD Drawing

Section of the facility

Figure 4 – Detailed AutoCAD Drawing

Ground floor

Figure 4 – Detailed AutoCAD Drawing

275

Survey

Please specify your faculty field.

1. FEA

2. OSB

3. FAS

4. FHS

5. FAFS

How do you usually get to AUB?

1. N/A: I live on campus

2. Driving private car (alone)

3. Carpooling

4. Bus/Van

5. Public Taxi

8. Walking

9. Motorcycle

10. Bicycle

If you do not drive to AUB by car, is the reason that you cannot find a parking spot?

- Yes

- No

Please specify the following:

Usual time of arrival and departure to AUB (This semester) MWF:

Arrival: _____
Departure:_____

Usual time of arrival and departure to AUB (This semester) TR:

Arrival: _____
Departure:_____

Where do you usually park your car (if used)?

1. Free on street parking

2. Paid on street parking

3. Off-street parking (in a parking lot or garage)

4. Other

   Please Specify: _____

If provided with a fully automated AUB parking facility seaside next to Olayan School of Business (OSB), will you drive to AUB and use this facility?

Fully Automated Parking is an innovative mechanical parking system that will park the car for you.

- Yes

- No

Do you think 100 U.S. $/Month is an acceptable parking fee to an AUB fully automated student parking facility?

- Yes

- No, please offer a more adequate pricing: _____

Harvesting Water from the Atmosphere

Serge Bechara, Ziad Beschir, Lara Jalwan, Nour Mdeihli, Nour Menassa

Civil and Environmental Enginering

American University of Beirut

Hamra, Lebanon

scb05@mail.aub.edu, zrb01@mail.aub.edu, lsj05@mail.aub.edu, nsm26@mail.aub.edu, nym02@mail.aub.edu

Abstract- Due to the great influx of Syrian refugees over the past few years, the energy and water supply crisis in Lebanon is becoming more pronounced. The project aims at solving this issue by harvesting atmospheric moisture found in the ambient air using the principles of renewable energy and condensation. This innovative idea, still at a research level, presents dew as a non-conventional water resource. The system uses electric energy from solar panels to drive a condenser that condenses vapor into dew. In order to design a highly efficient system, rating curves relating dew yield and energy contribution have been developed for different conditions of relative humidity and ambient air temperatures. The results have been used to conduct a design for both the cooling and the solar energy integrated system. A small prototype of the cooling system was also built using thermoelectric cooling units and tested in limited conditions. The study resulted in the design of a solar system that fuels the cooling units to be designed according to obtained design curves.

I. Introduction

Globally, according to the World Health Organization, more than one billion people around the world today have no access to clean water. Have you ever thought that the early morning droplets on the windshields of a car or blades of grass can actually be saving lives? Providing people with their right to have clean water is a growing challenge in the modern world. The available supply is further threatened by growing human activities, significant flux of refugees, groundwater pollution due to the accumulation of solid wastes, and sea water intrusion caused by illegal welling. In fact, with last year being recorded as the hottest year in history due to global warming, water availability is being highly threatened. What most people fail to acknowledge is that 95% of greenhouse gases are water vapor and not carbon dioxide [4]. This technique presents an innovative and sustainable water resource that can supports the limited supplies. Harvesting water from the atmosphere generates latent heat but at the same time reduces sensible heat potential thus creating a net positive cycle in terms of global warming.

Not all people have access to wells to pump ground water and neither do most people live near rivers and lakes. Atmospheric water is a resource available to all people equally, regardless of where they live. This is why this technique deposits itself as a futuristic approach to solve the issue of water shortage in many areas around the world. Common practices, like water and wastewater treatment, desalination, water decontamination, return impurities to the ecosystems; nevertheless, harvesting water from the atmosphere does not pollute. The extraction of water from air provides an almost unlimited source of clean drinking water at a low energy consumption without damaging the surrounding environment. In this paper, a system that could be a solution for the water shortage is presented. The system relies on harvesting water from the atmosphere: potable water is collected from water vapor found in the air by a process known as condensation. It generally occurs when air cools below its dew point temperature and loses its capacity to hold water vapor, so it condenses to form water droplets. The system worked on consists of an energy component, photovoltaic cells that replenish themselves solely from solar energy, and a cooling component that will be fueled by these solar panels. The paper at hand details the design of the energy component and presents rating curves to use in order to design the cooling component of the system.

II. Literature Review

The following section focuses on previous efforts to harvest dew as a nonconventional source of water. Since 2011, companies were trying to produce water from moisture in the air.

In 2011, Eole, a French company, came up with an idea to generate water in a nonconventional way in the desert of Abu Dhabi. The system designed claimed to generate more than 1400 liters of water each day. Eole needed a source of renewable energy and thus resorted to wind. The wind turbine installed is 24 meters high, and its 13-meter rotor turns at up to 100 rpm to run a 30-kilowatt generator [6]. The wind turbine's function was to provide electric energy to a cooling compressor that cools the air and condenses water out into a tank. However, there were some limitations regarding the electric energy needed and the resulting output of water. The wind turbine necessitated a minimum of 15 mph to generate the energy required to drive the compressor. The French company found out that the system was able to produce 350 Liters of water per day given a temperature of 95 degrees

Fahrenheit and a relative humidity of 30%. Further studies showed that the installation of a solar panel could power the turbine and could amend the output drastically. What rendered this effort obsolete was the immense cost of single turbines, which was around 650,000 USD at that time [6]. The massive cost of just one component of the system rendered the latter unaffordable for most developing countries.

In April 2013, The University of Engineering and Technology in Peru worked on a billboard, which produces hygienic water from atmospheric humidity. Electricity, needed to power the five condensers used to convert vapor to dew, was supplied from the power lines of Peru [6]. The concept behind the system lies in condensing vapor in the atmosphere into liquid state as the air comes in direct interaction with the condensing surfaces. The concept of reverse osmosis was implemented in order to purify the produced water. The system was found to generate a daily upshot of about 95 liters [6].

In June 2014, Ugo Bardi and Toufic Al Asmar observed how solar refrigeration could harvest a great amount of water by condensation from the air. As a result, they sought to develop this idea into a system that uses electricity from solar panels in order to actuate a condenser that collects humidity from the ambient air. They foreshadowed a net yield of two hundred liters of pure water per day. A photovoltaic panel, the source of electric energy that drives the condenser, makes the system autonomous, renders it versatile in remote locations, and eliminates the need for it to be linked to the grid. Also, the system proposed by Bardi and Al Asmar removes the need for pipeline systems employed in the usual desalination plants, what makes their system relatively cost-efficient in the long run.

Another idea came about in November 2014 when Kristof Retezar, an Austrian industrial design student created Fontus, which produces freshwater in areas where relative humidity is high. Fontus is made up of a condenser, Peltier Element, powered by a photovoltaic panel on top of it. Water generated is then collected in a bottle fixed underneath a bicycle [10]. The system is attached to the bicycle as to let air flow at a certain speed into a channel as the bike moves forward. By doing so, moisture in the air is cooled and therefore condenses to water. The system was tested and found to yield half a Liter of water per hour in areas where temperature and relative humidity values are high enough [10].

The Peltier Element used is apportioned in two. Once it receives electrical energy, the upper part of the Peliter cools down and the temperature in the bottom part increases. The higher the temperature increase in the part underneath, the lower the temperature in the upper part of the cooling device [10]. As the biker initiates his/her activity, air flows into the bottom compartment at a relatively high celerity, which cools it down. The system also accounts for air entering into the top compartment. The Fontus stops the air coming to the upper partition through "little walls perforated non-linearly" which halt the air and give it enough time to "lose its water molecules". It is important to note that the two compartments of the Peliter Element are detached and secluded from each other. The advantage of this cooling device is that it provides a large surface for condensation to occur; however, it uses the smallest space possible.

III. THE COOLING COMPONENT OF THE SYSTEM

A. Literature Review

The examination of the thermodynamics behind condensation and cooling is imperative in obtaining correlations that allow one to design a cooling system that harvests atmospheric water. The design of such systems that operate at a given temperature and relative humidity is limited either to a certain supply of energy or by an amount of water to be produced. A research conducted by Durran and Frierson on a cold can containing 350 grams of water placed in a hot and humid environment to monitor the release of latent heat when water condenses was examined. Durran and Frierson asserted that after 5 minutes, the condensed water formed a 0.1 millimetre thick layer around the 290 cm^2 can. The mass of water was calculated to be 2.9 grams and thus the temperature of the water inside the can, initially at 0°C, rose by about 5°C, due to latent heat. Durran and Frierson also came up with a graph that plots the change in temperature due to total and latent heat as a function of relative humidity at different ambient air temperatures (25°C and 35°C) [2].

B. Scope and Methodology

The rise in temperature of the water, inside the can, can be attributed to two factors: (a) heat transfer due to convection with the ambient air and (b) latent heat of condensation. Fig.1 presented by Durran and Frierson depicts the aforementioned argument since the upper increasing curve represents the change in total heat (i.e. the effect of both factors), the lower increasing curve represents the change in the latent heat (i.e. heat transfer due to condensation), and the heat transfer due to convection is accounted for by the constant distance between the two curves at any relative humidity [2].

The condensation of dew releases latent heat, so the latter was used to obtain the yield – m_{yield} – in g/m^2.hr at 25°C and 35°C through the following two equations:

$$Q_{latent} = m \times C \times dt_{latent}, \qquad (1)$$
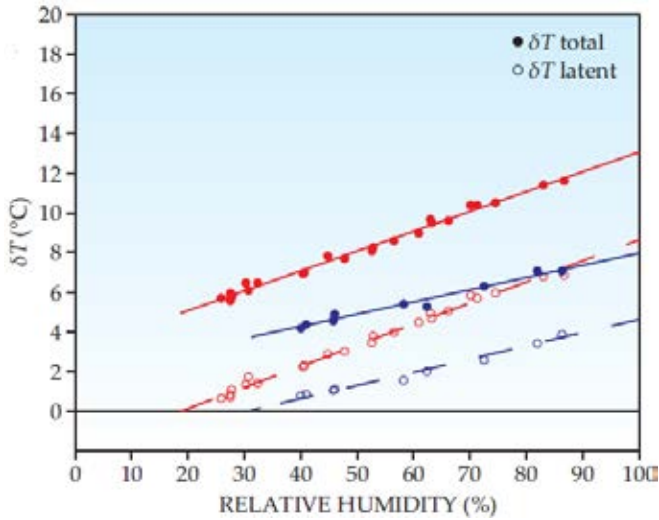
$$Q_{latent} = \kappa \times m_{yield}, \qquad (2)$$

Figure 1 - Relative Humidity vs Difference in Temperature [2]

where Q_{latent} is the energy released due to condensation (cal); m is the mass of the water inside the can (g); C is the specific heat of water (1 cal/g.^{o}C); dt_{latent} is the change in temperature due to latent heat (^{o}C) (obtained from the graph);κ is the latent heat of condensation (600 cal/g for water around 0 ^{o}C).

Then, the total energy needed to obtain the amount of water calculated was obtained through:

$$Q_{total} = m \times C \times dt_{total}, \qquad (3)$$

where Q_{tota} is the total heat transfer (cal); m is the mass of the water inside the can (g) ; C is the specific heat of water (1 cal/g. ^{o}C); dt_{total} is the change in temperature due to total heat ^{o}C (obtained from the graph).

The operations mentioned above were performed at 25°C and 35 ^{o}C ; therefore, linear interpolation was conducted to obtain the values of energy required and the yield per m^2 of cooling surface per hour for 0.5^{o}C intervals of temperature between 25°C and 35°C.

C.   *Design Curves*

The design curves in Appendix A were obtained.

Fig. 3 of Appendix A serves as a scientific model that shows the energy requirement (in kWh) to obtain a certain amount of water (in mL/ m^2.hr) at any combination of temperature and relative humidity. As an illustration, one can consider an area where the temperature is about 30°C and the relative humidity is 70%. In order to design a cooling system that condenses vapor into water, one needs 1.25 kWh in order to obtain 850 mL/m^2.hr of water.

Fig. 4 of Appendix A and Fig. 5 of Appendix A were derived from Fig. 3 of Appendix A to demonstrate the

obtained values from an engineering approach. To design a system, one will either be constrained with the amount of energy input or by the need of water and thus the yield. Fig. 4 of Appendix A shows the values of yield per unit of energy (kWh) at any combination of temperature and relative humidity. At 29°C and 67% relative humidity, 680 mL of water will be produced per kWh energy input. A fixed amount of energy available, say 2.5 kWh, will produce 1700 mL / m^2.hr. Fig. 5 shows the values of energy needed to obtain 1 mL of water. So, at 29°C and 67% relative humidity, 0.0015 kWh are required to obtain 1 mL. As a result, if you are constrained by the amount of water to produce, say 1700 mL, you will need about 2.5 kWh. The two figures represent the same data but from two different approaches: (a) designing when faced with limited energy input and (b) designing when a certain amount of water is required to be produced. They can be used to design the cooling component of the system proposed; however, one has to take into account the efficiency of the type of cooling scheme one will use since the values of the graph account for 100% efficiency. For instance, thermoelectric cooling units have a very low efficiency (10%) and are not ideal to use since they will require much more energy than cooling stechniques with 70% efficiency.

D.   *Experimentation: Building a Prototype*

A prototype was built not only to concretize what has been explained in the above section, but also to collect data to describe how the system will perform when put in practice. For that, a 30cmx30cm aluminum plate tilted 30 degrees from the horizontal, two thermoelectric cooling, *TEC,* units known as Peltier elements, a heat sink and a fan were used. Experimental measurements show that optimal yield is obtained at a 90 degree tilt due to gravitational load, and optimal wind influence and light-emission are obtained at 0 degree. Thus, the optimal tilt, which is 30 degrees is a good compromise between easy drop collection and weak wind influence [1].

The process goes as follows; the plate having its own initial temperature needs to be cooled down below the dew temperature for yield to be harvested. Since the natural cool-down by convection is a slow process, external help by cooling techniques is required. Thermoelectric cooling units or Peltier elements, though very minimal in energy efficiency (10%efficiency), were adopted because of their easy handling and maneuvering. Two Peltier units, each found experimentally to have a zone of influence of 9cm in diameter (as illustrated in Fig. 2 on the following page), were centered
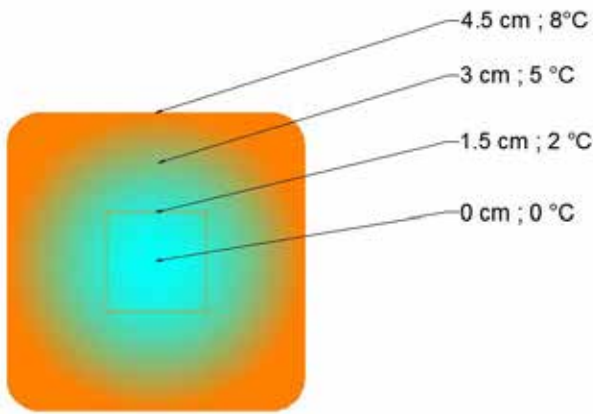
below the plate and separated by a 5cm gap.

Figure 2: Peltier Unit Influence Zone: Distance from Center and Temperature at Different Locations

Thermal paste was used to connect the peltier units to the plate which allows for a greater contact area and for a better heat conductivity between the aluminum plate and the peltier. Linked to the peltier is a heat sink that absorbs the heat generated by the hot side of the peltier, preventing a rise in temperature in the plate. To prevent any heat exchange between the plate and the heat sink, a thermal insulating sheet was placed between the two. The design of the sink is a function of the peltier's resistance and power; the peltier unit has a resistance of 1.7 ohms and a power of 62.2 watts. To get the thermal resistance (RQ) of the heat sink, the following equation was used:

$$RQ = (\Delta t)/(Q_{max}peltier + Rpeltier) = 0.45°/W \qquad (4)$$

where RQ is the thermal resistance of the heat sink; Δt is the temperature difference between the temperature of the aluminium plate and that of the ambient air; Q_{max}peltier is the maximum power of the peltier unit; and Rpeltier is the resistance of the peltier unit.

Typical RQ of a heat sink ranges from 0.05°/W to 0.9°/W. The value obtained is deemed acceptable since it lies within the range. Putting all components together, the system was tested under different temperature and different relative humidity and the results were promising.

IV. SOLAR ENERGY INTEGRATED SYSTEM

A. Literature Review

One of the main objectives of this project is to implement an innovative and sustainable technology.

The light and heat provided by the sun, the movement of water, and the wind are the main energy sources and main concern of sustainable studies. Each source offers a variety of advantages and a different way of converting energy into electricity. Both, wind power and hydroelectricity, are growing technologies that can fulfill considerable portions of electricity demands; however, this project focuses solely on solar panels.

The different components of a typical solar system are the photovoltaic cells, a charge controller, a battery bank and an inverter [9].

The Solar Panel

The photovoltaic cells are the basic components of a solar system. They convert the absorbed light into electricity. These cells are mostly made of Silicon, and they can be found in different forms [3]. When light shines on a solar cell, it passes through the cell and is then reflected. Part of the light might be absorbed. In fact, only the absorbed light contributes to the generation of energy.

The Battery Bank

The battery bank stores the electricity from the solar panel for later use. This part is one of the most complicated and costly components of the solar system [4].

The Charge Controller

The charge controller optimizes the lifetime and the production of the panels by keeping the batteries properly fed and safe for the long run. Charge controllers block reverse current and prevent battery overcharge.

The Inverter

The inverter converts direct current into alternate current which is typically needed by housing appliances.

B. Scope and Methodology

The first step in the study of the renewable energy integrated system is to understand the operation of a solar energy retrieving system and examine the most convenient module to the study. The group aimed at understanding the output of solar panels with respect to energy. In fact, the output differs depending on the module chosen and the array of cells. It is also important to account for the way the panel is exposed to sun i.e. the inclination angle [3]. The group chose to go with thin film solar cells, since they are very efficient in terms of space and offer a fair budget for study purposes [4].

Then, the group studied how the battery and the inverter work in order to incorporate both in the design. This task required basic knowledge in the electric system and the guidance of an expert in this field. Once the different components were explored, the group proceeded to the design phase of the energy system. The choice of a suitable solar panel was based on the expected energy required for the system at specified conditions of ambient air temperature and relative humidity. The battery was based on the amount of energy that the solar panel can generate. The solar panel generates direct current, so the use of the inverter will depend on the current needed by the cooling system.

The sizing of the solar panel, batteries, charge controller, and inverter can be calculated through the following steps:

1. The total number N of solar panels needed is given by:

$$N = TE/EC \qquad (5)$$

where TE is the total energy needed per day (kWh per day) and EC is the energy capacity of a solar panel throughout a day (kWh per day per unit).

2. The total area TA of panels needed is:

$$TA = A \; x \; N \qquad (6)$$

Where TA is the total area of the panels (m^2) and A is the area of one solar panel (m^2).

3. The required battery capacity is given by the following equation:

$$BC = (TE \; x \; 1000 \; x \; D)/(0.85 \; x \; V) \qquad (7)$$

Where BC is the battery capacity (Amp.h), D is the days of autonomy of the battery and V is the voltage of the battery (V).

C.   Design

The Photovoltaic Solar Panels

The solar panel should be implemented such that the angle towards the sun results in optimal output for the system. The optimal angle for the photovoltaic cells depends on the seasonal variations and location of implementations. It is also important to account for the way the panel is exposed to sun. Typically, according to existing databases, the angle is taken to be 30 ° in Lebanon. The solar modules of the system are best implemented in places with full exposure to sunlight i.e. open fields or rooftops.

During summer, the average sunlight time is around 15 hours; whereas, during winter there are 4-5 hours of sunlight. In Lebanon, there are 1400 effective hours of sunlight throughout the year. Calculations were thus conducted based on the average resulting hours of effective sun per year, which results in around 4 hours per day (1400 effective hours per year divided by 365 days per year).

The design is based on the following conditions:

- The design is conducted on the conditions of a typical summer day in Lebanon, during which the ambient air temperature is of T=30 °C and relative humidity is RH=70 %.

- Assuming a 100% efficient cooling system, according to figure 3 in Appendix A, the energy needed for these conditions is 1.38 KWh per hour. Since the solar panel is effective for 4 hours a day, the energy given by the solar panel is: 1.38 kWh x 4 hours= 5.52 kWh per day.
- A standard solar panel of power of 250 W with a size of 1 x 1.6 m is used. In Lebanon, the panel outputs, on average, an energy of 375 kWh per year, which results in almost 1 kWh/day.

Based on the conditions stated above, the number of solar panels needed per day was calculated to be around 6 panels:

$$N = 5.52/1.03 \cong 6 \; panels$$

This results in a total area of:

$$TA = 1x1.6 \; x \; 6 = 9.6 \; m^2$$

The Battery Bank

In order to design the battery storage, we will consider the following:

- Number of batteries
- Voltage of battery (V)
- Capacity of battery (Amp.h)

In the designed system, 12 V batteries are used and are assumed to have 12 hours of autonomy. The result is divided by 0.85 to account for battery losses.

$$BC = \frac{5.52 \; x \; 1000 \; x \; 0.5}{0.85 \; x \; 12} = 270 \; Amp.h$$

An average battery of 12 V 300 Amp.h, for half a day autonomy, is enough in the system; however, in this system, a battery is not necessarily needed. In fact, there is no essential need for back up in order to maintain power at times when the solar panels are not exposed to enough sunlight to produce energy.

V.   DESIGN AND PROJECT CONSTRAINT

A.   Economical Constraint

The implementation of the prototype is highly limited on the available budget. The choice of the various components is primarily based on their efficiency.

B.   Time Constraint

Another major constraint faced in the execution of the project is time. The group is limited by a short amount of time, for the members have several tasks to complete before the end of the final year. The testing of a prototype was done in March; however, due to time limitation the group was not

able to conduct tests for different conditions of temperature and relative humidity.

C. *Environmental Constraint*

The project itself is associated with an environmental constraint. The outcome of the system is very dependent on humidity and temperature. The effectiveness of the system can be best demonstrated during summer; nevertheless, due to the lack of available controlled environmental chambers in American University of Beirut, the design proposed could not be tested for a targeted temperature or relative humidity.

VI. CONCLUSION AND RECOMMENDATIONS

The project at hand is a radical innovation in the world of sustainability and renewable energy. The project relies on dew as a nonconventional water resource and on solar energy to power a condenser. The project is also based on a study done by futurist Thomas Frey, who forecasted based on trend monitoring and analysis that water harvesting from the atmosphere will be a job by 2030 [8]. The topic at hand is a breakthrough in the field of water resources management and energy resources management.

The authors of the paper have sought to develop design that allow to obtain the energy requirements, which yield a certain amount of water at any combination of temperature and relative humidity values. Based on the average weather conditions in Lebanon during summer, the energy value was extracted from the rating curves to design a compatible solar energy system. The choice of materials and sub-components was taken into consideration. The group researched the different types of solar panels and their respective configurations. For the energy system, the thin solar film cells were chosen because they require less space and are more affordable than other types of photovoltaic cells. As for the cooling surface, commercial aluminum was chosen because of its high conductivity. However, due to the limited amount of funding, cost had to be taken into consideration. To build and test a prototype, the Peltier cooler, although a low efficiency cooling system, was adopted because of its low maintenance cost and its high cooling capacity.

The use of dew as the water resource came from the idea of several studies showing that the other water resources being used at present times are being depleted. Also, dew occurs naturally and is generally potable, mainly in the countryside [5]. The use of solar energy was eminent due to the fact that solar energy is a renewable, sustainable, non-polluting source of energy. Also, studies show that it does not contribute negatively to the environment since it does not emit greenhouse gases. The combination of the two concepts is a major development in the joint advantage of Water and Energy Resource Management fields. Since these fields are both part of the Civil Engineering discipline, the profession should start considering "Harvesting Water from the Atmosphere" as a probable alternative to desalination and other techniques.

The group has not decided on a cooling system to be used. This part can be further developed as a joint venture with a mechanical engineering group to design an efficient cooling system based on the graphs obtained. The two groups can eventually work on optimizing the system as a whole.

VIII. REFERENCES

[1] Beysens, Daniel, Irina Milimouk, Vadim Nikolayev, Marc Muselli, and Jacques Marcillat. "Using radiative cooling to condense atmospheric vapor: a study to improve water yield." Journal of Hydrology 276.1-4 (2016): 1-11. Print.

[2] Durran, Dale R., and Dargan M. Frierson."Condensation, atmospheric motion, and cold beer." *Phys. Today* 66.4 (2013): 74. Web.

[3] Ewing.*Power R. with nature: Alternative energy solutions for homeowners*. Masonville, CO: PixyJack Press, 2006.

[4] Hieb, Monte. "Global Warming: A Closer Look at the Numbers." Global Warming: A Closer Look at the Numbers. 2 Mar. 2007. Web

[5] Maehlum.M. "Which Solar Panel Type is Best? Mono-, Polycrystalline or Thin Film?" May 2015. Web.

[6] Smith-Strickland, Kiona. "A Billboard That Condenses Water From Humidity." Popular Mechanics. N.p., 25 Apr. 2013. Web.

[7] Tomaszkiewicz, Marlene, Majdi Abou Najm, Daniel Beysens, Ibrahim Alameddine, and Mutasem El-Fadel. "Dew as a sustainable non-conventional water resource: a critical review." *Environ. Rev* 23.4 (2015): 425-442. Print.

[8] Thomas Frey. Futurist Speaker. Tapping into the Waterways in the Sky. 2013. Web.

[9] Washington State University *"Solar Electric System Design, Operation and Installation."* 2009. Web.

[10] Winter, Lisa. "Bicycle Bottle System Condenses Humidity From Air Into Drinkable Water." IFLScience. N.p., 18 Nov. 2014. Web.
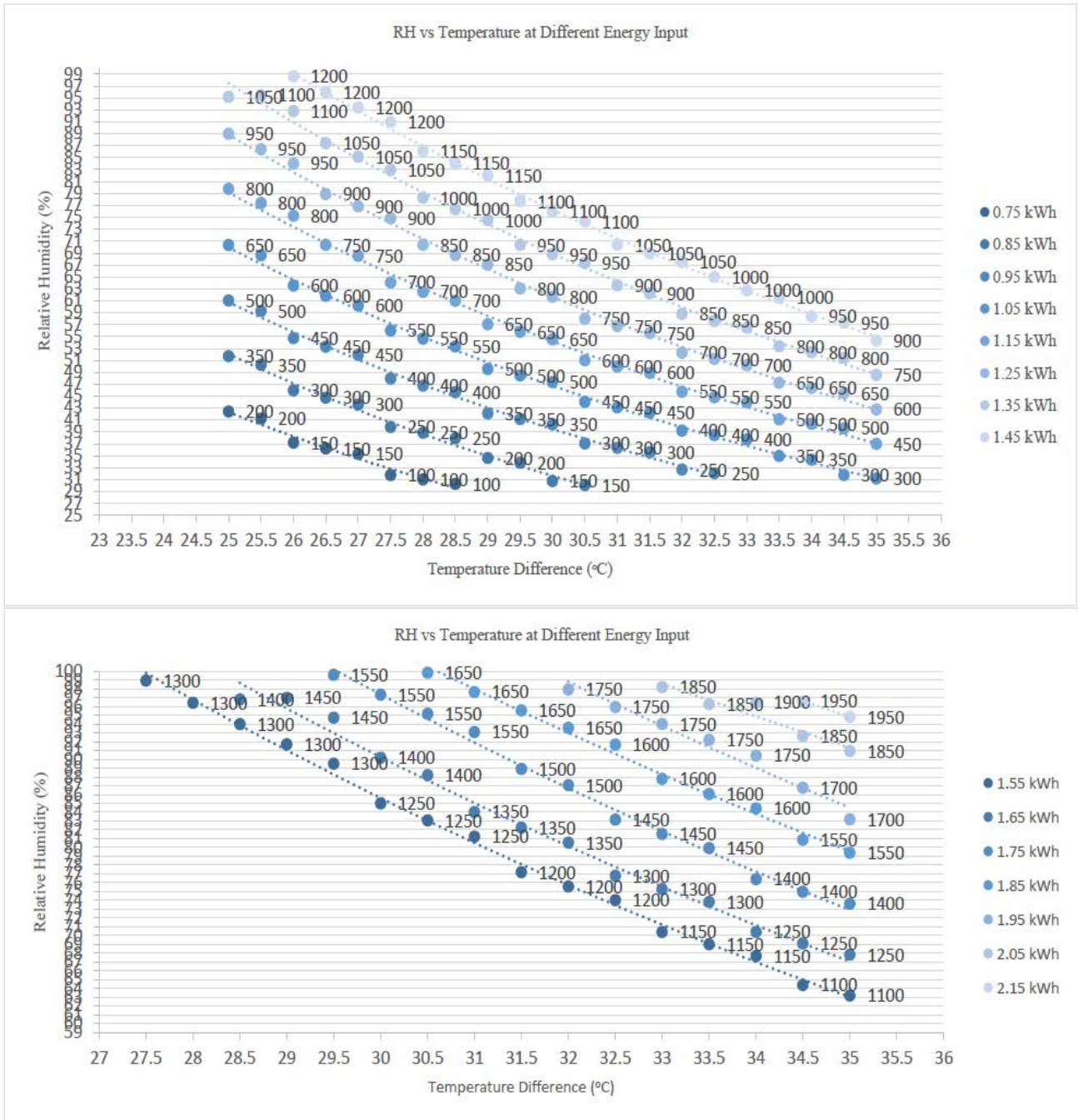
Fig. 3 – Yield at Different Energies for Different Temperatures and Relative Humidity Values
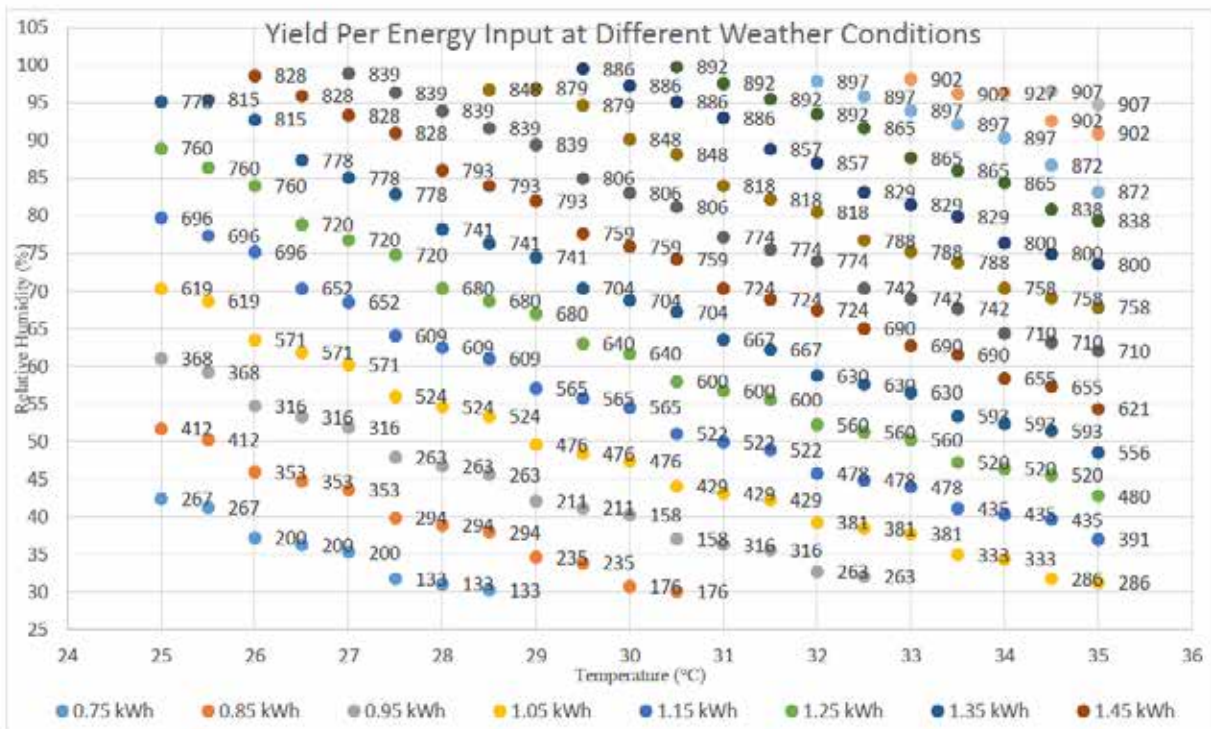
Fig. 4 – Yield Per Energy Input for Different Temperatures and Relative Humidity Values
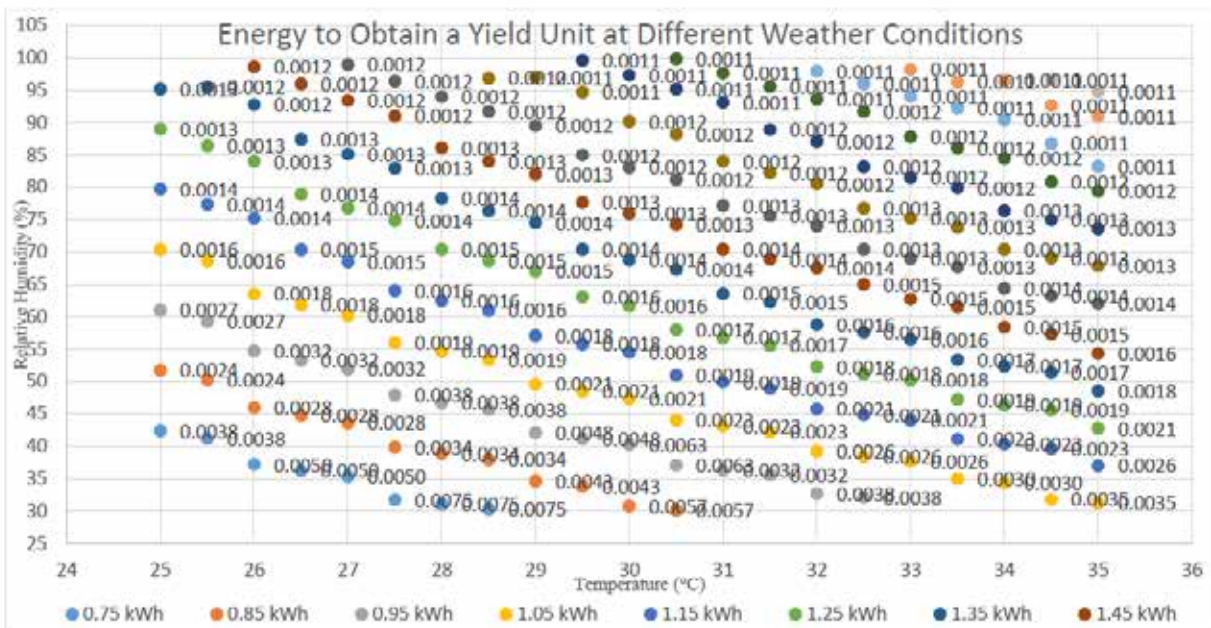

Fig. 5 – Yield at Different Energies for Different Temperatures and Relative Humidity Values

Investigation of Energy Piles as Sustainable Foundations in Lebanon

Houssam Ghandour, Hasnaa Housain, Reem Jaber, Isamar Mattar, Yara Teddy

Department of Civil and Environmental Engineering

American University of Beirut

Hamra-Lebanon

hmg06@mail.aub.edu, hsh17@mail.aub.edu, raj17@mail.aub.edu, iam12@mail.aub.edu, ymt03@mail.aub.edu

Abstract- **The use of energy piles is considered to be an innovative and sustainable method of energy conservation. These structural foundation elements act as thermal exchangers aiming to reduce the cost of energy required for heating and cooling demands. Due to the lack of established methods for the design of geothermal piles, this paper studies the effect of the heat exchange on the geotechnical capacity of piles in clayey soil. An experimental setup is being developed to study the effect of heat transfer on the load-response of model geothermal piles operating in cooling mode. A case study of a 10-story building located in Beirut, Lebanon is studied to analyze its geothermal system financially and environmentally. PileSim2 software is used to estimate the heating and cooling energy demands and the amount of geothermal energy utilized by the building. The results show a 30% reduction in the electrical energy consumption as well as a 2000 Tons saving in CO_2 emissions for 20 years of operation. At the end of the 20th year, the ground temperature reached nearly a constant value of 20.52°C starting from an initial value of 17°C. In addition, the payback period for the geothermal system is found to be between 5 and 7 years for interest rates of 3% and 10%.**

I. INTRODUCTION

The demand on energy has increased substantially over the past few decades, primarily as a result of the development and accelerated industrialization of many societies and nations. The main form of energy consumption in the industrial and power generation sectors is fossil fuel energy, which has been developed through the decomposition of plants and animals over the years. It can be predicted that this non-renewable energy resource will eventually diminish due to the excessive extraction. It is in that context that the attention is now directed towards the ability to harness renewable energy resources and decrease our reliance on oil and gas.

Developed countries started looking for renewable energy resources to sustain the continuous increase in energy consumption. One of the solutions which gained attention over the past few years is geothermal energy. Geothermal energy contains large amounts of the Earth's interior heat, making it a renewable resource that can be harnessed to operate power plants and turbines at deep depths. This paper, however, focuses on extracting geothermal energy at relatively shallow depths using heat pumps and geothermal piles.

Piles serve as foundation elements that support loads and stresses from buildings and other structures. They are used to transfer the load from the facility to the underlying soil through end bearing and resistance along the pile shaft. Moreover, piles may have an additional purpose as they can be used to exchange energy with the ground. These energy piles are equipped with pipe networks in which a circulating fluid undergoes continuous heat exchange with the surrounding soil. However, uncertainties lie in the effect of the temperature cycles on the geotechnical capacity of the pile and the properties of the surrounding soil. Thus, this paper includes a section aiming at studying the effect of heat transfer on the frictional and end bearing capacities of the piles through modeling and designing energy piles for controlled laboratory testing.

On the other hand, this paper presents a case study in Beirut, Lebanon determining the feasibility of implementing a geothermal system. PileSim2 software is used to simulate the behavior of the geothermal system through estimating the coverage of energy demands and predicting the ground temperature variation with time. The output of this software is used in a financial study and an environmental analysis to estimate Dollar savings and Green House Gases (GHG) emissions reduction.

The paper proceeds as follows: the next section introduces a literature review on the history and previous experiments conducted on geothermal piles. It is followed by a design and operation section which illustrates the thermo-mechanical behavior of pile in-situ. A material characterization section is then included whereby the properties of the soil are studied. The fifth section elaborates on the lab experiments that will be conducted, along with the objectives of each. Subsequently, the paper interprets a case study assisted by the PileSim2 software for the purpose of evaluating the performance of geothermal piles. A financial feasibility section is enclosed and an environmental impact section is then presented. Finally, the paper concludes with a summary of all the findings.

II. LITERATURE REVIEW

It was not until the late 19th century, that the idea of using geothermal energy became tangible [1]. The ideal setting that

distinguishes geothermal energy from other renewable energy resources is that it's not sensitive to climatic variations or any other environmental condition. While the Earth's temperature fluctuates seasonally in the first few meters, it becomes constant after a threshold depth with an average temperature range of 10°C - 18°C, depending on the region [2]. This fact makes it possible to benefit from the constant temperature during cold and hot seasons. In cold days, energy is harnessed from the soil by heat exchangers and can be used for heating purposes. In the hot days, the excess heat can be brought down to the underground that acts as a heat sink in this case.

The first application of harnessing geothermal energy by using structural elements started in 1980 in Austria and Switzerland with mat foundations. Afterwards, focus was directed on using geothermal piles, starting in Austria 1985 and becoming later the subject of many research efforts developed throughout the years in Europe [2].

In 1994, Morino and Oka were the first to utilize steel piles as heat exchangers. They evaluated the heat transfer properties of the piles through experiments and a 3D- finite element method. Concrete was favored over steel as material for heat exchanger [3]. Then in 1996, PileSim2 software which is a simulation tool for the energy piles, was developed by Pahud. Moreover, Laloui et al (2006) used finite element to analyze the effect of the increased stresses due to thermal exchange on the behavior of the geothermal piles. The results proposed a more pronounced effect of thermo-elastic strains as compared to mechanical strains [3]. In 2007, Sekine et al succeeded to develop a ground-coupled heat pump system using the drilled concrete piles with U-shaped pipes for heat exchange [3].

The studies continued to reach Lambeth College where pile-loading tests accompanied with temperatures cycles and extended period of loading were performed on 146 piles at a depth of 25m dipped in London Clay [4]. The strain profiles obtained suggest that the response of the geothermal pile is elastic, which means that the effects resulting from contraction and expansion of the pile seem to be reversible. It was concluded that the shear stresses induced due to thermal exchange are not large enough to cause failure [4].

In 2013, a physical model was simulated at Université Paris-Est. The load was applied in increments on a closed-ended aluminum tube, and heating was injected immediately by filling the pile with hot water. Afterwards, the vertical displacement was measured and analyzed. Results hat heating induces thermal dilation of the pile. Also, the magnitude of the applied load affects the pile behavior and consequently the factor of safety involved in its design. For loads exceeding 40% of the pile's ultimate resistance, any temperature increase in the pile is combined with long-term creep, thus further work is needed to detect these effects [5].

Throughout the next year, a paper by Akrouch aimed at targeting one of the areas that didn't receive much attention, namely the effect of temperature on the creep rate of geothermal piles. The results obtained yielded important conclusions mainly that the load induced due to heating is insignificant compared to ultimate values. As for the surrounding soil, the results reveal an increase in its viscous component due to the increase in temperature. This has increased the long-term displacement of the geothermal pile to about 2.35 times the displacement of a regular pile [6].

III. DESIGN AND OPERATION

Research is still on-going to determine rules and guidelines on designing geothermal piles. Published works, however, include recommendations for the design of piles and how to cater for their thermal operation. As material, Brandl recommends using reinforced concrete or thermally enhanced concrete along with water circulation through high-density polyethylene (HDPE) pipes fixed to the steel cage. In extreme cold weather, antifreeze solutions such as glycol mixtures or saline solutions have to be used to avoid freezing of the fluids, which may damage the pipe system [3].

As mentioned earlier, temperature variations could be utilized through fluid circulation in pipes to heat or to expel excess heat from a facility. During heating mode, the circulating fluid absorbs the heat from the ground and transmits it into the facility via the heat pump. However, in cooling mode, the heat delivered by the surrounding environment to the secondary system is exchanged with the ground using the primary pipe network. This exchange, shown in Fig.1, induces additional stresses and strains in the structural pile element, which should be accounted for during design. Limiting deformation due to constraints by the super-structure produces additional internal stresses [7]. During heating mode, the pile expands leading to the generation of an internal tensile stress. This causes the increase of the compressive stress and side friction resistance of the pile. As for cooling, an opposite effect occurs, leading to negative compressive stresses and reduction in skin friction [1]. However, it is important to note that for regular applications (temperature ranges of 20°C-25°C), the design of a regular pile needs no altering (i.e. no need to increase the factor of safety) to sustain thermal stresses. On the contrary, increasing the safety factor might impede the serviceability of the pile and induce additional costs [8].

According to Brandl, the overall system of the geothermal piles is composed of three main parts: primary system, heat pump, and a secondary system [3]. The primary system consists of the pipe network that either extracts or transmits the heat to or from the surrounding soil, while the secondary system is made up of a pipe network, which allows the circulation of the fluid inside the facility. The two systems are closed-loop systems, which are connected by the heat pump. The secondary system is also responsible to return the cold or the hot fluid to the heat pump in order to repeat the exchange cycle [9]. The heat pump transfers the thermal energy into the secondary unit. The latter unit consists of a delivery system, which is usually the HVAC system that undergoes cooling and heating processes inside the building. Sometimes, the heat pump is setup with a compressor that increases the temperature of the circulating fluid; this is done to meet the desired temperature of the facility.
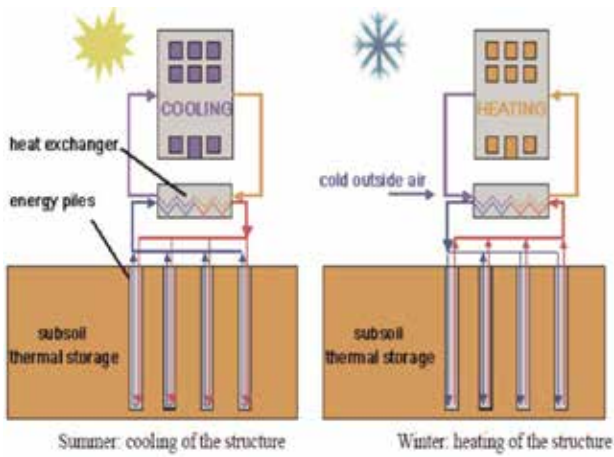
Figure 1. Energy flow during heating and cooling [9]

IV. MATERIAL CHARACTERIZATION

The soil is brought in buckets to the materials laboratory at the American University of Beirut (AUB) and dried for the purpose of sieving. Sieve analysis, hydrometer tests, Atterberg Limits tests, and specific gravity are performed. First, the composition of the soil is determined as follows: 57% fine material, 35% of which are clay and 22% are silt. Second, Atterberg Limits tests is conducted to conclude that the liquid limit (LL) occurs at a water content of 28.9%, whereas the plastic limit occurs at a water content of 17.7%. Third, the value for specific gravity is calculated to be 2.64, which conforms to the general range for clay.

In addition, 1-D consolidation tests are conducted on soil samples to produce *e-log p* curves in order to choose the most suitable water content. The water content is chosen based on the fact that clay needs a high amount of water to be workable enough. Thus three samples are studied: 100% LL, 80% LL, and 50% LL. Due to the similarity of the results of the three samples, a 100% LL is chosen for workability issues. The *e-log p* curve, found in Appendix A, shows that a pressure of around 40 KPa needs to be applied to yield enough consolidation before starting the experiment.

V. EXPERIMENTS

To model the real behavior of piles under the ground in clayey soil medium, experiments are conducted in the materials laboratory at AUB. The setup of the experiments consists of three geothermal piles and three control piles, each of 12 cm diameter and 60 cm length, which were casted at the materials lab at AUB. The pipe systems, configured with the help of the physical plant team, are of three types: double-U, triple-U, and spiral-shaped copper loops. Copper as a material is available in small diameters and is easier to mold with behavior similar to HDPE. In the spiral loop, water descends through the spiral then it exits through a vertical thermally isolated pipe to minimize losses. Fig.2 shows the double-U and spiral pipes.

Each experiment consists of three piles embedded in a 1m x1m x1.2m steel tank filled with clayey soil. Prior to embedding the piles, the tank is internally covered with geotextile around the sides for drainage purposes. Afterwards, the bottom is filled with a gravel layer of 20 cm thickness to model bottom drainage conditions. Then, the wet soil is inserted into the tank layer by layer. For consolidation purposes, the soil was loaded by 40 KPa and left for few days.

The cooling mode is modeled by choosing an inlet water temperature of 35°C by heating water in a water bath. Hot water is pumped to circulate through the pipe system to exchange heat with the 25°C clayey soil. In an effort to isolate any external heat effect and maintain this temperature, the tank is covered with foam around its external walls. The following three experiments are currently being conducted at AUB.

A. Pile load test

This experiment aims to compare the structural capacity (*end bearing + skin friction*) of a geothermal pile to that of a control pile. The piles are loaded with fluid circulating until steady state is reached. Steady state occurs when the temperature of the clay bed reaches a constant value measured by thermocouples spread along the pile length. The load is applied using the piston and geotechnical capacity of piles is measured after several cycles of heating until plunging failure is reached.

B. Skin friction

To determine skin friction capacity, the piles are pulled in tension. For this purpose, steel casings are introduced at top and bottom of the piles. The casings are plates of 12 cm diameter and 1 cm thickness and they are connected using two 16 mm diameter rods. The test will be done at steady state, and the end bearing capacity can be back calculated using pile load test results.

C. Efficiency

This experiment aims to check the effect of flow rate and pipe loop configuration on the efficiency of heat exchange between the pile and the soil. The flow rate will be varied in each trial measuring the variation of temperature along the pile depth for few hours. The flow rate corresponding to the largest ground temperature is the optimal flow rate. The same procedure is used to assess the effectiveness of the different pipe system configurations. Measuring the temperature variation along the pile depth, the most efficient system will yield the largest temperature gradient.
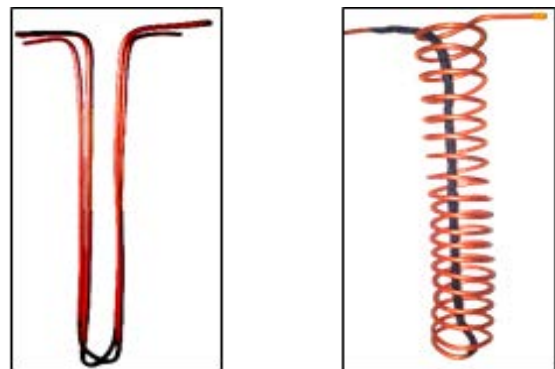

Figure 2. Double-U and spiral pipes

VI. CASE STUDY

The feasibility of implementing a geothermal system is studied for a building located in the city of Beirut, Lebanon. PileSim2 software is used to evaluate the performance of the heating/cooling system of the building with energy piles. The long term effect of the heat exchange on the ground temperature is simulated for 20 years. The results of this study are used to determine the operational and installation costs of implementing a geothermal system against a traditional one. Furthermore, the environmental benefits are considered through evaluating the reduction of CO_2 emissions.

The building under analysis is located in the North of Beirut, Lebanon. It is a 10-story building with 3 basements. The footprint is approximately 700 m^2 lying on a flat terrain. Each floor consists of 2 apartments; each has an area of around 200 m^2 and wide terraces. The soil profile consists of three main layers. The first layer, fill layer, ranges in thickness from 5 to 7 meters. The second layer is a 3 m silty sand underlain by an infinite sandy clay layer. Appendix B shows an AutoCAD plan of the building along with the soil profile.

The design calculations recommend the use of 70 drilled shafts, each is 40-meter long and 1 meter in diameter. The spacing between the shafts is assumed to be 3 m. Each energy pile consists of 6 U-shaped pipes allowing water to circulate as a heat exchange fluid; the pipes are 4 cm in diameter. This primary heat exchange system is connected to the heat pump.

A. PileSim2 Input Parameters

PileSim2 is used to simulate the behavior of the heat exchange system composed mainly of heat pump and geothermal piles. This software was developed by the Swiss Federal Institute of Technology in Lausanne (EPFL) and validated by measurements of existing systems. The heating/cooling system in the building is modeled as a "heating and geocooling" system. A percentage of the heat demand is covered by a heat pump coupled to the energy piles; the cooling requirements are partly satisfied by geocooling without the use of a cooling machine. The software computes the amount of geothermal energy exchanged with the surrounding soil, the auxiliary energy needed for heating and cooling, the electric energy used by the heat pump as well as the variation of the ground temperature during the 20 years of operation.

The weather data and ground characteristics are essential inputs for PileSim2. The weather data of Beirut, obtained from the airport weather station, were inputted as hourly values of temperature and solar radiation for the year 2014. The initial undisturbed temperature of the ground is set at 17°C equal to the ground temperature at 40 m below NGL. The thermal conductivities of the silty sand and sandy clay layers are 3.34 W/mK and 1.61 W/mK respectively [10].

B. Results and Analysis

The heating and cooling energy demands of the building are computed after running PileSim2. The following assumption is followed in order to determine the activation period of each mode: heating mode is turned off when the outdoor air temperature is greater than 20°C, whereas cooling mode is turned off when the outdoor air temperature is lower than 20°C. Fig.3 shows the monthly cooling and heating energy demands of the building. The total energy needed for cooling is 427.5 MWh /yr, while it is lower for the heating mode and equals to 88 MWh /yr. The cooling to heating energy ratio is around 5:1 representing the weather situation in Beirut.

The monthly variations of the ground temperature for years 1, 2, and 20 are plotted in Fig.4. The initial ground temperature decreases slightly in the first two months of the year since the heating mode is active during this period. When the hot season begins, the ground temperature increases as heat is injected into the ground at higher rate to fulfill the cooling demand. It reaches a value of around 19.9°C at the end of the first operational year. In the second year, the ground temperature slightly fluctuates between 19°C and 20.52°C. During the 20th year of operation, slight variation of the ground temperature is observed reaching an end value of 21.2°C. For further observation, the annual average ground temperatures during the 20 years are reported in Fig.5. The average ground temperature significantly increases in the first 5 years, and then its rate of increase reduces during the next years. Eventually, the temperature reaches nearly a constant value of 20.5°C.
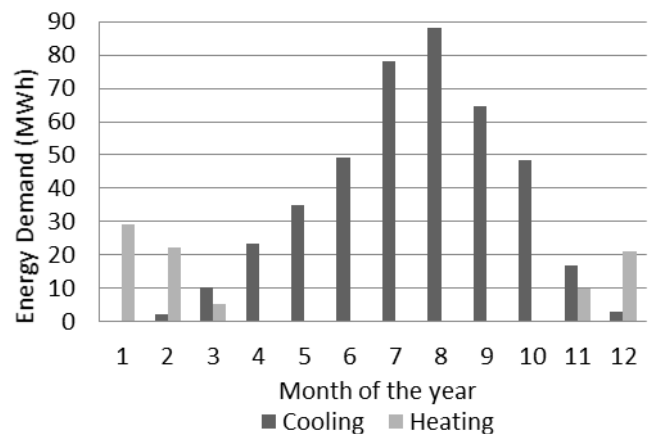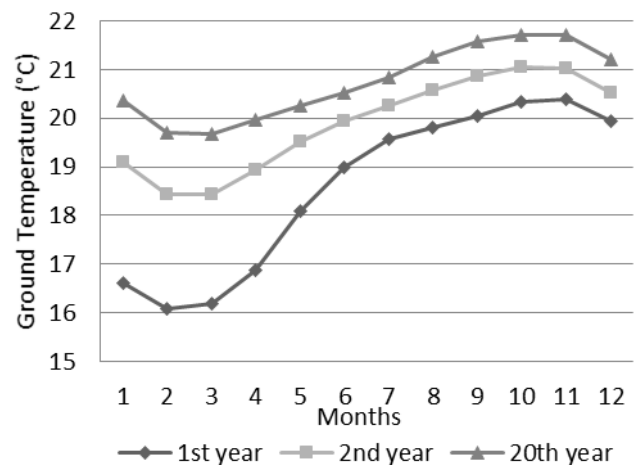

Figure 3. Heating and cooling energy demands

Figure 4. Ground temperatures during the 1st, 2nd, and 20th years

Figure 5. Average annual ground temperature (°C)

The graphs in Appendix C show that the heating demand is constantly covered by 80% since the heat pump needs electrical energy to operate. On the other hand, the cooling energy coverage decreases with time as the efficiency of injecting heat decreases due to the increase in ground's temperature. The cooling coverage is around 20.5%, while the average energy supplied by the whole system is around 30%.

The initial annual ratio of injected to extracted energy decreases with time and eventually follows an asymptote with a value of 1 as shown in Fig.6. This asymptotic direction represents the equilibrium state of the system where the amount of the injected heat is equal to that extracted from the ground. Furthermore, the geothermal system of the building will operate at a constant minimal efficiency as the ground temperature reaches its maximum value. At this state, the heating demand is still covered completely, whereas the cooling coverage will reach its lowest value. Referring to Fig.5 and Fig.6, the period during which the average ground temperature starts to be somehow constant coincides with the period at which the ratio of injected to extracted heat is around 1. This is consistent with the fact that the ground temperature reaches constant value at equilibrium.

VII. FINANCIAL FEASIBILITY

A financial analysis of the geothermal system of the building is performed based on the comparison of the Net Present Value (NPV) of geothermal and traditional systems. The system of higher NPV is considered to be better investment. The entire life of both systems is set at 20 years and interest rates of 3% and 10% are used. The geothermal system includes higher initial investment costs due to the additional required material and machines such as pipes, heat pump, valves, and compressor. However, based on PileSim2, the use of energy piles can save around 30% of the annual total heating and cooling costs. In this section, NPV analysis is presented and the Payback time is determined.

The investment costs are reported in Table I for the geothermal and traditional systems. The initial over cost of the geothermal system compared to the traditional one is estimated to be equal to $ 98,520. The cost of HDPE pipes is calculated considering 70 piles of 40 m long, each containing 6 U-shaped pipes and considering a cost of 0.67$/m. The cost of the installation depends on the labor time needed to finish the work; the labor time is based on a specialized contractor Elie Rafie. The price of the additional machines used in the geothermal system is estimated based on Association Française des Professionnels de la Géothermie [11].

Moreover, the annual electric energy needed to operate both systems are computed via PileSim2, and cost of the amount electric energy is calculated according to the pricing policy of Electricité du Liban institution shown in Appendix D. Finally, the net present values with different interest rates is presented in Fig.7. It is shown that the payback period of the geothermal system is around 5.6 years for 3% interest rate and 6.8 years for interest rate of 10%. After 20 years of operation, the net saving of the geothermal system equals to around $ 166,000 for 3% interest rate, while it is much lower for 10% interest rate and equals to around $67,000.
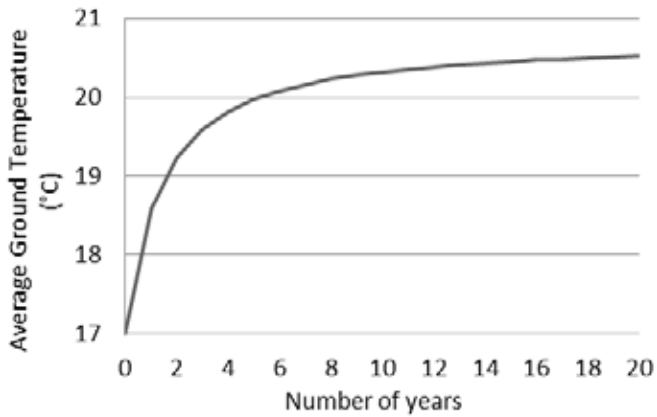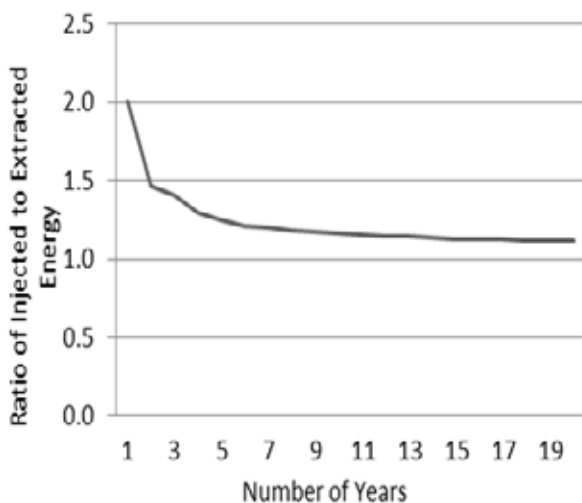

Figure 6. Annual ratio of injected to extracted heat energy
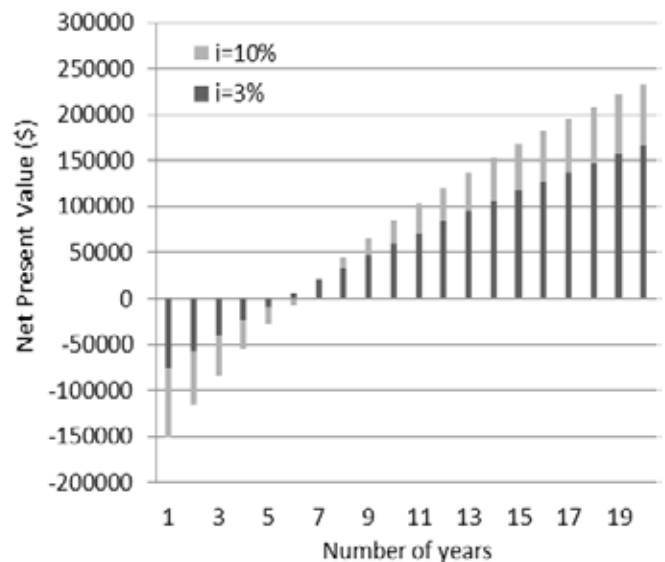

Figure 7. Net present values with interest rates of 3% and 10%

289

TABLE I
INVESTMENT AND ANNUAL COSTS OF BOTH SYSTEMS

Costs ($)	Geothermal System	Traditional System
Pipes	23,520	-
Installation	40,000	-
Heat pump	35,000	-
Annual Electric Energy	38,000	55,000
Total	136,520	55,000

VIII. ENVIRONMENTAL IMPACT

The environmental analysis is evaluated by assessing the savings in Green House Gases (GHG) emissions annually and during the 20 years simulation period. Using BLCC 5 software, consumption of 1 KWh releases around 0.65 kg of CO_2, 3.3 g of SO_2, and 0.97 g of NO_x.

The following table represents the GHG emissions to satisfy cooling and heating requirements by the traditional and geothermal systems. Based on the results of PileSim2, the amount of KWh energy was converted to amount of gases to calculate annual and life-cycle reductions. It can be concluded that geothermal system decreases GHG emissions by 30% on average as shown in Table II. In addition, about 2000 Tons of CO_2 are saved during the 20 years analysis period, which is equivalent to the emission of CO_2 from 25 cars circulating for 20 years continuously.

TABLE II
GHG REDUCTION

	Traditional System	GHE System	Annual Saving	Total Saving
CO_2 (T)	336	236	101	2,020
SO_2 (T)	1.7	1.2	0.5	10.2
NO_x (T)	0.50	0.35	0.15	3.02

IX. CONCLUSION

Due to the increase in energy consumption, the search for sustainable energy resources has become one of the most critical issues worldwide. Significant research is driven towards geothermal energy since it presents itself as one of the most abundant, renewable, and clean resources. Many processes can be used to harness this energy, one of which is the usage of geothermal piles, whose history, operation, and previous applications were presented in this paper.

Besides, this paper targets two main goals. The first is to introduce a series of experiments that are conducted for the purpose of modeling the behavior of energy piles and the effect of heat transfer on pile capacity. The second focuses on a case study, through which PileSim2 software is used to estimate the financial and environmental benefits of using geothermal piles.

The results of this study show a net saving of 30% in the electrical energy consumption which leads to a 30% reduction in the emissions of Green House Gases. The software also shows that the ground temperature increases from 17°C to reach nearly a constant temperature of 20.5°C at the end of the 20th year. As for the effect of heat on the geotechnical capacity, research is still on-going and results are to be presented during April 2016.

REFERENCES

[1] Abdelaziz, S., Olgun, C., and Martin, II, J. (2011) Design and Operational Considerations of Geothermal Energy Piles. Geo-Frontier 2011: pp.450-459.

[2] Brandl, H. (2006). "Energy foundations and other thermo-active ground structures" Geotechnique56, No. 2, 81–122.

[3] Bach et al. (2010). "Technological advances and applications of geothermal energy pile foundations and their feasibility in Australia". Renewable and Sustainable Energy Reviews 14, 2683-2696.

[4] Bourne-Webb P. J., Amatya K., Amis T., Davidson C., and Payne P. (2009). "Energy pile test at Lambeth College,London: geotechnical and thermodynamic aspects of pile response to heat cycles". Géotechnique 59, No.3, 237-248.

[5] Kalantidou, A., Minh Tang, A., Pereira, J., Hassen, G.(2012). "Preliminary study on the mechanical behaviour of heat exchanger pile in physical model."Géotechnique, Thomas Telford, 62 (11), pp.1047-1051.

[6] Akrouch, G., Sánchez, M., Briaud, J. (2014)."Thermo-mechanical behavior of energy piles in high plasticity clays" DOI: 10.1007/s11440-014-0312-5

[7] Laloui, L., Di Donna, A. (2011). "Understanding the behavior of energy geo-structures." Federal Institute of Technology, EPFL.

[8] Mimouni, T., & Laloui, L. (2013). Towards a secure basis for the design of geothermal piles. Acta Geotechnica,9(3),355-366. doi:10.1007/s11440-013-0245-4.

[9] Cervera, C.P. "Ground Thermal Modeling and Analysis of Energy Pile Foundations." 2013.

[10] Hamdhan I.N., Clarke B.G., (2010). "Determination of Thermal Conductivity of Coarse and Fine Sand Soils". Proceedings World Geothermal Congress. Bali,Indonesia, 25-29 April.

[11] Boissavy C.(2015). "Cost and Return on Investment for Geothermal Heat Pump Systems in France". Proceedings World Geothermal Congress. Bali, Indonesia, 19-25 April.

[12] G. Eason, B. Noble, and I.N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529-551, April 1955.
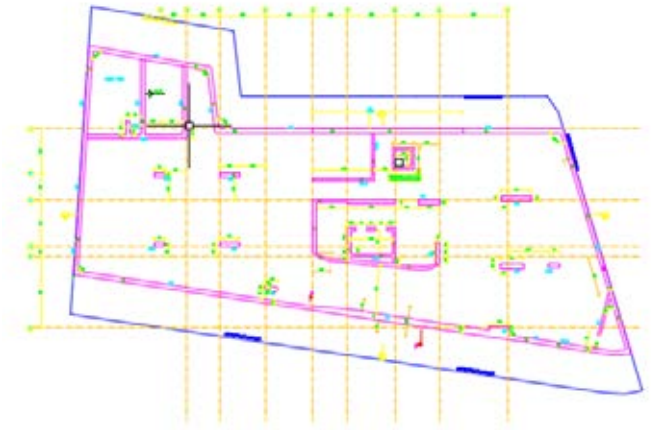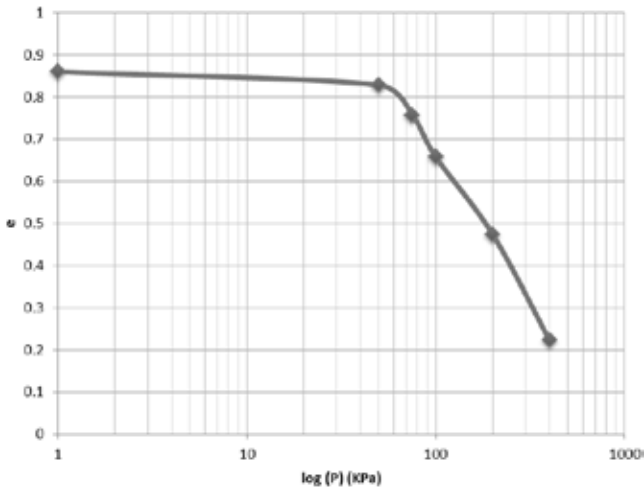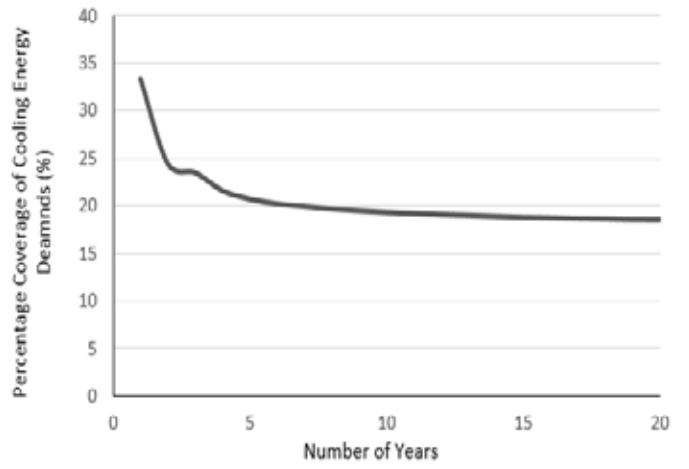
WC= 100%*LL

Fill Layer

5.0

3.0

Silty Sand
Cu= 10 KPa
Ø= 35°
γ= 17 KN/m3
λ=3.34 W/mK

Sandy Clay
Cu=30 KPa
Ø= 10°
γ= 22 KN/m3
λ=1.61 W/mK

291

APPENDIX C- CONTD'
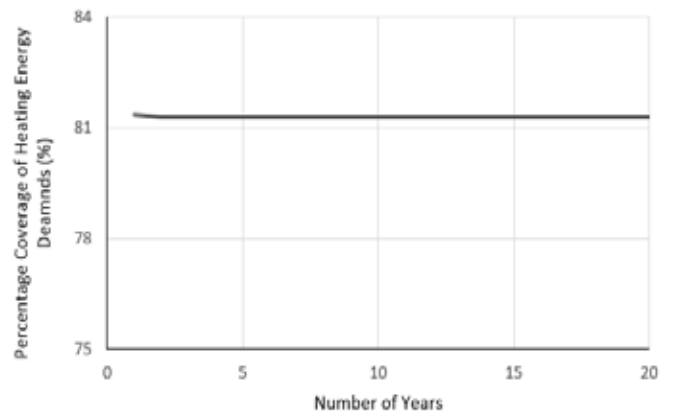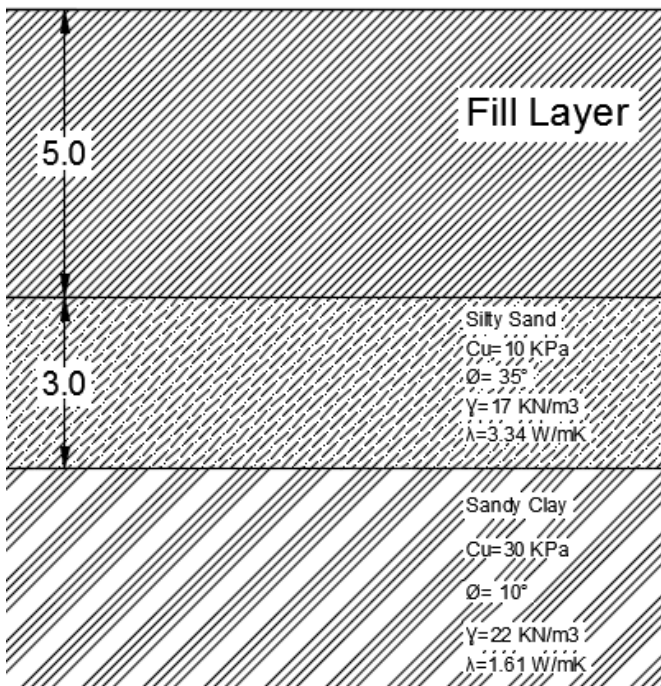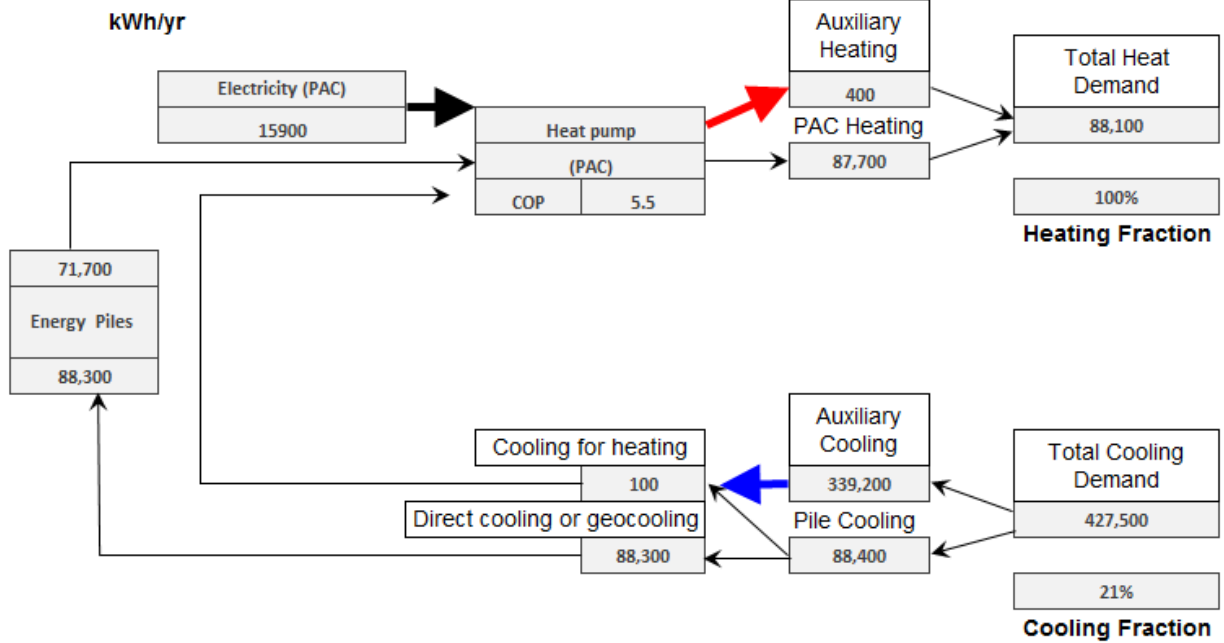
APPENDIX D
EDL Electric Pricing

	On monthly basis bill			On two months basis bill	
Tran.	Energy consumed (kwh)	Price L.B.P /Kwh	Tran.	Energy consumed (Kwh)	Price L.B.P /Kwh
1	100	35	1	200	35
2	200	55	2	400	55
3	100	80	3	200	80
4	100	120	4	200	120
5	> 500	200	5	> 1000	200

292

Mix Design and Automation of Concrete 3D Printer

Hanin Arja, Lynn Fayed, Dana Itani, Toni Narciss, Nayirie Skayan
Civil and Environmental Engineering
American University of Beirut
Beirut, Lebanon
haa127@mail.aub.edu, lhf06@mail.aub.edu, dmi10@mail.aub.edu, twn04@mail.aub.edu, nss22@mail.aub.edu

Abstract-With the trend shifting towards automation, the fast-growing industry of construction has been gradually adapting to this change. Engineers are urged to develop innovative techniques and original inventions to meet demand complexity, and overcome construction challenges. One of the latest revolutionary methods, though still under study, is 3D concrete printing. The automated mechanism and its product have various advantages of assisting engineers in minimizing human errors, achieving higher productivity, and designing architectural elements with complex geometry. The main focus is to consider integrating various design components and material characteristics in order to obtain a coherent concrete 3D printing system operating as a single entity. The scope of work is divided into several categories: design of the optimum concrete mix, fabrication of the extrusion nozzle with the integrated pumping mechanism, and coordination between material properties and mechanical parameters. The compatibility of the concrete mix design with the machine functional parameters is clearly assessed through numerous experimentations in which flowability, buildability, open time, and optimum additive ranges have been established. Flow of concrete is experimentally tested to function in accordance with the rate of motion of the machine, and the rate of extrusion from the head. The preparatory experimentations on the concrete mix were adjusted to fit the machine parameters. The final result was a coherent synchronized system able to operate and print several interconnected layers of different geometrical shapes. The machine generated a wall of 21 layers, a semi-circle and an arch shaped structures as successful outcomes. . The testing outcomes have shown that the ultimate concrete mix for this specific printing application consists of 600g of cement, 650g of sand, 250g of fine aggregate, and a water to cement ratio of 0.4. Addition of 3g of superplasticizer and 0.5g of fiber is indispensable to ensure that the concrete is continuously extrudable from the nozzle over all the printing period. The Archimedes screw, rotating at a speed of 20 rotations per minute, provide the needed pressure for concrete to continuously flow from a height no greater than 2 cm, and a machine feed rate along the three axes of 3cm/sec. From a managerial perspective, this innovation reduces construction time, manual labor and carbon footprint which comply with the sustainability requirements nowadays.

I. INTRODUCTION

Construction is one of the largest sectors in economy. The escalating complexity of the construction industry is causing increased difficulties in the execution process, direct and indirect costs, and project delays. In the last few years, with the technological improvements and growth in population, the human needs were amplified. More than half of the world's population is now occupying metropolitan regions. Investors have thus started capitalizing assets in a wide range of projects to serve the flourishing building demands. For them to succeed, contractors have to integrate and optimize cost, quality, and safety parameters into a single well-coordinated system. New techniques in the construction field should aim at mitigating the problems of low labor efficiency, deteriorated work quality, cut and trim waste, safety, and site control. Concrete 3D printing, which is based on a layered manufacturing process (LM), is a rising technique that has recently been taking a dominant position in the construction realm. The structure is built layer by layer using an automated machine extruding concrete [1]. The Contour Crafting (CC) technique was initiated and developed by Behrokh Khoshnevis, a professor of Industrial &System Engineering at the University Southern California (USC).

In field applications, the Chinese Company Winsun used the 3D printing technology to build ten houses out of recycled materials in twenty four hours, with a cost less than 5000 USD. Winsun's method is based on building blocks in a central factory and then assembling them on site. Very few labors are needed in the assembling work hence decreasing the cost of construction [2]. Also, Lewis Grand Hotel in the Philippines is the first operational 3D printed hotel in the world (130 square meters) that took approximately 100 hours of discontinued work to print. Yakich hopes in the upcoming years that 3D printers will enhance the production of low income housing in Philippines [3].

Furthermore, Amsterdam-based DUS Architects developed their latest 3D printer "The Kamer Maker". DUS Architects described Kamer Maker as their new "craftsmanship". They believe 3D printing will encourage governments to build affordable homes because of savings in time and cost. Therefore, one significant difference between traditional construction methods and 3D printing is efficiency [4].

The main advantages of such execution process are to increase the margin of creativity in the structural geometries, to reduce the constraints originating from the conventional use of formworks. In addition, 3D printing aims at minimizing cost, time and waste along with maximizing the client's added value [5].

For the aim of bridging some gaps left behind by previous researchers and overcoming further challenges, past implementations are further assessed and extended. Through the use of an automated version of the concrete 3D printer, many achievements become more tangible. The ultimate goal of using the CNC concrete machine is to transform any drawing inputs into an endurably stable structure of any geometrical shape. Extensive preparatory laboratory testing is carried out to optimize the concrete mix and identify the suitable material properties. Integrating the machine parameters, nozzle design, and operation control, and coordinating the entire system as a whole are primordial stages in the fabrication process. The end results of concrete structures are obtained and analyzed according to specific criteria illustrating flow and extrusion measurements, as well as open time and overlap of extruded concrete layers.

II. METHODOLOGY

In attempt to achieve a single coherent concrete 3D printer able to transform any drawing inputs into an endurably stable structure, it was essential first to optimize the part before coordinating the whole. Consequently, the design of the concrete mix, extruding nozzle, Archimedes screw, the machine dimensions, and the feed rate is independently examined. In the absence of conventional and exclusive testing methodologies for concrete printing application, numerous available laboratory tools were used to design relevant testing experiments and quantify factors of interest.

A. Preparatory Mix Testing

The testing methodology primarily tackles the ultimate concrete mix design, the proportions for printing application, and the effect of addition of aggregates, admixtures and fibers on the rheological properties of cement paste.

The optimal concrete mix should meet composite performance requirements rendering it fit for extrusion and printing. The self-consolidating concrete (SCC) used must be able to flow throughout the hose and out of the nozzle. Nevertheless, its properties from mixing stage to final setting might seem conflicting. The biggest challenge lies in printing layers displaying minimal deformation under their own weight. In addition, lower layers should bond to upper ones to form a strong homogenous system. The SCC is to be extruded while maintaining a consistent flow rate, and unless vibrators are introduced to the system, the mix must keep its workability through the entire "open time", that is the time from mixing initiation to printing finalization. The extruded material should consequently be flowable rather than stiff so as not to block the 2 cm nozzle head. Moreover, if the workability of the extruded SCC exceeds a certain threshold, this detrimentally affects the characteristics of the upper layers. Explicitly,

printed layers must not deform excessively under their own and other layers' weight. In case they lose their bearing capacities, the buildability of the entire structure is negatively affected. Hence, it is important to maintain an appropriate balance between the contrasting properties of SCC for the success of concrete printing application.

B. Machine Parameters

As opposed to laser engraving or plastic printing, concrete integration into the printing field is rather incipient. In the absence of Computer Numerical Control (CNC) machinery dedicated for concrete printing, it was necessary to customize the dimension and parameters of the printer to fit particular requirements. For instance, high precision, temperature control and slow motion are irrelevant parameters with limiting advantages in concrete application. Therefore, as compared to any other types of 3D printers, concrete needs considerable requirements of which is the machine's relatively large scale. Elevation is a crucial dimension, mainly because the CNC is supposed to build a relatively high structure. For the previously mentioned reasons, the preferential working zone of the printer is 1.2m in length, 1.2m in width and 1m in elevation. Fig.1 shows the preliminary sketch of the machine based on which it was manufactured. The maximum motion speed of the machine is 2500mm/min. The minimum lifting capabilities of the arm is requested to be greater than 15 kg for it to carry the weight of the concrete filled nozzle. The section below further elaborates on the importance of the previous number.

C. Nozzle Design

Designing the extrusion head is a major contributor to the concrete flow. Knowing the lifting capabilities of the CNC machine, the maximum lifting weight is estimated. The motion along the vertical axis is the main parameter to consider, simply because stepper motor controlling motion will have to accommodate forces resulting from lifting the entire system, and weight forces carrying the head upward. It is important to note that the use of stepper motors has various advantages and disadvantages. Its main benefit is that it gives high terms of accuracy when it comes to motion. However, if exceeded, the torque capabilities of the motor might cause the system to vibrate at high elevations, as the motion created through a rack gear transmission can possibly slip.

Figure 1. AutoCAD drawing of the machine and the actual manufactured CNC

For the previous reasons, the design methodology will follow a backward approach to calculate the maximum weight capacity that can be lifted by stepper motors and is responsible for motion along the "z axis". Selecting different alternative design parameters is restricted to the available lifting capacities. The governing parameters are the fabrication material of the nozzle, the geometrical dimensions of the head, the volume of concrete that can be accommodated, and the extrusion mechanism of concrete from the nozzle. Based on the above parameters, studies were performed to decide on the proper nozzle dimension and size. Fig.2 below provides detailed specifications of the Archimedes screw and head casing to be later fixed on the CNC machine.

The Archimedus screw is known to be an ideal choice when it comes to lifting considerably large water quantities. The main advantage of this screw pump is that it adapts to the volume input without the need of any control mechanism. When designing the nozzle, the main objective was creating a medium of constant pressure, hence facillitating the extrusion of consistent concrete layers. The main challenge at this stage is to adopt the same water mechanism to other viscous fluid such as concrete. Theoretically speaking, the Archimedus screw is supposed to rotate at different speeds in order to satisfy three major criteira: (1) flowability chatecteristics of the extruded material (2) the feedrate of the nozzle and (3) the consistency in the printed layers. Because its operation does not require any sophisticated control mechanism, and because of its relatively simple manufacturing, the Archimedes screw is assumed to be a feasible and tangible option in this project.

D. Operation Control

In order to create constant pressure in the nozzle, the Archimedes screw is supposed to rotate at fixed speed. Previous experimental results states that, for optimum extrusion and flow, the desired rotational speed of the Archimedes screw must fall in the range of 15 to 60 rpm. As for cleaning considerations, water must flow at even higher velocity to provide high pressure for removal of residual concrete. Therefore, given the necessity for varying the rotational speed of the Archimedes screw, the need of a stepper motor was essential and irreplaceable. After establishing the type of motor providing the sufficient kinematic energy input, the next step is to determine its required torque capacity. An experiment was designed accordingly to specify a safe motor operation speed. Throughout this experiment, the nozzle was intentionally

clogged. The mix used for clogging the nozzle was composed of a 3.4g to 1kg cement of superplasticizer and 25% of fine aggregate. This mix replicates the worst case scenario that might be encountered throughout the printing application. Finally, a spring scale was attached to the nozzle's handle, and manual forces were applied to unblock the bottleneck created in the nozzle. The respective force module was recorded. The three readings taken were as follows: 4N, 6N and 11N. Finally, these results were averaged and then multiplied by a factor of safety of 1.5. Hence, the ultimate force required was found to be 11N.

The mechanism used to operate the Archimedes screw was to couple it to a stepper motor, hence enabling the variation of the rotations per minute (rpm) and the rotation polarity. For this purpose, an electric circuit was designed based on the following main electric components (1) Arduino (2) voltage regulator (3) stepper driver and (4) power supply.

The following steps denote the major functional technicalities of the designed circuit:

1- The power supply transforms the input of an accelerating 220V (AC) to a direct 24V (DC)
2- The voltage regulator takes the 24V (DC) input and transforms it to a 5V (DC) output needed to operate the Arduino.
3- The Arduino is the brain that controls the stepper driver. This can be achieved by sending a set of instructions using Arduino programming language (based on wiring) to the micro-controller on the board.
4- The stepper driver controlled by the Arduino drives the stepper motor. The stepper driver is configured to achieve micro-stepping. In the latter mechanism each electric impulse will be transformed into a micro step which is equivalent to 1.8 degrees over 4. Achieving a system operating by micro-stepping is a must in the application of concrete printing, mainly to decrease the resulting oscillations during printing.

The standard CNC software takes input drawings, slice them into layers, and generate a G-code which is the major determinant of the machine motion. Nevertheless, all standard available software do not fit concrete 3D printing application, in the sense that it requires temperature, filled pathways and many other irrelevant inputs. Moreover, most slicer programs generate inapplicable codes, basically because they cut layers from top to bottom rather than from bottom to top, as per 3D engraving software. One further limitation is that the machine in the laboratory is only capable of operating based on coordinate inputs rather than absolute distances. Consequently, all the previous factors urged the authors to program a C++ code that takes dimensions from the user, and generate a G-code that is readable and operative by the CNC machine.
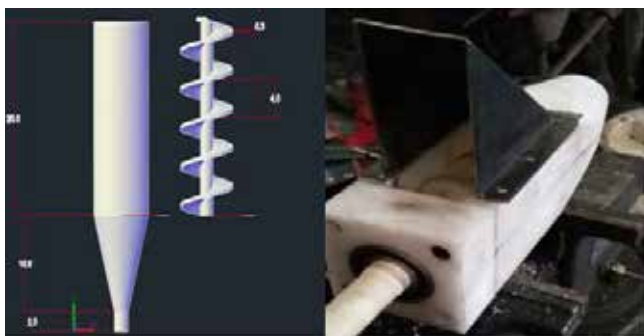

Figure 2. AutoCAD drawing of the Archimedes screw with the nozzle container and the actual manufactured nozzle

E. System Coordination

Once all the previous specifications have been considerably addressed, the main concern at this stage is to proceed with the assembly of the different entities into a single system. The critical task here is to implicitly integrate all the conflicting concrete mix properties into the picture through adjusting the machine parameters to the optimum mix characteristics. The procedure to follow is elucidated in Fig.6. To initiate operation of the system, three major inputs must be furnished: the preparatory concrete mix from earlier experimentations, the 2D drawing, and the rotational speed of the screw. The first checkpoint is to verify that the mix is still extrudable under controllable automated settings. In the case where concrete was not driven out of the nozzle, rpm must increase to provide the necessary consistent pressure for the mix to flow. Secondly, once pushed out of the nozzle, concrete must be printed in a continuous manner. Any observed discontinuities in the layer indicate that the machine feed rate is immoderately fast. Nevertheless, if the flow from the nozzle is intermittent, a different approach must be adopted: the incremental addition of superplasticizer is required, and thus the preparatory mix must be adjusted. Finally, in terms of buildability, the successive layers are supposed to adhere to each other, hence suggesting that the interlayer delay must be lengthy enough for the structure to support itself, yet short enough to ensure the necessary cohesion between layers. This parameter can be determined through assessing the open time of the mix. The output of this procedure is a well-coordinated system which parameters are matched to the benchmarked concrete mix.

III. RESULTS ANALYSIS AND DISCUSSION

The concrete mix that should be achieved to serve the concrete printing application must account for the following characteristics (1) flowablilty, (2) extrudablity, (3) workablity, (4) and strength. For the previous criteria to be met, numerous mixtures are tested to find the optimal proportions of sand, fine aggregate, superplasticizer, and fibers guaranteeing the minimum strength requirements. Initially, the water to cement ratio is fixed at 0.4 by all means to guarantee a minimum strength of 40 MPa in compression to satisfy the initial strength requirements. Before proceeding with the addition of aggregate, the starting point was to determine the proportion of sand relative to cement in attempt to achieve a reference material paste. By first assuming a 1 to 2 cement to sand ratio, segregation of the paste within the nozzle was inevitable. It was necessary then to reduce the quantity of sand relative to the binder so as to avoid segregation and cracks within the extruded filament once the paste dries. Next, the amount of aggregates was gradually varied through incrementing its percentage relative to dry matter. It is important to note here that the maximum aggregate size used is no greater than 2.36 mm, around one tenth the nozzle head diameter to overcome any complications originating from the clogging of the nozzle. One major consideration to account for at the stage is to make sure that the poured filament does not display any discontinuities due to the presence of aggregate. Once the
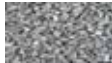
optimal ratios are fixed, the "control batch" is ready and hence must not be altered later on. Next, the effect of addition of superplasticizer (Viscorete) on the mix is closely assessed. Admixtures with varied dosages are added to the control batch until reaching a stage where the final mix design meets the extrudibility, flowability and workability criteria. Finally, to intercept any potential cracks and increase ductility, fibers are incrementally added until the mix is no longer extrudable. The resultant fiber percentage is hence the maximum allowable amount not clogging the nozzle. The previously described methodology sets practical guidelines and indications for establishing a mix design that is extrudable and workable.

A. Flow and Extrusion Measurements

In the preliminary study, it was necessary to assess the functionality of the designed nozzle before attaching it to the machine body. The experiment carried to define the variation of nozzle efficiencies with the changes in mix design is summarized in the graph of fig.3. The Archimedes screw was connected to a rotating arm, and the manual rotations were put in sync with periodic beeps. The amount of concrete flowing during a fixed interval of 15 seconds was determined using a volumetric flask. The same experiment was carried out at different rotating speeds (rpm) and superplasticizer percentage. The aim was to primarily define the optimum mix in terms of acceptable flow and ease of extrusion of the nozzle head. Fig.3 below summarizes the result of the experiment.

In order to determine the optimal rotational speed for concrete mixes with various superplasticizer percentages with respect to binder, the test was carried at different rpms. The objectives are to primary express the variation of flow with the rotational speeds, and to capture the mix which results are the most consistent flow. These are major considerations, mainly Looking at the graph, mixes with superplasticizer percentages

Table I
Optimal Concrete Mix Ranges

Material	Range	
	Minimum	Maximum
	1 to 1 (sand to cement ratio)	2 to 1 (sand to cement ratio)
	10% of dry fines	25% of dry fines
	0.2% of cement	0.4% cement
	3.4g/1kg cement	5.75g/1kg cement
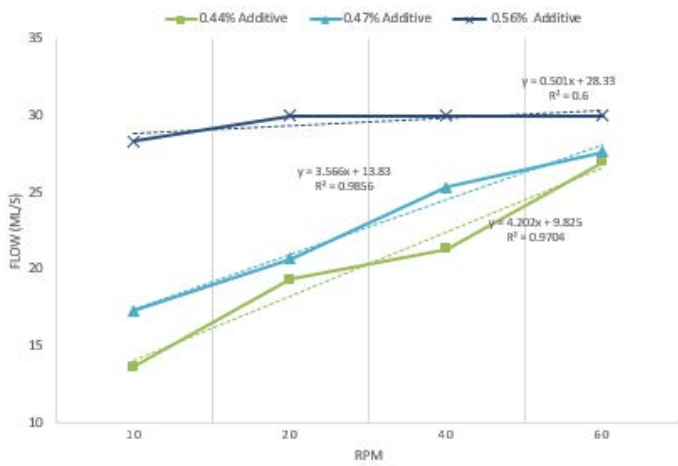
Figure 3. Variation of flow rate with respect to rpm

of 0.44 and 0.47 are rendered equally likely to fit the flow and extrusion for printing requirements. The general pattern observed indicates that the faster the rotation, the more is the volume of concrete extruded from the nozzle head. Nevertheless, when the amount of plasticizer is relatively large, it is believed that the screw mechanism has little effect, if any, on the concrete flow. When looking at the 0.56% curve, it is noticed that no matter what the increase in rpms, the flow rate of the extruded material remained 30ml/sec. Consequently, although it is highly preferable to use flowable concrete for its ability to diffuse in the hose, it makes the concrete extrusion from the nozzle uncontrollable. Note also that for lower percentage of superplasticizer, the results were predictable for low rpm values. It is important to mention that because the nozzle arm was manually rotated; all the above numbers and values are subject to further adjustments under automated controllable settings.

B. Offset Test and Open Time Measurement

The following bond test was conducted to determine the effect of delay in time on the cohesion between two overlapping concrete layers. The method adopted was printing four sets of two adjacent lines at $t = 0$ followed by printing two overlapping lines above each set at $t = 10, 20, 30, \& 40 min$ respectively. Each set was cut into a section of $4cm \times 8cm$. The bottom of the first layer and the top of the overlapping one were glued with a contact area of $24cm^2$, using epoxy glue of strength 25 MPa to steel plates and left to dry for 24 hours. The two plates were subjected to tensile forces to gauge the cohesive strength between the two overlapping layers.

The results obtained for all the samples exhibited a failure of the epoxy glue at the top layer under a stress of 3.3 MPa without any apparent cracks between the two layers, which can be contributed to the circular shape of that layer and the incomplete solidification of the epoxy glue. Furthermore, when examining the cross-sections there were no visible lines separating the layers; at $t = 40 min$ complete cohesion was still achieved.

Figure 4. Bond test apparatus (left) and epoxy failure (right)

The following test was conducted to determine the optimal center to center offset(d_{off}) between two adjacent layers. Given that the approximate width of the single bottom layer is measured to be 3cm, the method adopted was printing four sets of two adjacent lines while varying d_{off} by increments of 3mm {27, 24, 21, 18 mm}. In each set, three parameters were measured: total width, Δh, and height of interlock.

Table II
Offset Parameters

	Δh (mm)	total width(mm)	height of interlock(mm)
27mm	0	59	13
24mm	1	57	16
21mm	2	55	18
18mm	3.5	50	20

The results tabulated show low interlock values in sets 1 and 2 which result in weak lateral cohesion and a visible notch. Also a significantly high Δh value in set 4 was recorded which can affect vertical buildability negatively by unbalancing the overlapping layer. Therefore set 3 was chosen to be an optimal balance between balance and lateral cohesion.

C. Major Findings

Based on the printing system described, and given that, at all time, the mix was flowable and no extrudibility complications were found, the simplest geometric shape to start with was a linear wall. Varying the interlayer delay from 32 seconds to 92 seconds with an optimum rpm ranging from 8 to 20, and a federate of 3cm/sec, it was noticeable that after several iterative trials, the printing mechanism should proceed with a batch system. This system varies from 6 to 10 layers based on the complexity of the printed geometric shape. The need for a batch system was an advantageous option, mainly because it gives time for the lower layers to set and withhold the structure at higher elevations. However, the main drawback lies within the intrinsic variability between the batches. After several trials, the main finding for the linear mode was a 30cm double wall with 16 layers' elevations. As for the circular operation mode, more complexities were found due to the minor deviations of the machine pathway and thus those of the extruded material. The major finding was a double layered semi-circle with an average diameter of 30cm and a 6 layers-height. Finally, raising the level of complexity, the two modes where combined when printing a double layered self-standing arch with two 12cm footings, two 30cm columns and

a semicircle of 30cm diameter. Fig.5 illustrates some of the shapes printed by the machine.

Figure 5. Layered samples printed by the machine

IV. 3D CONCRETE PRINTING AND SUSTAINABILITY

Concrete 3D printing technology is assessed based on environmental, economic and social performances as compared to with traditional construction methods. The safety hazards on site, the material wastes as well as harmful emissions urge the immediate shift into more sustainable, eco-friendly and ameliorated technologies [6]. This highly precise automated process offers construction with reduced waste, pollution, noise and labor injuries on site. Thus, along with the increased level of human safety and environmental welfare, construction cost is reduced significantly. The percentage of CO_2 emission is mitigated by approximately 75% and the total energy by 50%.

V. SCALING UP

A starting point for scaling up this application is printing two level houses using precast frames. The simplicity of the structure and the low level of tensile stresses allow for the entire level to be printed as one unit. All structural and non-structural elements will be printed using only fiber reinforced concrete of 40MPa compressive strength. The method suggested is printing the entire level from a plan view, including walls with embedded window/door openings, columns, and other architectural features of the floor. The top part of the first level and the bottom part of the second shall include a set of printed notches that are mirrored in the intermediate slab to create a key and lock mechanism. Step two is printing the slab separately on the ground and lifting it using a crane to position it in the aforementioned interlocking mode.

The above system's weak structural integrity limits its scope significantly and could not fully replace traditional methods. The targeted plan includes a system of modular precast structural elements integrated with in-situ automated printing system. The structural elements including columns, beams, and slabs will be made of precast non-printed reinforced elements to resist fully lateral and vertical stresses. The remainder of the level will be printed in-situ using a hydraulically jacked printer (similar to the Grocon Jumpform system), including the walls, and other architectural features present. This method is highly favorable because the automation of the printing process offers speed and a precise estimate of completion time per activity. The knowledge of precise information allows for a more lean construction process. This is implemented by ordering the precast structural element on time and keeping small on-site buffers. Consequently, adopting a takt time process will decrease the variability of the schedule/labor force demand.

VI. CONCLUSION

The optimization of the CNC machine into a coherent and automated concrete 3D printing system requires merging several design, material and structural components. Starting from scratch, the mechanism of the machine and its parameters are modified in compliance with the printed concrete mix properties.

The extruded concrete from the nozzle is assessed based on the ease of extrusion, continuity of the flowing layer, the supported system and the stability and strength of the end structure.

As perceptible results, a coherent synchronized system is able to operate and print different geometrical shapes, building up several interconnected layers. Further testing and improvements will be carried out in order to achieve the remaining desired goals of improving buildability, finding more accurate delay time between batches of layers. Adopting a construction management approach will be useful when it comes to weighing the benefits of this innovation in terms of cost, time and environmental feasibility.

REFERENCES

[1] Koshnevis, B., Yeh, Z., Yao, K., & Hwang, D. (2006). Mega-scale fabrication by contour crafting.

[2] Blain, L. (2014, April 23). Chinese company uses 3D printing to build 10 houses in a day. *Gizmag*.

[3] Krassenstein, E. (2015, September 8). EXCLUSIVE: LewisGrand Hotel Erects World"s First 3D Printed Hotel, Plans to Print Thousands of Homes in the Philippines Next.

[4] Holloway, J. (2013, March 25). 6-meter tall KamerMaker to 3D print Amsterdam house by year's end. *Gizmag*.

[5] Rouhana, C., Aoun, M., Faek, F., Eljazzar, M., & Hamzeh, F. (2014). The Reduction of Construction Duration by Implementing Contour Crafting (3D printing).

[6] Crafting Technologies, Mansour Rahimi, Mahdi Arhami and Behrokh Khoshnevis "analyse the environmental impact of contour crafting technology as compared to that of concrete masonry units
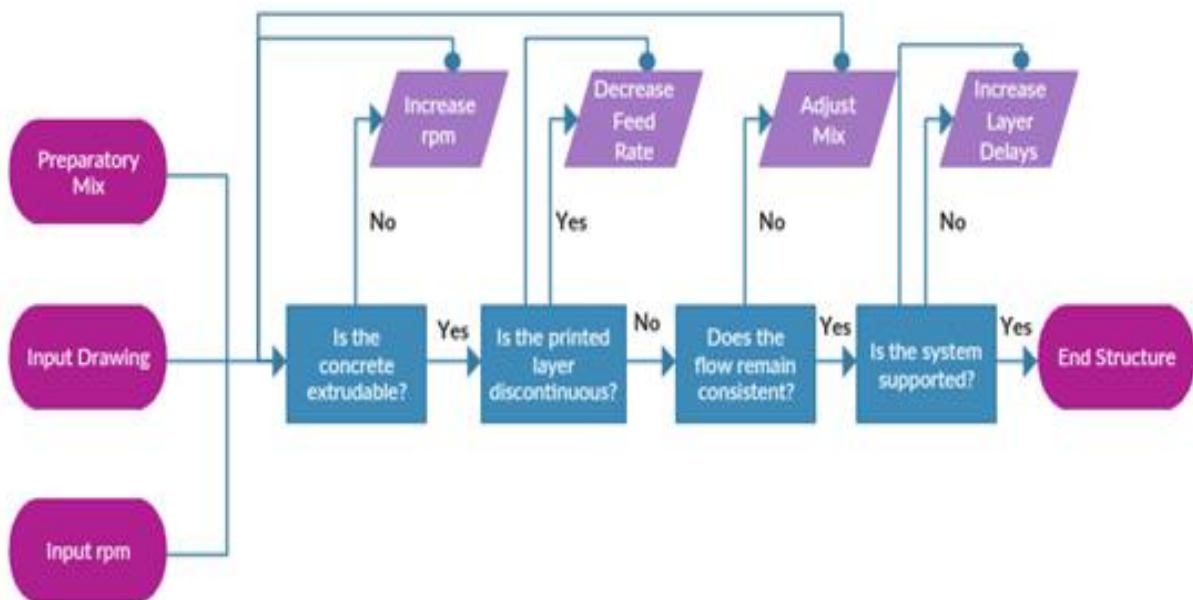
Appendix

Figure 6. The printing set up

Figure 7. Flow chart explaining the methodology to follow in assembly of the machine and the design of the optimum mix

Figure 8. Some of the architecture shapes printed by the machine

Figure 9. Open time test samples

Noise Mitigation in Urban Areas: AUBMC 2020 Case Study

Abstract- **This paper studies noise pollution in a local context and aims at finding feasible mitigation measures taking AUBMC area as a case study. After thoroughly researching the topic to better understand the causes and implications of noise, as well as suggested solutions to alleviate its severity, the group moved towards applying these findings to the study area. A professional sound meter was used to measure noise levels at different locations and at different times of the day, which allowed, after analysis, the preliminary identification of noise patterns and their causes. High levels of noise were noticed compared to the regulations of both the Ministry of the Environment (MOE) and World Health Organization (WHO). Solutions were developed tackling each problem. Noise barriers were tested using ArcGIS (Geographic Information System) software and found effective in alleviating the problem. Additional regulatory solutions related to construction and transportation were also suggested.**

Basma Hallak, Fawzi Khalifeh, Nabeeha Shokor, Reina Ezzeddine, Tamer Ladan

Department of Civil and Environmental Engineering

American University of Beirut

Beirut, Lebanon

bkh02@mail.aub.edu, fgk02@mail.aub.edu, nhs19@mail.aub.edu, rfe16@mail.aub.edu, tml02@mail.aub.edu

I. INTRODUCTION

Environmental issues facing the world nowadays often cause noise pollution to be marginalized. However, the perception of sound and the extent to which it varies in duration and level have considerable effects on human well-being. World Health Organization recognizes noise as a major factor in several physical and psychological problems, the most common of which is noise-induced hearing impairment, affecting perception of speech and communication between individuals, especially in the presence of background noise. WHO also states that noise levels of 65 dB(A) at Leq (24 hr) cause cardiovascular and hormonal disruptions including heart rate increase and changes in blood pressure. It is also stated that if the exposure to noise is temporary, the physiological system can over recover, which is not the case if exposure is consistent and intense. This project aims at assessing noise pollution around the area of AUBMC, where the ongoing AUBMC 2020 plans have been raising questions related to noise [1]. In a survey done both online and in print on 94 random people questioned about noise pollution in that area, over 88% were AUB/AUBMC students or employees, the majority of those passing around the AUBMC area almost daily. Over 90% of survey takers answered "yes" to a question asking about whether or not there is a noise problem in the area. For those survey takers, the number representing the severity of the problem on a scale from 1 to 5, 5 being the most severe was an average of 3.8/5. This shows that the problem at study is significant. The most recurrent answers to a question about the sources of noise were construction work and transportation, specifically vehicle horns. Frequent

suggestions for solving the problem were the following: using glass barriers, reducing site work on peak hour, re-routing, improving public transportation, using new transportation techniques. Noise annoyance prevalence in the area accentuates the need for the quantification of the problem and finding solutions.

II. METHODOLOGY

A. Noise Measurements

A noise meter (B&K 732A) along with an iphone app (NoiseTube) were used to record noise levels in dB(A), after calibration for consistency between devices and another calibration with white noise in the AUB mechanical engineering labs for accuracy. The reliability of the noise app was validated during a study between the AUB Neighborhood Initiative and Dr. Zaher Dawy from the AUB electrical engineering department [2]. Noise measurements were carried three times a day on a span of several days. Noon (11:00 to 13:00 hr), afternoon (15:00 to 18:00 hr), evening (19:00 to 21:00 hr), and midnight (23:00 to 00:00 hr) readings allowed the identification of noise levels at different times of the day to allow the identification of noise levels at peak hours. The times chosen were representative of rush hour noise levels, which proved the presence of a noise pollution problem as will be discussed later. Other than temporal variation, spatial variation was accounted for by representing the space surrounding AUBMC by taking measurements at different locations. These were at point A, on Abdelaziz Street near Casper and Gambini's, point B near the Childrens' Cancer Center, and point C at the emergency exit at the heart

of AUBMC. A dynamic measurement M (mobile) was taken also walking around the area directly surrounding AUBMC. Measurements were taken at ten minute durations. A total of 80 measurements was taken, 20 measurements for each location A, B, C, and M. Average values of noise levels at each location and at each time of the day were calculated and tabulated.

B.   Traffic Counts

Aiming at relating noise levels measured to traffic conditions, spot speed measurements and calculations of traffic flow at each of the three static points was performed. A specific trap length was considered on every road, and the time taken by the vehicle to cover the determined distance was measured and therefore the speed was calculated. Moreover, the number of vehicles passing by a fixed position at the three locations was recorded every 30 seconds 100 times. The flow and speed were measured on several days at different time intervals.

C.   GIS Modeling

GIS was used to map noise levels resulting from the ACC construction site and roads in the area as a base case, followed by introducing noise barriers shielding from both sources by 7 dB(A) at the barrier. A detailed GIS modeling methodology is described in title III. Theories and Assumptions

III.   THEORIES AND ASSUMPTIONS

A.   Noise Level Definitions

Throughout this project, the units of noise levels are represented in decibels (dB), which is the unit used to measure intensity of sound. dB(A) is a way of representing this intensity taking into consideration the range of frequencies at which the human ear is sensitive to sound. The "A" weighted curve is used for this purpose. Equivalent level (L_{eq}) is the continuous sound level that produces the same effect on the human ear as that of the actual noise during measurement. It's a way of simplifying representation since noise level are variable over time. L_{eq} can therefore substitute all these different values by a single number [3].

B.   GIS Assumptions and Detailed Methodology

i.   Sound- Distance Relationship

The equation used to find the sound intensity L_2 at any distance r_2 from the source, whether construction or transportation, is given below [4]:

$$L_2 = L_1 - \left| 20.\log\left(\frac{r_1}{r_2}\right) \right| \quad (1)$$

Fig. 1 illustrates the relation and the parameters involved in it.
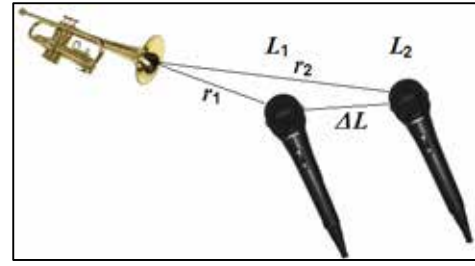
Figure 1. Sound intensity emission

Where:

r_1: Distance from object 1 to the source, object 1 being the location of the sound meter or measuring device. In this case, it is 3 m from vehicles and 7 m from the ACC construction site.

r_2: Distance from object 2 from the source, that is the distance between any point in the area under study and the sources of noise

L_1: Sound intensity at a certain distance r_1 from source in dB(A). For vehicles, it was assumed to be 70 dB(A) at a distance of 3 m while for the construction site, it was 75 dB(A) at 7 m.

L_2: Sound intensity at a certain distance r_2 from source in dB(A)

ii.   Effect of the Sound Barrier

A noise barrier is an obstacle implemented between a noise source and a receiver to control sound waves propagation and to limit noise levels. A typical improvement caused by these structures is to decrease noise levels by 5 to 10 dB(A) depending on the materials used and the design executed [5]. In this case, a green barrier was assumed to be decreasing the noise caused by the road network around AUBMC by 8 dB(A) on average at the source, so the noise reduction in the total area is a function of this 8dB(A) and the distance from the noise sources.

iii.   Effect of the Land Cover

The area around AUBMC is completely urban with a small area being green. Buildings in this area are mainly built from concrete so the reflection caused by such structures was found to be 65% [6], meaning that concrete absorbs 35% of the sound waves that hit it while reflecting back 65%.

This property of a certain material is defined as the Noise Reduction Coefficient (NRC) [7]. NRC is the percentage of sound that a surface absorbs. In the case of concrete, NRC is 0.35 as previously stated.

iv.   Noise at Sources

The two main sources of noise that were assumed are the road network in the region and the construction site.

Concerning the road network, a case where 70 dB(A) at source was considered due to the vehicles flow and this was in accordance with the measured sound levels [8]. Also, since the area is a small one respective to noise assessment studies

of cities, all the roads within it were assumed to have the same source noise level.

On the other hand, the ACC construction site was given a source noise level of 75 dB(A) which was taken from site measurements and literature review of typical noise levels in similar construction sites [9].

C. GIS Methodology

To assess the present problem of noise in the vicinity of AUBMC, a simplified model using the previously mentioned assumptions was implemented into ArcGIS.

The creation of the model involved sketching of the existing buildings in AUBMC in 2D as sound receivers and the road network in the area along with the ACC construction site as sound sources.

Following the sketching phase, the ArcGIS model builder (Fig. 2 and 3) was used to simulate how sound propagates with distance and how it is affected by the sound barrier and the land cover. Raster analysis was used throughout the process.

Figure 2. Model used to simulate noise propagation around AUBMC
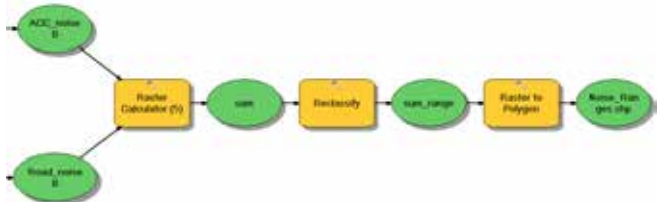
Figure 3. Model used to simulate noise propagation around AUBMC (Cont'd)

The first step in the model was to get the distance between each point in the region of AUBMC and the noise sources, which are the road network and the construction site. The Euclidean distance tool was used to do this and results were then used to set the sound-distance relation.

Values in the rasters were generated according to the relation, starting by the source noise levels which are 75 dB(A) and 70 dB(A) then decreasing as points are further away using the Raster Calculator tool.

The following step was to account for the land cover. Since the area is an urban area, a concrete structure was assumed to exist every 10 meters and the NRC was implemented into the formula using the Raster Calculator tool too.

Resulting values were averaged as to account for both noise emissions from transportation and construction sources.

Finally, the results were arranged into ranges using the Reclassify too in order to compare with allowable ranges.

Two cases were carried out in order to know how a barrier implementation along the road can affect noise levels in the region.

IV. DISCUSSION

A. Analysis

Averages for noise levels collected during experimentation are as follows:

TABLE I
AUBMC AREA MAXIMUM NOISE LEVELS

Location	Noise Level dB(A)			
	Noon	Afternoon	Evening	Midnight
A	69.4	73.3	64.3	53.5
B	67.9	67.8	66.4	51.8
C	67.9	73.0	67.5	54.1
Mobile	70.4	70.3	67.1	51.0

TABLE II
AUBMC AREA AVERAGE NOISE LEVELS

Location	Noise Level dB(A) $\mp ST.DEV$			
	Noon	Afternoon	Evening	Midnight
A	66.7 $\mp$ 1.79	67.9 $\mp$3.05	61.9 $\mp$2.14	49.2 $\mp$1.31
B	65.4 $\mp$1.75	63.5 $\mp$2.39	62.7 $\mp$2.93	47.5 $\mp$0.98
C	64.8 $\mp$1.95	65.7 $\mp$3.79	63.0 $\mp$3.90	48.6 $\mp$1.00
Mobile	66.7 $\mp$2.61	67.1 $\mp$1.73	62.8 $\mp$3.28	48.2 $\mp$1.24

TABLE III
LEBANESE AMBIENT NOISE LIMITS FOR INTENSITY IN DIFFERENT LAND ZONES [10]

Land Use	Noise Standard dB(A)	
	Day Time (7:00 a.m. to 6:00 p.m.)	Evening Time (6:00 p.m. to 10:00 p.m.)
Commercial, administrative, or downtown	55-65	50-60
Residential/commercial centers on highways	50-60	45-55
City Residential areas	45-55	40-50
Suburbs with light traffic	40-50	35-45
Country residential areas, hospitals, parks	35-45	30-40
Heavy industries	60-70	55-65

Level	Effect	Area
USEPA		
Leq(24)< 70 dB(A)	Hearing	All areas (at the ear)
Ldn< 55 dB(A)	Outdoor activity interference and annoyance	Outdoors in residential areas and farms and other outdoor areas where people spend widely varying amounts of time and other places in which quiet is a basis for use
Leq(24)< 55 dB(A)	Outdoor activity interference and annoyance	Outdoor areas where people spend limited amounts of time, such as school yards, play grounds, etc.
WHO		
Leq(24)= 55 dB(A)	Serious to moderate annoyance	Outdoor living area
Leq(24)= 70 dB(A)	Hearing impairment	Industrial, commercial shopping and traffic area, indoors and outdoors

The area under study can be considered as a city residential area and a hospital area. For ease of comparison, points A and B will be taken as residential and C as a hospital area. According to table III, noise levels should range between 45 and 55 dB(A), and those at C between 35 and 45 dB(A) during daytime, with values of 5 dB(A) lower during the evening. The difference between the required levels and those present clearly show a noise pollution problem. Table IV shows that present levels cause activity interference and annoyance and could cause problems to the hearing system.

The results show maximum levels of noise varying between 65 and 75 dB(A), with total noise levels averaging at values around 65 dB(A). Fig. 1 below better relates spatial and temporal variation of measurements.
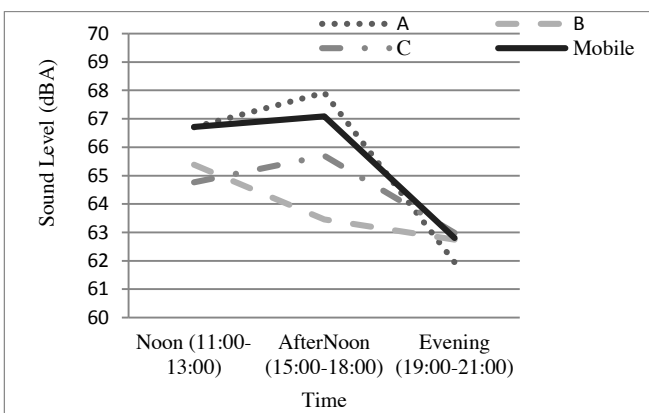
Figure 1. Variation of Average Noise Levels with Time and Place

The graph shows peak hours during the afternoon, with graduate decline until the evening.

Note that measurements during midnight were performed to obtain base values of noise to which noise levels at other times are compared. Levels at midnight are all around 48 dB(A). Points A and B maximum values exceed those of midnight by around 19 dB(A) and those of C by around 17 dB(A). Values also show that during the daytime, noise is most prominent at point A, B then C, while during the evening C becomes the most affected by noise and A the least. This could be due to the fact that during daytime, all areas are subject to construction and traffic noise, with point A the most affected by the AUBMC 2020 site and B by traffic on the road, especially that both points lie on two heavily used roads. Point C on the other hand is only affected by hospital activity and traffic. However, noise levels during the evening are present only due to hospital activities, where traffic and construction work are diminished. This shows that the two main constituents of additional noise are traffic and construction work.

As for traffic calculations, the table below shows values flow and velocity at points A, B, and C:

TABLE V
AVERAGE VEHICLE FLOW AND VELOCITY AT THREE STUDY POINTS

Location	A		B		C	
Parameter	Flow	Vel.	Flow	Vel.	Flow	Vel.
Noon	486	28	365	26	103	35
Afternoon	598	20	424	21	172	29
Evening	389	31	297	33	87	37

Flow in veh/hr and velocity in km/hr

All the results showed a peak flow of vehicles and a relatively low speed in the afternoon interval with respect to the other intervals which can be attributed to the fact that, at this time of the day, the traffic flow is concentrated because of employees daily leave time. This leads to the increase in the density, resulting in decrease of vehicle speed. Comparing the obtained results with the previously mentioned noise levels, it can be observed that as flow increases and velocity decreases, noise levels increase. It is important to note that noise due to traffic is not directly related to these factors through tire pavement-interaction, since this is usually applicable at highways and freeways. Rather it is due to the car horns and vehicle engines.

B. Suggestions
Since construction is the main contributor to high noise levels, regulations related to site work should be implemented.
Table IV presents conventional working hours at different days of the week:

TABLE VI
RECOMMENDED WORKING HOURS

Days	Hours
Monday to Friday	8:00 to 18:00
Saturday	8:00 to 13:00
Sunday and holidays	No work allowed

Failure to work within the given working hours should result in a warning followed by a written request for the closure of the site for a couple of days.

In addition to that, regulations must be followed within the site to reduce the nuisance resulting from machinery and equipment by regular (monthly) examination of the machines and trucks used, due to the fact that the more they are maintained the less nuisance they produce, in addition to implementation of noise reducing devices that act as barriers in equipment.

In addition, ArcGIS was used to test the implementation of barriers around the construction site and the roads.

In the first case, no barrier was implemented and the run gave values that are in accordance with values measured during experimentation, as shown in fig. 2 below.
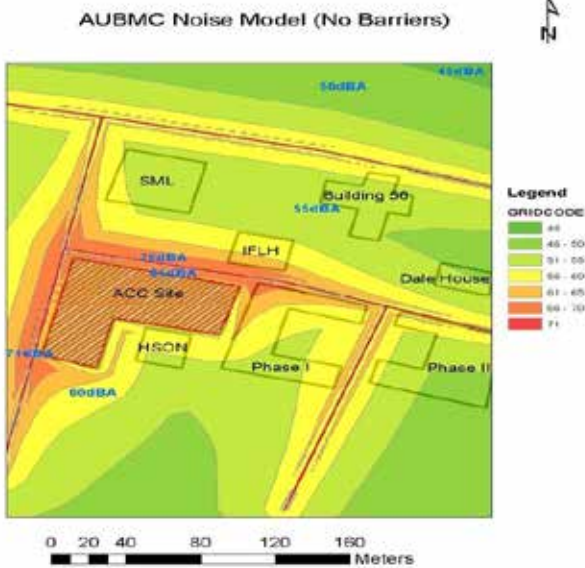
Figure 2. ArcGIS Output with no Barriers

In the second case, upon the implementation of the barriers, noise levels decreased to more acceptable ranges, as shown in fig.3
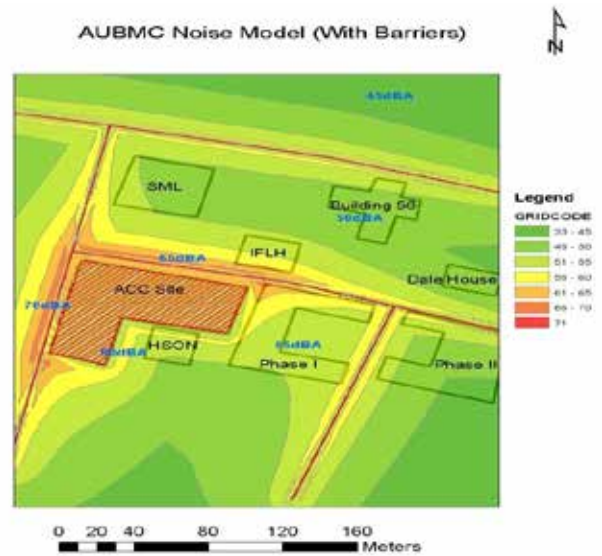
Figure 3. ArcGIS Output with Barriers

The figures show significant decrease in noise levels after the implementation of barriers, where acceptable ranges are achieved in comparison to those without implementation. The basis on which the characteristics of noise barriers were used and the extent to which they decrease noise was taken from general specifications of barriers.

Further studies are to be performed on the design and the implementation of these barriers, as in the materials used and the dimensions and specific locations.

To expand on the transportation sector, public transportation should be improved to pave the way for decreasing the usage of individual cars along with their associated negative implications. Regulations related to excessive use of vehicle horns should be enforced since the resulting sound is the highest contributor to noise from traffic.

For decreasing noise annoyance specifically at point C, where AUBMC employees and medicals complained the most about car horns, a logical suggestion would be pedestrianizing the road right next to the ER door, as shown in fig. 4 below (Road at point B, Cairo Street from Souraty to Maamari intersections). Vehicles heading towards Abdelaziz Street are forced to avoid this area, which could result in a better hospital environment especially that this road is at the main door of AUBMC. Patients cannot access the ER or the entrance on foot can be provided with a drop of area where they can be carried using a wheelchair or an ambulance car. This would definitely increase traffic flow at the entrance of Abdelaziz Street, but it will ultimately become the same as that before pedestrianizing the ER road, because the number of cars would still be the same, i.e. there is only dislocation of vehicles from the ER road to the entrance of Abdelaziz Street.
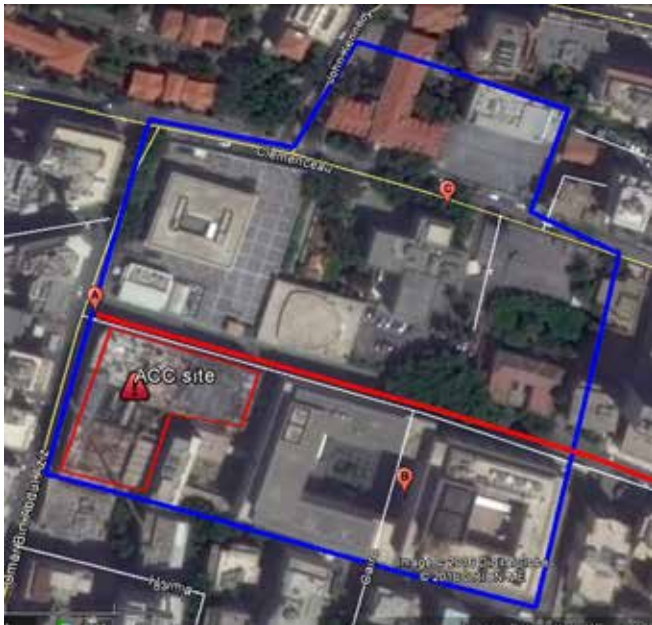
Figure 4. AUBMC Area Map

V. CONCLUSION

Work done enabled the quantification of noise levels around the AUBMC area and identification of their major sources. This was necessary to better understand the magnitude of problem. Noise levels are beyond the minimum levels recommended by health officials, which requires definitive action measures. The theoretical implementation of noise barriers reduced noise levels a between 5dB(A) and 10 dB(A) depending on the location of the point of study. Pedestrianizing the road next to the main entrance can drastically decrease annoyance in the area. Implementing these measures and others needs coordination between the construction and transportation fields as well as urban policy enforcement. To better serve the AUBMC 2020 area, future plans should take into consideration the problem on hand to ensure sustainability.

ACKNOWLEDGMENT

REFERENCES

[1] "Guidelines for Community Noise." Ed. Brigitta Berglund, Thomas Lindvall, and Dietrich Schwela. World Health Organization, 1999. Web. <http://www.who.int/docstore/peh/noise/guidelines2.html>.

[2] Dawy, Zaher, 2014, "Noise Measurement Techniques and Studies"

[3] Calixto, Alfredo, Fabiano B. Diniz, and Paulo H.t Zannin. "The Statistical Modeling of Road Traffic Noise in an Urban Setting." *Cities* 20.1 (2003): 23-29. Web.

[4] E. Sengpiel.(2014) "Damping of Sound Level vs. Distance" [Online]. Available: http://www.sengpielaudio.com

[5] "A Guide to the Reduction of Traffic Noise". N.p.: vicroads, 2003. Ser. 01224. p. 4.

[6] "Noise Control Design Guide". Owens Corning Acoustic Products. 2004.

[7] "Standard Test Method for Sound Absorption and Sound Absorption Coefficients by the Reverbation Room Method". ASTM Standard C423 – 09a. 2009.

[8] "A Guide to the Reduction of Traffic Noise". N.p.: vicroads, 2003. Ser. 01224. p. 2.

[9] "Construction Noise". S. Eaton. Workers' Compensation Board of BC. 2000.

[10] Lebanese MOE (Ministry of Environment): 1996, Resolution No. 1/52, Official Gazette Issue No. 45, 12/9/1996.

[11] USEPA (United States Environmental Protection Agency): 1974, *Information on Levels of Environmental Noise requisite to Protect Public Health and Welfare with an adequate Margin of Safety.*, EPA/ONAC 550/9-74-004.

[12] WHO (World Health Organization): 1999, *Guidelines for Community Noise*, WHO-expert task force meeting held in London, April, UK.

Progressive Structural Failures due to Elevated Temperatures

Bekdash Amin – El Fil Hala – Fakhreddine Mohamad – Kachouh Hussein – Lahoud Makram
Department of Civil and Environmental Engineering
American University of Beirut
P.O. Box 11-0236 / Riad El-Solh / Beirut 1107 2020 / Lebanon
aab41@mail.aub.edu – hke10@mail.aub.edu – mhf13@mail.aub.edu – hkk12@mail.aub.edu – mfl03@mail.aub.edu

Abstract- **With the increase of industrialization in Lebanon, the need for an efficient and safe work environment has become crucial. However, very little awareness is present concerning these issues from the social, engineering, and academics aspects. Major fire incidents in industrial buildings are occurring at an average pace in Lebanon. Incidents such as the Kesrouan cotton factory fire that occurred on the 15th of October 2015 due to an electric short circuit urged us to investigate the aftermath of the blaze. The structural and material damages seen, such as the slab deflection, crumbled concrete covers, and weakened concrete material were used as a base to our investigation. The purpose of our study is to examine the structural behavior of buildings subjected to fire loading, which can lead to progressive structural collapse of concrete due to elevated temperatures. The scope of the project was divided into three main parts. First, a parametric study was carried out to observe the influence of different design parameters of the structure's property subjected to ambient and elevated temperatures. Second, the induced thermal forces were obtained from the von Mises stresses by modeling the slab and beam under different scenarios using the finite element software, *Abaqus*. Third, an experiment was performed in the lab to simulate the fire event and record the behavior of concrete and its strength. The data was obtained from site visits and information was collected from the literature review. With this information, a SAFE model for the building was built and evaluated for slab deflections. Using the obtained results from the simulations and experiments, design guidelines were highlighted from the structural and materials perspective. Upon the research done, it was seen that there are fire codes in Lebanon, however, due to the large amount of corruption, the involved people do not abide by it. This study will be used as a basis for possible future studies that will encourage and contribute in implementing and abiding by the fire safety building codes in Lebanon.**

I. INTRODUCTION

Fire in structures has catastrophic effects due to elevated temperatures, which may cause structural failure leading to casualties. A building undergoes progressive collapse when one of its main elements fails due to fire. This is frequent in industrial buildings, where tons of combustible items are stored, fire can easily propagate across the building, making the job for firefighters very challenging.

In terms of design, very little attention is given to the detrimental effects of fire on the structural elements, whereas the main focus is on other hazardous incidents, such as earthquakes and wind. However, fires happen much more frequently than earthquakes and hence the design of anti-fire structures is crucial. According to UPSeis, 5.5 magnitude earthquakes and above on the Richter scale cause slight damages to buildings and other structures, with an average of 1400 of these earthquakes happening worldwide every year [3]. On the other hand, around 370,000 buildings catch fire per year in the U.S. alone, showing a large significant difference between the number of earthquakes and fires that occur annually [8]. Hence, due to the large number of fires that occur every year, fire-building codes should be implemented to prolong the fire resistance of the facility.

II. LITERATURE REVIEW

Looking at the history of fire induced progressive structural collapse, it is very rare that commercial tall buildings will face complete failure due to a normal accidental fire [7]. This is mainly due to bracing the frame against high thermal stresses, equipping the building with sprinkler systems, and designing the structure so that the heat generated will not surpass the material's threshold. Regardless of these precautions accounted for in the preliminary design of various structures, the Faculty of Architecture building at the Delft University of Technology, and Windsor Tower in Madrid, Spain, experienced mild or critical structural failure due to fire-induced events.

A fire triggered in the Faculty of Architecture Building at the Delft University of Technology, Netherlands, quickly escalated into a blaze that lasted for hours. The fire caused a major portion of the composite structure to collapse leading researchers to investigate the cause and aftermath of this occurrence. With the data collected, researchers could perform heat transfer simulations and thermal stress calculations of the fire event [5]. Upon aggregating the required data, researchers can understand the fundamental development of the time-temperature history. By investigating the thermally induced stresses developed via simulations, researchers were capable of analyzing the structure under fire conditions and comment on the spalling of the columns from the fire-induced incident. Hence, simulations via finite element software could be used to model fire events and account for induced forces from thermal stresses that will be investigated in the study using Abaqus.

The Windsor Tower in Madrid, Spain, was subjected to a fire due to an electric malfunction. This caused the entire composite structure to burn for more than twenty hours, resulting in the failure of the top portion of the exterior steel frame [6]. The building was structurally designed to

redistribute the load upon any failures, which explains why the entire frame did not collapse from the fire. Further investigation showed that the facility lacked fireproofing in the upper portion of the building because the following were still being installed [6]:

1) Fire protection of the perimeter steel columns using a boarding system

2) Fire protection of the interior steel beams using spray protection (epoxy coating)

3) Sprinkler system.

4) Aluminum cladding system to reduce heat conduction

The lack of fire proofing of the steel and concrete members, sprinkler systems, and fire compartmentation was the reason behind the rapid cause and spread of fire. Thus, one of the feasible solutions to increase buildings' fire resistance would be to add fire protection coatings such as epoxy coating, fire extinguishing, or retardant systems.

In an investigation conducted by Behnood and Ghandehari [2], high strength concrete samples with and without polypropylene (PP) fibers were heated to temperatures of 600°C to observe their compressive and splitting tensile strengths. Compressive and splitting tensile strengths of the samples with PP fibers were higher than the ones without PP fibers at high temperatures. However, concrete samples that contain PP fibers were more sensitive in tension than they are in compression at high temperatures [2]. Thus, the mechanical properties of concrete showed significant improvements at high temperatures with the addition of PP fibers.

III. ABAQUS MODEL

Concrete structures are known to resist fires, concrete is non-combustible and has a low thermal conductivity, it tends to remain in place during a fire, with the concrete cover protecting the reinforcing steel, and the cooler inner core continuing to carry the load. Calculations of the behavior of concrete structures in fire depends on many factors, the most being loads on the structure, the elevated temperatures in the concrete and the mechanical properties of the steel and concrete at those temperatures. When a concrete structure is exposed to fire, the temperatures of both steel and concrete increase, leading to increased deformation and possible failure, depending on the applied loads and the support conditions [4]. A section of the Daraya cotton factory building was modeled using Abaqus and a finite element analysis was performed to simulate the induced thermal stresses.

A. Material Properties

A linear static model was used. The ambient-temperature mechanical properties used are Poisson's ratio for elastic behavior equal to 0.2, Young's modulus (E) for elastic behavior equal to 23500 (kg/cm^2), the concrete compressive strength f_c' equal to 250 (kg/cm^2), and the expansion coefficient alpha equal to 6E-006 [9]. Those properties were varied with temperature starting from room temperature assumed to be 20°C to 600°C.

B. Loading and Boundary Conditions

The section understudy was subjected to a uniform loading of dead and live load equal to 0.0003 of magnitude, which represents the axial force imposed by the floors on the column that is holding them. The boundary conditions details of the finite element model are presented in Fig.1.

C. Concrete and Reinforcing Temperatures
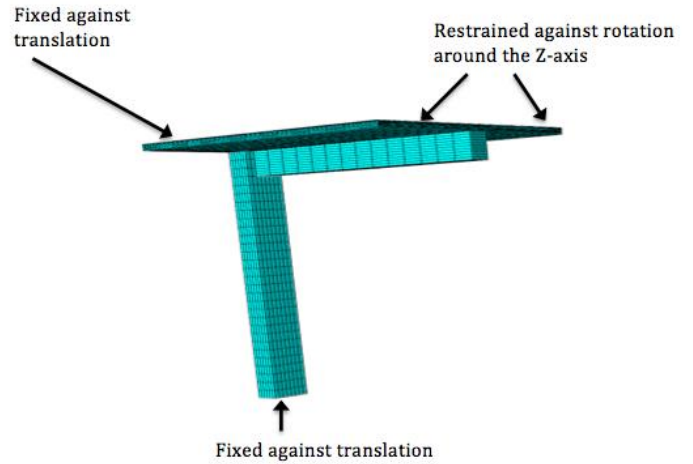
Figure 1. Boundary conditions.

In order to make the analysis at hand realistic, design charts that give thermal gradients in slabs could be used. For simplicity, it has been proven that for simple members of normal-weight concrete, empirical hand calculation means are available. These empirical methods are derived from computer-based thermal analysis [4].

Wickström's method for calculating temperatures in a concrete slab for a standard fire is based on fire exposed surface temperature T_w.

$$T_w = \eta_w T_f \qquad (1)$$
$$\text{Where, } \eta_w = 1 - 0.0616 t_h^{-0.88}$$

T_f is the fire temperature and t_h is the time in hours. At any depth x in meters into the slab, at time t_h, the concrete temperature T_c is a factor of η_x of the surface temperature T_w, η_x given by the following equation.

$$\eta_x = 0.18 \ln (t_h/x^2) - 0.81 \qquad (2)$$

The concrete temperature T_c is given by the following equation:

$$T_c = \eta_x \eta_w T_f \qquad (3)$$

These formulas were utilized to obtain a heat transfer analysis so that results would be more real.

The heat gradient analysis and the formulas listed above are only used for concrete slabs, there are other formulas for beams and columns (heat conduction would be in two directions), but they were not used in the analysis.

From the fact that the slab understudy spalled, we can deduce that the temperature at the cover which was 3.5 cm was about 250°C. From this temperature, the temperature at the surface of the slab (exposed to the fire) was back calculated; this is how the slab was divided into 10 parts each having its own temperature as shown in Fig. 2.
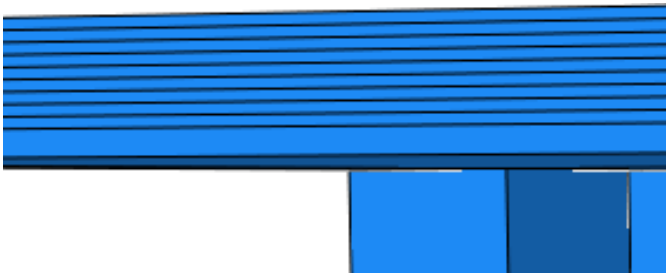
Figure 2. Slab compartments showing temperature variability.

D. Concrete and Reinforcing Temperatures

After conducting several runs on the model, it is worth noting that the cover spalled as show in Fig. 3.

Figure 3: Spalled slab.
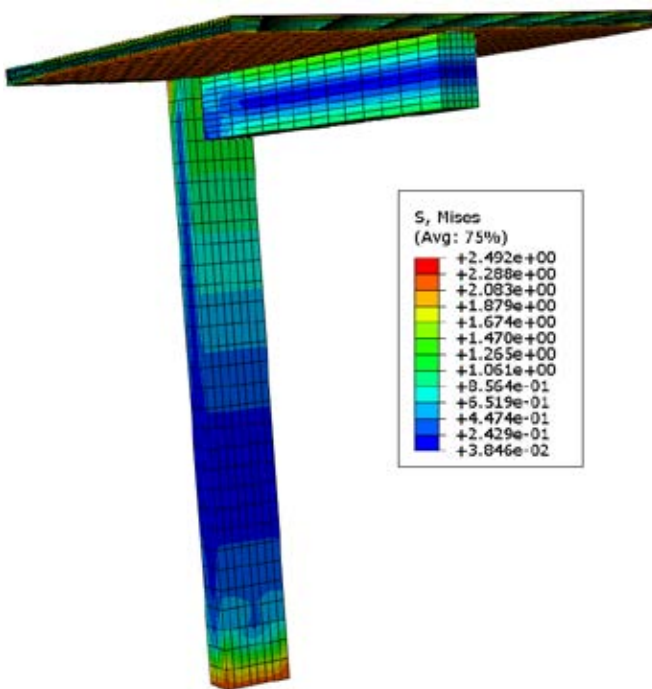
Figure 4: Von Mises showing the slab spalled.

Moreover, axial forces were calculated from the Von Mises stresses by multiplying the stresses by the area of intersection between the column, slab and beam (Area=25500 cm^2).
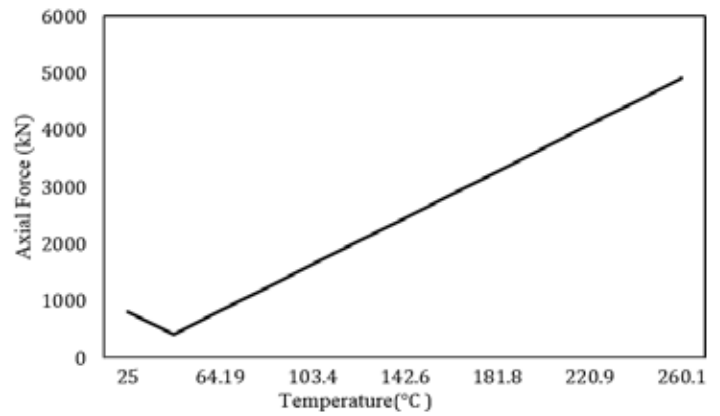
Figure 5. Axial force as a function of temperature.

The axial forces obtained from the Von Mises stresses increase linearly as temperature increases. This is reasonable because once a fire incident occurs, the stresses on a member increase and the stresses are directly proportional to axial forces.

IV. LABORATORY EXPERIMENT

This section includes the experimental procedure that was undertaken to test both ordinary and modified Portland cement concrete samples against increased ambient temperatures. The motive of the experiment is to obtain the values of the maximum compressive strength f_c' of these samples before and after being subjected to high temperatures. The contribution of polypropylene fibers admixture and intumescent epoxy coating to fire resistance is yet to be determined.

The set-up comprises of a high temperature oven chamber, a heating oven for removing moisture content, 7cmx7cmx7cm molds, Portland cement, course aggregates, fine aggregates, polypropylene fibers, intumescent epoxy coating, and water. A burning chamber was not available in the laboratory, instead a high temperature oven chamber was used to obtain the desired temperatures for testing the concrete samples. In order to test different kinds of concrete samples for comparison, the samples of concrete were categorized as follows:

1) Plain Portland cement concrete

2) Portland cement concrete containing polypropylene fibers

3) Portland cement concrete with intumescent epoxy coating

4) Portland cement concrete containing polypropylene fibers and intumescent epoxy coating

The concrete design mix for the experiment without any admixture was designed to obtain an ideal concrete compression strength of 290 kg/cm^2. Having a cube volume of 343 cm^3, the following proportions were used per concrete cube for the concrete design mix as presented in table I.

TABLE I
CONCRETE DESIGN MIX WEIGHT PROPORTIONS

Material	Weight per cube
Portland Cement	105 g
Coarse Aggregates	326 g
Fine Aggregates	320 g
Water	81.9 g

It must be noted that the coarse and fine aggregates had to be oven dried at 105°C overnight in the heating oven to completely remove the moisture content from them. Moreover, the original water to cement ratio intended for use was equal to 0.53. However, additional water had to be added, yielding a water over cement ratio of 0.78, to take into account the high absorption rate and empty voids of the oven dry aggregates. After preparing the concrete design mix, two quantities that yield 7 concrete cube samples each were prepared, one batch containing polypropylene fibers admixture and one without.

To obtain the quantity of polypropylene fibers needed for the batch with fibers, 2.2% by weight of the Portland cement were added to the mixture, yielding to around 2.4 grams of fibers needed for the entire batch [1].

After the two batches were mixed, the mixtures were poured into the cube molds to create the samples. Once the concrete samples were dry, the cubes were cured for 28 days to obtain 99% strength of the concrete, and the intumescent epoxy coating was applied on the samples that needed coating.

The samples are now ready to be tested under different elevated temperatures. In order to compare the samples with one another, each batch was tested with its own control sample, both with and without fibers at its room temperature of 22°C. As for the others, the samples were heated at the desired temperatures of 420°C, which was the temperature attained in the case study, 550°C, and 670°C. Note that each sample was heated for one hour after the surface temperature of it reached the desired temperature. After the heating was completed, the samples were allowed to cool to their normal room temperature before applying the compressive strength test. The results of the compressive strength test along with the type of sample and temperatures applied are summarized in the following table II.

TABLE II
RESULTS OF CONCRETE SAMPLES AT ELEVATED TEMPERATURES

Sample	Temperature (°C)	Area (cm^2)	Maximum Weight (kg)	Compressive Strength (kg/cm^2)
Control (Plain)	22	49.7	14500	291.7
Plain (Non Coated)	420	51.48	8000	155.4
	550	50.77	4000	78.8
	670	50.77	1500	29.5
Plain (Coated)	420	50.77	8500	167.4
	550	51.55	6000	116.4
	670	51.48	2000	38.9
Control (Fibers)	22	50.98	14500	284.4
Fibers (Non Coated)	420	50.98	8000	156.9
	550	51.12	4800	93.9
	670	51.12	2500	48.9
Fibers (Coated)	420	51.47	8300	161.3
	550	52.20	5500	105.4
	670	51.33	2700	52.6

The results obtained show significant improvement in the compressive strength with the addition of the coating. The coated samples retained 40% of their original strength at 550°C, whereas the non-coated samples retained only 27% of the original f_c'. The coating added a resistant shield against the elevated temperatures, protecting the concrete material from heat. It can be inferred from the results above that the effect of coating diminishes as the temperature reaches high values, the difference between f_c' of the coated and non-coated samples at 670°C is small.

The fibers added to the mix help in reducing micro cracks in the concrete and in preventing heat to seep through it. The samples with fibers added showed no improvement in strength at ambient temperature, yet at 670°C the samples retained 17% of their original strength while the samples without fibers reached 10% of their f_c' at similar temperatures. This value is also higher than that of the coated samples without fibers. Hence, it can be inferred that the fibers reduce cracks that help concrete to resist elevated temperatures even without the use of coating. It is also important to note that the samples with fibers showed minimal spalling, even at 670°C, while the samples without fibers had many cracks and spalled severally.

The results show equal compressive strength values between the coated samples with and without fibers. Yet, at high temperatures the samples with fibers showed higher values of compressive strength. Although the effect of coating is small at very high temperatures, the fibers in the samples still were able to minimize cracks and spalling.

This shows the importance of using similar products in construction. They help the concrete to retain higher values of their compressive strength, increasing the structure's resistance to temperature and heat for longer periods. The coatings provide an additional layer that would decrease the heat flow into the concrete. In addition, the fibers reduce the cracks in the concrete and do not let the heat seep into the material. Thus, under high temperatures such products are necessary for the integrity of any structure, helping the concrete to resist deterioration and spalling for longer periods to allow for a longer evacuation time during fire events.

Figure 6. Concrete samples in the high temperature oven furnace.

Figure 7. Plain concrete and fibers concrete samples after the compression strength test, from left to right.

V. PARAMETRIC STUDY

A parametric study was carried out to observe the influence of different design parameters of the structure's property subjected to ambient and elevated temperatures. The parameters varied were span length, concrete compressive strength, slab thickness, live load, and column size. The column capacity was obtained using the following equation acquired from ACI-building codes.

$$Capacity = \emptyset \times .065 \times (f_c' \times 0.85 \times A_g + \rho \times f_y) \quad (4)$$

Values for f_c' at elevated temperatures were obtained from previous research carried out by Buchanan [4]. Slab deflections were checked through SAFE models. Column capacities were analyzed at ambient and elevated temperatures and compared to the demand (service loads). The parametric study includes an iterative process that continues until the demand upon certain elements overcomes the capacity and a failure occurs. Results showed that a typical column will fail at a temperature exceeding 500°C regardless of parameter changes such as span length or member dimensions. At 500°C the corresponding f_c' is 60% of the original compressive strength. After this decline in strength, the structure would fail under service load.

VI. SOCIAL ASPECT

Fire safety practices aim to prevent the upsurge of fire and limit a fire occurrence after it starts. Following the building code that is usually complemented by the fire code does this practice. In this paper, we will focus on fire safety measures that are planned during construction of a building and not those taught to occupants of a building. To be eligible in obtaining a permit that states that the built structure is safe in a fire occurrence, the client should go through two stages:

1) Check whether the designs and specifications comply with the codes present in the country.

2) During construction, the structure should not fail any inspection conducted by authorities hired by the Civil Defense.

General fire precaution techniques needed will vary from one site to another and also during the construction process, leading structural engineers to consider the following:

i) Escape routes and fire exits
ii) Compartmentation
iii) Fire detection / firefighting equipment
iv) Emergency signs

In the Kingdom of Saudi Arabia, civil defense institutions assign a third party company to inspect the fire safety legislation of a construction site. The latter ensures that drawings and specifications are according to code before the start of construction. During construction, a person from the civil defense is assigned to evaluate if the fire safety measures in the specifications are implemented on site. At the end of the construction phase, the building is either considered as safe and can start operating normally, or is not given a "can start operation" permit until it is in accordance with the code.

With fires in industrial buildings increasing in the past years, it is evident that fire codes are not being well implemented in Lebanon. According to a local business owner, Mr. Karout, "nothing here requires you to follow any regulations", adding that since he opened in 1989, there was not any inspection made to ensure that the structure is fire-safe. The only way a building is considered as safe from fire is through hiring a private insurance company that will do the job. On paper, it is mandatory for owners to pass an inspection annually, however, in reality it is not the case.

The decisions taken by clients can have a great influence on the health and safety of a project. Abiding by the fire code can prolong the life of the structure and decrease human fatalities drastically. Depending on the severity of the fire, the client is obliged to fully reconstruct or rehabilitate the structure, largely increasing the costs of maintenance.

VII. CONCLUSION

From the obtained results in our study, it can be indicated that industrial and large-scale buildings should use high strength concrete so that the columns can withstand elevated temperatures. From the Abaqus model, it has been shown that it is preferable to have symmetrical industrial buildings for stability and integrity. The reason behind this is that the axial forces obtained from the Von Mises stresses are relatively small due to the symmetrical plan of the building at hand. It is desirable to have reasonably large sizing of members and large concrete covers to withstand the stresses resulted from the elevated temperatures, increasing the factor of safety of the building. Concerning the experiment held, it was concluded that fire coating adds a longer fire resistance to the material. The fibers on the other hand, prevent spalling, protect the steel reinforcement, and increase integrity. With structural and construction modifications to the building, reliable fire systems, efficient evacuation plans, and raising awareness through education, we can prevent building damages, reparation costs, and most importantly, casualties.

REFERENCES

[1] Anthony Nkem Ede, & AbimbolaOluwabambi Ige. (2014). Optimal Polypropylene Fiber Content for Improved Compressive and Flexural Strength of Concrete \n. *IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE)*, *11*(3), 129–135. Retrieved from http://www.iosrjournals.org/iosr-jmce/papers/vol11-issue3/Version-4/S01134129135.pdf

[2] Behnood, A., & Ghandehari, M. (2009). Comparison of compressive and splitting tensile strength of high-strength concrete with and without polypropylene fibers heated to high temperatures. *Fire Safety Journal*, *44*(8), 1015–1022. http://doi.org/10.1016/j.firesaf.2009.07.001

[3] Earthquake.usgs.gov (2015). Earthquake Facts and Statistics. Retrieved 4 December 2015, from http://earthquake.usgs.gov/earthquakes/eqarchives/year/eqstats.php

[4] H. Buchanan, A. (2001). *Structural Design for Fire Safety.* Wiley.

[5] Meacham, B., Park, H., & Straalen, I. J. (2008). Fire and Collapse , Faculty of Architecture Building , Delft University of Technology : Data Collection and Preliminary Analyses. *Architecture*, 14.

[6] Parker, D. (2005). Madrid tower designer blames missing fire Protection for collapse. *New Civil Engineer.*

[7] Scott, D., Lane, B., & Gibbons, C. (2003). Fire induced progressive collapse. *Prevention of Progressive Collapse: Report on the July 2002 National Workshop on Recommendations for Future Efforts*, 17. Retrieved from http://www.buildingsmartalliance.org/client/assets/files/mmc/Scott's revised paper.pdf

[8] U.S. Fire Statistics. (n.d.). Retrieved December 3, 2015, from https://www.usfa.fema.gov/data/statistics/

[9] Wong, Y.-L. (2005). Spalling of concrete cover of fiber reinforced polymer reinforced concrete under thermal loads. *Materials and Structures*, 0–0. http://doi.org/10.1617/14306

Re-Design of the Holiday Inn Hotel

Elie Mansour, Elio Ibrahim, Ghadi Barhouche, Karim Adaimi, Samia Abdallah
Department of Civil and Environmental Engineering
American University of Beirut
P.O.Box 11-0236 Riad El-Solh / Beirut 1107 2020
ejm04@mail.aub.edu, eji02@mail.aub.edu, gab04@mail.aub.edu, kma46@mail.aub.edu, sna43@mail.aub.edu

Abstract- **This paper integrates a design of The Holiday Inn hotel in Ain el Mreissi. It incorporates a structural design of the hotel, geotechnical study, parking design, transportation analysis, and construction management study along with an environmental concern. The structural design takes into account the design of the columns, slabs and beams of a typical floor, and the design of the staircases, the basement walls and the shear walls. It was done through either hand-on calculation samples or software design of each element followed by the modeling of the building using ETABS to analyze the structure under static and seismic loads. Moreover, the team will present a design alternative with pre-stressed concrete slabs that will be compared with the traditional approach. An analysis of geotechnical studies of nearby sites was done to get the needed soil characteristics for the foundation design. A parking design was carried out with the aim of satisfying the hotel's future demand and a traffic analysis was done to assess the impact that the Hotel will have on its surroundings. The team took also into consideration the use of recycled materials to reduce negative impacts on the environment, and tried to integrate in design and construction as many LEED requirements as possible. Last but not least, the group submitted a bill of quantities and costs and a schedule for all the activities needed to construct this Hotel. Also, this paper mentions the reason the group opted for the redesign approach rather than renovation.**

I. INTRODUCTION

The Beirut Holiday Inn Hotel, which opened its doors two years before the civil war broke out in 1975, is not only a 24-floor shattered building but also a ghost of separation between east and west. As the civil war divided the city, the two main rivals, the Lebanese Front representing Christian rightwing militias backed by the Lebanese army and the National Movement representing Lebanese leftist parties backed by Palestine's PLO, competed to dominate the area. "Seen by militants as a strategic military asset, the Holiday Inn became a trophy in the battle" (Reference [1]). This iconic building located between the coastal neighborhoods of Ain el-Mreisi and Mina al-Hosn standing next to the Phoenicia Hotel and facing the St George Yacht Club was one of the most luxurious buildings that attracted rich visitors from Middle East and Europe. People used to come enjoy its rooftop restaurant rotating at the top. Nowadays, in order to access the building and walk through a skeleton of concrete, one needs a permit from the army. Standing at the bottom hardly able to capture the colossal size of the building in one glance, a person is confronted with a cadaverous, dented shell of history. This hotel holds a lot of memories. Plenty of people want them erased, and for them, the only way to accomplish that, is by demolishing the building. On the other hand, because of these memories, the good and the bad, other people want it to stay. They want to renovate it and turn it into the striking hotel it once was.

The primary focus of the project was to study the feasibility of demolishing and redesigning, or renovating the hotel then deciding on the better option from a cost, social and environmental perspective. However accessing the site early on turned out to be a major constraint and the permits for conducting the necessary tests needed for renovation, were refused. Given these circumstances, along with other main reasons mentioned later in this paper, the team decided to redesign the building from scratch. Fortunately the members were able to put their hands on limited architectural drawings, which were used to redesign the structure, and were able to access the site lately.

Consequently, the goal of this project is to relate the knowledge acquired from five different disciplines in civil engineering in order to use the abstract architectural drawings and come up with a sound and feasible design, all the while taking into account the new ACI code requirements, as well as the new standards for a five-star hotel.

II. LITERATURE REVIEW

A. Potential Impacts due to Demolition

Building materials account for a big portion of waste materials generated around the world. Rules and regulations are being enforced as much as possible to limit the environmental impacts associated with any construction or demolition project.

The impacts associated to demolition in this case can be classified in three main categories that are:

1. **Physical Impacts**
2. **Socio-economical Impacts**
3. **Transportation and Disposal Waste Impacts**

The various types of pollution (Air, Noise, Dust, Land and Water), demolition waste, vicinity of the pedestrians, vibration disturbance and impacts resulting from the parking and vehicle movement form the Physical Impacts. Socio-Economical impacts are related to labor safety and inconvenience to drivers and pedestrians near the site. Last but not least, the impacts associated with the transportation and disposal of waste material are dust, noise and road safety.

B. LEED requirements

In order to make the building as eco-friendly as possible, the team researched about the LEED requirements. Since the building is going to be built according to existing architectural drawings, not all of the requirements could be implemented. The requirements that can be applied to the hotel along with the points earned are:

Sensitive Land protection (1 point):
Leed Requirement: Locate the development footprint on land that has been previously developed.
Result: The Holiday Inn is designed in a land that has already been developed.

High priority site (2 points):
Leed Requirement: Locate the project on an infill location in a historic district.
Result: The holiday Inn is located in the historic district of Beirut, where its land has already been developed before.

Surrounding Density and Diverse uses (5 points):
Intent: To conserve green land and wildlife habitat by encouraging development in areas with existing infrastructure.
Leed Requirement: Locate on a site whose surrounding existing density within a ¼-mile (400-meter) radius of the project boundary meets the values in Table 1.

Table 1 Points for average density within 1/4 miles of project

Combined density	Separate residential and nonresidential densities		Points BD&C (except Core and Shell)	Points BD&C (Core and Shell)
Square feet per acre of buildable land	Residential density (DU/acre)	Nonresidential density (FAR)		
22,000	7	0.5	2	2
35,000	12	0.8	3	4

Result: The Holiday Inn is located in Downtown Beirut, the 85th densest city in the world (2800 hab/km^2).

Green Vehicles (1point):
Leed Requirement: Designate 5% of all parking spaces used by the project as preferred parking for green vehicles and enforce for sole use by green vehicles. Discounted parking rate of at least 20% for green vehicles is an acceptable substitute for preferred parking spaces. The discount must be publicly posted at the entrance of the parking. Preferred parking areas and signage for carpool and vanpool vehicles and green vehicles may be combined if 10% of total parking capacity is reserved with this signage and both Reduced Parking Footprint and Green Vehicles credits are achieved.
Result: the group provided 111 parking spots restricted to preferred parking and reduced its parking rate by 50%, in addition to a private van minimize the use of staff parking spaces (elaborated in the parking study section).

C. Recycled and Reused materials

According to Reference [2], in a developed country like Belgium, the amount of construction and destruction (C&D) waste is around 230 tons/km2 per year from which 86% is either recycled or reused. Such precise data is not available in Lebanon, but an approximate number can be obtained by using a C&D waste average density of 1.5t/m3 and by assuming that every 100 m^2 of built up area generates 50 m^3 of waste

According to the Order of Engineers, the total built up area in Lebanon for the year 2012 was 8,484,500 m^2. So, using the previous estimation method, the amount of C&D waste generated would be:

$$C\&D \; Waste = 8484500 \; / \; 100 * 50 * 1.5 = 6,363,375 \; tons$$

Of which zero % is Recycled or Reused.

The use of recycled concrete aggregates is expanding all over the world with infrastructural laws by developed countries in order to promote the usage of these materials more prominently. According to Professor Issam Srour, about 6.1 million m^3 of demolition waste was generated in the July and Naher El Bared wars only, all disposed in illegal landfills. Srour noted that approximately 0.9 million tons of construction waste were generated in the Beirut area between 2009 and 2010, where more than 85% of the total amount by mass was formed of concrete, hourdis and masonry. These three cement-based materials can be reused or recycled.

Unfortunately, there is yet no legislative law in the LBC (Lebanese Building Code) that permits the use of recycled concrete aggregates (RCA) in the construction industry. Hence, a modification of the Order of Engineer's Building Code has to be issued in order to authorize the use of RCA in precise percent rates. Many previous international publications proposed different method statements to determine new mixing combinations of concrete using recycled materials. It is up to the Lebanese authorities to base themselves on a relevant study that reflects the same condition in this country, or to develop a new research on the topic in order to legalize it.

One relevant publication is Ref. [2]. First, a sieve analysis was carried for both crushed concrete and natural aggregates and the authors were able to find a resemblance in the particle sizes of the crushed concrete aggregates and that of natural aggregates.

Reference [3] also shows that both Specific Gravity and Water Absorption are satisfactory results. As for the average compressive strength, a decrease of 8-16% was noted when RCA were used. The compressive strength is a function of demolished concrete grade, replacement ratio of recycled aggregates, and the water cement ratio. Thus the Study concludes that RCA could be used for up to 30% in Portland cement concrete without affecting the functionality of the structure created.

III. METHODOLOGY

A. Site Visits

Pictures of damaged concrete columns and damaged shear walls were taken and pinpointed for reference. The group visually investigated the steel to check for potential corrosion or weakness. In addition, the team inspected the current structural elements of the hotel, including size of the columns, thickness and type of slabs, clear height of typical floors and underground floors, shear walls and staircases to check their compliance with the five-star hotel requirements. Moreover,

the concrete strength was measured using a Schmidt hammer borrowed from the AUB lab.

B. Structural Design

The team started the design based on the architectural drawings at hand. These drawings were dimensionless, and thus required the use of Google Maps in order to scale the building and acquire the proper dimensions. The typical floor plan is shown in figure 12 in the appendix

Using the ACI code 318-11, and inferring from the market about the type of materials used in the construction industry nowadays, the group prepared several designs alternatives and chose the most suitable one. The design of a typical floor, public floor, and underground basement entails the dimensions of slabs, columns, beams, shear walls, and staircases.

The assumptions taken for the dead and live loads are based on Uniform Building Code:

For a hotel typical floor and a parking floor:
   L.L=200 kg/m^2
   S.D.L=500 kg/m^2
For a hotel public floor:
   L.L=500 kg/m^2
   S.D.L=500 kg/m^2
For stairs:
   L.L=500 kg/m^2
   S.D.L=170 kg/m^2

The group then ran simulations on ETABS, Ref. [4], in order to see how the structural elements react to different types of loads. Lateral loads were also integrated to make sure that the building can sustain seismic activities. The properties that provided the basis of the seismic parameters were extracted from the ACI code to account for lateral loads.

Factors:
According to Table 12.2-1 of Reference [5], using **F** Shear wall-frame interactive system with ordinary reinforced concrete moment frames and ordinary reinforced concrete Shear walls:
Response Modification: R = 4.5
System Over strength: Omega = 2.5
Deflection Amplification: Cd = 4
Occupancy Importance: I = 1 (Since the building represents a low hazard to human life in the event of failure, including agricultural, temporary, and minor storage facilities.)

Seismic Coefficients:
According to Reference [6]: Ss = 1.5 (for Moderate to high seismicity zone-rest of the Lebanese territory)
S1 = 0.45 (for Moderate to high seismicity zone-rest of the Lebanese territory)
Tl = 8
Site Class:
According to section 20.3.4 of Ref. [5], since the rock is limestone, the site is classified as B.

C. Geotechnical Study and Foundation Design

A proper soil investigation is mandatory in order to check the soil condition especially that the site is also close to the sea.

The fact that the building exists; a soil investigation cannot be achieved. Thus, the team relied on information of a nearby site, Reference [7], and assumed the same soil properties. These properties provided the basis of the seismic parameters, which were extracted from Ref. [8] to account for lateral loads. In order to decide on the appropriate foundation type and design, the group will refer to the allowable bearing pressure, amount of settlements, and the loads imposed by the structure.

Based on the Geotechnical investigation Reference [7], information obtained from field works and laboratory testing indicated that the subsurface stratum is uniform consisting of two main strata as follows:

- A top layer of Fill Material consisting of light brown to creamy well graded gravelly sand; this layer extends from the ground surface level to a maximum depth of 60cm and extending throughout the entire site.
- Underlying the Fill Material is limestone bedrock. While the intact rock cores are moderately strong to strong, the rock mass is closely to extremely fractured and moderately to highly weathered.

For the very poor quality rock mass, the allowable bearing capacity of bedrock stratum can be estimated using the Hoek-Brown Criterion which takes into consideration the lithology, compressive strength, degree of fracturing, weathering state and groundwater condition of the bearing bedrock.

The equation is of the form:

$$Qa = \frac{[c \times s^{0.5} \times \sigma] \times \{1 + [m \times s^{0.5} + 1]^{0.5}\}}{F}$$

Qa: Allowable bearing pressure
c: foundation factor
s &m: lithology & fracturing factors
σ: uni-axial compressive strength
F: factor of Safety

The maximum expected settlements are going to be estimated using the formula of elastic linear theory for relatively homogeneous layers. The equation has the following form:

$$S = cqB \times \left(\frac{1 - v^2}{E_m}\right)$$

S: settlement
c: Foundation shape factor
q: Bearing pressure
B: Foundation width
v: Poisson's ration
E_m: Deformation modulus

D. Parking Design

In this part, the team will analyze the parking demand according to the newest Lebanese law criteria for parking and garages based on the number of rooms, restaurants and exhibitions areas. The first step is getting the total number of parking spots required. The team referred to the Lebanese parking lot criteria to infer the number of car parking needed

given the number of rooms, plus the number of cars to be provided for recreational areas.

Moreover, the team allocated spaces for the handicapped by referring to the American with Disabilities Act.

In addition to the aforementioned, parking spots were allotted to special cases like loading trucks, and electronic cars.

Lastly, the members conducted a parking management strategy whereby they assumed parking fees according to the type of users

E. Traffic Analysis

Intersection studies are important to determine streets where traffic signals are needed, study the quality of traffic services and control the intersection movements by setting adequate traffic signal timing to ensure good traffic flow.

The traffic analysis was conducted on the intersection leading to the hotel at peak traffic time (5pm). The Holiday Inn is located on the main road that leads from Downtown-Kontari to Ain El Mreissi, which is one of the most congested lines in Beirut. Any activity across the road will definitely lead to an increase in traffic. In order to reduce the traffic impact, a detailed study on traffic flows and directions during the day (during peak hours) took place on the surrounding road network in order to come up with the least harming solution.

This experiment consisted of analyzing the intersection delays and pedestrian crossing time on this particular intersection. On one hand, the vehicular delay by counting the number of stopped vehicles is obtained in each approach during 15-second intervals and multiplying it by the length of the interval. On the other hand, the adequacy of the actual pedestrian crossing time is checked by referring to its theoretical value.

The time a vehicle is stopped, i.e. waiting in a queue on an approach to the intersection, represents the stopped delay. In order to achieve this task, the team was divided on each approach to count the number of vehicles stopped in intervals of 15 seconds. One has to note that some vehicles remained stationary for more than one interval, and were thus counted on the following interval as well. The team kept on following this methodology for a total of 15 minutes, where each minute was divided into four-15 seconds intervals

The level of service (LOS) of the approach can be determined using the calculated average stopped delay per vehicle.

As for the pedestrian part of this study, using the crosswalk widths of the intersection and the effective crosswalks being smaller than 3.05 meters, the team checked the adequacy of the measured pedestrian crossing time using the following equation:

$$G_P = 3.2 + \frac{L}{S_P} + 0.27 N_{PED} \qquad (1)$$

F. Bill of Quantities

Last but not least, the group did a quantity takeoff to come up with the Bill of Quantities (BOQ). The way this was done was through hand-on quantifying of items and materials on AutoCAD. The team measured room and corridor floor areas for tiling; wall areas for plastering, tiling, painting and waterproofing; wall perimeters for masonry blocks; column and slab volumes for concrete pouring.

The BOQ was then used to price activities based on unit material and labor cost. Furthermore the BOQ was also used to develop a schedule for a typical floor based on Lebanese market productivities. The productivity rates were collected from several local contractors. One has to note that unlimited number of resources was assumed in the resource allocation process.

IV. RESULTS

A. Site Visit

After accessing the site, the team concluded that the hotel was not heavily damaged structurally. Some columns and shear walls were slightly damaged by shells, but most importantly, the team noticed that the hotel does not meet the current standards of a five stars hotel. The clear height of typical floors is 2.85m without tiling, so there is not enough space to install proper mechanical materials. The underground parking clear height is equal to 2m, which makes it impossible to install proper mechanical materials for ventilation. Also, the passage of trucks, vans, or high 4x4s is impossible. All the staircases in the hotel are very narrow and circular (1m radius), which makes them very uncomfortable. It is also clear that the existing design didn't account for fire escape. Finally, the design study, (explained in part B-7) has shown that the existing shear walls design did not account for seismic loads.

The team noted that the exposed steel in some cracked columns and walls showed severe signs of corrosion. The extent of the damage is shown in figure 6,7, and 8 in the appendix. From the second visit, the team acquired an idea about the compressive strength of concrete due to non-destructive testing. The test was performed on various columns, slabs and beams from different floors. The obtained strength values revealed a high variability ranging from 8 MPa to 59 MPa.

B. Structural design

The following material properties were used for the design:
For steel: fy = 4200 kg/cm2
For concrete: f'c = 35MPa and $\rho \ balance = 0.0361$
Concrete for the columns: f'c = 40MPa.

1) Ribbed Slab for typical floor:
Span Length = 6.5m
Ultimate weight = Wu = $1.55 T/m^2$ = 0.88T/m/rib
Maximum ultimate moment demand = Mu = 3.8 T.m
Slab thickness= ln/21 = 6.5/21 = 30cm to prevent deflection based on ACI code 9.5.2.1 table.
We get: b = 15cm, d = 27cm and h = 30cm with
$\rho \ critical = 0.27 \ \rho \ balance$
Longitudinal reinforcement: 2T16 for bottom reinforcement and 2T14 for upper reinforcement.
Shear reinforcement: $2\emptyset 6mm$ at s = 13cm up to L/5 = 1.3m from support, rest at s = 20cm.

2) Beams:
Continuous beam with non-equal spans

Longest span length = 8.5m and shortest span = 5.8m
Ultimate weight = Wu = 11.6T/lm
The team drew the continuous beam on SAP2000 to get the ultimate moments.
Maximum ultimate moment demand = Mu = 74.2 T.m
$\rho\ desired = 0.55\ \rho\ balance = 0.01986$
Beam dimensions obtained: b = 120cm, d = 31cm and h = 35cm
Longitudinal reinforcement: 13T32 for bottom reinforcement and 12T32, or 7T25, or 6T20 for upper reinforcement depending on the changing moment.
Shear reinforcement: 4T10 at s = 13cm up to L/5 = 1.3m from support, rest at s = 20cm.

3) Two-way solid slab:
The team members are learning how to design a two-way slab on SAFE. The design will be completed soon.

4) Columns:
Using $\rho = 2\%$ and using hand-on calculations, the tributary areas of each of the 2 types of columns (interior and exterior columns) were calculated.
The ultimate load demand on each column was then calculated by adding the loads from 4 parking floors, 7 public floors and 24 typical floors. The columns sections were calculated; the column self weight was then calculated and added to the total load to get the final column section.
Tables 6 and 7 in the appendix summarize the different column sections and their respective steel reinforcement requirements.

5) Staircase:
The team researched the ACI design procedure for a two-flight stair spanning longitudinally and implemented it.
The design resulted in 11 steps per flight with risers of 16cm, goings of 30cm, and landing of 1.5m width.
The steel reinforcement for the staircase is as follows:
9T14 at 15cm longitudinally and T10 bars at 10cm intervals transversally.
The following drawing shows the staircase plan along with its steel reinforcement.

Figure 1 Staircase drawing

6) Basement wall
The team members are learning how to design a basement wall. The design will be completed soon.

7) Shear walls
The shear walls were analyzed on ETABS as per the old dimensions measured from the architectural plans of the hotel. The walls thickness was equal to 60cm. After running the analysis, it appeared that the current shear walls could not resist the lateral loads and that seismic design was not taken into account when the hotel was designed. So the shear walls design had to be adjusted in order to resist seismic loads.
After consulting with the advisor, the team decided to add 4 additional T-section shear walls of thickness equal to 30cm at the place of respective partition walls.
Also, the original shear walls thickness was decreased to 50cm on the 13th floor and above to get a less conservative design.
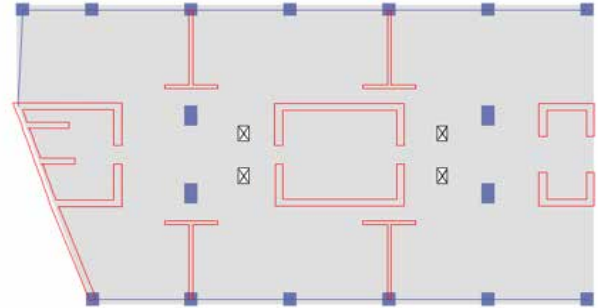
Figure 2 Original and added shear walls on ETABS

C. *Parking Study and Design*

According to the parking decree for hotels in Lebanon shown in figure 9 of the appendix, 1 car must be provided for every four rooms. In addition, one should account for the cars allocated for restaurants or recreational areas in the hotel and thus providing 50% of the number of cars allocated for these recreational areas in the decree.

The Hotel accommodates 41 rooms (37 singles, 1 double, 3 suites) per floor in the first 12 floors in addition to 22 rooms (1 single, 20 double, 1 suite) for the rest of the building. The surface area of one floor of the hotel is about 2100 m^2. Restaurants are present on the top two floors and four exhibitions and entertainment floors are present between the parking basements and the hotel's lobby. Each exhibition floor has a net area of 10200 m^2. As mentioned in the decree, one parking spot is needed for every 30 m^2.

Calculation of parking spaces:
- Parking spots for rooms = $\frac{\#\ of\ Rooms}{4} = \frac{41 \times 12 + 22 \times 13}{4}$ = 194.5. 195 parking spots are needed for the rooms.
- Parking spots for restaurants and entertainment areas = $0.5 \times (\frac{2100 \times 2}{30} + \frac{4\ x\ 10200}{30}) = 750$

750 parking spots are needed for the exhibition and entertainment areas.
- Total Parking spots = $195 + 750 = 945$ parking spots

As such, a total of 945 car parking spots are needed for the Holiday inn building.

Parking Design:
The parking is designed to accommodate a total of 1110 cars distributed into 4 floors.
The parking designs are found in the figures below. The first image shows the parking design on the first basement floor,

and the 2nd shows the three other parking floors on the second, third and fourth basements since they have the same design.

Figure 3 First basement floor parking

Figure 4 Second basement floor parking

The stall dimensions, turning radii and isles distance are all designed according to the Lebanese building code.

The Lebanese law does not provide minimum spaces for the handicapped. According to ADA (American with Disabilities Act) requirement parking criteria, it is required to provide 2% of the total parking spaces to the disabled people.

As such, Handicapped parking spaces $= 0.02 \times 1110 = 22.2$

23 parking spots for handicapped are needed.

The 23 handicapped parking spots are placed next to the elevators as seen in the design figures in blue.

The remaining available parking spaces can serve two other users: the hotel and the Credit Libanais staff, since both buildings share the same parking.

D. Traffic Impact Analysis

The following AutoCAD drawing is a representation of the intersection geometry along with its directional movements.
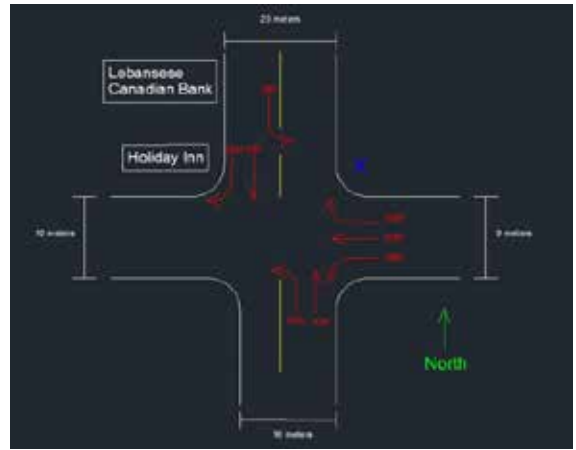
Figure 5 Intersection Geometry with directional movements

The data obtained for each of the Northbound, Westbound and Southbound approaches are summarized in the appendix. The table below however summarizes the average delays and percentage of vehicles stopped based on the recorded data.

Table 2 Average stopped delay per vehicle for each approach

	NB	*SB*	*WB*
Total number of stopped vehicles	82	67	469
Estimate of total stopped delay (veh.sec)	1230	1005	7035
Number of vehicles departing approach	131	88	148
Average stopped delay/vehicle (sec/veh)	9.39	11.42	47.53
Number of non-stopped vehicles	49	32	251
% of vehicles stopping	62.6	63.6	46.5

LOS criteria for signalized intersections are as follows:

- For the North Bound approach, the LOS is A as the average stopped delay per vehicle is equal to 9.39 sec/vehicle which is less than 10 sec/vehicles.
- For the South Bound approach, the LOS is B since the average stopped delay per vehicle is equal to 11.42 sec/vehicle, which is between 10 and 20 sec/vehicles.
- The LOS for the West Bound approach is D, having an average stopped delay per vehicle equal to 47.53 sec/vehicle, which is between 35 and 55 sec/vehicles.

An LOS of A for the northbound approach and B for the southbound approach are considered good. One can conclude that no key modifications or improvements are needed for those approaches with regards to renovating the Holiday Inn. However, the westbound approach yielded a LOS level of D, implying major improvements to be undertaken. As far as the demolition and reconstruction of the Holiday Inn is considered, it will surely increase the traffic in the surrounding areas which will in turn lead to the need of improving this intersection. This undertaking requires adding more green time to the westbound approach for the cars to pass. Another solution would be to add more green time on the southbound approach of the subsequent intersection since the majority of the westbound traffic goes left. Those suggestions will ultimately clear and facilitate the

route of the vehicles, which will in turn pave the way for the improvement of the LOS for this approach.

As for the pedestrians, the following data was obtained:

Table 3 Green time measured for the different crosswalks

Crosswalk	Width (m)	Measured green time for Pedestrians (sec)
WB	9	24.12
WB	10	19.63
SB	20	29.92
NB	16	28.12

Using equation (1), the table below summarizes the minimum green time needed for 15 pedestrians to cross the different crosswalks.

Table 4 Theoretical green time needed

	WB-9m	WB-10m	SB-20m	NB-16m
Green time	16.25	17.25	27.25	23.25

By comparing the actual green time available for pedestrians and the calculated minimum green time needed by pedestrians to cross, it is found that there is more than enough time for 15 pedestrian to cross at each green light at all the approaches.

It is important to notice that there is no heavy pedestrian activity on this intersection at the moment. But the traffic will surely increase considerably during major construction activities and after the reconstruction of the hotel, and may become a crucial criterion for the design of this specific intersection, resulting in changes in green time pedestrian-allocation for each approach.

E. Bill of Quantities and Schedule

The total cost for civil related tasks was estimated to be 499,199$ per typical floor. The hotel consists of 24 typical floors. The total cost of civil related tasks would then be $24 \times 499,199\$ = 11,980,776\$$ for the hotel without the lobby, restaurants, basement and exposition floors.

To get an estimate of the complete project cost, the cost of civil related works should be added to the costs of the geotechnical studies, mechanical and electrical works, furniture and interior decoration, contractor's overhead, contingency and profit…

It would take 115 days roughly per floor to complete the tasks listed in the BOQ. The estimated lag time between floors is equal to 30 days since the activities are overlapping. So the total construction time estimated for the hotel's typical floors is equal to $24 \times 30 + 115 = 835\ days = 2.3\ years$.

The full bill of quantities along with their prices and the schedule and sequence of activities of the project can be found in figures 10 and 11 respectively in the appendix.

V.   FURTHER DEVELOPMENT

The team's original plan for the final year project was to assess whether to demolish the hotel or renovate it. However, many problems and delays were encountered in order to access the site and perform the necessary tests that help decide on the best alternative. So the team decided to re-design the hotel from the original architectural maps that they managed to draw.

The team was given permission for site access in the middle of the Spring semester, and they were able to assess the different areas were the current structure fails to meet the new 5 stars hotels standards, and they took them in account for the new design.

As for the further works, a foundation design is going to be presented in the final report along with the design of a two-way solid slab. Also, a pre-stressed concrete slab design alternative will be completed for both the typical and public floors of the hotel. Two pre-stressed concrete slabs alternatives will be studied: Precast ribs and precast hollow core slabs with pre-stressed beams. Then, a thorough comparison between the traditional design approach and the two pre-stressed concrete design approaches will be achieved, comparing the advantages and disadvantages of each (cost, time of construction, fire safety, possible need for drop beams, different slab thicknesses, etc.) The best or preferred alternative will then be chosen and will be implemented in the hotel.

VI.   ACKNOWLEDGMENT

VII.   REFERENCES

[1] ACI Committee 318. (2011). Building code requirements for structural concrete: (ACI 318-11). Farmington Hills, MI: American Concrete Institute.

[2] ASCE. 2010. *Minimum Design Loads for Buildings and Other Structures.* ASCE/SEI Standard 7-10.

[3] Berkeley, C. S. I. "Computer Program ETABS Ultimate 2015." *Computers and Structures Inc., Berkeley, California* (2013).

[4] Haddad Engineering. "LOT 526 – EIN AL MRAYSI GEOTECHNICAL INVESTIGATION" (2004)

[5] Huijer, Carla, Mohamed Harajli, and Salah Sadek. "Re-evaluation and updating of the seismic hazard of Lebanon." *Journal of Seismology* 20.1 (2016): 233-250.

[6] Monier, V., et al. "Service contract on management of construction and demolition waste." Report EU Commission (2011)

[7] Nayel. "Beirut's bullet-riddled Holiday Inn". *a history of 50 cities in 50 buildings, day 28*. The Guardian (2015)

[8] Nilson, Arthur. *Design of concrete structures*. No. 12th Edition. 1997.

[9] Sonawane, Mr Tushar R., and Sunil S. Pimplikar. "Use of recycled aggregate in concrete." *International Journal of Engineering Research and Technology*. Vol. 2. No. 1 (January-2013). ESRSA Publications, 2013

Table 5 Areas of use of RCA

Recycled Material	Use	Areas of Application
Crushed Concrete/Brick	As Aggregates	Concrete roads and aprons
		Drainage Works
		Pipes and Culverts
		Sewage/Water Treatment Plants
		Permeable Backing to Earth Retaining Structures
		Building Partition Walls
		Floors and Foundation
Crushed Concrete/Brick	As Unbound Base Course	Runways, Taxiway and Aprons
		Parking Lots and Yards
Crushed Concrete/ Brick or Recycled Aggregate	As Fill Material	Cable Trenches

Table 6 Interior Column

	Basement	1st Floor	13th Floor
Interior Column section (cm²)	100 x 165	100 x 120	100 x 65
Steel Section (cm²)	330	240	130
Steel Bars	42T32	30T32	17T32

Table 7 Exterior Column

	Basement	1st Floor	13th Floor
Exterior Column section (cm²)	100 x 110	100 x 80	100 x 40
Steel Section (cm²)	220	160	80
Steel Bars	28T32	20T32	10T32

Figure 6 Steel reinforcements exposed

Figure 7 Damaged column

Figure 8 Damaged shear wall with corroded reinforcements

Figure 9 Lebanese Parking Decree

	Description of Work/Items	Unit	Quantity	Unit Cost	Total Cost	Productivity/Crew/Day	Number of Crews	Duration
1	**Concrete Works**	m3			$ 184,687.50		5	6
	Slab	m3	606.87	$250.00	$ 151,717.50	25		
	Columns	m3	131.88	$250.00	$ 32,970.00	25		
2	**Block Work**	m2			$ 23,301.50		5	15
	15 cm Blocks	m2	1217.00	$11.50	$ 13,995.50	30		
	10 cm Blocks	m2	1034.00	$9.00	$ 9,306.00	30		
3	**Plaster**	m2			$ 65,505.00		7	28
	Room Wall Plaster	m2	2433.00	$11.00	$ 26,763.00	30		
	Room Ceiling Plaster	m2	763.00	$11.00	$ 8,393.00	30		
	Bathroom Wall Plaster	m2	1225.00	$11.00	$ 13,475.00	30		
	Bathroom Ceiling Plaster	m2	213.00	$11.00	$ 2,343.00	30		
	Common Areas Wall Plaster	m2	460.00	$11.00	$ 5,060.00	30		
	Common Areas Ceiling Plaster	m2	861.00	$11.00	$ 9,471.00	30		
4	**Paint**	m2			$ 57,060.00		3	25
	Room Wall Paint	m2	2433.00	$15.00	$ 36,495.00	50		
	Room Ceiling Paint	m2	763.00	$15.00	$ 11,445.00	50		
	Bathroom Ceiling Paint	m2	213.00	$15.00	$ 3,195.00	50		
	Common Area Wall Paint	m2	395.00	$15.00	$ 5,925.00	50		
5	**False Ceiling Common Areas**	m2	862.00	$25.00	$ 21,550.00	15	4	14
6	**Waterproofing Bathroom**	m2	214.00	$15.00	$ 3,210.00	20	3	4
7	**Tiling**	m2			$ 91,025.00		10	12
	Room Floor Tiling	m2	763.00	$25.00	$ 19,075.00	25		
	Bathroom Floor Tiling	m2	214.00	$35.00	$ 7,490.00	25		
	Bathroom Wall Tiling	m2	1226.00	$35.00	$ 42,910.00	25		
	Common Areas Floor Tiling	m2	862.00	$25.00	$ 21,550.00	25		
8	**Sanitary Fixtures**	-			$ 32,660.00		2	10
	Bathtubs	-	40.00	$500.00	$ 20,000.00			
	Lavabo	-	40.00	$150.00	$ 6,000.00			
	Pipelines	m	1110.00	$6.00	$ 6,660.00			
9	**Doors and Windows**	-			$ 20,200.00		2	10
	Room Doors	-	40.00	$200.00	$ 8,000.00	6		
	Bathroom Doors	-	40.00	$130.00	$ 5,200.00	6		
	Windows	-	40.00	$175.00	$ 7,000.00	6		
					$ 499,199.00			125

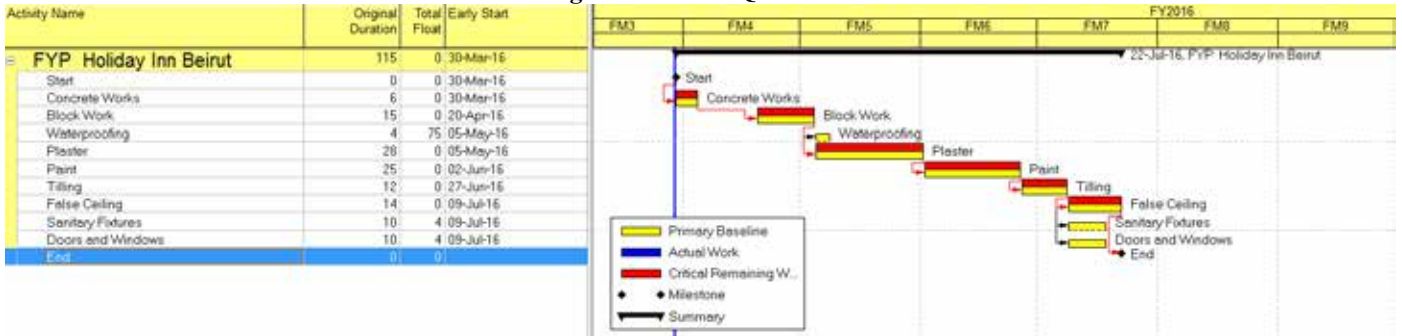**Figure 10 Bill of Quantities**

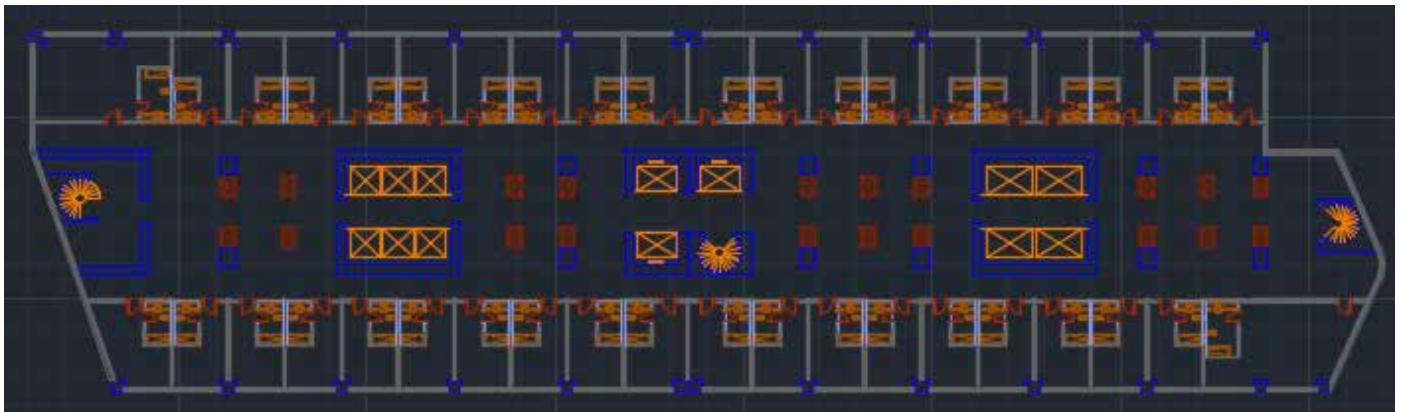**Figure 11 Construction Schedule**

Figure 12 Typical floor plan

Rehabilitation of Taxiway Alpha at Rafic Hariri International Airport

Amani Allan, Riad Bou Ghannam, Ghali El Samad, Maher Harmouche, Ali Nasser El Dine

Department of Civil and Environmental Engineering

American University of Beirut

P.O.Box 11-0236, Riad El-Solh / Beirut 1107 2020, Lebanon

ama155@mail.aub.edu, rsb24@mail.aub.edu, ghe02@mail.aub.edu, mgh24@mail.aub.edu, ahn11@mail.aub.edu

Abstract- **The current condition of taxiway alpha raises serious safety concerns that call for rehabilitation. The taxiway has served previously as a runway and was converted into a taxiway after the construction of runway 03-21. This paper presents several scenarios in design and construction to tackle the rehabilitation of the existing taxiway with minimal interference with the airport operations, while abiding by the Federal Aviation Administration (FAA) recommendations and standards. A site investigation analysis of the taxiway's current situation has led to the conclusion of two main scenarios which are: the complete replacement of the existing taxiway or overlaying the existing concrete layer. Both would provide a serviceability level which extends to more than 20 years. Design models and construction schedules were created using several software that were executed to simulate different scenarios in order to determine the optimal solution to present the destruction and reconstruction phases of the project. All the construction works are timed to comply with the airport operations and the FAA standards in terms of quantifying and controlling the movement of construction equipment and material. The result of this project will include several design alternatives, with their corresponding bill of quantities (BOQ) and construction schedules, which will be analyzed based on a life cycle cost analysis (LCCA) and a sustainability analysis to study the cost, feasibility and environmental impact for each scenario. The alternatives were compared and the asphalt scenarios were chosen due to their prevailing optimal feasibility.**

I. INTRODUCTION

The Rafic Hariri International Airport was first open to air operations in 1954. The airport suffered major damage at its terminal and other facilities throughout the wars it has witnessed. Being the gateway to Lebanon, there were efforts from the government to reconstruct the damaged areas of the airport. Renovation of the two primary runways was implemented between 1982 and 1984. After the civil war, a master plan was integrated in 1991 to improve the airport to accommodate 16 million passengers per year by 2035 [1]. The primary stage included the construction plan of the new east runway. A new runway designated 03-21 was constructed parallel to the old east runway, which was converted into the alpha taxiway. The current condition of the taxiway parallel to runway 03-21 is raising safety concerns. According to sources at the airport, the taxiway was originally a Portland Cement Concrete (PCC) runway and was overlain with an asphalt layer 25 years ago. Hence, the prevalent concerning condition of the taxiway will be analyzed and the decision to rehabilitate or reconstruct the taxiway will be considered relying on the results obtained in reference to its performance. A new alternative pavement will be designed based on the FAA standards, taking into consideration expected future increases in air traffic.

Figure 1. Airport layout – taxiway alpha highlighted in red

II. LITERATURE REVIEW

In airfield pavement, the critical loading is usually perceived to take place at the touchdown zone of the runway. However, the impact is relatively a minor one. The critical case occurs when the aircraft is travelling at slow speed with maximum fuel load. The highest load concentration usually occurs on taxiways, where lateral position variation of the aircraft around the taxiway's centerline is negligible [2]. The design should ensure limited deformation where the slopes of the pavement's cross-section must be 1-2% to avoid water ponding on the taxiway. Enough frictional force between the pavement and the wheels of the aircraft must be provided to ensure safety especially during harsh weather conditions. Therefore, airfield pavements have to be textured to provide enough skid resistance. Skid resistance is not an issue in concrete pavements since the texture is achieved by grooving or any similar technique.

Airports can't afford shut downs or reductions in capacity. Accordingly, essential parts of the airport such as runways and taxiways must be designed taking into account future increase in demand in order to reduce maintenance, which might hinder the airport's operations. While designing a pavement, there are some underlying engineering principles that must be kept in mind to execute a sustainable design. One vital aspect is subgrade protection. The existing ground is not strong enough to withstand the subjected traffic load which might cause the soil to deform and rut. Hence, the pavement must be thick enough to distribute the subjected load imposed by the wheels of the aircraft over a wide area that is sufficient enough to prevent subgrade deformation. Moreover, the materials constituting the layers in the pavement structure must provide

ample stability and resistance to stresses that are being transmitted at each level. The strength of the pavement layers should be adequate enough to prevent cracking under imposed traffic loads. The selection of material is also of paramount importance due to the pressing necessity for the pavement to withstand the local climate conditions. In addition, the design must guarantee that maintenance can be executed within a reasonable cost and without disrupting the traffic operations for too long.

Another prominent aspect in design is traffic data. Traffic loading is the primary input that governs the design of a pavement. An airport taxiway that accommodates large aircrafts requires a different design from a highway due to the different magnitudes of loads subjected on each segment of the pavement. The load magnitude, contact pressure, load groups, and loading speed are all aspects that must be taken into consideration in order to properly evaluate the stresses imposed on the pavement.

The contact area between a tire and the surface of the pavement is not easy to represent due to the textured surfaces of both the tire and the pavement. This reduces the area into a series of discontinuous areas. For simplification, pressure is assumed to be uniform and the contact area is assumed to be a circle. Also, wheels are sometimes located close enough to cause a combined effect from both tires on a certain location on the pavement. For the dual, dual-tandem and dual-tridem wheel arrangements of aircrafts, the increased load from the combined effect of the wheels can breach the critical design load capacity and cause serious permanent damage to the pavement. Thus, the wheel arrangement has to be determined for the purpose of identifying the critical stresses imposed on the pavement.

III. Evaluation of Current Taxiway Pavement

A site visit to the airport was done in order to determine the severity extent of the cracks that are present in the pavement. After examining the taxiway's pavement, it was noticeable that the pavement suffered from several types of cracks including:

1. Block cracks: connected cracks that come together to form blocks. It is caused due to the shrinkage and expansion of asphalt due to temperature oscillations. It occurs in response to the aging of the asphalt binder or poor choice of asphalt binder in the mix design.

2. Fatigue cracks: fatigue is the weakening of the material due to the implication of repetitive loading. When the loads exceed a certain threshold, cracks will start to propagate to the surface [3]. The cracks will evolve parallel to the direction of traffic.

3. Crocodile cracks: the connecting of fatigue cracks with one another which resembles the pattern present on the skin of a crocodile. Crocodile cracking occurs when the pavement fails due to the implication of repetitive loads. If the subbase becomes weak, cracks will propagate to the surface. A crocodile crack is considered to be a severe type of fatigue crack.

4. Reflective cracks: reflective cracks give an overview about the condition of the underlying concrete layer. The cracks can signify that the concrete layer is deteriorated and has a significant amount of cracks. The cracks present in the asphalt might also be the result of cracks propagating from the joints of the concrete slab to the surface of the pavement. To accurately determine the condition of the concrete layer, nondestructive testing should be implemented.

5. Edge cracks: is the formation of cracks at the edge of the road due to insufficient support at the edges, weak shoulders and improper drainage of water. The accumulation of water at the edges can weaken the subbase and damage the pavement.

6. Potholes: in this case, are due to moisture damage that occur because of water penetration through the cracks. This phenomenon results in the stripping potential of the aggregates-binder interface causing them to form potholes.

7. Other problems: Usually, the pavement's slope must be between 1-2% to ensure proper drainage of water and to avoid ponding, which is undesirable on airfield pavements. Due to repetitive loading, the slopes might have decreased to an extent where drainage is not effective anymore.

Table 4, in the appendix, shows images of the different problems found in the airport's pavement.

IV. Scenarios

After visiting the site, it was obvious that the asphalt layer has immensely deteriorated and needs to be removed. It was impossible to derive conclusions about the underlying concrete layer due to the limitations of testing imposed by the airport authority. Therefore, different scenarios were constructed and analyzed to compensate for the uncertainties.

A. Complete Removal

Complete removal of the concrete slab is recommended when the pavement has structurally failed. For sustainability reasons, the removed slab is usually rubblized or partially rubblized depending on the required pavement elevation. Part of it is included in the new layer as recycled concrete.

B. Rubblization

Rubblization involves the crushing of the PCC pavement into a compatible flexible base having an appropriate modulus. This tight angular interlocked material provides a stable base that is not susceptible to moisture damage. Concrete pavement rubblization should be considered when one or more of the following conditions are met [4].

- More than 20% of the concrete joints are in need of repair.
- More than 20% of the concrete surface has been patched.
- More than 20% of the slabs exhibit slab breakup.

C. Bonded Overlays

Bonded concrete overlays consist of a thin concrete layer bonded to the top of an existing concrete surface. This method is used for pavements in good condition with little deterioration in order to increase their structural capacity. The bond causes the overlay and existing pavement to act as one unit [5].

D. Asphalt Overlays

This rehabilitation technique is used when the need for fast track construction is essential. However, asphalt overlays can't be placed over damaged pavement because cracks might reflect from the damaged layers to the asphalt overlay [6].

V. DESIGN METHODOLOGY

In this project, the airport pavement design was based on the guidelines and recommendations of the Federal Aviation Administration's Advisory Circular 150/5320-6E, which recommends the use of the FAARFIELD software to study the stress created by the expected aircraft traffic mix.

A. Aircraft Traffic Mix

Previous advisory circulars published by the FAA used to convert all aircraft departures into one design aircraft. The design aircraft was selected by choosing the most critical aircraft with respect to load and number of departures. Currently, the FAA assesses the cumulative damage done by all types of aircraft. The Advisory Circular recommends considering the predicted annual departures for each aircraft to determine the thickness of the pavement. The FAA only considers departures for design and ignores annual arrivals because of fuel consumption which reduces the planes' load. Thus, the group members tried to contact the airport in order to obtain such data. Unfortunately, the airport claimed that traffic mix data is not collected. Therefore, the data was acquired from www.flightradar24.com. Since the website only displays data for each day, the data was collected on daily basis for a couple of weeks and then extrapolated for 1 year while taking into consideration growth patterns of certain aircrafts and addition of new aircrafts like the A380. The design traffic mix is summarized in table 1.

Note: A380 assumed to start using the airport after 10 years since it does not meet the required standards at the moment.

TABLE 1. AIR TRAFFIC MIX

Aircraft	Annual Projection
Airbus A310	190
Airbus A319	1075
Airbus A320	17250
Airbus A321	2660
Airbus A330	3470
Airbus A340	710
Airbus A380	100
Boeing 737	4810
Boeing 747	265
Boeing 757	380
Boeing 767	460
Boeing 777	960

B. Concrete Pavement Design

For rigid pavement, FAARFIELD takes into consideration the maximum horizontal stress at the bottom edge of the concrete slab as a basis for design.

C. Asphalt Pavement Design

For flexible pavements, FAARFIELD takes into account two modes of failure in determining the structural life of the pavement. The two modes of failure are the maximum vertical strain at the top of the subgrade layer and the maximum horizontal strain at the bottom of the asphalt layer [7]. FAARFIELD presents the required thicknesses for each layer in order to withstand the subjected traffic load. However, FAARFIELD doesn't take into account asphalt rutting and assumes a constant elastic modulus regardless of traffic and temperature. Asphalt is a viscoelastic material and its dynamic

modulus depends on the material mix, the temperature and the speed of traffic subjected on the pavement. To provide a feasible design, mechanistic-empirical pavement design guide (MEPDG) software was used in addition to FAARFIELD to calculate the required asphalt thickness.

The design for a flexible pavement consists of an asphalt layer placed on a base course or pre-existing pavement which is supported by a subgrade. The asphalt layer must ensure the stability of the structure under the imposed loads and halt water penetration. Also, it is to provide a smooth well-bonded surface, which has to be free from loose particles that might damage the aircrafts. To fulfill the requirements, the asphalt layer should contain the right combination of aggregates and bituminous binders. As recommended by the FAA, dense-graded asphalt mixture item P-401 (Plant mixture bituminous pavements) is proved to achieve all the essential requirements and is used in the design of flexible pavements [7].

VI. CURRENT DESIGN

The taxiway's current design plans were obtained from Dar Al Handasah indicating the cross-section of the taxiway. The taxiway's cross-section, Fig. 4 in the appendix, is made of a 45 cm PCC layer which is casted on a 20 cm cement treated base. The shoulder's cross-section is made up of a 7.5 cm asphalt concrete layer. Asphalt is casted on a 30 cm crushed aggregate base course layer and the entire pavement structure rests on a weak-to-average soil subgrade.

VII. DESIGN

A. Rigid Pavement Design
1) New PCC Pavement-Complete Removal

Taxiway pavements may require complete reconstruction if it is heavily worn out. The new PCC pavement was designed using FAARFIELD based on a design life of 30 years using a concrete flexural strength of 4.83 MPa. The result was a 43 cm thick PCC surface layer over a 30 cm cement treated base. For construction purposes and elevation concerns, the PCC layer will be taken to be 45 cm. In the case of the complete removal scenario, the taxiway will be upgraded from Code letter E to F. The current taxiway has pavement width of 23 m, which follows Code E. It does not apprehend to Code letter F aircraft like the A380.

2) PCC Bonded Overlay

FAARFIELD was used to calculate the thickness of the PCC bonded overlay. During the design procedure, an undefined material was placed on top of the PCC surface to resemble the action of an overlay and was assigned an elastic modulus value of 27,579 Mpa to resemble the properties of PCC. The obtained thickness (7.05 cm) was not sufficient to satisfy the minimum thickness requirements of FAA. Hence, the minimum thickness (7.5 cm) was taken and rounded up to 10 cm for construction purposes. The total thickness was obtained to be 75 cm which complies with the current level of the taxiway.

B. Flexible Pavement Design

For flexible pavement, MEPDG was used to determine the thickness of the asphalt layer. MEPDG only designs for a critical aircraft. Thus, iterations were conducted using FAARFIELD to determine the number of aircrafts needed to

generate a PCC thickness of 45 cm. After determining the number of departures needed for each aircraft to generate a 45 cm PCC slab, the critical aircraft was determined based on the lowest departures needed, which in this case is Boeing 777-300 ER. A conversion factor was computed for each aircraft to convert it to the critical aircraft. This was done by dividing the number of departures of the critical aircraft by the number of departures of a certain airplane. The conversion factor was then multiplied by the actual annual departures of each airplane to convert all the airplanes to the critical airplane.

1) Asphalt Overlay on Existing PCC Pavement

Asphalt overlay can be applied to existing rigid pavements given that the PCC slab isn't too damaged to ensure that cracks won't reflect to the asphalt layer. The overlay thickness was determined to be only 5 cm using FAARFIELD.

After all the aircrafts were converted into the critical aircraft, the gear of the B777-300 ER aircraft can be inputted into MEPDG. Taking into account Beirut's climate, the asphalt thickness was computed again on MEPDG to determine whether rutting is acceptable. For the asphalt layer, the PG grade for the binder and asphalt mix were selected according to what is used in Lebanon and complying with the FAA standards. PG 64-16 was chosen in this case. The asphalt mix gradation was chosen according to the specifications for item P-401 (plant mix bituminous pavements). Also, the air voids were taken to be 6%, so that after compaction they will be reduced to the required 4%. The asphalt mixture is composed of 5% of asphalt binder and 95% of aggregates. The asphalt thickness was computed to be 10 cm. The structure was strong enough to withstand fatigue cracking which was approximately 2% over the design period of the pavement. The structure also passed rutting, surface down cracking and IRI or international roughness index, which is a roughness measure for the surface of the pavement.

2) Asphalt Overlay on Rubblized PCC Pavement

PCC slabs that acquire a lot of cracks and deteriorated sections may be rubblized and used as a subbase for an asphalt overlay. The elastic modulus for rubblized PCC layers ranges between 100 to 400 ksi (690 to 2757 Mpa) depending on the depth of the PCC layer [8]. For a rubblized layer of 450 mm, the asphalt layer required was computed to be the minimum thickness regardless of the modulus of elasticity. To determine the appropriate thickness for rutting, the thickness must be computed by MEPDG.

The properties of the asphalt layer are the same as the previous scenario. While checking for fatigue, it was noticeable that fatigue was also low in this scenario. This is due to the thick rubblized layer which has a relatively high elastic modulus. The pavement also passed rutting, surface-down cracking, and IRI. Asphalt experiences almost no rutting while the subgrade has the highest contribution to the total rutting in the structure. The asphalt thickness was computed to be 15 cm.

3) Shoulder Design

The taxiway's shoulder was designed based on one annual departure for the most critical aircraft that uses the pavement for a design life of 10 years. Since the shoulder is designed for one annual departure, it was designed as a flexible pavement because it is cheaper [7]. Using FAARFIELD, it was found that a 10 cm thick asphalt layer needs a 30 cm crushed aggregate base. However, the crushed aggregate layer will be taken as 37.5 cm due to elevation constraints. The shoulder design is applied to all sceanrios. The shoulder was designed again on MEPDG to check for rutting, IRI, and surface down cracking. The asphalt layer barely experiences rutting and most of the rutting is due to subgrade rutting which is similar in all the different asphalt design scenarios.

VIII. CONSTRUCTION

A. Phasing

In order to not disturb the airport traffic, the construction process will be divided into 3 phases. The first phase starts from the beginning of the taxiway to 65m before the second entrance (total distance of 445m). Thus, allowing aircrafts to access runway 03-21 through the second entrance while complying with FAA's safety standards. Phase one is the most critical phase since it blocks access to the nearest entrance to the runway. Hence, the section included in phase one must return back to operation as soon as possible to reduce taxing. The second phase extends from the second entrance to 65m before the third entrance (total distance of 600m). Finally, the third phase extends from the third entrance to the end of the taxiway (total distance of 1335m). The reconstruction process must occur when runway 03-21 is used for takeoff and runway 16-34 is used for landing. This is done to reduce the interference with the planes entering runway 16-34. If reconstruction is to take place while runway 16-34 is used for takeoff, construction at phase 3 will hinder takeoff operations by reducing the available runway length. Landing requires less runway length to achieve the required deceleration rate. The most critical aircraft (B777) requires 2000m of landing distance which is available even when construction is taking place at phase three. The phases are shown more clearly in fig. 5 in the appendix.

B. Scenarios

1) New PCC Pavement - Complete Removal

Complete removal is necessary if the concrete pavement is badly damaged. The asphalt layer will be removed by scrapers and loaded to trucks. Then, the concrete layer will be rubblized using a resonant machine. The rubblized material will be removed using excavators and later be used as aggregate for the shoulder's base layer. Next, a PCC slip-form paver will lay 45 cm of 60 Mpa PCC and insert the dowels in 3 stages. The core is divided to 3 parts: two sides (10 m each) and the center (5 m). The sides will be paved to provide a standing ground for the paver to be able to finish the last 5 m located in the center. The paving of the center section will take place after 7 days from paving the sides. Saws will be used to cut the concrete into sections to prevent thermal, extension and shrinkage cracks from occurring [9].

2) PCC Bonded Overlay

For the PCC bonded overlay, the existing layer of asphalt is to be removed and replaced by a PCC overlay. Asphalt will be scraped using loaders and will be loaded by back hoe excavators into trucks. Similar to the complete removal scenario, a PCC slip-form paver will lay 10 cm of 60 MPA PCC in 3 stages. Finally, saws will be used to cut the concrete into sections to prevent thermal, extension and shrinkage cracks from occurring.

3) Asphalt Overlay on Existing PCC Pavement

In this case, the apparent layer of asphalt is to be removed and replaced by a new one. First, the asphalt is to be scraped using front loaders or bulldozers and then loaded in trucks using back hoe excavators. Since the shoulders already have an asphalt layer instead of concrete, the lowest layer will be removed during the process and replaced by base aggregates (7.5cm) to compensate for elevation difference and maintain a total asphalt thickness of 10 cm. The milling machine will pass once over the present PCC layer to create a bonding surface between PCC and asphalt.

4) Asphalt Overlay on Rubblized PCC Pavement

In this scenario, all asphalt layers will be scraped using front loaders and loaded into trucks. The concrete layer will be rubblized and will form a subbase for the new asphalt layer. 5 cm of the rubblized concrete will be removed to correct for the difference in elevation. A 15 cm layer of asphalt will be constructed in 2 equal lifts of 7.5cm.

5) Shoulder

Shoulder pavement is composed of a Hot Mix Asphalt layer. Standard construction procedures for asphalt will be applied. After the removal of the shoulder material, a 7.5 cm layer of crushed aggregates will be added above the existing layer. The existing layer will be tested for the required density and will be compacted if needed. The new layer will be leveled and graded, using a grader, and will be compacted and tested on site to achieve the required density. Finally, the asphalt layer will be placed on the crushed aggregates layer using pavers.

IX. COST ESTIMATION

Cost estimation is crucial in determining the ideal scenario for the rehabilitation process. Since the rehabilitation of the taxiway is a public project, the lowest bidder is usually chosen to commence with the execution. In light of that, an approximation of the total cost of each solution must be computed to discern the optimal solution. Costs of material, machinery, and labor were collected from local contractors and used to calculate the BOQ for each scenario. By comparing the initial construction costs of the different scenarios, it was concluded that the asphalt scenarios are more cost-effective than the concrete scenarios as shown in table 2. However, after conducting a life cycle cost analysis of 60 years and a maintenance costs, the asphalt overlay scenario turned out to be more expensive than the PCC bonded scenario. On the other hand, the asphalt over rubblized scenario was more cost effective than the total removal scenario even after the integration of maintenance costs such as cut and patch, surface treatment, and crack and seal.

TABLE 2. COST ESTIMATION

Scenario	Initial Cost	Total Cost (life cycle)
Total Removal	$ 5,927,200	$ 6,527,200
Asphalt Rubblized	$ 2,120,400	$ 4,770,400
PCC Bonded	$ 2,511,000	$ 3,084,000
Asphalt Overlay	$ 1,897,200	$ 4,147,200

X. TIME SCHEDULING

The time schedule is based on the destruction and reconstruction processes of the taxiway. The time allocated to each process was collected from local contractors. The formula shown below was used to calculate the duration.

$$Duration = Production\ rate \times quantity \qquad (1)$$

On the other hand, the volume of material needed for each activity was calculated based on the taxiway plans obtained from Dar Al Handasah. Table 3 shown below displays the time needed for each scenario. As stated earlier, phase one is the most critical phase since it blocks access to the nearest entrance to the runway, which increases the taxiing distance for aircrafts. Accordingly, it is necessary for construction to follow the designated schedule to ensure minimal interference with aircraft maneuvering. The asphalt scenarios require less construction time since asphalt does not require curing nor sawing. This reduces the time required for the taxiway to go back to operation.

TABLE 3. ESTIMATED TIME FOR DIFFERENT SCENARIOS

Scenario	Phases	Time (days)
Total Removal	1	36
	2	48
	3	62
	Total	146
Asphalt overlay	1	10
	2	15
	3	30
	Total	55
PCC Bonded	1	26
	2	33
	3	42
	Total	101
Asphalt over Rubblized	1	11
	2	17
	3	33
	Total	61

XI. ENVIRONMENTAL IMPACT

When it comes to engineering projects, it is vital to assess the environmental impact in order to minimize the production of hazardous material and chemical substances. The environmental impact for each scenario is provided in the following sections.

A. Asphalt Overlay VS PCC Bonded Overlay

PALATE excel sheet was used in order to determine the environmental impact for the different scenarios by inputting the quantity of material needed. In these two scenarios, the asphalt overlay and the asphalt layer of the shoulder are the only materials that are removed since the underlying layers are assumed to be in a good condition. Figure 2 shown below displays the comparison between the two scenarios. As illustrated, the PCC scenario emits a significant amount of environmentally ominous gasses compared to the asphalt scenario.
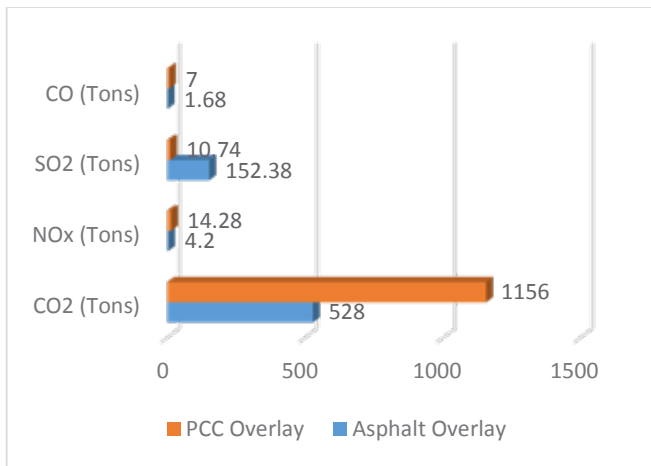
Figure 2. PCC overlay vs. asphalt overlay

B. Asphalt over rubblized concrete vs. new PCC layer

For the complete removal scenario, the asphalt overlay in addition to the existing PCC layer were removed. In the asphalt scenario, the PCC layer was rubblized so only the asphalt overlay was removed. The rubblized PCC layer was used as an aggregate base course for the asphalt layer. Similarly to the previous scenarios, the PCC alternative emits more environmentally ominous gasses than the asphalt scenario. In conclusion, the asphalt scenarios are considered to be environmentally feasible compared to the PCC scenario.
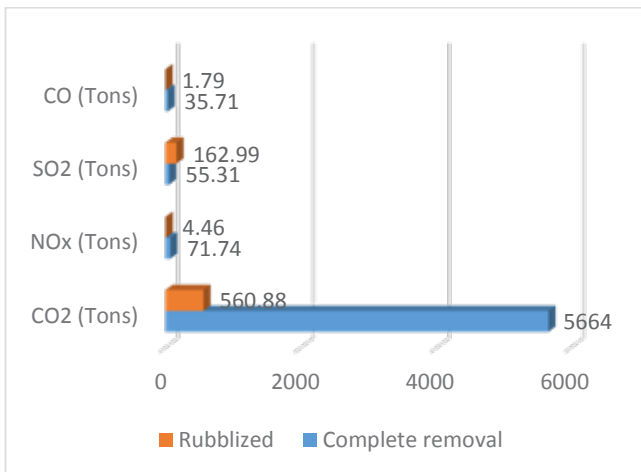

Figure 3. Rubblized vs. complete removal

XII. CONCLUSION

Projects that touch on the rehabilitation of structural elements are complicated but specific in the purpose they are done to serve. The rehabilitation of taxiway Alpha in Beirut International Airport is a prominent project to complete. Looking at the current state of the taxiway, the need for rehabilitation was obvious. Based on extensive research and comparison to projects of the same nature, many scenarios seemed to fit the case of taxiway Alpha. The rehabilitation process includes a significant service life factor that is expected to last without the frequent need for maintenance. Bearing that in mind, the design is based on the most critical aircraft that can use the facility, the B777-300 ER aircraft.

This design takes into consideration the alarming state of the current taxiway and tackles the goal of not encountering

such damage again and if so; the damage is meant to be kept to the minimum with proper regulated maintenance. While assuming the underlying PCC layer to be in a good condition, the optimal scenario was chosen to be the asphalt overlay. Although it costs more than the PCC bonded overlay due to the induction of maintenance costs, the time required to complete the rehabilitation is much less than the PCC overlay, which is an important aspect when dealing with projects like the airport. In addition, the maintenance of the asphalt overlay takes less time than that of the PCC overlay.

Assuming the PCC layer is in a deteriorated state, the asphalt over rubblized PCC scenario is chosen to be the optimal solution, since it is executed faster and at a lower cost compared to the complete removal scenario. Also, the asphalt scenarios generate less ominous chemical substances which make them more favorable than the PCC scenarios.

In conclusion, this project attends to all practical aspects that adhere to construction of this magnitude and nature. Operating airports rehabilitation requires a strict ethical commitment from all concerned parties, and it poses a significant pressure on the contractor to deliver the promised end product as scheduled to avoid fines and legal responsibilities.

REFERENCES

[1] N. Hage Ali, "Beirut Rafic Hariri International Airport (BRHIA), A Hub for the Middle East Once Again", Michigan, 2010.
[2] N. Thom, Principles of Pavement Engineering, 2nd ed. London: ICE Publishing, 2012, pp. 25-30.
[3] "Flexible Pavement Distress | Pavement Interactive", Pavementinteractive.org, 2016. [Online]. Available: http://www.pavementinteractive.org/category/pavement-management/pavement-distresses/flexible-pavement-distress/. [Accessed: 06- Jan- 2016].
[4] Transportation Research Board, "Rubblization of Portland Cement Concrete Pavements", Washington DC, 2006.
[5] American Concrete Pavement Association, "Guidelines for Bonded Concrete Overlays", 1990.
[6] "Adding Layers but Staying Thin – Thin Asphalt Overlays | Pavement Interactive", Pavementinteractive.org, 2016. [Online]. Available: http://www.pavementinteractive.org/2013/09/16/adding-layers-but-staying-thin-thin-asphalt-overlays/. [Accessed: 31- Feb- 2016].
[7] Federal Aviation Administration, "Airport Pavement Design and Evaluation", FAA, 2009.
[8] N. Gucunski, "Elastic Modulus of Concrete from Compression Tests", ACI Materials Journal, vol. 863, p. 33, 2008.
[9] "Early-Entry Sawing of Portland Cement Concrete Pavements", FHWA, 2007. [Online]. Available: http://www.fhwa.dot.gov/pavement/concrete/pubs/07031/. [Accessed: 31- Jan- 2016].

Table 4. Types of pavement defects found in BRHIA

Type	Image 1	Image 2
Block Cracks		
Fatigue Cracks		
Crocodile Cracks		
Reflective Cracks		
Edge Cracks		
Other Problems (Potholes, Ponding)		

Table 4. Types of pavement defects found in BRHIA
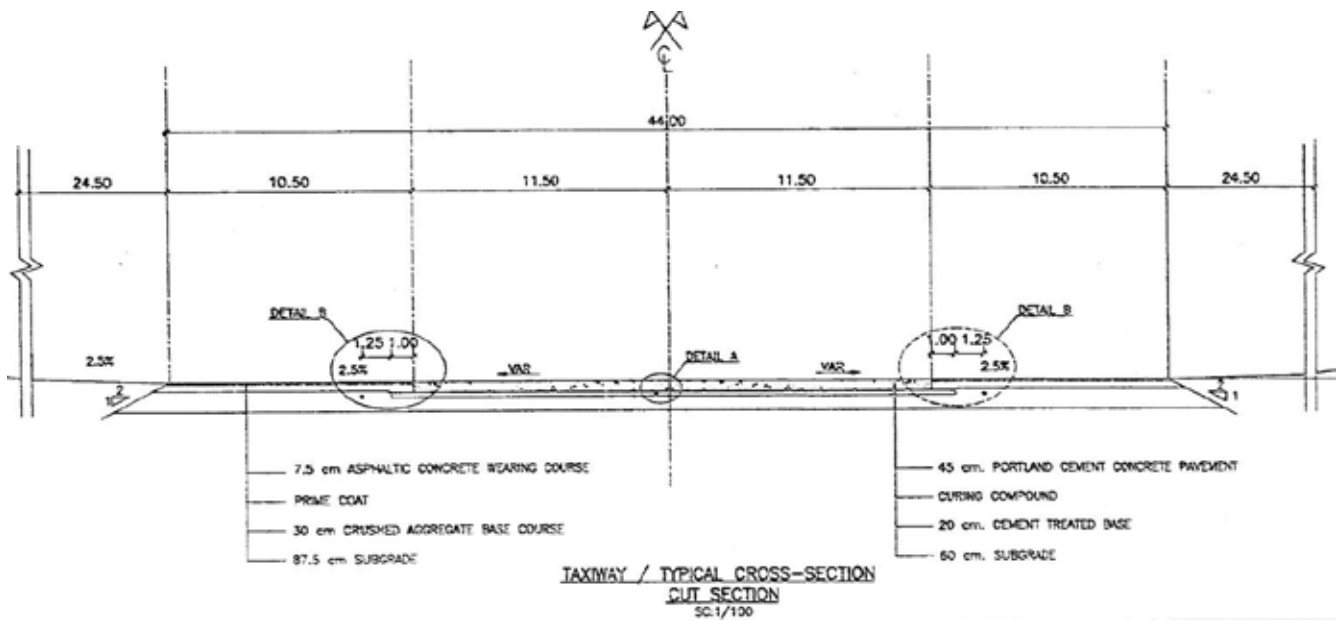
Figure 4. Current design cross section
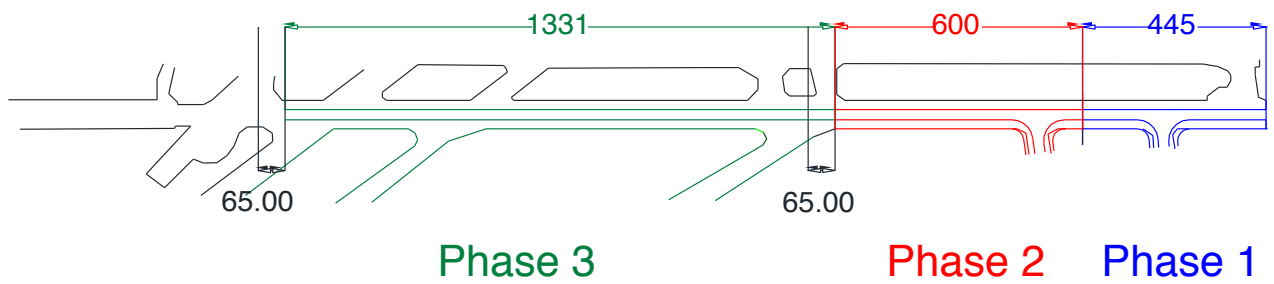
Figure 5. Phasing plan

Short Sea Shipping in Lebanon: Economic Benefits and Business Case

Nawras Akroush, Bassel Halabi, Ousama Moughraby, Bassel Sadek, Cesar Yahia

Civil and Environmental Engineering

American University of Beirut

Riad El-Solh, Beirut, Lebanon

nsa54@mail.aub.edu, bnh04@mail.aub.edu, unm00@mail.aub.edu, bas11@mail.aub.edu, cny00@mail.aub.edu

Abstract - **Following the expansion of the Beirut port in 2012, the high demand for land freight transportation exacerbated the congestion levels along the coastal highways. Short Sea Shipping (SSS) is proposed as a sustainable freight transport mode that is capable of alleviating congestion levels along the coastal highway and reducing the environmental impacts associated with freight transport by trucks. A multi-step aggregate model is developed to obtain the expected share from the national Lebanese freight transport using the coastal highway the SSS system would capture based on its service attributes (time and cost) compared to existing freight modes. Then, the economic benefits and costs as well as the financial viability of the proposed system are evaluated. The study results indicate that for a competitive service fee, the short sea shipping market share could be as high as 55%, and represents a viable long-term business proposition with an internal rate of return for the service operator of close to 30%. In addition, the proposed service is expected to make significant economic benefits to society based on reductions in congestion, gaseous emissions, pavement maintenance costs, and accidents.**

I. INTRODUCTION

Since the start of the civil war in 1975, Lebanon has witnessed severe economic and congestion costs due to its deteriorating infrastructure system. Heavy vehicles such as container trucks constitute a small proportion of the traffic mix. However, they contribute significantly to congestion levels due to their increased physical dimensions and their inferior acceleration performance [1]. Therefore, the need for an alternative freight transport system is apparent.

Short sea shipping (SSS) had recently emerged as an alternative freight transport mode in Europe. Paixao and Marlow characterized the SSS network in Europe by detailing the type of vessels to be used, the dependence on other multiple modes of transport, the operations and types of services, as well as the geographical area in which the system would operate [2]. However, a "strict definition" of SSS that classifies the type of vessels, distance to be covered, and type of cargo transported is not provided in the literature [3].

Sambracos and Maniati evaluated the financial feasibility of SSS between Greek ports; they deduced that sea transportation is significantly more competitive than the truck mode [4]. Additionally, Paixao and Marlow indicate that the financial costs incurred in developing a SSS network are lower than infrastructure costs associated with road and rail projects [2]. On the other hand, Mulligan and Lombardo evaluated the economic benefits of SSS that pertain to reductions in travel time and reduction in fuel consumption; they concluded that SSS significantly reduced harm to the environment [5]. Therefore, SSS is an attractive financial and economic alternative.

In Lebanon, the need for an alternative mode of freight transport along the coastal highway was highlighted in [6]. Other than that, the flow of goods between zones in Lebanon was evaluated using a distribution model [7]. Particularly, the amount of freight transported between zones was computed by developing a method to generate a national freight origin-destination matrix [7].

The aim of this work is to propose a framework for evaluating the viability of using a short sea shipping network along the Lebanese coast (S3L). A multi-step aggregate model was developed in order to obtain the quantity and type of freight that will be transported using the S3L system. The aggregate model is comprised of a freight generation model, a distribution model, and a mode split model. First, the generation model is used to determine the demand and supply of goods at different Lebanese zones. Then, the distribution model was implemented in Matlab to obtain the flow of goods between zones. Afterwards, the mode split model was used to obtain the likelihood of choosing the S3L alternative as compared to the road transport alternative. Finally, the results of the aforementioned generation, distribution, and mode split models are used to determine the flow of each commodity by mode of transport between any two zones. The multi-step aggregate model framework is shown in Fig.1.

Furthermore, the economic and financial feasibilities of the S3L system were evaluated. The economic feasibility was determined by quantifying the benefits resulting from reduced accidents, decreased maintenance for pavements, reduced travel time, and reduced emissions. Meanwhile, the financial feasibility was determined by comparing the costs incurred in shipping to the costs incurred while transporting freight by truck. The paper proceeds by detailing the aggregate model framework. Then, the financial and economic feasibilities are evaluated. Finally, a discussion of the obtained results is provided.
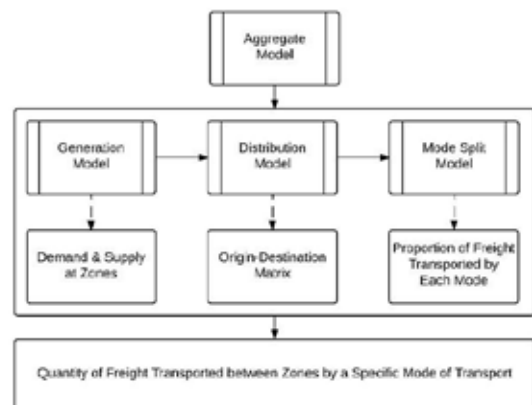

Figure 1 Aggregate Model Framework

II. Aggregate Modelling Of National Goods Movements

Aggregate models are usually developed to simulate the flow of national and international goods. The main purpose of using these models is to determine the flow of a certain type of freight that departs from traffic zone x and reaches another traffic zone y, using transport mode m [8]. The multi-step aggregate model procedure constitutes of 3 steps which are generation, distribution, and mode split. Afterwards, the quantity of flow by mode is determined using (1). Q_x^{sh} is obtained from the generation model and it refers to the average quantity of freight type 's' in the time period 'h' produced by traffic zone 'x'. $p[\,y/xsh\,]$ is obtained from the distribution model and it refers to the probability that the freight of type 's' in the time period 'h' produced by traffic zone 'x' goes to traffic zone y. Therefore, the product of $p[\,y/xsh\,]$ and Q_x^{sh} would give the amount of freight transported between zones; this would enable the generation of an origin-destination (O-D) matrix. Afterwards, $p[\,m/xysh\,]$ is obtained from mode split model and it refers to the probability that the freight of type 's' in the time period 'h' on the x-y pair is transported by mode 'm'. Finally, $Q_{xy}^{sh}[m]$ could be evaluated and it refers to the average flow of freight type 's' in the time period 'h' between the x-y pair that is transported by mode 'm'.

$$Q_{xy}^{sh}[m] = Q_x^{sh} \cdot p[\,y/xsh\,] \cdot p[\,m/xysh\,] \qquad (1)$$

A. Freight Generation

The internal zones for generation and distribution of goods were considered to be the 26 districts ('Kada's') in Lebanon such that each district represents a zone. Fig. A1 in the Appendix shows all the districts in Lebanon. For the short sea shipping system under study, only two categories of goods were taken into consideration: goods that are produced in Lebanon and exported, and goods that are imported and consumed in Lebanon. Products that are produced and consumed in Lebanon were disregarded since it was not possible to find any data concerning the inter-zonal flow of Lebanese goods.

The amount (in tons) of each imported/exported product at each customs office in the country was obtained from the Lebanese Customs [9]. In order to determine the supply and demand at each internal zone, the distribution of factories in Lebanon was obtained from a report by The Directory of Exports and Industrial Firms [9], [10]. Afterwards, the demand at each internal zone was determined by evaluating whether the commodity was consumed by the general population or a specific industry. If the commodity was consumed by the general population, the distribution of goods was based on the percent population in each zone. Otherwise, the demand for commodities by industry was obtained by calculating the percentage of each industry in each zone and multiplying that percentage by the total amount of imports. A similar procedure was used to obtain the supply from each zone based on the number of factories present. The commodities considered in this study were food and beverages, iron and steel, chemicals and plastics, stone and plastering material, mineral fuels and mineral oils, fertilizers, and wood.

B. Freight Distribution

The gravity model was utilized to predict the flow of freight between demand and supply zones as a function of the transport cost [11]. The demand zones for imports were considered to be the aforementioned districts, while the supply zones were considered to be the ports and the Lebanese border checkpoints. On the other hand, for exports, the supply zones were considered to be the districts, and the demand zones were considered to be the ports and checkpoints. Equation (2) shows the gravity model in its general form [12]. T_{ij} represents the flow of goods from zone i to zone j. B_i is a balancing factor that ensures the total flow from i equals to the number of trips produced at i. C_j is a balancing factor that ensures the total flow into j equal to the number of trips attracted to j. S_i is the supply at node i. D_j is the demand at node j. F_{ij} is the friction factor which represents the transport cost between i and j.

$$T_{ij} = B_i * C_j * S_i * D_j * F_{ij} \qquad (2)$$

However, a common form of the gravity model assumes that the model is production constrained only ($C_j = 1$) [12]. The formulation of the production constrained gravity model is shown in (3). M represents the total number demand zones such that k and j correspond to a specific demand zone. This modified version of the model was used iteratively along with a demand balancing procedure that minimizes the difference between actual demand and calculated demand [7]. First, the flow into a particular demand node, j, from each supply node, i, is obtained by applying (3) repeatedly across all supply zones. Thereafter, the calculated demand at j is obtained by summing the computed flows from each zone to j and is given by $\sum_i (T_{ij})_n$. Then, equation (4) is used to minimize the difference between calculated demand and actual demand. The actual demand at zone j is given by D_j', this value is fixed based on generation model data. $(D_j)_{n+1}$ is the estimated demand, after balancing, at iteration n of the balancing procedure. The estimated demand, $(D_j)_{n+1}$, is assigned to D_j, $D_{k=j}$, and $(D_j)_n$ at iteration n+1 of the balancing procedure. Thereafter, the aforementioned procedure is applied iteratively to update the flow values into node j that are obtained from (3). The procedure stops when the difference in estimated demand from two consecutive iterations $((D_j)_{n+1} - (D_j)_n)$ is within a certain threshold. This procedure is repeated for all demand zones to generate an origin-destination matrix. The initial values of D_j, D_k, and $(D_j)_n$ were taken to be equal to the actual demand D_j'. The gravity model was implemented in a Matlab code that relies on three dimensional matrices to accommodate all the commodities involved.

$$T_{ij} = \frac{S_i * (D_j) * F_{ij}}{\sum_{k=1}^{M}(D_k) * F_{ik}} \qquad (3)$$

$$(D_j)_{n+1} = (D_j)_n * \frac{D_j'}{\sum_i (T_{ij})_n} \qquad (4)$$

Furthermore, the friction factors were taken to be the inverse of the square of the distance between every two zones. The distances were obtained by writing a Python script that uses Google Maps API to calculate the 'waypoint distances' travelled by trucks on the existing road network in Lebanon. The aforementioned road network is shown in Fig. A2 in the Appendix. The center of each zone was considered to be near the center of business activity instead of the geographical center in order to improve model accuracy [13]. Fig. A3 in the Appendix shows the results of the gravity model, the import and export flows were overlaid to show the total flow in tons on a certain link.

C. Mode Split Model

The mode choice model used is based on Random Utility Theory, it assumes that the users are rational and that they aim to maximize their utility by choosing a particular mode. The choice set that was considered in this case study consists of the truck mode (road) and the vessel-truck mode (road-sea). Trucks are required with the vessels in order to provide door to door services. The systematic utility functions for the alternatives are based on level of service and performance attributes [8]. Equation (5) shows the systematic utility function used in this study. The cost and time attributes associated with the mode are considered as the major determining factors for mode choice. V_m^{od} is the systematic utility of using transport mode m on the origin-destination (O-D) pair. α is an alternative specific constant which is used to capture the preference to a particular mode, given that all other parameters are equal across modes. T_m^{od} is the travel time of transport mode m on the O-D pair. C_m^{od} is the travel cost of transport mode m on the O-D pair. $\beta_{r,m}$ represents the model parameters, where r corresponds to the attribute under consideration.

$$V_m^{od} = \alpha + \beta_{c,m} \cdot C_m^{od} + \beta_{t,m} \cdot T_m^{od} \qquad (5)$$

In order to evaluate the travel time by road along each O-D pair, a Python code was developed to extract the travel time from Google Maps API. On the other hand, calculating the travel time by sea required determining the navigation routes on which the vessels are permitted to ferry. The map showing the navigation routes was obtained as a hard copy from the Director General of Land and Maritime Transport. Afterwards, the map was digitized as shown in Fig. A4 in order to develop an appropriate scale and to obtain the distance traveled along the routes. The travel time was determined by dividing the distance by vessel speed. However, the total travel time for the road-sea alternative was determined by adding the travel time by sea, the time needed for loading/unloading, and the travel time required for door to door service.

The cost of transporting freight by the road alternative was obtained from summary sheets established by the Cargo Master Group [14], [15]. The sheets provided transportation rates based on container size (20 ft. or 40 ft. container) and distance traveled. According to the Deputy CEO at Gezairi Transport, Mr. Fouad Bawarshi, 60% of the Lebanese goods are transported in 20 foot containers while the remaining 40% are transported in 40 foot containers. For every O-D pair, the travel cost by truck was calculated assuming a weighted average container size based on the above proportions and based on the distance between the two nodes. On the other hand, the travel cost of the road-sea transport for the mode choice model was broken down into two categories: sea transport costs and road transport costs. The sea transport cost represents the price of the service charged by the S3L operator. Meanwhile, the road transport costs for providing door to door service was also obtained using the rates provided by Cargo Master Group [15], [14]. Therefore, the total cost for the user to use the road-sea alternative would be the summation of the sea transport costs and the aforementioned road costs.

The model parameters, $\beta_{j,m}$ were estimated by relying on a model calibration performed in [8] for freight transported by road-sea and by road in Italy. However, a sensitivity analysis was performed on the aforementioned model parameters in order to understand the implications of the adopted assumption. The results of the sensitivity analysis indicated that the variation in model parameters did not have a significant impact on the market share for 92% of the commodities considered (all commodities except exported iron and steel, and salts). Fig. A5 in the Appendix shows the results of the sensitivity analysis. After obtaining the necessary parameters for (5), the systematic utilities for each alternative are computed. Then, the logit model shown in (6) was used to obtain $p[\,m/xysh\,]$ in (1).

$$p[\,m/xysh\,] = \frac{\exp(V_m^{xy})}{\exp(V_{road}^{xy}) + \exp(V_{road-sea}^{xy})} \qquad (6)$$

Finally, after identifying all the necessary parameters of (1), $Q_{xy}^{sh}[m]$ could be computed. Thereafter, it is required to assess the values of the variables that influence the mode split. As previously mentioned, the time and cost are the parameters considered. The cost of transport by truck are given at fix rates, the time for travel by road is provided by Google API, and the time for transport by sea is dictated by the speed of the vessel chosen. Therefore, the only variable that could be altered to influence the results of the aggregate model is the price of the service. The expected demand for the short sea shipping service is estimated from aggregate model to be in the range of 55% of the total freight transported along the coast. This result was based on varying the price of the service between $150 and $250 per container. However, the price of the service will also influence the financial and economic feasibility of the proposed system. Therefore, a framework was developed in order to obtain the optimal price based on its impact on the mode split, financial feasibility, and economic feasibility. Fig. 2 shows the proposed framework.
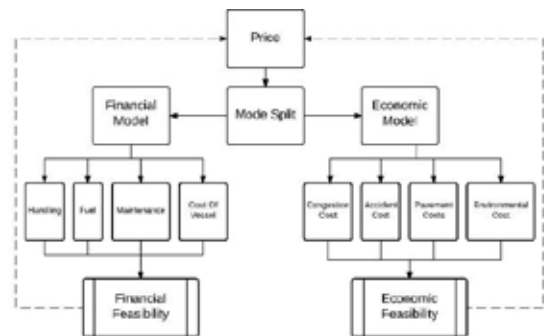

Figure 2 Impact of Price on Financial & Economic Feasibility

III. Financial Feasibility & Infrastructure Costs

A. Financial Feasibility

The financial feasibility of the proposed short sea shipping system was considered from the perspective of the operator of the service. The operator would have to incur terminal costs, fuel costs, vessel costs, and vessel maintenance costs. In return, the operator would ship containers between ports along the coast for a certain fee per container. The terminal costs per container were obtained to be $125 per container [14], [15].

The B203 Arctic Cargo Vessel manufactured by Remontawa was chosen to operate the system due to its capacity of 108 TEUs which is commensurate with the level of demand expected. Additionally, the use of the B203 vessel eliminates the terminal handling costs associated with using the cranes at the port since it is a geared vessel. The capital costs of the B203 vessel is estimated to be eight million dollars, and the life span is estimated to be 30 years based on comparison with a vessel described in [16]. Furthermore, assuming 320 days of operation, the annual operating and docking costs were approximated to be $2,200,000 [17]. Note that the aforementioned operating costs include maintenance costs, crew costs, and insurance costs.

As for the fuel costs, the amount of fuel consumed by the vessel per nautical mile was obtained from [5]. Thereafter, assuming that every liter of fuel costs $1, the fuel costs were estimated to be $3.67 per nautical mile. However, determining the total fuel cost requires defining the frequency of operation. Therefore, an optimization model that aims to minimize the distance travelled while satisfying the average demand for S3L at different locations in Lebanon was developed. A binary decision variable was set to each link connecting the different locations considered (North of Lebanon, Beirut, and South of Lebanon). The objective function of the proposed optimization is shown in (7). $X_{(i,j)}$ is a binary variable that assumes a value of 1 if the link is traversed and 0 otherwise. $D_{(i,j)}$ is the distance corresponding to the considered links.

$$Min \left(Total\ Distance = \sum_{All\ Links\ (i,j)} X_{(i,j)} \times D_{(i,j)} \right) \quad (7)$$

The objective function shown in (7) was subject to a set of constraints. The first constraint was that the average daily demand and supply of containers at each port must be met. Second, the vessel was assumed to operate for a maximum of 16 hours per day. The operation time included shipping and loading/unloading times. The last constraint applied was that the number of containers on any traversed link must not exceed the vessel capacity. Afterwards, the Excel Solver tool was utilized to determine the set of links that determine the most viable service while satisfying all the aforementioned constraints. Results of the optimization model indicated that the links connecting North to Beirut, and Beirut to South, must be traversed twice per day for optimal operations. In other words, the vessel can complete a whole round trip from North to Beirut to South and backwards from South to Beirut to North per day. Therefore, since one roundtrip corresponds to 167.8 nautical miles, the total fuel consumption per day would cost $616. Note

that the ports considered for that roundtrip are the ports of Beirut, Tripoli (North), and Sidon (South).

Finally, the costs for transporting freight by sea were obtained by adding the terminal, maintenance, fuel, and vessel costs. The price of the service was determined by adding to the total cost a certain profit. Table I below shows the internal rate of return, net present value, and the number of years to break even considering several service prices. The life span considered for the financial feasibility was the service life of a vessel which is 30 years. In addition, it was assumed that the discount rate is 12%. For a price of $175 it was shown that the net present value of the project is zero. This implies that the revenues generated over the life span of the project are equal to the costs. Increasing the price of transporting a container using S3L drops the percent of freight transported by sea along the coast as compared to road transport. Nevertheless, it is shown that a price of $190 per container would generate enough revenues such that the costs incurred initially are reimbursed in 8 years. Additionally, at $190 per container, the return on investment is in the range of 22%. However, it can be observed that IRR is not significantly influenced by the change in price per container. This could be attributed to the low absolute value of $\beta_{c,sea}$ in (5) which was obtained from a study of freight transport in Italy [8]. In order to obtain accurate values for the parameters, a survey is required to obtain the stated preference of the potential service users. Then, the parameters should be calibrated based on a maximum likelihood estimation.

TABLE I
FINANCIAL FEASIBILITY - NPV, IRR, AND YEARS TO BREAK EVEN

Price ($/container)	Profit ($/container)	NPV ($)	IRR (%)	Years to Breakeven
175	48	0	12	30
190	60	5,980,000	22	8
200	70	11,070,000	30	5

B. Port Infrastructure

The Ports of Beirut and Tripoli are currently equipped to accommodate the SSS system [18]. On the other hand, according to the Director General of Land and Maritime Transport, a new port is being constructed in Sidon. Yet, a coastal breakwater has to be designed and implemented at the new port. An integrated platform of ArcGIS software and Python coding language was used to develop a modular program that is capable of designing the aforementioned breakwater. Fig. A6 in the Appendix shows the breakwater designed along with its characteristics.

IV. Economic Benefits

The economic benefits of the proposed short sea shipping network were evaluated in order to assess the benefits to society. In addition, the economic benefits entail the contribution of the system to sustainable mobility [19].

A. Environmental Benefits

Heavy duty diesel vehicles are the main contributors of NO_x and particulate matter which cause respiratory problems, cancer

risks, premature deaths, breathing diseases, and the creation of ozone [20]. In order to quantify the amount of emissions, the emission factors that are used were adopted from a study that measured emissions per kilometer for a heavy duty truck manufactured in 1998 [21]. On the other hand, the fuel consumption for the vessel was obtained from [5]. Furthermore, the vessel emission factors for CO_2 and NO_x were obtained as 3110g/kg-fuel and 87g/kg-fuel, respectively [22]. Then, the cost of emissions was determined by referring to a study that quantified the costs emitted by evaluating the impact of emissions on a population of 1.53 million people [23]. The reduction in environmental costs were computed by multiplying the truck-kilometers alleviated from the coastal highway by the emission factors in ton per kilometer and the costs per ton of emission. Then, the environmental costs associated with S3L were obtained by multiplying the amount of emissions from the vessels by the corresponding costs per ton. The net environmental benefit was obtained by subtracting the vessel environmental costs from the environmental benefits due to truck reductions along the road.

B. Travel Time Savings

Travel times are estimated through the use of travel time models that are usually a function of volume and roadway capacity [24]. The most commonly used travel time model is the BPR equation which is given in (8).

$$tt = tt_0 \left(1 + 0.15 \times \left(\frac{V}{N \times C}\right)^4\right) \qquad (8)$$

The hourly traffic volumes along the coastal road were obtained from TEAM International; the morning and afternoon peak hours were selected for evaluating the travel time benefits. Afterwards, the traffic composition, vehicle occupancy, value of time, and passenger car equivalence (PCE) for vehicles were obtained from a recent study that evaluates road transportation in Lebanon [25]. The provided traffic counts were then multiplied by traffic composition proportions to determine the volumes for each vehicle type which were converted into PCE (V). The free flow travel times (tt_0) were obtained using Google Maps, the number of lanes on each segment (N) were obtained from Google Earth, and the roadway capacity (C) was assumed to be 1200 PCEs/hr./lane. The unit travel time of each vehicle type (tt) was then calculated using (8). Afterwards, the product of the annual passengers, unit travel time, and value of time per passenger for each vehicle type resulted in the travel time costs for each type of vehicle. Then, following the same procedure, the travel time costs that are incurred along the coastal highway after implementation of the S3L were computed. The travel time benefits represent the difference between the travel time costs without and with the S3L.

C. Benefits from Reduction in Pavement Maintenance

This study is concerned with two classes of vehicles: the 2-axle tractor and trailer (20 ft. container carrier) and the 3-axle tractor with a 2-axle trailer (40 ft. container carrier). The vehicular characteristics, namely, the empty and gross vehicular weights were obtained from Cargo Master Group [14], [15]. The corresponding ESALs for each vehicle type were obtained from

a study conducted by [26]. Results of the distribution model indicated that around 201,424 fully loaded trucks currently utilize the coastal highway every year. In the case of data unavailability, it is safe to assume that idle trips account for 33 % of the total trips made [26]. The yearly number of trips is thus 300,632. The distribution of trips among the different vehicles and conditions (empty/loaded) was then determined along with the corresponding ESALs/year. Afterwards, the structural number (SN), which is a measure of the pavement strength, was calculated assuming that the pavement along the coastal highway follows a standard Lebanese design of 5 cm hot-mix asphalt, 7 cm aggregate base, and 10 cm aggregate sub-base thicknesses. The SN was then employed in a pavement deterioration model to estimate the expected number of ESALs for pavement failure, defined as unacceptable ride quality [26]. The expected number of ESALs to failure is 1,451,381.

The overlay frequency is defined as the ratio of the ESALs required for pavement failure to the applied ESALs/year. Results showed that implementing the S3L prolongs the pavement overlay frequency by 5 years. The cost of pavement overlay is approximated to be around 12.25 $/m² for an overlay thickness of 2'' [27]. The pavement area was calculated for each segment along the coastal road where the segment length and number of lanes were obtained using ArcMap.

D. Benefits from Reduction in Accidents

A transport study conducted by [25] classified the accidents based on the degree of severity and estimated the cost per accident for each category. Accident rates were measured in personal injury accident per million vehicle kilometers [25]. The accident rate was used along with the yearly reduction in truck-vehicle kilometers to arrive at an estimated number of accident reduction per year. Furthermore, approximate proportions of fatal, severe, and slight accidents comprising the total number of accidents were provided by [25]. These proportions were used to estimate the yearly reduction in accidents for each accident type. The product of the accident reductions with the cost per accident yielded the total benefits accrued from each accident category for the current year. However, import/export activities are expected to increase at a rate of 2% per year, leading to an equivalent increase in road freight transportation and truck-vehicle kilometers [18].

E. Economic Feasibility Results

The aforementioned benefits were projected for the next 30 years, and the NPV of the benefits was calculated using a discount rate of 12%. Table II below shows the results.

TABLE II
ECONOMIC FEASIBILITY RESULTS (NPV OF BENEFIT CATEGORIES)

Price ($/container)	Congestion Benefits	Environmental Benefits	Pavement Benefits	Accident Benefits
175	$64,462,515	$3,062,229	$13,617,618	$1,981,869
190	$63,981,689	$3,033,733	$13,617,618	$1,968,423
200	$63,478,217	$3,003,911	$13,617,618	$1,949,599

V. Discussion & Conclusion

This paper provided an evaluation framework for the proposed S3L system. The framework consisted of 3 main parts: the aggregate model, the financial feasibility, and the economic feasibility. The aggregate model comprised of generation, distribution, and mode split models. The result of the aggregate model revealed the quantity of freight that would be transported by each mode along the coast in Lebanon.

From the perspective of the S3L service operator, the financial feasibility indicated that for a service price of $190 per container the IRR would be around 22%. The demand for S3L was shown to be relatively inelastic to change in price; this could be attributed to the low absolute value of $\beta_{c,sea}$. Nevertheless, the cost of transporting freight by sea over certain links is significantly less than cost of transporting by road. For example, transporting freight from Sidon to Tripoli would cost $280 less if transported by sea.

Additionally, the market share of S3L from total freight transported along coast was 55% for a service price of $200 per container. This reduction in trucks from the coastal highway results in economic benefits that are worth $82,049,000. The economic benefits are sufficient enough to justify public investment in project infrastructure and they correspond to reduction in travel time, reduction in environmental emissions, reduction in pavement maintenance, and reduction in accidents. It is worthy of note that the political risk associated with the S3L system should be considered. Nevertheless, the proposed short sea shipping network could be extended to incorporate ports in Syria and Turkey; this would facilitate economic development along the Eastern Mediterranean coast and enhance the viability of the SSS network.

References

[1] A. Al-Kaisy, "Passenger Car Equivalents for Heavy Vehicles at Freeways and Multilane Highways: Some Critical Issues," *Institute of Transportation Engineers. ITE Journal,* vol. 76, no. 3, pp. 40-43, March 2006.

[2] A. Paixao and P. Marlow, "Strenghts and Weaknesses of Short Sea Shipping," *Marine Policy,* pp. 167-178, 2002.

[3] A. M. Arof, "Determinants for a Feasible Short Sea Shipping: Lessons from Europe for ASEAN," *Asian Social Science,* vol. 11, no. 15, 2015.

[4] E. Sambracos and M. Maniati, "Competitiveness Between Short Sea Shipping and Road Freight Transport in Mainland Port Connections; The Case of Two Greek Ports," *Maritime Policy and Management,* vol. 39, no. 3, pp. 321-337, May 2012.

[5] R. F. Mulligan and G. A. Lombardo, "Short Sea Shipping - Alleviating the Environmental Impact of Economic Growth," *WMU Journal of Maritime Affairs,* vol. 5, no. 2, p. 181–194, 2006.

[6] DMJIM+Harris, "Beirut Suburban Mass Transit Corridor Feasibility Study," 2000.

[7] I. Kaysi, R. Ramadi and A. El-housseini, "Computer Based Modeling of National Goods Movement," in *Computing in Civil and Building Engineering,* 2000.

[8] A. Nuzzolo, U. Crisalli and A. Comi, "An Aggregate Transport Demand Model for Import and Export Flow Simulation," *Transport,* pp. 43-54, 2013.

[9] "http://www.customs.gov.lb/customs/trade_statistics/yearly/search.asp," Lebanese Customs, 2012. [Online].

[10] "Statistical Results and Conclusions from Field Survey - النتائج والخلاصات الإحصائية للمسح الميداني," The Directory of Exports and Industrial Firms - دليل الصادرات والمؤسسات الصناعية اللبنانية, 2013-2014.

[11] J. E. Anderson, "The Gravity Model," *Annual Review of Economics,* pp. 133-160, 2011.

[12] E. I. Pas and M. A. Todes, "A note on the balancing factors of gravity models," *Socio-Economic Planning Sciences,* vol. 15, no. 5, p. 249–253, 19 March 1981.

[13] J. Duanmu, P. Foytik, A. Khattak and R. Robinson, "Distribution Analysis of Freight Transportation with Gravity Model and Genetic Algorithm," *TRANSPORTATION RESEARCH RECORD,* vol. 2269, no. 2269, pp. 1 - 10, 2012.

[14] "Import to Lebanon General Cargo," Cargo Master Group, 2014.

[15] "Exports from Lebanon General Cargo," Cargo Master Group, 2014.

[16] H. A. Tvete and A. Brandsaeter, "ReVolt; the unmanned, zero emission, short sea ship of the future," DNV GL Strategic Research & Innovation.

[17] R. Greiner, "Ship Operating Costs: Current and Future Trends," Moore Stephans, LLP, 2015.

[18] "Analysis of Lebanon's Maritime Transport," BankMed, 2015.

[19] W. R. Black, Sustainable Transportation: Problems and Solutions, New York: The Guilford Press, 2010.

[20] G. Lee, S. You, S. Ritchie, J.-D. Saphores, M. Sangkapichai and R. Jayakrishnan, "Environmental Impacts of a Major Freight Corridor: A Study of I-710 in California," *Transportation Research Record: Journal of the Transportation Research Board,* vol. 2123, pp. 119-128, December 2009.

[21] X. Wu, S. Zhang, Y. Wu, Z. Li, Y. Zhou, L. Fu and J. Hao, "Real-world emissions and fuel consumption of diesel buses and trucks in Macao: From on-road measurement to policy implications," *Atmospheric Environment,* vol. 120, pp. 393-403, November 2015.

[22] Y. Du, Q. Chen, X. Quan, L. Long and R. Y. Fung, "Berth allocation considering fuel consumption and vessel emissions," *Transportation Research Part E,* vol. 47, no. 6, pp. 1021-1037, 2011.

[23] J. Berechman and P.-H. Tseng, "Estimating the environmental costs of port related emissions: The case of Kaohsiung," *Transportation Research Part D,* vol. 17, no. 1, pp. 35-38, January 2012.

[24] E. Demir, Y. Huang, S. Scholts and T. V. Woensel, "A Selected Review on the Negative Externalities of the Frieght Transportation: Modeling and Pricing," *Transportation Research Part E,* pp. 95-114, 2015.

[25] M. Omran, J. Ojeil and Y. Fawaz, "Economic Impacts of Adopting a Sustainable Transport System in Beirut," Sustainable Transport Series, Beirut, 2015.

[26] N. Sathaye, A. Horvath and S. Madanat, "Unintended Impacts of Increased Truck Loads on Pavement Supply-Chain Emissions," *Transportation Research Part A,* pp. 1-15, 2009.

[27] "Site Work & Landscape Cost Data," RS Means, 2014.
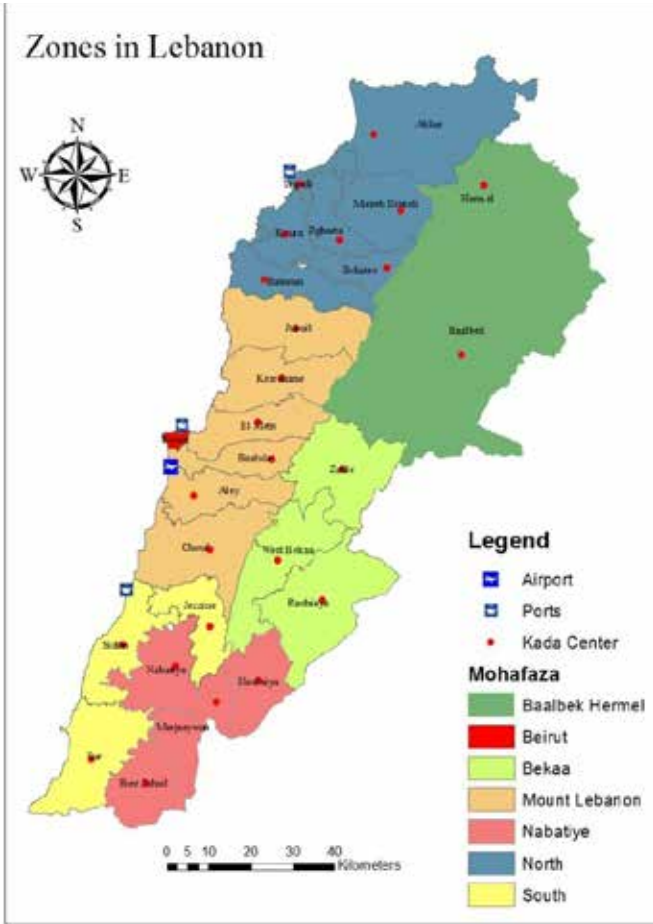
Figure A1 Distribution of Districts in Lebanon
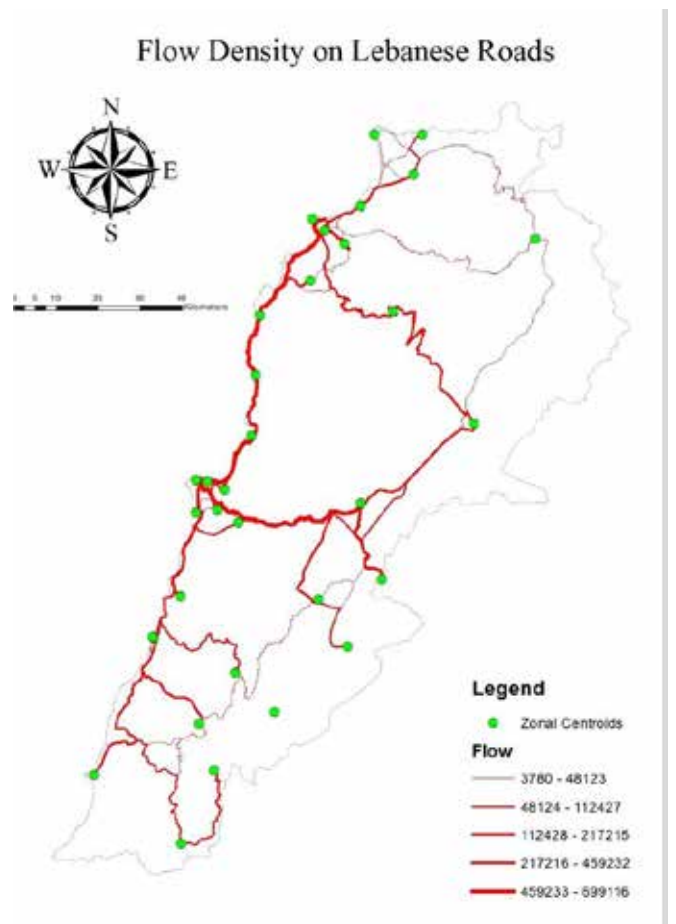
Figure A3 Flow Volumes between Zones
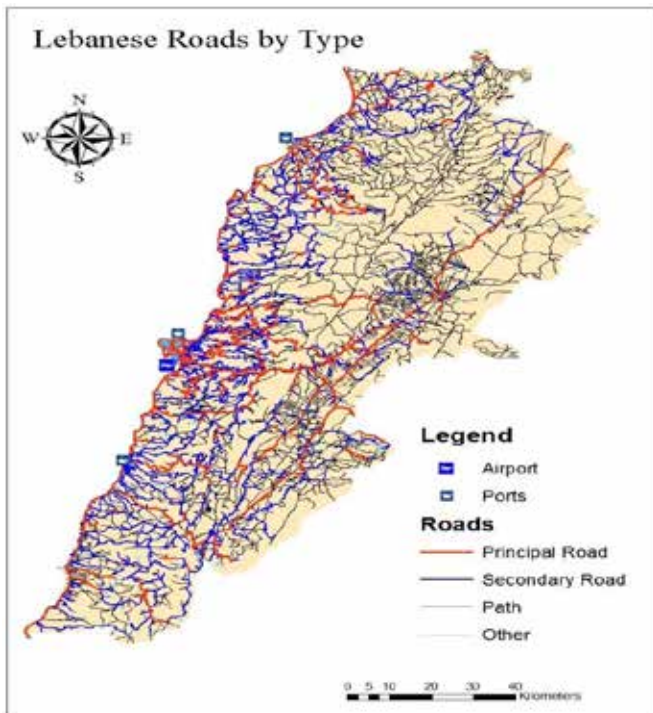
Figure A2 Lebanese Road Network
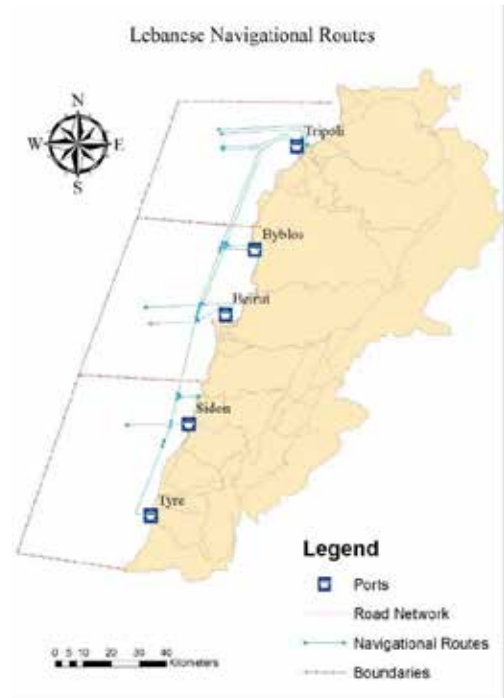
Figure A4 Lebanese Navigational Routes
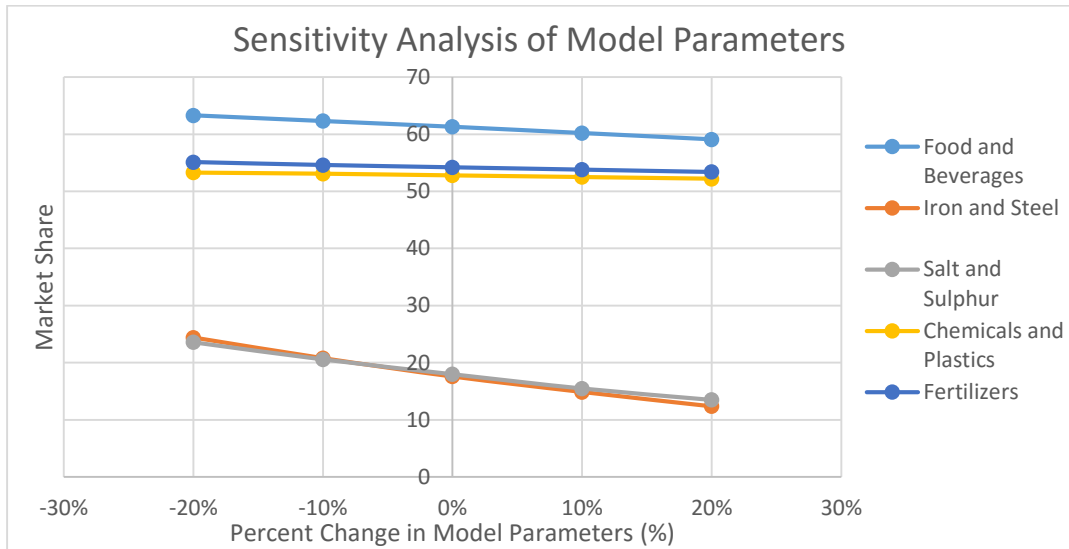
Figure A5 Sensitivity Analysis of Model Parameters

Wave Length (L) (m)	57.50
r/L (unit less)	3.58
Coefficient of Diffraction (K_d)	0.59
Wave Height (m)	4.04
Decision	wave height in the harbor is not acceptable
Breakwater Height above S.W.L (m)	14.50
Overtopping rate (m^3/s/m)	0.09

Figure A6 Breakwater Design Results and Plan View

Solid Waste Management in Nabatieh:
An Integrated Plan

Mohammad Hasan Senan - Yara El-Habbas – Reem Mehdi – Sally Smouneh – Maher Salameh
Department of Civil and Environmental Engineering
American University of Beirut (AUB)
Beirut, Lebanon.
Mas128@mail.aub.edu - Yoe00@mail.aub.edu – Rkm10@mail.aub.edu – Sms78@mail.aub.edu – Mhs54@mail.aub.edu

Abstract- **This project introduces a complete eco-friendly design of a solid waste management facility that is capable of recycling and processing the wastes generated within the Nabatieh Mohafaza in order to solve the waste current waste crisis.**

I. INTRODUCTION

Solid waste is a major issue facing developing countries worldwide. In Lebanon, this problem was never properly solved; previous governments always chose short-term solutions that led Lebanon to the current unbearable situation it is going through now. The need for a proper solid waste management plan is not only an urge to overcome the current disaster but also an opportunity to recover resources and reduce waste accumulation. The world today is heading towards sustainability and environmentally friendly projects; thus scientists and researchers are providing new concepts, ideas, and techniques that help achieve the goals set behind such projects.

Currently, Lebanon is facing a waste crisis that is not being adequately addressed due to the absence of well-designed and maintained waste management strategies. "Beirut and Mount Lebanon were flooded in trash last year when protesters blocked the entrance to the landfill until police intervened and shut down the protest."[1] These shortcomings are due to different reasons including political, economic and social issues that pushed subsequent governments into adopting non eco-friendly policies that increase the problem rather than solve it [1]. This project will tackle the main causes of the current crisis and provide a full plan of solid waste management. It targets a designated area and integrates the newest technologies in this field while taking into consideration all the technical, economic, political, and social conditions and constraints. Upon success, this plan will be a model plan to be implemented all over Lebanon in order to put an end to the current situation.

The main goal behind this study is to find the eco-friendliest techniques able to reduce the impacts of the present practice of random open dumping.

Choosing the suitable site for the proposed facility includes a detailed site selection, which compels us to check its geotechnical conditions, environmental soundness, and socio-economic setting. Excavation upon construction of the facility will provide large amounts of excavated materials that can be used in construction as well as paving roads used to connect facility buildings and the site with the major roads and highways in the area. Connecting the facility with the major roads and highways will allow for a better organization of the truck movement in the study area, without compromising on the level of service of the nearby highways. Construction management imposes proper scheduling and financial planning of the construction of the facility from the design phase of the latter until its operation.

II. METHODOLOGY

Solid Waste Management Facility Type Selection

Different technologies are present in the field of solid waste management. Some are inevitable due to environmental constraints. Each of these technologies, which will be briefly discussed, has its pros and cons.

A. Source separation

Source separation is inevitable to facilitate separation processes in subsequent facilities of the system. This may also supply recycling and composting facilities with initial volumes of recyclables and organic materials. This would be implemented by reducing taxes with every kilogram of separated wastes. A 5 % source separation percentage may increase with time as the population gets used to this policy which in turn will enhance the system's efficiency and reduce the overall cost.

B. Treatment and disposal

Treatment of waste can be done through different methods. Those methods are done in various facilities. Those include Material Recovery Facilities for both organic wastes and recyclables, remanufacturing facilities (for paper, glass, metals etc.), composting facilities, incinerators, and landfills for ashes and mixed wastes.

C. Material recovery facilities

Material recovery facilities allow mechanical separation of the collected wastes. Those may then be transferred to manufacturing or incineration. At this point, organics are routed to composting plants, while paper and cardboard, glass, and plastics are routed to remanufacturing facilities. This step in the model ensures extracting the optimum useful outcome out of the wastes and providing reusable materials at lower costs.

D. Composting

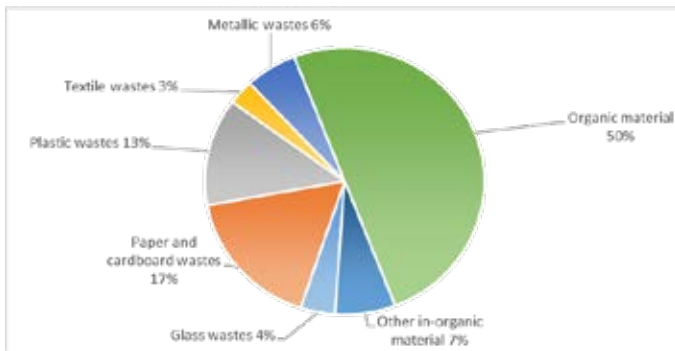Reference [2] claims that nearly 50% of the received wastes are organic wastes (Fig. 1).

After passing through pretreatment process in the material recovery facilities, composting starts simultaneously with monitoring the amount of volatile solids outgoing.

Composted wastes may be incinerated, landfilled, or used in agriculture projects. Using composted wastes instead of commercial fertilizers reduces the impacts of commercial fertilizers on the environment and decreases user costs as well. This is of course due to the low price of composted wastes compared to the commercial fertilizers.

E. Incineration

Incineration uses energy generated from wastes for heating and producing electricity, which minimizes the volume of waste. This process emits some gases and produces ashes, sludge, gypsum, and wastewater. The high moisture content of the wastes to be treated decreases the efficiency of the incinerator by limiting its productivity and increasing sludge and wastewater. Ashes and sludge require a landfill while the produced wastewater requires treatment in a specialized wastewater treatment facility. Due to these factors, in addition to a previous experience of the Lebanese community with such technology in different areas, it is the least desirable technique. This fact is reflected in the assigned weights later in the proposed model in order to choose the facility

F. Landfilling

After the separation process, materials are either routed to incinerators then produced ashes are landfilled, or routed to composting plants and then to landfills, or to the landfills directly. After landfilling, leachate is generated from the decomposition process of organic content. Leachate infiltrates and thus gets collected and treated to prevent its seepage into underground water layers below the landfill. Some gases produced can also be collected to be vented or used for combustion in engines, which supplies a portion of electricity for the facility. The size of the landfill depends on the quantity of wastes to be landfilled. Incineration process requires smaller landfill while direct land filling requires larger one.

Table 1 Criteria Used for Site Selection

Aspect	Criteria
Size	• Minimum of 1.2 Km
Topography	• Preferably in a depression (valley)
Distance From Mohafaza	• Within the Nabatieh Mohafaza
Surrounding Land Use	• Agricultural or Unallocated land
Land Use Separation Distances	• Residential Areas : 1000m separation • Airports : 9 km separation
Environmental Separation Distances	• Groundwater : 20 m separation • Surface water bodies: 150m
Site Access	• Less than 1.5km away from an existing highway.
Services	• Electricity, landline, and water services
Geology	• 100m away from a fault line • Leveled land – minimum slopes

Site Selection

Site selection process involves a set of environmental, infrastructure, planning, and social factors. These criteria are listed in Table 1.

G. Facility Design

After the selection of the type of solid waste management facility to be constructed, the design of the landfill will be modeled. However, a landfill will be needed irrespective of whether incineration is adopted or not. Yet the size of the landfill will vary depending on the alternatives chosen.

The landfill design will include the selection of materials to be used as well as the quantities of the different needed components to achieve the complete structure. After completing the construction management study, adequate modifications will be made in order to come up with a financially, structurally and technically sound solid waste treatment facility. The latter will account for long-term settlements and have a conservative factor of safety. This design will include an evaluation of the slope stability of the land where the facility will be built as well as the components of the dumping facility accounting for leachate collection, the use of geotextile, geo-membranes, and accounting for methane gas generation.

H. Optimizing transportation

The aim of this study is to find the most suitable ways to transport the produced wastes into the treatment facility without disturbing the existing transportation network. To be precise, we will consider each village of the Nabatieh Mohafaza to be a collection node. Waste generation data will be obtained as well as the frequency of waste collection.

Afterwards, the road networks involved in transporting waste from each node to the solid waste management facility will be defined. Later on, the number of the vehicles specialized for waste transportation, will be calculated, given the waste quantities and frequency of collection as well as the capacity of waste gathering per vehicle. Moving on, the effect of the waste transportation with regards to all the mentioned factors on the existing network will be determined. Alternatives that do not disturb the existing network will be formulated and the one that is more convenient in terms of cost and time will be selected.

The first step in the transportation optimization requires settling for a determined type of trucks. Based on local sources, we chose rear-end loaders of capacities ranging between 13, 10, 8.4, and 6 tons. The next step would be extracting the population of each village, and taking into account a population growth factor of 3% per year. This is due to the fact that the provided information dates back to 1998. Afterwards, we computed the waste quantity generated in every village per day by multiplying the population of each village by the mass of waste generated per capita per day according to the following simple equation:

Q_{total} = Population x Waste per capita (Kg)

Based on a fruitful interview arranged with both of NTCC's head of administration and head of technical operations, we were able to adopt the following information:

1- The dimensions of one bin are: 1.2 length 1.2 width 1.3 depth
2- The stoppage time of one truck is 2 minutes per bin
3- The average cleaning time is 30 minutes per truck
4- The average truck speed is 40 km/hr

In order to simplify the task on GIS, we divided the whole Nabatieh area into 3 zones for waste collection. In each zone, we were able to locate the combination of villages that will make up every route. This was accomplished based on the computation of total waste quantities in accordance with the capacities of trucks at our disposal while taking into account significant geographic boundaries.

In order to find the optimal route for each combination, we used Network Analyst tool in GIS.

Quantity of waste per route is the sum of the waste quantities of all the villages constituting the considered route:

$$Q_{route} = \sum Q_{per\ village} \qquad (1)$$

Then, we extracted the length of each route to and from the villages to be able to estimate the time needed by each truck to complete the tour around the targeted villages. In order to compute the time, we assumed an average velocity of 40 km/h. So the time spent on each route is obtained through the equation:

$$Time = T_1 = \frac{Route\ Length}{Velocity} \qquad (2)$$

Bin dimension: 1.2 x 1.3 x 1.2 implies a capacity of 1.12 T. The number of bins is assumed as the total waste of each route divided by the capacity of one bin. The stoppage time (according to NTCC) is 2 minutes per bin.

The stoppage time for the collection of waste from the bins is obtained by the following equation:

$$T_2 = \left(\frac{Route\ Capacity}{Capacity\ per\ bin}\right) \times 2 \qquad (3)$$

The total time for each truck is equal to the average velocity multiplied by the route length (calculated using the lengths determined from GIS) plus the stoppage time.

$$T_3 = T_1 + T_2 \qquad (4)$$

After finalizing the total time needed the routes, we determined the type and number of trucks needed for each one of them by combining the routes that need the same capacities and therefore the same type of trucks, while taking into consideration an 8hours operation time for each truck. In addition to that, we assumed a 75% efficiency rate in the operation.

Max Operating Time = 8 hrs x 0.75 x 60 min/hr = 360 min.

Based on that number and the total time needed for each route, we determined the number of trucks needed.

III. MODEL AND RESULTS

Trucks Count

Based on the calculations done, the following trucks (Table 2) were found suitable according to the available waste quantities and routes.

Table 2 Trucks Needed for Waste Collection

Capacity (tons)	Number of trucks
13	6
10	7
8.4	2
6	5

Waste management system

Model

The model presents four scenarios where different techniques are used in order to choose the most economic and efficient one. The cost of each scenario is determined per ton of wastes based on the rates of treatment obtained from Annex E of the Urban Development Series – Knowledge Papers (Table 3) published by the World Bank[4]. For calculation purposes, the upper bound of the cost range was taken.

The cost of land acquisition is common for all four scenarios since all combinations of facilities are assumed to spread over the same area. Material Recovery Facilities are required in all four cases. Those are presented in percentages of techniques used multiplied by their respective costs per ton in order to get the total rate per ton of wastes.

Table 3 Estimated Solid Waste Management Cost by Disposal Method

Technique	Cost (US$/Ton) for Lower Mid income Countries (Income/Capita = $876-3,465)
Collection	30-75
Sanitary Landfill	15-40
Composting	10-40
Incineration	40-100
Recycling	5-11.5

The cost and the social acceptance of the used technique are both considered.

A 30% waste recycling percentage policy was adopted in all scenarios. Also, collection rates were mentioned and assumed to be equal in all cases. It's worth mentioning that an additional 10% was added to the landfilling process percentage due to the presence of this technique following incineration or composting (Scenarios 2, 3 and 4). Scenarios (Fig. 2) ranged from full dependence on a landfill (Scenario 1) to a mixture of incineration, composting, and landfilling (Scenario 2) or a mixture of composting and landfilling (Scenario 3, presented in Fig. 3, and Scenario 4).

Selected Technique

Financially, Scenario 1 seems to be the most economical scenario to use and apply. However, Scenarios 1, 3, and 4 have close rates (110.45$/ton), which leads us to the social acceptance of the technique to be implemented. The community refuses options that include incinerators due to an unpromising experience [5]. This leads us to immediate discarding of Scenario 2. On the other hand, the community favors scenarios that include composting techniques due to its benefits in agriculture, which therefore excludes Scenario 1. Also, the high organic content of the municipal wastes (Fig. 1) urges us to select the scenario that makes use of this huge quantity of organic matter. Thus, Scenario3 is selected (Fig. 3).
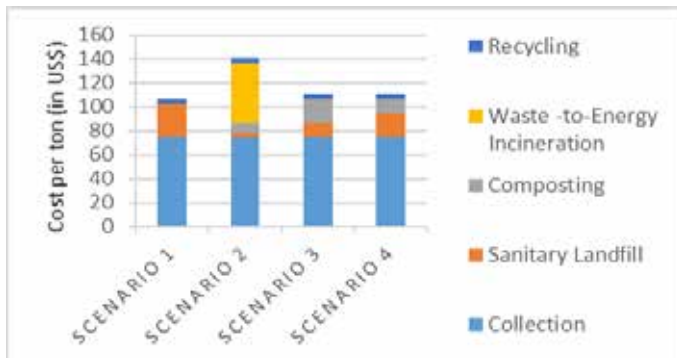
Fig. 2 Summary of Scenarios
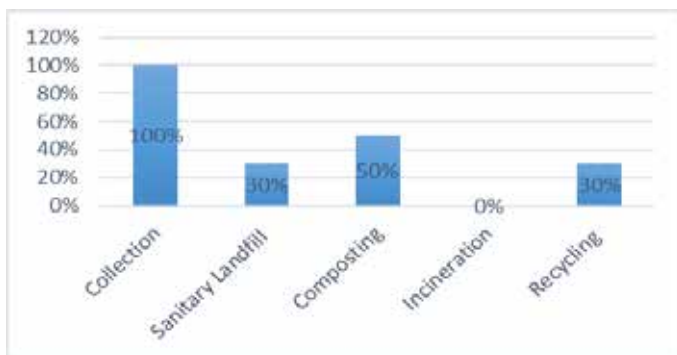
Fig. 3 Percentage Split of Scenario 3

Geospatial Mapping

According to the site selection criteria, a Geospatial Model was build using GIS software which allowed the selection of the site through several built-in tools used to analyze the available data (Fig. 9 found in Appendix 1).

This permits analyzing through combination of criteria and identifying suitable sites. The process yielded different sites. The most suitable site was then selected and identified in the figure below [6].

Fig. 4 Satellite Map of the Proposed Site

IV. FACILITY DESIGN

A. *Landfill Design*

The preliminary design of the landfill was done based on the wastes generated from all over the Nabatieh Mohafaza. The design of the landfill and calculations are detailed in the Appendix 2. At present, about 568T/day of wastes are generated in the Nabatieh Mohafaza considering a 2Kg/capita/day generation.

About 30% of the daily-generated wastes are to be landfilled after processing, which translates to 62415 T/year. The water table is found to be on a minimum of 25 meters all over the facility, thus a limit of 20 meters in ground depth of landfill was considered. Assuming an annual increase of 3% in population of the Nabatieh Mohafaza, the total volume needed for the landfill is about 85000 m^2. The landfill was designed in a rectangular shape with a length of 315 meters and width of 285 meters which yields to 89775 m^2.

The side slope of the landfill is 3:1 below ground. The landfill is designed to operate for 15 years divided to 15 phases (trenches) of equal areas (315m x 20m), each used for a period of one year. Each phase consists of daily cells, daily cover, liner, and leachate and gas collection systems (Fig. 5).

As per Environmental Protection Agency of Ireland regulations (EPA) [2], 6 gas wells are to be placed in each trench on grid separated by a distance of 45 meters. Also, 62 20-cm-diameter perforated pipes are provided at a spacing of 5 meters to facilitate collection of leachate generated in the landfill. The leachate is to be treated in a specialized treatment plant in the facility to avoid soil and ground water contamination.

341
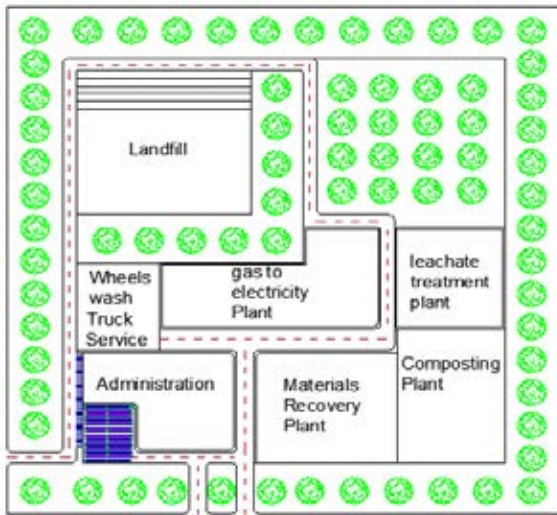
Fig. 5 Proposed Facility Layout
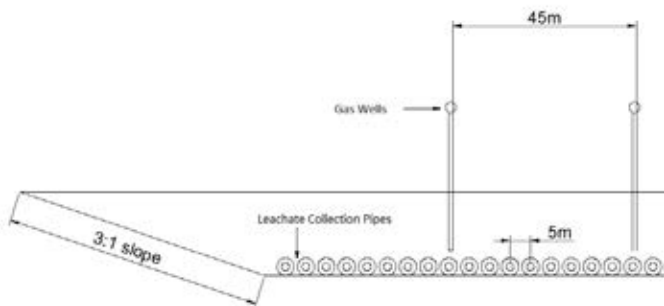
Fig. 6 Proposed Cross-Section Design of Landfill

Fig. 7 Facility Layout Projected on Satellite Map of Nabatieh Mohafaza

The landfill is provided with a final cover system that reduces infiltration and controls gas emissions. The proposed design of the landfill cross-section is shown in Fig 6.

A. *Phasing and Cells*

The landfill should be divided into multiple phases, each of 12months life span. Cells are smaller divisions of phases that are separated by cell bunds. Minimizing the cell size reduces leachate generation and decreases the thickness of needed cover.

B. *Bunding*

Bunds are used for dividing a phase into multiple cells as separators. A 2m height bund is located on the base of the landfill. The primary liner is to be placed under the bunds, separating the cells in order to prevent leakage of leachate.

C. *Covering*

The cover is an important element in constructing and operating a landfill. Cover decreases odor and possibility of flies and birds gathering on wastes. A 15 cm layer of soil cover is put over the landfilled wastes at the end of each day.

D. *Landscaping*

The facility should have a good, clean, and acceptable appearance. Trees should be planted all over the project sides to attenuate the negative environmental impacts of the facility. The fence of the facility should be lined up with tall trees that resist wind and prevent odor penetration.

E. *Liner*

A landfill liner acts as a low permeable barrier intended to prevent leachate from migrating into the ground and polluting ground water. It is located at the bottom of the landfill site. There are two types of liners: compacted clay liners and geosynthetic clay liners. The type adopted for the design of the Nabatieh Mohafaza landfill is the geosynthetic liner. Geosynthetic clay liners present advantages over compacted liners. They are easy to construct and install, cost less than compacted clay liners, and help save on construction time of the landfill. Also, due to the fact that they are prefabricated, geosynthetic clay liners are uniform throughout the entire length of the liners; there are no variations in performance between the different spots.

A liner system is composed of three parts: the geosynthetic liner and 2 layers of geotextile and gravel where the leachate pipes are present. The geosynthetic clay liner is directly installed on a smooth subgrade to prevent any rock from penetrating into the liner. As a matter of fact, the particle size in the first 150 mm of subgrade should not exceed 12mm, with 80% of the particles finer than 1mm. The geosynthetic clay layer is manufactured and installed directly on site. It is made of bentonite and is 8mm thick. The most important feature of the geosynthetic clay liner is its hydraulic conductivity which is the rate at which a liquid passes through a material that should range between 10^{-12} and 10^{-10} m/s. The low hydraulic conductivity results in a low permeability which slows the rate of seepage out of the landfill and thus controls leachate percolation.

Fig. 8 Liner System

Bentonite is used because it is an extremely absorbent clay. It is used in geosynthetic clay liners due to its ability to swell in volume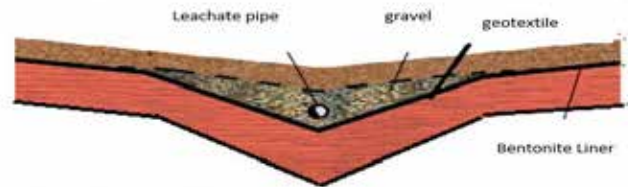 as it absorbs liquid particles thus forming a barrier between the waste and the ground when in contact with leachate. The geosynthetic bentonite layer is fixed between two geotextile layers. Each of the geotextile layers is 3 mm thick woven or non-woven sheet materials made of polyester used in order to provide tensile stress and protect the liner from mechanical stresses. Also, they provide support and strength to the liner which becomes vulnerable due to erosion, wind and disturbance by workers during installation. The type of geotextile chosen is woven in order to be able to connect the geotextiles to the liner by needle-punching, which is only possible with woven material. Needle-punching provides additional internal shear strength, acting as a reinforcement to the liner system.

F. Site Layout

The facility design should note the ease of movement of trucks and personnel between the different buildings and parts of the facility. The facility has three gates, of which two are restricted to waste collection trucks and the third is restricted to staff automobiles that leads into restricted parking that serves around 400 cars. The stalls in the parking and roads connecting plants in the facility are designed according to the Lebanese standards of parking. The trucks enter the facility and empty in Material Recovery plant where wastes are transferred into composting or landfilling after processing. A composting plant is located right next to the Material Recovery plant. Wastes to be landfilled are then transferred to the landfill area that is connected to the Leachate treatment plant and Gas to Electricity plant. Trucks then proceed to wheel washing and servicing (if needed) then park in their designated parking lot that serves up to 38 trucks.

G. Gas to Electricity Plant

Methane Production is estimated to increase as the project proceeds in its useful life. The yearly estimated produced amounts of methane are presented in Table 2 and Fig. 7. Those estimations are based on calculations presented in the Appendix 2. The projected amounts could be used to generate electricity for the facility and supply the Mohafaza needs of natural gas for chillers and heaters.

Table 4 Methane Estimated Emissions in Project's Useful Life

Year	Methane Emissions (m^3/year)	Year	Methane Emissions (m^3/year)
2017	0	2025	6723000
2018	2839000	2026	6935000
2019	4334000	2027	7149000
2020	5165000	2028	7366000
2021	5667000	2029	7588000
2022	6010000	2030	7816000
2023	6276000	2031	8051000
2024	6507000	2032	8293000
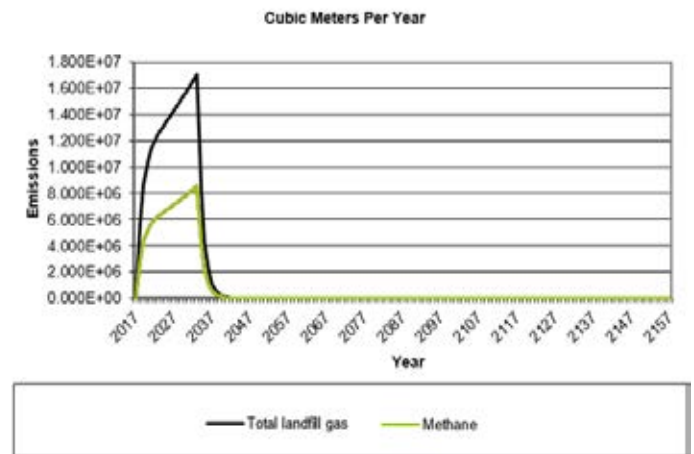
Fig. 9 Total Gas and Methane Emissions from Landfill

V. CONCLUSION

The presented waste management system aims to prevent the accumulation of wastes on roads using advanced eco-friendly techniques. Nabatieh Mohafaza was selected due to the random waste disposal. However, this model can be implemented in other areas as well. The model took into consideration all the required facilities and equipment needed as well as optimizing their usage to prevent failure in other inter-related systems such as the existing transportation networks. The system solves the current crisis and supplies part of the Nabatieh Mohafaza's gas needs.

REFERENCES

[1] Sidahmed, M., "Landfills or Incinerators for Lebanon?" McClatchy - Tribune Business News. http://search.proquest.com/docview/1643403056?accountid=8555, 2015.

[2] The Environment Agency, "How Do We Regulate Landfill Sites?" http://apps.environment-agency.gov.uk/wiyby/37823.aspx, 2015.

[3] United Nations Statistics Division, 2009.

[4] World Bank, "Annex E Urban Development Series". http://siteresources.worldbank.org/inturbandevelopment/resources/33638 7-1334852610766/annexe.pdf, 2015.

[5] Abou Najm M., El-Fadel M., "Computer-Based Interface for an Integrated Solid Waste Management Optimization Model, Environmental Modelling & Software, Vol. 19", 2004.

[6] Sadek S., El-Fadel M., Freiha F., "Compliance Factors Within a GIS-Based Framework for Landfill Siting", International Journal of Environmental Studies, Vol. 63, Issue 1, 2006.

Appendix 1

Fig. 9 Total Gas and Methane Emissions from Landfill

344
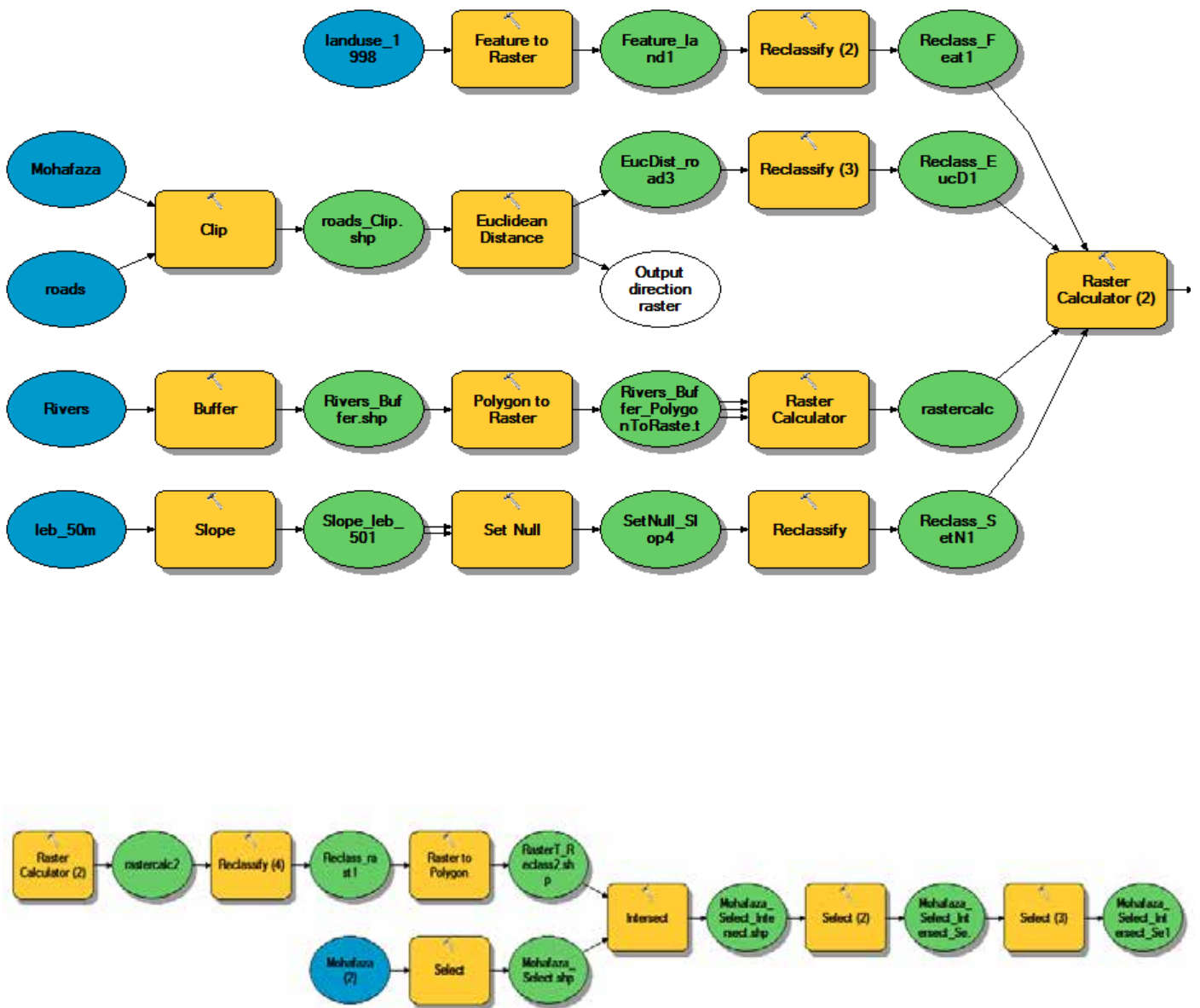
Appendix 2

Basic data

- Location: Nabatieh Mohafaza
- Waste: 570 Tons per day (current)
- To be landfilled : 30%
- Design Life: Active Period = 15 years
- Topography: Flat ground (of selected site)
- Water-table: 25 meters below ground surface

Landfill capacity, height and area

(a) Current (to be landfilled) waste generation per year, W = 62415000 kg

(b) Estimated rate of increase of waste generation per year (rate of growth of population), x = 3%

(c) Estimated waste generation after 15 Years
$$= W (1 + x)^n \qquad (5)$$
$$= 62415000 (1+0.03)^{15}$$
$$= 97240536.31 \text{ kg}$$

(d) Total Waste Generation in 15 Years, T
$$= 0.5 (W+W (1+x)^n) n \qquad (6)$$
$$= 0.5(62415000+62415000 (1+3\%)^{15} *15$$
$$= 1197416.522 \text{ T}$$

(e) Total Waste Volume (assumed density 0.85 t/m³),
$$V_w = T /0.85 \qquad (7)$$
$$= (1197416.522) /0.85$$
$$= 1408725.32 \text{ m}^3$$

(f) Volume of Daily Cover,
$$V_{dc} = 0.1 V_w \qquad (8)$$
$$= 0.1*1408725.32$$
$$= 140872.532 \text{ m}^3$$

(g) Volume of Liner and Cover Systems,
$$V_c = 0.2 V_w \qquad (9)$$
$$= 0.2*1408725.32$$
$$= 281745.0641 \text{ m}^3$$

(h) First Estimate of Landfill Volume, C_i
$$= V_w + V_c \qquad (10)$$
$$= 1408725.32+281745.0641$$
$$= 1690470.384 \text{ m}^3$$

(i) Possible Maximum Landfill Height = 20 m

(j) Area Required = 1690470.384/20
$$= 84523.52 \text{m}^3$$

(k) Approximate Plan Dimensions = 315m x 285 m
$$=89775 \text{ m}^2$$

Landfill section and plan

Landfill Section and Plan is evaluated by providing a side slope of 3:1 below the ground surface. The material excavated from the site is used as the material for daily cover, liner and final cover.

Landfill Phases

(a) Active life of landfill = 15 years

(b) Duration of one phase = one year

(c) Number of phases = 15. (Each phase extends from base to final cover.)

(d) Volume of one phase = landfill capacity/15
$$= 1690470.384 / 15$$
$$= 112698 \text{ m}^3$$

(e) Plan area of phase = 315 m * 19 m

(f) Number of daily cells = 365

(g) Plan area of one cell (on the basis of 1m lift of each cell)
$$= (\text{Volume of one phase})/365$$
$$= (315*19*20)/365$$
$$= 328 \text{ m}^2$$
$$= 17.25* 19 \text{ m}$$

Methane Generation

First-Order Decomposition Rate Equation:

$$Q_{CH_4} = \sum_{i=1}^{n} \sum_{j=0.1}^{1} kL_o \left(\frac{M_i}{10}\right) e^{-kt_{ij}} \qquad (11)$$

Where,
- QCH4 = annual methane generation in the year of the calculation (m3/year)
- i = 1-year time increment
- n = (year of the calculation) - (initial year of waste acceptance)
- j = 0.1-year time increment
- k = methane generation rate (year-1)
- Lo = potential methane generation capacity (m3/Mg)
- Mi = mass of waste accepted in the ith year (Mg)
- tij = age of the jth section of waste mass Mi accepted in the ith year (decimal years, e.g., 3.2 years)
- Model Parameters from User Inputs:
- k = 0.700 year-1
- Lo = 96 m3/Mg

StPaul Youth Village - Lehfed

Program Development and Technical Design

Karl Eid, Alain El Khoury, Jad Khalil, Reham Musaed and Theresea Yazbeck

Civil and Environmental Engineering Department

American University of Beirut

Riyadh El-Solh, Beirut, Lebanon

This project is a community-based project that consists of the program development and technical design of a Youth Village, bound for implementation, in Lehfed, in close coordination with Saint Paul – Charity Mission, a youth movement with a wide outreach in Lebanon and abroad. Through a participatory approach, the team followed an inclusive design for individuals with special needs and collaborated tightly with technical and social experts. The team was able to significantly intervene in the development of the Youth Village to include a Youth Facility, a Children Facility, Playgrounds, a Forest, a Garden and a Walking/Bike Trail. And to support these components, suitable infrastructure was developed, while strongly integrating sustainability into the overall project and to each component as well. Additionally, technical engineering design was carried out and included a detailed structural design for the Youth Facility planned for construction at the initial phase. As a final step, the team calculated the approximate cost of each component in the Youth Village and the cost of the overall Village and suggested a phased construction of the project. All in all, the team made fundamental modifications/improvements at the concept design and detailed design levels that will assist Saint Paul in positively and aggressively promoting the Youth Village's viability and funding. Finally, the Youth Village's process of development and deliverables outcome serve as inspirational models for blending academic learning with civic engagement.

I. INTRODUCTION

This paper presents a community-based project consisting of the program development and design of an actual Youth Village in Lehfed, Jubail Caza, in partnership with Saint Paul – Charity Mission (known as StPaul), a forceful youth movement that exerts itself in championing the cause of love and brotherhood in the world, outreaching to thousands of individuals in Lebanon and abroad on a yearly basis. The movement is increasingly growing, currently reaching about 500 active members from all over Lebanon, and, in order to foster its growth and advance its service to the community, it has already initiated the project and acquired a 10,842 m^2 land for it, seeking to take the project from the planning phase to the construction phase in the current and upcoming years. The Youth Village will be the first center in Lebanon to be professionally designed and dedicated to educational, spiritual and recreational youth activities. Composed of a Youth Facility, a Children's Facility, Playgrounds, a Forest, a small Bike/Walking Trail, and a few Bungalows, the project will be able to host primarily youth, children and even families through camps, day trips, retreats, meetings, workshops, fundraisers and conferences. Before the team's intervention, with the exception of the Youth Facility for which architectural drawings had been prepared and for which construction works are to start during the upcoming summer, the project's facilities were still mainly proposed at concept levels (Fig. A.1).

In this project, the team worked on the different aspects of the Youth Village while maintaining close coordination with StPaul. Firstly, the team developed the program of the Youth Village based on a participatory approach that involved the various stakeholders. Secondly, in alignment with StPaul's aspirations, a key target of the project was sustainability that was assiduously considered in laying out the facilities, in elaborating the landscape, and in developing the infrastructure: power supply, water supply, wastewater treatment, solid waste management, circulation and parking. In assessing alternatives, financial viability was also taken into account. Moreover, sustainability was implemented at the level of the construction of the project through different practices, and at the level of its operation by developing efficient systems as well as awareness guidelines and activities. Thirdly, attention has been given to accommodate individuals with special needs in the Youth Village, by relying on a universal and inclusive design for its indoor facilities and outdoor landscapes. After a preliminary design was reached for the overall Youth Village, technical engineering design was carried out and included a detailed structural design for the youth facility planned for construction at the initial phase. Lastly, the team underwent an initial costing and phasing for the project with the help of an experienced contractor. Throughout this paper, the different disciplines in Civil and Environmental Engineering have been noticeably integrated to the Youth Village, allowing a deeper appreciation of the profession from a sustainable and community engaged perspective.

II. APPROACH

The team had first to understand the concept and subtle particularities that StPaul aspires to have in the Youth Village. The team also had to trace StPaul's development of the project's different components hitherto. This was mainly done at the initial stage of the project through continuous communication with the board of StPaul and the project's lead architect who is also a member of the Charity Mission.

In order to involve the users and better cater for their needs, the team adopted a participatory approach for the design of the Youth Village:

- The team presented the Youth Village to members of StPaul, discussed it with them, and around 30 experienced members filled a paper-based survey that inquired about concerns faced in previous facilities and about key sustainable points proposed by the board and the team (Member information and answers are shown in Table. A.I, Table A.II and Table A.III).

- The team visited parishes in Daychounieh, Dekwaneh, Dbayeh, Deir al Qamar and Ferzol during StPaul weekly activities and asked about 140 kids to draw mental maps of their dream playground (Fig. A.2).

- The team visited the village of Lehfed and met with villagers in order to collect their feedback. Their comments were very positive and supportive.

Based on the above, the team and the architect introduced new elements and several modifications to the project.

All throughout its multidisciplinary development of the Youth Village, the team referred to experts in the fields of Civil, Environmental, Architecture, Landscape and Electrical Engineering. These meetings were thoroughly documented as a solid base for any future development.

III. PROJECT PLANNING AND DEVELOPMENT

The team worked on the planning of the whole Youth Village. Sustainable development was conducted for each facility. In addition, all facilities were modified and designed so to account for people with special needs.

A. Building Facilities

Youth House

The Youth House design process was already underway before the team's involvement in the project. Hence, upon the team's engagement with StPaul, the team was given the architectural drawings of the Youth House (Fig. 1). This presented the team with potential challenges including some reduction in flexibility with regard to the structural system and sustainability (since the massing and orientation had already been decided). It was to the team's advantage that environmental responsiveness was in line with the architect's initial intent. With the massing and orientation of the Youth House already set, the team checked whether the orientation of the facility was in accordance with the sun path diagram of the region; the architect oriented the building southwards as to reduce insolation on vertical facades in order to reduce the facility's heat gain in summer. The team also recommended the integration of overhangs on the southern façade to protect against the sun at mid-day during the summer while allowing it in in the winter. The integrated passive design strategies will help reduce the need for heating and cooling through

mechanical systems during the operation of the building.

The team aimed to decrease the consumption of fossil fuels through the use of energy efficient systems. Moreover, the installation of solar thermal panels on the roof, to heat the water was studied and recommended.

Due to the remoteness of the site, the mission would have to buy cisterns from the local domestic water provider to supply its needs. This action neglects the availability of natural rainwater which falls on the Youth House. The team studied and recommended the use of rainwater.

The wastewater output must also be treated and managed. Greywater and blackwater will be separately treated. The greywater accumulated from the showers and sinks will be treated and reused for toilet flushing, and its excess can be used for irrigation. The blackwater will be treated through a compact wastewater unit that is discussed in more detail later on in the paper.

The accessibility of the facility to people with special needs was given special attention by the team as this is becoming common-practice in "sensitive" projects. The original architectural drawings were modified with a specialist to allow easy access, mobility and living in the Youth House. This has been achieved through the installation of an elevator for wheelchair users (Fig. A.5) in addition to guidance signs, special furniture, bathrooms and bedrooms as well as parking spots.

Children House

The Children's Facility is the largest facility in the Youth Village, and it is expected to house around 150 persons at a time. However, due to its large scale, its implementation is postponed until the construction of the Youth House is completed. It is still in its conceptual design stage, therefore allowing for more flexibility in the design.

The first concept design for the Children House was two parallel buildings having their longer façade directed in the north-south direction in accordance with the topography. However, this orientation was not optimal for passive solar design. After conducting the survey, the team worked with the lead architect to change the design to one L-shaped building to incorporate the different program alterations such as an amphitheater, an indoor playground, a library and other features. Also, the design was made inclusive of people with special needs (Fig. 2 and Fig. A.3).

The climate responsive measures that were implemented in the Youth House will be applied to the Children House as well.

B. Landscape and Outdoors

The program development of the Youth Village considers the landscape to serve not only an aesthetic, but also

Figure 1. Youth House

Figure 2. Children House

educational and ecological purposes. The landscape was divided into four main components.

The Walking/Bike Trail

The internal circulation within the Youth Village was planned and designed in a way that prioritizes pedestrians and bikes over cars and buses. Therefore, the team designed the horizontal and vertical alignment of a walking/bike trail on AutoCAD Civil 3D in order to connect all of the Youth Village (Fig. A.4). The trail was integrated into the forest so as to make use of natural shade. It has a width of 3m and a length of around 300m, with resting spots distributed every 50m. In spite of steep topography, the team was able to limit the slope of the trail to 8% to accommodate people with special needs while also minimizing cut and fill. The team provided the trail with a superelevation of 2% and a side channel in order to collect rainwater, serving as an example of sensitive site treatment. Concrete tiles were suggested for the trail since they are suitable for wheel chair users, are of relatively low cost and require little maintenance. For the retaining walls, rock from the facility excavations will be used in order to use materials from the site and reduce costs as initial resource, labor and transport.

The Playgrounds

Based on the participatory exercises, the team was able to identify the need for a basketball court and a mini-football court that were designed and located in suitable areas around the Children House. Additionally, the team was able to identify the need for a playground specifically tailored for little kids. Inspired by the playground implemented by CatalyticAction, the team collaborated with a colleague from the architecture department at AUB for its design.

The playground was designed to be inclusive. It was also positioned and designed so as to provide shade for the kids all throughout the day. The material for the playground was chosen as wood so that the members could build the playground themselves. Also, this facility will have an educational aspect for the kids, where various educational, spiritual and intellectual messages will be written in key positions in the playground.

The Forest

A large number of members and kids communicated the need for green spaces for activities. The team suggested the idea of a forest that could be planted by the youth and the members of StPaul themselves. Ecosystem preservation was one of the main concerns, and the importance of using native or naturalized flora was focused. The team researched the AUB landscape database [1] and discussed its findings with a landscape horticulture expert to select the proper trees. The

main criteria for the team's choice were the amount of shading provided, minimum pruning requirements and a low water demand. Non-bearing trees were more feasible than fruit trees, and the selection was narrowed down to the following: Judas's tree, Red Maple, Swamp Maple and Turkish Pine.

The Garden

The Garden was suggested for educational purposes where the youth could plant and grow edible vegetables and fruits. Due to limited space, the objective of this area was not to provide self-sufficiency, but rather to serve as an educational activity area. The northern side will be planted with fruitful trees, while the southern part will be composed of raised beds for growing seasonal vegetables. This layout was suggested based on sun considerations. The fruit trees suggested were the following: Ziziphus Jujube, Apricot, Apple, Peach, and Pear (Table A.4). The diversity included in the fruit tree selection provides a wider range for education. For the vegetable area, the technique of raised beds is adopted for the delimitation and protection of the planted areas and for its low implementation cost. The plantations were divided into two categories based on their growth season: spring and summer. In spring, the raised beds will be used for growing special herbs from seeds. As for the summer season, the spring herbs would be collected and vegetable seedlings would be planted. Lastly, it is important to emphasize that the Garden needs more maintenance than the Forest, not forgetting the need for natural fertilizers. This maintenance is required only in the summer season, when the Youth Village witnesses a high level of activities.

C. Infrastructure

In order to support the facilities, the team sought to devise sustainable infrastructure in the Youth Village.

Power Supply

In collaboration with individuals in the field, the team focused its analysis on the Youth House. Initially, the team reduced the load through energy efficient systems using metal halide lights, appliances certified by entities such as Energy Star, solar thermal water heating and through natural ventilation, important in slashing the energy bill, alongside small awareness signs. Then, the team calculated the power demand for the Youth House and investigated 4 alternatives in order to assess the feasibility of using solar energy. The analysis was mainly financial, and the 4 alternatives were: Stand-alone PV Panels, EDL Power Supply with PV panels, EDL Power Supply with Diesel Generators, and Stand-alone Diesel Generators. Present worth analysis was performed, the study period was taken as 20 years (equal to the lifespan of the PV panels), and the interest rate was taken as 4%. Initial costs,

Figure 3. Kids Playground

Figure 4. Youth Village Development Overview

operational costs and maintenance costs were approximated based on preliminary designs and on current private and public sector prices. The costs are shown in Table A.V.

Power demand is modest during the winter and fall seasons, and highest during the summer season when StPaul's activities peak. This variation considerably affects the design of PV's since additional batteries and panels will have to be provided for the summer season, while they will be left unused during most of the year leading to a waste of resources. Additionally, this will lead to a low average consumption of fuel over the year. This explains the relative high costs of PV alternatives when compared with generator-based alternatives. While typically, PV panels prove justifiable on the long-run, for the Youth Village's current occupancy pattern, the team's study shows that it would be more sustainable and financially feasible to invest in connecting the Youth Village to the public grid and to use diesel generators during rationing outages. And since the demand is already considerably reduced, the diesel generators used can be of relatively small size, they would be turned on only when needed, and their emissions filtered. Nonetheless, for a sustainable development, the team suggests that StPaul makes use of the Youth Village efficiently during winter as well: if not through its own activities, then by allowing schools, parishes or other movements to benefit from it as well; what would significantly alter the occupancy pattern of the Youth Village and restore to solar energy its long-term advantage.

For outdoor lighting, the team recommends the use of stand-alone PV panels mounted on the light poles to reduce demand, cable installation costs, maintenance costs and power losses.

Water Supply

Firstly, the team reduced daily consumption by installing water aerated faucets and dual-flush toilets and by promoting awareness through small signs. The demand for domestic water was initially calculated for both facilities. However, domestic and potable water are not supplied by the municipality at the site. So the team studied and proposed environmental alternatives to provide for water supply.

Rainwater will be collected from the roof of the Youth House, the roof of the Children House, Playgrounds and the Bike/Walking Trail, so that StPaul would have benefited from the location of the Youth Village which is situated in the area having the highest yearly rainfall in Lebanon [2]. Water collected from the roofs will be dedicated for domestic use, after applying filtration to it since minimal air pollution exists in the project's area.

For the Youth House, the team calculated the water storage space needed to benefit fully from rainwater harvesting, and it proved the need for additional water tanks for which the team suggested key positions. Harvested water from bleachers, playgrounds and bike trail can be used for irrigation. Its use for domestic applications could be feasible but requires further study.

Additionally, collaborating with experts in hydrogeology and looking into hydrogeological maps [3], the team confirmed the feasibility of a water well to provide the Youth Village with the remaining of its water needs, in spite of the possible need for multiple trials due to karstification. Furthermore, during the site visit, the team learned of several domestic wells that hit underground conduits at about 200m depth.

Wastewater Treatment

After assessing wastewater disposal practices in Lebanon with an environmental consultant, the team identified that it would not be suitable for the Youth Village to follow common practices that rely on open septic tanks. The team determined that a compact wastewater treatment unit would be the best solution for the facility. The system includes 3 compartments: aeration, a trickling filter and chlorination/ozonation. The system will result in low BOD water that can be used for irrigating.

Solid Waste Management

First of all, the team reduced waste generation by recommending reduce and reuse policies. Moreover, the Village will integrate a simple form of sorting, recycling and composting, seeking a zero-waste environment.

After meeting with local solid waste specialists, three types of labeled bins will be spread throughout the facility: organic material, recyclables, and rejects. For the recyclables, a specialized company will collect them for processing as they include glass, plastics, metals and papers. Zero-Waste Act is an active recycling company that collects sorted waste from the Jubail Caza and was recommended to the team.

Composting will be used to treat organic wastes on site at minimal costs. The team decided to implement the wooden box system for its high capacity, ease of use, and construction which uses pallets that are construction material wastes; its design incorporates wooden planks, an interior wire mesh and a cover to protect it from snow and rain. The volume of organic waste supplied by the Youth House was calculated and its actual implementation was studied.

The daily generation of organic waste is assumed 0.6 Kg/day/person with a density of 500 Kg/m^3. A maximum of 50 people will be using the Youth Facility for a maximum 135 days per year. These assumptions lead to a yearly volume of organic waste of 8.1 m^3. The team suggested to install four wooden boxes with 1.5x1.5x1.5 m^3 dimensions, knowing that the boxes should be at maximum 75% full for maintenance purposes. The wooden boxes need minimal upkeep. The compost will be used to improve the soil in the Forest and the Garden.

Circulation and Parking

The two vehicle roads, already fixed by the authorities, and shown in Fig. 4, were used for parallel parking, allowing for approximately 70 parking spaces. The upper road was connected at its end to the basketball court in order to use the court for parking in peak times instead of increasing the footprint of the project just to accommodate rare instances of peak parking demand. The court will fit 30 cars in total.

Moreover, a parking space for people with special needs was also provided for each facility. The total number of parking spaces provided is 100.

D. Special Needs

For the facilities, the Youth House was modified to accommodate individuals with special needs. As for the Children's House, in collaboration with an expert in inclusive design, the team reviewed its proposed architectural drawings and suggested adjustments.

The trail that will connect the Youth House to the Children's facility was also based on the principle of inclusion and hence the slopes adopted do not exceed 8% to allow safe and easy commuting in the Youth Village.

The playground that was designed by the team includes ramps, elements of play, chairs and tables [4]. Through its universal design, the team sought to institute social equity among the kids.

IV. ENGINEERING DESIGN DEVELOPMENT

A. Geotechnical Testing

A preliminary geotechnical investigation was required to assess the ground which will support the different facilities. Therefore, the team underwent an initial desk study. It showed that Lehfed's rock was most likely made of C4 Sannine limestone which could harbor some fissures as well as abundant karstification features based on Dubertret's maps [5]. Therefore, the team conducted a site visit where very shallow rock was apparent with a dominance of limestone and conglomerate rocks. The team brought three rock samples from different parts of the site as well as soil samples. The rock samples were tested in the AUB lab giving satisfactory results for unconfined compression tests. As for the soil, it was identified as clayey. In spite of the reassuring results considering the relatively small loads that will be carried by the foundations, a conservative estimate of the allowable strength ($q_{allowable}$) was assumed to be 3 kg/cm^2 after test results were discussed with specialists in the geotechnical field. In addition, the team surveyed several residents of Lehfed, the villagers affirmed that they faced no particular issues neither with landslides nor cracks.

B. Structural Design of Youth House

With the guidance of a design office and professionals in the field, comprehensive gravitational and seismic structural analysis was done for the Youth House via software and manual verifications according to ACI 318-11 [6] and ASCE 7-10 [7]. The basic assumptions and design criteria, are:

- Concrete compression strength (f'$_c$) is 25 MPa and steel yielding strength is 420 MPa.
- Super imposed dead loads (SDL) were taken as 500 kg/m^2 and live loads (LL) as 300 kg/m^2 for both horizontal and inclined slabs.

The main challenge was the continuous coordination and iterations with the architect and MEP, which were an essential dimension of the project planned for implementation. As a result of the meetings held and the structural analysis done for the building, a feasible and optimized structural design was provided to StPaul along with structural execution drawings.

Slabs
a) Typical Slabs

The Youth House consists of one slab on grade, three horizontal suspended slabs and an inclined roof. StPaul was considering having hourdi slabs considering that it has a lower construction cost compared to solid slabs. But after taking a closer look at the sunken slabs imposed by the bathrooms, solid slabs seemed to be more practical. Sunken slabs imposed the addition of several beams to the hourdi slab which complicated its constructability and made it more costly. Serviceability and strength requirements were investigated, resulting in an optimized slab thickness of 25 cm. SAFE software version 12 was used for slab ultimate strength design and for serviceability checks. A two-way bottom mesh of T14 at 15 cm spacing and a two-way top mesh of T12 at 15 cm spacing were used with some additional reinforcement over the columns.

b) Inclined Roof

Due to sophistication in the design of the inclined roof, SAP2000 version 18 was used. Fig A.6 shows the moments about the lateral direction.

Vertical Elements

A preliminary analysis of tributary areas was manually conducted using an excel spreadsheet in order to minimize the number of iterations for vertical elements design. The dimensions obtained by manual calculations were confirmed by checking the demand over capacity ratio for each column. Consequently, a model was created using ETABS version 15 (*Fig. 5*), which allowed the consideration of seismic design.

A shear wall (R=4.5) system was adopted for lateral loads. As part of the optimization process, the basement walls had their sizes reduced from 25 cm to 20 cm.

Footings

A sensitivity analysis was conducted in order to decide whether to have a raft foundation or isolated footings by comparing the corresponding quantities. The footing design was done using SAFE and checked manually. Isolated footings proved to be the choice.

C. Structural BOQ

A Structural BOQ requires detailed drawings, which were only available for the Youth House. The BOQ included the quantities of steel, concrete and water proofing materials in addition to their workmanship and construction cost. Based on
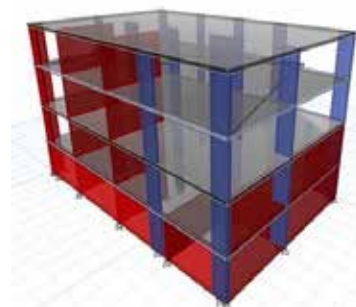

Figure 5. Youth House ETABS Model

rates that were taken from an experienced contractor, the structural BOQ cost for the Youth House is estimated to be around 85 000$.

D. Rainwater Collection

Accurate calculations for water demand and supply were done at the level of the Youth House. Assumptions were taken into consideration:

- Annual rainfall in Lehfed is 1400 mm based [2]. In order to account for climate change, annual rainfall will be considered as 1200 mm.
- Demand = 150 liters/person/day.
- 50 persons will be using the Youth House.
- The Youth House will be used three full months during the summer and five days per month in the remaining parts of the year.

The annual demand for the Youth House is 1013 m^3, while the annual supply from harvested water from the roof is 290 m^3 i.e. around 30% of the demand. The needed water storage was estimated to be 142 m^3. However, the available water tank in the basement provides a storage of 56 m^3. Therefore, there will be a need for an additional water tank. The team recommended having an underground rectangular reinforced concrete water tank to the southern side of the Youth House with an area of 5 x 4 m^2 and a depth of 5 m. It would be a onetime initial investment from which the Youth House would benefit from significantly on the long-run.

As for the remaining components, only the demand can be currently calculated. Then, based on the previous assumptions, the total water that could be harvested is summarized in Table A.VI. The total amount of water that could be harvested from the whole Youth Village is around 4805 m^3.

V. BUDGET AND PHASING

After consulting an experienced local contractor, a rough cost for the construction of the Youth Village was estimated. The table below shows the estimated cost for each facility in addition to the total cost of the project. For the facilities a price 1000 $/m^2$ of built-up space was used, and special considerations were taken for each feature.

The Youth Village will be constructed under several phases, with staged financing. First, the Youth House will be constructed since it is the smaller of the two facilities and will allow StPaul to start using the Youth Village early on while also promoting it. Provided with financing, its construction is expected to take two years, with the forest planted in parallel. Following which, the Children House along with the playgrounds and garden are planned for the following 5 years. The final step would be the construction of Bungalows which is considered a long term project for StPaul.

TABLE I. ESTIMATED CONSTRUCTION COST OF EACH FACILITY

Facility	Construction Cost ($)
Youth House	864,000
Children House	4,000,000
Basketball and Football Court	100,000
Children House Playground	28,000
Bike Trail	235,000
Total Cost	**5,227,000**

VI. SUMMARY AND CONCLUSION

Relying on a participatory approach and on close coordination with experts, the team alongside StPaul planned, designed and developed the Youth Village's Facilities, Landscape, Infrastructure and Systems, deeply embedding sustainability into the overall Youth Village and into each component, while ensuring an all-inclusive design. The team made fundamental modifications and improvements at the concept design and detailed design level. Consequently, blending academic learning with civic engagement, StPaul's Youth Village will serve as a model which process of development and deliverables outcome can be motivational. The project will finally serve as an essential component in promoting the project and fundraising for its further development and complete realization.

REFERENCES

[1] S.N. Talhouk , M. Fabian, R. Dagher, Landscape Plant Database. Department of Landscape Design & Ecosystem Management, AUB, 2015 (http://landscapeplant.aub.edu.lb, 29 Mar 2016).

[2] *Atlas climatique du Liban*, Service Météorologique, Ministère des Affaires Publics, Beirut.

[3] *Carte Hydrogeologique du Liban,* United Nations, 1967.

[4] *CAE (2010). Stairs, ramps and escalators: Inclusive design guidance*. Centre for Accessible Environments. London. RIBA Publishing.

[5] L. Dubertret, *Carte Geologique du Liban, 1956*

[6] ACI 318-11, Building Code Requirements For Reinforced Concrete and Commentary, American Concrete Institute, Farmington Hills, Michigan, USA, 2011.

[7] ASCE 7-10, Minimum Design Loads for Buildings and Other Structures, American Society of Civil Engineers, Reston, Virginia, USA, 2013.
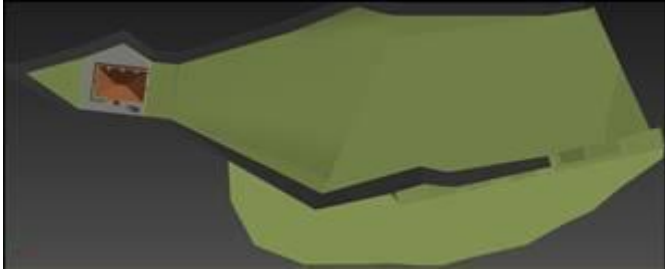
Figure A.1. Initial Development of the Youth Village

Table A.I. Surveyed St Paul Members Information

Number of Years in StPaul	0-2 Years	2-6 Years	More than 6
Number of Members	0	11	15
Level of Involvement in StPaul	Highly Active	Intermittent	
Number of Members	23	3	

Table A.II. StPaul Members Problems, Suggestions and Innovative Ideas

Problems encountered at previous camps
Power Cuts, Water Shortages, Lack of parking spots, Absence of indoor and outdoor playgrounds
Suggestions for the Youth Village
Make use of solar energy for water heating, Harvest rainwater, Indoor playgrounds, Green spaces, amphitheater, Outdoor shaded playgrounds, Chapel, Library, Agricultural spots
Innovative Ideas
Sorting and Recycling of generated wastes, Inclusive design, Participatory approach in construction, Smoke free village, Bike path within the village,

Table A.III. StPaul Members Sustainability Feedback

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	N/A
Sorting at the source	73%	23%	0%	0%	0%	4%
Use of Solar Power	73%	23%	0%	0%	0%	4%
Rainwater Harvesting	58%	27%	8%	0%	0%	7%
Planted Forest with Fruitful Trees	81%	12%	4%	0%	0%	3%
Environmental Activities	58%	31%	8%	0%	0%	3%
Educational Signs for Youth	58%	31%	4%	0%	0%	7%

Figure A.2. Playground Mental Maps

Figure A.3. Children House Basement Architectural Drawing
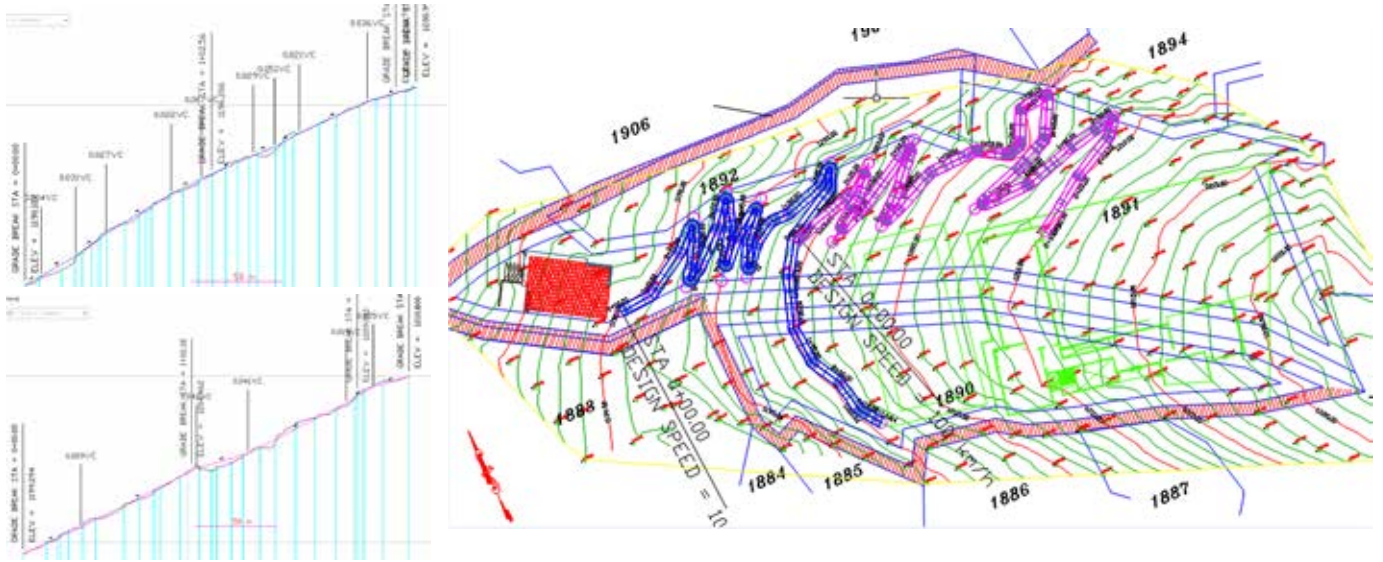
Figure A.4. Walking/Bike Trail Vertical and Horizontal
Alignments (in Blue and Pink)

Table A.IV. Fruitful Trees Characteristics

Technical Name	Common Name	Water required	Pruning Needed	Season of Bearing	Spreading Parameter	Growth Rate
Peruvian Pepper Tree	Peruvian Pepper	Low	Little	Spring	8-10m	Fast
Olea Europaea	Olives	Low	Little	Spring	10-15m	Slow
Ziziphus jujuba	Chinese Date	Moderate-Low	Needed	Spring	3-5m	Moderate
Prunus armeniaca	Apricot	Low	Little	Summer	8-10m	Moderate
Prunus dulcis	Almond	Moderate-Low	Needed	Summer	5-8m	Moderate
Malus domestica	Apple	Moderate-Low	Needed	Spring	5-8m	Moderate
Prunus persica	Peach	Moderate	Needed	Spring	5-8m	Fast
Prunus domestica	Plum	Moderate	Needed	Summer	3-5m	Moderate
Pyrus communis	Pear	Moderate	Needed	Spring	5-8m	Moderate
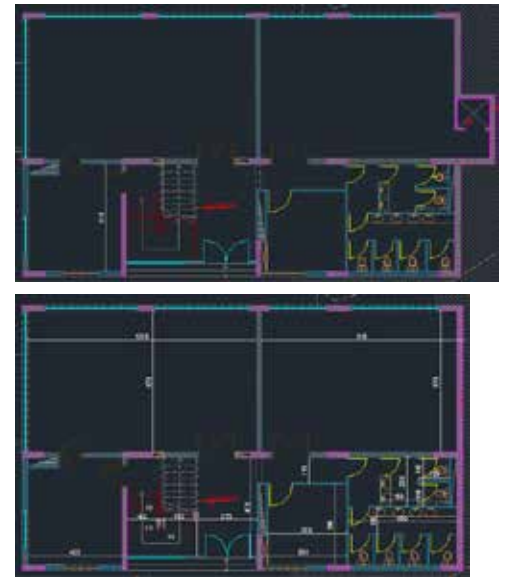
Figure A.5: Youth House Initial and Final
GF Architectural Drawings

TABLE A.V
COST OF POWER SUPPLY ALTERNATIVES FOR YOUTH HOUSE

Alternatives	Present Worth ($)	Initial Cost ($)	Yearly Cost ($)
PV	146,000	42,000	7,500
EDL & PV	118,000	43,000	5,500
EDL & Diesel Generators	72,000	23,000	3,600
Diesel Generators	107,000	16,000	7,700

TABLE A.VI. MAXIMUM RAINWATER COLLECTED FROM EACH FACILITY

Facility	Harvested Water (m³)
Children House	1344
Basketball Court	524
Football Court	775
Children House Playground	523
Bleachers	269
Bike Trail	1080
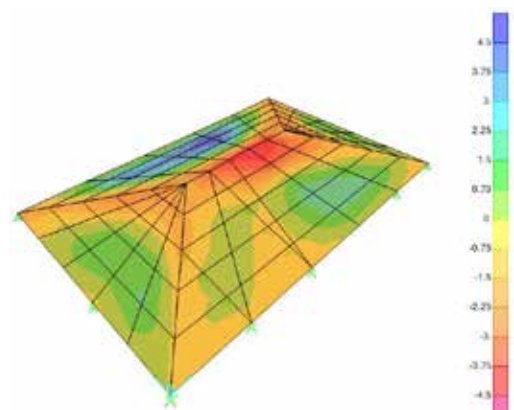
Figure A.6. Roof FEM in Lateral Direction

353

Study of the New Port of Sidon

Abdul Rahman Mikati, Hassan Damerji, Leen Sinjaba, Rayan Asaad, and Yara Kayali El-Alem

Civil and Environmental Engineering Department

American University of Beirut

Bliss Street, Beirut, Lebanon

aam65@mail.aub.edu, hmd12@mail.aub.edu, lms30@mail.aub.edu, rha55@mail.aub.edu, yfk02@mail.aub.edu

The port of Sidon is currently a small seaport that accommodates for break bulk, yet also serves a small number of containers and fishermen. Due to the urge for a new port in Sidon, and due to the congestion of the port of Beirut after its several expansions, the Ministry of Public Works and Transport of Sidon already accomplished phase 1 of a new seaport located south of the old port. A preliminary design of a port with break bulk and commercial yacht development is established, yet subject to change. This study focuses on establishing a new design for the new port which incorporates a cargo terminal, and a break bulk terminal. A geotechnical study is carried out for the suggested design and run on the vb.net software, for quantities of dredging and reclamation. Moreover, a coastal study is done to assure the safety of the breakwater built in phase 1, and to establish a design of the new vertical wall. The design is carefully studied to establish a capacity of 205,000 TEU, which is comprehensive of the existing port capacity, and the additional containers that will be transferred from the port of Beirut to serve the south. To further test the design, a benefit-to-cost analysis is carried out yielding a ratio greater than 1, and so proving the necessity of the port.

I. INTRODUCTION

Lebanon accommodates four main sea ports that serve as a gateway for its external trade. The ports are located from South to North in the following order: port of Tyre, port of Sidon, port of Beirut, and port of Tripoli. Although the port of Beirut is the largest sea port in Lebanon, it is reaching capacity and is congested. The three other ports in Lebanon are not capable of accommodating the increased commercial activities, so a new port in Sidon is now being constructed to serve Sidon and the South trade.

The old port of Sidon is a small port that accommodates fishermen and can only handle ships of considerably small size. Many incidents have been noted where ships find it difficult to enter the port. This is due to the fact that the port is insufficient in terms of size and quality and cannot accommodate more than one large ship at a time (Zaatari, 2004). The new port is being constructed to provide the necessary capacity to handle the ships, and to aid the port of Beirut. The port is expected to boost Sidon's trade, increase job opportunities, and nourish the area. The commercial terminal will serve as a solution to the congestion of the port of Beirut, by handling the ships intended for the South Governorate region, which is about 10 % of Lebanon's area. The project is handled by the Ministry of Public Works and Transport and the contractor is Al Jihad for Commerce and Contracting. Phase I of the project, which is the construction of the breakwater systems, has already been completed.

This study of the new port in Sidon will accommodate for the growing trade in Lebanon. The aim of the study is to propose a new design for the new port and test its various elements. The analysis will be based on the design itself: geotechnical study, coastal study, and handling capacity levels in terms of the ships and trucks' flow, as well as the terminal processes (loading/unloading and transporting). A queue analysis for the ships and trucks inside the port will ensure the port's operations are under proper movement. Moreover, a benefit-to-cost analysis will prove that the design is feasible in monetary terms. Not only are these criteria used to assess the impact on the port's performance, but also their impact on the old port of Sidon.

II. DESIGN OF THE NEW PORT

A. Geotechnical Study

Dredging and land reclamation are two major geotechnical activities for port construction. For the new port of Sidon, dredging to a level of 13m below the mean sea level is needed for serving the demand of the port. Since the dredged volume is large, the dredged seabed materials (carbonate sediments with sands and clay intrusions) are used to reclaim the land servicing the port. The quantities of the dredged and fill materials are obtained using the cross-section method applied on the contour maps of the port. Table 1 below shows the quantities dredged, used, and disposed for the new design.

The reclaimed land has moderate to low strength. This is low compared to the needed strength for handling large loads from the cargo and equipment, so a study for soil improvement was carried out using wick drains. The wick drains are geotextile jackets installed in the soil to boost up the rate of drainage and consolidation. For that design purpose, a user interface software using vb.net was developed and used to solve the design problem. The soil properties needed for the software are taken from geotechnical studies of the Tripoli port expansion, since the seabed layers in Tripoli are similar to those of the Sidon port site.

Tables 2 and 3 show the design parameters of wick drains (width, spacing etc.), settlements, and time needed for a consolidation degree of 95%.

Table 1 Dredged Material

DREDGED (M^3)	FILL (M^3)	EXCESS (M^3)
1,650,875	760,003	890,872

Table 2 Software output

DESIGN PARAMETERS	VALUE
Spacing (m)	1.61
Dw (mm)	60
De (function of spacing)	1.13 x Spacing
N	30.34
F(n)	2.67
T	180 days
U(=S'_total/S_total)	95.01%

Table 3 Software Output

LAYER	TH	U (%)	S'
Layer 1	1.61	99	0.33
Layer 2	0.71	88	0.166
Layer 3	0.62	85	0.01
		S'_Total	0.506

After achieving the settlement as a result of the soil improvement, 1m of clean sand layer is laid over the reclaimed land.

B. Coastal Study

Break Water Analysis

During Phase I of the project, the breakwater with the characteristics shown in Table 4 was constructed. The run-up value of 3.55 m is less than R_c = 5.50 m, which is the distance from the crest of the structure and the sea-water level. Thus, there's no overtopping and the breakwater is well constructed to protect the commercial port that will be built in the leeward of the breakwater.

The Design of Vertical Wall

With the new proposed design, a vertical wall is needed at the quay primarily to withstand wave attack and act as an earth retaining structure for ships to be berthed at. Since it's a vertical wall, the reflection coefficient will be taken as the most critical case, which is Cr=0.9 (Knovel: Coastal Engineering Manuel, 2006). Using linear wave conditions, the wave crest is $\frac{H}{2} = \frac{0.1432}{2} = 0.0716$ m (Knovel: Coastal Engineering Manuel, 2006, and the run-up will be equal to

$$(\frac{H}{2}) \times (1 + Cr) = Run - up \qquad (1)$$

Table 4 Characteristics of the Breakwater

CHARACTERISTICS OF THE BREAKWATER	
Material	Rock-armored accropode
Slope	3:4
Length of Breakwater	920 m
H.W.L Elevation	+5.00 m
R_c	5.5 m

Table 5 Design Parameters of Vertical Wall

DESIGN PARAMETERS	
Height (m)	13.5
Width (m)	8
Material	Concrete

The run-up value for the vertical wall is around 14 cm. Thus, the minimum h_c, that is the distance between the seawater level and the crest of the structure, should be 14 cm, but will be taken as 50 cm for safety and structural measures. The minimum width needed for the overturning moment to be equal to the moment due to the weight of the wall is 7.707 m. A value of 8 m is taken as the width of the vertical wall with a safety factor of 1.03. The design parameters of the vertical wall are summarized in Table 5.

III. DESIGN OF THE INSIDE FACILITIES OF THE PORT

A. Container Yard

Quay Cranes

Considering the maximum ship size that can safely enter the port (the Panamax 1980), and the crane outreach required to handle it, the 1st Generation PANAMAX STS Crane was found to be the most suitable for the new port of Sidon. A total of 4 STS cranes will be employed, since a maximum of two container vessels can be handled at a time.

Landside Container Handling

Four high straddle carriers will be used for landside container handling. The main advantage of straddle carriers is that they are independent from any other equipment and are able to perform all the different handling operations: transporting, stacking, and the loading/unloading of trucks and rail cars. Since a ratio of 4 or 5 straddle carriers are needed to every STS crane, a total of 18 straddle carriers will be used for the whole port, 6 of which will be dedicated for horizontal transport and container stacking within the yard. As for the remaining 12 straddle carriers, they will be divided into pairs, each assigned to one container yard to handle the loading/unloading processes of the container trucks.

Container Yard Layout and Operation

The STS Crane places the containers onto the apron from where the straddle carriers transport and stack them in the container yards. Linear stacking was adopted since the straddle carriers need space to pass between the containers. In front of each container yard, there exists a truck service area (TSA). This is where the container trucks wait to be serviced (loaded/unloaded) by a straddle carrier.

B. Break Bulk Warehouse

Sidon is an important importer/exporter of break bulk cargo. The major category of goods transported include: pharmaceuticals, electrical machinery, iron, steel, plastics…

All of these goods are transported and stored in pallets. Pallets come in standard dimensions of 600 cm by 400 cm. a configuration of 4 pallets stored in metal frames is built, with a three story capacity space for storage.

In order to carry out the transport process of the break bulk, the break bulk are unloaded from a ship by a mobile crane. Two mobile cranes will be available to carry out the loading/unloading process at a faster pace, or in the case of two ships arriving simultaneously. After the goods are unloaded, fork lifters act as the mean of transport from the break bulk terminal to the warehouse, where the pallets are either stored or loaded onto trucks. The reverse process takes place when fork lifters transport goods out of the warehouse and the mobile cranes load them onto the ships for export. Five fork lifters will be used, each with a capacity of transporting loads to and from a three shelf pallet system.

C. Parking Systems

A parking is one of the most important and essential facilities that must be provided to enhance the transportation characteristics of any seaport. The parking's capacity, orientation, and aesthetics have a significant impact on the economic activity of ports. Hence, the main objective in designing the parking facility was to provide a well-defined searching path that allows a good circulation.

The proposed design abides by the Lebanese Standards for a surface parking area of greater than 800 m^2. Those Standards include the following:
- Design vehicle dimensions
- Motorcycle parking
- Aisle width
- Horizontal curves dimensions

The proposed design is efficient in a sense that it has a clear searching path which eliminates confusion and allows ease of circulation within the parking area.

D. Other Facilities

Other than the facilities discussed, the port will have areas of: truck parking, equipment parking, and equipment repair. In addition, as a necessity for proper administration, an administrative and operation building will be located between the car parking and terminals for proper supervision. WCs and firefighting stations will also be included.

E. Port Gate Terminal

The gate is studied carefully to insure proper inspection management and customs jurisdictions. All external truck drivers and vehicles are required to obtain permission at the gate to enter the port, which creates large queues if not managed properly.

The gate has six inbound and six outbound lanes with static scales for weighing. When a truck driver arrives at the terminal, the driver must show authorization to enter the port. This is the worker identity verification credentials. Once all requirements are satisfied, the driver is granted access and a gate pass is issued. A gate pass is a printout showing the date and the time of the access, the name and ID number of the driver, and a list of the equipment brought into the port (Gate to Gate: What Happens When a Truck Picks Up a Container, 2014).

If there were problems with the paperwork, the truck is sent to the customer service to check with its company and the port administration office if needed. Once everything is verified, an interchange is printed which includes the location of the container to be picked up/dropped, and the truck is cleared to proceed into the terminal. The truck has to first pass through the RFI Scanner for safety checks and to ensure the truck is registered in clean truck data base, and it can then proceed to the transfer location with the directions provided.

Upon arriving at the specified transfer location, the truck waits for the straddle carriers for service. Once the truck is ready to exit, it heads to the outbound lanes that have static scales as well for those trucks that might need weighing before leaving. In the outbound process, once the gate pass number is entered, the officer checks the information and image of the driver to validate that he/she is the same person as the one who entered. Once validated, the driver exits (Hernandez, 2009).

It is important to note that cargo mainly requires weighing on the way in and out, whereas containers don't require any. Also, imports and exports in Lebanon to/from Arab countries are assigned neither a fee nor a VAT based on agreement between Arab countries.

IV. Capacity and demand

A. Demand

The new port of Sidon is carefully designed to accommodate the forecasted demand of the South, which is comprehensive of the current demand of the existing port of Sidon and the additional containers transferred to the South from port of Beirut.

Currently, the existing port of Sidon handles 25,000 TEUs per year (Chami, 2008). Since the old port of Sidon will serve fishermen only in the future, all of these TEUs will be transferred to the new port of Sidon.

In addition, the port of Beirut handles 1.2 M TEUs per year, 15% of which is transferred to the South region (BCTC, 2016). Since the new port of Sidon cannot accommodate for ships larger than the Panamax vessel, it will not be able to capture all of this percentage. However, the percentage captured will be assumed to be 15% as a safety measure to accommodate for further expansions. Thus, the forecasted total demand at the new port of Sidon will be 205,000 TEUs per year.

B. Capacity

To check if the new port can handle the forecasted demand of 205,000 TEUs per year, the berth capacity of the new port of Sidon is calculated based on (2):

$$Berth\ Capacity = n \times \emptyset \times T_{year} \times P \qquad (2)$$

Where:
n: number of berths, represented by (3).
∅: Acceptable berth occupancy ratio, represented by (4).
T_{year}: Hours the terminal is operational per year (hrs.).
P: Annual average productivity of vessel at berth.

$$n = \; Length \; of \; berth/(Length \; of \; vessel \; x \; (100\% + K \; separation)) \qquad (3)$$

and

$$\emptyset = number \; of \; ships \; arriving/service \; rate \; at \; berth \times n \times 365 \qquad (4)$$

The length of the berthing facility is 375 m, and the length of standard vessel is that of the Panamax, 250 m (FINCANTIERI, n.d.). The berthing gap or distance between vessels at berth is zero, since only one Panamax can berth at a time. This results in **one** berth at the port of Sidon, according to (2).

Since one vessel needs two days, the number of ship arrivals per year is 175 vessels, and the average number of ships each berth can service per day is 0.5. This gives an occupancy ratio of 0.958 from (3).

It is assumed that there are three shifts, 8 hours per shift, and 360 working days per year, totaling to 8,640 operating hours/year of the terminal.

The vessel productivity is correlated to the STS productivity i.e. 20 containers/hour (Kaysi, 2016). Using (2) the berth capacity is 207,124 containers/year, and it exceeds the forecasted demand on the new port of 205,000 TEUs per year, justifying the design.

V. Cost Estimations

For the cost estimations, a work breakdown structure (WBS) was done, and based on it, a bill of quantities (BOQ) for the material, cost, and quantity of work needed was established, as shown in Table 6. Later, these costs were used as inputs for the benefit-to-cost analysis.

A. Work breakdown Structure (WBS)

For the new design of Sidon Port, a WBS was done on the levels of: offshore works, landscape/sidewalks, pavement construction, equipment placement, and structures construction. The five major tasks are then divided into activities and sub-activities. For example, the offshore works were divided into coastal construction activity and dredging with backfilling activity, then the coastal construction activity was divided into two sub-activities: break water construction and vertical wall construction.

B. Bill of Quantity (BOQ)

The bill of quantity for the new design of Sidon Port project was developed on the level of sub-activities. Costs for all sub-activities were calculated to come up with a rough estimate of the project's cost. Table 6 below shows the BOQ developed for the five main activities of the project.

The BOQ yields a total infrastructural cost of $72,903,085 for the project.

Table 6 Bill of Quantities of the Project

ITEM DESCRIPTION	QUANTITY (M³)	UNIT PRICE ($)	ITEM COST ($)
Site Preparation (Dredging, Reclamation)	700,000	46.60	32,617,900
Landscape/Sidewalk Design	7200	31	223,934.72
Pavement/Quays Construction	78,000	133	10,404,000
Equipment Placement	50	564,500	28,225,000
Structures Construction	3500	409.21	1,432,250

VI. Scheduling

A. Primavera Output

The primavera software is used to schedule the project and determine its total duration. A 7-day work week is assumed in the port as it consists of three shifts with 8 hours per shift. For each activity defined in the WBS, its duration is added on primavera. Then relationships between activities are defined by assigning to each activity its predecessors using a finish-to-start relationship. All the activities of ordering the equipment consist of a start-to-finish relationship with a delay factor.

The final step consists of scheduling the project to get a project duration of around **2 years**.

The network diagram is depicted in the appendix displaying the scheduling of the project and defining its critical activities.

VII. Benefit-to-Cost Analysis

A benefit-to-cost (B/C) analysis is a necessity in a project of any scope. Investing a large sum of capital calls upon benefits to be weighed against those costs in order to verify the project. The simple equation comes to use when calculating the B/C factor, which in this case yields a value of 8.15, greater than 1. This can act as an indication for a go ahead decision.

$$B/C = Benefits \; (\$)/Costs \; (\$) \qquad (5)$$

A. Benefits

The benefits of building an entirely new seaport cover a wide span. The difference in the benefits according to the number of twenty foot equivalent (TEU) containers handled by the old port and the new design of the new port is the incremental benefit calculated in the benefit-to-cost analysis. The first category of benefits arises from the reduced cost per TEU due to the larger amount of TEUs transferred using the Panamax (3000 TEU capacity), as compared to

the 500 TEU ships that reached the old port of Sidon. Taking into account the daily running costs, daily costs at the berth, and the port charge, the cost per TEU is $83.43 less. Multiplying this value by the difference in the capacities of the old and new port, 180,000 TEUs, and changing it to annual payments for 20 years, with an interest rate of 10%, yields a value of $97,738,645.8.

The inland transportation savings also need to be considered. For every truck that used to transfer the TEUs dedicated to the South from the port of Beirut, a truck with a driver were the means of the transfer. Without these trucks, environmental costs of the operations of the trucks, and value of time gained would be recognized as benefits. The benefits associated to the trucks include the fuel cost, and hence carbon dioxide, the tire cost, and the maintenance cost. The vehicle operating cost for a large vehicle is hence calculated. All of the costs depend on the distance traveled, 48 km from the port of Beirut to the port of Sidon. The fuel cost is the fuel price multiplied by distance traveled, and by the number of trucks per year that will no longer be traveling from Beirut to Sidon. The value calculated is about $48,388. For a tire cost of $0.009 on average per vehicle, and a maintenance cost of $0.021per vehicle, a total vehicle operating cost for all the vehicles is $25,162,546.6.

The reduction in trucks allows for environmental factors to be shed. These trucks no longer burn carbon dioxide and affect the climate. For an average social cost of carbon price of 83 $/ton of carbon dioxide, and an air pollution cost of 0.012 $/veh km, the total benefits calculated result in $135,206.3 saved.

B. Costs

The total costs to construct and operate the new port are calculated. Costs include the fixed costs of the infrastructure capital and equipment. They also include the annual costs of the employment, insurance of equipment, fixed and variable maintenance of equipment, maintenance of infrastructure, repair costs, and maintenance on dredging. The expected increase in maintenance cost of equipment and dredging is also taken into account.

For the equipment costs, the new port requires 4 STS cranes, 18 straddle carriers, 5 fork lifters, 4 empty container handlers, 16 yard tractors/trailers, 216 pallets, and 2 mobile cranes. The number of equipment is based on the ratio needed per STS Crane (Kaysi, 2016). Multiplying each number by its corresponding unit price yields a value of $ 28,225,190 M, along with other miscellaneous equipment. In addition, the major cost of the new port is the capital cost of the infrastructure, which is already calculated in the BOQ as $ 44,678,085.

The difference in employment cost for the new port of Sidon is $892,800 per year. The insurance for equipment is assumed to be fixed yearly value equal to 1.25% of the purchase value, yielding a value of $352,814 per year. The annual maintenance of equipment comprises of a fixed cost and a variable cost per TEU. Fixed costs are estimated at $320,000 per STS Crane per year, and variable costs are valued as $2.77 per TEU (Kaysi, 2016). Inflation is

applied to maintenance and repair costs. Dredging work require annual dredging maintenance, with annual inflation of 5.5%.

$$Dredging\ Maintenance\ cost(\$/year) = 50,000 + 0.8 \times (Annual\ TEUs)$$

(6)

This results in a value of $214,000 per year, which results in $225,770 per year after inflation (Kaysi, 2016).

The fixed costs add up to a value of $72,903,275. The annual operational costs are taken as the incremental costs to the old port of Sidon. The total annual costs of the whole project are represented in (7):

$$Total\ annual\ costs\ (\$/year) = 2,666,395.72 + (A/P, 10\%.20) \times 72,903,275 + (A/G, 10\%, 20) \times 672,957$$

(7)

The interest rate is taken as 10% by the Central Bank of Lebanon (Lebanon Interest Rate, 2016). This results in a value of $15,609,285 per year to be included in the cost-to-benefit analysis.

VIII. QUEUE ANALYSIS

A. Truck Queue Analysis

Modeling Truck Arrival

According to the new port's demand, the arrival pattern of the trucks will be approximated by a Poisson function with an average arrival rate of 205,000 TEU/year. Assuming 360 days per year,

$$\lambda = 570\ TEU/day$$

and 24 hours per day,

$$\lambda = 24\ TEU/hour = 24\ container\ truck/hour.$$

Modeling Truck Service at Inbound and Outbound Gates

Assuming that a truck entering the port requires 10 mins of inspection on average (paperwork check, weighing, and scanning), the inspection rate at each inbound gate hence is expected to be $\mu_1 = 6$ trucks / hr. Note that the queue system at the gates does not include trucks with incorrect paperwork, as these trucks will not wait in the queue but they will rather be deviated to the customer service area, after having a trouble ticket issued for them.

As for the outbound gates' service, it is expected to be faster than that of the inbound gates since no scanning is required. Thus it will be assumed that a truck will require 8 mins to be served on average, which gives an average inspection rate at each outbound gate μ_3 of 7.5 trucks / hr.

Modeling Truck Service at the Container Yard

For an automated straddle carrier, the loading and unloading of a road truck requires 4 to 6 mins (Intermodal Handling, n.d.). However, the straddle carriers employed at the port of Sidon will not be automated. The time per truck increases consequentially. Also, it is important to account for container shuffling in case a container to be loaded was not on the top of the stack. To account for the previous remarks, the straddle carrier service time will be assumed to be 15

mins/container which is equivalent to a service rate μ_3 of 4 container trucks/hr.

Queue Network Assessment

After combining the three models above (trucks' arrival to the port, their service at the inbound and outbound gates, and their service at the container yards), a complete queuing network was established and assessed. The results are shown in Table 7.

These numbers clearly show that it is unlikely to have very lengthy queues in any of the three stations. This is of great importance since it reduces the probability of having a spillover of trucks onto the main road, when these trucks are waiting at the inbound gates.

In addition to that, the average truck visit time for the new port of Sidon can be computed as:

$$Average\ Truck\ Visit\ Time = 30 + 20 + 17 = 67\ mins$$

(8)

The truck visit time at the new port of Sidon was compared to that of other ports in the world. Despite the fact that the new port of Sidon is considerably smaller in size and handling capacities than the ports considered, it could still be concluded from this comparison that an average truck visit time of 67 mins is quite acceptable.

B. Ship Queue Analysis

The handling services between the new port of Sidon and the port of Beirut (BCTC, 2016) are expected to be more or less the same since both of them will be using similar labor force and similar handling units, ship-to-shore cranes (STSs). The port of Beirut container movement by vessel type for each of the 12 months was noted. In order to estimate the service time at the port of Sidon, the data was first filtered and then a correction factor was applied so that the average service rate can be estimated in a more precise way.

After running the Queue Model on excel, the results were as follows: on average, the utilization rate or efficiency of the container terminal is 60.26%, the probability that there are no ships present at the container terminal is 24.97%, the probability that the system is full and a ship waits is 45.50 %, the time for a ship in the queue is 20.48 hours, or 0.85 days, and the time a ship spends at the port is 2.24 days.
This shows that the proposed design for the port of Sidon is acceptable. The port can easily handle the expected demand given the proposed handling services

IX. Conclusion

In conclusion, the design proposed allows the port to serve containers, and so handling the original demand intended for the South. This is proven from the results obtained from the queue models of the trucks and ships arriving and departing the port. The models result in a mean time of 67 mins spent by the trucks in the terminal, and 2.24

Table 7 Queue Network Assessment

	STATION		
	Inbound Gates	Container Yards	Outbound Gates
Mean number of trucks	2.0	1.3	1.1
Mean time in queue (mins)	2.0	5.0	9.1
Mean time at station (mins)	30	20	17

days for the ships to stand in queue and operate. The B/C analysis yields a ratio greater than 1. Although this value is significant, the dredging environmental effects are expected to add to the costs that will undermine the value.

References

[1] BCTC. (2016). Retrieved from http://www.bctc-lb.com/Terminal.aspx
[2] Brinkmann, B. (2011). Operations Systems of Container Terminals: A Compendious Overview. Handbook of Terminal Planning, 49th ser. Retrieved from Chapter 2 Operations Systems of Container Terminals: A Compendious Overview.
[3] Chami, M. (2015, September 16). Port of Saida Regain Some Trading Activity. Retrieved 2015, from http://blog.blominvestbank.com/port-of-saida-regain-some-trading-activity/
[4] Commercial Ports of Lebanon Main nodes in the Mediterranean network. (2015). EUROMED INVEST. Retrieved from http://www.silbcn.com/ponencias2015/puertos/rami_semaan_tms_consult_lebanon.pdf
[5] FINCANTIERI: Bulk carrier cargo ship / Panamax. (n.d.). Retrieved from http://www.nauticexpo.com/prod/finctierani-cantieri-navali-italiani/product-31009-319738.html
[6] Gate to Gate: What Happens When a Truck Picks Up a Container? (2014). Retrieved from https://www.youtube.com/watch?v=P9IJN1yIIJ4&feature=youtu.be
[7] Hernandez, H. (2009). Port of Miami Miami-Dade County Seaport Department. Integrated Cargo Gate Vehicle Processing System. Retrieved from http://aapa.files.cms-plus.com/PDFs/ITAwards/2009/Miami_IT_Award_09.pdf
[8] Intermodal Handling. (n.d.). KALMAR. Retrieved from http://www.toyota-forklifts.ro/SiteCollectionDocuments/Intermodal_Handling.pdf
[9] Kaysi, I. (2016, March). Term Project. Retrieved from http://moodle.aub.edu.lb/pluginfile.php/304055/mod_resource/content/3/CIVE665 - Term Project Description final.pdf
[10] Knovel: Coastal Engineering Manual. (2006). U.S. Army Corps of Engineers.
[11] Lebanon Interest Rate. (2016). Retrieved from http://www.tradingeconomics.com/lebanon/interest-rate
[12] Martín Soberón, A. M. (n.d.). The Capacity in Container Port Terminals. UNCTAD: Ad Hoc Expert Meeting on Assessing Port Performance.

Scale 1:5000
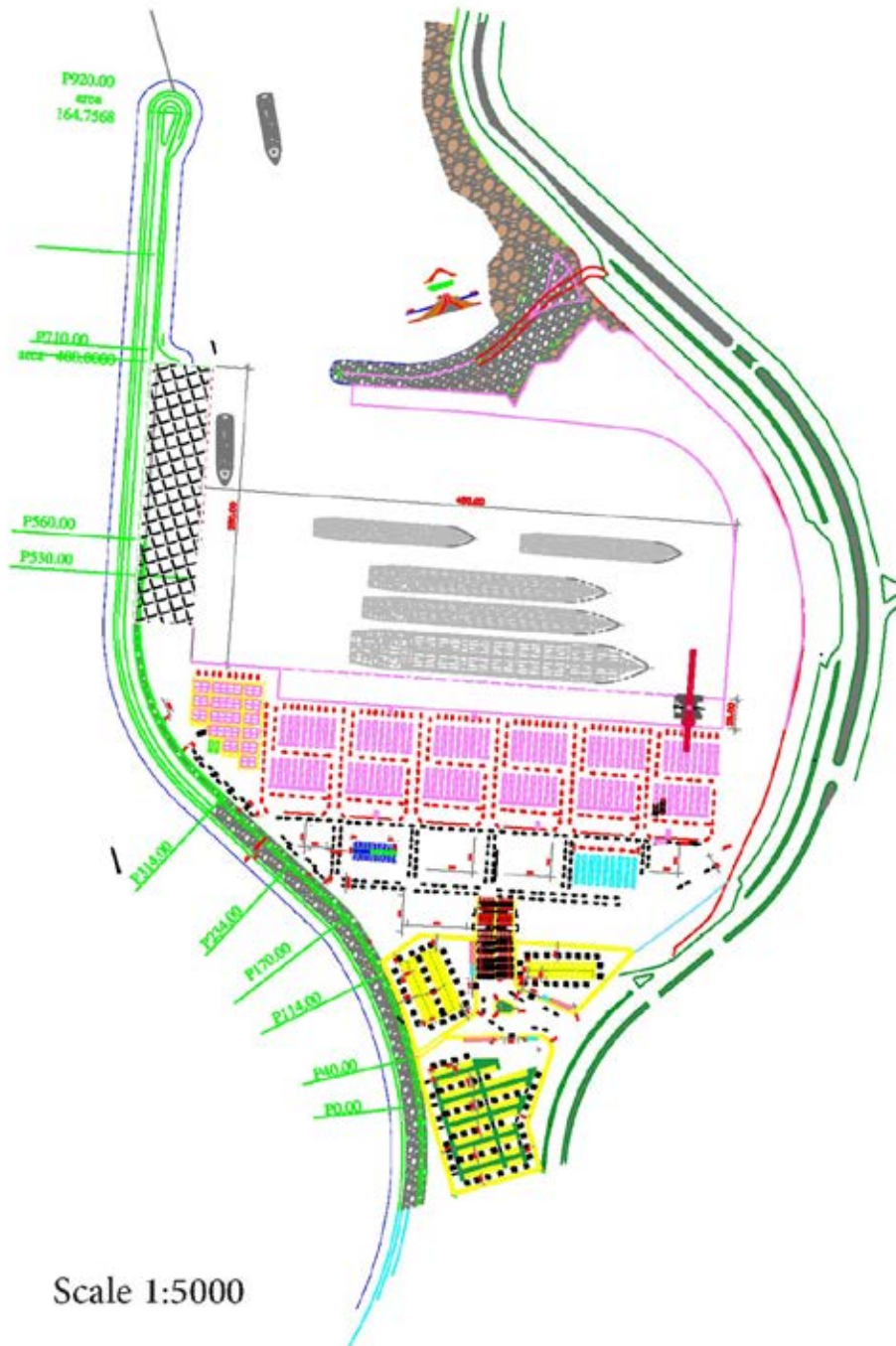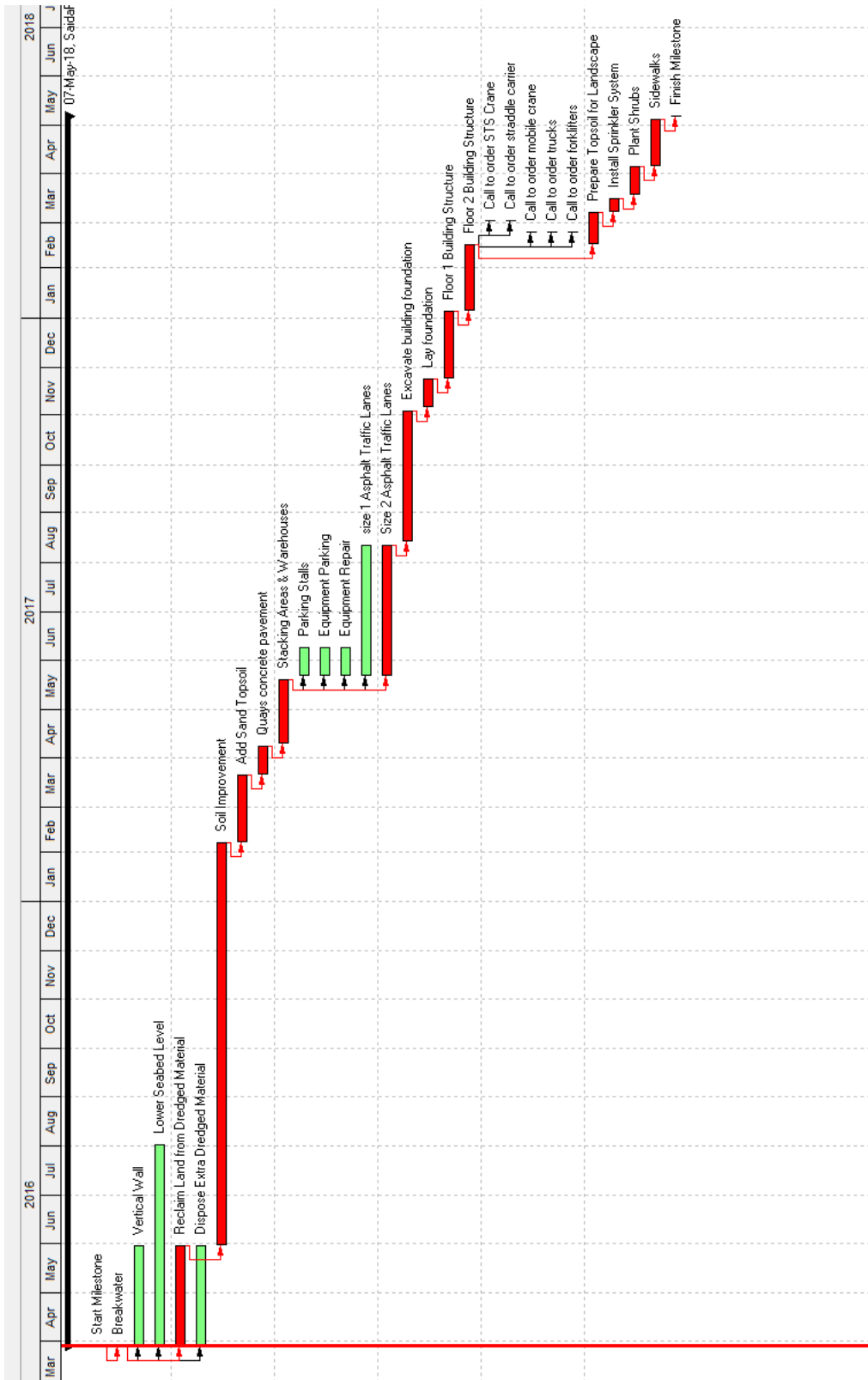
Sustainable Design Solution for the Bsalim Recreation and Care Center

Tarek Abdul Hadi, Sarah Abou Dargham, El Moataz Mohamad El Sayed Ahmad, Anthony Khoriaty, Joy Saad
Department of Civil and Environmental Engineering
American University of Beirut
Riad El-Solh / Beirut 1107 2020
tza08@mail.aub.edu, sra26@mail.aub.edu, eoe00@mail.aub.edu, ajk15@mail.aub.edu, jas24@mail.aub.edu

Abstract- **This paper presents the structural design and detailing of the Bsalim Recreation and Care Center. The sustainable materials hemp-reinforced concrete, recycled aggregate concrete, and biochar concrete were proposed as alternatives to be used in concrete masonry blocks and plaster. A reduction in the carbon footprint due to the implementation of each material was observed to be 25.8%, 40.6% and 33% along with a decrease in thermal conductivity of 29.7%, 33.6% and 2.6% respectively. In addition, a rainwater harvesting system was designed and a cost analysis was performed.**

I. INTRODUCTION

This paper proposes a holistic design solution for the Bsalim Care & Recreational Center while incorporating sustainability. The project was developed by the American University of Beirut alumnus Tessa Al Sakhi as her Final Year Project. Due to its promising design features it received the AREEN award in 2015. The facility serves as a therapeutic and recreational center for psychologically ill children. The development follows the topography of the hill it is located on in order to minimize excavation. It comprises large green spaces, clinical and supervisorial spaces, and non-therapeutic spaces (such as Library/ Cafeteria/infirmary) distributed over eight irregular floors each serving a different function. This paper will cover the transition from an unfeasible architectural concept to a structurally sound system that incorporates sustainable and optimized design solutions. The slabs, columns, retaining walls, foundations were designed while minimizing changes in the architectural drawings.

In order to promote the sustainability of the development, three construction materials were investigated: recycled concrete, biochar and hempcrete. These materials could reduce natural resources depletion, lessen the solid waste problem and solve the concrete demolition waste issues in Lebanon. Moreover, their effect on the carbon footprint of the building as well as the energy consumption of the structure will be investigated. A Revit model was developed in order to accurately evaluate the quantities of the different materials used and estimate the energy consumption. Moreover, a water harvesting system was designed.

II. STRUCTURAL DESIGN

A. Reading Architectural Drawings

The first step in the design was providing a sound structural system based on the architectural drawings. The provided drawings constitute eight different floors with large spans, asymmetrical slab limits, columns disconnected from one slab to another and irregular basement wall curves.

Partition walls were distinguished from structural walls by adopting a maximum allowable span of 8m between columns to minimize deflection. After that, the slabs were superimposed on each other to identify planted columns, or any other inconsistencies in the walls or the slabs. Figure 1 depicts the transition of Floor 3.

Figure 1: The transition of floor 3

Other modifications were made. For example, Floor 6 was divided into two separate slabs, with close proximity and a retaining wall that separates both slabs. However from a construction perspective it is more sustainable to join them into one coherent slab instead of having an internal retaining wall as shown in Figure 2.

Figure 2: Transition of floor 6 to one coherent slab

Many other examples, include joining discontinuous columns over the connected slabs posed limitation to spaces used in the one of the slabs, a column supporting slab 4 when extended to slab 5 will end up in the middle of a spacious room. Later this was solved in coordination with the architect who proposed to extend the slab limit and divide the room into two separate rooms.

In conclusion, the transition from the architectural drawings to feasible structural drawing with minimal alterations proved to be challenging especially while avoiding expensive alternatives such as drop panels, transfer beams, and pre-stressed concrete.

All structural design in the following sections was based on the ACI-318-11 and ASCE 7-10 requirements.

1. Slab design:

The structure includes nine slabs of different layout. Solid slabs were adopted as they were adequate for the available spans, and easier to construct. A common thickness of 30cms was chosen to satisfy both strength and serviceability requirements. Nine SAFE models were generated in order to model the behavior of the different slabs. Sample models are shown in Figure 3:
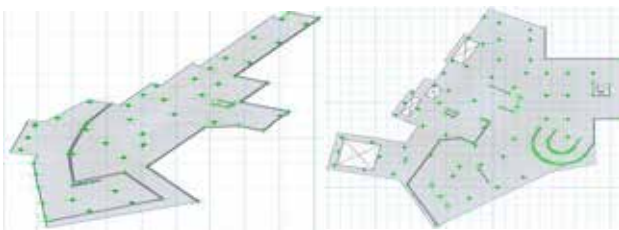

Figure 3: Sample models

The Direct design method described in ACI was used in order check the results generated by SAFE. Punching shear was checked for both the columns and foundations. Punching Shear was satisfied in most of the columns in places where it was critical reinforcement was provided in accordance with the ACI requirements. Using drop panels was avoided because of architectural constraints.

2. Columns design:

The project contains 125 columns were required in order to maintain the average span length that was set. 15 different column dimensions were used because of architectural constraints. Most of the columns were chosen to be circular with a diameter of 60cms. An excel sheet was created in order to estimate the strength of the different types of columns along with the required reinforcement. An ETABS model featuring the entire structure was generated in an attempt to model the behavior of the building and to account for seismic loading as shown in figure 4:
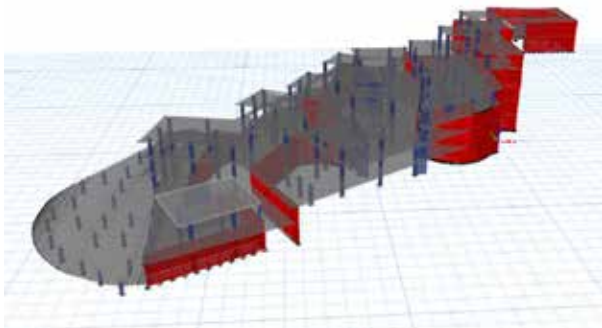

Figure 4: ETABS model

Slenderness, moments and forces the columns were subject to were all evaluated manually and by using ETABS. The reinforcement for each column was determined accordingly.

B. Geotechnical Design

3. Basement/retaining walls:

The project includes retaining walls that are connected to the slabs on different levels and are therefore different from traditional basement walls or retaining walls. The most critical basement/retaining wall was modeled on SAP2000 as shown in the picture below:

The slabs connected to the wall were modeled as pin connections and the footing was considered as fixity. This wall is subject to a triangular distributed load varying from 0 to a maximum value that can be determined according to Mayne, P.W. and Kulhawy, [1], as follows:

$$\text{Maximum stress} = 1.6 \times \gamma \times D \times K0$$

Where γ is the density of the soil, D is the depth at which the bottom of the wall is located, and K0 is the lateral earth pressure coefficient (at rest). Detailing of the wall along with its supporting footing was designed accordingly.

4. Global Stability:

Local stability was assured by designing the basement/retaining walls against most types of failure. The global stability of the structure should be investigated as well. The structure can fail as a whole by either sliding or overturning since it is subjected to lateral forces coming from the mountain. Overturning is not a concern in this project since the structure in extremely wide in the direction of the lateral load. Sliding of the building is on the other hand possible and should be investigated. If the total lateral force exceeds the friction force that is holding the building in place, the structure will slide. According to Rankine, [2], the lateral force can be calculated by:

$$\text{Total lateral force} = 0.5 \times \gamma \times D^2 \times K_0 \times L$$

Where L is the total length of the critical retaining wall, γ is the density of the soil, D is the depth at which the bottom of the wall is located, and K0 is the lateral earth pressure coefficient at rest.

K0 is used here since it is the most conservative approach. Ka (active Lateral earth pressure coefficient) could replace K0 in the equation above depending on the displacement of the structure under lateral forces. However, if no failure occurs by using K0, no further investigation is required since K0 is greater than Ka.

The resisting force can be calculated by using the following equation according to P.J. Blau, [3]:

$$\text{Total resisting force} = W \times f$$

Where W is the building's own weight and f is the friction coefficient.

The resisting force is therefore equal to 69747 KN which is greater than 38400 KN for the lateral load. The structure will therefore not fail by sliding or overturning.

III. SUSTAINABLE MATERIALS

The process of concrete manufacturing produces large quantities of carbon dioxide and causes the depletion of natural resources. Cement production releases vast amounts of carbon dioxide while natural resources are depleted as natural coarse and fine aggregates are sourced from quarries. "Green" or sustainable concrete aims to mitigate the environmental impact of concrete by partially or fully replacing one or more of the constituents in the concrete mix. Performance requirements, as stated by standards such as the American Society for Testing and Materials (ASTM), should be met. The focus in sustainability is on life-cycle cost (rather than solely on initial cost) as well as durability.

A. Hemp-reinforced Masonry Blocks

Locally available industrial hemp fibers were explored as a viable alternative in a study conducted at the American University of Beirut [4]. Hemp belongs to the family of plants called Cannabis. When it is not produced as a drug, hemp is often referred to as industrial hemp. It is used in industries of plastics, food, and textiles. The research done by Awwad proposes that the demand for industrial hemp can increase dramatically if used in concrete leading to an incentive for farmers to grow the plant, especially opposed to its "illegal" sister plant. The value of industrial hemp could amount to $192 per dunum[1] while the cost of production amounts to $79 [4].

Hemp fibers can be used in non-load bearing elements such as concrete masonry units, mainly used as partition and cladding walls in the region.

Hemp can basically be added as a raw material when used in hollow blocks [5]. Reduction of the use of aggregates in the mix, lighter mass of the blocks, and better thermal performance of the masonry units was the objective of the study.

Adding cement and hemp fibers equivalent to 0.7% each of the total weight of the control design mix resulted in about 25% more blocks. The proportion of hemp fibers in the concrete mix was 3.4 Kg/m^3 [5]. In addition, the thermal conductivity was measured to be 0.984 W/m.K [5].

B. Recycled Aggregate Concrete

The use of recycled aggregates in concrete mix design is another solution to protect natural resources. The waste from demolition of concrete structures is one of the most important sources of recycled concrete aggregates. The fact that this would resolve the issue of waste management and disposal of demolished concrete further adds to the environmental aspect of this solution. A large number of concrete demolition waste is being disposed of at sea in Lebanon. Therefore, the use recycled aggregate concrete can be proposed as another alternative for use in the Bsalim Recreation and Care Center.

Similarly to hemp-reinforced concrete, recycled aggregates can be used in the formation of masonry blocks. The amount of recycled aggregates should not exceed half of the total amount of aggregates in the mix for concrete masonry units in order to obtain acceptable results [6].

The thermal performance of masonry units with recycled aggregate concrete should also be investigated. In general, when the replacement ratio of recycled to natural aggregates increases, the thermal conductivity of concrete decreases [7].

The thermal conductivity of a concrete masonry block with 40% recycled concrete aggregate (by weight of total aggregates) can be estimated at around 0.93 W/m.K [7]. It is important to note that the presence of air increases the thermal conductivity of hollow blocks. The block whose thermal conductivity was estimated above had three rows of holes.

C. Biochar Concrete

Given the solid waste crisis that Lebanon is currently facing, controlled pyrolysis seems to be an alternative for solid waste management. This process consists of decomposing waste by exposing it at a very high temperature in the absence of oxygen [8] (N.J. Themelis et al, 2011). One of the byproducts of pyrolysis is biochar. This material is still relatively new and there is a lot of ongoing research about ways to make use of it. One alternative that this paper suggests is incorporating biochar in non-structural concrete. By doing so, this byproduct could be used to reduce the carbon footprint generated in the construction process of a building. Moreover, energy consumption could be reduced due to the new thermal conductivity that the concrete will have. At the moment, there is only one plant that does pyrolysis in Lebanon. It is located in Saida and produces on average 2 tons of biochar for every 20 tons of treated waste.

In order to assess the environmental impact of using biochar in non-structural elements, three concrete mix designs were introduced in accordance with the American Society for Testing and Materials (ASTM) recommendations. Two of them contained biochar with a replacement ratio to sand of 25% and 50% and the third was used as a control sample. Thermal conductivity tests were performed in the Department of Mechanical Engineering at the American University of Beirut in order to estimate their impact on the energy required to heat or cool the building. The thermal conductivity of the sample with 50% replacement ratio was measured to be 0.490 W/m.K. The table below shows the concrete mix design for the three samples.

TABLE I
CONCRETE MIX DESIGN FOR THE THREE SAMPLES

Item	Control	25% Biochar	50% Biochar
Cement (g)	1250	1250	1250
Water (g)	625	625	625
Sand (g)	3438	2578.5	1719
Biochar (g)	-	859.5	1719

IV. CARBON FOOTPRINT ANALYSIS

Carbon footprint is defined as the total amount of greenhouse gas (GHG) emissions produced directly or indirectly during human activities. For buildings, this would be over the entire lifespan of the structure, from the beginning of

---

[1] 1 dunum is equivalent to 1000m^2

the construction phase and throughout the maintenance and operation phase.

"Building Carbon Estimator or (BCE)", developed at the American University of Beirut, was used to estimate the footprint of the Bsalim Recreation and Care Center. This was done at first without incorporating any of the sustainable materials. The materials typically used in the construction industry in the region today were inputted to the calculator.

The calculator estimates the carbon emissions of the raw materials used in the construction process, transportation from suppliers to the site location, and on-site activity [9]. Raw materials' carbon footprint is called embodied carbon. It entails the carbon emitted during the manufacturing process of the materials. The calculator has built into it a database for the major suppliers in Lebanon. The calculator is divided similarly to the Bill of Quantities (BOQ) of a project. The MasterFormat, developed by the Construction Specifications Institute (CSI), was used as the basis for the division [9]. The result is the carbon footprint of a building in Kg CO_2.

For the Bsalim Recreation and Care Center, the BOQ was generated after a quantity take-off exercise using a detailed model generated with Autodesk Revit. The model is shown in the figure below.

Figure 5: Autodesk Revit model of the Bsalim Recreation and Care Center

The quantities generated were used as input for the BCE. The calculator can also provide a rough approximation of carbon emissions throughout the operational lifetime of a facility. The average energy consumption per person in Lebanon is around 5720 kWh/year [10]. The development is expected to have 100 to 150 occupants. It is important to note that mechanical and electrical works were considered out of scope of this study and, therefore, carbon emissions result exclude these parameters. The carbon emissions calculated using the BCE, as well as the estimated quantities, are shown in the table below.

TABLE II
CARBON FOOTPRINT FOR THE BSALIM CARE AND RECREATION CENTER

Item	Quantity	Carbon Footprint (T CO_2)
Concrete (m^3)	5254	3882.4
Steel (T)	824	1457.6
Masonry (m^3)	2023	821.4
Finishes (m^2)	55627	90.3
Operation Phase	-	1500.4

The carbon footprint of the development using the proposed alternative materials should be calculated as well. Since hemp-reinforced concrete and recycled aggregate concrete will be used in concrete masonry units only, the study will be limited to the footprint of masonry blocks only as all other items of work will remain unaffected.

The hemp fibers used in concrete masonry units require very little to no processing. Furthermore, all plants can be considered to be carbon negative, since they absorb carbon dioxide and release oxygen. In fact, the hemp plant's fast rate of growth produces a lot of oxygen. Given these parameters, the carbon footprint of this raw material will be assumed to be equal to zero and only transportation and site activity CO2 emissions will be considered. The transportation distance, in general, can be considered from the Bikaa Valley to the manufacturing plant. Similarly, biochar is a byproduct of the process of pyrolysis. In other words, the aim of pyrolysis is a means of disposing of waste rather than producing biochar. Therefore, its embodied carbon will also be assumed equivalent to zero.

Recycled aggregates, on the other hand, do require processing. Concrete needs to be cleared from other demolition waste such as debris and steel. The concrete also needs to be crushed to serve as a replacement for natural aggregates. The processing of recycled concrete aggregates consumes around 0.04 million BTU[2] of energy per short ton[3] [11]. To convert this to carbon emissions, a calculator provided by the Environmental Protection Agency (EPA) was used. Using a replacement ratio of 40% by total weight of aggregates, the carbon emissions for processing this raw material is around 9.2 T CO2. The table below shows the difference in carbon emissions when hemp and recycled aggregates are used in concrete masonry blocks.

TABLE III
CARBON FOOTPRINT FOR DIFFERENT ALTERNATIVES OF CONCRETE MASONRY BLOCKS

Proposed Alternative	Carbon Footprint (T CO_2)	Percent Difference
Normal concrete	821.4	-
Hemp-reinforced concrete	609.6	25.8%
Recycled aggregate concrete	488.1	40.6%

Biochar was proposed as an alternative for use in plastering work. The following table shows the percent difference in carbon footprint:

TABLE IV
CARBON FOOTPRINT FOR BIOCHAR IN PLASTERING WORK

Proposed Alternative	Carbon Footprint (T CO_2)	Percent Difference
Normal Plaster	12.8	33%
Plaster with Biochar	8.6	

[2] 1 British thermal unit is equivalent to about 1055 joules

[3] 1 short ton is equivalent to 907 Kg

V. THERMAL ANALYSIS

Thermal conductivity, or the ability of a material to conduct heat, is the primary parameter used in this study to measure thermal performance. In general, it is better to use materials of lesser thermal conductivity in order to lighten the load on heating, ventilation, and air conditioning (HVAC) systems. A smaller thermal conductivity means that less heat enters the building space during the summer and less cold enters during the winter.

Autodesk Green Building Studio was used to simulate energy analysis on the Bsalim Recreation and Care Center. This cloud-based service imports the Revit model into its servers and analyzes the performance of the building elements. The climate of the location where the project is located is a primary factor taken into account in the analysis. The thermal properties of the materials of interest were all used in the energy performance analysis. The cooling load for the development throughout the year is shown in the figure below. None of the proposed sustainable materials were used in the model.
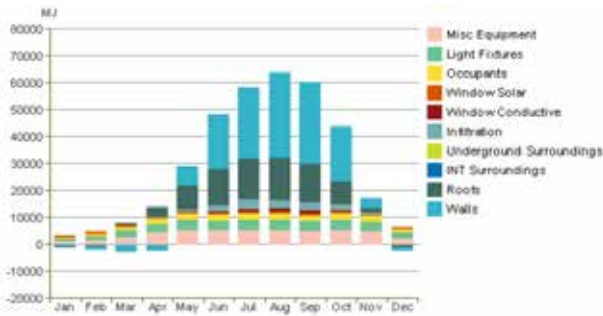

Figure 6: Cooling load for the Bsalim Recreation and Care Center

One of the many architectural innovations of the project is its terraced shape which makes it look as if it forms part of the side of a hill. In order for this to occur, there are around five story levels of retaining walls to sustain the load of the hill on one side. Since there is a lot of space in the building which is connected to the hill, passive energy can help control the temperature in the rooms of the facility. Passive energy, as opposed to active energy, is generated without the use of mechanical systems or equipment. As the temperature of the earth is stable, the presence of the hill throughout one side over the height of the entire building can help to steady the indoor temperature. To check on whether the hill does indeed have an effect, a Revit model was generated without modeling the mountain or the topography of the land. The resulting cooling load for the structure, for example, is shown in Figure 7.
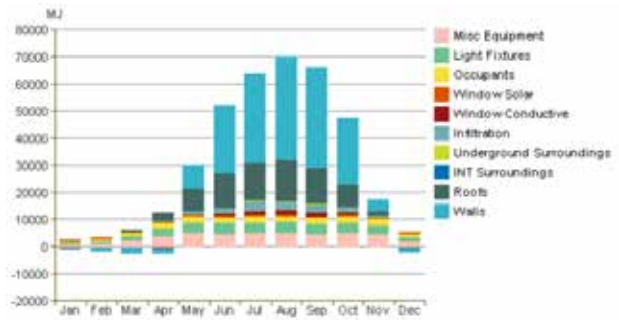
Figure 7: Cooling load for the Bsalim Recreation and Care Center neglecting the effect of passive energy

Both heating and cooling loads decreased due to passive energy. The thermal conductivities for the proposed alternatives were described in previous sections. The thermal conductivity for concrete masonry blocks can be estimated at 1.4 W/m.K [7]. The thermal conductivity of a plaster mix was found to be equal to 0.503 W/m.K after a test was conducted for this study at the American University of Beirut. The thermal conductivities for the materials used in the proposed alternatives were described in previous sections. The table below shows the difference between those values and the ones used in usual concrete mixes.

TABLE V
THERMAL CONDUCTIVITIES FOR DIFFERENT CONCRETE MIXES

Item	Thermal Conductivity (W/m.K)	Percent Difference
Concrete Masonry Block	1.4	-
Hemp-reinforced Masonry	0.984	29.7%
Recycled aggregate concrete masonry	0.930	33.6%
Plaster	0.503	-
Plaster with biochar	0.490	2.6%

The decrease in thermal conductivity of the hollow blocks led to a decrease in both heating and cooling demands for the facility. The cooling load if hemp-reinforced masonry blocks are used is shown below.
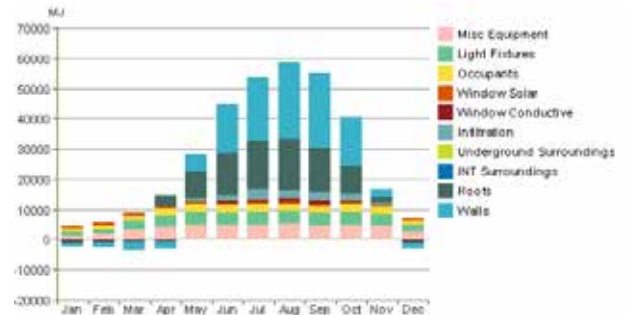
Figure 8: Cooling load for the Bsalim Recreation and Care Center when using hemp-reinforced masonry blocks

VI. WATER HARVESTING SYSTEM

Rainwater harvesting is defined as the collection of water from surfaces on which rain falls, and subsequent storage of this water for later use [12]. For the Bsalim Recreation and Care Center, the rainwater harvesting system will consist of Flo-Cell and Flo-tank module developed by "Atlantis Corporation Australia Pty Ltd". First of all, the amount of rain precipitation was acquired from the weather station closest to the project which is Beyrouth-Jdeide with an average precipitation of 856 mm per year. It was assumed that the amount of rain that precipitates on the center was only within its area and no runoff from other locations. The total volume of rainwater per year was calculated as follows:

Average precipitation per year x concrete areas x runoff coefficient + average precipitation per year x green areas with geodrains (Flo-Cell) x coefficient = 3,363 m3/year.

After that, the specifications of both "Flo-Cell" and "Flo-tank module" were studied to design a proper water harvesting system. Flo-Cell efficiently drains excess water while retaining a moisture level which is ideal for promoting plant growth in roof garden applications. One of the benefits of the Flo-tank module is its high void to solid ratio which maximizes the water storage capacity. Flo-Cell will be placed on the green roofs of the center and an inside earth tank will be built using an impermeable layer and the Flo-tank module.

Finally, the size of the tank was determined based on the precipitation rate. The tank will accommodate 1000 m3 of water which is equivalent to 1,000KL. According to the specifications, each Flo-tank unit has a storage capacity of 576.1 L. Hence, the needed amount of units to store the rain water is 1736 units. The surface area of these units is 485 m2 with a height of 2.17 m. The green roofs will be covered by Flo-Cell with 1% gradient. Pipes will take the harvested rainwater to the Flo-tank where it will be stored for later usage.

VII. COST ANALYSIS

The Bsalim Recreation and Care Center is characterized by a footprint of 4550 m^2 with a total built up area of 10275 m^2. The green roofing and gardens which are a main characteristic of the building form 45% of its total built up area. This is due to the fact that only 5651.25 m^2 of the project are allocated for indoor spaces leaving what is left for green roofs and indoor gardens. A cost build analysis was necessary to achieve the project at the lowest feasible price. The cost estimation was based on a unit price approach to follow the effect of every division of then MasterFormat on the total project outcome. There are six main divisions estimated in the section: Site Work, Concrete, Masonry, Opening, Finishes, and Specialties.

The site work has minimal effect on the total cost since the shallow foundation adopted was designed to follow the topography of the land to minimize any environmental disturbance and pollution. However, the type of the foundation plays an important role in the cost of the project. Choosing a raft foundation over isolated footings would increase the concrete division price by 32% and the project total cost by

14%. Another important agent in the cost is the amount of glass used in the building. The curtain glass walls constitute 40% of the project partition walls making a total surface area of 2544 m^2. If the glass material was optimized to the highest standards it would increase their total cost by 32% and the project initial cost by 6%. Finally the green roofing was assumed to be extensive. However if it was optimized to intensive, its price would dramatically increase by 233% and the project cost would rise by 20%.

The estimated cost under normal standards is around 5.6 million dollars. This price reflects the lowest construction cost under the current regional economy. However, if the regional economy retrieved its former status thus increasing the price of concrete and steel the project total price would increase to 8.7 million dollars.

VIII. CONCLUSION

This paper provides several design approaches which meets sustainability requirements for the different stages of the building. Future work will include implementing an integrative system of passive and active use of geothermal energy, and devising methods for managing construction waste.

REFERENCES

[1] Mayne, P.W. and Kulhawy, F.H. (1982). "K0-OCR relationships in soil". Journal of Geotechnical Engineering, Vol. 108, 700-900.

[2] Rankine, W. (1857) On the stability of loose earth. Philosophical Transactions of the Royal Society of London, Vol. 147.

[3] Blau, Peter J. "The Significance and Use of the Friction Coefficient." Tribology International 34.9 (2001): 585-91. Web.

[4] Awwad, E., Hamad, B., Mabsout, M., & Khatib, H. (2014). Structural Behavior of Simply Supported Beams Cast with Hemp-Reinforced Concrete. ACI Structural Journal, 111(6).

[5] Awwad, E. A., Choueiter, D., & Khatib, H. (2013). Concrete Masonry Blocks Reinforced with Local Industrial Hemp Fibers and Hurds.

[6] Matar, P., & Dalati, R. E. (2012). Using recycled concrete aggregates in precast concrete hollow blocks. Mat.-wiss. U. Werkstofftech. Materialwissenschaft Und Werkstofftechnik, 43(5), 388-391.

[7] Zhu, L., Dai, J., Bai, G., & Zhang, F. (2015). Study on thermal properties of recycled aggregate concrete and recycled concrete blocks. Construction and Building Materials, 94, 620-628.

[8] N.J. Themelis et al. "Energy and Economic Value of Nonrecyclable Plastics and Municipal Solid Wastes that are Currently Landfilled in the Fifty States" Columbia University Earth Engineering Center

[9] Harmouche, N., Ammouri, A., Srour, I., Chehab, G., & Hamade, R. (2012). Developing a Carbon Footprint Calculator for Construction Buildings. Construction Research Congress 2012.

[10] Dagher, L.,Ruble,I. (2010). "ChallengesforCO2 mitigation in the Lebanese electric- power sector". Energy Policy 38, pp.912–918.

[11] EPA. (2003). Background Document for Life-Cycle Greenhouse Gas Emission Factors for Clay Brick Reuse and Concrete Recycling. EPA530-R-03-017. Washington, DC: U.S. Environmental Protection Agency.

[12] "What Is Rainwater Harvesting? Rainwater Harvesting Methods And Techniques". *Sustainable.com.au*. N.p., 2016. Web. 21 Mar. 2016.

SUSTAINAIBLE SUSPENDED VILLA

Georges Akl, Jad Barouki, Nader Nasreddine, Somar Swaid, Mohammad Yassine

Civil and Environmental Engineering

American University of Beirut

gha22, jbb06, nsn08, sms79, mky06

As the tourism industry in Lebanon has historically contributed tremendously to the local economy, it stands today as a major source for revenue to the country, yet fails to deliver on innovative planning for future expansionary development beyond its luring naturalistic terrain.

Although the construction industry presents itself as a contingent pillar for tourism in Lebanon, recently, no serious attempts have been made to keep up the pace with the global trend of green buildings, new construction methods, and innovative designs. For luxury seekers, typical buildings are now off the table as means for them too seek ownership in the country. Hence, motives are clear to investigate innovative environmentally friendly alternatives capable of breaking the asymptotic trend endangering the country's development.

Our team proposes a concept of a high-end sustainable villa suspended from a cantilevering raft on a rocky cliff in Chekka. The mounting action exploits several cutting edge technologies in civil engineering, such as the use of a composite structural steel frame alongside typical cast concrete methods, use of tension piles, and investigating appropriate construction methods.

I. SOIL CHARACTERISTIC/ GEOTECHNICAL ASSUMPTIONS

An Important aspect of our design project is the geotechnical aspect. Many assumptions and correlations were made, about the nature and strength of the in situ soil (rocks) and their characteristics, in order to make a well-informed decisions concerning:

- The type of foundations to use for the "Raft/Slab" structure.

- Study the possibility of using drilled tension piles.
- Assess the necessity for soil improvement.

The way we proceeded, due to equipment and budget restrictions (Incapacity of doing boreholes, SPT, etc.…), was:

Analyze geological maps of the Chekka shoreline, available in the Engineering Library of the American University of Beirut. First inspection of the maps indicates the presence of sandstone dunes.

Compare these maps with online resources of geological mappings of the Lebanese ground and contact local authorities to gather extra available information about the nature of the soil, revealing that we can assume, with confidence, that our chosen location effectively contains sandstone dunes.

Investigate the characteristics of Sandstone. It is a sedimentary rock group, mostly consists of small grains of quartz. Physically, they are very hard, compact, fine-grained, equi-granular laterally and vertically homogeneous rocks of sedimentary nature. Laboratory tests and research data of other physical characteristics are presented in the table below:

Hardness	6.5 to 7 on Mohr's Scale
Density	2.3 to 2.4 Kg/cm3
Compressive Strength	90 to 140 N/mm2
Modulus of Rupture	16-40 N/mm2
Water Absorption	1.0 - 1.2 %
Porosity	Low to very low.
Weather Impact	Resistant

Table 1- Existing Soil Characteristics

After doing the aforementioned "desk study", multiple site visits were conducted in order to gather more information about the nature of the ground.

With all the possible information that we could gather about the rock condition and characteristics on our site, the following assumptions were made:

1. The rock present on site is sufficiently dense, resistant to weathering, with relatively high strength and few/no major cavities – If major cavities are to be discovered during the construction process, high-pressure cement grouting will be used to fill the cavities.
2. No need for major soil improvement in order to prevent liquefaction in case of an earthquake or to improve strength of in situ soil.
3. Excavation to a depth of three meters will not constitute a major challenge and could be done traditionally.
4. The Raft/Slab structure will be constructed on the rocks directly after the excavation is done and the site is leveled.

Manual calculations, optimized using Excel, eliminate the need for drilled tension piles, whereas a raft foundation of specific dimensions will sufficiently counter balance the structural weight of the villa, help distribute the weight on the existing rock bed and provide a good compacted layer for the parking lot.

The calculations are performed as follows:

The service load due to the building weight is 1335 Tons.

The moment due to this load on the cross-section of the raft on the cliff side is calculated as follows:

Global Moment $_{(Cliff\ Side)}$= 1335 x 8 = 10680 Ton.m

Taking X as the length of the raft foundation. Solving for X to get the sufficient length of the reinforced concrete raft foundation of 2m depth to counter balance the weight of the structure while taking an overturning factor of safety of 2.5:

1335 x 8 x 2.5 = 20 x 2 x 2.5 x 0.5 x X^2

The required raft dimension is 20x23 meters.

II. STRUCTURAL DESIGN:

Codes of Practice, Standards:
The buildings straining forces (gravitational and lateral) and the capacity of the structural resisting elements were determined in accordance to the following codes of practice:

- **"ACI 318-14"** for the determination of loads combinations, the design and detailing of various concrete elements (slabs, beams, columns, walls and foundations).

- **"ASCE-14" code:** for wind loads and analysis

Software:

The design was done with the aid of the following software programs:

- Autodesk AutoCAD: Draw and plan and detail any needed figure, with 2D features.
- CSI ETABS: ETABS is a sophisticated special purpose analysis and design program developed specifically for building systems. It is mainly used for modeling, and mainly the design of vertical elements.
- CSI Safe: Design of slabs, beams and foundations, reinforced concrete.
- S-Concrete: S-concrete is a stand-alone product that investigates, designs, and graphically details reinforced concrete beam in our case.

Design Assumptions:
In order to be able to start our design, we must start from a definite point, where we determine the main materials that are going to be used. Also we should recognize the structural elements presented in the building, and give a starting dimensions for each element to be checked then. Finally we have to load each member by the code's recommended load related to its type.

Materials:
The materials utilized in the construction phase of our building are concrete, steel and glass facades. The concrete sections are designed based on concrete strength of 30MPa. And for the steel the tensile yield strength fy=420MPa.

As we stated previously, materials that are typical for any construction project are steel and concrete. Now, employing these in our design was the main concern. For that an extensive design chain was adopted throughout the planning stage. Knowing that the structure is uncommon, thus cracking its challenges requires multiple iterations.

Starting with the leading design consideration mechanism of transferring the building loads safely to the rocky cliff. The first approach that was assessed is designing anchors drilled and placed in the cliff embedment to withhold the building loads. The design for this option was highly dependent on the availability of such structural members in our regions and the pieces of equipment was of great concern. Having this option held in mind, predesign studies were required. Those studies followed this chain of events:

First of all the conceptual plans of the villa were sketched by hand satisfying all the residential spaces requirements.

The dimensions of the structural elements were assigned logically. As any alternative structure, our structure contains the following structural elements: slabs, columns, walls, beams, and footings.

- Slab: slabs assumptions are concerned about its type and thickness. Clearly the designer prefers less thickness that offers him less cost. These assumptions depend mainly on the spans found through the slab, and the type of support used. Due to long spans found between supports (columns), we decided to use a two-way solid slab (flat plate). We will use a two-way solid slab with 30cm thickness. Later on analysis results are obtained to check the deflection and reinforcement.

- Columns: column sections were assumed with consistent dimensions and distributed with maximum spans of 6 m.

- Footings: Usually the thickness and dimensions are related to loads and bearing capacity of supporting soil. But, in our case the typical footings are not applicable. So, those factors are used to design the anchors withholding all the loads and shifting them to the rocky cliff bed.

The different loads associated with the structure were calculated for Dead and Live loads:

- Dead Loads:

The dead loads of the building include the self-weight of the structural components of the building being based on the first round dimensioning we did and on the materials specific unit weight. In addition to that, it incorporates the superimposed dead loads from finishing, fixtures and partition and was taken 0.5 T/m2.

- Live Loads:

The live load was obtained from ACI table 1.2 that was given to be 0.25 T/m2 on average for residential-type buildings.

After finalizing the predesign studies. The work shifts to the design of the anchors. So, literature review was done on anchors loaded laterally. The latest studies was done on anchors acting as restraints for retaining walls not for a block of concrete building. So, certain correlation was done to determine the dimensions of the anchors that tend to be very large and unfeasible in the available size. In addition, to that several inquiries were done with academic professionals who recommended diverging from the proposed option. Since it requires finite element modeling and detailed soil investigation to obtain a full simulation of the present rock formation. So, we altered the conceptual design as to involve a raft foundation situated on the rocky cliff face.

From the slab, concrete beams will cantilever from the cliff side, from which steel columns supporting the villa in tension are designed.

However, along the way we have studied various other options of the structure, such as a tapered slab, steel beams, and box sections. Due to different factors like deflection, the concrete beams option, mentioned above, was adopted. Below is a 3D Max model for better visualization of the concept.

Figure 1- 3D Max Model

III.   SUSTAINABILITY

In this concept of a suspended villa, our group tends to employ the tools at hand, embracing latest technological advancements, to come up with a virtually near off-grid design. Here, we investigate the technical aspects, feasibility and efficiency of a simple, yet innovative design of a trough solar power

system with the goal of maximizing energy self-sufficiency, along with other façade and structural considerations with the same goal as the latter.

Rotating Concentrated Solar Power RCSP:
The concept consists of a hyperbolic -highly polished- extruded aluminum reflective mirror of a specific radius, reflecting condensed sunlight onto the focus where a cylindrical pipe (receiver tube) is placed along the north-South axis. The fluid running in the receiver tubes heats up to around 550 degrees Celsius, which in turn generates steam in special compartments used by conventional steam turbines to produce electricity. Finally, a drive motor for rotation, programmed to track the sun's pre-determined path will be installed for optimal energy yield rates with around 75% of collector efficiency and overall annual efficiency of 15%. Originally, a similar design of vertical solar PV towers was done by one of the group members during his training period. In principal, they were designed to accommodate acceptable energy production rates utilizing vertical space instead of flat surface areas in constrained master plans. Moreover, it is necessary to note that historical data advocates the scalability of such systems into large-scale projects (solar farms) especially in large open areas, similar to the project's site. However, the RCSP will be used in this project for more reliability.

Operational Phase:
Weather statistical data in Lebanon shows that the systems will utilize 2720 sunhours with maximum efficiency and thus expected to operate around 31% percent of the year.

Based on historical data of existing plants, along with other statistics for energy consumption, greenhouse gasses production and sunlight average density, the following empirical results are expected upon utilizing the full available roof area of (320 m^2):

- The system will provide electricity for around 1.684 average Household. The remaining 0.684 will be stored in thermal storages in saturated salts.
- The operations will reduce the amount of greenhouse producing gasses by 7.57 tons annually.

However, utilizing full roof area is not feasible for structural and aesthetics reasons. Preliminary results suggests a maximum utilization of 156 m^2 which will result in:

- 82% energy efficiency.

- Annual reduction of 3.7 tons greenhouse producing gasses.
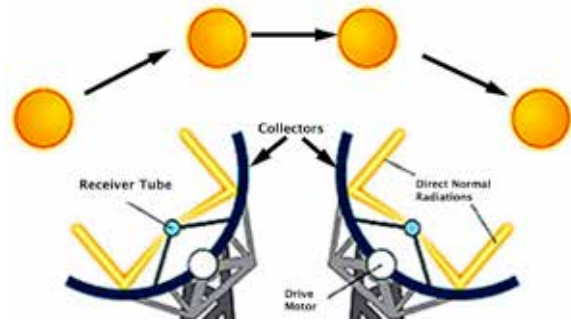
Figure 2- RCSP System

In conclusion, theoretical calculations show that these systems will hit high figures of self-sufficiency for the proposed design in terms of energy consumption.

Other structural and architectural considerations will ensure an increased energy conservation rates and include:

- Façade orientation.
- Open façade will ensure maximum light penetration.
- Choice of materials with high R-values, along with wall and façade sections detailing.

Other sustainability considerations include materials and ventilation.

Natural Ventilation (Breathing Building):
Natural ventilation -wind driven- will ensure lower AC use and better air quality circulating the building through dilution of pollutants, and follow the following guidelines:

- Urban Planning Conditions: Given the open space ensuring maximum airflow surrounding the structure.

- Flexibility of Openings (windows): this will ensure an inflow and outflow openings in each room, resulting in a better non-mechanical airflow circulation.

However, vertical ventilation pipes might be needed for better airflow within inner rooms like chimneys.

Materials:
Locally produced materials will help reduce the carbon footprint of the project and thus lower the environmental impact.

Local suppliers for different raw/modular materials are available for steel, cement, sand, aggregates, glass, and wood.

However, the marine platform will be imported from Saudi Arabia due to its critical use during the construction phase.

IV. PLANNING AND MANAGEMENT

The site is accessible by the main Chekka sea road coming both from the North and South. The site location is shown in Figure 2 below. The area under consideration is going to undergo site clearing and removal of trees to be able to create easy access from the main Chekka sea road and another side road to the site through a hard sandstone surface not requiring any further work to able to make a way through for the construction equipment and materials laydown.
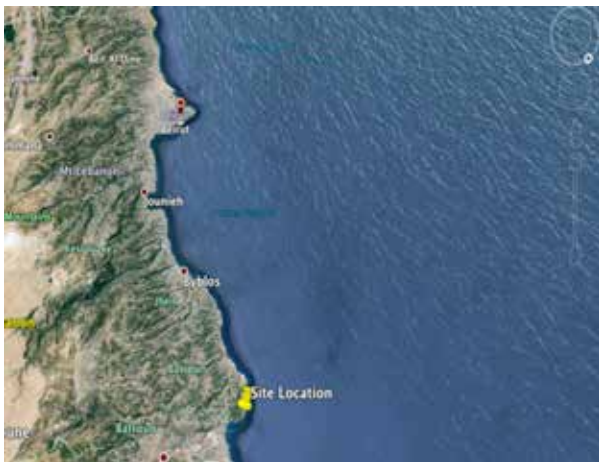
Figure 2 - Site Location (Google Earth)

Safety
The construction site will be strategically prepared and equipped to provide a safe construction environment for the workers and various construction machinery. This is provided by different safety procedures to be applied on site such as perimeter fencing, safety signs and instructions, emergency equipment, personal protective equipment, first aid kits, fire extinguishers, guardrails, security (surveillance cameras, watchmen, etc.), and proper electrical equipment.

The construction process of this structure will be supervised by an international safety control organization whose main objective is to control the construction process and ensure that the final product complies with the design criteria and safety standards, with other responsibilities such as fire protection, seismic protection and other hazards.

Marine Platform:
As for the marine platform, the most suitable option is called "Self Elevating Platform Fleet Expands" shown in Figure 4 below. As the platforms are of modular units, transport is easy and can be put to work any place in the world.

Some characteristics of this platform are:

First its Size is 18.5m by 24.5m but can be extended as required and its capacity is 400m/tons and the maximum height of elevation for operation is 22 meters.

However, other options were investigated. The second option is DP2 Apollo Self-Elevating Platforms, with dimensions of 87.5 m by 42 m having a capacity of 13250 tons and a crane capacity of 800 tons with a leg length that's 82.4 m. Another option is DP2 Innovation Self-Elevating Platforms having a dimension of 147.5 m by 42 m and a capacity of 31,440 tons with a crane capacity of 1500 tons and a leg length of 89m.

In conclusion, and based on our academic advisors recommendations, the first option (Self Elevating Platform Fleet Expands) will be installed and used for construction.

Figure 4- Self Elevating Platform Fleet Expands

ACKNOWLEDGMENT

REFERENCES

[1] Geological Maps in the American University of Beirut Library.
[2] Building Code Requirements for Structural Concrete: (aci 318-99) ; and Commentary (aci 318r-99). Farmington Hills, Mich: American Concrete Institute, 1999. Print.
[3] http://www.maritimejournal.com/news101/marine-civils/port,-harbour-and-marine-construction/self_elevating_platform_fleet_expands
[4] https://flexiblelearning.auckland.ac.nz/rocks_minerals/rocks/sandstone.html
[5] http://www.ce.udel.edu/courses/CIEG%20486/scheduling_notes103002.pdf
[6] https://weather-and-climate.com/average-monthly-Rainfall-Temperature-Sunshine,Beirut,Lebanon
[7] http://www.tradingeconomics.com/lebanon/electric-power-consumption-kwh-per-capita-wb-data.html

The Design of a Bioreactor Waste Landfill in Metn, Lebanon

Hanadi Hachem, Khaled Sakaya, Zaid Shtayyeh, Jad Zalzal
Department of Civil and Environmental Engineering
American University of Beirut
P.O.Box 11-0236 / (Department of Civil Engineering)
Riad El-Solh / Beirut 1107 2020
Lebanon
hmh58@mail.aub.edu, kjs04@mail.aub.edu, zws04@mail.aub.edu, jfz01@mail.aub.edu

Abstract: **The purpose of this project is to present the design of a bioreactor landfill for the Metn district in the Mount Lebanon governorate. The scope of this project reflects the amount of solid waste produced by the district, approximately 1.05 Kilograms per person per day of municipal solid waste [1]. It is assumed that a solid waste collection system entailing of evenly distributed waste bins around the district and the adequate number of waste trucks needed to collect the waste at a regularly maintained schedule to ensure the system leaves no excess waste. The system will require sorting of waste upon disposal (metals/glass, paper/cardboard, and most importantly organic) separately. This project aims at also handling the design of the bioreactor landfill itself given the several criterions needed to be considered, namely: socio-economic impact, climate conditions, political assessment, traffic impact, environmental impact, and sustainability. Structural components of the landfill will be looked at. These include: the liners, roads accommodating truck traffic, leachate collection and recirculation pipes, gas wells, number of cells in the landfill, and the ground water monitoring wells. Site visits and soil maps will aid in the geotechnical study required for design. The design and operation of the landfill require certified references; therefore the Environmental Protection Agency (EPA) landfill design manual will be an influencing reference to this project. It is expected that the landfill will operate as soon as completed and in approximately 6 years there will be significant energy produced from the circulated methane to generate sufficient power suitable for user demands. The site selection phase provided four suitable locations for the Bioreactor landfill that was much larger than the actual area required to meet demands. Therefore, there is ample land available for the creation, operation, and management of recycling plants for all non-organic materials in the vicinity of the landfill. This is to point out that a fully articulate solid waste management scheme is a plausible option and spatial requirement isn't a concern.**

I. INTRODUCTION

Twelve years after its planned closing day, the Naameh landfill was shut down at last, leaving the capital Beirut, and Mount Lebanon overflowing with trash with no prepared contingency by the government to handle what is known today as Lebanon's worst trash crisis in history. Population growth, urbanization and dwindling land areas are worsening the solid waste situation in the country, thus increasing the burden on the other existing - inefficiently operating dumps and landfills. This state is causing adverse effects on the environment and health, rendering the situation unbearable, and alerting for an immediate solution.

A possible solution to respond to the current trash crisis in Lebanon is to have a separate waste management system for each district. The scope of this project proposal consists of the feasibility study and design of an anaerobic bioreactor landfill for the entire Metn district (i.e. an distinct area in Lebanon) in the Mount Lebanon governorate. The feasibility study of this project will project the solution into the future, and clarify whether a similar thinking can be applied for the rest of the districts in all of Lebanon's governorates.

The concept of the anaerobic bioreactor landfill is relatively new to the emerging countries, including Lebanon; however, it has shown a positive impact on the environment and society in developed countries such as the United States. A bioreactor landfill is a landfill where leachate is recirculated in order to increase the degradation of wastes, and thus increasing the wastes' settling speed and methane generation to be used as an energy source. This bioreactor landfill will be able to provide power for some villages surrounding the landfill.

II. METHODOLOGY

The tasks achieved so far are the following:

- Site selection process using ArcGIS program, which allowed the overlaying of different maps to facilitate the analysis.
- Collection of data on solid waste composition, accurate waste volumes generated, existing systems and their efficiency. The data were collected from SWEEP-NET (solid waste exchange of information and expertise network).
- Population projection for the Metn district.
- Site visit to assess site conditions and collect soil samples, accompanied by lab tests to identify the soil characteristics.

- Design of the various components of the bioreactor landfill given the compacted weight density: area required for the landfill, final height of landfill, depth of excavation, liners and final cover system, slope stability, landfill gas generation and collection system, leachate generation and collection system, leachate recirculation, and access road pavement design. This is followed by an electricity generation study.
- Optimized layout of the landfill site including its different administrative components.

The remaining components of the design are surface water management, water monitoring wells, and landfill operations.

Figure 1. Anaerobic bioreactor landfill [3]

III.   ANAEROBIC BIOREACTOR LANDFILL CONCEPT

Bioreactors are still a novel technology so they are still in the development phase. Hitherto, only a few bioreactor landfills have been implemented, especially in the United States and Canada. Bioreactors are the same as sanitary landfills, except that they try to increase the biodegradation of waste thus increasing methane generation, increasing the rate of settling of the waste and decreasing the cost related to leachate management as leachate is recirculated in the waste. By monitoring the moisture and oxygen levels in the landfill, we can create conditions suitable for microbial activities.

There are three types of bioreactor landfills: aerobic, anaerobic and a hybrid (using both aerobic and anaerobic methods). All these mechanisms involve the recirculation of collected leachate through the waste in order to maintain a certain moisture. Thus the microorganisms responsible of degrading the waste will tend to work at an increased rate.

Anaerobic bioreactors seek to optimize the conditions of anaerobes present in the waste in order to increase the rate of the degradation. Anaerobic conditions are present in nearly all landfills without any intervention. The optimal moisture content for waste degradation is around 35 to 45%, however in typical landfills the moisture content is around 10 to 20%. Therefore additional moisture is introduced into the waste under the form of collected leachate or under the form of storm water and sewage sludge in case the leachate volumes collected weren't enough. The degradation in bioreactor landfills can be twice as fast as in a traditional landfill, therefore methane generation will happen in a shorter period of time [2]. Because of the shorter period of methane generation, the generation rate of methane will be greater in the case of bioreactors, so the gas wells must be designed to handle big peaks in methane concentrations. "Fig. 1" represents a typical anaerobic bioreactor landfill process.
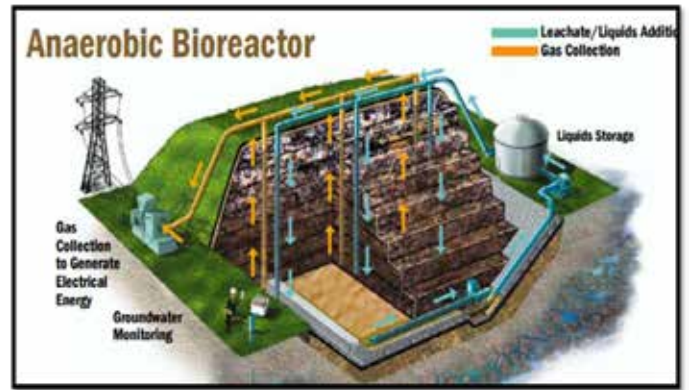
IV.   DESIGN COMPONENTS

This section explains the different design components covered so far, the methods used and the findings.

A.   Site Selection

The first step was to eliminate all areas that did not meet the requirements of a landfill site e.g. distance from water bodies, distance from cities and villages, distance from roads, slopes, elevation from sea level, etc.

Then, due to inaccuracy in data further investigation had to be carried out using Google Earth to make sure that no buildings fall within the acceptable areas. This resulted in four potential landfill sites: three in Mtein and one in Baskinta as shown in "Fig. 2".

Afterwards, the four sites were compared and graded in terms of 17 criteria such as: area, water vulnerability, type of soil, visibility, etc. And site 1 was selected as the final landfill site with an area of 1.09 Km^2 and at an elevation of 1400 m. The site selection process follows the same procedure followed in a similar project in Serbia [4].
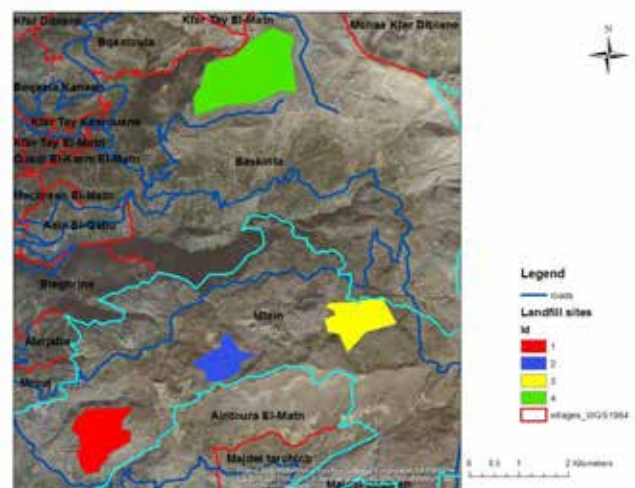
Figure 2. Potential landfill sites

B. Waste composition and characteristics

According to SWEEP-Net's findings [1] of Lebanon's Municipal Solid Waste (MSW) in 2015, the average Lebanese person generates 1.05 Kg of MSW per day, with 68.5% being organic material (including 16% paper and cardboard), 20.5% of non-organic recyclable material leaving only 11% medical or hazardous waste which has to be dealt with separately. For the purpose of designing the volume of the landfill, only the organic and paper/cardboard constituents will be considered and the compacted density of MSW in the landfill is required. This leaves a total of 68.5% by weight of the MSW contributing to the generation of landfill gas and leachate, as can be seen in "Fig. 3". According to Zekkos (2006), the average in-situ compacted density of MSW is 0.614 T/m^3 (assuming good compaction) [5]. The compacted density of the waste is important for the estimation of the total volume of the landfill. The waste is first compacted in the garbage truck, then further compaction is required on site using trucks that drive over the waste, and finally the waste is further compacted due to the load of the new waste dumped over it.
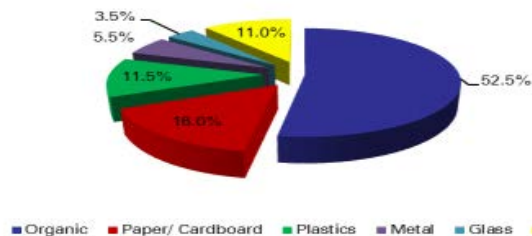
C. Population projection

Metn is the third most populated district in Lebanon. Therefor its population forecast is similar to that of Lebanon. This method is used where the future growth is expected based on past growth curves of a similar but larger city. A plot of population versus years was drawn in order to compare between logarithmic, exponential and linear growth curves. The exponential method was the best fit for Lebanon. "Fig. 4" displays the projected population of Metn for the next 35 years.

D. Landfill dimensions

The volume of the landfill is the result of the combination of: volume of waste, volume of daily cover (15% of the volume of the waste) and the final settlement of the waste (20% for bioreactors) [6]. This gives the total volume of 3,776,352m^3. 0.5Km^2 of the total area is allocated to waste disposal, this will allow for future expansions. The height was calculated to be around 19m. The landfill was divided into 7 cells (5 years per cell) in order to optimize landfill gas production and collection. The dimensions of each cell are 500x150mx19m.

Figure 3. MSW composition in Lebanon [1]
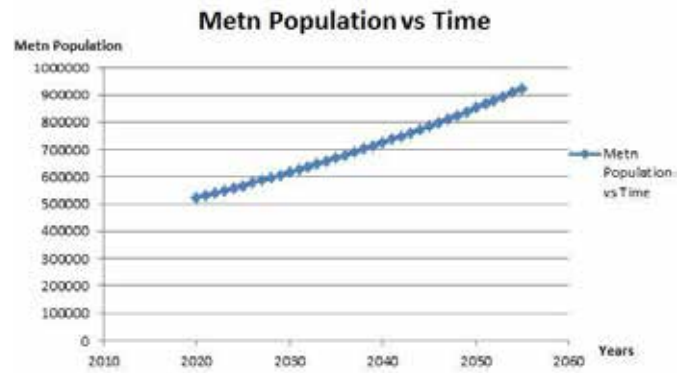
Figure 4. Metn population projection

E. Soil

Soil samples from the site were directly taken and tested in the lab according to ASTM standard tests for grain size distribution (Sieve Analysis) and plasticity (Atterberg limit tests). The results obtained showed that the soil was 95% silty clay with a low hydraulic conductivity and plastic behavior. This enhances the safety of the landfill against contaminated water (leachate) leaking from the landfill and into the groundwater. From investigation of nearby construction sites, information obtained regarding the soil profile showed that the top soil consists of 3 meters of soil sitting on top of a very deep limestone layer. Plus the water table level is very deep, so it wouldn't play a role in the design.

F. Bottom Liner System

Since the top soil is underlain by a layer of limestone rock which is very porous and fractured, further protection measures are taken in the form of double composite liners. This is a necessity given larger leachate production in a bioreactor system. A liner is a protective layer of low hydraulic conductivity soil and synthetic sheets to prevent leachate from seeping into the groundwater. The liner system is composed of:

1. 0.6m of soil layer as protection for collection pipes
2. Geosynthetics for separation and filtration, to protect the pipes from clogging
3. Leachate collection layer: 0.3m of sand or gravel K>10^{-3} m/s (includes the perforated pipes)
4. Geomembrane 60mm thick
5. 0.3m of sand leakage detection layer or gravel (K>10 4m/s)
6. Geomembrane 60mm thick
7. 0.6m of clay layer compacted in 4 lifts (0.15m per lift) (K<10^{-9}m/s)

Both geomembranes used are HDPE (High Density Polyethylene) sheets because they have a higher chemical resistance than other types of membranes, and can withstand high stresses. The chosen membranes have a density of 0.97g/cm^3, a tensile strength of 40.3KN/m, and a tear resistance of 0.33KN [7]. Adjacent geomembranes will be seamed together (attached together) using a hot wedge that can be done at a rate of 100m/hour [7].

G. Final cover

The final cover is the impermeable layer that is added after the closure of the landfill to prevent water from rainfall from infiltrating into the waste in order to control degradation. The design was chosen based on EPA guidelines, and is as follows:

1. 15cm top soil for freeze protection
2. 30cm drainage layer made of sand
3. 40mm HDPE geomembrane
4. 60cm clay
5. 30cm gravel for gas collection, gas distribution and leachate recirculation pipes.

The HDPE geomembrane used is thinner than the ones used in the bottom liner system because it is only exposed to rain, unlike the bottom liner that is exposed to leachate. So fewer chemicals and a smaller factor of safety are required for the final cover. A cross section of the final cover is shown in "Appendix A".

H. Slope stability

Slope stability is a crucial phase of the landfill design, as it has to be studied in great detail to prevent any slope failures. Failures can happen while excavating the cells or when the landfill is closed. Both of these failures will cause the liners to break, thus allowing leachate to seep into the groundwater if not repaired or replaced. The analysis was done using Slope/W, which uses the limit equilibrium method to assess the stability of the slope. The stability of the slope had to be assessed under the worst loading case, which is when one cell has just reached its capacity, and the cell adjacent to it has just been excavated. The inputs to the software were the following:

- Clay Sub-base: 1) Unit Weight= 20KN/m³, 2) Cohesion= 13 KPa, 3) Friction angle= 20°.
- Liner system: 1) Unit Weight= 4.71 KN/m³, 2) Cohesion=0 KPa, 3) Friction angle= 40°
- Waste (assuming 40% moisture content): 1) Unit Weight= 6.025 KN/m³, 2) Cohesion= 19 KPa, 3) Friction angle= 20° [8].

Since the software was a student version, the option of modeling the liner system was unavailable, so it was represented by a weak soil layer. The output of the software is the most critical slope along with the Factor of Safety of the slope. In this project, the slope was taken to be 2h to 1v for both the excavation and the slope of the waste, and the factor of safety turned out to be 2.3 (higher than the accepted value of 1.5). "Fig. 5" represents the output of the software.
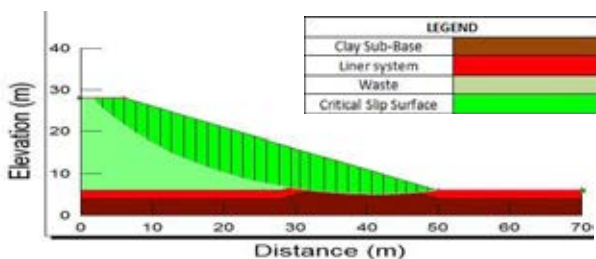
Figure 5. Output of Slope/W

I. Leachate collection system

The leachate produced from the deposited solid wastes must be collected in properly engineered system to prevent its buildup in the landfill, and to recirculate it in the landfill to fulfill its bioreactor function. The leachate collection system within one cell consists of a mesh of 20 horizontal pipes sloping downwards at a slope of 2% in a crisscross manner (10 towards each side of the cell), and 8 vertical pipes sloping downwards at a slope of 2% in a zigzag manner as well. The mesh will feed 8 vertical pipes outside the cell (along the longitudinal dimension of the cell) which will feed in turn 4 main pipe connected to 4 tanks that will collect the leachate produced.

The quantity of leachate produced is calculated as a function of the following parameters:
1. Rainfall quantity entering the cell the first five years = 187.2 lb.
2. Rainfall quantity entering the cell after its closure (starting year 6) = 46.8 lb.
3. Amount of leachate recirculated in each cell (starting year 6) = 600 m³/month.
4. Dry weight of each lift, their cover and their moisture content.
5. The capacity of each lift to hold water.
6. The soil specific weight of the cover material = 3000 lb/yd³.
7. The total weight of each lift (including dry weight of lift, water reaming and cover).
8. The average weight above each fill.
9. The quantity of gas produced.
10. The quantity of water consumed in the production of gas.
11. The quantity of water vapor in the gas.
Note: the calculations were carried out using the correct unit conversion to be consistent with the units.

The calculations of the produced leachate are made for each year throughout the life span of the landfill. Figure 6 shows the results of the calculations [9].

The calculations result in a peak amount of leachate produced at the end of year 5, which is equal to 17376.16 m³. This obtained number was used to design the pipes of the leachate collection system. The pipes are sized as follows:
1. The 20 horizontal HDPE pipes within the cell are sized with a 30 cm diameter.
2. The 8 vertical HDPE pipes within the cell are sized with a 20 cm diameter.
3. The 8 vertical HDPE pipes outside the cell are sized with a 40 cm diameter.
4. The 4 main horizontal HDPE pipes outside the cell are sized with a 60 cm diameter.

Figure 6. Leachate production

377

J. Leachate recirculation

Very little research is done concerning the optimal leachate recirculation systems and leachate recirculation rates. Therefore, the recirculation syste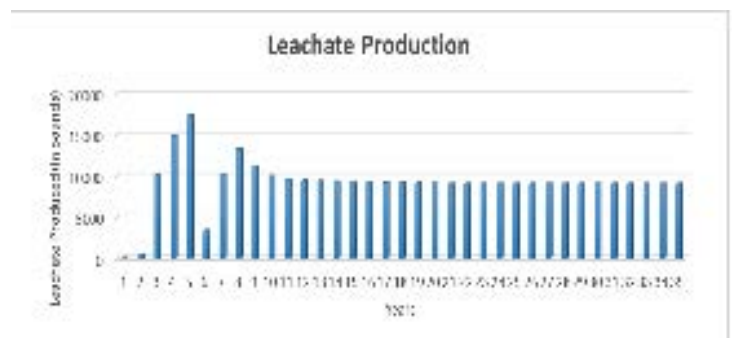m as a whole will be similar to a system implemented by Veolia in a bioreactor in France. This bioreactor had an area of 2.25 hectares and a height of around 21m. In the tests conducted by Veolia, the leachate recirculation rate that optimized gas generation was 200m³/month, with a methane generation of 35m³ in 200 days [10]. By scaling the bioreactor to the Metn bioreactor, the system will become as follows: a series of 75mm diameter, perforated PVC pipes along the width of each cell spaced at 25m from each other. These pipes will be located at the bottom of the gas collection gravel layer. The recirculation rate is 600m³ of leachate and storm water per month per cell. The storm water will be collected in a lagoon that will be mixed with the leachate produced from the landfill to provide the required amount of leachate to be recirculated (in case the leachate produced is not enough for recirculation).

K. Landfill Gas (LFG) collection

Every pound of waste deposited into the landfill will produce a cumulated 0.749 m³ of landfill gas as seen in "Fig. 7", which consists of 60% Methane and about 38% Carbon Dioxide, the remainder are traces of other harmful compounds [10]. The gas production curve in "Fig. 7" was developed using a triangular gas production function. It is calculated that a total weight of 14,707,506,000.0 pounds of waste will produce 10,379,297,092.000 m³ of LFG over 35 years at STP.

The landfill gas collection system aims at minimizing the risk of LFG migration outside of the landfill and minimizing the impact on air quality; this effectively reduces risk of landfill fires. By effectively controlling LFG, energy recovery from the landfill becomes permissible. The LFG collection system is affected by the composition of the waste, the leachate level, the applied suction, and gas well type. This project will require vertical gas wells drilled down to within 2 meters of the liner for active extraction of the gas. The connections to the overall network will provide flexibility to allow for waste settlement. The gas collection system will comprise of:

1. 700mm diameter HDPE perforated pipes placed at 30-meter intervals.
2. A collection network designed to handle the surge flow rate of 70 m³ / hour at year 5 per active cell per individual gas well.
3. Gas wellheads fitted on every well made of Polyethylene to control and monitor gas extraction and routing to the collection network.
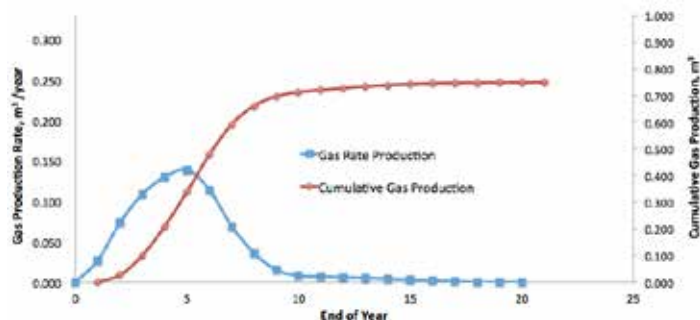
Figure 7. Gas production rate and cumulative gas production curves per pound of deposited waste showing stability at year 20 onwards.

4. The upper portion of the vertical gas well must be sealed with 1 meter of Bentonite from the surface downwards to provide impervious sealing and prevent LFG leakage into the atmosphere.
5. Perforated slot widths should be equal to 5 mm.
6. The pipes must be surrounded by 30mm well-rounded aggregates, each.
7. Condensation removal stations to 'dewater' the gas and prevent restriction of flow, this returns to the waste as leachate.
8. Extraction pumps fitted with centrifugal compressors providing a suction pressure of 20 KPa.
9. LFG monitoring wells placed around the perimeter of the landfill to check for unlikely ground migration.
10. A gas collection tank is provided to store the gas produced in the first 2 years of operation (3,538,540 m³). The methane will be stored under a pressure of 30 MPa in a tank with a volume equal to 11,950 m³. After which, the stored gas will be sent to electricity generation and any further gas production will be directly routed towards electricity generation.

L. Pavement Design

The amount of waste to be transported to the landfill creates a considerable volume of truck traffic that must not be neglected. Therefore pavement design was done using the AASHTO 1993 Method [13] in accordance with PaveXpress software for flexible pavement design.

Given the volume of waste to be hauled into the landfill over the operational lifespan of 35 years, an average of 63 trucks a day will be entering and leaving the landfill. The flexible asphalt pavement is designed taking into consideration:

1. The predicted loading on the pavement.
2. The reliability of the pavement.
3. The pavement structure, which is the Structural Number (SN) expressing the actual strength of the pavement.
4. And the Serviceable life of the pavement (35 years).
5. The strength of the soil beneath the pavement.

The results were conservative and provided a reliable design to account for further increase of loading on the pavement in the likely case that the landfill is expanded for future accommodation of additional MSW. The structure of the pavement facilitating the trucks is as follows:

1. Aggregate Subbase of 100 mm.
2. Aggregate Base of 100mm.
3. And Surface Asphalt layer of 140 mm.
4. Resting on a subgrade soil having a Resilient Modulus of 103 MPa.

Providing a 95% reliable design over 35 years it is safe to say the pavement will handle the continuous traffic load.

M. Site layout

The landfill site which can be seen in Appendix B, consists of the following units:

1. Administrative building: 10x30m
2. Garage for storage and repair of equipment and landfill trucks: 30x20m
3. Parking for employees and visitors: 20x20m
4. Gas tank for storage of landfill gas before transfer to power plant: 20x20m
5. Leachate storage tank for storage of leachate before treatment (included inside treatment plant)
6. Leachate treatment plant, to treat saturated leachate that cannot be recirculated anymore: 60x40m
7. 7 cells where the waste will be dumped: 500mx150mx19m each
8. Access roads for garbage trucks to reach the active cells: 10m wide
9. Drainage trenches surrounding the entire landfill to reroute surface water to an open lagoon to collect the storm water for later use

While a cell is being dumped in, the native soil that will be used as daily cover will be piled on the adjacent, unexcavated cell.

N. Electricity Generation

The bioreactor gas collected at the landfill site contains high amounts of methane compared to the relative maximum amount of around 55 %. This gas is going to be transformed to electricity using a double fuel internal combustion engine or gas turbines whenever small installations exist (i.e. up to 5 MW) and steam turbines whenever large installations exist.

➢ According to the World Bank, the electric power consumption in Lebanon per capita was estimated to be around 3499.37 kWh in 2011 per year [11] equivalent to 291.614 kWh per month.
➢ Assuming each household contains on average 4 people.
➢ Given that 1 MWh of electricity needs 700 m^3 of methane [12].
➢ The total landfill gas volume produced over 35 years is 10,379,297,092 m^3.

Thus, 14,827,567,270 kWh of energy is produced over the 35 years. For one year, a total of 3499.37 kWh is needed per household; therefore a total of 34993.68 kWh is needed per household for 10 years. Accordingly, a total of 423,721 houses near the landfill can be served for a period of 10 years.

REFERENCES

[1] Sweep-net.org,. 'Lebanon | SWEEP-Net'. N.p., 2015. Web. 1 Dec. 2015.
[2] Fei, X., Zekkos, D., & Raskin L.,(2015), Archaeal community structure in leachate and solid waste is correlated to methane generation and volume reduction during biodegradation of municipal solid waste, Waste Management 36, 184-190
[3] WM_Waste Management. (2014). THE LANDFILL The Bioreactor Landfill : The Future of Landfill Management, 1–5. Retrieved from http://www.wm.com/sustainability/pdfs/bioreactorbrochure.pdf
[4] Josimović, B., & Marić, I. (2012). Methodology for the Regional Landfill Site Selection. Sustainable Development–Authoritative and Leading Edge Content for Environmental Management, 513–538.
[5] Zekkos, D., Bray, J. D., Kavazanjian, E., Matasovic, N., Rathje, E. M., Riemer, M. F., & Stokoe, K. H. (2006). Unit Weight of Municipal Solid Waste. Journal of Geotechnical and Geoenvironmental Engineering, 132(10), 1250–1261. doi:10.1061/(ASCE)1090-0241(2006)132:10(1250)
[6] Carey, P., Carty, G., Donlon, B., Howley, D., & Nealon, T. (2000). Landfill Site Design. Environmental Protection Agency, 154.
[7] Sharma, H. & Lewis, S. (1994), Waste containment systems, waste stabilization, and landfills.
[8] Vajirkar, M. M. (2004). Slope Stability Analysis of Class I Landfills With Co Disposal of Biosolids Using Field Test Data.
[9] Tchobanoglous, G., Theisen, H., & Vigil, S. (2010), Integrated Solid Waste Management, Engineering principles and management issues.
[10] Beavan, R., Knox, K., & Powrie, W. (2009). A technical assessment of leachate recirculation. Environment Agency. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/291196/scho1109brjc-e-e.pdf
[11] The World Bank. (2010). Electric power consumption (kWh per capita). Http://Data.Worldbank.Org/, 7–10. Retrieved from http://data.worldbank.org/indicator/EG.USE.ELEC.KH.PC?cid=GPD_28
[12] Budisulistiorini, S. H. (2007). Electricity generation from landfill gas. Presipitasi, 3(2), 9–15. Retrieved from http://eprints.undip.ac.id/502/1/hal_9-15__sri_hapsari_.pdf
[13] Huang, Yang Hsien. *Pavement analysis and design*. 1993.

APPENDIX:

A. The figure below shows the leachate collection and liner systems cross-section and materials along with the appropriate layer thicknesses.

Leachate Collection and Liner Systems Cross Section

French Drainage Protection Layer 100mm
Leachate Collector Pipe HDPE Φ 300 mm

Soil Protection Layer 600 mm
Geotextile Filter Membrane 30 mm
Gravel Drainage Layer 300 mm
Geomembrane HDPE 60mm
Sand Leakage Detection Layer 300 mm
Geomembrane HDPE 60mm
Clay Layer 600 mm

B. The figure below displays an AutoCAD drawing of the landfill layout showing the buildings, roads, and overall boundary of the landfill and its cells. The drawing is to scale but lacks dimensions for the purpose of prioritizing the components' spatial distribution on the project site.

PLANVIEW LAYOUT
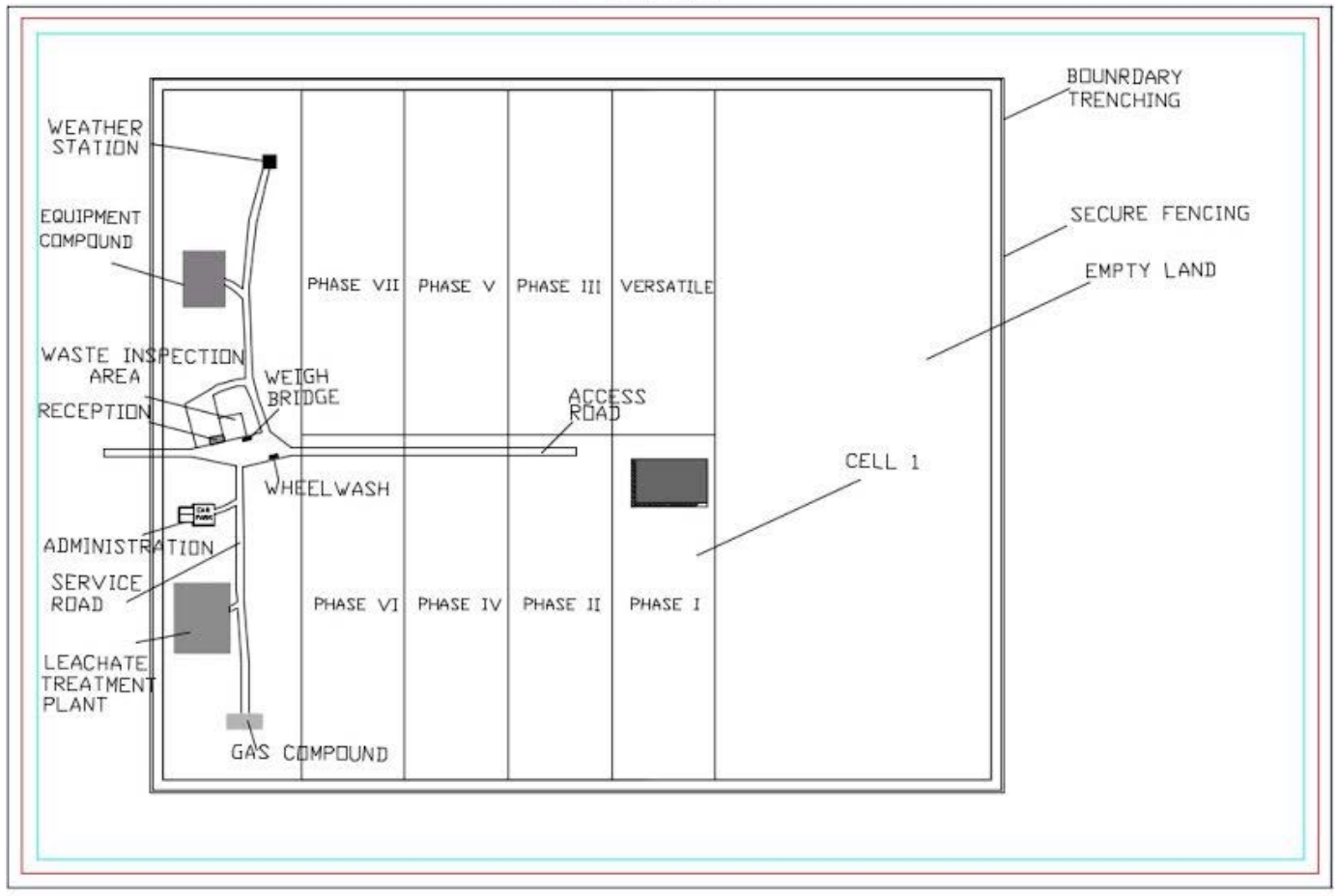
C. The figure below is showing the designed access road within the project site based on the AASHTO 1993 method. It is a simple cross section with the structural components to the road drawn to scale. The service road also is shown which will be the facilitating road leading to the active cell where trucks will daily commute to deposit waste and leave.

ACCESS ROAD
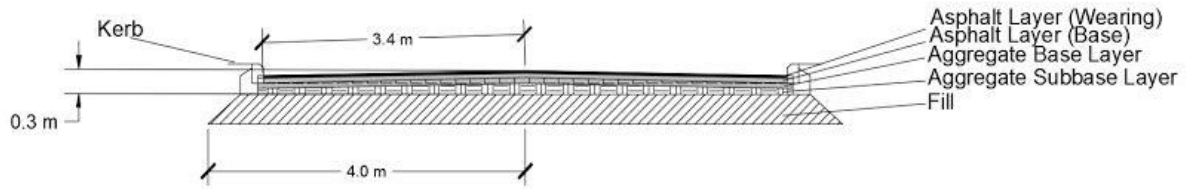
SERVICE ROAD
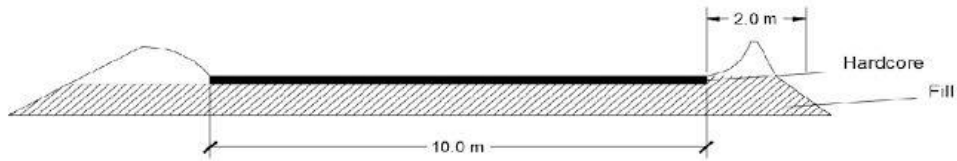
Traffic Estimation Using Implemented and Proposed Sensors in the Greater Beirut Area

Omar Masri, Omar Nahas, Nour Nachabe
Electrical and Computer Engineering Department
American University of Beirut
Beirut, Lebanon
oym00@mail.aub.edu, omn02@mail.aub.edu, njn07@mail.aub.edu

Abstract- **The Lebanese Transportation sector is in complete disarray leading to social, environmental and economic problems. The aim of this project is to alleviate some of the heavy traffic load from the Greater Beirut Area roads. Under the supervision of the National Committee for Scientific Research (NCSR), a traffic estimation algorithm will be developed in order to be implemented in an application to alleviate the deficiency in the transport sector. The proposed method is to fuse data from a speed sensor or GPS probe with the data from the counters upstream and downstream the link. Travel time estimated will then go towards the pursuance of the development of the Advanced Traveler Information System (ATIS). The system would inform users of what roads are congested and which ones are not. It also informs them of the duration of the trip on a certain route given a source and destination. This will allow policy makers to produce informed and helpful decision that will hopefully solve the Lebanese congestion problem. The impact of such a system on the environment and specifically with the gaseous emissions was assessed in order to emphasize the benefits if such a system is implemented in Lebanon.**

Overall the project is aimed at creating a comprehensive framework to mitigate the effects of the major problems, traffic congestion and pollution, from which Lebanon has been suffering over the last three decades.

I. INTRODUCTION

Congestion and lack of efficient road transportation is a major social, economic and environmental problem in Lebanon. The project aims to address this problem, part of the many problems, which plague the transportation sector. By having correct traffic estimation data, officials would be able to take more appropriate decisions and hopefully alleviate some of the congestion that currently infects the Greater Beirut Area (GBA).

The National Council for Scientific Research (NCSR) is currently working with Dr. Azzam Mourad who is an associate professor of Computer Science at LAU to develop solution for the Lebanese traffic problem. Dr. Mourad is also the Coordinator of Associated Research Unit on Intelligent Transport & Vehicular Technologies. This final year project is part of a greater project that is currently being undertaken by seven other universities in Lebanon and abroad. The traffic estimation algorithm that is being developed is going to go towards a mobile application that will serve as an Advanced Traveler Information System (ATIS).

II. MOTIVATION

The transportation sector allows for mobility and accessibility. It allows people and freight to be mobile so they can go from location A to location B by using any of the modes that are available within the country or globally. Moreover, it eases the accessibility of certain services for people. For example, a person currently at home cannot have access to all the food supplies in a market without the use of the transportation sector. So with transportation, people, services, industries, companies, corporations etc. are connected. This allows the production and consumption of products to occur at different locations. Accordingly, economic growth has always been dependent on increasing the capacity and rationality of transport.

In Lebanon, the road network consists of 22,000 Km of roads, of which only 6,380 Km are considered as paved roads while the rest are governed by municipalities. The land transport sector suffers from major problems including lack of organization. Thus major cities especially the GBA suffer from high amounts of congestion. Most arterial roadways, highways and intersections suffer from severe under-capacity and delays. In addition, the daily passenger trips are expected to exponentially increase [1].

The Lebanese vehicular fleet is composed of 1.58 million vehicles of which 85% are passenger cars, 0.9% are buses, 8.9% are trucks and 5.2% are motorcycles [2]. Moreover, the Lebanese fleet is mostly compromised of old vehicles with more than 70% of the cars on the road being over 10 years old (as can be seen in Figure 1 below).
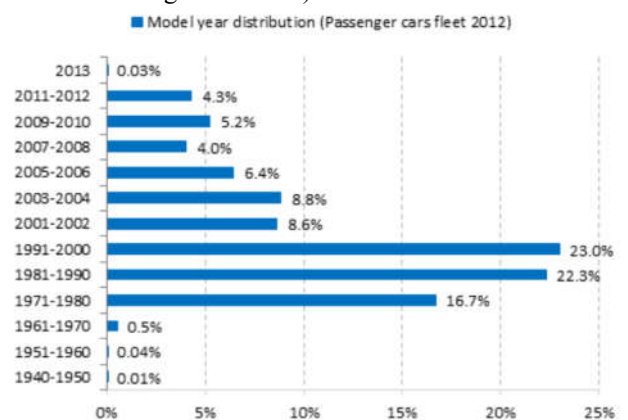

Figure 1: Private Fleet Age Distribution. Source: [2]

Lebanon suffers from a lack of public transportation infrastructure. For example, only 0.9% of vehicles on the road are buses. The bus grid is characterized by being insufficient, lacking in infrastructure (such as bus stations and depots) and possessing old and polluting buses. Furthermore, the old Lebanese railroad system is out of commission. Lebanon therefore suffers from a ratio of 526 cars/1,000 persons [1]. This high vehicle to person ratio characteristic of Lebanon leads nicely into the next motivation, which covers congestion. Congestion is the condition within the road network when the capacity of a section of the network is exceeded by the demand at a certain point in time. It is characterized by increased travel time, very slow speeds, erratic stop-and-go driving, and increase in operation costs [3]. The results of the congestion in the GBA are: significantly higher than average fuel consumption and higher levels of carbon monoxide emissions. Congestion is a major problem in the Lebanese network and is becoming harder to ameliorate. According to Lebanon's Second National Communication to the UNFCCC in 2011, the GBA accommodates more than 5 million daily passenger trips. It was shown that more than 50% of these trips have a distance of around 5 Km and a 15% of the travel time spent stationary [1]. A clear pattern emerges from Figure.2 pertaining to a high level of accelerations at low speeds. Most of the road going traffic in the GBA is concentrated at -0.5 to 0.5 m/s^2 on the x axis and at 0 to 20 Km/h on the y axis. This means that drivers spend most of their time on the road either braking or accelerating at low speeds. This leads to the conclusion that most cars in the GBA suffer from stop and go traffic (congestion).
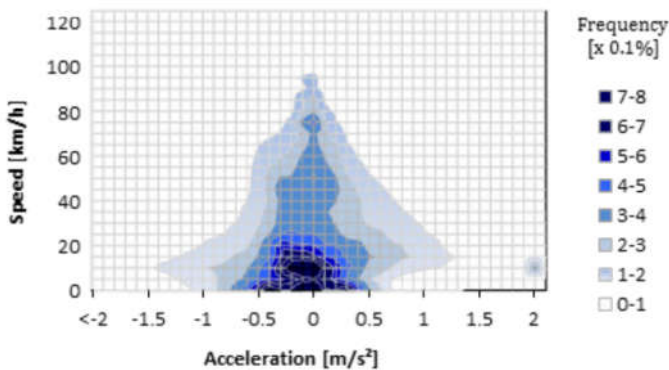
Figure.2. Acceleration distribution in the Greater Beirut Area.

Source: [2]

The overall motivation for the traffic estimation algorithm is to provide the foundation for the NCSR team to implement the mobile application that will serve as an Advanced Traveler Information System. This application will give users specific information in real-time about traffic and congestion problem. This will in turn influence the users' road choice, departure time and even the decision of whether to get on the roads or not. Therefore, the aim is to improve the efficiency of the transportation sector, reduce congestion and

traffic and decrease overall fuel (energy) consumptions in Lebanon.

III. Solution Methodology

To estimate traffic in the Greater Beirut Area is a problem to be tackled piece by piece. Thus it was branched and travel time within the links was given the most focus.

The travel time within the link is a hard to estimate value. The classical analytical procedure for travel time estimation is a typical method that makes use of sensors upstream and downstream the link in order to plot cumulative car counts at the two positions. In Figure 3, these are U(t) and D(t) [19]. Assuming: a) First-In-First-Out (FIFO) discipline is respected for all vehicles traversing from u/s to d/s (i.e., there is no vehicle overtaking); and b) the vehicles are conserved (i.e., there is no loss or gain of vehicles between u/s and d/s). The vertical distance (along Y-axis) between the two plots at time t defines the instantaneous number of vehicles (n) between the two locations. The horizontal distance (along X-axis) for count i defines the travel time (tt$_i$) for the i$_{th}$ vehicle [4]. The classical analytical principle for average travel time estimation defines the total travel time for all the N vehicles departing during the travel time estimation interval (T$_{EI}$) (from the location d/s) as the area (A) between the two cumulative plots. Average travel time per vehicle is the ratio: A/N [4]. Even if FIFO discipline is not respected, the area (A) between the two plots represents the total travel time as long as all the vehicles which arrive at upstream during time t$_1$ and t$_2$ actually depart at downstream during time t$_3$ and t$_4$, and vice versa.
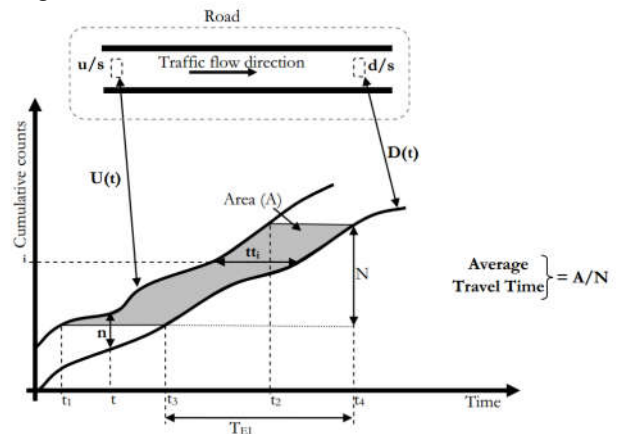
Figure 3: Classical analytical procedure for average

travel time estimation Source: [1]

This method is vulnerable to relative deviation. This condition is difficult to achieve in real life situations due to the following: Detector error and mid-link sources and sinks. In addition, the procedure doesn't allow the differentiation of the flow between different link movements since it is a cumulative average. Another source of data can be a travel times of vehicles equipped with GPS devices that are called probes. The data from both sources is redundant and complementary and thus can be used in a solution that merges multi-sensor data in order

to estimate travel times [5]. Both Bhaskar and Hage developed methods that combine both the probes and vehicle counts in order to come up with an algorithm to provide better travel time estimates. Data fusion is the processing tool that takes into account the quality of the data provided by each source with the aim of increasing the accuracy, reliability and robustness of the estimation [5]. Ashish Bhaskar in his thesis [6] estimated the cumulative number of vehicles plots on the upstream and downstream of a link. These numbers (or cumulative plots) are deterministically corrected by adding the data from the probe vehicles.

As part of the project, a speed sensor was tested as a supplementary data source in order to remedy the relative deviation within the classical analytical procedure.

Thus, in order to tackle the travel time issue within links, a speed sensing system was implemented. Extensive testing was done in a closed area as well as on public roads for the sensors in order to affirm its reliability and integrity.

Further testing took place when a full fledge test was conducted. During this test, we collected data from ongoing traffic as if simulating NSCR sensors. In addition, the speed sensing system and the GPS probe were tested in the same environment and at the same time thus allowing us to compare their respective results.

Finally, the experiments discussed above yielded the required data to produce and test a traffic estimation algorithm based on travel time within individual links.

IV. IMPLEMENTATION

A. Speed Sensor Implementation

In this section of the paper, the overall preliminary testing and implementation of the speed sensor system will be taken into consideration.

An Arduino board serves as the programmable controller. The advantage of this board is that it allows programming with high level languages such as Java. For each sensor installed, an LED was installed, indicating if the sensor is triggered or not.

Finally, it was time to test both sensors running at the same time and observe the results. Figure 4 highlights the circuit that was employed by the team. The distance between the sensors can be adjusted as needed since the cables are quite long (around 1.5 m). In this particular case, the IR sensor on the left has detected an object passing in front of it. Hence the left LED is illuminated. The timer is triggered at the moment that the left sensor detected an object. The timer is stopped when the object passes through the right sensor's field of view.
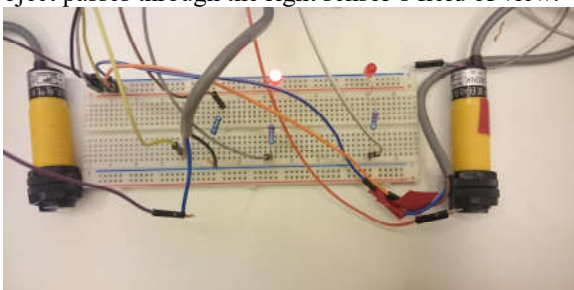
Figure 4: Preliminary speed sensing system test

B. Traffic Estimation Algorithm Implementation

Implementing the Classical Analytical Model in MATLAB

In order to investigate on how to fuse the data form the counters and the speed sensor data, the classical analytical model was to be implemented in code. The model takes in two plots and calculates the area between them given a fixed Y axis interval (as seen in Figure 5). The area is then divided by the same Y-axis interval in order to attain the travel time per vehicle.

The model was implemented in MATLAB since the need for it was for simulation and since MATLAB has built in functions that ease the process. The main built in function to be used was "trapz" which calculates the area under a graph using the trapezoidal method, which is sufficient to our purpose. The function takes a list of abscissas and a list of ordinates in the format trapz(X, Y). The input to the function was inputted in reverse order i.e. trapz (Y, X). The function then calculates the area between the Y-axis and the graph not the area between the X-axis and the graph. In order to calculate the area between the two graphs, two areas were calculated from the Y-axis towards the two plots respectively and then their difference would represent the area between the two plots (as shown in the Figure 5 below).
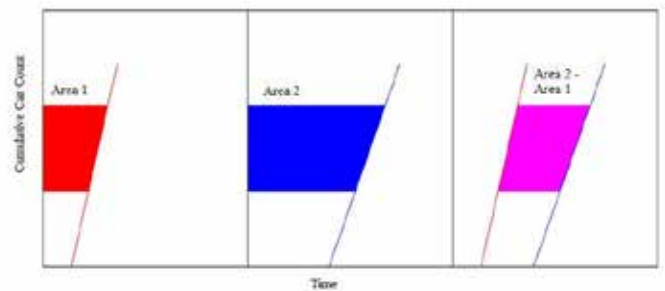
Figure 5: Illustration on the implementation of the

Classical Analytical Model

Using the Speed Sensor to improve the results from the Classical Analytical Model

As discussed before, the Classical Analytical Model is mainly sensitive to the conservation of flow. The method would result in erroneous travel time estimations if there are mid-link sources or sinks. These mid-link infrastructures can act as sources or as sinks dependent on the time of day and the day of the week [7]. A significant proportion of the flow can be lost or gained by these infrastructures. This proportion is dynamic but on average 10% of the flow is lost or gained [7]. Figure 6 shows the effect 10% loss in a mid-link sink on the classical analytical procedure.
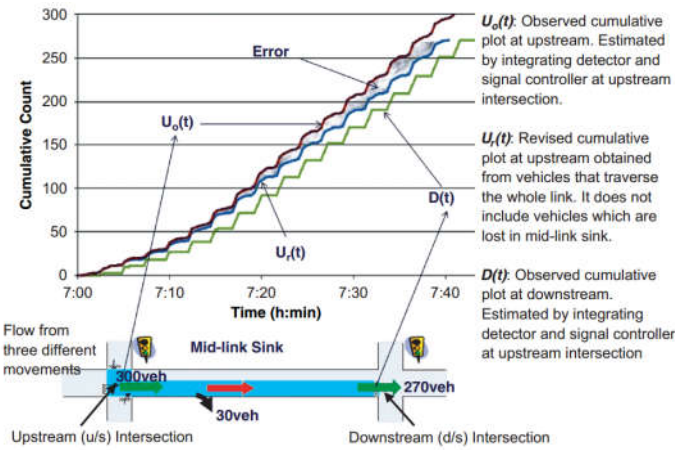
Figure 6: Effect of mid-link sink on the classical analytical procedure. [21]

The Upstream count $U_0(t)$ contributes to both the downstream count $D(t)$ and the vehicles lost in the mid-link sink. The main concern is to identify the vehicles in the upstream count $U_0(t)$ that only contributed to the downstream count $D(t)$. That is represented in the $U_r(t)$ plot where only the vehicles that contributed to the downstream count are considered. $U_r(t)$ is unknown and the travel time is calculated using $U_0(t)$ and $D(t)$ which results in the error represented in the shaded area between $U_0(t)$ and $U_r(t)$. Leaving this relative deviation unchecked and uncorrected the error might grow exponentially. Considering the case of a mid-link source, the upstream count will become less than the downstream count: $U(t) < D(t)$. Then the area between the upstream and downstream the link will become negative and thus the travel time cannot be calculated. This situation was observed when the field test was conducted. This forces the addition of vehicles to the upstream count $U(t)$ in order to act as upstream vehicles for the extra count coming out of the link in $D(t)$.

The issue is to be addressed by integrating supplementary data to that of the counts in order to remove the relative deviation between the two plots. The initial attempt was to use data from a mid-link speed sensor in order to remedy the issue.

Problems associated with such an attempt:

1. Speed Sensors attain spot speed and that might not be representative of the average speed on the link.
2. Speed sensor developed might not get the speed of all the cars crossing the link.
3. Speeds cannot be associated with time given the design of the speed sensor algorithm.

Assumptions associated with this attempt:

1. The point in the upstream count $U_0(t)$ to be adjusted is only the top point in the graph Ymax.
2. The average speed brought by the speed sensor is to be used to adjust the point in the graph.
3. First In First Out (FIFO).

The algorithm developed by Bhaskar in his paper [7] will be used, given the assumptions above, in order to improve the travel time by the counters (classical procedure). The method is modified in order to be used with the speed sensors.

The modified algorithm applied on the data from the Speed Sensor:

1. Cumulative plots are defined by integrating detector counts with signal timings.
2. Average travel time is calculated from the speed sensors.
3. Point through which U(t) should pass are defined at the largest point in D(t).
4. U(t) is redefined by vertical scaling and shifting the plots so that it passes through the points defined in Step 3.
5. Finally, average travel time is defined as the ratio of the area between the plots and number of vehicles departing.

How to Redefine U(t):

Supplementary data sources are used to identify points which U(t) should pass in. Take the point $[t_{ref}, U(t_{ref})]$ as the point where there is confidence that U(t) should pass in (U(t) already passes through this point) and point $[t_p, Y_p]$ through which U(t) should pass (U(t) doesn't pass through this point yet).

1. Before point tref there is no correction applied: correction= zero.
2. Between point tref and tp U(t) is scaled vertically.
3. After point tp U(t) is shifted vertically such that it stays parallel to the previous curve.

$$U(t) = U(t) + \text{correction}$$

$$\text{correction} = \begin{cases} 0 & \forall t \le t_{ref} \\ (\text{scale} - 1) * (U(t) - U(t_{ref})) & \forall t_{ref} < t < t_p \\ (\text{scale} - 1) * (U(t_p) - U(t_{ref})) & \forall t \ge t_p \end{cases}$$

$$\text{scale} = \begin{cases} \dfrac{Y_p - U(t_{ref})}{U(t_p) - U(t_{ref})} & \text{if } U(t_p) \ne U(t_{ref}) \\ 1 & \text{if } U(t_p) = U(t_{ref}) \end{cases}$$

NOTE: what is being corrected is the rank of the vehicles coming into the link. The presence of a mid-link source or sink only affects the rank of the vehicles in the plot, which results in the relative deviation.

Figure 7 shows an example of a mid-link sink where the vehicles C and D didn't exit the link normally and thus are not counted in D(t). Thus in U(t) the ranks of points E, F, and G are 2 slots higher than what they are. In order to remove the relative deviation, their ranks should be decreased accordingly by taking the point B as the reference point. From point B to point E the graph is scaled. From point E to G the graph is shifted down two slots and kept parallel to the original plot.
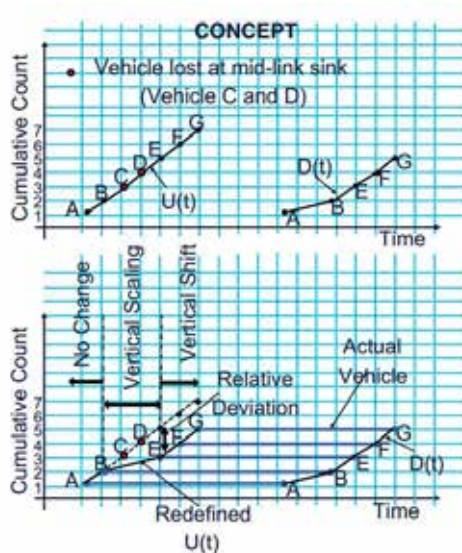
Figure 7: Example on Redefining the U(t) [21]

How to Define Points Where Plot Should Pass:

The probe data comes in as two lists $[t_u]$ and $[t_d]$ where the first is all the time stamps when the probe reached upstream the link and the second is the time stamps when the probe reached downstream the link. The jth element in both lists corresponds to the jth probe.

In order to identify the points where the plot should pass the two lists are sorted in ascending order. $[t_d]$ is sorted since the points are fixed to D(t). $[t_u]$ is sorted to make sure the redefined U(t) is monotonically increasing and satisfies SUM(delta t)=0. The sorted lists are then used to identify the points where U(t) should pass: $[t'_u, D(t'_d)]$.

Using the speed sensor, the average speed is used in order to find the point where the plot should pass. The point where the plot should pass in becomes $[t_{max}-t_{speed}, Y_{max}]$. This is the max point in D(t) $[t_{max}, Y_{max}]$ with a modified abscissa.

How to Define Reference Points:

The initial reference point is defined in free flow conditions and then the second point is fixed according to that one. So if $[P_1, P_2, P_3, ..., P_n]$ is the list of n points from where U(t) should pass, then for redefining U(t) for point P_i, the reference point is P_{i-1}. Thus the last point where U(t) should pass becomes the reference point when there is a need to redefine U(t).

Using the Probe to improve the results from the Classical Analytical Model

Probe vehicles (e.g., taxi fleet), are vehicles equipped with vehicle-tracking equipment (e.g., Global Positioning System) and can provide data for a vehicle's trajectory (time stamp and position coordinates) and hence its travel time. They represent the random sample from the population of the vehicles traversing the link. Therefore, average travel time for all the vehicles traversing the link can be estimated by statistical sampling techniques [7].

Data is fused from the detector and probe vehicles, with the goal of improving the accuracy and reliability of the estimates on travel time. This would relieve the errors induced by interaction with external controls such as signals and significant proportions of flow from and to mid-link sources and sinks found in urban areas.

The algorithm as by Bhaskar who developed it as part of his CUPRITE procedure is as follows:

1. Cumulative plots are defined by integrating detector counts with signal timings.
2. Probe information, list of [tu] and [td], is defined by fixing real probe data with D(t).
3. The above list is appended by virtual probe data if conditions for virtual probe are satisfied.
4. Points through which U(t) should pass are defined.
5. U(t) is redefined by vertical scaling and shifting the plots so that it passes through the points defined in Step 4.
6. Finally, average travel time is defined as the ratio of the area between the plots and number of vehicles departing

V. EXPERIMENT DESIGN AND TESTING

A. Speed sensor testing

The first test was done in order to verify the integrity and validity of the Arduino code for the speed sensor. The speed sensor at its core was designed to measure the time required to cross a predefined distance between the two sensors and then eventually come up with the speed given the distance.

Limitations in the Sensor:

- The speed sensor works for a single car at a time.
- The speed sensor works on a single lane at a time.
- The speed sensor acts in a single direction only.
- The speed sensor only outputs time when the IR sensors are activated correctly.

The limitations exist due to the algorithm followed within the Arduino chip. The timer is only activated when the first IR sensor, which is supposedly found upstream, is activated. It is only stopped when the second IR sensor is activated and only if the first was already activated. The main cause of these limitations is that the upstream IR sensor cannot be triggered after it was initially activated since it would be waiting on the stimulation of the second sensor to terminate the timer. A simple remedy to that would be the usage of a queue data structure. Each time the upstream IR sensor is activated a time stamp is pushed in the queue and when the downstream sensor is activated the top time stamp is popped from the queue and compared to the current time stamp to give the time elapsed.

Methodology of Testing

The procedure is as follows:

1. The sensor was setup
   a. The Arduino was placed in a convenient place and connected to a laptop
   b. The two IR sensors are placed 2 meters apart on two stools.
   c. The Serial Com is activated on the laptop.

2. A vehicle was driven around in front of the sensors inside the range of the IR sensors.
3. The readings were taken from the laptop.
4. Readings were taken from the speed-o-meter inside the vehicle.
5. Comparison between the speed sensor results and actual speed-o-meter reading was performed.

Results and Analysis

The experimental procedure was performed several times in order to verify the integrity of the sensor. Table 1 summarizes the results:

Table 1: Speed sensor experimental results

Time (sec)	Sensor Speed (km/h)	Speed-o-meter (km/h)
1.47	4.9	5
1.07	6.7	10
0.80	9	10
0.96	7.5	10
0.59	12.3	15
0.51	14	15
0.52	13.8	15

The first column is the raw data from the sensor as a form of time in seconds. The second column is the speed calculated from the sensor. The last column is the speed read from the speed-o-meter.

Illogical readings were appearing when the speed sensors were placed at a small distance from each other. Therefore, a distance of 2 meters was used between the two sensors. This distance acts as a buffer zone in order to minimize speed-sensing errors. It should be noted that the first data given from the sensor on a car pass is not erroneous since the first activation of the sensors happens on the front bumper. The upstream IR sensor even though it's constantly activated by the vehicle passing it, it is not actually making any trouble since it is waiting for the second sensor to be activated to report time. After the second IR sensor is activated the first time, erroneous data starts coming in from the sensor. Getting the IR sensors far enough to allow the whole car to pass the upstream sensor before activating the downstream IR sensor solved this.

Table 1 shows that the speed sensor is giving accurate data. A 14.5% error can be extrapolated from the results. This percentage in error can be attributed to human error in reading the car's speed-o-meter. Errors between the speed-o-meter readings and the sensor are minute given the intrinsic error found in the speed-o-meter. The speed sensor is functioning well and is giving good readings as it is.

B. Traffic Estimation Algorithm Testing

The second test was made in order to simulate the data coming from the counters and be able to test the MATLAB code, which applies the Classical Analytical Model. In addition, further data was gathered regarding positioning the speed sensor inside the link. Data was also gathered in order to test the use of GPS probes to remedy the relative deviation in the Classical Analytical Method. Moreover, the street chosen

contained a minor mid-link source/sink (mostly acted as a source during the test time) which allowed us to test the difference between the speed sensor and the GPS probe in improving the travel time estimation.

The link chosen is the one linking American Eagle Outfitters to Napolitana Pizzeria in Hamra as seem in Figure 8. The actual road length is 102.3 meters.

Figure 8: Field Test area

Testing Prerequisites

Before testing and in order to facilitate the testing procedure and allow simulating the counter data, an android app was developed/implemented. The app allows getting a time stamp of the current time at a button press. Thus it allows us to get the exact time when a car crosses a certain gate representing the counter. Each time stamp corresponded to a single vehicle. The data the app collects is stored in a text file on the phone, which allows easy data recovery. The app was deployed on two phones to simulate upstream and downstream counters. The interface of the app is shown in the Figure 9.
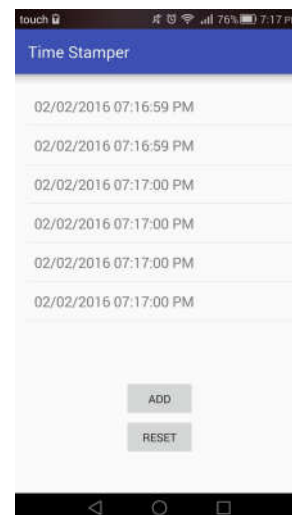
Figure 9: User interface of our android counter application

The UI as seen contains three items. The first is the ListView, which shows the time stamps that got recorded and stored on the phone. The two buttons ADD and RESET allow us to add time stamps to the phone at each press of the button ADD and allows resetting the ListView at the press of the button RESET.

Observed Conditions at Time of Testing

The following conditions seemed prominent on Thursday 21st January at 10:30 am:

- The link had a consistent flow of vehicles.
- Pedestrian crossings were minimal.
- Shared Taxis were stopping on the right but with little effect and frequency.
- Double parking was noticeable.

Methodology of Testing

The procedure stated in this section was followed in order to achieve the outcomes required for this test.

1. The team moved to the link with the app deployed on two android phones.
2. The Tasks were divided
3. One member was to get time stamps of cars upstream the link
4. Another member was to get the time stamps downstream the link
5. The last member was to find the speed of the cars using a speed trap at the middle of the link
6. An extra person was to drive with the GPS probe in the vehicle around the area passing through the link multiple times throughout the test duration
7. The two members upstream and downstream had to wait for the same car to cross the road in order to start. (Same Y-axis)
8. After a certain count the speed trap was moved slightly upstream and speeds were recorded
9. The test terminated after 40 minutes corresponding to about 380 vehicles/counts

Results and Analysis

After the testing ended, the data was extracted from the phones and aggregated. The data was aggregated to further mimic the counters that are available as per the constraints stating that the data come in at 10-minute intervals. The data from the speed trap was used to attain the average spot speed at the two spots. The data from the application on the phones was aggregated and summarized in Table 2.

Table 2: Phone application data summary

Upstream Time (Seconds)	Upstream Counts (Aggregated)	Downstream Time (Seconds)	Downstream Counts (Aggregated)
0	0	24	0
600	101	600	97
1200	197	1200	192
1800	297	1800	295
2400	384	2400	391

The average speeds that were attained from the speed sensors are as follows:

- Speed Sensor 1: 20.42 km/h
- Speed Sensor 2: 23.95 km/h

Assuming a uniform speed throughout the link yields the following travel times coming from the sensors:

- Speed Sensor 1: 18.04 seconds
- Speed Sensor 2: 15.38 seconds

Probe data collected in a raw format is as follows:

Table 3: Raw data coming from the probe

Tu (Seconds)	Td (Seconds)
722.26	745.26
1507.86	1538.15
2185.25	2206.25
2383	2400

The probe data was modified in order to be used in the algorithm in order to remedy the relative deviation caused by the mid-link source. The data is represented in Table 4.

Table 4: Probe data modified to be used

Position	Travel time
120	23
250	30
360	21
391	17

The Data stated above was used in order to achieve travel time estimations. Data from the counters on its own cannot be used to get the travel time. This is due to the fact that the two plots cross as seen in Figure 10.
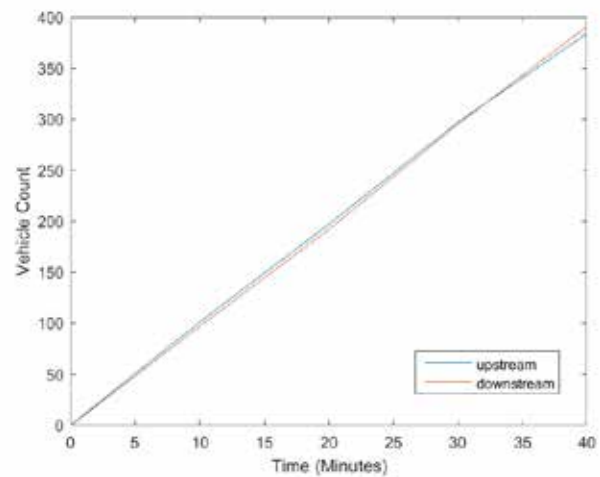
Figure 10: Plot of the unmodified vehicle counts from counter data

The counter data was modified using the speed sensor as discussed in the implementation. The following plot was obtained.
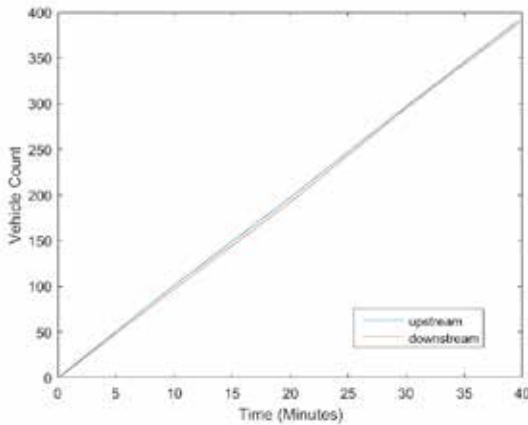
Figure 11: Plot of the vehicle counts with the data using the speed sensor

The data was then modified using the GPS Probe data and the upstream graph was modified as in the implementation.
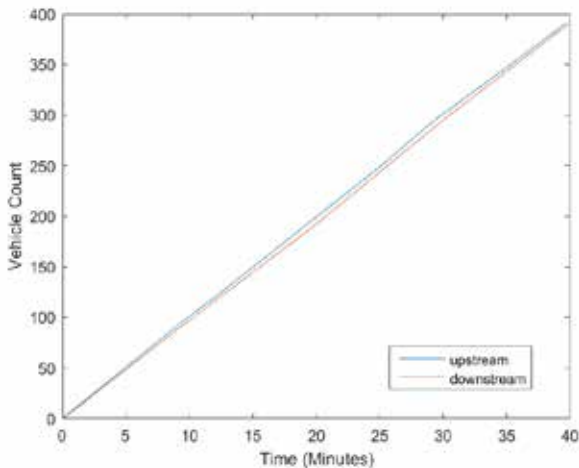
Figure 12: Plot of the vehicle counts with the data modified using the speed sensor

Applying the MATLAB code in order to get the travel times from all the two latter plots. The results are as follows:

- Travel time after correcting the relative deviation in the cumulative plots using Speed trap 1: 19.4 sec
- Travel time after correcting the relative deviation in the cumulative plots using Speed trap 2: 19.02 sec
- Travel time after correcting the relative deviation in the cumulative plots using the GPS Probe data 26.7 seconds

It should be noted that the difference in distance between the two speed traps is only 13 meters, which is not a very significant amount.

It was observed that the number of vehicles exiting the link is higher than the number of vehicles entering the link. It was found that 384 entered the link and 391 departed downstream the link. Thus the mid-link source/sink was acting more like a source than a sink and thus cause the departure of 7 extra vehicles from downstream the link. In addition, during the first 10 minutes of the test it was observed that most vehicles took on average slightly below 20 seconds to go through the link whereas in the next 20 minutes the time to go through the link increased to about 30 seconds and then in the final 10 minutes the travel time decreased but to values closer to 25 seconds rather than 20 seconds. The average traversal time within the link was about 25 seconds given that the later 30 minutes of the test pushed the average beyond the 25 seconds mark.

The speed sensor showed highly varying results when it was moved 13m downstream. The reason lies in the fact that the speed sensor allows the measurement of the spot speeds, which are usually not representative of the average speed on a link. The speeds throughout the link vary due to the intersections and interruptions at the start and end of the link and throughout it. Even though the speed sensor showed accuracy of measurement it was not reliable as a supplementary data source to the classical analytical model. The speed sensor is too sensitive to the fact that it gives a single value every time interval without randomness and thus it cannot be always statistically representative of the situation compared to a more stochastic data set that can be seen in the data coming from the probe.

Even though the travel time from the corrected plots resulted in values that are closer to the control value of about 25 seconds, they still remained unrepresentative of the real situation. The values contained a large error. The major setback was the fact that the speed sensor results varied tremendously and the data didn't represent the overall travel speed in the link.

The GPS probe data in contrast contributed well to the classical analytical model. The data coming from the probe is fixed to a given vehicle and thus it gives assurance in the departure and arrival times within the link. The values are highly representative of the travel time given that the probe vehicle has to pass through the link in order to gather the data required in the link. In addition, the data comes in at different times with different ranks within the D(t) and this adds reliability to the data. The travel time calculated by the classical model given the ramifications of the GPS probe were closer to the true values than were the ones coming from speed sensor corrections. The value contained minor errors but if presented to the traveler this data would be sufficient to give the traveler a good demonstrative idea about the situation within that link.

VI. IMPACT ASSESSMENT

A. Economic Effects:

According to [9] the cost of travel of one passenger inside a conventional passenger vehicle of 1 kilometer in Beirut taking into consideration the costs of congestion, travel time,

pollution and the accident cost is estimated to be around 0.48 $. The travel time cost accounts for the highest portion of the total cost.

Moreover, The Lebanese Ministry of Interior and Municipalities estimates in 2012 a 1.3 million total ground fleet approximately.
The travel time cost is estimated to be 0.052$/veh.km. Assuming the ATIS will allow its users to travel 20% faster, and assuming a linear relationship between travel time and travel time cost, the travel time cost will decrease by 20% to be 0.0468$/veh.km.
Assuming a daily ground fleet presence of 60% in Beirut hence 780 000 vehicles circulating in Beirut every day, the travel time cost improvement will cause a total saving of:
 780000 x 0.0468 = 8112$/km travelled. Figure 16 shows the percentage of travel time reduction versus the mobility cost savings expressed in $/km.

Figure 13: Mobility Cost Savings
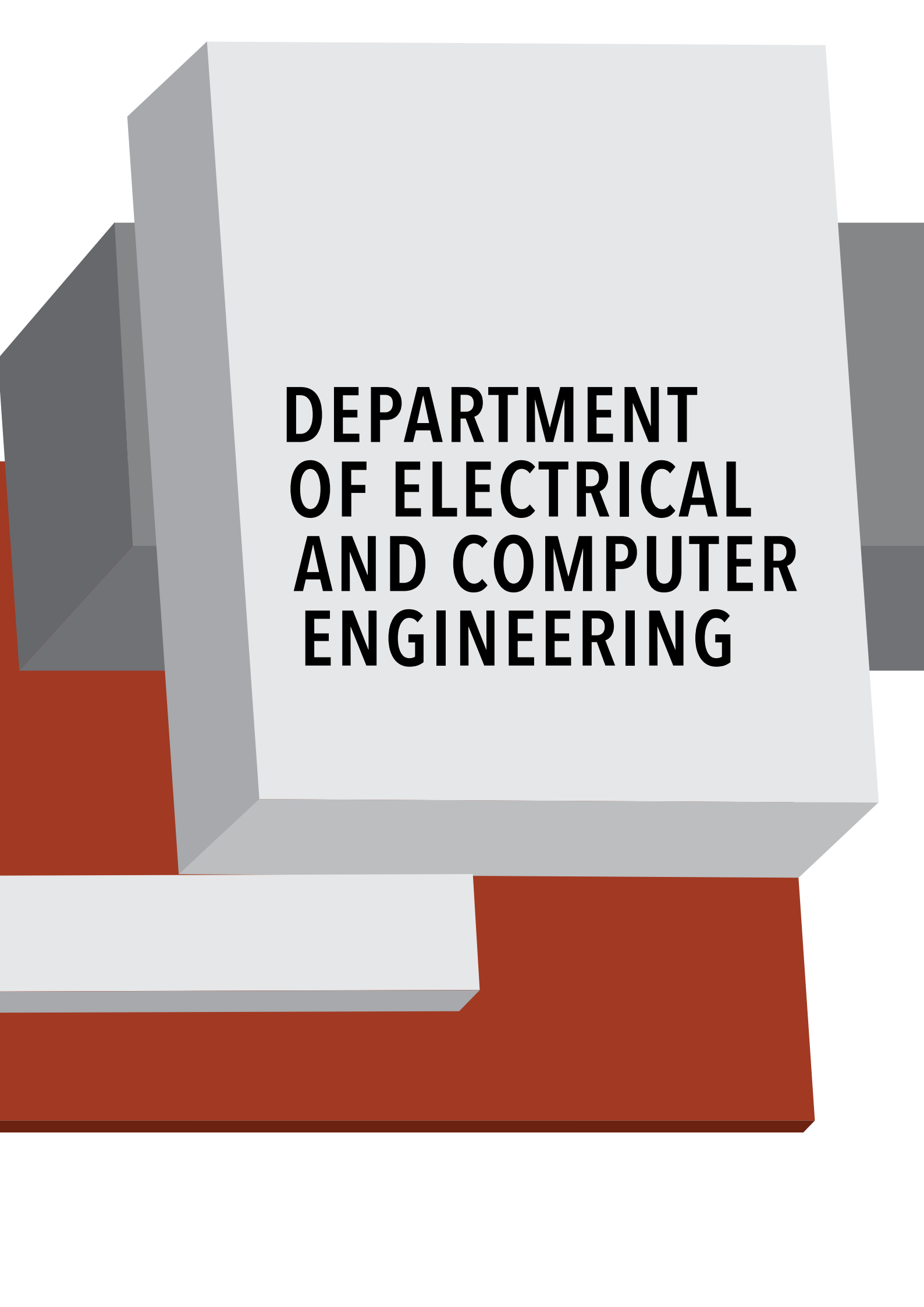
Extrapolating this data by multiplying it by the total number of kilometers travelled/day will result in considerable estimated savings that are due solely to travel time cost improvement ignoring the pollution and congestion costs the ATIS integration will bring to the transportation network. The study is generalized to include reduction in travel time spanning from 5% to 25%.

VII. ACKNOWLEDGEMENTS

VIII. REFERENCES

[1] Ministry of the Environment, "Lebanon's Second National Communication to the UNFCCC,"2011. [Online]. Available at: http://unfccc.int/resource/docs/natc/lebanon_snc.pdf. [Accessed: Feb-2016].

[2] CEDRO, "). Sustainable Transport: Grouped Publication," Jan-2014. [Online].Available at: http://cedro-undp.org/content/uploads/publication/141217083314434~groupedpublication-sustainabletransport.pdf [Accessed: Feb-2016].

[3] M. S. Moghrabi, A framework for assessing intelligent transportation systems deployment: a case study of Lebanon. 1999.

[4] Ashish Bhaskar1, Edward Chung2, André-Gilles Dumont, Arterial travel time estimation: Revisiting the classical procedure, Australasian Transport Research Forum 2011 Proceedings 28 - 30 September 2011, Adelaide, Australia.

[5] R'e-Mi Hage, Eric Le carpentier, Fran¸cois Peyret, Dominique Meizel. Link travel time estimation in urban areas by detectors and probe vehicles fusion. 2nd International Conference on Models and Technologies for Intelligent Transportation Systems, Jun 2011, Leuven, Belgium. 2011.

[6] Ashish BHASKAR, A Methodology (CUPRITE) for Urban Network Travel Time Estimation by Integrating Multisource Data. Thesis NO 4416, 2009 EPFL

[7] Bhaskar, Ashish, Chung, Edward, & Dumont, Andre-Gilles, Estimation of travel time on urban networks with midlink sources and sinks. Transportation Research Record, 2121, 2009, pp. 41-54

[8] V. Sathe Yogesh. "Air Quality Modeling in Street Canyons of Kolhapur City, Maharashtra, India", Universal Journal of Environmental Research and Technology, Volume 2, Issue 2: 97-105. Available Online at: www.environmentaljournal.org. [Accessed: 15- March- 2016]

[8] Mobility Cost, a case study for Lebanon, the Ministry of Environment – United Nations Development Programme, May 2015

DEPARTMENT
OF ELECTRICAL
AND COMPUTER
ENGINEERING

CONTENT

A Dynamic Network that Optimally Distributes Wi-Fi Connectivity

Abdel Wahab Turkmani, Ibrahim Youssef
Department of Electrical and
Computer Engineering
American University of Beirut
Beirut, Lebanon
Email: ast11@mail.aub.edu, iyy01@mail.aub.edu

Jeffry Zalloum
Department of Mechanical
Engineering
American University of Beirut
Beirut, Lebanon
Email: jizz00@mail.aub.edu

Abstract—**Modern Wi-Fi networks suffer from their static nature, which prevents them from delivering quality Internet service at their maximum capacity. This inherent immobile nature prevents these networks from adjusting to load variations. Additionally, ground-based Wi-Fi networks are limited to their places of installation, which prevents redeployment upon demand. To address this, we propose a solution that employs unmanned aerial vehicles equipped with Wi-Fi utilizing a 4G back-link. This offers the ability for the network to be relocated on demand while dynamically adjusting router locations for optimal connectivity. This is achieved through the design of a fleet of blimps fitted with 4G Wi-Fi routers. A microcontroller monitors the position, altitude, and orientation of the UAVs and uses the data to reposition the blimp. Target locations for the UAVs are coordinated via a web-server base station. Bandwidth demand estimates are drawn from individual router traffic, and are used in tandem with an optimization algorithm running on the base station. Together, these will be used to decide on the ideal router distribution that best services the current demand. In order to validate this system, a single blimp is constructed and used to demonstrate that Wi-Fi can indeed be delivered from an airborne router, and that the blimp can be controlled from a base station. The operation of an entire fleet of blimps is tested via simulation to evaluate the performance of the optimization algorithm.**

Keywords—Wi-Fi, 4G, Dynamic Networks, Optimized Networks, Genetic Algorithm, Blimp.

I. INTRODUCTION

The aim of this paper is to study the performance benefits of implementing a dynamic Wi-Fi network instead of a traditional static network of Wi-Fi routers. As opposed to its counterpart, a dynamic network has the advantage of being able to constantly readjust its configuration according to changing demand patterns. The system proposed in this paper is a fleet of airborne blimps fitted with Wi-Fi routers with a 3G/4G back-link. Given estimates in network demand, a genetic algorithm is implemented in order to repeatedly reposition the blimps to maximize user connectivity. Variations of the genetic algorithm have been applied to problems in Wi-Fi network design[1] and in UAV positioning[2]. Our aim is to implement a genetic algorithm for continuous network reconfiguration using blimps as the router carriers. The algorithm is executed on a central base station, and the new blimp positions are sent to each blimp in the fleet as a pre-formatted string via a TCP connection established with the base-station every 10 minutes. The location, altitude, and

orientation of the blimps are determined through a number of position monitoring techniques[3] that process data coming from GPS modules, altimeters, and magnetometers. As shall be demonstrated, the implementation of this dynamic system results in a network that can adapt to variations in demand, giving it an advantage over static Wi-Fi networks. Such a system could be applied to a wide range of practical scenarios, such as traffic monitoring, providing service at remote events, and providing communications support in disaster stricken areas. Previous research has been conducted into the deployment of static broadband communication systems in disaster stricken areas[4], highlighting the need for temporary and quickly deploy-able systems that can provide Internet service in the event of major infrastructure damage. The organization of this paper is structured as follows: Section II will discuss the hardware design including construction of the airship and it's gondola as well as the choice of controller. Section III will discuss the process of estimating network demand. Section IV will explore the optimization of the network distribution. Section V will provide a brief overview of the communication setup utilized and Section VI will provide a discussion of additional modules possible and reflect on the flexibility and modularity of the system. Section VII will conclude the paper.

II. HARDWARE DESIGN

Many restrictions arise from building an airship than carries payloads for a long period of time. Possible options are an: airplane, helicopter, quad-copter, blimp and a quad-blimp combination. A comparison of airship design along with the advantages and disadvantages of each are tabulated and each specification is ranked in Table II.

Property	Airplane	Helicopter	Quad-copter	Blimp	Quad-blimp
Power Cost	2	1	1	3	2
Control Cost	2	1	1	2	2
Payload/Volume	3	2	2	1	2
Maneuverability	2	3	3	2	2
DOF	1	3	3	1	2
Stationary flight	1	3	3	3	3
Low speed	1	3	3	3	3
Endurance	2	1	1	3	2
Miniaturization	2	3	3	1	2
Indoor Usage	1	3	3	2	2
Total	17	23	23	21	22

TABLE I.   COMPARISON OF AIRSHIP DESIGNS

The quad-copter and helicopter designs are can lift much

more weight than the blimp design, which is important given the weight constraint of our project, and the weight requirements of the hardware. However, the high energy consumption is a major problem with this design. The selected UAV design should maintain a considerable flight time to be practically implementable. Moreover, having a quad-blimp combination needs a lot more equipment on board the blimp, reducing the allowed payload capacity for the communication and control hardware. As such, a regular blimp design is preferable.

A. Gondola Design

The gondola attached to the balloon will hold all the equipment needed to operate the blimp. The gondola will contain the batteries, the controller, the 4G to WI-FI adapter and the electronic speed controllers (ESCs). Moreover, we decided on a tri-motor design to manoeuvre the blimp: two on each side of the gondola to navigate in the x-y plane, and a third attached to the bottom of the gondola to control the altitude of the blimp. The final design of the blimp is displayed in Fig. 1

Fig. 1.   Complete gondola and propellers.

B. Choice of Controller

The requirements on the controller of choice are:

1)   Ability to connect to the Internet.
2)   High-speed control loop.
3)   Large memory to buffer sensor data.
4)   Support for parallel execution.

Internet connectivity rules out controllers such as the Arduino and PIC. In addition, the high-speed and parallel execution requirements render the Raspberry Pi as infeasible. As such, a suitable controller that abides by the above requirements while maintaining a low footprint and weight is the National Instruments myRio controller. With it's built-in Wi-Fi connectivity, high speed control loop and true parallel execution offered by the FPGA in addition to it's capability to buffer large amounts of data render it the controller of choice.

C. Choice of Blimp Shell Size

The size of balloon needed is dependent on the payload required. A list of components and their respective weights are tabulated in Table II-C.

Component	Weight (g)
Gondola Box	550
Horizontal Bar	400
myRio Controller	400
WI-FI adapter	200
Sensors	100
Battery 2200 mah (x2)	400
Battery 11800 mah	150
ESC (x3)	150
Brush-less DC motor (x3)	150
Propeller (x3)	150
TOTAL	2700

TABLE II.   WEIGHTS OF COMPONENTS

The total required payload of the blimp is approximately 2.5 kg. Therefore, to determine the amount of helium needed to deliver this payload, a survey of balloon sizes available in the market was merited. Local suppliers deliver balloons with either a 1m or 2m diameter balloon. A calculation of the payload capabilities of these balloons follows: For a 1m balloon, the volume is $V = \frac{4}{3}\pi r^3 = 1.05m^3$. The net lifting capability calculated for this type of balloon is around 1kg/m3. This leads to a net lifting mass of 1kg, which is not enough for the gondola. For a 2m balloon, the volume is $V = \frac{4}{3}\pi r^3 = 4.19m^3$. This leads to a net lifting mass of 4kg and allowing enough of a safety margin to lift the gondola in different temperatures and altitudes.

III.   DEMAND ESTIMATION

One great difficulty in operating this system is estimating user demand [5] in order to optimize the locations of the blimps. An exact measurement of user demand is difficult to do without a dedicated router running custom firmware such as DD-WRT [8], and as such we have chosen to rely on stochastic methods to approximate the demand incurred by users in the field of operation. We assume that the amount of bandwidth demanded by each user is a random variable D, with a mean and variance determined empirically at start-up. To compute the mean, elementary bandwidth statistics procured from the Mobi router are used. Total bandwidth consumption for each router (blimp) divided by the number of IP addresses connected to the network is used to estimate the mean. Computing the standard deviation cannot be done as cleanly; we shall employ a running estimator that compares the mean to the amount of bandwidth demanded by each additional user, or the bandwidth freed up by each disconnecting user.

$$s = \sigma = \sqrt{[\mu - (\mu * N - Measured\ Demand)]^2}. \quad (1)$$

Then the bandwidth demand incurred by N users can be approximated as:

$$Approximate\ Demand = \sum_{i=1}^{N} D_i \quad (2)$$

IV. GENETIC ALGORITHM

A genetic algorithm is a nature-inspired search algorithm that is often used in optimization problems, particularly in cases where the function to be optimized is highly non-linear or heuristic in nature. The algorithm borrows many concepts from the process of natural selection such as the phenomena of inheritance, mutation, selection, and recombination [6][7]. This algorithm operates by taking an initial population of solutions and progressing through several generations. At every stage in the process, every individual in a population is analyzed for its survival fitness using a fitness function specific to the current problem. Individuals are then selected based on this fitness concept and modified through recombination and mutation to produce a second generation. This process occurs until either the desired fitness level has been reached or a maximum number of iterations has been exceeded. This algorithm is highly appropriate for our application, due to the stochastic nature of our demand estimation. The nature of our demand estimation means that the fitness function is highly nonlinear, therefore making the Genetic Algorithm a good match for our problem.

A. Genetic Algorithm Formulation

In order to implement a genetic algorithm, we are required to have two things: a way to genetically represent the solution set, and a fitness function. The solution set shall be generically represented as a vector of addresses indicating the positions of the fleet of blimps. The fitness function provides a measure of how fit a particular solution is, and it is this quantity that we wish to maximize. Let us begin by dividing the overall field of operation into sectors. Each sector is assigned a unique binary address, A_i:

$$A_i = a_1 a_2 a_3 ... a_k, a_i \in [0,1] \tag{3}$$

The sector containing a blimp is assigned a value of 1, and all other sectors are assigned values of 0. Within each sector there is some number of users seeking to access the network. The demands throughout the sectors is represented by a demand vector, :

$$\theta = \left[\sum_{i=1}^{N_1} D_i \sum_{i=1}^{N_2} D_i \sum_{i=1}^{N_3} D_i ... \sum_{i=1}^{N_k} D_i \right]^T \tag{4}$$

Where k is the number of sectors in our field, and the Ni's represent the number of users in each sector. Essentially, θ is a matrix that we estimate and that we have no control over. For any given optimization, θ therefore behaves as a constant. A solution is a vector of addresses containing the proposed location for every blimp in the fleet:

$$\alpha = \begin{bmatrix} A_1 A_2 A_3 ... A_x \end{bmatrix}^T \tag{5}$$

Where x represents the number of blimps in the fleet. Each blimp in the fleet is capable of providing a fixed amount of bandwidth, 8 Mbps. In order to ensure that we are properly servicing as many customers as possible, we introduce the constraint to the optimization that the entire field must be covered by the network, even if there is no more capacity for

a blimp to provide service. This constraint allows us to avoid potential blind spots where hopeful users are not serviced due to the fact that they are simply invisible to all blimps in the fleet.To ensure that this constraint is met, we shall consider that any sector that is partially covered by a blimp is an uncovered sector, and must therefore be covered by another blimp.

We can now use the above definitions to formulate our fitness function. The fitness of a solution is the total amount of demand serviced by the blimps in our fleet. At the same time, a blimp configuration that fails to cover the entire field of operation should receive a fitness value of 0. The result is the following nonlinear, piecewise fitness function that applies for a system containing x blimps and k sectors:

$$F = \begin{cases} \alpha * \theta & all\ sectors\ covered \\ 0 & otherwise. \end{cases} \tag{6}$$

Which can be expanded to:

$$F = \begin{cases} A_1\theta + A_2\theta + A_3\theta.. + A_x\theta & all\ sectors\ covered \\ 0 & otherwise. \end{cases} \tag{7}$$

B. Motion Penalty

An improvement to the proposed system is the introduction of a motion penalty factor P to the fitness function. This addition would ensure that minimum energy is expended in repositioning the blimps during a configuration update. In addition, the a penalty gain is introduced which allows for the trade-off between maximum service and minimum motion to be customized. In applications where perhaps some connectivity could be sacrificed in favor of energy conservation by keeping blimps stationary, this gain can be increased. On the other hand, the gain can be reduced in order to increase the mobility of the system.

$$P = \beta * Total\ blimp\ displacement \tag{8}$$

The overall fitness function can now be re-written.

$$F = \begin{cases} A_1\theta + A_2\theta + A_3\theta.. + A_x\theta - P & all\ sectors\ covered \\ 0 & otherwise. \end{cases} \tag{9}$$

C. Simulation

In order to test the advantages of this system over a static system, simulations are conducted on a small scale for both cases. An arbitrary square field of side length L is divided into 9 sectors of side length $\frac{L}{3}$. Three blimps shall be deployed over this region, and three static Wi-Fi routers, each with the same bandwidth capabilities of 8 Mbps, and distribution radius of $\frac{L}{2}$. A blimp is positioned at the center of each sector, and the static Wi-Fi routers are placed in the following configuration, which meets the constraint of covering the entire field of operation.

$$\begin{bmatrix} 3 & 4 & 2 \\ 4 & 0 & 1 \\ 0 & 0 & 17 \end{bmatrix} \quad \begin{bmatrix} 3 & 2 & 2 \\ 4 & 0 & 1 \\ 0 & 0 & 17 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 10 & 2 \\ 4 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 14 & 0 & 0 \end{bmatrix}$$

Fig. 3.   Test Cases
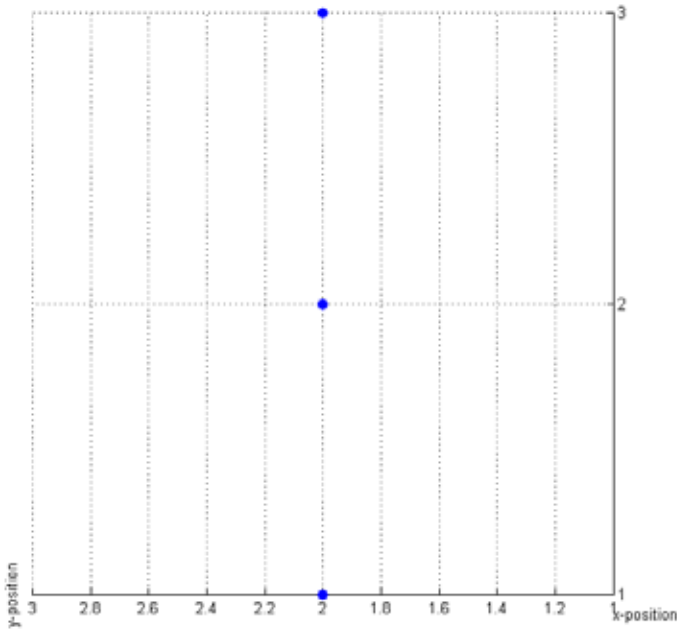
Fig. 2.   STATIC CASE

On the other hand, the three blimps are controlled by the dictates of a Genetic Algorithm and are able to relocate as needed to maximize user connectivity.

In order to test the performance benefits of the dynamic system over a static system, the following test cases are used, where the elements represent the amount of bandwidth being demanded by each sector.

For each case, the total amount of bandwidth serviced is compared between the static and dynamic scenarios in order to highlight the advantages of the dynamic network.

D. Simulation Results

1) Case 1: For the dynamic case, a fitness of magnitude 24 Mbps is measured, and is achieved by the following distribution of blimps. For all of the following figures, the X and Y axes represent the position of the blimps in space, while the Z axes represents the amount of bandwidth being demanded in each sector.
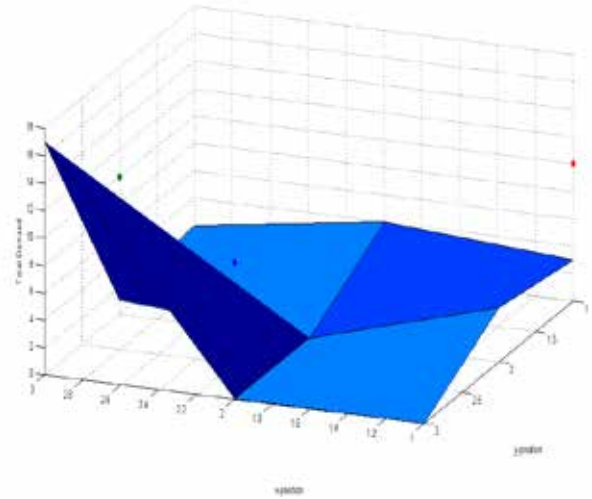
Fig. 4.   CASE 1

In this instance, the static network is able to match the performance of the dynamic network, also being able to provide 24 Mbps.

2) Case 2: For the dynamic case, a fitness of magnitude 24 Mbps is measured, and is achieved by the following distribution of blimps. Notice that in this case, no reconfiguration was needed. For this demand pattern, this is the most that three blimps can service while maintaining coverage of the entire field.
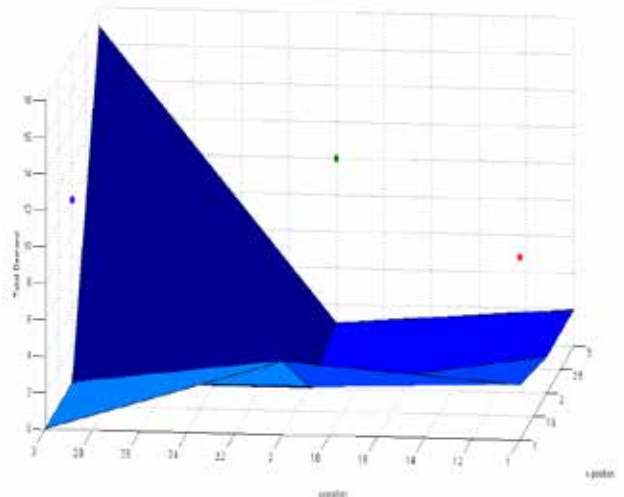
Fig. 5.   CASE 2

On the other hand, the static network performs more poorly than the dynamic one in this case, providing 20 Mbps of service.

3) Case 3: For the dynamic case, a fitness of magnitude 20 Mbps is measured, and is achieved by the following distribution of blimps. Notice that in order to achieve this level of fitness, the network had to re-adjust its configuration.
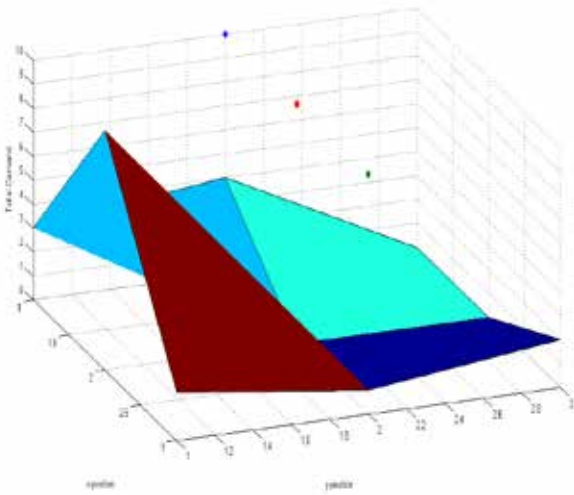
Fig. 6. CASE 3

Comparing the two types of networks,we notice that the static network can supply at most 18 Mbps.

4) Case 4: For the dynamic case, a fitness of magnitude 14 Mbps is measured, and is achieved by the following distribution of blimps. Once again, we are able to achieve this fitness through a re-configuration of the blimps.
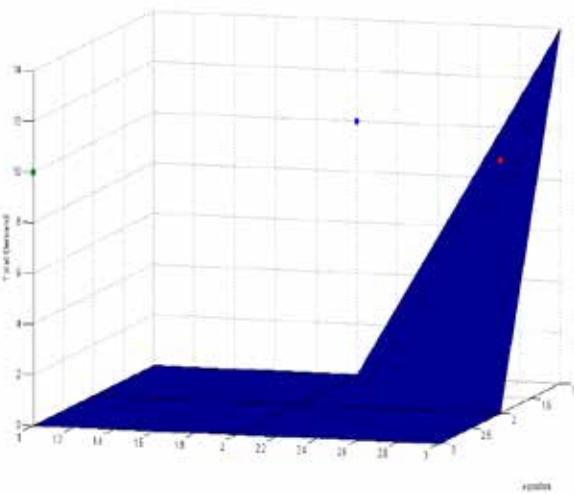
Fig. 7. CASE 4

Finally, we compare this with the static case that is only able to achieve a fitness of 8 Mbps, since only one blimps is able to service the 14 Mbps of demand.

The results are summarized in the following table. We observe that the performance of the dynamic network, in Table IV-D4, is consistently better than or equal to that of the static case.

Configuration	Static Fitness	Dynamic Fitness
1	24	24
2	20	24
3	18	20
4	8	14

TABLE III. COMPARISON OF STATIC VERSUS DYNAMIC ALLOCATION FITNESS (MBPS)

V. COMMUNICATION

With the optimization algorithm running on a base-station, a communication and command transfer protocol is needed between the controllers on the blimps and the base station. Data from the sensors, and demand estimates per blimp are buffered on the myRio. Every 10 minutes, a TCP connection is opened between the myRio and the base-station with the myRio transmitting data in the outgoing message, and receiving update instructions in the reply message, before closing the connection. Commands from the controller to the base-station are formatted as displayed in Table V.

Commands from the controller are formatted with the command name first, a colon, followed by the command-specific data as such:

$$command : data1, data2, data3$$

Command	Data1	Data2	Data3
Heartbeat	x-position	y-position	altitude
New User	longitude	latitude	total demand
Deleted User	longitude	latitude	total demand

TABLE IV. COMMANDS FROM CONTROLLER TO BASE STATION

Commands from the base station are formatted with the command name first, a colon, followed by the command-specific data as such:

$$command : data1, data2$$

A list of commands are in Table V.

Command	Data1	Data2
Location update	x-position	y-position
Recalibrate	none	none
Return to base	x-position	y-position
Stop	none	none
Deploy	x-position	y-position

TABLE V. COMMANDS FROM BASE-STATION TO CONTROLLER

VI. ADDITIONAL MODULES

Additional modules implementable in such a system are documented below.

A. Weather Data Collection

Such a system offers the ability for expansion to allow for reading weather data. Basic weather measurement sensors we intend to include are temperature sensors, barometers for atmospheric pressure, hygrometers for moisture content of the atmosphere, and photocell sensors for measuring amount of sunlight.
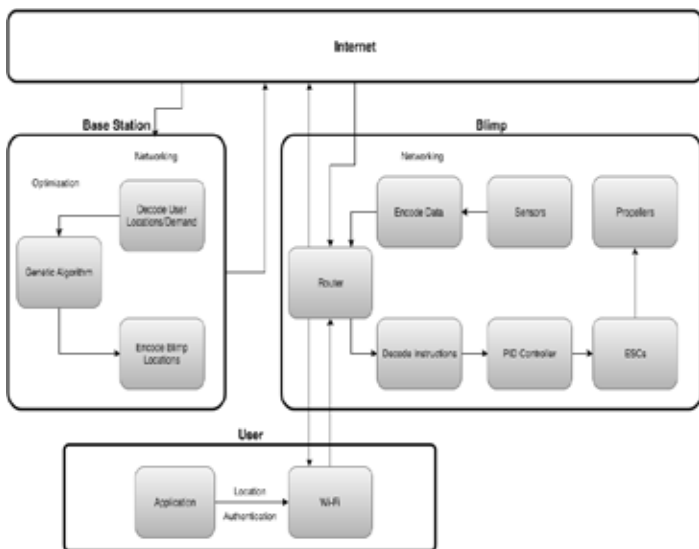
Fig. 8. Overall system

B. Vision

Additionally, a camera will be installed on a balloon, that will transmit basic information that can be used for crowd-control and monitoring. Examples include population density estimates, car density estimates, traffic estimation, etc.

VII. CONCLUSION

In this paper we displayed a possible solution to implement a dynamically adjusting Wi-Fi network. Employing blimps as an airship, this system can be deployed on demand, and can be used to offer support to overloaded Wi-Fi networks, supply connectivity in remote or disaster stricken areas, and serve as temporary respite for unconnected zones. A major advantage of this model is the optimization on the router positions to deliver maximal bandwidth connectivity to each user, while ensuring all users remain connected. A flow chart of the overall system, indicating how its constituent parts interact with one another is shown in Fig. 8

ACKNOWLEDGMENT

REFERENCES

[1] E. Agustin-Blas, S. Salcedo-Sanz, P. Vidales, G. Urueta, A. Portilla-Figueras and M. Solarski, "A Hybrid Grouping Genetic Algorithm for citywide ubiquitous WiFi access deployment", 2009 IEEE Congress on Evolutionary Computation, 2009.

[2] O. Sahingoz, "Flyable path planning for a multi-UAV system with Genetic Algorithms and Bezier curves", 2013 International Conference on Unmanned Aircraft Systems (ICUAS), 2013.

[3] J. Borenstein, H. Everett, L. Feng, S. Lee and R. Byrne, "Where am I? Sensors and Methods for Mobile Robot Positioning". The University of Michigan, 1996.

[4] M. Scott, B. Charles, "Rapidly-Deployable Broadband Wireless Networks for Disaster and Emergency Response", Presented at The First IEEE Workshop on Disaster Recovery Networks (DIREN 02), June 24, 2002, New York City, NY.

[5] M. Zhong, P. Hu and J. Indulska, "Revisited: Bandwidth estimation methods for mobile networks", Proceeding of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks 2014, 2014.

[6] Jonathan E. Rowe. 2007. "Genetic algorithm theory." In Proceedings of the 9th annual conference companion on Genetic and evolutionary computation (GECCO '07). ACM, New York, NY, USA, 3585-3608.

[7] D. T. Larose, "Genetic Algorithms," in Data Mining Methods and Models , 1, Wiley-IEEE Press, 2006, pp.240-264

[8] X. Ye, B. Zeng, H. Pu and L. Lv, "Remote monitoring wireless router system based on SaaS model", 2010 2nd IEEE International Conference on Information Management and Engineering, 2010.

A Reconfigurable Antenna Design with an 'AUB150' Shape

Fatima Alzahraa Bassam Asadallah ,
Electrical and Computer Engineering Department
American University of Beirut
Beirut, Lebanon
fba16@mail.aub.edu

Abstract— **In this paper, a new reconfigurable antenna design is presented. The reconfigurable antenna has a patch with the shape of the "AUB 150" logo. The antenna operates at GPS, 3G, WiFi, GSM and 4G frequencies. The multiband operation is accomplished by the multi-resonant property of the 'AUB150' slots that were etched of the patch and the ground plane of the antenna structure. The antenna reconfiguration is achieved using two PIN diodes switches. The antenna is proposed for integration into various wireless communication platforms.**

I. INTRODUCTION

New wireless communication requirements impose various constraints on the design of modern antenna systems. In the world of handheld devices, there is a need for a communication systems that can satisfy all constraints and fullfill our multimedia daily usage. Reconfigurable antennas appear to present a suitable solution that can achieve an efficient solution that provides improved gain communication over various frequencies. Reconfigurable multi-band antennas are based on multi-band topologies while adding reconfiguration capability to their operation. In order to cover multiple bands of frequencies for wireless applications, many related printed flat-plate antenna designs have reached multi-band operation using various techniques. For example, a compact multiband and a reconfigurable Printed Inverted F Antennas (PIFA) are designed to operate at all needed communication frequencies [1]. In addition, meandered PIFAs and folded loop PIFAs are implemented in [2] and [3] respectively to achieve multiband operation, whereas in [4] a capacitive slot is used to achieve a more compact antenna topology. Reconfigurable antennas are incorporated   for additional resonances and added functionality [5-6]. Reconfiguration in these antennas can be realized using PIN diodes, Microelectromechanical Systems (MEMS), or GaAs field transistors.

In this paper, a new reconfigurable 'AUB150' patch antenna design is proposed. The antenna caters for all the frequency operational needs that are required by wireless communication protocols. In section II, the full design is presented then section III discusses the design results. Concluding remarks are drawn in the last section.

II. ANTENNA DESIGN

The antenna design and prototype are presented in Fig.1 and Fig.2 respectively. It is composed of three layers; the top layer is circular with a radius of 29 mm constituting the patch from which slots in the shape of the "AUB 150" logo are etched. The middle layer is composed of a 7x7 cm^2 Roger Duroid (RO 5880) substrate with a dielectric constant of 2.2 and a thickness of 1.6 mm. Finally, the bottom layer is composed of a full ground plane with "AUB" logo shaped slots etched from it. The proposed antenna is designed to operate at GPS (1.54 GHz), GSM (1.8 GHz) and 3G (2.1 GHz) frequencies. It is fed through a power divider splitting the input power into two directions to feed the antenna at two different strategic points. Two SMP1320-079LF PIN diodes (S1 and S2) are integrated within the feeding network in order to reconfigure the antenna operation. The various states of the PIN diode switches control the power flow of the feeding network. Their biasing networks are also integrated within the antenna structure. Two RF chokes and a shorting sheet are placed on the top layer of the design to connect the PIN diodes to the DC power supply for biasing purposes. The reconfiguration of the feeding network allows the antenna to switch on demand from its original operation into an operation at GSM (0.9 GHz), WiFi (2.4 GHz), UMTS (2.66 GHz) and 3G (3.55 GHz).

(a)

S1 S2

(b)
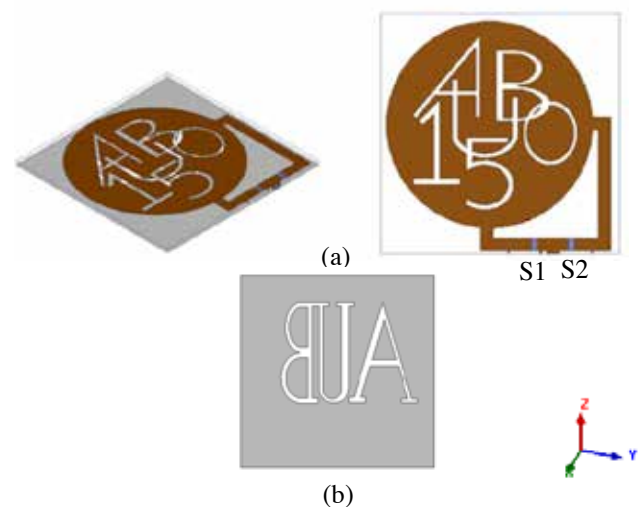
Fig.2 The antenna prototype

III. DESIGN RESULTS

The antenna is simulated using Ansys' HFSS [7] and then measured using Agilent Network Analyzer. The reflection coefficient results for the different switch configurations are shown in Fig.3. Varying the switch configurations leads to different operating frequencies. This is due to the variation of the power intensity at each feeding location. The two PIN diodes allow four different states of operation. However in this design, only three configurations are sufficient for reaching the required performance. It can be noticed from Fig.3 that the ON/ON state of the switches allow the antenna to operate at 0.9 GHz, 2.4 GHz, 2.66 GHz and 3.55 GHz whereas the OFF/OFF state shows resonance at 2.1 GHz . When the switch S1 is OFF and the switch S2 is ON, operation takes place at 1.54 GHz and 1.8 GHz. A wider bandwidth operation is observed within the targeted frequencies, this is due to the incorporation of a slotted ground plane. In addition, the cross feeding technique used allows circular polarization that is realized at 2.4 GHz and 3.55 GHz frequencies. The slots that form the "AUB 150" logo are variations of U and L shaped slots which have been extensively used in literature for multiband antenna designs [8]. In this design, A, B and 5 can be considered as a modified U shape slots thus the same performance can be expected. The other slot topologies add to the multiband operation of the antenna due to the variation and multitude of current paths. Fig. 4 shows the current distribution over the patch and the ground plane at 1.54 GHz. It can be observed that this frequency resonates due to the A and B slots shapes. The gain of this antenna is maintained at acceptable levels for various frequencies. The radiation pattern of the frequencies and the maximum gain are shown in Fig.5.
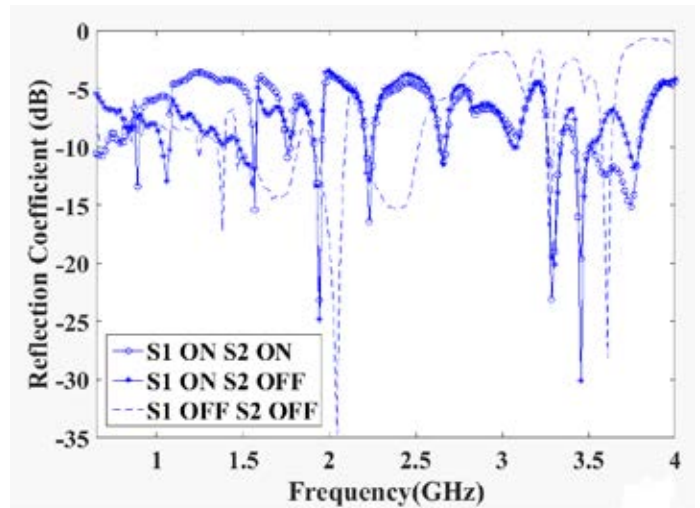
Fig. 3. Reflection coefficient of the measured reconfigurable 'AUB150' design with different switches configuration
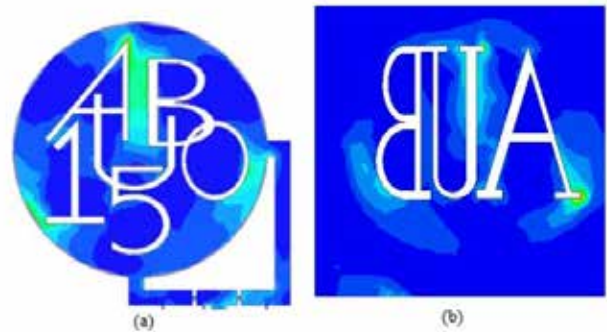
Fig. 4. The current distribution at 1.54 GHz frequency (a) for the patch and the (b) ground plane

Realized Gain Plot

– xz plane /– xy plane

(a)   (b)

Fig. 5. The realized gain plot of
(a) 0.9 GHz with 0.7 dB, (b) 1.54 GHz with 7.5 dB, (c) 1.8 GHz with 7 dB
(d) 2.1 GHz with 7 dB, (e) 2.4 GHz with 9.5 dB, (f) 2.66 GHz with 8.5
and (g) 3.55 GHz with 6.5 dB as a maximum gain.

IV. CONCLUSION

An 'AUB150' shape reconfigurable antenna is proposed in this paper. The antenna covers frequency bands needed for wireless communication in various applications. The antenna design presents a great advantage in terms of achieving a wide range of operation while maintaining a good gain.

V. ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my advisor Prof. Joseph Costantine. I really appreciate his continuous support and non-hesitating help. After that, I would like to thank my colleagues Aline Eid and Mahmoud Abdallah for their fruitful cooperation.

REFERENCES

[1]   F. A. Asadallah, J. Costantine, Y. Tawk, F. Ayoub and C. Christodoulou, "A Multiband and Reconfigurable PIFA for Mobile Devices," IEEE International Symposium on Antenna and Propagation, June 2016.

[2]   M. K. Kärkkäinen, "Meandered Multiband PIFA with Coplanar Parasitic Patches," IEEE Microwave and Wireless Components Letters, vol. 15, pp. 630–632, October 2005.

[3]   M. Zheng, H. Wang and Y. Hao, "Internal Hexa-Band Folded Monopole/Dipole/LoopvAntenna Wwth Four Resonances for Mobile Device," IEEE Transactions on Antennas and Propagation, vol. 60, pp. 2880-2885, June 2012.

[4]   C. Rowell and E. Y. Lam, "Multiple Frequency Band and High Isolation Mobile Device Antennas Using a Capacitive Slot," IEEE Transactions on Antennas and Propagation, vol. 60, pp. 3576-3582, August 2012.

[5]   K. R. Boyle and P. G. Steeneken, "A Five-Band Reconfigurable PIFA for Mobile Phones," IEEE Transactions on Antennas and Propagation, vol. 55, pp. 3300-3309, November 2007.

[6]   A. C. K. Mak, C. R. Rowell, R. D. Murch, and C. Mak, "Reconfigurable Multiband Antenna Designs for Wireless Communication Devices," IEEE Transactions on Antennas and Propagation, vol. 55, pp. 1919-1928, July 2007.

[7]   Pietzak, Amy. "ANSYS - Simulation Driven Product Development". Ansys.com. N.p., 2001. Web.

[8]   P. Salonen, M. Keskilammi and M. Kivikoski, "Single-Feed Dual-Band Planar Inverted- Antenna with

U-Shaped Slot," IEEE Transactions on Antennas and Propagation, Vol. 48, No. 8, August 2000.

On-site Assessment of Photo Voltaic Panels

Jad K. Zeineddine, Omar R. Al-Halabi, and Mostafa A. Atweh
Department of Electric and Computer Engineering
American University of Beirut
Bliss Street, Hamra, Beirut, Lebanon
Jkz02@aub.edu.lb – ora02@aub.edu.lb – maa221@aub.edu.lb

Abstract- **The assessment of photovoltaic panels' performance is a pressing concern for the continuation of work and improvement of any panel; analyzing its Current-Voltage characteristics indicates the status of the panel. Repetitive and detailed measurements and analysis, which will evaluate the parameters of the panel model, are done in different weather conditions and on a number of panels, always taking the lab test results of a new panel as a reference to do the comparison. The prototype will be a device that can be connected to any panel regardless of the open-circuit voltage; it also gives an indication of the action that must be done based on the assessment of the panel. The study is supported by LABVIEW and MATLAB coding which will allow accurate results to be calculated and approximations of the parameters of the model used.**

I. INTRODUCTION

Upon the decaying of non-renewable energy material and the peaking of renewable technologies since the beginning of the 21st century, assessing and supporting the current technologies are key concerns for Power Engineers. Photovoltaic Panels are considered the primary renewable energy source; however, every panel depreciates 0.5% of its efficiency per year. The best efficiency reached so far is around 40%. The main problem power engineers face is the inability of assessing the performance of a photo voltaic panel after being exposed to different weather conditions, which leads to a decrease in its efficiency.

W. De Soto, S.A Klein, and W.A Beckman [3] talked about the factors that affect the output power of the PV panel, which are the incident solar radiation, the cell's temperature, the solar incident angle and the load resistance. The main problem that they are discussing is the fact that manufacturers provide electrical parameters of the panel at only one operating temperature.

The approach included the analysis of the data provided by the manufacturer, the absorbed solar radiation, the temperature of the cell and some empirical formulas to predict the remaining coefficients. This approach helped in predicting the I-V curve of the cell and comparing it to the experimental one. The advantage of the proposed approach is that it gives all the specifications of the corresponding panel without depending only on the manufacturer. Additionally, this model can predict the I-V curve that could be very similar to the experimental one.

The cons of this approach remain in the fact that it does not give a solution that can be used every single time of the day, and does not take into consideration some small alterations in the atmosphere, which can change the experimental curve.

In our approach we will be using a similar logic using MATLAB to predict the I-V curve, but the difference will be that the model we are working on has the ability to create the curve instantaneously at any operating condition and at any time, and is capable of detecting the changes on the PV panel such as dust.

B. P. Dougherty, A. H. Fanney and M. W. Davis [4] talked about building integrated PV cells and the problem that they were talking is the lack of validated predictive simulation tools that can measure the performance of the cell and create its characterization curve. The characterization of the different panels was carried perpendicular to the sun rays and at normal temperatures and specific other temperatures. Using this approach they were able to plot the I-V curve using a prediction method, and curve fitting using the characteristic equation.

The pros of this approach include giving a very accurate view about the curve of the panel at a certain temperature and radians. The cons are that it uses a prediction method that is conducted at specific time and other normal conditions which is not always the case. The difference in our approach is that we will be plotting the characterization curve at any time without setting any of the variables.

The problem presented by V. Nanjannavsr, P. Gandhi and N. Patel [7] was how to get an accurate characterization curve during any period of time under any condition of temperature and solar radians. The approach included varying the temperature and solar radians on the panel and observe how the characterization curve will be affected.

The solution to that was to create LABVIEW software that is interfaced with a data acquisition board and get the results in 2 stages: the first was to set the temperature and change the solar radians energy, and the second is to set the solar radians and change the temperature. These results show the variation in the PV characteristics curve and the variation in the power output of the panel. The results that they got are shown in the graphs below in Fig. 1 and Fig. 2.
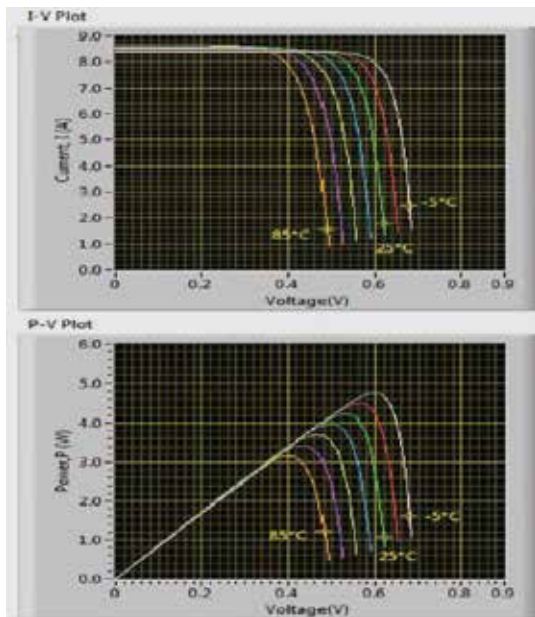
Figure 1: Variation of I-V curve with respect to solar radians
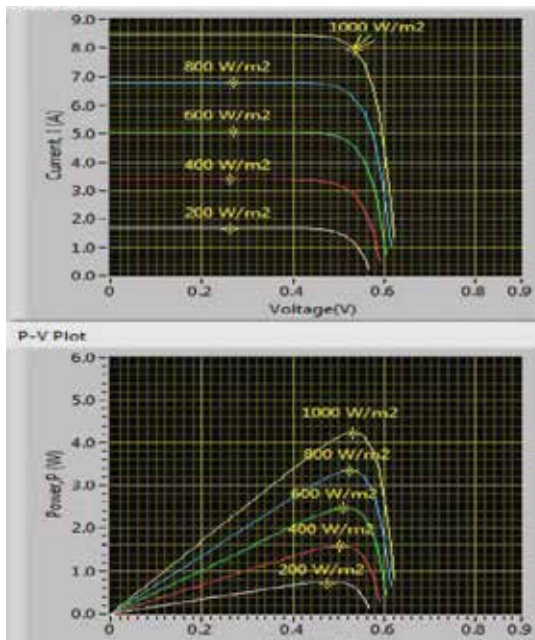
Figure 2: Variation of I-V cure with respect to temperature

This type of modulation is very effective and accurate because it uses a new software that is capable of detecting every single change in voltage. This can ensure the proper reading of the panel's curve at any time no matter of the corresponding radians or temperature. This model is very robust but it does not take into consideration the alteration in weather.

Our approach is somehow similar to this approach but the difference is that we are using such an approach to get the exact experimental result of the panel's curve. The predicted curve is to be made using the inputs we got from the panel like ambient temperature and irradiance and plot it using MATLAB and the characteristic equation of the panel. The difference between both curves is our area of interest in such a way that if this difference is significant, then there is something wrong in the panel. This error could be due to the dust accumulation which we should clean, by setting a cleaning schedule.

IJAREEIE (International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering) highlighted that a capacitor in any model, which would represent the IV characteristics, has the role of charging and discharging. As the capacitor charges and discharges it draws the IV characteristics. The size of the capacitor plays a role in determining the time constant and hence in determining how smooth is the IV characteristics. Therefore, the capacitor provides a solution for the problem of smoothening the characteristic. The following characteristics is shown in Fig. 3.
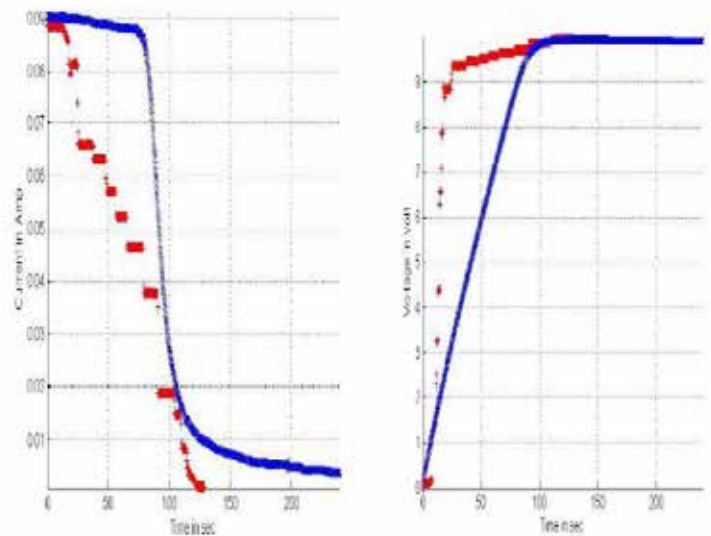
Figure 3: Current vs. Time and Voltage vs. Time using super capacitor (Blue Graph)

The use of an extremely huge capacitor in the range of 5 farads, also known as super capacitor, has the following Pros and Cons.

Pros:
- Precise, linear approximation of PV panel characterization
- It could be used to control and predict the tracing time by choosing the value of super capacitor
- Cost effective
- Less circuit complexity

Cons:
- As the capacitor size increases to reach the huge size of a super conductor, the tracing time will be very large. In our case, using super capacitor would result in weak results because the temperature and the amount of radians may change between readings.

In our approach we had to choose a value of capacitor to compensate the tracing time and smoothness of the curve. We chose a capacitor of 220uF, which presents compensation between the tracing time (Time to read values) and the smoothness of the curve.

E.D Aranda, J.A. Galan, M.S. De Cardona, and J.M Rquez [6] were able to point out that the main problem is the ripples that result when the IV characteristic is graphed. We faced a similar problem in our project. The problem of ripples is a main problem engineers face when with the sampling process is done. Using the single-ended primary inductance converter (SEPIC) is optimal for this application, and the following Pros and Cons summarize its use.

Pros:
- Direct graphical display of the IV characteristics
- It traces the I-V curve in the two directions (From I_{SC} to V_{OC} and vice versa)
- Partial reproduction of the IV graph around three main points (MMP, I_{SC} and V_{OC})
- The ripples are minimized

Cons:
- The use of parallel SEPIC will result in much higher circuit complexity and increased number of components used

Our approach to minimize the ripples found in our IV graph is different. We had a grounding problem that was calculated by 1/ (distance between two ripples). It happened that the value calculated is approximately 50 Hz. This proved the theoretical analysis. We had one ground for the full system, and the ground is connected it to the common channel of the DAQ (Data Acquisition Card).

D. L. King, W. E. Boyson, and J. A. Kratochvil [8] highlight the problem that includes poor performance within low irradiance (low amount of radians). It is crucial to make all measurements in all the conditions. When measurements are taken, poor performance due to the low amount of radians is depicted. This problem is important to our design since we need to solve the dust accumulation problem that will decrease the amount of radiance received.

Some solved this problem by making an empirical relation between AM functions and AOI-function used to compensate the influence of the solar spectrum and solar angle of incidence (AOI), and then create these relations by mathematical modulations to reach PV array and invertors. Others solved it by using maximum power tracking.

For the maximum power tracking approach, the advantage is that this method gives us the needed point to always have the maximum power point that the PV panel would supply. Its disadvantage is that it would not solve the problem of the poor performance in the condition of the low irradiance.

For the other approach, the important advantage is that it could solve the problem of the poor performance in low radian conditions by some electrical parameters and functions modulations to get PV arrays. Its disadvantage is that it's costly and there are many errors associated with measurements.

Our approach includes solving the problem of dust accumulation that will lead to a poor performance due to the low amount of radians. We believe that a maintenance schedule and optimization is very good choice to solve the problem of low irradiance.

II. Perceived Contribution of the Paper

The model explained in the paper will be based on readings of voltages and currents, along with solar radiation using pyranometer and the temperature. The software used is a combination of MATLAB and LABVIEW coding. The model will output the I-V curve characteristic of the panel(s), a visual user interface to denote the functionality and the efficiency of the panel, and the action needed to be taken by the user regarding the panel.

III. Methodology

The first step in our work was drawing the scheme that will illustrate the electric circuit of our hardware part in the project. After the components are chosen, drilling and fabrication on the PCB (Printed Circuit Board) is done. The second step, is the setting and connections of the devices we have, like the photovoltaic panel used, the temperature and the pyranometer for measuring the solar irradiance and the temperature of the panel. We installed the second panel to be ready for the connection. The software used in this project is LABVIEW, in addition to the DAQ, which has a role of an interface between the hardware and the software used. The DAQ gives channels to measure the current and voltages in the circuit to be seen as graphs on LABVEIW. In the circuit we have, there are two modes of operation: charging and discharging of the capacitor. A five-pin relay is used to convert from charging to discharging and vise-versa. This relay is powered by an external biasing circuit. MATLAB is used for parameter estimation of the panel, which minimizes the errors found in the experimental results compared to the theoretical ones.

A. Hardware

The design is composed of a combination of resistors and capacitors that are placed based on circuit analysis to allow us to study and take readings from the PV Panel.
The hardware is based a primary design that consists of three main circuits.

1) Charging Circuit:

The charging circuit consists of the PV Panel with two branches in parallel to it. The following design allowed us to take accurate measures of current and voltage after taking reading from the junctions into a Data Acquisition Card and in turn to LABVIEW. The charging circuit is based on the use of

a capacitor that will have a maximum voltage for the minimum current and vice versa.

2) Discharging Circuit:

The discharging circuit consists of a capacitor and two series resistors. The resistor allows the capacitor to discharge and allows a junction point to read the current through the capacitor. Through the charge and the discharge of this capacitor we will be able to plot the I-V curve of the panel.

3) Switching Circuit:

The switching circuit consists of a 5-pin relay and a BJT that is controlled using a junction signal from a DAQ device interfaced with LABVIEW.
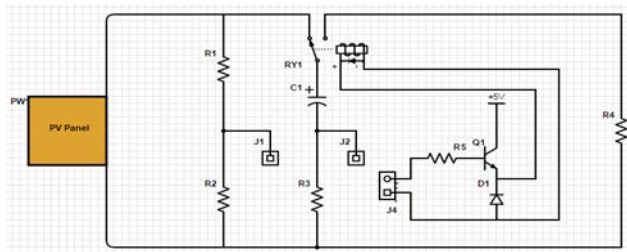
Figure 4: Circuit diagram of our design

B. Software

1) LABVIEW:

The LABVIEW blocks are used to receive the readings from the junctions labeled in Fig. 3. Furthermore, a series of logic arithmetic is used to overcome the technical constrains of the Data Acquisition Card. The LABVIEW blocks allowed us to take readings, save them into an Excel file, and plot graphs to analyze the panels' status. The over-all I-V Characteristic Graph will be used to assess the efficiency of the panel and the actions needed to be taken by the user.

The looping block in LABVIEW and the time delay allowed us to check the readings of the voltage and current of the panel, automatically at any time of the day.

2) MATLAB
(a) Modelling

PV Panels are modeled by an equation that best fits its characteristic curve. Every equation has a series of parameters that are dependent on several variables that are related to external conditions on the PV Panel. Most importantly, the temperature of the panel and the solar radiation falling on it along with its internal resistance determine the characterization of a panel that consists of a certain number of cells.

The model used has the following equation:

$$V = V_{oc} + N_s V_t m \ln\left(1 - \frac{I}{I_{sc}}\right) - IR_s \quad (1)$$

$V = Voltage\ of\ the\ Panel$
$I = Current\ of\ the\ Panel$
$V_{oc} = Open - Circuit\ Voltage$
$I_{sc} = Short - Circuit\ Current$
$N_s = Number\ of\ Cells$
$m = Ideality\ Factor$
$R_s = Internal\ Resistance\ of\ the\ Panel$
$V_t = Thermal\ Voltage$

Note that V_t is 25 mV at 298 °K and will change proportionally with temperature (°K)

Parameter Estimation

From the equation mentioned in the previous section, the only parameters we know are the voltage and the current which were saved in an excel file from the LABVIEW blocks.

The goal is to estimate the rest of the parameters using MATLAB at any time of the day, i.e. at any given temperature of the panel and at any solar radiation falling on the panel. The parameters we need to estimate are V_{oc}, I_{sc}, m, R_s.

Calculating the parameters allows us to estimate the efficiency of the panel by comparing the parameters calculated by the ideal values of these parameters that were calculated in ideal (lab) conditions.

(b) The MATLAB Process

The Excel File that contains the corresponding values of currents, voltages and temperature is given to the MATLAB functions. Initial values of all the parameters are entered. For every value of current there will be a corresponding voltage value using the modeling equation and using the initial values of the internal parameters. The calculated voltage value will differ from the experimental one by an error ε because of using non-accurate internal parameters. Then for all values of currents, ε will be represented as a column of values.
The next step is to compute the sum of errors squared in order to get accurate results of the parameters. The sum of errors squared will be computed as $\varepsilon^t * \varepsilon$ that will be given as a single value.
We will then use this value in addition to other parameters to the built-in MATLAB function "fminunc" that will return the set of the panels internal parameters that will correspond to the least error.

IV. EXPERIMENTAL RESULTS

The following test was done on Sunday, December 6, 2015 at 3:00 pm. The screenshots are taken from LabVIEW's graph blocks.

The process was taken with a delay of 10 minutes.

A. First test:
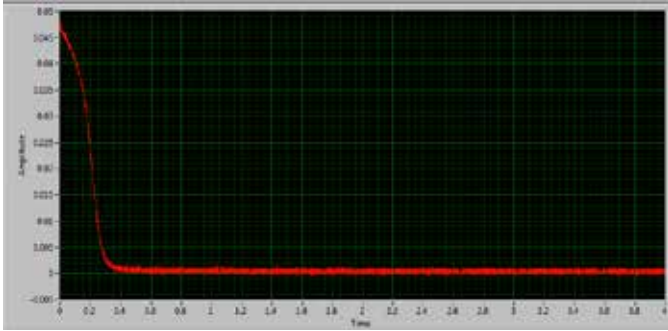
Figure 5: Current vs. Time Graph
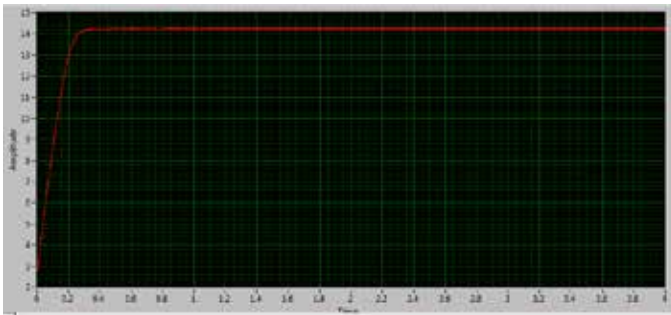
Figure 6: Voltage vs. Time Graph
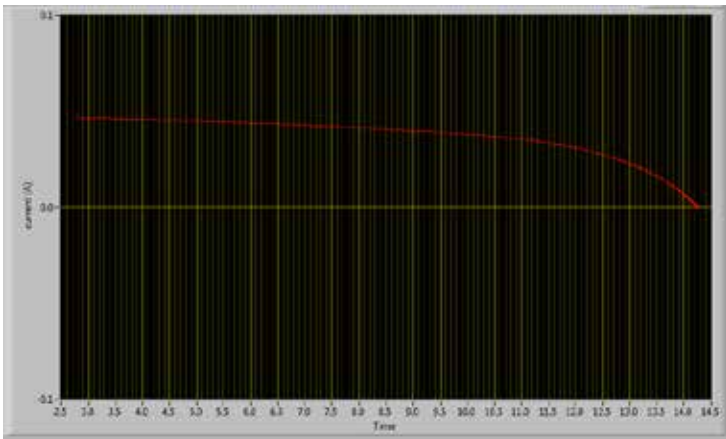
Figure 7: I-V Characteristic Curve

B. Second test
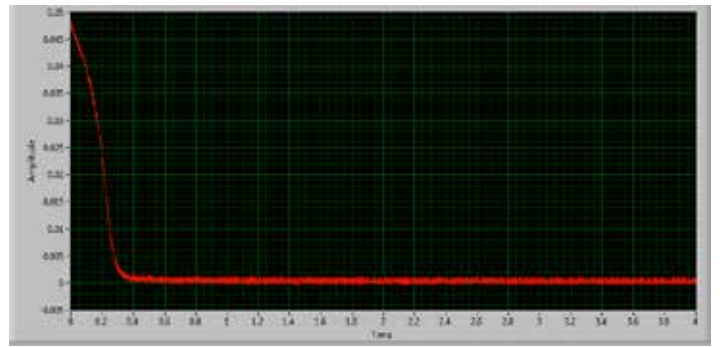
Figure 8: Current vs Time Graph
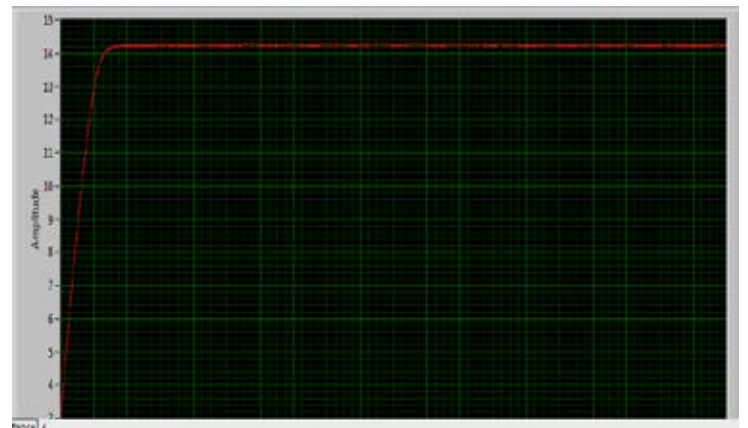
Figure 9: Voltage vs. Time Graph
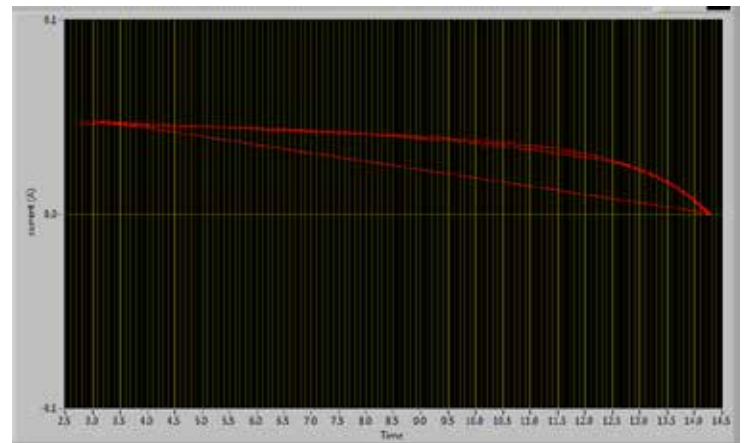
Figure 10: I-V Characteristic Curve

407

V. Numerical example result using Matlab

We took the above shown values and entered them as an array of voltages and current to our MATLAB function. Our function computed the error between the computed voltages using the characteristic equation and initial parameters' values. The error will be in the form of a column vector of rows equal to the samples of voltage and current entered (in our case 1000). The sum of errors squared is then computed and using our MATLAB function the best parameters will be the following:

V_{oc} = 14.2 V, I_{sc}= 0.047 A, m=0.92, R_s= 0.5 Ω.

Here we should note that m is low and R_s is high because the panel we are using is old and is not in good condition, which due to limited lab equipment provided to the research.

VI. Acknowledgment

This paper is the result of a collective effort of a team of three. Team work was essential to get the work done. Special thank you to Dr. Sami Karaki for guiding us through the project. Finally, we would like to thank the Labs of Engineering at the American University of Beirut for providing us lab equipment allowing this research to be conducted.

VII. References

[1] W. Durisch, "*Characterization of photovoltaic generators*", European Journal of Physics, Vol 32, No 3, 2011

[2] D. L. King, "*Photovoltaic Module and Array Performance Characterization Methods for All System Operating Conditions*", NREUSNL Program Review, AIP Press, pp.347-368, 1996

[3] W. De Soto, S.A Klein, and W.A Beckman, "*Improvement and Validation of a Model for Photovoltaic Array Performance*", Solar Energy journal, Vol 80, No 1, 2006

[4] B. P. Dougherty, A. H. Fanney and M. W. Davis, "Short-term Characterization of Building Integrated Photovoltaic Panels", Journal of Solar Energy Engineering, Vol 125, No 1, 2003

[5] T. Belel *"On The Role of Super capacitors towards Characterization of PV Generators*," Journal, Available FTP: http://www.ijareeie.com/upload/2014/august/37_On.pdf (2007*).*

[6] E.D Aranda, J.A. Galan, M.S. De Cardona, and J.M Rquez, (2009). "*Measuring the I-V Curve of PV Generators*," [Magazine], Available FTP: http://www.ieeexplore.org/xpl/recentissue.jsp.1932-4529/933882

[7] V. Nanjannavar, P. Gandhi, and N. Patel, "*LABVIEW Based PV cells characterization and MPPT under varying temperature and irradiance conditions*", Nirma University International Conference on Engineering (NUiCONE), pp 1-6, 2013

[8] D. L. King, W. E. Boyson, and J. A. Kratochvil, "Analysis of factors influencing the annual energy production of photovoltaic systems, Photovoltaic Specialists, IEEE Conference - PVSC , pp. 1356-1361, 2002

Biometrics Safe

Farah Nasreddine, Ahmad Diab, Elie Kouz
Department of Computer and Communications Engineering
American University of Science & Technology
Beirut, Lebanon

Abstract—Safety boxes are commonly used to store valuable items and possessions that need to be protected from theft or any other peril. Locks, with different complexities, can achieve a certain security level; however, biometric security; employing human characteristics to provide identification and authentication, is the trendier practice. This project proposes a dual biometric security that is based on both iris and voice recognition to secure a safety box. As an additional security measure, an SMS is sent in case an attempt has been made to access the safety box, which is interpreted as a mismatch with the owner's biometric data.

A camera placed on the safety box acquires the intricate details of the iris, and using MATLAB, specific features are extracted and encoded. As for the voice recognition; once the voice is captured using a microphone, MATLAB extracts the acoustic patterns and encode them. Mel-Frequency Cepstral Coefficients (MFCCs) along with vector quantization are used for feature extraction and pattern matching, respectively.

I. INTRODUCTION

Trespassing and hacking into personal devices are common crimes in today's world. Many security measures are being taken to avoid such issues. "The three main types of authentication are something you know (such as a password), something you have (such as a card or token), and something you are (biometric)" [1]. As Rene Ritchie mentioned in his article entitled The Future of Authentication: Biometrics, Multi-factor, and Co-dependency: "Today your basic password takes mere minutes, if not seconds, to break through. A string of letters and numbers known by just you is not enough to keep your accounts and devices secure". That is why a more reliable solution is biometric recognition, whereby the human distinctive features are the password.

Biometrics provide a protected method of authentication and identification, as they are hard to replicate and steal. Once the biometrics are used in conjunction with something you know, we reach what is known as a two-factor authentication. Along this principle, this paper introduces the Biometrics Safe; a box designed to protect personal belongings from unintended users through the joint identification of the iris and voice of the user.

Iris/Voice identification systems consist of four main stages:
(i) image/voice acquisition. (ii) pre-processing, (iii) features extraction, and (iv) matching procedures. For the iris identification process, the image acquisition includes iris localization, image normalization, and unwrapping, whereas, for the voice recognition process, it includes windowing, filtering, and Fast Fourier Transform (FFT). The features of the iris are extracted using Circular Hough Transform, while the voice features are extracted following the Mel-Frequency Cepstral Coefficients methodology (MFCC) [4]. Hamming distance and vector quantization are used for matching for the iris and the voice recognition procedures, respectively.

II. LITERATURE REVIEW

"The term biometrics is a composite word and it is formed by "bio", which means living creatures and "metrics" which means to measure an object, generally in a quantitative way, but in some occasions it refers to a qualitative criteria" [5]. Biometric verification is any means by which a person can be uniquely identified by evaluating one or more distinguishing biological trait. Unique identifiers include face, voice, fingerprints, hand geometry, earlobe geometry, retina and iris patterns, DNA, signatures etc. In addition, biometrics are classified according to two categories which are type and trait [6].

The selection of a certain biometric technology for user recognition or authentication depends on many factors, such as the security requirements of the application, the environment and the frequency of usage, the type of access required, the cost factor, the long term accuracy, etc…

Fingerprint authentication systems are simple to use and require inexpensive equipment. Besides they have shown a high verification speed and accuracy. Their main disadvantage is that they are not private: we all leave fingerprints almost everywhere.

The facial recognition is a user-friendly design and can perform massive identification but it suffers from an average performance that degrades when facial expressions vary due to age or other factors [7].

In the mid-1980s two ophthalmologists, Drs. Leonard Flom and Aran Safir, claimed that no two irises are alike, even in twins, thus making the iris a good biometric. This belief was based on their clinical experience where they observed the distinctive features of irises including the "many collagenous fibers, contraction furrows, coronas, crypts, color, serpentine vasculature, striations, freckles, rifts and pits" [8]. Iris identification took years to be developed and applied. Many recent studies show that iris-based authentication systems are one of the most secure systems since they can offer very low False Acceptance (FAR = 11 %) as well as False Rejection Rates ((FRR = 10.5%) [9]. Nowadays iris identification technology is spread throughout the world and is used especially in high security environments as in many airports [10] [11].

On the other hand, in order to protect secure buildings from unauthorized access, the people voice may replace many kind of devices traditionally used for such applications: pin pads, keys, smart cards, etc. However, the security performance produced by this biometric feature alone is not enough to ensure the high security required for some particular applications [12] [13].

Using biometric information in security systems has been a field of research largely studied with a large number of algorithms proposed in the recent years. Each technique has its advantages and drawbacks. The present work proposes to use a combination of two biometric procedures in order to control the access to a safe.

III. ACQUISITION

A. *Iris Acquisition*

Image acquisition is considered the most critical phase in our project, since all the subsequent stages depend highly on the quality of the image. Thus, a high focus camera is used for this purpose. The resolution is set to be 640 x 480 and the image to be saved with a .png extension. Also, the images are converted from RGB to gray levels, and from eight-bit to double precision, thus facilitating the manipulation of the images in later stages.

B. *Voice Acquisition*

The voice acquisition is performed through a microphone. In order to minimize the effects of aliasing that may occur during the analog-to digital conversion, the sampling rate used is 22050 Hz.

IV. PRE-PROCESSING AND FEATURES EXTRACTION

A. *Iris Pre-processing and Extraction*

Localization.
Localization is the process of separating the pupil from the iris by detecting, using a 2D spatial gradient edge detectors, the pupil boundary. Two sets of spatial masks are assigned respectively; the first is 3 x 3 for pupil's edge detection and the second is 5 x 5 for iris edge detection purposes [14]. Each set consists of four 90° shifted kernels that are applied separately to the input image.
To check the presence of an edge, a comparison between the convolution results and a specific threshold is conducted. Moreover, contour identification is conducted using a simple circular summation of the previously detected edge intensities. A contour is the circle with maximum summation. Finally, during the localization step, iris extraction is to be applied. However, in order to reduce the effects and errors behind the interference of the eyelids and eyelashes, only 80% of the iris region is to be extracted. Extraction consists of many stages: Detection Pupil Edges, Identifying Pupil Contour, Detecting Iris Edges, Identifying Iris Contour and Iris Region
Extraction. Circular summation of the intensities is the method applied with the aim to identify a contour, where a contour is the circle with the highest sum of intensities.
To identify first a contour, a set of specific circles (specific radii and centers) are assigned. One of these circles will sum up the edges detected in the earlier steps. This collision will result in the maximum sum which indicates the presence of the pupil. Finally, the center and radius coordinates of the pupil or iris are saved.

Normalization and Unwrapping.
To extract the region of interest, the iris should be changed into rectangular shape, considering the pupil center as the reference point, and circular vectors go through the iris area. The iris region is unwrapped by picking a consistent number of pixels along every radial line at a specific point, and subsequently creating a 2D 10 x 40 array. Taking into consideration the probability of the pupil to dilate, it is instructive to change the coordinates system used, by first eliminating the lower part of the iris and representing all the points within its limits into their polar equivalent. While unwrapping the picture, we tend to use the bilinear interpolation to obtain the intensities of the points in the new picture.

B. *Voice Pre-processing and Extraction*

Windowing
After converting the analog speech signal into a digital form, the windowing process consists of decomposing the signal into overlapping equal frames, where the signal is supposed to vary weakly, and thus it can be considered as quasi-stationary. The duration of each frame is about 20 to 40 ms, which is equivalent to 256 samples. The hamming windowing function defined as below is used for this purpose:

$$w(n) = 0.5 - 0.4 \cos\left(\frac{2\pi n}{255}\right), \qquad 0 \le n \le 255$$

Fast Fourier Transform
The 256 samples are converted from the time domain to the frequency domain using Fast Fourier Transform (FFT), which is implemented using the below equation:

$$X[k] = \sum_{0}^{255} x(n)e^{-2\pi jkn/256}, \qquad k = 0,1,2,\dots,255$$

Filtering
Filter bank processing takes the energy output of a series of adjacent triangular bandpass filters that simulate the frequency-selective mechanism in the human cochlea. The spacing of the center-frequencies of the filters as well as the bandwidth are determined by a constant mel-frequency interval that mimics the critical band of the human auditory system. The mel-frequency scale is a linear frequency spacing below 1000 Hz and a logarithmic spacing above 1000 Hz. As a reference point, the pitch of a 1 kHz tone, 40 dB above the perceptual hearing threshold, is defined as 1000 mels. Therefore we can use the following approximate formula to compute the mels for a given frequency f in Hertz [4]

$$mel(f) = 2595x \log_{10}(1 + \frac{f}{700})$$

The modified spectrum of S(w) thus consists of the output power of these filters when S(w) is the input. The number of mel spectrum coefficients, K, is typically chosen as 12 or 20 [4]. In our design, we set K = 20.

Cepstrum
Consider the mel power coefficients represented as:
$$S[k] = 1,2,\dots,K$$
Then, MFCC can be calculated using the below formula:

$$C[n] = \sum_{k=1}^{K} (logS[k]) \cos\left[\frac{n\left(k - \frac{1}{2}\right)\pi}{K}\right], \qquad n = 1,2,\dots,K$$

Because the mel coefficients are real numbers, then the Discrete Cosine Transform (DCT) is used.

V. MATCHING PROCEDURES

A. *Iris Matching Procedure*

2-D Gabor filter is the main filter utilized here to encode the iris zone since the amplitude contains the data we are worried about. While demodulating the iris pattern utilizing 2-D Gabor Wavelet, the pattern is encoded into 3002 bits iris code. The Hamming separation methodology is then used for synchronizing purposes; the Hamming distance between two strings of equal length is the number of positions at which the corresponding symbols are different. In our configuration, before the encoding step, we ended up with a normalized unwrapped iris region of size 10 x 40. This zone is convolved with the 2D complex Gabor wavelet of the same size, and consequently we get 19 x 79 = 1501 complex qualities. Every complex value is encoded into 2 bits, to obtain a 3002 bits iris code prepared to be coordinated utilizing the hamming distance approach.

B. Voice Matching Procedure

To apply feature matching, the Vector Quantization (VQ) method is utilized. VQ is a procedure of mapping vectors from an expansive vector space to a limited number of locales in that space. Every locale is known as a cluster and can be represented by its corresponding codeword. The collection of all codewords is known as a codebook. In the preparation stage, a speaker-particular VQ codebook is produced for each known speaker by clustering his/her preparation acoustic vectors (set of MFCCs). The resulting codewords are called centroids. The distance from a vector to the nearest codeword of a codebook is known as a VQ-contortion. In the acknowledgment stage, an information expression of an obscure voice is "vector quantized" using each prepared codebook and the aggregate VQ distortion is processed. The speaker correlated to the VQ codebook with smallest aggregate contortion is recognized.
See Figure. 1.
We implement the VQ technique with 8 training vectors and 100 codebook vectors, by doing the following

1) Plan a 1D-vector codebook which is the centroid of the whole arrangement of preparing vector.
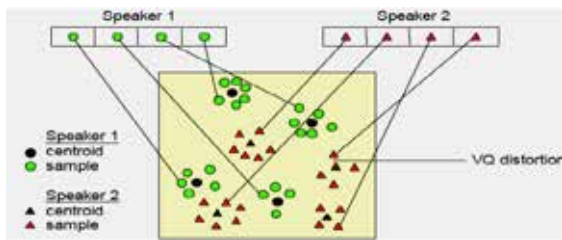
Figure 1. Vector quantization codebook formation

2) Double the size of the codebook by splitting each current codebook.
3) For each training vector, we find the code-word in the current codebook that is closest, and assign that vector to the corresponding.
4) We update the code-word in each cell using the centroid of the training vectors assigned to that cell.
5) Iteration 1: repeat steps 3 and 4 until the average distance falls below a preset threshold.
6) Iteration 2: repeat steps 2, 3 and 4 until a codebook size of 100 is designed.

VI. DESIGN FRAMEWORK

A. Functionality

After the development and validation of both identification techniques, it is the time to practically implement them on the safe. The two biometric systems are combined together through an Arduino Uno. First, the digital output extracted from the MATLAB is connected to the Arduino through serial connections. These two results will occupy the two of the three input pins of the Arduino, while the remaining one will be used for the password alternative. All the three input pins will communicate with each other using an AND gate function. The functionality of this combination is to open the safe, when the result of the AND gate is one in binary, which is the same as the three input results. If one of the results mismatches then another chance is given by an implemented loop inside the Arduino. When the second attempt ends with a failure, this will trigger an impulse for the GSM shield to function as an alert for the user and send him/her an SMS stating "Watch out Someone is Hacking Your Safe". In the case when all output verifications are true leading a true AND combination, this will output a binary one as a result that will be the input to the solenoid of the safe. But due to the circulation in this process, the output signal would be of high voltage and low current. That's why, we are in need to a relay in order to increase the current. This increase is provided by an external power supply of 5V and 0.5A specifications. The resulting signal is then more powerful and leads to open the safe by the solenoid.

B. System Flowchart

The flowchart of the system is provided in Figure. 2. It shows the different steps in order to access the protected safe.
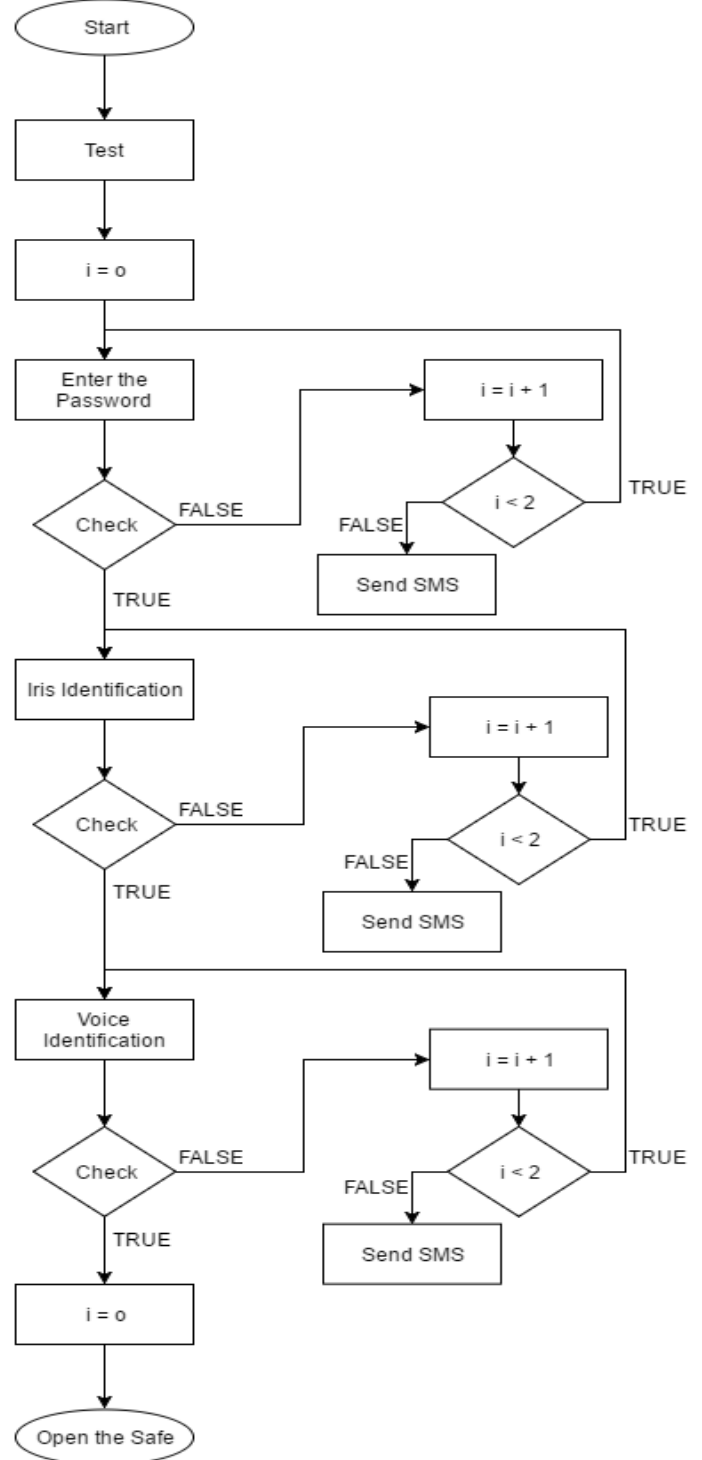
Figure 2. Design Flowchart

VII. EXPERIMENTS AND RESULTS

A. Experimental Environment

Biometrics Safe is a system developed using the MATLAB platform. Every algorithm implemented is tested separately and each has its own measurement techniques. The performance of any biometric system can be measured by the accuracy and speed of the system. For both systems, a sample of each biometric identity was stored in the database.

Both templates are stored in the ideal condition; the iris sample in the normal light background and the voice sample in the laboratory condition.

Iris Enrollment Conditions

The enrollment procedure was accomplished in the following conditions: low light background, normal light background, high light background, same individual's left eye, with eyeglasses, and a new person. The system is tested five times in each condition.

Voice Enrollment Conditions

The enrollment procedure was accomplished in the following conditions: laboratory condition, noisy surrounding, with TV, new person in noiseless condition, and new person in noisy condition. The system is tested five times in each condition.

Iris Implementation Result

After five testing attempts in each condition, the results are shown in the table below.

TABLE I.   IRIS SYSTEM TESTING

Conditions	Successive Attempts	Failure Attempts
Low Light Background	2	3
Normal Light Background	4	1
High Light Background	3	2
Same Individual's Left Eye	1	4
With Eyeglasses	0	5
New Person (in normal conditions)	0	5

As noticed from the second row in the above table, 80% of the attempts are successful. These attempts were experimented under the normal light effect, which is considered as the ideal circumstance. The failure is due to the deforming of the eye shape of the same individual. T
The last row of the table shows that the system succeeds in rejecting all the attempts of a strange individual (100%).

Voice Implementation Result

After five testing attempts in each condition, the results are shown in the table below.

TABLE II.   VOICE SYSTEM TESTING

Conditions	Successive Attempts	Failure Attempts
Laboratory Condition	5	0
Noisy Surrounding	2	3
With TV	3	2
New Person in Noiseless Condition	0	5
New Person in Noisy Condition	1	4

As for the voice identification system, 100% is the percentage of successful attempts (TAR) in the ideal environment (first row) and 0% is the percentage of FAR (fourth row): an intruder never succeeded to open the safe. Yet, 20% is the error rate where the system responded to a same gender intruder in a noisy surrounding row 5).

Biometrics Safe: System Testing

To test the accuracy of the whole system, the previous databases are joined together. For 20 different conditions based experiments:
- 14 trials resulted in true acceptance (TAR).
- 3 trials resulted in false rejection (FRR).
- 1 trials resulted in false acceptance (FAR).
- 2 trials resulted in true rejection (TRR).

The decrease in the performance of the combined system (70 % as TAR) is due to the fact that the system was tested in more real conditions, where the individual ones were validated in laboratory conditions. It can be noticed also that the "Biometrics Safe" diminishes the False Acceptance Rate to 5%, highlighting the impact of the proposed technology.

VIII. CONCLUSION

Implementing the Biometrics safe gave us a wide knowledge about various topics. This experience gave us the opportunity to master the topics of iris and voice recognition, and heavily introduced us to the different approaches involved in the topic of eye and voice anatomies in order to practically combine multiple biometrics. The requirement for secure techniques for validation is turning out to be progressively essential in the corporate world today. Passwords, token cards and PINs are all dangers to the security of an association because of human instinct. Our powerlessness to recollect complex passwords and inclination to record these alongside losing token cards or overlooking PINs, all add to the conceivable breakdown in security for an association. Implementing the Biometrics safe unified different approaches used in biometric security and provided a more accurate authentication measure.

REFERENCES

[1] Khaw, Penny. Iris Recognition Technology for Improved Authentication. SANS Institute Infosec Reading Room.

[2] Richie, Rene. The Future of Authentication: Biometrics, Multi-factor, and Co-   dependency.

[3] Phil Nickinson said: "In a world of biometric security, you are the password", 2008K. Elissa, "Title of paper if known," unpublished.

[4]   G. Min, X. Zhang, J. Yang, and X. Zou, "Speech reconstruction from mel-frequency cepstral coefficients via L1-norm minimization," 2015 IEEE 17th International Workshop on Multimedia Signal Processing (MMSP), pp.1–5, Oct. 2015.

[5]   Manuela Marzotti, Cristina Nardini. "Biometric Authentication using Voice". TDDC03 Projects, Spring 2006

[6]   Biometrics institute. "Types of Biometrics" Crows Nest NSW 1585, Australia.   For   reference, http://www.biometricsinstitute.org/pages/types-of-biometrics.html

[7]   Rupinder Saini, Narinder Rana. "COMPARISON OF VARIOUS BIOMETRIC METHODS". International Journal of Advances in Science and Technology (IJAST) Vol 2 Issue I (March 2014)

[8]   Iris-scan.com.   Iris   Recognition: Issues.   URL: http://www.iris/scan.com/iris_cautionary.htm (11 February 2002).

[9]   C. R. Prashanth1, Shashikumar D.R.2, K. B. Raja3, K. R. Venugopal3, L. M. Patnaik4. "High Security Human Recognition System using Iris Images". International Journal of Recent Trends in Engineering, Vol. 1, No. 1, May 2009.

[10]   "Backgrounder". *Canadian Air Transport Security Authority (CATSA)*.

[11]   "Iris scans at Amsterdam Airport Schiphol". Amsterdam Airport Schiphol.   Retrieved 2013-07-14.

[12]   Hairol Nizam Mohd. Shah*, Mohd. Zamzuri Ab Rashid, Mohd. Fairus Abdollah, Muhammad Nizam Kamarudin, Chow Kok Lin and Zalina Kamis. "Biometric Voice Recognition in Security System". Indian Journal of Science and Technology, Vol 7(2), 104–112, February 2014.

[13]   Wahyudi, Winda Astuti and Syazilawati Mohamed. "Intelligent Voice-Based Door Access Control System Using Adaptive-Network-based Fuzzy Inference Systems (ANFIS) for Building Security". Journal of Computer Science 3 (5): 274-280, 2007.

[14]   Jamil Abou Saleh, Abdallah Bou Saleh, Samer Sharafeddeen, Professor Mohamad A. Al-Alaoui. "Iris Recognition". American University of Beirut.

C-RAN clustering in highway access

Diana Ammar, Layale Hachem
Department of Electrical and Mechanical Engineering
Saint Joseph University of Beirut (ESIB)
CST, Mar Roukos, Beirut, Lebanon
diana.ammar@net.usj.edu.lb layale.hachem@net.usj.edu.lb

Abstract- **By separating Baseband Units (BBUs) from Radio Remote Heads (RRHs) and running these BBUs on a cloud, a new era of mobile communications has started where radio cells can be dynamically adjusted to adapt to the network traffic demand. In this paper, we tackle the problem of RRH clustering in the case of linear highway deployment with different traffic conditions. The clustering problem aims at finding an optimal partition of RRHs where each subset will constitute a cell associated with one BBU in the cloud. We model the traffic dynamics by a Jackson network of queues and use this model to derive the quality of service metric. We then propose a utility function taking into account the quality of service, the signaling load, and the energy consumption. This study seeks to minimize the use of energy and BBU resources while maintaining users' satisfaction. Therefore, several heuristics to achieve the clustering are presented and compared to the optimal but complex exhaustive-search based algorithm. The comparison is two-dimensional; traffic load density and type of users' distribution in the system are taken into consideration. Finally, we develop a Matlab simulator to test our approaches in different traffic scenarios and state the efficiency of each algorithm.**

INTRODUCTION

Traffic loads on mobile network are in a continuous exponential growth. According to Ericsson Mobility Report [1], mobile data traffic is growing 65 percent year over year. Further, Cisco Visual Index (VNI) [2] reports that the global data traffic is forecasted to increase nearly eightfold between 2015 and 2020, with the growing number of connected devices. Therefore, mobile network operators have to increase the capacity and network coverage to satisfy users' demand either by using heterogeneous networks or adding more cells. However, inter-cell interference levels will increase and every new cell will add to the capital and operational costs of the operator. In addition, growth in data also severely impacts power consumption, with consequent cost burdens. Most of the power consumption is in the radio access networks, specifically at base stations. Hence, new technologies that optimize cost and energy consumption become a necessity for mobile operators.

To address these challenges, Cloud-based Radio Access Network (C-RAN) architectures have been adopted by different mobile operators. C-RAN architecture includes Remote Radio Heads (RRHs) connected to a centralized base band unit (C-BBU). The RRHs are simple, light-weight radio units with antennas responsible for the filtering and amplification of the signal. Unlike traditional RAN, the BBU is separated from RRHs and performs the signal processing functionality of the RAN. A BBU can control one or more

RRHs depending on the limit of data volume a BBU can handle. In addition, a number of BBUs can be aggregated to form the centralized base band unit (C-BBU) as shown in Figure 1; it is a new RAN architecture with two-dimensional clustering of RRHs and BBUs. Moreover, C-RAN defines a virtualized BBU-pool that can change the configurations of the clusters (change the combination of BBUs and RRHs while not exceeding the upper limit of BBU resources). Centralized signal processing reduces the number of equipment rooms at sites. As a result, the power consumed by air conditioning and other site equipment is significantly reduced. Furthermore, the connections between RRHs and BBUs are changed dynamically according to the traffic demand so the number of BBUs is decreased. Thus, the power and energy consumption are reduced resulting in a lower cost of network operations.

A recent study [3] proposes algorithms that determine configurations maximizing the traffic demand satisfied on the RAN while optimizing the utilization of base band resources and minimizing the energy consumption. The authors in [4] formulate the clustering problem as a bin packing problem with a utility function comprising both the user's quality of service and the power consumption. However, this study does not take mobility into account. In [5], the radio resource clustering problem is jointly considered with the users scheduling in order to enhance the benefits of CoMP (Coordinated Multi-Point) transmissions. Here also, the mobility non-uniform user distributions are not taken into account. In the context of heterogeneous C-RAN, a dynamic frequency reuse scheme is derived in [6]. A graph coloring method is used to perform an adaptive bandwidth allocation based on the traffic demand. The BBU-RRH mapping is dynamically changed leading to a decrease in the energy consumption due to the reduce number of active BBUs.

In this paper, we propose different algorithms that tackle the problem of RRHs clustering in the case of linear highway access with different traffic conditions. The importance C-RAN in highway coverage has been highlighted in [7]; in a high speed environment, frequent handovers will lead to voice and data services breakouts. This should be combatted by an adaptive clustering taking into account, not only the system capacity, but the network performance as well. The purpose of our work is to create dynamic configurations of clustering in order to find the optimal partition of RRHs where each subset will constitute a cell associated with one BBU in the BBU-pool. This study seeks to find an energy-efficient clustering

configuration while reducing BBU resources and maintaining users' satisfaction.

This paper is organized as follows. We introduce the system model in Section II and present the heuristics in Section III. We summarize our simulation results in Section IV and we conclude in Section V.

II. SYSTEM MODEL

A. Highway, Network, and Traffic Models

We consider a linear deployment of n RRHs on a highway access as shown in Figure 1 where we study different heuristics of RRHs clustering. We assume that the RRHs are covering only one direction of the highway, but the model can be easily generalized to deal with bi-directional vehicles traffic. Vehicles can enter or leave the highway.

We denote by $R = \{R_i, i = 1..n\}$ the set of all RRHs numbered according to their linear location. $B = \{B_k, k = 1..b\}$ is the set of virtual BBUs (clusters) that are allocated to RRHs and where b is simply the number of clusters (or cells) verifying $1 \le b \le n$. $i \in k$ means that R_i is served by BBU B_k.

In the following, the term session designates a communication session initiated and terminated by a user. We adopt a session-level traffic model where we assume that for each RRH R_i, we know the following traffic parameters:

- $\lambda_{i,new}$: The arrival rate of new sessions to R_i, assumed to follow a Poisson process. By new session, we mean a session initiated by a user inside R_i or a session that just entered the highway through R_i.
- h_i: The rate of sessions leaving R_i and entering R_{i+1}.
- λ_i: the total session arrival rate to R_i. $\lambda_i = \lambda_{i,new} + h_{i-1}$.
- μ_i: the parameter of the exponential distribution governing the session duration. In other terms, the average session duration in R_i is μ_i^{-1}. Note that a session leaves R_i if it is terminated, or if the vehicle exits the highway, or if the vehicle moves to R_{i+1}.
- $\rho_i = \lambda_i/\mu_i$: Traffic intensity in R_i. This is also the average number of sessions in R_i.
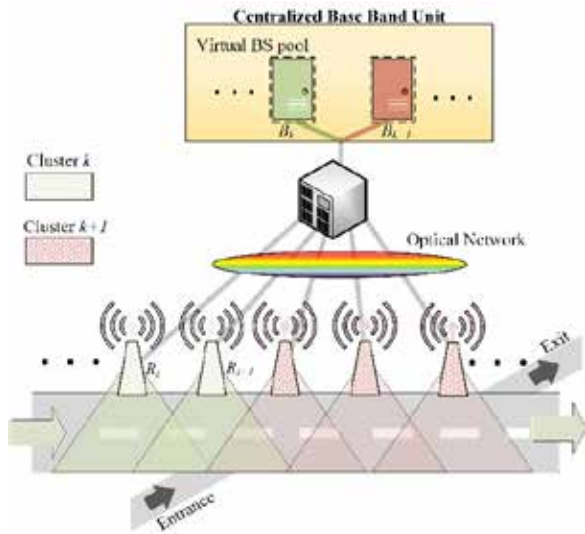
Figure 1. C-RAN architecture and highway system model.

B. Quality of Service Model

In this work, the Qality of Service (QoS) indicator in R_i will be related to the congestion probability G_i which is the probability that the number of ongoing sessions in R_i exceeds a given threshold m.

In order to derive G_i, we model each R_i as an infinite-server queue where the total session arrival rate is λ_i and the session service rate is μ_i. Therefore, the system is Jackson network of n queues of type $M/M/\infty$. Jackson theorem states that each queue acts as in independent $M/M/\infty$ ([8]) queue where we can define the congestion probability as:

$$G_i = \sum_{j=m}^{\infty} \frac{\rho_i^j}{j!} e^{-\rho_i} \qquad (1)$$

We define then the QoS indicator Q_i as:

$$Q_i = \rho_i(1 - G_i) \qquad (2)$$

C. Traffic and QoS models for clusters

The previous models were derived for individual RRHs but are generalizable to any cluster. In fact, we can easily show that the clusters form a Jackson network where each cluster is a $M/M/\infty$ queue whose traffic intensity is the sum of the traffic intensities of its constituent RRHs. The previous traffic parameters and QoS formulas remain the same and we just replace ρ_i with ρ_k the sum of the traffic intensity of the RRHs that form a cluster, hence,

$$\rho_k = \sum_{i \in k} \rho_i \qquad (3)$$

D. Energy Consumption Model

The energy consumed by one BBU depends on the number of sessions served. So, we define in this paper the energy consumed by one BBU, B_k as follows:

$$E_k = \alpha + \beta.D_k.\rho_k \qquad (4)$$

where α is the minimum value of the energy consumed by a BBU (when there is no session to serve), and $\beta D_k \rho_k$ is the total amount of energy added to α that depends of the number of arrival sessions, with D_k denoting the average bit rate switched by B_k, and ρ_k denoting the average number of sessions served by B_k.

In our study, each session is a real-time traffic. We denote by d the bit rate generated by each session and D_{max} the maximum bitrate that a BBU can offer to its constituent RRHs. We assume that D_{max} is a multiple of d with $D_{max} = m.d$. Since we consider real-time traffic:

- If the total number of sessions in B_k is larger than m, then each session will have a bit rate equal to $\frac{D_{max}}{total\ number\ of\ sessions}$.
- If the total number of sessions in B_k is smaller than m, then each session will have a bit rate equal to d.

415

Therefore, based on [9], the average bit rate for each session is given by :

$$D_k = d(1 - Erl\_B_k)(1 - G_k) + \frac{Dmax G_k}{\rho_k} \qquad (5)$$

Where: Erl_B_k represents the Erlang-B formula which gives the blocking probability in the Erlang loss model of B_k:

$$Erl\_B_k = \frac{\frac{\rho_k^m}{c!}}{\sum_{i=0}^{m} \frac{\rho_k^i}{i!}} \qquad (6)$$

and G_k represents the congestion probability in B_k:

$$G_k = \sum_{j=m}^{\infty} \frac{\rho_k^j}{j!} e^{-\rho_k} \qquad (7)$$

E. Utility Function

The decision taken in order to cluster several RRHs is based on the utility function that takes into consideration three parameters: energy consumption, signaling loads and quality of service. We propose the following formula for the utility that is a weighted sum of the three parameters:

$$U = u \frac{\sum_{k=0}^{b} E_k}{Pmax} + v \frac{\sum_{k=0}^{b} Q_k}{\sum_{k=0}^{b} \rho_k} + w \frac{\sum_{k=0}^{b} h_k}{\sum_{k=0}^{b} \rho_k} \qquad (8)$$

Q_k is the QoS indicator of B_k obtained using eq. (2). E_k is the energy consumed by one BBU obtained using eq. (4) and h_k is the handover rate between R_k and R_{k+1}. It is equal to the rate of sessions leaving R_k and entering R_{k+1}.

III. ALGORITHMS AND HEURISTICS

In this section we describe the optimal algorithm as well as the heuristics we derived in order to find a sub-optimal solution. In heuristics, during each iteration, we compare the previous value of the utility function of the global system with the new one obtained after altering the clustering. If the new utility value is greater than the previous value, the new configuration of clustering is adopted; otherwise the old configuration is maintained.

A. Optimal Algorithm

The optimal algorithm takes into account all possible partitions of RRHs. The starting point is individual RRHs, and with each partition the utility is evaluated. At the end, the partition presenting the best utility is adopted. This means that this algorithm goes through every single solution possible and selects the best one. Thus, all other approaches shall be compared to it when evaluating performances.

B. Two_Ways Algorithm

The two ways algorithm forms clusters from initial aligned RRHs. The approach starts by creating a cluster formed by one RRH, the first one. Then, the cluster tries to grow itself by adding adjacent RRHs. Once the utility deteriorates, the cluster will be considered completed and another cluster starts to form from the next first single RRH and so on. The same procedure is repeated on the single RRHs but from the other side, meaning if the first time the algorithm is executed from

left to right it will be repeated the second time from right to left and eventually the best configuration will be selected.

C. Adaptive Algorithm

- *Adaptive 1*

Starting from an initial configuration state where a certain number of remote radio heads is assigned to each cluster, a first adaptation approach consists of each cluster trying to snatch a RRH or more from the cluster right next to it. The RRH rearranging stops once the utility deteriorates. In extreme cases, this strategy allows a cluster to steal all RRHs of a system leaving all other clusters with 0 RRHs. This method runs through adjacent clusters in both ways starting from left to right then it repeats itself in the other way.

- *Adaptive 2*

In an alternative approach, a cluster B tries stealing only one remote head from the pool of RRHs that contains both clusters A and C adjacent to B. The RRH selection is made when the highest value of utility is achieved. Understandably, segment end clusters only perform this technique on one adjacent cluster. Each cluster undergoes the same procedure until all clusters have participated in this game of loss and gain. Ideally, this approach should be repeated x number of iterations until achieving full convergence, with x determined according to the convergence condition. Our algorithm currently stops after the first iteration. When searching for a convergence condition, the user can stop the computation after a fixed number of iterations determined statistically or when detecting an unchanged combination of RRHs.

IV. SIMULATION RESULTS

We evaluate the performance of the proposed algorithms comparing with the optimal configuration and the conventional cell deployment. We studied the algorithms on different cases and came up with the following results.

The comparison between the optimal algorithm and the other algorithms is two dimensional, the first dimension being the density of user distribution and the second one being the type of user distribution (homogenous – traffic is of same level in all areas or heterogeneous - some areas contain more traffic than others) . The assumed number of RRHs is ten. Therefore, in a conventional cell deployment, we need ten RRHs and ten BBUs.

We fix the maximum number of sessions that can be served by one BBU to 100 ($m = 100$) and the traffic load for each RRH represents a ratio of a number T. T goes from 1 to 100 representing off peak hour when T is at low values and peak hours when T approaches 100. In a homogeneous network, this fraction is approximately the same in all cells, in a heterogeneous one this fraction varies greatly from one cell to another. We fix the handover rate as a ratio of the traffic load. We test the algorithms with fixed parameters' values. These are as follows:

- m = 100 Sessions
- d = 2 Mbits/s
- α = 50 Joules
- β = 2

The reader should note that we are considering the average number of sessions in the system, therefore the session duration is taken into consideration when averaging this number.

A. *Evaluation of the Utility function*

The utility function is evaluated in different cases. The traffic load of each RRH is generated randomly as a ratio of the density of user distribution. The vertical axis represents the value of the utility while the horizontal axis represents the arrival rate of sessions.

When comparing the utility of the different algorithms as shown in Figure 2, we can notice that the optimal configuration represents the best utility for low and high traffic density in both homogenous and heterogeneous distribution. The two ways algorithm is equivalent to the optimal algorithm in both case where users are homogenously and heterogeneously distributed across the highway and when traffic is at off peak hour meaning users' density is low. Considering homogeneous distribution for high traffic, both Adaptive algorithms (Adaptive 1 and Adaptive 2) have demonstrated good but yet not efficient performance compared to the optimal configuration. Therefore we can conclude that for a network where users are homogeneously distributed, two ways algorithm delivers better performance than other algorithms during off peak hours, Adaptive 1 and Adaptive 2 for on peak hours. As for heterogeneous distribution, the two ways algorithms must be adopted for off peak and on peak hours. The reader should note though that this is a scenario where probably each RRH would perform better on its own, this is why any attempt of clustering will deteriorate greatly user's experience.

b) Heterogeneous Distribution
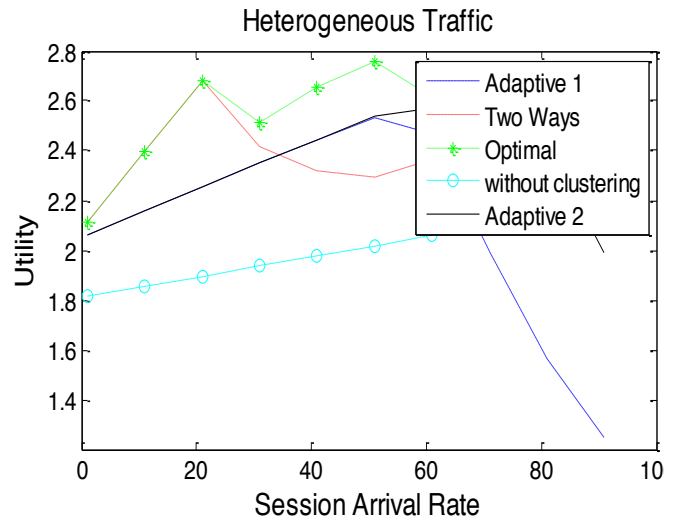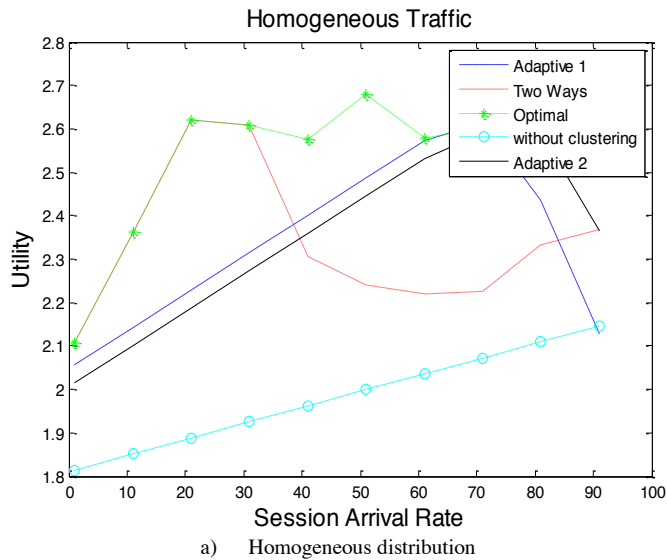
Figure 2-Evaluation of the Utility Function for a) Homogenous and b) Heterogeneous Distribution

B. *Evaluation of the energy consumption*

During the next step, we needed to evaluate the energy consumption of the clustering proposed by each algorithm. As we can see in the Figure 3, the optimal configuration of clustering consume energy the least with both distribution during off peak. All other approaches present higher energy consumption than the optimal configuration but lower than initial cell deployment without clustering during off peak hours. We can conclude from the graphs below that during off peak hours, the two ways algorithm is power efficient; for peak hours, the adaptive algorithm. With these algorithms, we were able to reduce the energy consumption in the radio access network, thus achieving C-RAN objective for having a power-efficient radio access network.

a) Homogeneous Distribution

b) Heterogeneous Distribution
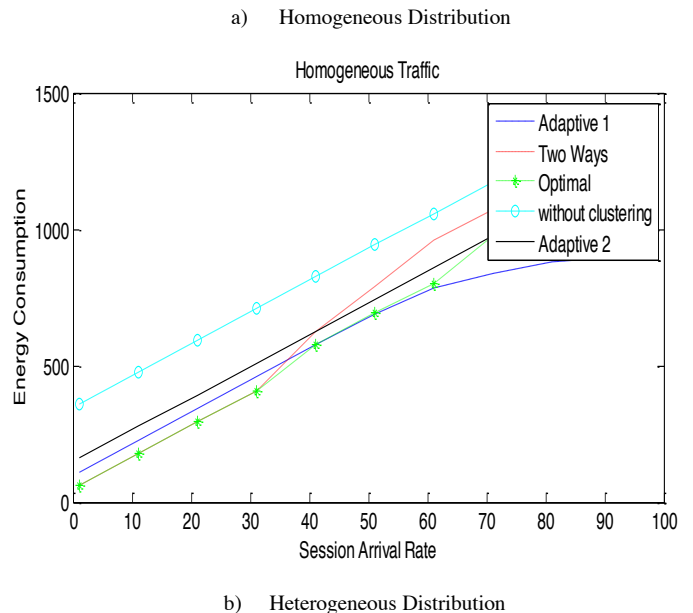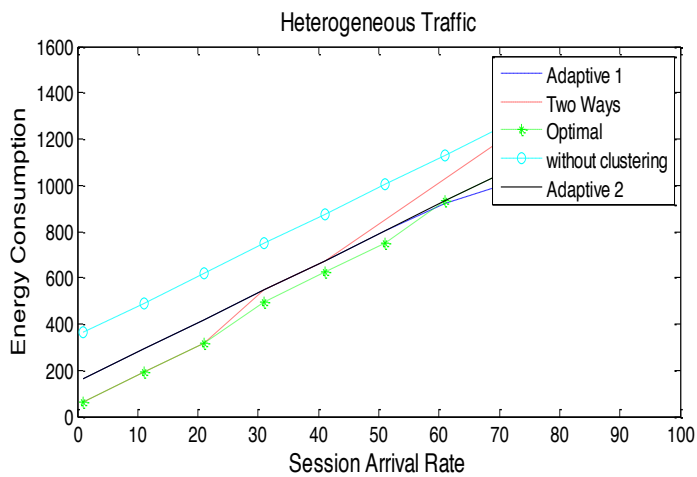
Figure 3-Evaluation of the Energy Consumption for a) Homogenous and b) Heterogeneous Distribution

V. CONCLUSION

C-RAN technology is progressively becoming a very important part in today's mobile network architecture considering its several advantages. Nowadays, multiple approaches and configuration algorithms exist but the technology is still considered as optional. In 2019, C-RAN will impose itself as a significant manner for mobile networks deployment. Consequently operators will be forced to study, test and implement different configuration approaches. Choices will be made according to geographical factors and will possibly change depending on time intervals. Approaching optimal configuration is widely possible but will result in decreased utility value, alternatively adopting an optimal configuration that maintains the highest level of utility will require a reliable central processing unit and will introduce a higher level of complexity.

The choice is purely subjective. Finally, it's up to each operator to decide upon the compromise between the level of satisfaction it wishes to offer its users and the total cost of ownership.

REFERENCES

[1] Ericsson, "Ericsson Mobility Report. On the pulse of network society", November 2015.

[2] Cisco Visual Networking Index Forecast, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015–2020", White Paper, 2016.

[3] K. Sundaresan, M. Y. Arslan, S. Singh, S. Rangarajan, and S. V. Krishnamurthy, "FluidNet: A Flexible Cloud-based Radio Access Network for Small Cells", *IEEE/ACM Trans. Netw.*, vol. PP, no. 99, pp. 1-14, 2015.

[4] K. Boulos, M. El Helou, and S. Lahoud, "RRH clustering in cloud radio access networks", *International Conference on Applied Research in Computer Science and Engineering (ICAR)*, Beirut, 2015.

[5] Y. Du and G. De Veciana, "Wireless Networks Without Edges: Dynamic Radio Resource Clustering and User Scheduling", In *Proc. IEEE Conference on Computer Communications (INFOCOM)*, pp. 1321 – 1329, 2014.

[6] K. Wang, M. Zhao and W. Zhou, "Traffic-aware Graph-based Dynamic Frequency Reuse for Heterogeneous Cloud-RAN", In *Proc. IEEE Global Communications Conference (GLOBECOM)*, pp. 2308 – 2313, Austin, 2014

[7] Next Generation Mobile Networks, "Suggestions on Potential Solutions to C-RAN", technical document, 2013.

[8] L. Kleinrock, *Queueing Systems*, Volume 1: Theory, Wiley-Interscience, 1975.

[9] T. Bonald and M. Feuillet, *Network Performance Analysis*, Wiley-ISTE, 2011.

Direction Finding and Localization of Transmitters

Ghassan Fakher, Karim Kaafarani
Electrical and Computer Engineering
American University of Beirut
Bliss Street, Beirut, Lebanon
gkf01@mail.aub.edu, karimkaafarani@gmail.com

Abstract- Spectrum monitoring, in particular RF Frequency monitoring, detects any interference caused by illegal transmissions over a large area. It is well known that interference on the spectrum causes a degradation in the QoS, thus causing immense revenue losses for network. In this paper, a direction finding tool was developed to determine the azimuth angle of arrival of sources using NI-USRP. A controllable SP4T antenna switch was developed on Agilent ADS to reduce the number of USRPs.

I. INTRODUCTION

With the increase in the number of mobile broadband subscribers reaching 5 billion by 2016 [1] and the full occupancy of the available spectrum, mobile operators are required to fully maintain their allocated spectrum to ensure a good quality of service for their customers. Spectrum monitoring, in particular RF Frequency monitoring, detects any interference caused by illegal transmissions over a large area. It is well known that interference on the spectrum causes a degradation in the QoS, thus causing immense revenue losses for network operators whose band is being jammed. Therefore, operators are in desperate need for a tool that can effectively identify and locate illegal transmissions to reduce costs and provide a better service for customers. The aim of this project is to implement a direction finding algorithm on an antenna array to accurately determine the location of any illegal transmitter in a limited time duration, thus helping Touch, a Lebanese network operator and a sponsor of this project, to maintain its spectrum more effectively.

II. MOTIVATION

The emergence of new data services, such as HD video streaming, that require high bit rates and less interference has forced operators to improve their quality of service by either increasing the bandwidth, which is not a feasible option since bandwidth is a scarce resource as mentioned earlier, or decreasing the interference imposed on their allocated spectrum which is possible to do thanks to our proposed solution. Illegal transmitters installed by end users without consent of the operator contribute a great portion of the interference imposed on the network [2]. It is a serious problem network operators suffer from in terms of decrease in QoS, revenue losses, loss of customer trust and increased cost of maintaining the network; according to Touch, the call drop rate increases to a high 3.12%. Hence, to solve this issue, network operators need a direction finding tool that can effectively determine the location of any illegal transmitter. According to ABC news, illegal transmitter locators are present today in the market but do not provide an effective solution [3].The reason behind that is those devices require manual movement of the antenna to determine the direction from which the maximum interference emerges. In addition, existing devices are not time and cost effective. Therefore, the emphasis of the project lies in detecting, tracking, and managing interferences in order to reach an improved service quality and cost control for network operators.

III. DIRECTION FINDING TECHNIQUES

A. Radio Direction Finding

Radio direction finding or RDF for short, is the action of measuring the direction or location of an unknown transmitter. Systems of RDF try to obtain the Angle of Arrival (AOA) of a signal by computing the phase or the time of arrival of the incoming wave. Almost all modern applications nowadays use the data from the AOA to determine the location [4].

1) Amplitude Comparison:

A design with a Category I have a system that changes the amplitude response of the antenna elements into voltages in order to acquire the AOA data. Figure 1 below shows the two kinds of methods used to compute the AOA data. θ shows the calculated Angle of Arrival.
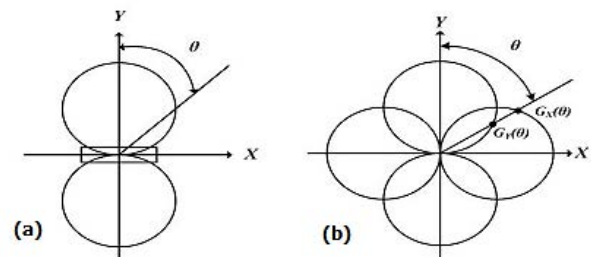
Figure 1: (a) Direct (b) Comparative Amplitude Techniques "Source : [4]"

2) Phase Delay and Time Delay:

In order to reach the design set by the two categories, a minimum of two antennas must be present with a fixed space between them. Figure 2 below shows the basic configuration for the two above mentioned Categories [4].
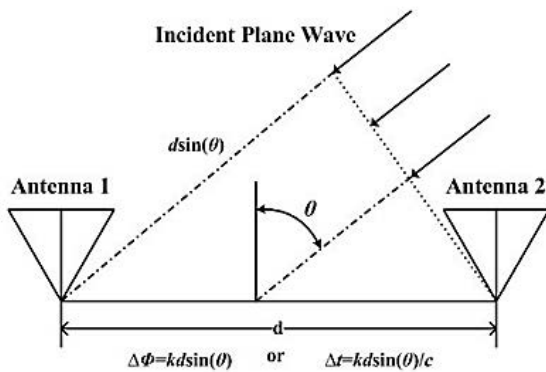
Figure 2: Basic Phase-Delay and Time-Delay Techniques "Source: [4]"

Since the waves encounter unalike propagation distances from a transmitter to the antennas, the waves form different phases and hence the waves arrive to each individual antenna at different time. Also since the phase or time differences are dependent on spacing and wavelength, we can then easily compute the AOA data.

3) The Maximum Likelihood (ML) Method:

This method approximates the Angle of Arrival from a given set of Phase data by maximizing the Log Likelihood function. The Maximum Likelihood approaches return almost ideal performance. Even when the SNR and n are small, the method still achieves a good result. But unfortunately the disadvantage comes in terms of computation as this technique requires a huge amount of resources allocated to it.

B. Eigenspace Techniques:

These techniques depend on the Covariance properties of matrices. Its Eigen vectors can be divided into two separate spaces, the signal subspace and the noise subspace when the sources monitored are incoherent [4]. The vectors that represent the direction of the transmitter are orthogonal to the noise subspace. In general this method uses the direction of these vectors that are orthogonal to the noise subspace and are enclosed inside of the signal subspace. One famous algorithm that uses Eigen-space methods is The Multiple Signal Classification (MUSIC).

1) Estimate of Signal Parameters via Rotational Invariance technique (ESPRIT)

It has a powerful advantage when compared with other techniques and that it has a much reduced computation and storage characteristics. This is due to the fact that it is implemented by using two identical arrays, where it applies the geometrical properties of arrays to calculate the displacement of elements that are paired with elements to find distances and directions [4]. The only disadvantage is that it's considered to be a complex method to deal with.

2) Multiple Signal Classification Algorithm

The Multiple Signal Classification (MUSIC) algorithm aims to give estimates about arbitrary locations and directions. It determines mainly the number of signals, the angle of arrival (AOA) and the strength of the noise / interference [5]. It is used for multiple antennas transmitting over various frequencies. MUSIC algorithm operates as follows:

1. collects the transmitted data
2. constructs the matrix S from the resulting data and calculate its Eigen structure
3. determines the number of signals
4. introduces the polarization effect
5. chooses the maximum value of the resulting array

IV. SIMULATIONS

The simulation are results of a system implemented on LabVIEW to test the correctness of the MUSIC algorithm when implemented in real environment. The system was composed of 4 Antennas, displaced by half the wavelength of the operating frequency, and connected to 1 USRP each [6]. Each USRP is feeding the LabVIEW code with the received signals to form a matrix on which the MUSIC algorithm will work to detect and locate a special interferer which is transmitting independently.

The system composition and displacement can be seen in the Figure 3.
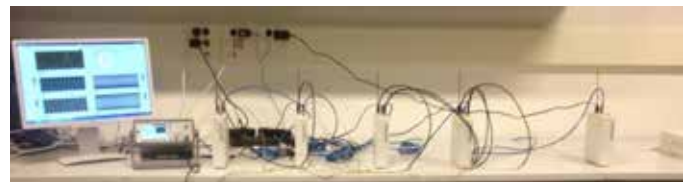
Figure 3: MUSIC Algorithm testing on Labview

After detecting an interference caused by an illegal repeater, MUSIC Algorithm is used to find the angle of arrival of the interfering signal. As shown in the Figure 4, the angle corresponds to the peak power.
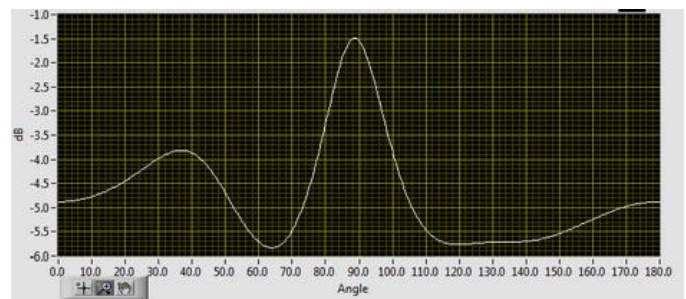
Figure 4: AOA using MUSIC Algorithm

V. Novel Antenna switch

A. Motivation

At first, the setup consisted of 4 Antennas each connected to a USRP reading and inputting to the LabVIEW to construct the matrix and run the MUSIC algorithm on it. Therefore, we had to use 4 USRPs to form the antenna array, 1 USRP for the synchronization and 1USRP to emulate the transmission of an illegal repeater. This novel design and use of an antenna switch allows to use 1 USRP only, thus eliminating the need for synchronization. The receiving switch is single pole 4 throw (SP4T) because 4 antennas are being switched to one output. Hence, reducing the cost and the complexity of the system.

B. Switch Setup

The Antenna Switch is connected to the 4 receiving antennas from one side to the USRP on the other. It is controlled by an Arduino that has analog output pins providing +5V and 0V [7]. PIN Diodes placed in a series and shunt manner were used to realize the RF switch. PIN diodes were chosen as the principle switching elements because an intrinsic (I) layer with very high resistivity material is present between positively (P) charged and negatively (N) charged material [8]. When an Arduino pin provides a positive voltage to the diode, the series diode is forward biased and acts as a resistor, and the shunt diode is reversed biased acting as a capacitor [8], [9]. The opposite will occur when the attached Arduino pin delivers negative voltage; the shunt diode is forward biased and the series diode is reverse biased [9]. The isolation between RF ports is determined by the series diode capacitance when reverse biased and resistance of shunt diode when forward biased, and the insertion loss is determined by the series resistance of series diode when forward biased and capacitance of shunt diode when reverse biased [9]. To have an isolation greater than -15dB and insertion loss less than -1dB, a PIN diode with minimal capacitance and resistance. Therefore, SMP1320-011LF diode was chosen because it has a typical capacitance of 0.23pF and resistance of 0.75Ohms [10].

C. Switch Simulation

Using Agilent ADS, the switch was realized on schematic in Figure 5, and S-parameter simulations were conducted. The switch provides a required high isolation of approximately -23 dB between ports 1 and 2 when port 1 is reverse biased and port 2 is forward biased. In addition, the circuit minimizes output return loss reaching almost -80dB at 1.4GHz.

In summary, a single USRP will be able to read one signal at a time thus mimicking the presence of several synchronized USRPs. The received signals, after being forwarded to the USRP, will apply MUSIC algorithm to give the sources' azimuth angles of arrival on the antennas. Finally, using coverage maps from operators, the illegal sites will be differentiated from the legal ones.
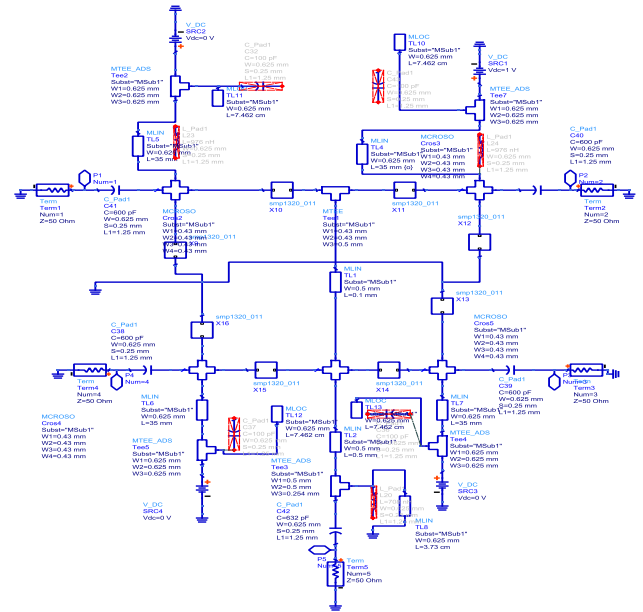
Figure 5: Antenna Switch schematic

Conclusion

In conclusion, building an antenna array using a controllable antenna switch and USRP to apply MUSIC algorithm will give a good solution for a serious problem affecting the way people communicate negatively. The degradation in the quality of services offered by network operators causes tremendous losses on subscribers and operators themselves. However, with our solution that covers all cellular networks protocols and uses the flexible NI-USRP and LabVIEW will ensure the fair utilization of the spectrum by ensuring that no illegal transmitters are operating.

References

[1] Ericsson, "LTE An Introduction," 2011. [Online]. Available: http://www.ericsson.com/res/docs/2011/lte_an_introduction.pdf. [Accessed 17 October 2015].

[2] Globe, "About Illegal Repeaters," Globe, [Online]. Available: https://www.globe.com.ph/help/about-illegal-repeaters. [Accessed 17 October 2015].

[3] B. Mark, "Illegal mobile phone signal boosters causing problems for other network users," ABC News, 9 March 2015. [Online]. Available: http://www.abc.net.au/news/2015-03-07/mobile-repeaters-disrupting-mobile-phone-signal/6287256. [Accessed 25 October 2015].

[4]   Jui-Chun Chen, "AVirtual RSNS Direction Finding Antenna System", Noval Postgraduate School, December 2004.

[5]   Schmidt, R.O., "Multiple emitter location and signal parameter estimation," in Antennas and Propagation, IEEE Transactions on , vol.34, no.3, pp.276-280, Mar 1986.

[6]   M. Willerton, "Direction Finding," National Instruments, [Online]. Available: http://sine.ni.com/cs/app/doc/p/id/cs-15016. [Accessed 26 November 2015].

[7]   Arduino, "Arduino," Arduino, [Online]. Available: https://www.arduino.cc/. [Accessed 23 February 2016].

[8]   Agilent Technologies, "Understanding RF/Microwave Solid State Switches and their Applications," [Online]. Available: www.agilent.com. [Accessed 30 January 2016].

[9]   Skyworks Inc, "RF Diode Design Guide," [Online]. Available: www.skyworksinc.com. [Accessed 15 Februrary 2016].

[10]   Skyworks, "SMP1320_Series," 7 August 2015. [Online]. Available: http://www.skyworksinc.com/uploads/documents/SMP1320_Series_200 047Q.pdf. [Accessed 7 Februrary 2016].

Energy Efficient Sensor Design with RF Wake-Up

Farah Issa, Hala Khodr, Mou'men Mohsen
Electrical and Computer Engineering, Faculty of Engineering and Architecture
American University of Beirut
Beirut, Lebanon
fri00@mail.aub.edu, hjk04@mail.aub.edu, mmm95@mail.aub.edu

Abstract- **The great interest in the Internet of Things (IoT) initiative has given life to many different applications where small mobile computing units are embedded in objects that surround us and which we interact with on a daily basis. These IoT devices are standalone devices that need to have a long lifetime. Consequently, using energy-efficient communication becomes crucial to insure the self-sustainability of these devices since keeping the sensor always on will decrease dramatically its lifetime. Using RF-triggered wake up is one of the efficient methods to reduce the sensor energy consumed by leaving the sensor asleep in a very low power state and only activating it when needed.**

Index Terms—**Internet of Things (IoT), OOK (On-Off Keying), dB (decibels), TI (Texas Instruments), RTD (Resistance Temperature Detectors), mAh (milliAmpere Hour)**

I. INTRODUCTION

THE electronic IoT movement is set to facilitate transferring data over a network with less human intervention to create a seamless data logging process. And some of the pioneer projects in this domain were energy harvesting and radio triggering which created a shift in battery usage of electronic equipment for applications that need higher power over the life of the device [1]. The main goal of the project is to develop prototype of a sensor device based on RF-triggering to locate the source of a fire. So, the aim is to have only one main sensing unit that is frequently retrieving data from the surrounding environment and can thus assess whether there is a need to wake up the other units to check whether an incident has occurred. This can be related to our target application of detecting wildfires where the main unit would continuously sense for conditions that may cause a fire incident. If these conditions were met, the other units in sleep mode are triggered to test for fire, and send the location of this fire if it were to occur before entering the inactive state again.

This method of conditional triggering was done in other projects via close range technologies using optical or ultrasound hardware that requires hardwiring to build. However, this application brings the application to a wider range of operation using WiFi signals as a standard of communication. This allows for a big development in the possible applications that could incorporate long range wireless communication without having both ends stay active to communicate.

The prototype designed will be put under power consumption supervision to prove that it can offer a solution for the Internet of Things problem of achieving more efficient systems that aim to reach self-sufficiency. Also, future work includes integrating different modules that will allow for different methods of sensing for applications beyond the scope of the work currently done on fast response to fires.

The design of the system can be described by a central sensor, always powered and connected to a base station for monitoring, connected wirelessly to a number of sensors. At the other sensors' side, the first interface with the environment is an antenna that if detects the correct signal transmitted, allows for the rectifying circuit to turn on the other circuit in which the sensing activity takes place followed by the required control of the obtained signals to send back a signal to the base unit with the necessary reporting when it comes to the data under study. This design can be better illustrated by Fig.1 which shows how the different sensing units communicate together in the set environment.
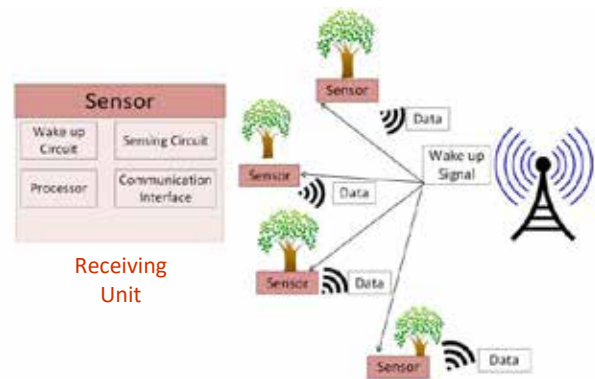
Figure 1: Schematic illustrating the scope of the project

The objective of the paper is to demonstrate the different stages of the design process of the receiving units and what goes into the integration of different sub-processes where each takes its own implementation process to later be inserted between other finished system units. In what follows, section II will discuss the different design elements and decisions that were taken in order to satisfy the functionality constraints to be listed alongside the specifications needed for the system to function as desired. Section III will describe the testing procedures and methodology used in the different stages for complete integration.
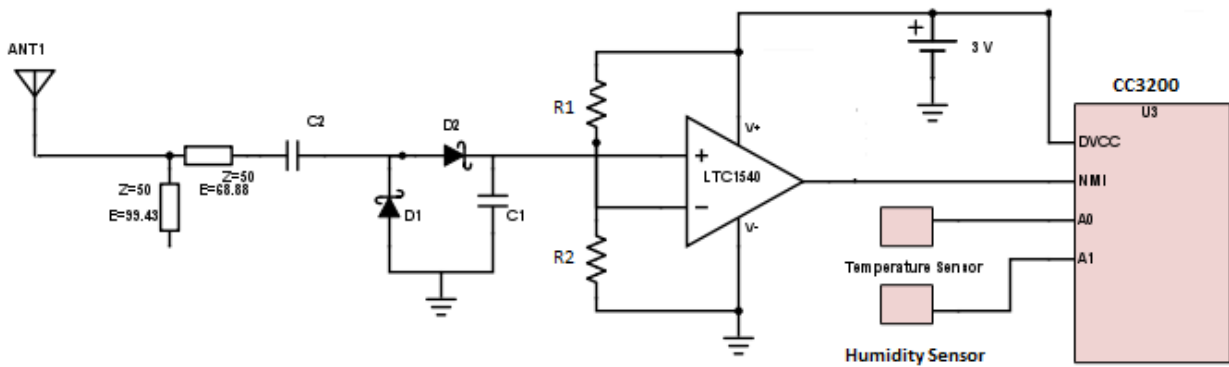
Figure 2- Scheme illustrating overall design layout

II. DESIGN

The design of the sensor with RF wake-up functionality can be divided, as seen in the above section, into three main categories: Wake-Up circuit, sensing function circuit, and communications interface for data uploading. In each subcategory, there exists different design alternatives that can be categorized as high or low level where the high level choices contributing in shaping up the scheme of the project. This is furthermore supported by a set of specifications desired from the system and to which the different design choices will refer. Figure 2 shows a general view of the system elements and how they integrate as part of the system.

A. Specifications

The system will have its specifications set on an array of technical and non-technical constraints it must verify to output the desired functionality.

1) Power Management: Saving power by using RF triggering can be considered to be the main feature of the system's design. But since some components found in the circuitry of the receiving unit have to be always turned on, this implies very little power consumption in the order of microwatts which creates a constraint for the design of the power supply of the standalone receiving units or sensors where if a battery were to be used, its capacity has to accommodate for this regular small supply until the next wake up cycle. This has to take into consideration as well the time the unit can remain in its idle state without the need to change the included batteries. So a desired consumption of current in the order of less than 0.1 Amperes is desired in the active state, and in the order of microamps in the idle state.

2) Range: The system design involves wireless communication between the sensors and the base unit which creates a constraint for the maximum distance each sensor can be put away from the transmitting end to keep normal functioning. It is desired to have a distance much larger than a couple of meters for the nature of the project and the desired outcome to have a system that can detect the source of emergency in a big area where this detection will make a difference from sensing from a single unit. But even though a larger distance is desired, the project is bound by the wireless standard to be used in addition to other constraints like the size of the antenna. So for the sake of the application at hand, a minimum distance of 10 meters[1] is required between the receiving and base units.

3) Latency: The preliminary design outline includes transmitting the signals at a standard commercial frequency. And on that note, we would have a lot of interference between the signals with the same frequency which gives rise to the need to find a method to distinguish the desired signals from the others already being transmitted in the location of the system's implementation.

Based on the given constraints, the receiving unit's specifications list includes the following: frequency transmission and reception at 2.4 GHz IEEE 802.11 b/g standard which will allow to achieve the desired range of more than 10 meters, a power supply rated at 0.1 Amperes and 3 Volts DC input provided by standard AA or AAA batteries, two external antenna with to have both reception and transmission modes, and a maximum temperature tolerance of 80 degrees Celsius as a simulation for proof of concept.

B. Design Elements and Choice

The system can be considered as an integration of three main elements: a wake-up circuit that will receive the triggering wireless signal, the processing unit with the sensing elements, and the communication interface that will send back any information that needs to be reported back to the base unit.

1) Wake-Up Circuit: The wake up circuit functionality is to detect the wake-up signal and accordingly activate the processor to accomplish its function. As illustrated in the Figure 2, the circuit is made up of: an antenna to pick up the RF power an impedance matching network to guarantee a maximum power transfer in the system, and a rectifier to convert the RF power to DC voltage.
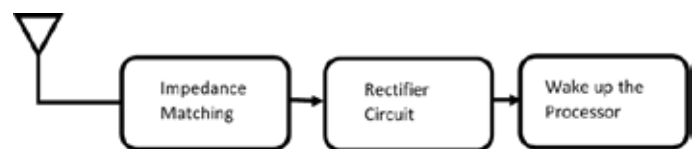
Figure 3: Schematic illustrating Wake-up Circuit Layout

Antenna: For the antenna, the key specification is that it should be able to detect the frequency of interest. Also, it

should be of small size and have good radiation characteristics. It is crucial to note that the size is inversely proportional to the frequency [2]. The antenna implemented in the circuit and is at the front of all interactions with the receiving unit has the following specifications: Frequency range between 2.4-2.5 GHz, Gain of 5 dBi, SMA type connector, Impedance of 50 Ohms, Omni-type radiation, and a length of 163 millimeters.

Rectifier: The rectifier circuit is the essential part of the wake up circuit which generates the DC output voltage to wake up the processor. A basic design for the rectifier is to use 2 diodes and 2 capacitors. This configuration is a charge-pump structure. In the negative cycle, the current flows in diode D1 and charge the capacitor C1. Whereas, in the positive cycle, the current flows in diode D2 while diode D1 is off; therefore, the output voltage will be equal to twice the input voltage minus the turn-on voltage of the diode. It is advisable in this application to use zero- bias Schottky diodes instead of a simple p-n junction diode. Schottky diodes are based in a metal-semiconductor barrier. In the zero-bias types, this barrier is very low, consequently the forward voltage is very small. This is desirable since the voltage won't be lost in the diodes.

To be able to wake up the processor from deep sleep, the output voltage of the rectifier should be above a threshold voltage able to turn on the microcontroller. Using a comparator after a one stage amplifier as illustrated below steps up the dc voltage to the required level. Figure 4 shows a complete scheme of the rectifier circuit. It should be noted that the comparator is an active component, leading to the need of an external power source. On the other hand, there exist comparators with nano-power consumption minimizing the mentioned drawback.
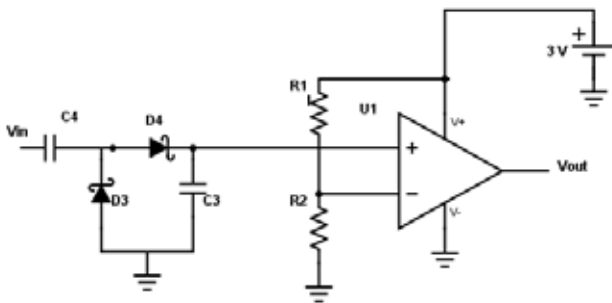
Figure 4: Schematic showing the one-stage rectifier with a comparator

Impedance Matching: The impedance matching circuit is very crucial to insure maximum power transfer. This part should match the output impedance of the antenna to the input impedance of the rectifier. Stubs are used in the matching circuit. Stubs are lengths of transmission line that can be either open or close circuits. The stub length can be varied from zero to half a wavelength making it more convenient to tune it to any value of reactance that satisfy the matching property. This can be troublesome when using inductors and capacitors since we will be restricted to standard values of L and C.

2) Sensing Circuit: This circuit follows the wake-up circuit whose output is a DC voltage that could trigger on a controller that in its turn will have sensing inputs and data outputs to oth-

er sources. So, the two main components can be considered to be the processor or controller and the different sensors linked to it.

Processor: The main requirements needed from the processing unit is to have a low voltage input supply range to accommodate for the low voltage output from the wake-up circuit, in addition to low power consumption during idle and active states. In addition, the processing unit must have enough ports to allow for the sensors to interact with minimal latency. The Texas Instrument-made CC3200 Launchpad which includes a wireless interface in addition to a microcontroller (MSP430) with an antenna embedded inside the board. This board have 1.8-3.6 V input voltage range, 400 nanoamperes power consumption in the idle state, 200 microamperes in the active state, 16 KB of flash memory to store the instructions, and a response time to wake up of less than 5 microseconds [4].

Sensors: For the sensing of the undesired conditions the system is made to notify us about these conditions, we will need a set of sensors that will allow us to account for the change of conditions in the environment. The conditions that will stay under control are conditions of temperature and humidity.

For temperature control, RTD sensors will be used for the added linearity in operation as simpler operation is desired for the purpose of our system [3]. Even though RTDs do not have the largest range of operating temperatures, but it will suffice for the projected high temperature of 80 degrees Celsius described before. This sensor of choice is DHT-11 as it outputs digital data instead of analog which is recommended for the CC3200 functioning. This sensor also includes humidity sensing giving percentage of vapor found in the air. These data can be retrieved from the output pin of the sensor and interfaced on the proprietary software of the CC3200 called Energia. The 5V input voltage for the sensor is provided from the board itself.

3) Communication Interface: The function of the communication interface is to upload the data collected by the sensor to the appropriate destination (base station in this case). A low supply voltage and a low power consumption are essential in choosing between different modules. CC3200 offers the Wi-Fi interface that will hook up to the microcontroller and at the same time is able to communicate with other units via a host server established for the system and application at hand [5]. Wi-Fi as a standard offers the best wireless communication method to try and extend the range as much as possible to meet the constraints established at the beginning.

III. IMPLEMENTATION AND TESTING

After determining the different elements that will go into the design in Section II, the integration of the different components in the sub-processes and testing of the sub-processes will be studied and discussed in aims of getting to the required output and results. This section will include simulation results in addition to preliminary testing output with different hardware and software tools.

A. Simulations

All software simulations were done on ADS and using PSPICE models.

One Stage Rectifier: The zero bias Schottky diode used is chosen to be the SMS 7630 which has a typical forward voltage of 260 to 320 mV. To use this diode in the simulation we imported the PSPICE model to the software. The PSPICE parameters can be found in the datasheet [6]. The results when sweeping the received power in dBm versus the output DC voltage are as seen in Figure 5 which also includes experimental results to be discussed later. We notice that the output voltage is increasing as the power received increases with an efficiency of 13.62% at -20 dBm power received.
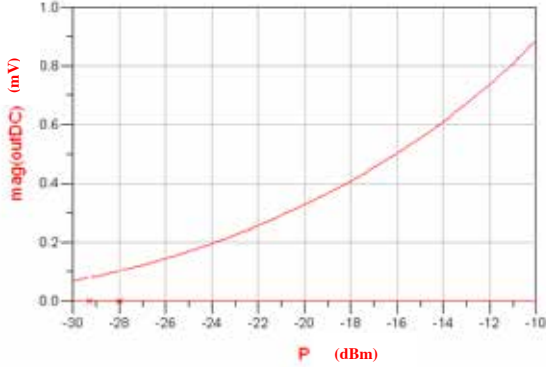
Figure 5: Graph showing output DC versus power received

Comparator: By referring to the comparator circuitry in Section 2, the minimum threshold to turn on the comparator is set to 50 mV which is a reasonable number below which the noise is significant. To get the threshold voltage, a voltage divider formed by R1 and R2 is done to get the value for the reference voltage equal to 50 mV for an R1 of 1 MegaOhms, an R2 of 47 KiloOhms, and Vcc of 3V. High resistance was used to minimize power consumption. Sweeping the received power in dBm versus the overall output voltage, the results are shown in Figure 6 with a minimum received power that should be equal to -28 dBm. Power during the ON state is measured at 29.11 microwatts and during the OFF state at 1.1 microwatts.
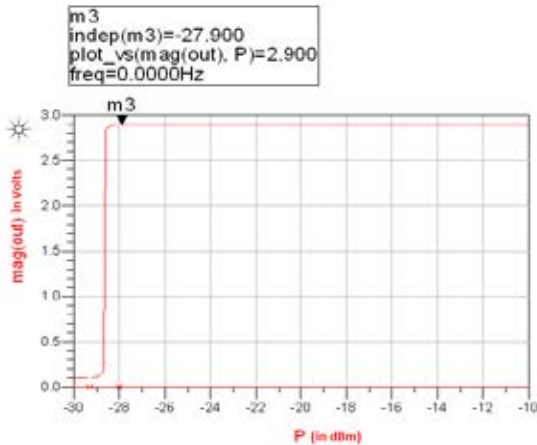
Figure 6: Graph showing output DC versus power received

Distance: The relation between the power received and distance shown in (1) allows for the calculation of the range achievable given the specifications and plotted in Figure 7 which shows a distance of 22 meters achieved at a power of -27.9 dBm. This meets the range constraint.

$$Pr = \frac{Pt\, Gs\, Gr\, \lambda^2}{(4\pi.D)^2} \qquad (1)$$

Pt is the transmission power, Gs is the transmission antenna gain and Gr is the receiving antenna gain, lambda is the wavelength of the EM wave, and D is the distance between the transmitting and receiving node. The wavelength is equal to the ratio of c to the frequency which gives 0.125 meters, and Gs and Gr have values of 3.162 corresponding to 5 dBi.
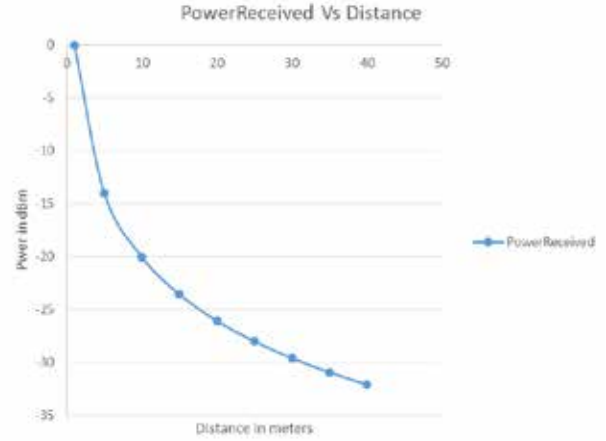
Figure 7: Graph showing power received versus distance

Power Consumption: The lifetime of the system during its active and non-active cycles is a big concern when it comes to the discussed constraints. To measure this lifetime a relation between the consumption and the size of the battery included gives out the results as in (2) with the inclusion of a multiplication factor of **0.7** accounts for the external factors that would give worse numbers for the estimated time [7].

$$Estimated\ Time = \frac{Battery\ Capacity\ (mAh)}{Device\ Consumption\ (mA)} * 0.7 \qquad (2)$$

Having a battery size of 2500 mAh and assuming that the system turns on for a total of 30 minutes per day, the total endurance of the system was found to be for more than 6 months before a battery replacement is needed. And that is mainly due to the efficiency of the microcontroller during idle state.

Communication Interface: To establish the connection between the receiving unit and the base unit. Energia was used as a software to program the CC3200 launchpad. In order to send the data to the cloud, dweet.io was used as a server. Changes in the data can be viewed through a platform called Freeboard.io.

Sensing elements: the temperature and humidity sensors were input to the processor with preceding resistive circuits. Then, the sensors are calibrated to replicate the desired levels. The temperature sensor was calibrated to output an alarm indicator at a temperature above 80 degrees Celsius. And for the humidity level, DHT-11 sensor a predefined level of vapor that simulates the presence of a fire. So, that threshold was used in order to have sensing elements able to detect conditions of fire.

B. Experimental Results

After the fabrication of the different circuits, the units were independently tested in the controlled environment which includes a signal generator to obtain the desired frequency of signals and test how the various systems react to different signals to get indications of performance and efficiency. The different circuits under the experiments can be found in Figure 8 which also shows the sequence in which the different units must be inserted to achieve the system functionality.

Figure 8: Image showing the circuit in the testing process

Starting with the testing of the rectifier circuit, Figure 9 shows the comparison between the achieved results and those retrieved from the simulations. It can be seen that both results do not stray away majorly with results around the lower end of power received being very close, which is the region where data matters the most for our application.
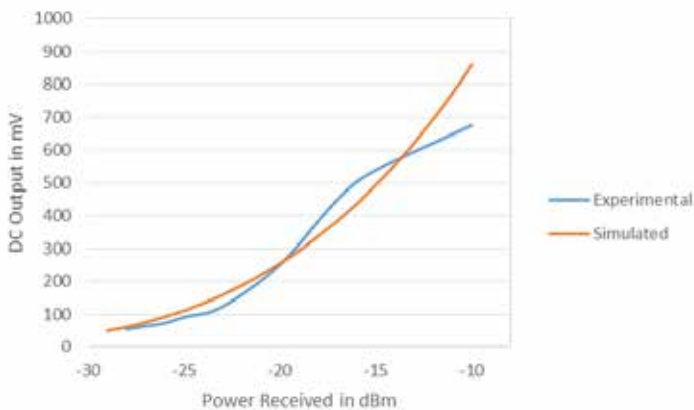
Figure 9: Graph showing DC Output versus Power Received

Figure 10 illustrates the reflection coefficient of the simulated and tested results. A peak of almost -20 dBm is seen at a frequency of 2.4 GHz which indicates that the antenna radiates best at this frequency which matches the requirements.
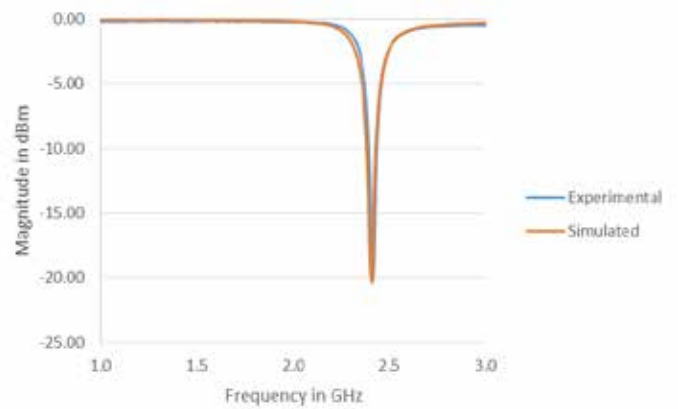
Figure 10: Graph showing reflection coefficient for the system

As for the comparator results, Figure 11 demonstrates how the voltage level varies after the comparator stage to reach the desired voltage level illustrated by the experiment by about 2.9 Volts
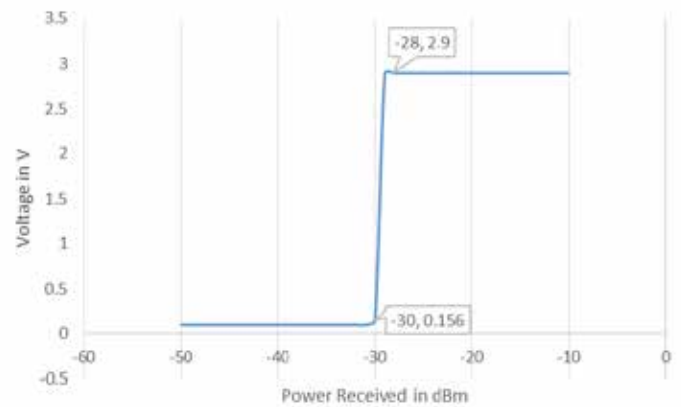
Figure 11: Graph showing DC output after comparator

C. Communication Interface:

The below figure shows the interface where the temperature and humidity sensing data are uploaded to the cloud.

Figure 12: Dashboard showing data transmitted from the sensor to the cloud

IV. Conclusion

In this paper, a method of wireless triggering was studied and proposed as a solution within the IoT movement framework. Wireless triggering has yet to show any results in any field other than that of close quarter (less than one meter). The system has entered the testing phase with undergoing processes of fabrication and hardware testing. Early test trials show promising results with output being very near to projected desired data. The next step includes the continuation of the testing process that will be carried over all the processes especially when it comes to the integration of the different units and creating a user friendly medium of monitoring the different tested data in real time or to check previous logs. After all functionalities are present in the system, additional functionalities can be studied like addressing of the wireless signal in order to avoid any false alarms and differentiate the desired triggering signal from interfering ones in the same medium in addition to other potential applications.

References

[1] H. Ostaffe, "Sensors Online," 13 October 2009. [Online]. Available: http://www.sensorsmag.com/sensors-mag/rf-energy-harvesting-enables-wireless-sensor-networks-6175. [Accessed 28 Novermber 2015]

[2] J. Wang, Q. Gao, Y. Yu, H. Wang and M. Ji, 'Radio-Triggered Power Management in Wireless Sensor Networks', *Wireless Sensor Networks - Technology and Applications*, 2012.

[3] Ultra Electronics, 'RTD vs. Thermocouple comparison chart « Ultra Electronics Nuclear Sensors & Process Instrumentation (NSPI)', *Ultra-nspi.com*. [Online]. Available: http://ultra-nspi.com/information-central/rtd-termocouple-comparison/. [Accessed: 25- Oct- 2015].

[4] Texas Instrument. (2013, May). "MIXED SIGNAL MICROCONTROLLER" [Online]. Available: http://www.ti.com/lit/ds/symlink/msp430g2553.pdf [Aug 12, 2015]

[5] Texas Instrument. (2015). "CC2500" [Online]. Available: http://www.ti.com/lit/ds/swrs040c/swrs040c.pdf [Aug 12, 2015]

[6] SKYWORKS. (2015, May 7). "SMS7630" [Online]. Available: http://www.skyworksinc.com/uploads/documents/SMS7630_061_201295H.pdf [Nov 12, 2015]

[7] Digi-Key Electronics," [Online]. Available: http://www.digikey.com/en/resources/conversion-calculators/conversion-calculator-battery-life. [Accessed 22 Nov 2015].

Gravity: A Smart Façade Cleaning Robot

Adnan Arnaout, Ghassan Oueidat, Ibrahim Tall, Kassem Dabaje, Mohamad Husseini

Electrical Engineering and Mechanical Engineering Department

American University of Beirut

Beirut, Lebanon

aja28@mail.aub.edu, gio00@mail.aub.edu, ikt00@mail.aub.edu, kod00@mail.aub.edu, mfh23@mail.aub.edu

Abstract- **Most façade cleaning processes are performed by humans hanging off ropes. Since safety problems can arise from this traditional cleaning method, researchers have been triggered to come up with new automated solutions regarding skyscrapers cleaning. However, the proposed solutions deal only with continuous glass building facades with no ability to cross edges, if present. In addition, the current window cleaning robots are not fully autonomous and need human intervention. In order to solve these problems, a new façade cleaning robot, "Gravity", has been developed to enhance cleaning efficiency of facades regardless of their external structures. Gravity is an articulated wall climber and window cleaning robot. It is composed of two separate bodies connected together by a two-link actuated arm that allows it to cross edges. It adheres to vertical surfaces using vacuum-based active suction cups regulated to allow motion while ensuring sufficient amount of adhesion. The cleaning system is composed of a two-stage cleaning process. Gravity has a track-based locomotion mechanism which allows a better edge crossing and grip on smooth walls and glass surfaces. Gravity has two modes of operation, Remote Control and Autonomous.**

Keywords: Window cleaning, vacuum cups, vacuum pump, tracks, cleaning system, articulated arm, safety, control system.

I. INTRODUCTION

Cleaning skyscrapers' window panels is a dangerous task performed by human labors. In order to ensure an adequate and clean appearance of the skyscrapers and to lower the risk of injuries and fatalities of window cleaners, robot intervention is needed. In fact, a study conducted in 2011 by the Department of Labor of the United States indicates that every ten seconds, on average, a high-rise window cleaner dies [8]. For this reason, researchers have been motivated to develop window cleaning robots. For instance, in paper [1], researchers have developed a window cleaning robot that consists of a vacuum pump, sealing mechanism and driving mechanism, with only one suction cup. It is also formed of a single body, limiting its application to only continuous glass buildings. Moreover, Qian et al. [2, 3] has developed a robot with a locomotion system utilizing a crane installed on the roof using non-actuated dual suction cups. In addition, Elkmann et al. [4-6] have developed the SIRIUS_c which uses a crane-based system installed on the roof. Also, it has an advanced sliding frame mechanism which can perform vertical and horizontal motion. Cranes can be difficult to install

on many high-rise buildings due to their external structure. Furthermore, Moon et al. [7] have developed a window cleaning robot similar to those mentioned previously.

To avoid such complications, "Gravity" has been developed. This robot is intended to help save people's lives and to guarantee better results of cleaning. Gravity is a fully automated robot capable of cleaning walls and glass surfaces without the need of cranes and humans. It is able to move on the window on its own, once placed on it, ensuring the cleanliness of the whole window. The robot starts from an initial position placed by the user and returns back to the same spot at the end of the cleaning process. After cleaning the window panel, the robot will move to the next window by climbing edges between glass panels of a certain thickness. In addition, the robot has a Remote Control option to clean surfaces that cannot be covered by the autonomous mode. A main feature distinguishing Gravity from other window cleaning robots is its innovative design. It is composed of two bodies and a link having multiple degrees of freedom. This design makes Gravity flexible and allows it to clean hard-to-reach spots, unlike other façade cleaning robots.

In this paper, the three main robot mechanisms: adhesion, locomotion, and cleaning are heavily discussed. Afterwards, a section is dedicated for the safety and edge crossing mechanisms. In addition, the motion control of the robot and the simulation results of the Autonomous mode are discussed. Moreover, the results and experimental conclusions are illustrated. Finally, a conclusion is reached, including potential future directions of Gravity.

II. ADHESION SYSTEM

A. Mechanism Overview

The adhesion system consists of a vacuum pump connected to suction cups, generating vacuum in them. The suction cups can slide on glass and on the edge surface. Since the robot consists of two bodies, one body can lift the whole robot by cutting off the vacuum from it and keeping the vacuum flowing in the second body. Two on/off valves are connected to the robot, thus allowing the control of the vacuum flow in the two bodies. The connections between the vacuum generator, separator, valves, and vacuum cups are illustrated in Figure 1.A that shows the entire robot adhesion mechanism.

B. Vacuum Pump

Vacuum pumps are characterized by the maximum pressure they can reach and the volume flow rate they provide. In order to find the proper vacuum pump, the free body diagram of the system is obtained with the help of the literature [7]. Figure 1.B illustrates the free body diagram with the forces acting on the

body. The equations governing the amount of pressure needed for adhesion are given as follows:

For the x-axis and y-axis respectively:

$$m_{robot}g - F_{friction} = 0 \tag{1}$$

$$F_{normal} - F_{vacuum} - m_{robot}g = m_{robot}a \tag{2}$$

Substituting with the different force relations, the literature [3] introduces the following equations:

$$\Delta p \geq \frac{m(g+a)}{(\mu_{tracks}\varepsilon A_{suction} - \mu_{sealing\ loop}\varepsilon A_{sealing\ loop})} \tag{3}$$

This equation results in the minimum pressure difference between the atmosphere and the vacuum cups for the robot to be able to move on glass. In equation (3), μ_{tracks} represents the friction coefficient of the tracks; ε is the ratio of the pressure of the sucker to that acting on the wheels having a value ranging from 0.8 – 0.9; m is the mass of the robot; g is the gravitational acceleration; a is the acceleration of the robot; $\mu_{sealing\text{-}loop}$ is the friction coefficient of the sealing loop; $A_{sealing\text{-}loop}$ is the area of the sealing loop of the vacuum cups. $A_{suction}$ is the area of the inner lip of the vacuum cup.

C. Vacuum Cups

The vacuum cups are chosen to have sliding ability, allowing a smooth and frictionless motion of the robot and avoiding unnecessary stresses that might arise from having non-sliding suction cups. The vacuum system of Gravity consists of twelve vacuum cups, each body containing six. The diameter of one vacuum cup is 78 mm and the placement of the vacuum cups is shown in Figure 1.C. The vacuum cups are placed such as to ensure that the suction force is at the center of mass of the body. Knowing $A_{suction}$, $A_{sealing}$, the mass of the robot, and the acceleration, the difference in pressure can be calculated using the following equation:

$$\Delta p = \frac{m(g+a)}{(\mu_{tracks}\varepsilon A_{suction} - \mu_{sealing\ loop}\varepsilon A_{sealing\ loop})} \tag{4}$$
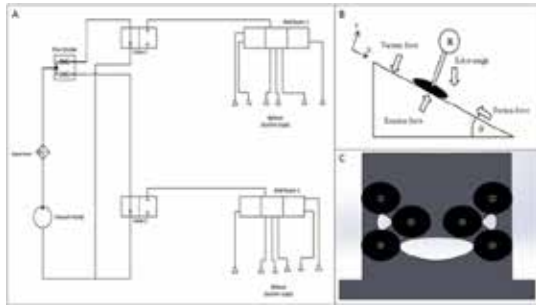
Figure 1. Adhesion System of Gravity. A) Vacuum system design consisting of a vacuum generator, two valves, and twelve vacuum cups, six for each body. B) Free body diagram of the robot. C) Placement of vacuum cups on the two bodies of the robot.

III. Locomotion System

The locomotion method used is track-based. Using tracks increases the contact area between the robot and the surface, thus providing a smoother motion. The main problem that should be tackled is that the motors driving the tracks should be able to overcome the adhesion force created by the vacuum pump and gravity. In order for the robot to move on vertical surfaces, equations found in paper [3] are used. First, the driving force is found by calculating the required ΔP which is the pressure needed to adhere to the surface. A lower and upper bound for the driving force is selected, and the final value is the average of these two:

$$F_{lower} = m(g+a) + \mu_s \Delta PA_{seal} \tag{5}$$

$$F_{upper} = \mu_t \Delta PA_{suction}\varepsilon \tag{6}$$

where m is the mass of the robot, g is the gravity (9.81 m/s^2), a is the acceleration, μ_s and μ_t are the friction coefficient of the suction cups and tracks, respectively. A_{seal} is the area of the suction cup that makes direct contact with the surface (Outer minus inner area), $A_{suction}$ is the suction area (inner area) and $\varepsilon = 0.85$.

The tracks are chosen to allow the robot to cross over small bumps easily. They have an inverted trapezoidal shape as shown in Figure 2.A. These tracks can allow the robot to gain a better traction with respect to edge-like obstacles. In many cases the robot will be able to cross edges without the need to uplift its two bodies since the tracks shape will allow the robot to increase the gripping force on the surface.
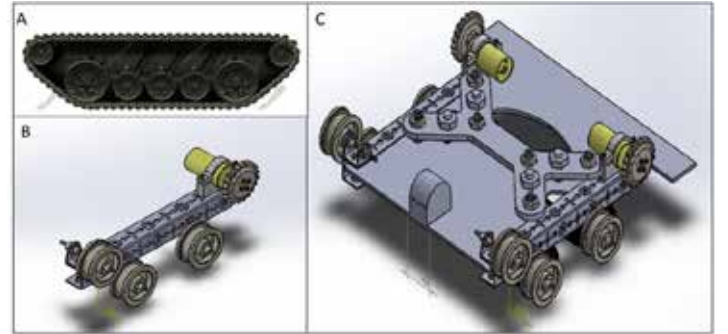
Figure 2. Locomotion System of Gravity. A) Image showing the shape of the tracks used. B) Schematic image showing the tracks design. C) Schematic image showing the design of one body of the robot.

IV. Cleaning System

The cleaning mechanism is divided into three stages: (1) water spraying and dirt removal using a high speed rotating brush, (2) dirty-water suctioning in the first body and (3) dirty-water suctioning and wiping at the end of the second body. The spraying system, consisting of three nozzles, is placed above the rotating brush at the front of the first body (see Figure 3). The nozzles are chosen to span the spacing between the two sides of the body. The head pressure is used to drive the water from a tank placed at the top of the building to the nozzles. The rotating brush is 35 cm long. One of its sides is connected to a DC motor gearbox and the other side is connected to a bearing fixture. The DC motor used for the cleaning has a rotational speed of 300 RPM and a torque of 0.6 Kg.cm. To ensure good cleaning process, the brush must be rotating at relatively high speed. For this reason, an external gear box of ratio 4.8 is used to connect

the motor to the brush's shaft. This way, the brush will be rotating at 1440 RPM; thus removing dirt easily. As of the suctioning process, two slits are used. The first one is placed at the front of the first body, behind the nozzles and the brush. The second one is placed right before the wiper in order to collect the dirty water at the end of the cleaning process. These slits are connected to another vacuum pump utilized only for water suctioning. Finally, a rubber wiper blade is placed at the end of the second body to remove any un-suctioned dirty water.
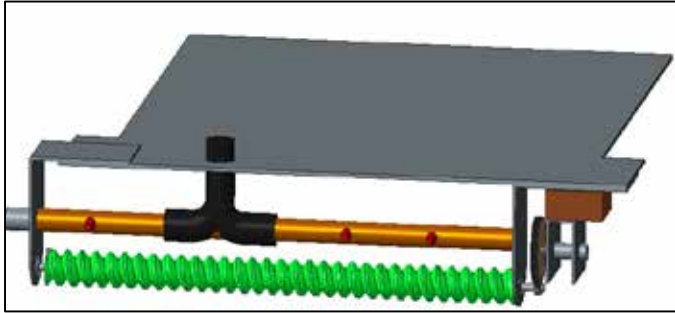
Figure 3. Cleaning System of Gravity.

V. LINK DESIGN

The link that provides the ability to overcome façade edges is shown in Figure 4. The dimensions of the link are selected to ensure the robot can withstand the bending moments and traverse, relatively, high edges. The large link is 30cm long and has a web height of 50mm, and the small link is 30cm long and has a web height of 40mm. Different materials such as carbon fiber, hardened steel and aluminum 1060 are tested while fixing two main constraints: the weight of the link and the factor of safety. Aluminum 1060 has provided the best results in terms of the weight and factor of safety. In order to have a safe design with no failure, a stress analysis of the link is conducted on SolidWorks. A 20Kg distributed mass is applied to the first free body while fixing the second body. The finite element analysis done using SolidWorks shows how the stresses are spread along the link's geometry. The analysis has given a maximum Von Mises stress of 2.213×10^8 Pa at the point of maximum stress. In order to ensure that a failure will not occur at this point, a safety analysis is conducted. The analysis has given a minimum factor of safety of 2.804 that is distributed evenly along the link's geometry which indicates that the link can withstand 2.8 times the force applied. Hence, Aluminum 1060 will be used for the link.
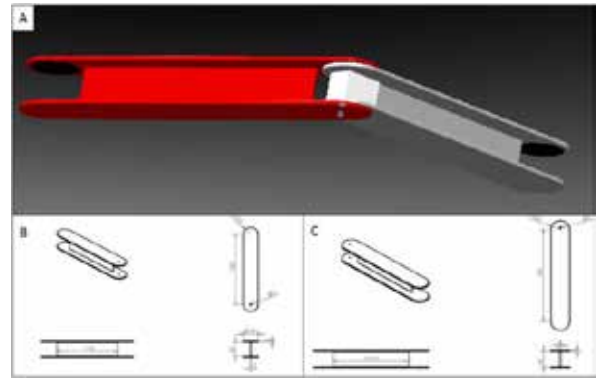
Figure 4. A) Illustration image showing the structure of the link. B) Dimensions of the large link. C) Dimension of the small link.

VI. SAFETY ROPE

In order to ensure a safe cleaning process, a safety rope will be used. It is composed of a 15m long 1.2cm diameter rope connected to a thrust bearing connection. From this connection, three smaller subdivisions of the main rope are linked to the robot; two of them are connected to the center of gravity of each body and the remaining one is connected to the link arm. The dimensions of the ropes are chosen according to the geometries of Bechtel building and the robot. Likewise, the rope material is chosen based on the forces affecting the robot in case of failure. The main rope is connected to a coiler that is placed on the roof of the building. This coiler is, at the same time, placed on a support casing with wheels that follow a linear track allowing it to stay on the same center-axis of the robot as it moves across the building. The safety rope follows the "extend and retract" mechanism used in cars' seatbelts. In case of failure, the coiler will get locked, preventing the robot from falling.

The final design of the robot is shown in Figure 5.
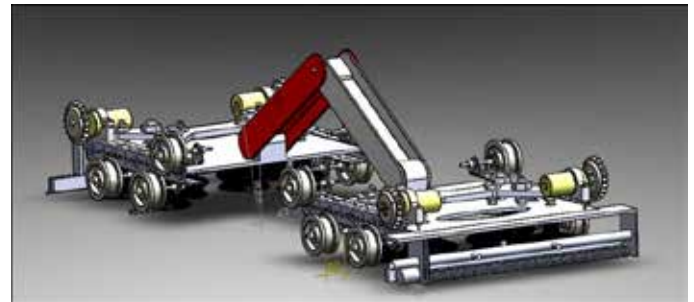
Figure 5. Final Design of Gravity.

VII. MATHEMATICAL MODELING

A. Link Modeling

The dynamics of the link are derived from the kinematics equations involving the kinetic, potential, and dissipation energies along with the torques applied to the three joints. These equations are given as follows:

$$KE = \frac{1}{2}\left\{ I_1\dot{\theta}_1^2 + m_1\left[\frac{d}{dt}(a_1 sin\theta_1)\right]^2 + m_1\left[\frac{d}{dt}(a_1 cos\theta_1)\right]^2 \right\} +$$

$$\frac{1}{2}\left\{ I_2\dot{\theta}_2^2 + m_2\left[\frac{d}{dt}(l_1 sin\theta_1 + a_2 sin\theta_2)\right]^2 + m_2\left[\frac{d}{dt}(l_1 cos\theta_1 + \right.\right.$$

$$\left. a_2 cos\theta_2)\right]^2 \right\} + \frac{1}{2}\left\{ I_3\dot{\theta}_3^2 + m_3\left[\frac{d}{dt}(l_1 sin\theta_1 + l_2 sin\theta_2 + \right.\right.$$

$$\left.\left. a_3 sin\theta_3)\right]^2 + m_3\left[\frac{d}{dt}(l_1 cos\theta_1 + l_2 cos\theta_2 + a_3 cos\theta_3)\right]^2 \right\} \quad (7)$$

$$PE = m_2 g a_1 cos\theta_1 + m_2 g(l_1 cos\theta_1 + a_2 cos\theta_2) +$$

$$m_3 g(l_1 cos\theta_1 + l_2 cos\theta_2 + a_3 cos\theta_3) \quad (8)$$

$$DE = \frac{1}{2}\sum_{i=1}^{3}(C_i(\dot{\theta}_i - \dot{\theta}_{i-1})^2) \quad (9)$$

$$T_1 = T_{m1} - T_{m2}, T_2 = T_{m2} - T_{m3}, T_3 = T_{m3}$$

where $T_{m1}, T_{m2,} T_{m3}$ are the motors torques and are expressed as follows:

$$T_{m1} = G_1 u_1 - I_{p1}\ddot{\theta}_1 - C_{p1}\dot{\theta}_1 \quad (10)$$

$$T_{m2} = G_2 u_2 - I_{p2}(\ddot{\theta}_2 - \ddot{\theta}_1) - C_{p2}(\dot{\theta}_2 - \dot{\theta}_1) \quad (11)$$

$$T_{m2} = G_3 u_3 - I_{p3}(\ddot{\theta}_3 - \ddot{\theta}_2) - C_{p3}(\dot{\theta}_3 - \dot{\theta}_2) \quad (12)$$

where u_1, u_2 and u_3 are the input voltages to the motors located at joints 1, 2 and 3 respectively.

Moreover, the mathematical model is derived using the Lagrangian equation of motion, which gives:

$$\frac{d}{dt}\left(\frac{\partial K}{\partial \dot{\theta}_i}\right) - \frac{\partial K}{\partial \theta_i} + \frac{\partial D}{\partial \dot{\theta}_i} + \frac{\partial P}{\partial \theta_i} = T_i \quad i = 1,2,3 \quad (13)$$

Rearranging the equations, the following mathematical model is obtained:

$$\widetilde{P}\begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \\ \ddot{\theta}_3 \end{bmatrix} + \widetilde{K}\begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{\theta}_3 \end{bmatrix} + \widetilde{F}\begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} + \widetilde{G}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (14)$$

where

$\widetilde{P}$

$$= \begin{bmatrix} J_1 + I_{p1} + I_{p2} & l_1(m_2 a_2 + m_3 l_2) - I_{p2} & l_1 m_3 a_3 \\ l_1(m_2 a_2 + m_3 l_2) - I_{p2} & J_2 + I_{p2} + I_{p3} & l_2 m_3 a_3 - I_{p3} \\ l_1 m_3 a_3 & l_2 m_3 a_3 - I_{p3} & J_3 + I_{p3} \end{bmatrix}$$

$$\widetilde{K} = \begin{bmatrix} C_1 + C_2 + C_{p1} + C_{p2} & -C_2 - C_{p2} & 0 \\ -C_2 - C_{p2} & C_1 + C_2 + C_{p1} + C_{p2} & -C_3 - C_{p3} \\ 0 & -C_3 - C_{p3} & C_3 + C_{p3} \end{bmatrix}$$

$$\widetilde{F} = \begin{bmatrix} [m_1 a_1 + (m_2 + m_3)l_1]g & 0 & 0 \\ 0 & [m_2 a_2 + m_3 l_2]g & 0 \\ 0 & 0 & [m_3 a_3]g \end{bmatrix}$$

$$\widetilde{G} = \begin{bmatrix} -G_1 & G_2 & 0 \\ 0 & -G_2 & G_3 \\ 0 & 0 & -G_3 \end{bmatrix}$$

The symbols used are available in Table 1 in the Appendix.

B. Motion Modeling

The motion dynamics of the robot are derived using a similar force model shown in page 4 of [9]. The model applies for the motion of the robot on walls and takes into consideration the friction between the tracked wheels and the wall. The linear and angular momentum balance can be written as

$$\sum_{i=0}^{2}(F_i - f_{2i})R(\theta)D_i - G\begin{pmatrix} 0 \\ (cos\theta)^2 \end{pmatrix} = m\ddot{P} \quad (15)$$

$$d\sum_{i=1}^{2}F_i - sgn(\dot{\theta})M_{r2} = J\ddot{\theta} \quad (16)$$

In the above equations, F_i ($i = 1, 2$) represents the magnitude of the force produced by the ith motor, $R(\theta)$ is the rotation matrix, D_i represents the direction vector of the ith tracked wheel, P is the position of the center of mass, G is the weight of the robot, m is the mass of the robot, f_{2i} is the rolling friction between tracked wheel and the wall, M_{r2} is the resistance moment produced by friction on walls, m is the mass of the robot, and J is its moment of inertia. Following the derivations shown in [9], the expression for the rolling friction and the resistance moments is obtained and the non-dimensional motion equations on walls are derived:

$$\begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{\theta} \end{bmatrix} + \begin{bmatrix} \dot{x} \\ \dot{y} \\ \frac{md^2}{J}\dot{\theta} \end{bmatrix} = u(\theta, t) \quad (17)$$

where, $u(\theta,t)$ is the control action of the form:

$$u(\theta, t) = P(\theta)\bar{U}(t) \quad (18)$$

$$P(\theta) = \begin{bmatrix} -sin\theta & sin\theta & sin\theta \\ cos\theta & -cos\theta & -cos\theta - \frac{G(cos\theta)^2}{2f_{21}} \\ 1 & 1 & -sgn(\dot{\theta})\frac{M_{r2}}{2df_{21}} \end{bmatrix} \quad (19)$$

$$\bar{U}(t) = \begin{bmatrix} \bar{U}_1(t) \\ \bar{U}_2(t) \\ \frac{2f_{21}}{\alpha U_{max}} \end{bmatrix} \quad (20)$$

According to [10], in equations (17)-(20), $U_1 = U_{u1}/U_{max}$ and $U_2 = U_{u2}/U_{max}$ are the control inputs for the two sets of motors, and $2f_{21}/\alpha U_{max}$ is the friction input from the walls. U_{u1} and U_{u2} are the actual voltages applied to the two sets of motors; U_{u1} is applied to the front and back left motor and U_{u2} is applied to the front and back right motor. U_{max} is the maximum voltage used for normalization purposes and α is a motor constant (N/V).

VIII. MOTION CONTROL

The robot has two modes, RC, driven by an operator, and Autonomous, once placed on its predefined initial position. The input/ output interfacing, the processing, and the acquisition are done using Arduino Mega. The Arduino mega specializes by having 54 I/O pins, 16 MHz clock, and 8-bit Atmega2560 microcontroller. The motion control of the robot under RC is done by direct mapping between the different buttons on the controller and the locomotion motors. For the Autonomous

mode, the motion control is achieved by applying Control Theory and final minor tunings in order to obtain perfect navigation.

A. RC Mode

The hardware schematic for the RC mode is shown in Figure 6 below.
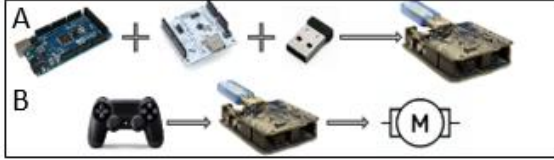
Figure 6. Hardware schematic for the RC Mode

For the RC mode, the main communication is between a Bluetooth dongle and a PS4 controller. The Bluetooth dongle is connected to the Arduino through a USB host shield in order to deliver the commands of the controller to the robot. These commands are delivered to the different actuators on the robot such as motors and pneumatic valves.

B. Autonomous Mode

In the Autonomous Mode, the robot is assumed to be positioned on the upper-left corner of the building as an initial condition. Moreover, the robot is to be faced rightwards with an angle assumed to be the $0°$ angle. Then, the process of Autonomous cleaning starts with the robot moving rightwards and regulating the yaw angle seen by the robot. The robot cleans the façade in horizontal lines that correspond to angles of $0°$ and $180°$. The robot makes its rotation from the $0°$ to the $180°$ once it reaches the end of a horizontal line and hits an edge.

The sensing considered for the Autonomous process include 10-dof IMU which has a 3-dof accelerometer, 3-dof gyroscope, 3-dof magnetometer, and 1-dof barometer. The IMU provides the heading angle, translational velocity and acceleration, rotational velocity and acceleration, and altitude of the robot. For sensing the end of the façade, IR distance sensors and bump sensors are utilized. Moreover, a glass detection sensor is employed for the cases where no edge exists at the end of the façade.

To control the motion of the robot, the state space model is derived. The derivation is obtained from Equation 17 by using Taylor series approximation and the employment of the Jacobian method for the linearization of the system. Moreover, small angle approximation is considered for the system.

$$\begin{bmatrix} \ddot{x} \\ \dot{y} \\ \ddot{y} \\ \dot{\theta} \\ \ddot{\theta} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \frac{md^2}{J} \end{bmatrix} \begin{bmatrix} \dot{x} \\ y \\ \dot{y} \\ \theta \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} -1 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & -1 & -1 - \frac{G}{2f_{21}} \\ 0 & 0 & 0 \\ 1 & 1 & -\frac{M_r}{2df_{21}} \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ f_f \end{bmatrix}$$

$$y = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{x} \\ y \\ \dot{y} \\ \theta \\ \dot{\theta} \end{bmatrix}$$

The above model is of the form

$$\dot{X} = AX + Bu$$
$$y = CX \tag{21}$$

where $X = \begin{bmatrix} \dot{x} \\ y \\ \dot{y} \\ \theta \\ \dot{\theta} \end{bmatrix}$ is the state vector of the system.

Assuming the façade is an x-y coordinate frame, $\dot{x}$ is the velocity along the x-direction, y is the height, $\dot{y}$ is the velocity along the y-direction, θ is the heading angle of the robot with reference to the horizontal, and $\dot{\theta}$ is the angular velocity of the robot.

The controllability and observability of the state space model are studied and the system as a whole is fully state observable but not fully state controllable. The lack of controllability of the system is due to the singularities that the system suffers at $\theta = 0°$ and $\theta = 90°$ when the robot is moving vertically and horizontally, respectively.

As there is a lack of controllability in the system, the autonomous control of the system is done by splitting the system into subsystems and using behavior-based control. The behaviors include horizontal movement where $\theta = 0°$, vertical movement where $\theta = 90°$, and the rotation behavior when moving from a horizontal line to a vertical one, and vice versa.

Two-degrees of freedom controller is designed based on the principle of full state feedback along with the principle of feed forward. Gain Scheduling is employed based on the different linear systems obtained at the different behaviors.

The settling time for the different behaviors is chosen in relation to the motors capability of delivering torque and speed. A settling time range of 3-5 seconds is considered for the behaviors.

IX. SIMULATION RESULTS

Simulations have been done to test each component alone, not as a whole system. Subsystems such as the locomotion, cleaning, and the adhesion work when tested alone.
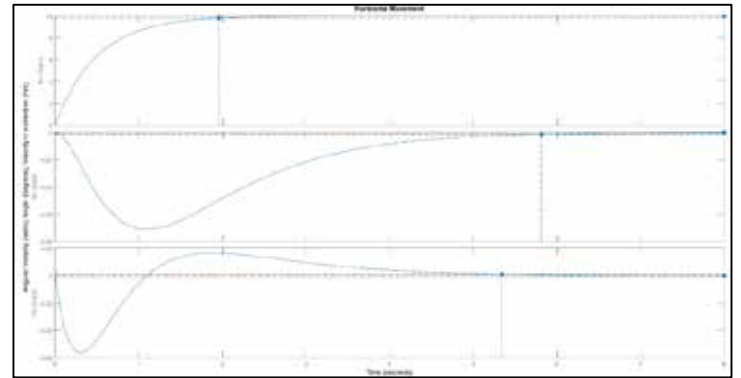
Figure 7. Autonomous Horizontal Motion

In Figure 7 three plots are shown, the first one represents the ability of the robot to track a desired velocity in the x-direction, 10 m/s in this case. The second plot represents the heading angle which is being regulated or set to zero. The third graph represents the angular velocity which is also being regulated.

In the same manner, the autonomous vertical motion for the robot is done along with the autonomous rotation as seen in Figures 10 and 11, respectively.

X. IMPLEMENTATION AND EXPERIMENTAL RESULTS

Experimental results cannot yet be obtained since the implementation is not finished. The robot is still in the manufacturing stage as shown in the figure below:

Figure 8. Gravity installed on one of Irani Oxy's windows.

Figure 9 shows a detailed schematic of the electrical circuit of the system. The circuit includes two converters, three valves, four H-bridges for the link and track motors, eight motors, an Arduino mega, and a vacuum pump. Moreover, twelve sensors are connected to Arduino. Four switching circuits have been designed in order to control the valves through Arduino. Furthermore, six low pass filters for the IR sensors are designed in order to damp the noise and have a more stable output. The total power consumption of the robot is around 907 W, a detailed power distribution can be seen in Table 2, shown in the appendix. All the electrical connections of the circuit are done through PCBs, shown in Figure 12 in the Appendix.

Figure 9. Electric Circuit of Gravity.

XI. CONCLUSION AND FUTURE DIRECTION

In this paper, Gravity, a fully autonomous façade cleaning robot is introduced. It is a one-of-a-kind robot, developed toward the commercialization for domestic use such as skyscrapers' window panes. In designing the mechanism of the Gravity, vacuum pump is used as an adhering method. Gravity's locomotion is tracked-based, allowing better grip on smooth walls and glass surfaces.

The robot has been tested on Irani Oxy Engineering Complex in the American University of Beirut which is shown in Figure

8. The robot is able to move on the glass with continuous motion while the brush is rotating and the vacuum pumps are turned on. Therefore, the project is deemed successful.

As a future work, more advanced control theories will be applied to achieve optimum control. In addition, computer vision and artificial intelligence will be implemented to improve Gravity's performance.

REFERENCES

[1] Y. Zhao, Z. Fu, Q. Cao, and Y. Wang, "Development and applications of wall-climbing robots with a single suction cup," *Robotica,* vol. 22, pp. 643-648, 2004.

[2] Z.-Y. Qian, Y.-Z. Zhao, Z. Fu, and Q.-X. Cao, "Design and realization of a non-actuated glass-curtain wall-cleaning robot prototype with dual suction cups," *The International Journal of Advanced Manufacturing Technology,* vol. 30, pp. 147-155, 2006.

[3] Q. Zhi-yuan, Z. Yan-zheng, F. Zhuang, and W. Ya, "Fluid model of sliding suction cup of wall-climbing robots," *International Journal of Advanced Robotic Systems,* vol. 3, pp. 275-284, 2006.

[4] N. Elkmann, D. Kunst, T. Krueger, M. Lucke, T. Böhme, T. Felsch, *et al.*, "SIRIUSc—Façade cleaning robot for a high-rise building in munich, germany," in *Climbing and Walking Robots*, ed: Springer, 2005, pp. 1033-1040.

[5] N. Elkmann, M. Lucke, T. Krüger, D. Kunst, and T. Stürze, "Kinematics and Sensor and Control Systems of the Fully Automated Facade Cleaning Robot SIRIUSc for Fraunhofer Headquarters in Munich," in *Field and Service Robotics*, 2008, pp. 505-512.

[6] N. Elkmann, T. Felsch, M. Sack, J. Saenz, and J. Hortig, "Innovative service robot systems for facade cleaning of difficult-to-access areas," in *Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on*, 2002, pp. 756-762.

[7] S. M. Moon, C. Y. Shin, J. Huh, K. W. Oh, and D. Hong, "Window cleaning system with water circulation for building façade maintenance robot and its efficiency analysis," *International Journal of Precision Engineering and Manufacturing-Green Technology,* vol. 2, pp. 65-72, 2015.

[8] Study: Every 10 Seconds A Skyscraper Window Washer Falls To His Death. (2011). Theonion.com. Retrieved 20 March 2016, from http://www.theonion.com/article/study-every-10-seconds-a-skyscraper-window-washer--19734

[9] R. Yue, J. Xiao, S. Wang, and S. L. Joseph, "Modeling and path planning of the City-Climber robot part I: Dynamic modeling," *2009 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2009.

[10] R. Yue, J. Xiao, S. Wang, and S. L. Joseph, "Modeling and path planning of the City-Climber robot part II: 3D path planning using mixed integer linear programming," *2009 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2009.

APPENDIX

l_i	Length of link i
a_i	Center of gravity of link i
m_i	Mass of link i
C_i	Viscous friction on joint i
I_i	moment of inertia of link i around its center of gravity
I_{pi}	moment of inertia of motor i reflected at the output shaft
C_{pi}	Viscous friction of motor i reflected on output shaft
G_i	Static gain of gearbox i
T_{mi}	Torque generated by motor i
T_i	Torque applied on joint i
KE	Kinetic energy
PE	Potential energy
DE	Dissipation energy
g	Gravity
θ_i	Angle of link i from horizontal
u_i	Input voltage to motor i

Table 1. Physical parameters used for the link modelling.

Equipment	Quantity	Power(W)
Arduino Mega	1	6
Motors for the link	3	180
Motor for the brush	1	8.4
Tracks motor	4	144
Vacuum Pump	1	559.5
IR range Sensor	6	1.65
Adafruit 10-DOF IMU Breakout - L3GD20H + LSM303 + BMP180	2	0.3
Adafruit water solenoid valve	1	5.16
Electric valves	2	2.4
	Total	907.41

Table 2. Total power consumption of the system.
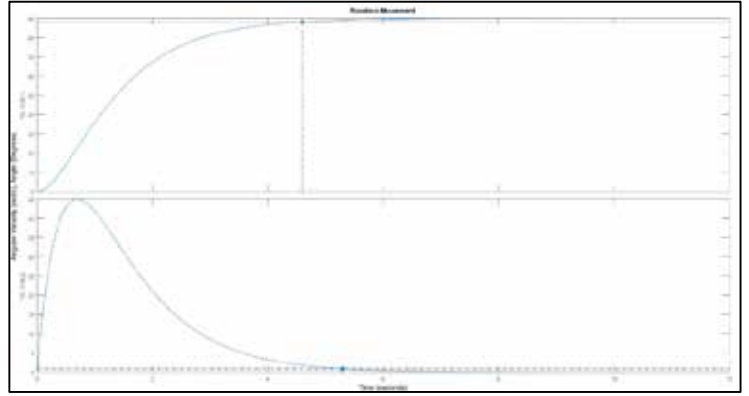
Figure 11. Autonomous Rotation
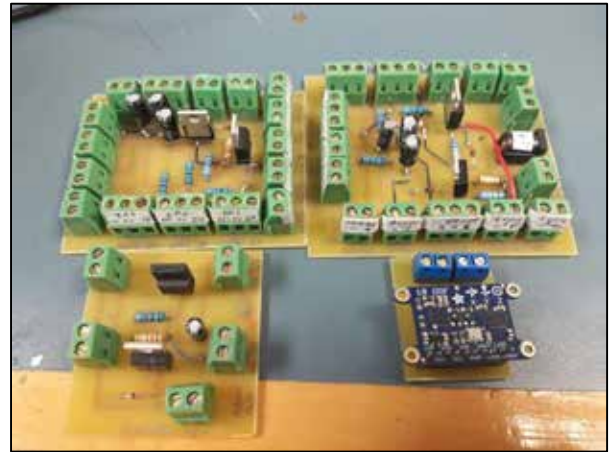
Figure 12. Printed Circuit Boards
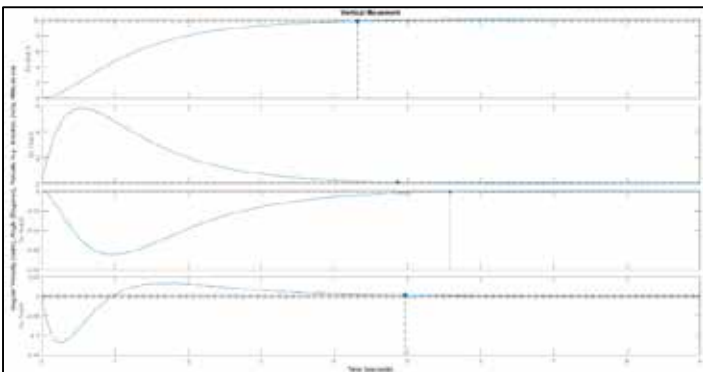
Figure 10. Autonomous Vertical Motion

Green Miles - An Intelligent Carpooling App for a Green Social Solution to Traffic and Parking Congestions

Mortada Haidar, Hussein Hammoud, Ousama Kanawati
Department of Electrical and Computer Engineering
American University of Beirut
Beirut, Lebanon
Emails: {mmh98, hah49, omk09} @mail.aub.edu

*Abstract –*Due to our public transportation infrastructure not being able to accommodate the increasing need for the population to commute from a place to another, researchers have been heavily focused on the implementation of social solutions, such as carpooling, to address the daunting traffic and parking shortage problems in urban environments. Carpooling is the process in which a driver and one or more passengers having semi-common routes share a private vehicle. Although many carpooling systems already have been already proposed and implemented, most of them lack various levels of automation, functionality, practicality, and solution quality, which made them unqualified to be used in our everyday life. Motivated to propose a solution for the traffic and parking shortage problem around the American University of Beirut (AUB) campus, we present in this paper "Green Miles", an android mobile application that provides, using its highly efficient and effective backend algorithm, the matching solution with the minimal travel distance, timely arrival, and maximum fairness while taking into account the riding preferences and constraints imposed by the carpoolers. The computational results and simulations show the ability of such a solution to have its influence on the problems mentioned above by either reducing traffic congestion or alleviating parking woes.

Keywords—carpooling; mobile solution; matching algorithm; intelligent transportation

I. INTRODUCTION

With more than two millions residents, Beirut, similarly to other crowded cities around the world, suffers from overpopulation of not only humans, but also cars. Due to the serious scarcity of land issue, Beirut faces the daily problem of parking shortage. For example, Ras Beirut Neighborhood, which contains different types of land uses (educational, health, hotel, residential, etc.), suffers from the very limited parking spaces available. This problem is not only affecting public streets, but is also affecting different private institutions in the neighborhood, including the American University of Beirut. And naturally, whenever cars are involved, air pollution becomes an issue. According to rideshare.com statistics, every car annually emits its own weight in carbon dioxide into the atmosphere. Different administration in the area, including AUB's, is seeking for solutions that would help addressing these two major problems that can affect the person's mobility, the area accessibility, the traffic flow, the safety of the public, and the pollution to a great extent. Finding a parking place near AUB can be a mission impossible, as described by AUB students and staff. In fact, according to [1], a survey which was conducted back in 2010 among AUB students, shows that more than 70 percent of the students spend more than 40 minutes daily searching for a parking spot. The same survey shows that around 77 percent of students who have cars are willing to pay 100$ a month to have a guaranteed parking space. Another survey that is usually done each 2 year by the civil and Environmental Engineering department at AUB displays amazing results that one needs to observe. As shown in figure 1 below, around 69 percent of AUB's community use vehicles for their transmissions, half of which come driving alone. Such results represent a great motivation for our project that comes as a solution for the parking problem. If Green Miles was able to match each 2 users together, we would reduce the traffic and the shortage parking problem by around 35 percent, which is very significant.
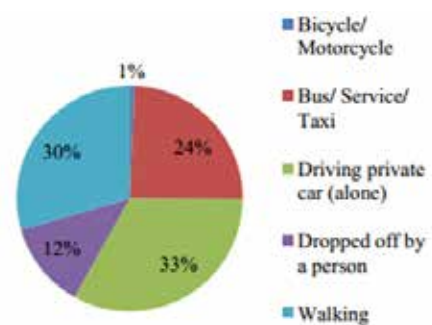
Figure 1: Modes of transportation to AUB, adopted from [13]

The same survey done in [13] studies the satisfaction among the students per mode of transportation. The results are shown in figure 2 below. It clearly shows that the satisfied portion of the community is the part that comes to campus walking. Most of the rest are unsatisfied. This again give us the impression of the need for such a project to be implemented, and that Green Miles could be accepted as a solution by the community to increase its satisfaction.
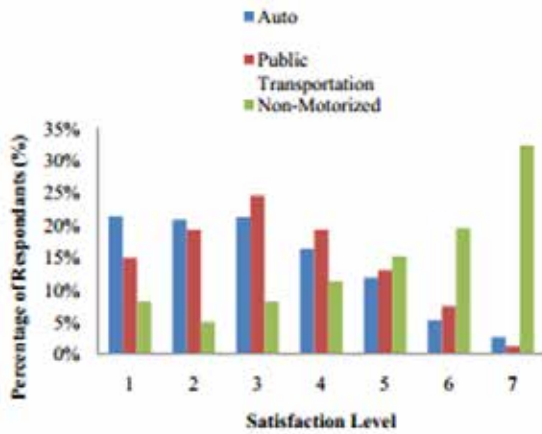
Figure 2: Satisfaction of the community per mode of transportation, adopted from [13]

In order to solve the problem of parking shortage faced daily by students, faculty, and staff at AUB in the surrounding neighborhood, the director of the Office of Financial Planning and Budget at AUB Drew Wickens states that: "the administration is keeping an open mind to workable and affordable solutions that would help students commute to AUB." [2] Also, the Neighborhood Initiative (NI) addressed the parking problem through its Neighborhood Congestion Studies, and is working on parking studies and thinking of alternate modes of transportation to commute to the campus. At the end of the Congestion Studies presented in [11], the NI literally proposed that "AUB commuters would best utilize a mass transport system which combines the professionalism and comfort of private taxis with the vehicle occupancy of shared taxis". This is exactly what Green Miles is going to offer.

Green Miles will be novel technique that contributes in solving different problem such as parking shortage, air pollution, traffic rates, and much more. It allows AUB students and staff member to carpool when coming or leaving from the campus. Each member joining the ride of someone else means that we have one less car on the road. A campus such as AUB, with more than 10000 student and staff will have a significant influence on all the problems listed above, since the app's target will be around 70 percent of the AUB community as discussed above. Green Miles is therefore a solution for AUB students and staff members to carpool. It constitutes an interface between different individuals in AUB's community to help coordinate, schedule, and reward all carpoolers. Given the users' schedules and locations, it will distribute the passengers on drivers in sort of an optimized way to minimize the number of car used and the total cost in terms of time and distance crossed.

II. LITERATURE REVIEW

A. Literature and Already Existing Systems

The carpooling market has been growing in the last couple of years due to the urgent need for a solution for the traffic congestion problem due to the enormous growth in the number of private cars in urban. In addition to the advances in technology that is providing huge opportunity to make our lives more connected, social, and smart. In this section we will explore the available carpooling application and platforms in terms of their functionalities, characteristics, target market, design, and model.

In the literature, matching between drivers and passengers have been classified as Long Term Carpooling problem or Daily Carpooling problem were the goal is to reduce the number of cars used by a group of people [17].

In the first problem, users could be classified by the algorithm as drivers or passengers for a specific day. Thus, each user should have a car and the algorithm will effectively allocate each day each user as a driver or passenger. Taking into account users' schedule, total travelled distance, and a measure of fairness in assigning tasks to each users on the long-term [4]. This automated model rely on a generic heuristic algorithms that optimizes the matching problem maximizing the number of matched users and minimizing the total travelled distance subject to assigning fairness and users time arrival constraints. Further, we will explore the state of the art algorithms in terms of their complexity, limitations, and differentiating factors.

Starting with the algorithm proposed in [16], an iterative algorithm that choose from a set of drivers the earliest time departure driver and find the efficient path from source to destination with respect to least path cost and un-serviced passengers' penalty. Finally, remove this driver and passengers from the list and loop again until the list of drivers is empty. In addition, this algorithm adapts quickly to new users and is quite simple and scalable. However, the maximum percentage of picked up clients at the end of the algorithm was reported to be at most 80%. Another algorithm proposed in [15] is based on Dijkstra Algorithm. The users in the drivers and passengers list are divided over a geographical network model divided into small area settled around the driver coordinates through its optimal path from source to destination. A requested ride is matched if a driver passes near him/her which is determined by the boundaries of the nodes. This algorithm complexity approaches $O(n^2)$ were n is the number of nodes and it allow to instantaneously process requested rides and serve them in real time.

In addition to the above, an adaptive generic algorithm is proposed in [7] to solve the long term carpooling problem. This approach rely on a tree search framework were at each stage a partial solution is reached and the algorithm tries iteratively to expand the bound of the solution to reach the optimal one. The optimality is achieved by applying four different operators which are divide, merge, swap, and move. In the divide operator, m cars with high travel cost could be divided into n cars such that n>m in order to achieve lower total travelled distance. In the merge operator, two non-full pools could be merged to decrease total number of cars in the system. In the swap operator, two car pools can swap passengers in order to achieve a lower travel cost. In the move operator, a passenger could be moved from a car pool into another non-full pool. Near optimal results can be possibly

obtained. However, this algorithm is very greedy and doesn't take into account the fairness between the users or their preferences.

Another attempt to solve to long term carpooling problem is presented in [5]. The carpooling model presented assumes same destination for all users in the system. The algorithm's input is a list of drivers coordinates, a list of passengers' coordinates, and the coordinates of the workplace. The algorithm tries to determine the routes for the drivers maximizing number of clients that are inserted in to car pools. Determining where and which pool to insert each passenger in is the main problem.

In [18] an optimized multi-matching system is presented. This system is based on Lagrangian relaxation focusing on the fairness dimension of the long term carpooling problem while maximizing the fairness of the drivers assigned in the system. This algorithm relaxes some constraints in order to achieve a higher flexibility within the system.

In the second problem, daily carpooling problem. Each day, each user declare him/her self as a passenger or driver on that particular day. Thus, the complexity of the carpooling problem is to assign passengers to drivers and to choose the efficient paths taken by the drivers in order to maximize the matched requested rides and minimize the total driven distance subject to users' time constraint. This problem is modeled as a routing problem which is a subset of the dial a ride problem [3,12].

B. Differentiating Factors

The proposed carpooling solution is different from already existing platforms on many dimensions. First of all, the application is schedule-based, i.e. the matching is done based on the schedules of the users. This is convenient especially for institutions such as AUB where all students, staff, and professors have predefined schedules. All the existing carpooling applications do on-demand matching, whereas we are the first to tackle this problem. Second, one of the most important feature of the app is its security. By being restricted for only the AUB community, it offers a very secure and trusted ridesharing. Finally, due to the localization of the application for AUB, the campus will always be present in the algorithm, either as source or as destination. This property will lead to realizing a significantly faster algorithm, due to the limited number of possible routes to be checked during the matching process.

III. PROPOSED MODEL

Our proposed model is sketched in the diagram given in Fig 3 on the following page. In what follows, we will discuss each part in details..

A. Client

The Client side for our solution consists of a smartphone application developed for the Android OS. It is the user's portal to the system. The application allows a new user to register providing all the required information such as full name, date of birth, gender, AUB username, phone number, car details, etc. After login, a main screen appears allowing the user to view a summary for his statistics and history of using the app. This includes the number of rides taken, the number of rides given, the total CO2 emission saved, the vouchers collected so far, etc. In addition, this screen allows the user to compare his statistics versus the total user's statistics, in order to see how much he saved in his total population. In a simple swipe from the left, the user will be able to navigate through the different pages of the application. To upload his schedule and location, he can go to the corresponding section, where he will be able to submit his location using two different ways either by GPS location or by placing a marker on a google map. The user can then upload his schedule for the academic term, and he can revise the input before submission using a view button. Another section in the app is the rides section, where the user will be able to view a summary of the scheduled rides for each week. Other screens exist such as the user profile screen, the vouchers screen, the help and about screens, etc. The smartphone needs an internet connectivity to exchange data with the server, as well as a GPS capability to feed its coordinates into the system in case google map is not preferred. Screenshots of the client application are shown in Fig. 4.

B. Server

The server, implemented initially on local host, has to respond to the mobile app's requests and queries. Generally speaking, it has to be able to provide login for the user, add user and driver info to the database, store the user's schedule into the database, and keep track of the whole matching process.

Based on the demands presented above, the database we designed consists of 5 tables presented in the below figure. We tried in our design to minimize the number of tables by joining the fields that have a one-to-one relationship in one table. This will reduce the overhead of creating new tables and will make the process of data retrieval simpler and faster.

1- User-Info table: This table stores the corresponding user's info, in addition to relative statistics and reviews.

2- Driver-Info table: If the user declares himself as a driver, he will have to upload info related about his/her car and stored in this table.

3- User's–Schedule table: As per semester, users will be asked to upload their scheduled to be used when performing the matching. This table will store all user's schedules based on day and semester.

4- Rides table: After the matching algorithm takes place, and upon the confirmation of the associated driver, a ride will be created with the specified driver matched to a number of passengers. This table will save the rides' information with the associated driver. Passengers associated to the ride will be stored in the user-rides table having the ride_id as a foreign key.

5- User-Rides table: This table will store the associated users to a specific ride and will keep track of the confirmations or declines of the users.
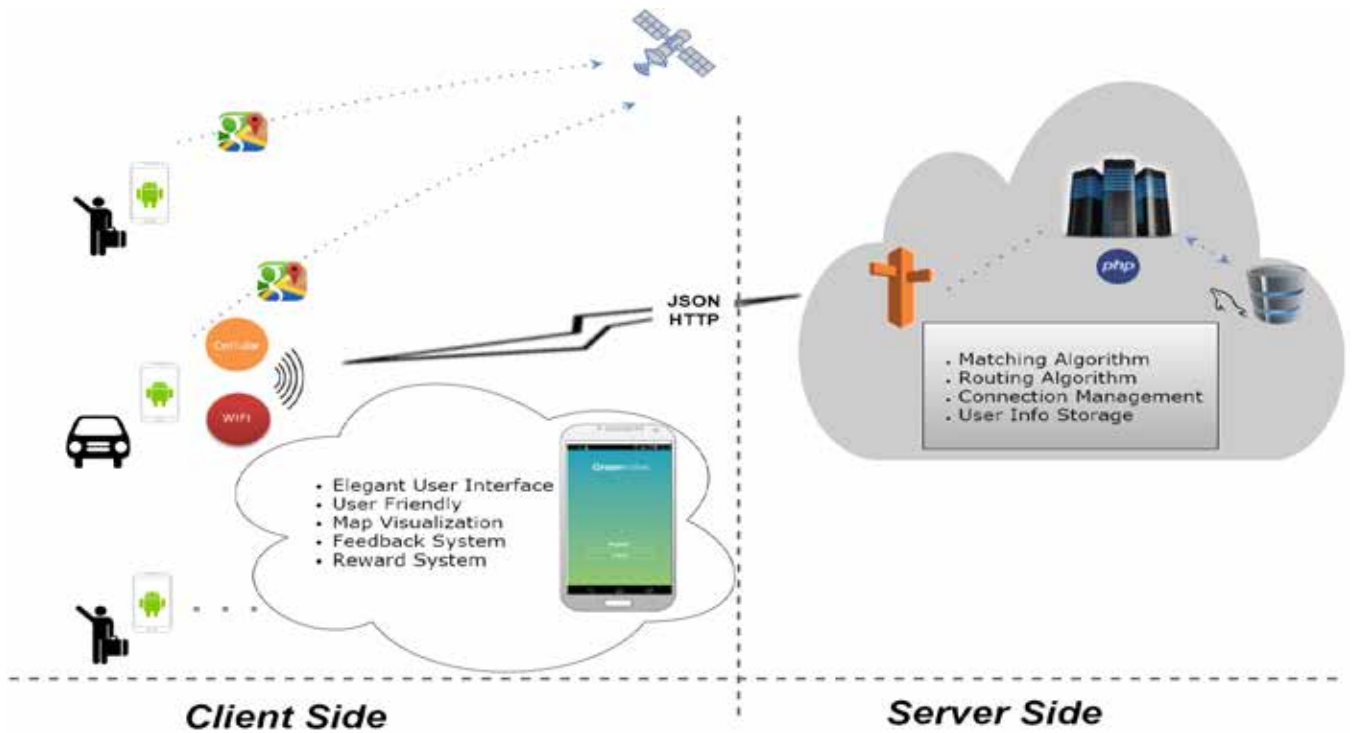
Figure 3: General project architecture

C. Algorithm

Our approach is based on game theory matching algorithms. Inspired by the stable marriage problem that seeks to find a matching pairs solution between two pools of equal size of men and women given for each member a preference list of the members in the other pool [19]. We have solved the drive/passenger problem by developing a stable matching algorithm were the solution will converge to the state were each driver will be matched to the absolute least cost passenger that proposed to the driver. The long term aim of our algorithm is to match all passengers with the available drivers, but also maintain the satisfaction of both drivers and passengers. This might end up having a number of passengers not being matched to any driver. These passengers are not yet tackled in our approach. Our future plan aims on tackling such passengers by providing short term matching which allows passengers to request for a ride on demand. The proposed algorithm can be summarized in the following pseudocode:

Algorithm DPSM

Phase I:

Initialize all d ϵ D & p ϵ P to free
For each d ϵ D
        Compute the extra mileage to each p ϵ P
        Normalize by maximum extra mileage (M)
For each p ϵ P
        Compute the arrival time difference to each d ϵ D
        Normalize by minimum arrival time (T)

Phase II:

While ∃ free p ϵ P who still have a driver to propose to {
d' = highest ranked driver to whom p haven't yet proposed to

if d' have an empty seat:
    (d',p) become matched
else
    compare p with least preferable passenger p' matched to d'
    if (d' prefer p):
        p is matched to d'; p' is free;
    else
        d' and p' remain matched
}
END

In the proposed algorithm, each driver will rank the passengers according to the extra mileage cost. The extra mileage will be calculated as the difference between the distance driven by the driver from the source to the destination and the distance driven by the driver from the source to the passenger location to the destination. Similarly, passengers will rank the drivers according to the arrival time difference cost. This time difference (T) reflect how long will the passenger arrive to the destination before the target arrival time. Arrival time difference can be calculated by the difference between the driver arrival time and the passenger arrival time. Our work ensure an efficient solution that improves the insufficiently addressed issue, such as user preference and fairness, which are the main measures that affect the enthusiasm of a user to contribute in a carpooling system. Our algorithm allows the
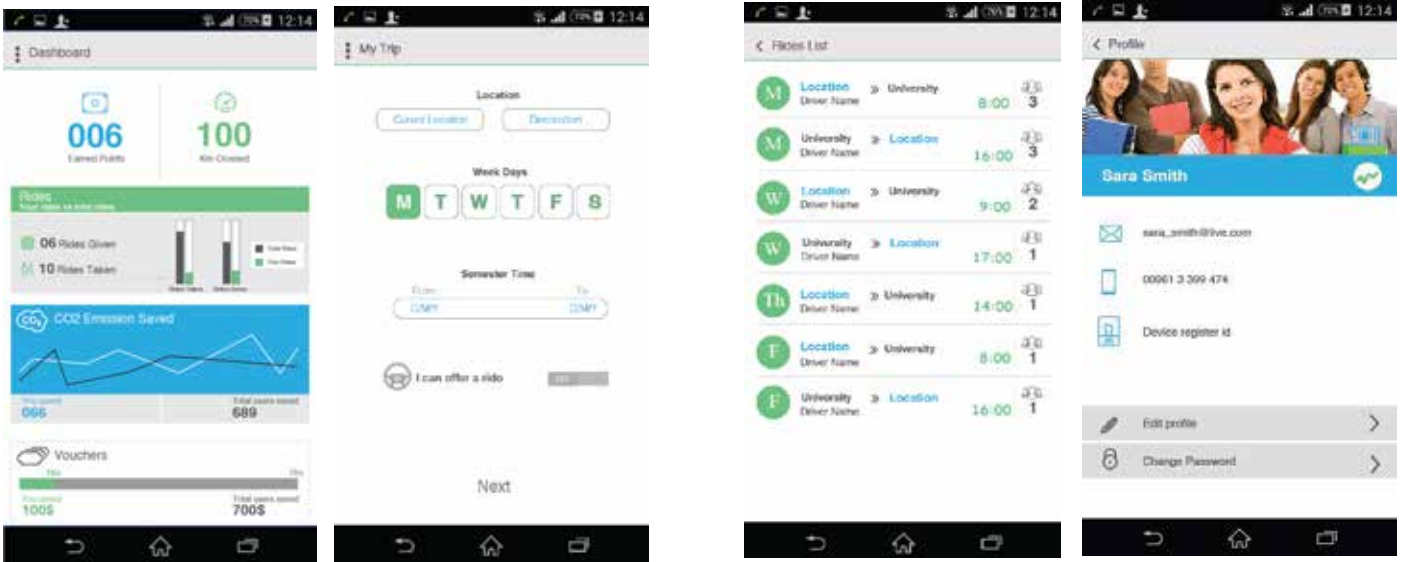
Figure 4: Application interface
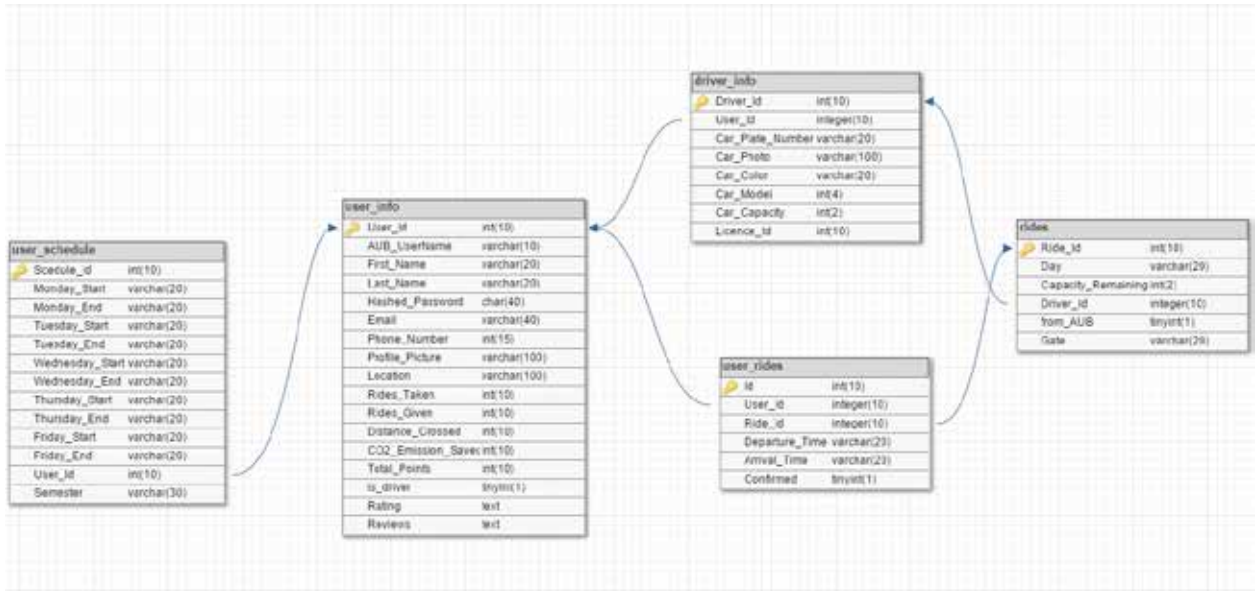
Figure 5: Database design

participants to set their own parameters in term of maximum extra mileage (M) and maximum time arrival (T). These parameters will define the set of passengers a driver is interested in and a set of drivers a passenger is interested in. After that, each driver will rank the passengers according to the extra mileage (extra distance crossed by the driver) normalized by the max extra mileage the driver is willing to drive and each passenger will rank the drivers according to the arrival time difference normalized by the maximum time arrival difference as shown in equation (1), (2) respectively.

$$weight_{d-p} = \frac{M - ExtraMileage}{M} \qquad (1)$$

$$weight_{p-d} = \frac{T - ArrivalTimeDifference}{T} \qquad (2)$$

where T is the minimum arrival time and M is the maximum mileage.

Note that a passenger is not interested in a driver whose arrival time is after the passenger's desired arrival time, or the time difference is greater than his/her maximum time arrival (T). A driver as well is not interested in a passenger if the extra distance crossed by the driver to get the passenger is greater than his/her maximum extra mileage (M). This is reflected by ignoring the parties having negative values from the set of the

first party. Moreover, the higher the weight is, the more satisfied the first party is in being matched to the other party.

As our aim is to maximize the satisfaction of both drivers and passengers, we define the matching weight between a driver and a passenger as a linear combination of both weights, shown in equation (3). The algorithm will work on maximizing the matching weight which reflects the weight (interest) of both the driver and the passenger.

$$matchingWeight = \alpha \times weight_{d-p} + \beta \times weight_{p-d} \quad (3)$$

The fraction $\gamma = \dfrac{\alpha}{\beta}$ reflects the amount of bias in the matching

weight of the driver to the passenger which should be carefully set. As our first attempt, we set it to one which indicates no bias for one party on the other. This will need further investigation later.

IV. SIMULATION RESULTS

The algorithm proposed has been tested heavily to judge its performance, and the optimality of the solution it is producing. A brute force test was performed on 100 test cases generated randomly, for a fixed 8 passengers and 3 drivers, varying the positions for the drivers and the passengers, the time constraints imposed by the drivers and the passengers, and the total allowed extra mileage that the drivers are ready to withstand in order to pick up passengers. The proposed algorithm has been tested on the same cases, and the results were compared. The results show that the algorithm was always able to generate a matching solution that is in the top 0.75% among all the other possible combinations. Specifically, among a minimum of 40,000 possible matching combination for each test case of the 100, the proposed algorithm was able to produce the optimal solution 44 times, second optimal 4 times, third optimal 3 times, fourth optimal 9 times, and fifth optimal 9 times. In total, out of all the possible combinations, the algorithm produced one of the top 5 best combinations with a percentage of 69%, one of the top 10 best combinations with a percentage of 76%, and one of the top 100 best combinations with a percentage of 97%. In what follows, we explore a simplified scenario on which we run the proposed algorithm on. Assume we have 3 drivers and 8 passengers. Each driver sets his/her maximum allowed mileage, while each passenger sets his/her maximum tolerable difference in arrival time. Assume we get the following tables.

Table 1: Drivers maximum mileage

Drivers	max mileage M (km)
D1	5
D2	5
D3	8

Table 2: Passengers' maximum time tolerance

Passengers	Max Tolerable difference in arrival time (hours)
P1	2
P2	2
P3	4
P4	3
P5	4
P6	3
P7	2
P8	2

In addition to the time and distance preferences, we have the exact location of the drivers and the passengers as given in Fig 6. The output of the algorithm is shown in Fig.7. One observation can be seen from this scenario is that P7 was matched to D3 even though D2 is closer to him, and this is due to two main factors. One factor is that the arrival time of D3 is closer to P7 than D2. Another factor is that D3's max mileage is higher than D2, so D3 is more flexible with moving extra distance than D2.

ACKNOWLEDGMENT

REFERENCES

[1] A. Noura. "Parking problems on campus", Now. Media, 28/01/2011, [Online]. Available at
https://now.mmedia.me/lb/en/reportsfeatures/parking_problems_on_campus
[2] A. Zeineddine, M. Harb, "Reducing AUB related Parking Problems through a Park-and-Ride System", Urban Research Methods, 2012, Jun.
[3] Baldacci R, Maniezzo V, Mingozzi A. An exact algorithm for the carpooling problem. Proceedings of CASPT-2000, 8th International Conference "Computer-Aided Scheduling of Public Transport", 2000.
[4] Boukhater, C.M.; Dakroub, O.; Lahoud, F.; Awad, M.; Artail, H., "An intelligent and fair GA carpooling scheduler as a social solution for greener transportation," in Mediterranean Electrotechnical Conference (MELECON), 2014 17th IEEE , vol., no., pp.182-186, 13-16 April 2014.
[5] Calvo R. W. , Luigi F. L. , Haastrup P. and Maniezzo v: A distributed geographic information system for the daily carpooling problem, Computers and Operations Research, 31 (2004) 2263-2278
[6] D. Gale and L.S. Shapley. College admissions and the stability of marriage. American Mathematical Monthly, 69:9–15, 1962.
[7] Guo, Yuhan, Gilles Goncalves, and Tiente Hsu. "A Clustering Ant Colony Algorithm for the Long-term Carpooling Problem. " International conference on swarm interlligence. Lille, France, 2011.
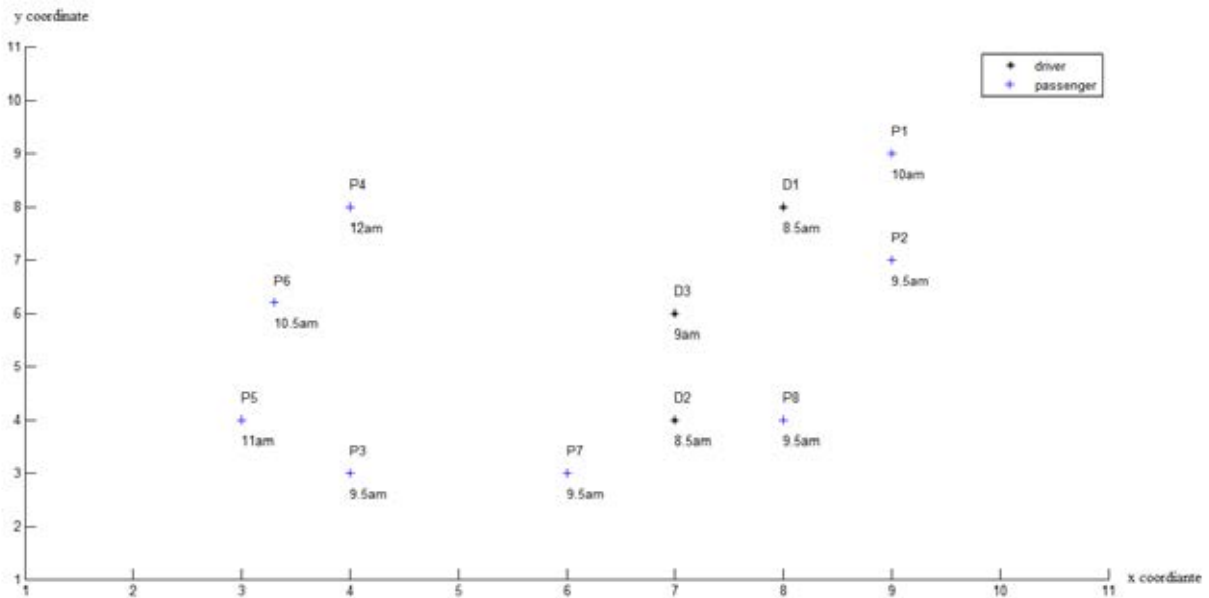
Figure 6: Locations for drivers and passengers
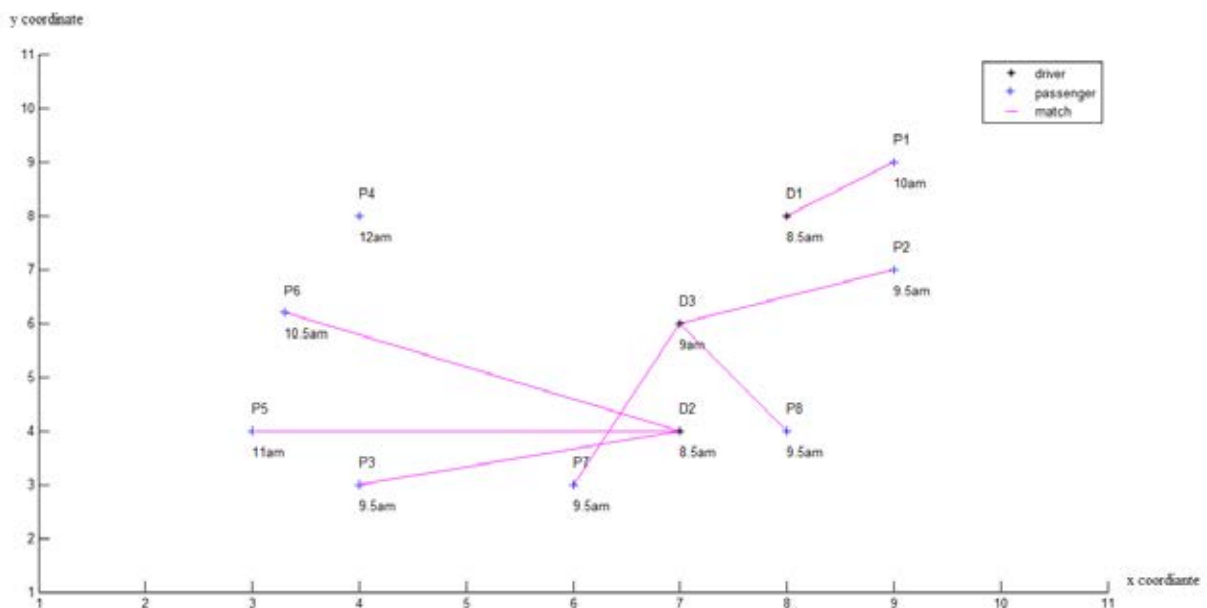
Figure 7: Output of the algorithm - possible matches

[8] J. Cogswell, "Google's Android Studio vs. Eclipse: Which Fits Your Needs? - Costum," Dice Insights, 2014. [Online]. Available at: http://insights.dice.com/2014/03/19/googles-android-studio-vs-eclipse-fits-needs/. [Accessed: 2015].

[9] K. Miklouk, "PHP vs. ASP.NET: Costs, Scalability and Performance," Udemy Blog, 2013. [Online]. Available at: https://blog.udemy.com/php-vs-asp-net/. [Accessed: 2015].

[10] L. Bridge, "Mobile Web Apps vs. Mobile Native Apps: How to Make the Right Choice," Lion Bridge, 2012. [Online]. Available at: http://www.lionbridge.com/files/2012/11/lionbridge-wp_mobileapps2.pdf. [Accessed: 2015].

[11] I. Kaysi, "Neighborhood Congestion", The Neighborhood Initiative, 2010, May, available at https://www.aub.edu.lb/NI/RESEARCH/Pages/NeighborhoodCongestion.aspx

[12] Mingozzi A, Baldacci R, Maniezzo V. Lagrangean column generation for the carpooling problem. Technical Report WP-CO0002, University of Bologna, S.I., Cesena, Italy, 2000.

[13] M. Khattab, "American University of Beirut Student Travel Survey ", 2010, Nov

[14] O. Dakroub, C.M. Boukhater, F. Lahoud, M. Awad, H. Artail, "An intelligent carpooling app for a green social solution to traffic and parking congestions", Intelligent Transportation Systems, 2013, Oct.

[15] Sghair, Manel, Hayfa Zgaya, Slim Hammandi, and Christian Tahon. "A Distributed Dijkstra's Algorithm For The Implementation Of A Real Time Carpooling Service With An Optimized Aspect On Siblings. " IEEE Annual Conference on Intelligent Transportation Systems. Madeira Island, Portugal, 2010.

[16] Son, Ta Anh, Le Thi Hoai An, Pham Dinh Tao, and Djamel Khadraoui. "A Distributed Algorithm Solving Multiobjective Dynamic Carpooling Problem." International Conference on Computer & Information Science. 2012.

[17] Yarrentrapp, K. , Maniezzo, Y. , Stutzle, T. : The Long Term Carpooling Problem: On the Soundness of the Problem Formulation and Proof of NP-completeness. Technische Universitat Darmstadt (2002)

[18] Y. Lin. Matching Model and Heuristic Algorithm for Fairness in the Carpool Problem. Doctoral Thesis of National Central University, Taiwan, 2009.

[19] D. Gale and L.S. Shapley. admissions and the stability of marriage. American MathemCollege atical Monthly, 69:9–15, 1962

Intelligent Automated Vehicle Gear Switching Simulation Software based on Fuzzy Reasoning

Ralph E. Abboud
Department of Electrical and Computer Engineering (ECE)
Lebanese American University (LAU)
36 Byblos, Mont Lebanon, Lebanon
ralph.abboud01@lau.edu

Abstract- **The goal of this project is to develop an automatic gear switching simulation agent which makes its decisions in a more intelligent manner compared with legacy systems. While legacy automatic gearboxes base their decisions on vehicle RPM, speed, and applied throttle levels using traditional mathematical models usually based on exact reasoning, this design employs approximate reasoning using fuzzy logic computations. This design also factors in additional parameters such as car gas levels (full, half-full, and/or empty tank) and terrain nature (uphill/downhill and/or turns) for more rational/subtle decision-making and optimal fuel efficiency and comfort/aggressiveness in driving (following the driver's preferences).**

I. INTRODUCTION

With the development of computer aided automation systems and their increasing use in automobile technology in the few past decades, car manufacturers have been able to significantly improve the automation of gearboxes: aiming to make driving simpler and more enjoyable. However, despite the progress made in recent years, automatic gearboxes still cannot conclusively best their manual predecessors, with many drivers finding that gearbox automation reduced their control of their vehicles and lowered their fuel efficiency. This has led to the introduction of alternate solutions, namely "tiptronic" gearboxes, which are a hybrid of both manual and automatic gearboxes. Reaching the next phase in automatic gearbox performance will require ushering in yet another technological breakthrough: Artificial Intelligence (AI).

This paper describes the design and development of an intelligent software agent simulating an automatic vehicle transmission system. This paper will explain the development process from start to finish, from the simulation of the chosen car (i.e., Corvette C5) to the development of additional features, such as a smart braking system that will assist the gearbox in ensuring an optimal and safe driving experience, especially for adrenaline-loving sports car enthusiasts. Preliminary simulation tests are promising.

II. BACKGROUND

2.1. Context

Automatic transmissions were developed relatively recently to simplify driving: these transmissions change the car's gear autonomously and automatically during travel, sparing the driver the burden of having to switch gears manually. These transmissions work with the objective of keeping the vehicle's engine in its optimal range of revolutions per minute (RPM) so as to drive maximum performance of out the car.

Legacy automatic gearboxes base their decisions on vehicle RPM, speed, and applied throttle levels, using traditional mathematical models usually based on exact reasoning [1].
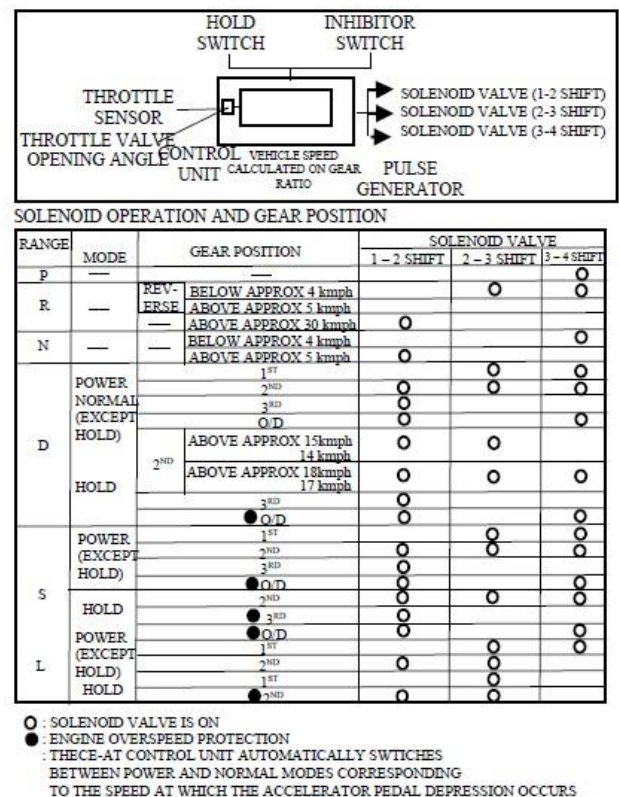
Figure 1: A Sample set of crisp rules governing the operation of a legacy automatic transmission.

This project aims to create a software that intelligently emulates automatic gear shifting, such that gear changes are made more efficiently and optimally, using approximate reasoning with fuzzy logic computations [2].

2.2. Prerequisites

Developing an automatic gearbox required knowledge of a car's inner workings, mainly the operation of an electronically controlled transmission (ECT) and the physical model of a vehicle's movement (acceleration and deceleration) [3].

<div align="center">III. PROPOSAL</div>

3.1. Building Blocks

At the most basic level, this software decides whether to raise, lower, or maintain the gear the simulated car is running on based on fuzzy logic. The software agent's inputs are: the vehicle's engine's revolutions per minute (RPM), throttle and brake values to emulate application of brake and/or throttle by the driver, as well as acceleration, and gas level. Since the agent is fuzzy, these parameters have to be fuzzified and converted into several fuzzy partitions. Hence, fuzzy membership functions had to be created for all the aforementioned inputs. The defined membership functions for three of the six inputs to the fuzzy agent (RPM, acceleration and throttle) are shown in the figures below:

Figure 1: Revolutions per minute (RPM) fuzzy membership functions.
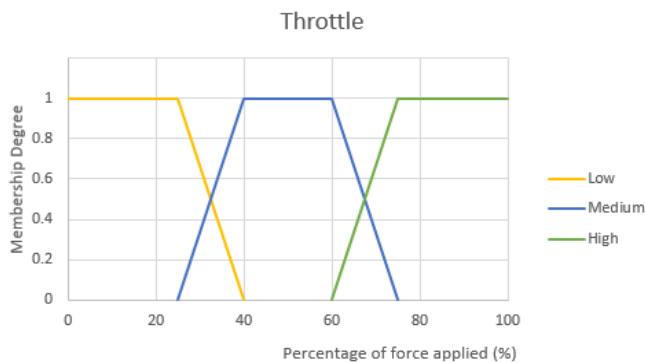
Figure 2: Throttle fuzzy membership functions.

Figure 4: Acceleration fuzzy membership functions.

A fuzzy agent handles its fuzzy inputs through condition-action rules, so this software utilizes over 80 different rules, some of which are given in Table I, to produce fuzzified outputs based on Larsen's product inference. When these outputs are computed, they are then aggregated using the weighted sum rule. The aggregated output is then defuzzified using maximum to the left defuzzification, which will produce a discrete decision (gear up, gear down, or maintain gear): due to the output's singleton membership functions, shown in Fig. 5, the output will be a set of 3 singletons of different weights. The singleton with the highest weight is chosen and its corresponding action is performed.

<div align="center">TABLE I:
Sample Condition-Action Rule.</div>

RPM	Throttle	Brake	Gas	Acceleration	Action
Redline	Low	High	Any	Any	Maintain
Optimal	High	Low	Any	Peak	Maintain
Optimal	High	Low	Any	Cruising	Gear Up
High	Medium	Low	Any	Intense Slowdown	Maintain
Low	Low	Low	High	Steady	Gear Down
High	High	Low	Any	Cruising	Gear Up
Low	Low	Any	Low	Intense Slowdown	Maintain
Low	Low	Medium	High	Steady Slowdown	Gear Down
Redline	Low	Low	Any	Peak	Gear Up
Optimal	High	Low	Any	Steady Slowdown	Gear Down

Figure 5: Gear change decision singleton functions.

It is important to note that, while the steepness of the road being driven on (uphill or downhill) is not a direct input to the fuzzy agent, it does indirectly affect decision-making, since this parameter can affect the car's movement via an additional gravitational force that supports/opposes the car's movement.

In order to make the project more realistic and intelligent, a Smart Brake agent was introduced to apply a brake in potential emergency situations, for instance when the car is speeding up with a sharp turn coming up or when the car is uncontrollably or dangerously accelerating downhill. When this agent is enabled, it assumes control of throttle and brake, but will cede it when its job is complete. Based on the situation, the agent will apply a low, medium, or high brake, also based on fuzzy logic. Initially, turn intensity (an angle in degrees) and vehicle speed (in km/h) are fed into the system and fuzzified based on membership functions shown in Fig.6 and Fig. 7. The agent then uses the fuzzified intensity of the turn (Slight, Regular, and Sharp turns) and fuzzified speed (Low, Medium, High) to produce, based on some logical rules, fuzzified outputs based on Larsen's product inference. These outputs are aggregated using the weighted sum mechanism and the aggregated output is defuzzified using center of gravity defuzzification, which will produce a certain intensity with which the brake is applied while, of course, setting throttle to zero. After that, smart brake is turned off and control is returned to the driver.

The Condition-Action Rules of this agent, as well as the membership functions for turn intensity and speed, are shown in the table and figures below:

TABLE II:
Smart Brake Agent Condition-Action Rules.

Turn Intensity	Speed	Brake Applied
Slight	Low	None
Slight	Medium	Low
Slight	High	Medium
Regular	Low	None
Regular	Medium	Low
Regular	High	High
Sharp	Low	None
Sharp	Medium	High
Sharp	High	High
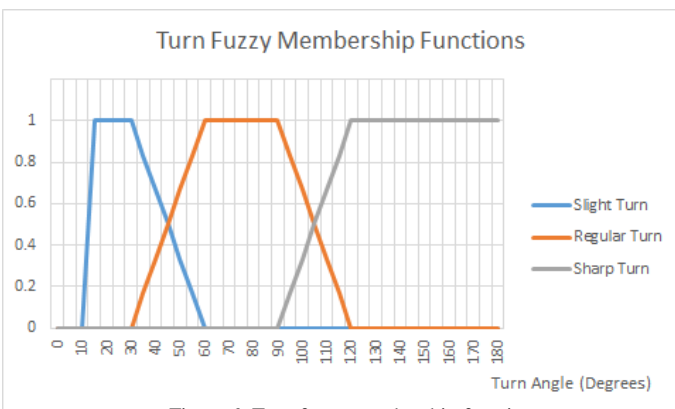
Figure 6: Turn fuzzy membership functions.
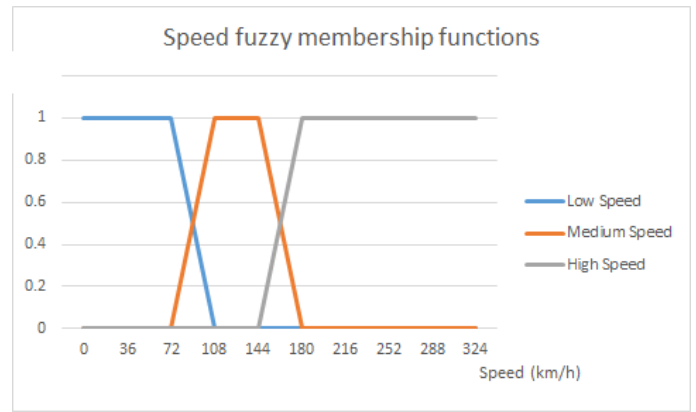
Figure 7: Speed fuzzy membership functions.

3.2. Design and Conceptual Model

This model is based on fuzzy logic since a gear change decision in itself is fuzzy (non-crisp) and depends on several inputs that cannot be processed analytically. Some inputs are more important than others in decision making (For example, engine revolutions per minute are more important than gas level.).Therefore, these inputs are represented and handled with more rules and linguistic variables. The design of the fuzzy membership functions corresponding to these linguistic variables required great attention to detail and meticulous testing, since inaccurate or unreasonable functions would lead to undesirable agent behavior (such as excessive gear switching and alternation). The fuzzy membership functions were also normalized. In other words, for any given value for an input, the sum of all membership degrees for all corresponding linguistic variables will equal one. This will simplify the conceptual analysis of the system and render it more intuitive and relatable to what traditional logic offers. With the functions in place, the rules had to be set in a way that would not undermine the agent's operation. This task proved to be particularly difficult: Since this agent can be modeled as a 6-Degrees-Of-Freedom system, it must handle a 6-dimensional input vector, and do so perfectly, lest the software get stuck in a limbo state (continuously switching between two neighboring gears) or behave incorrectly. Hence, the rules had to be set such that they do not interfere with one another or overlap on each other's trigger state.

Should this software be implemented for a real-life application on the streets, the Performance/Environment/Actuators/Sensors model (i.e., PEAS model) will be as follows:
- Performance Measure: Percentage of decisions produced that the driver agrees with during agent testing.
- Environment: Roads that can be flat, downhill or uphill, which can also be straight or left and right turns of variable intensity.
- Actuators: a transmission system similar to the one implemented in automatic vehicles is used for the basic agent and smart valves connected to the hydraulic pressure that in turn are connected to the wheels are used for the Smart Brake agent.

- Sensors are required to measure inputs: a speedometer for speed, a tachometer for RPM, valves to measure the levels of throttle/brake applied, an accelerometer for acceleration and a gas sensor for gas measurement. Road curvature and steepness can be measured and used for gearbox control via speed comparison done at the wheel level (on all vehicle wheels) [4]

3.3. Implementation

This project was developed from the ground up. The first part implemented was a mathematical model to simulate the Corvette C5's operation. Data concerning engine performance and torque are withheld by Corvette, so the model had to be based on empirical data found for the Corvette's acceleration, deceleration and braking, in a way that minimizes error [5]. With the car model in place, a manual gearbox using the C5's actual gear ratios was coded: later on, this gearbox will be controlled by the intelligent agent. With the gearbox now ready, the building blocks for the fuzzy agent (aggregation/ condition-action rules/defuzzification, etc.) were combined so as to eventually be used to make gear switches automatic. Next, a user interface was developed such that a user can evaluate all aspects of the car's operation (nature of the upcoming road, car's current gas level, speed and rpm) and take control of throttle and brake, as any driver can on the roads. The interface offers data similarly to the cockpit of an actual car.

The main agent now complete, the second intelligent agent, the Smart Brake, designed to make the vehicle safer especially given the high speeds which it can reach, was implemented using the previously discussed fuzzy logic techniques and membership functions. This smart brake agent will act as a last resort defensive agent that will apply the car's brake should the driver put himself in a dangerous situation (like speeding at a sharp turn). For the sake of testing and demonstration, turns, downhill, and uphill are generated based on a Poisson distribution in the code, so as to emulate real turns and slopes in driving as accurately as possible.

Finally, for the sake of testing and flexibility, all agents in this software can be disabled. In other words, the user can drive his car in manual mode, disable smart brake and even disable the random generation of turns and uphills/downhills.

All in all, the final project consists of four components:

1) A car mathematical model which simulates the Corvette C5's operation and produces acceleration, RPM, speed and gas values from the gear, throttle and brake inputs
2) A Poisson-based random terrain generator to test the car on several road configurations of variable difficulty
3) The main fuzzy logic-based intelligent gearbox, which controls the vehicle's selected gear
4) The additional smart brake agent, which can optionally control the vehicle's braking in emergency situations.

A diagram illustrating the project design and its inner components is shown in Fig. 8:
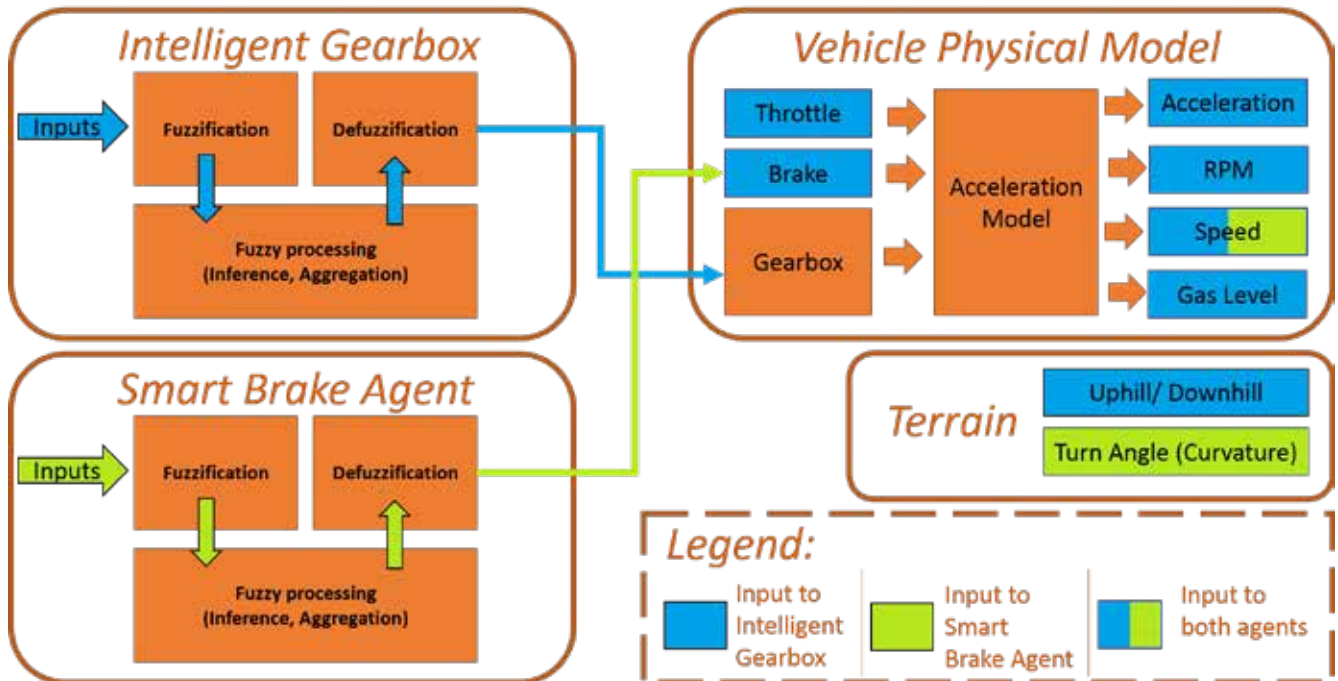
Figure 8: Project Operational Diagram.

IV. TESTING

This software makes decisions which can potentially place drivers at risk, so testing its decision-making is a crucial part of development. To test this software, manual simulation of critical situations had to be made in its code. The agent's response was then observed and recorded. If the action observed, if any, is logical, sensible and, crucially, fast, then the agent will have successfully handled the situation and changes would not need to be made. However, should the agent act incorrectly or slowly, then the condition-action rules responsible for this, as well as any computational overhead slowing the system down are to be reviewed, evaluated and rethought.

4.1. Test Cases,

Testing encompassed a very large range of situations in which both agents (Gearbox and Smart Brake) are called into action. The intelligent gearbox was continuously tested throughout its development phase under a variety of different conditions, to very satisfying results, some of which are shown in the following table:

TABLE III:
Sample test cases and results

Situation	Response
Redline RPM, low throttle and a regular brake applied	Agent maintains same gear
Low RPM, throttle and gas level. Car decelerating rapidly	Agent maintains gear
Low RPM, throttle and high gas level. Car decelerating rapidly	Agent lowers gear
RPM in optimal range, high throttle application, low (no) brake, but car decelerating	Agent lowers gear
RPM in optimal range, high throttle application, low (no) brake, car is cruising	Agent raises gear

The agent's response with respect to the fourth test mentioned in Table III is also shown in Fig. 10. As for the smart brake agent, it was tested in critical situations and yielded highly satisfactory results. For instance, if the car redline RPM and is going downhill, the smart brake agent is activated to protect the engine. This situation is shown in Fig. 9.

Crucially, neither system lagged when responding to a particular situation. Indeed, both intelligent agent successfully handled test cases without any delay in response (all responses seem to happen in real-time), which is vital for an application in which changes occur in fractions of a second.

V. CONCLUSION

The gearbox agent, car model, smart brake agent, and Graphical User Interface (GUI) were each designed with a view to making the most of the Corvette C5's power while maintaining safety and ease of use. Implementing this project required a lot of reading about cars' internal mechanisms and tiresome number crunching for the computation of physical constants, and this is without mentioning the challenge of implementing a coherent and logical agent for the car which can successfully handle multiple inputs of different nature and

dimension. Yet ultimately, all the work bore fruit and eventually, a reasonable set of condition-action rules and a mathematical car model simulating the Corvette C5 with remarkable accuracy were developed, which led to a solid and reliable intelligent agent. This project is a testament to the power of fuzzy logic as a foundation for intelligent agents whose applications can certainly go beyond vehicle gearboxes.

Figure 9: Smart Brake activates at advent of risk.

VI. FUTURE WORKS AND PERSPECTIVES

Looking ahead, the scope of this project will be extended further. For instance, one could consider adding a learning element to the gearbox, such that it would "learn" the driver's style. This learning would be done by updating the membership functions or adding weights to the condition-action rules based on the learning agent's inputs. For inputs to the agent, the simplest and most effective means of input would be a learning phase in which the driver drives the vehicle in manual transmission or, should the driver opt otherwise, to allow driver feedback (the driver overrides a decision he deems wrong) that is used to update the gearbox model. Beyond this, we plan to extend this software to handle any vehicle model: Through a simple change in physical constants (gear ratios, friction coefficients, etc.), new adapted membership functions would be generated, allowing the agent to make intelligent decisions for a completely new vehicle. Hence, this would hopefully pave the way for a universal vehicle gearbox, which would simplify the maintenance and driving of all cars in the years to come.

Figure 10: Test Case Number 4: From screenshot 1 to 2, one can notice the car decelerating despite full throttle being applied. In screenshot 3, the agent's response is shown. The agent lowers the gear allowing the car to regain acceleration.

REFERENCES

[1] What-When-How. *What-When-How*. [Online] [Accessed: November 15, 2015.] Available at: http://what-when-how.com/automobile/electronically-controlled-automatic-transmission-automobile/

[2] Russell, S.and Norvig, P., 2010. Logical Agents. *Artificial Intelligence: A Modern Approach*. s.l. : Pearson Education Inc., 2010, pp. 234-365.

[3] Monster, M.. 2003. Car Physics for Games. *awakicki.info*. [Online] 2003. [Accessed: 11 15, 2015.] Available at: http://www.asawicki.info/Mirror/Car%20Physics%20for%20Games/Car%20Physics%20for%20Games.html.

[4] Gruhle, Wolf Dieter, et al. 1998. *Process for Controlling an Automatic Vehicle Gearbox*. U.S Patent 5748472, issued May 5, 1998.

[5] Automobile Catalog. *Automobile Catalog*. [Online] [Accessed: 11 15, 2015.] Available: http://www.automobile-catalog.com/make/chevrolet_usa/corvette_c5/corvette_c5_coupe/1998.html.

Interactive Face Robot

Fady C. Elias-Sleiman, Hasan A. Ezzidine, and Nourhane H. Haidar
Department of Computer and Communications Engineering
American University of Science & Technology
Beirut, Lebanon
{fady.e.sleiman, hasanezzidine, nourhanehhaidar}@gmail.com

Abstract- **The interaction between humans and robots has increased and become more personalized in this century. Robots are being used in various fields such as medicine, manufacturing, automation, etc.; however, they lack the interaction capability that is found in a human mainly because more than 60% of human communication is non-verbal, and thus, is hard to replace him. This project presents an interactive face robot which is able to mimic the six universal human facial expressions: happiness, surprise, fear, sadness, anger, and disgust. It can interact with the user using speech recognition in order to mimic a career/major advisor. The Interactive Face Robot is made out of a rubber mask to simulate the skin. Facial expressions are created by moving muscle-like structures based on a facial mass-spring model. Nine servo motors move the different action units in the face to create the different expressions. The Interactive Face Robot mimics a career/major advisor using speech recognition software, the BitVoicer, which recognizes the user's speech and communicates with a LabVIEW™ program that evaluates the answers and calculates the best career/major choice. The LabVIEW™ program, which runs on a laptop computer, also converts the pre-defined questions that the robot asks into voice, and communicates with an Arduino Mega the expression that the face is to generate. Using pulse-width modulation, the Arduino Mega controls the needed servo motors to create the necessary facial expression. The Interactive Face Robot has proven to be user-friendly, and interacted with humans as pre-programmed. This Interactive Face Robot can replace a career/major advisor, and can be programmed as a receptionist, front-desk operator, or can be used in therapeutic applications. Facial expressions provide robots with more capabilities for human-like interactions, making them more adequate for applications requiring interactions with humans.**

I. Introduction

As the use of service robots continues to grow rapidly, an increase in the human-robot interaction (HRI) is required [1]. More than 60% of the human interaction is inducted nonverbally [2]. Thus a human-friendly interface should be developed to communicate information the same way we do with humans. For a robot to be effective in the human world, it must respond to the human emotional state by generating humanlike expressions back.

This project is able to mimic five universal expressions and benefits users in traditional roles such as the servants, assistants, or companions. In a typical work day, an employee might be facing certain personal issues or might be exhausted; a state that is reflected on the individual's facial expressions. Though facial expressions play a major role in communication, this might send the wrong signal to the other person. The proposed robot face is a user friendly system. However, its true worth appears in the therapeutic characteristics. Interactive

robots have been proved to play a major role in improving the users' relaxation and motivation levels by reducing stress [2]. Moreover, it was found very efficient in helping children with autism by engaging them in social interaction interface [1].

The significance of this project is to infer a broad scope of uses in the field of HRI. The system is a fully automated real-time one. The objective is to deliver a complete system that generates facial expressions, recognizes speech, and generates speech in real time. The proposed technique is made up of three primary key points: artificial expressions generation, speech generation, and speech recognition. The proposed scheme can be applied in different industries and circumstances and for different purposes. The goal proposed in this paper is to make the face robot an administrator at a university. Students usually have a hard time deciding what major they should go into. Therefore, the face robot will undertake a career test with the student to find out what the student likes and what not, and accordingly tells the candidate the major that most fit him/her.

II. State-of-the-Art

A Japanese teacher called Hiroshi Ishiguro developed his own robotic doppelgänger. The doppelgänger can act as the teacher by explaining lessons and answering questions. [3] Also, Hisashi Ishihara, Yuichiro Yoshikawa, and Prof. Minoru Asada of Osaka University in Japan developed a child robot that can generate realistic facial expressions. [4]

Within the recent decades, many researches have been trying to compose strategies to automatically classify facial expressions. Various methods have been used. The representative research-work based on this approach includes the work done by Hanson Robotics [5]. The newest addition to their family was Han, a face robot that can engage in conversation and generate facial expressions and make eye contact [5]. On the other hand, the Einstein robot head used 31 servo motors, 27 of them are used to generate facial expressions [5]. What remarks this project from others is that the facial expressions are detected by the movement of the muscles under the skin to be learned by the robot and then generated when necessary. The project describes how each servo motor is used with every particular expression [5].

Moreover, Karsten and Jochens paper is about the construction of the robotic head ROMAN. It shows the three major steps in this project: first how the robot is able to realize facial expressions, then the mechatronic system of this robot

including the neck and eye construction, and finally the software architecture of ROMAN [6].

III. MATERIALS AND METHODS

The project's design can best be described in three parts: hardware, model construction, and software. Figure 1 shows the block diagram of the project, detailing the interaction between the hardware and the software parts which will be described in more details hereafter.

A. Hardware

Every step that went into the construction of the hardware is presented in Figure 2. To begin with, the face was divided into three parts: the upper, middle, and lower. The upper part contains the eyebrows and their corresponding mechanism. The middle part contains the eyes, the upper lip, and the fixed upper jaw. Finally, the lower part contains the lower lip and moving lower jaw, and will be rotating to model the mouth as it opens and closes.

The eyes are ping pong balls that were reshaped and colored to look like the human eyes. The position of the eyes required several calculations. The human eyes are positioned at the same distance from the right and the left nostrils, and to the line perpendicular to the mouth. Moreover, when looking from the top plane of the face, the distance in which the eyes are pushing forward or backwards should be taken into consideration. After fixing the eyes, the back of every eye was attached with an L shaped plastic, as shown in Figure 3, that was fabricated using a 3D printer. The end of the L shape is connected to a joint, as shown in Figure 4, attached to the servo motors arm. This allows the eyes to move right and left simultaneously.

To apply a real lips movement, a guitar string is fixed around the lower and upper teeth. Other guitar strings are fixed at two points on the lower jaw and three points on the upper jaw with a round string connecting each part to model each lip. Every fixed point should be tightened by a servo motor.

Figure 1. System's block diagram.

Figure 2. Making of the hardware.
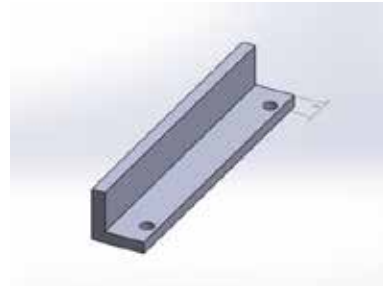
Figure 3. L-shape 3-D CAD design.
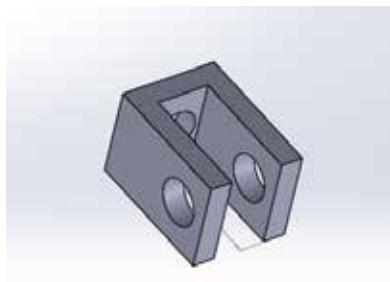
Figure 4. Joint 3-D CAD design.

Influenced by the idea behind the bicycles' brakes, the guitar strings are navigated through plastic pipes that are fixed on the mannequins head. The end of the pipe is immobile, allowing that end of the steel string to be tightened using servo motors.

The lower jaw needed two servo motors to allow easy movement. Moving to the upper jaw, knowing that there are three fixed points on the round guitar string bridge that modeled the upper lip, three servo motors are used. Due to lack of space, a polyamide body was designed to hold the three motors as seen in Figure 5. The motor wings hold a plastic rod that fixes the guitar strings on it in order to move the lips. In addition to that, a cable conjunction is fixed in front of the polyamide body in order to fix the plastic pipes with a screw. The conjunction itself is fixed by a screw so that the pipe, the rods on the motor end and the openings above the lips are on the same line of sight.

Figure 5. Upper jaw.

Connecting the lower jaw and the upper jaw should be done in a way that allows the mouth to open and close. Therefore, the two parts are fixed using an iron joint as seen in Figure 6.

To allow rotation, first the lower jaw is connected with a loosened screw, and second it is smoothened from its back-right and beck-left to make an angle of approximately 30 degrees with the upper jaw. The movement mechanism is done using two steps. The first is to fix three springs from the front end of the lower jaw to the upper jaw. While the second is to connect the back end of the lower jaw to a servo motor fixed on the back of the head, using a plastic thread. This will pull the jaw and make the required rotation. To return the jaw to its initial position, a spring is attached to pull it back.

The next step covers the eyebrows. The eyebrows are fixed 2 cm above the eyes. They are made up of fiber masked by a nylon sheet. The edges are smoothed and later drilled and fixed from their middle to be able to rotate easily to perform happy, angry, and neutral emotions. This mechanism is made up of plastic threads that are connected from the center of the eyebrows, up to the upper part of the head, then into the head to be connected to servo motors. This will allow the eyebrows to move up to 60 degrees anti-clock wise.

As for the neck, an iron rod is attached on a base platform that is connected to a rotating disk around a fixed axis. The base can be adjusted to the left and right using a servo motor. Such a mechanism will not be affected by the robot's weight.

Figure 6. The connection of the jaws.

B. Model Construction

After assembling all the parts together, there comes the binding stage between the rubber face and the skull. This was too risky because of the incompatibility between them. Knowing that facial expressions are produced due to fixed AUs and flexible ones, the mask is secured in some places, and left loose in others. Also, the cheeks are filled with sponge to achieve the smooth touch of the face and give it a realistic form. The nose is also thickened by sponge and compressed to have the desired shape. After fixing all the face, the motion of every motor was tested. You can see the before and after of the Interactive Face Robot in Figure 7.

C. Software

The project's system is based on three software: LabVIEW™ (National Instruments, Austin, TX, USA), Arduino™ (Arduino, Northern Italy), and BitVoicer™ (BitSophia, So Paulo, Brazil) which are connected via a serial communication (Figure 8). The LabVIEW™ program is called the decision maker since it connects the BitVoicer™ to the Arduino™. The LabVIEW™ program gets its input from the BitVoicer™, and then according to a predefined algorithm, it will send a command to the Arduino™ microcontroller which will drive the corresponding servo motors.

Figure 7. Before and after of the interactive face robot.

Figure 8. Software block diagram.

451

The face robot can be used in a variety of fields. One of them is to act as a university administrator. One of the hardest things that a student encounters when moving from the school phase to the university phase is to choose the right major. That is why, the previously mentioned program was edited to be able to do a career test with new students. A sample career test was performed. Eleven majors were taken into considerations, which are: Radio TV, Interior Design, Journalism, Public Relations, Computer Science and Informatics, Finance, Management, Marketing and Advertising, Computer and Communications Engineering, Biomedical Engineering, and Clinical Laboratory Sciences. The sample career test displayed in Table. I consists of eighteen questions. The users' answers are a simple yes/no answers. Whenever the user answers yes on a certain question, the majors related to the question should increment once. The majors are distributed between the majors equally with no favor of one over the other. Finally, when the test is over, the program should be able to know the major with the highest score and relay the information as speech to the user. The Arduino code controls the motors in the face to produce an emotion. Table II displays the angles the ten motors rotate to create every emotion. There are one motor for the eyes, one for the neck, one for every eyebrow, one for the jaw, three for the upper lip, and two for the lower lip.

TABLE I
SAMPLE CAREER TEST

	Question	Incremented Majors
1	Do you like mathematics?	Computer and Communications Engineering; Computer Sciences and Informatics; Biomedical Engineering; Finance; Interior Design
2	Are your communication skills good and do you like to meet new people?	Journalism; Marketing and Advertising; Public Relations
3	Do you find out-of-the-box solutions for your problem?	Computer and Communications Engineering; Computer Sciences and Informatics; Marketing and Advertising
4	Are you confident and do you like to show yourself to the public?	Journalism; Radio TV; Public Relations
5	Do you like to learn about living things?	Biomedical Engineering; Clinical Laboratory Scientist
6	Are you good in analyzing and interpreting data?	Finance; Management
7	Do you have an analytical mind and do you organize your responsibilities?	Finance; Management
8	Are you interested in acting, film making, or media production?	Journalism; Radio TV
9	Did you spend most of your high school time in laboratories?	Biomedical Engineering; Clinical Laboratory Scientist
10	Do you have the ability to recognize general principles in certain situations?	Clinical Laboratory Scientist
11	Do you need the freedom to create and be creative?	Interior Design
12	Do you like designing and artwork?	Interior Design
13	Are you appealing?	Public Relations
14	Do you like tinkering with electronics and robots?	Computer and Communications Engineering
15	Do you like to sit in front of your laptop?	Computer Sciences and Informatics
16	Do you interact with people in patience and congeniality?	Public Relations
17	Do you have the ability to link things together?	Management
18	Do you like to do heavy researches?	Marketing and Advertising

TABLE II
EMOTIONS VS ACTION UNITS

Emotion	FACS Code	
	Muscle Description	Associated AUs
Normal	Poker face	Neck : 50°; Left eyebrow: 120°; Right eyebrow: 30°; Eyes: 180°; Jaw: 130°; Left lower lip: 150°; Right lower lip: 90°; Left upper lip: 170°; Middle upper lip: 170°; Right upper lip: 170°
Surprised	Eyebrows raised; Mouth open; Eyes open; Lips protruded	Neck : 50°; Left eyebrow: 90°; Right eyebrow: 0°; Eyes: 180°; Jaw: 50°; Left lower lip: 120°; Right lower lip: 130°; Left upper lip: 170°; Middle upper lip: 170°; Right upper lip: 170°
Angry	Nostrils raised; Mouth compressed; Furrowed brows; Eyes wide open	Neck : 50°; Left eyebrow: 140°; Right eyebrow: 40°; Eyes: 180°; Jaw: 130°; Left lower lip: 150°; Right lower lip: 90°; Left upper lip: 170°; Middle upper lip: 170°; Right upper lip: 170°
Frustrated/Contempt	Lip protrusion; Upper lip raised; Eyebrows raised	Neck : 50°; Left eyebrow: 90°; Right eyebrow: 0°; Eyes: 180°; Jaw: 130°; Left lower lip: 150°; Right lower lip: 90°; Left upper lip: 170°; Middle upper lip: 170°; Right upper lip: 170°
Sad	Mouth depressed; Eyebrows partially raised	Neck : 50°; Left eyebrow: 115°; Right eyebrow: 10°; Eyes: 180°; Jaw: 130°; Left lower lip: 120°; Right lower lip: 90°; Left upper lip: 170°; Middle upper lip: 170°; Right upper lip: 170°
Happy	Upper lip raised at corners; Eyebrows partially raised; Jaw partially opens	Neck : 50°; Left eyebrow: 100°; Right eyebrow: 0°; Eyes: 180°; Jaw: 100°; Left lower lip: 150°; Right lower lip: 90°; Left upper lip: 100°; Middle upper lip: 170°; Right upper lip: 100°
Move left	Turn neck to the left	Neck: 130°
Move right	Turn neck to the right	Neck: 30°
Move eyes	Move eyes left then right	Eyes: from 0° to 180° with delay 50ms for every 5°
Talk	Open and close mouth	Neck : 50°; Left eyebrow: 120°; Right eyebrow: 30°; Eyes: 180°; Jaw: 130°; Left lower lip: 150°; Right lower lip: 90°; Left upper lip: 170°; Middle upper lip: 170°; Right upper lip: 170° / Delay (200ms) / Neck : 50°; Left eyebrow: 120°; Right eyebrow: 30°; Eyes: 180°; Jaw: 90°; Left lower lip: 120°; Right lower lip: 120°; Left upper lip: 150°; Middle upper lip: 150°; Right upper lip: 150°

IV. Testing and Results

This project presented results for a fully automated real-time expression generation system able to generate the five universal expressions: normal, happiness, sadness, surprise, and anger (Figure 9).

The primary goal of this work is to solve an engineering problem: how to make the human robot interaction as humanly as possible. Facial expressions play a major role in this interaction. Therefore, the appearance of human facial muscles was approximated with available motors. The career test was experimented on several graduate friends, and almost all of them got the result of their university major or something that they have always wanted to do, but never got the chance to. Moreover, they enjoyed taking the test since the robot was interacting with them non-verbally. The robot was smiling and making funny expressions so that the user felt comfortable and happy. It was observed that users interacted more with the robot than with either a laptop or with a human.

V. Conclusions and Future Work

Because of the demand to study human robot interaction, a robot platform was produced to allow this requirement. The exact task of this project is to have a robot face that is human-like and produces facial expressions similar to those in humans. The methodology followed is simple. The robot takes as an input the users' speech. Then, the robot interacts while generating facial expressions by feeding the Arduino with specific command that will thus move the servo motors to generate the desired expression by stretching and releasing the rubber face on the AUs that were generated in the testing phase. Moreover, in order to make this interaction more realistic, speech was added for the robot to talk back with the user. The main task of the robot is as an administrator at a university. It asks new students questions to conduct a career test. According to the students' responses, the robot will tell him/her what major he/she should get into.

The scope of this project is promising. It is accurate to say that such a scheme can be used for different purposes easily. For example, the project can be used in the therapeutic domain to help people with autism; also, it can be useful in the education application or even in aged care. For this project to reach its foremost goal we should allow it to interpret the facial expressions of the human facing it and act accordingly. Also, artificial intelligence should play a role so that the robot will know how to act without predefined scenarios. Moreover, the speech domain should become broad enough to make sure that the robot can reply to most of what he is receiving.

Figure 9. Results.

References

[1] M. A. Goodrich1 and A. C. Schultz, "Human–Robot Interaction: A Survey," *Foundations and TrendsR in Human–Computer Interaction,* vol. 1, no. 3, pp. 1-20, 2007.

[2] D. Feil-Seifer and M. J. Matarić, "Human–Robot Interaction," *Encyclopedia of Complexity and Systems Science,* R. A. Meyers (eds.), Springer New York, 4643-4659, 2009.

[3] E. Guizzo, "Hiroshi Ishiguro: The Man Who Made a Copy of Himself," April 2010. [Online]. Available: http://spectrum.ieee.org/robotics/humanoids/hiroshi-ishiguro-the-man-who-made-a-copy-of-himself.

[4] N. Kageki, "Meet Affetto, a Child Robot With Realistic Facial Expressions," Feb 2011. [Online]. Available: http://spectrum.ieee.org/automaton/robotics/humanoids/meet-affetto.

[5] Hanson Robotics. "Face Robot," 2015. [Online]. Available: http://www.hansonrobotics.com/robot/han/

[6] K. Berns and J. Herth, "Control of facial expressions of the humanoid robot head ROMAN," *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3119-3124, 2006, October 9-15, Beijing, China.

Model United Nations
MUN Simulation Software

Joumana Samir Moodad
Department of Electrical and Computer Engineering
Lebanese American University (LAU)
36 Byblos, Lebanon
joumana.moodad@lau.edu

Abstract - **This paper describes the Model United Nations (MUN) Simulation Software which is an automated software tool dedicated to the Lebanese American University (LAU) – MUN organizing committee members. The software is used in the Global Classrooms International Model UN Conferences for Middle Schools and High Schools that will take place at the Grand Hyatt and the UN Headquarters in New York City on March 31 – April 2, 2016 and May 12 – May 14, 2016 respectively [1]. It allows LAU-MUN organizing committee members to more effectively organize and process the data needed to launch MUN simulation exercises and conferences, including school/country data storage, organization, and automated preference matching using optimization solutions based on the transportation problem.**

I. INTRODUCTION

Model United Nations is an internationally known simulation model of the UN and its agencies. Global Classrooms is the United Nations Association of the United States of America's (UNA-USA) flagship education program, a program of the United Nations Foundation. Through this program, students of middle schools and high schools explore current world issues, improve their skills, and acquire necessary attitudes for active citizenship. Over the past decade, Global Classrooms has worked in 24 major cities around the world, and has engaged more than 25,000 students and teachers annually [2]. The LAU MUN High School program was first launched in 2005-2006. For the past 10 years, more than 1,000,000 hours of teaching peace building and diplomacy were offered to a total of 16,000 Middle School and High School students from around 174 schools throughout Lebanon and from selected schools in the Arab world [3].

A. Challenges

When organizing an MUN exercise, organizing committee members are faced with many data processing challenges, mainly: i) acquiring information regarding participant schools from all around the world, in a timely and organized manner, ii) storing and organizing school information, including: each school's preferences (in terms of the countries – and/or committees - to be delegated) as well as the number of prospective participants, iii) matching school preferences with available countries (committees) profiles in order to assign each school one or more countries, and/or vice versa. The existing solution used by the organizing committee to handle the mentioned challenges consisted of a bunch of disconnected manual and semi-automated processes, which induced various problems/bottlenecks. For instance, information was acquired through the use of Google documents filled using a primitive Web interface, without any linguistic pre-processing or data cleansing. This produced data full of typos, mistakes, and sometimes ambiguities (e.g., school names or country names written differently). Concerning data storage and organization, it was typically handled in spreadsheets which are disregarded at the end of each simulation exercise, making way for the upcoming one. As for school/country matching and preferences, the task was conducted manually by organizers, which required considerable manual effort and consumed a huge amount of time.

B. Objectives

The goal is to provide a fully automated software-based solution which addresses the three main limitations mentioned above: i) designing a dedicated Web interface, using the sophisticated technology in linguistic pre-processing techniques to clean the data prior to storage, ii) designing a dedicated database structure, handled via a light weight Web-based DBMS (Database Management System) in order to properly store normalized data, iii) allowing to keep track of: not only the current data concerning the current simulation exercise, but rather of historic data regarding previous exercises, allowing advanced applications such as temporal querying (querying the past) and data mining functionality (e.g., discovering and analyzing school preference/country selection patterns), iv) implementing an automated solution to perform school/country matching, adapting existing assignment and/or transportation problem algorithms to handle the specific data at hand.

In the remainder of the paper, we first describe the context of the project (Section II) and its requirements (functional and non-functional in Section III), then we present the proposal (Section IV) including implementation details (Section V), as well as preliminary experimental evaluation (Section VI), before wrapping up with future improvements (Section VII).

II. CONTEXT

The simulation programs have different conditions and requirements. The software is meant to operate on the following programs: Model United Nations (MUN), Global

Classrooms International Model UN (GCI MUN), and Model Arab League (MAL).

A. *Model United Nations Conferences (MUN) & Global Classrooms International Model UN Conferences (GCI MUN)*

In each conference, schools are assigned countries to represent them. The rules and regulations for country assignments: i) the world is divided into 5 regions, ii) each school has to rank these regions from 1 to 5, iii) then, a school needs to choose 2 countries/region and rank them as priority 1 and priority 2, iv) a school has to provide the total number of students that they are applying.

Based on the data provided by each school, the organizers of the conference start by: i) First Come First Serve Basis Rule: the school that registers first will be considered first (through checking the timestamp), ii) Regional Preference Rule: the region that has the highest rank (rank of 1) is first considered, and the region that has the lowest rank (rank of 5) is least considered, iii) Country Preference per Region Rule: the country that has the highest priority (priority #1) in the region is first considered, if the conditions are not satisfied, the second country within the same region (priority #2) is then considered, iv) Delegation Size Match Rule: the number of students applying per school should be matching with the number of representatives of the country it will represent.

One of the main conditions in MUN and GCI MUN is that countries cannot be assigned twice; a country should be assigned to one and only one school. A school should also get only one country.

B. *Model Arab League (MAL)*

The MAL conferences follow the same procedure of MUN and GCI MUN conferences, however with some minor differences in the assignments conditions.

In MAL, a country can be assigned several times to several schools. In other words, a country that has been assigned previously and still lacks a certain number of representatives, can be assigned a second time to a different school than the one it has been assigned before. Similarly, a school can be assigned several countries. For instance, if a school has been assigned to one country which requires lesser number of representatives than the number of students applying, the school can be assigned a second (third, fourth, etc.) country.

Moreover, MAL conferences differ from MUN and GCI MUN conferences in the way schools provide their priorities. In MAL, there is no concept of regions; in other words, a school directly provides ten countries prioritized from the highest to the lowest instead of ranking five regions and choosing two countries within a region.

III. REQUIREMENTS

This section specifies the functional and non-functional requirements of the software. This is done through establishing the services that the users require from the system, in addition to the constraints under which the system operates.

A. *Functional Requirements*

a. *School Application Interface*

Functional requirements are specified in terms of both user and system requirements. One of the requirements which are common to all types of conferences is to have an application interface that is sent to schools to collect data from them and store this data in a database.

b. *MUN and GCI MUN Application Interface*

This interface is to be used by the MUN and GCI MUN organizers to do the school-country assignments. The interface should have as input fields: i) An excel sheet of all applying schools and their regional and country priorities, ii) An excel sheet of all schools with their corresponding number of applying students, iii) An excel sheet of all countries and their corresponding number or representatives, iv) location of the output country-assignment sheets (output excel sheet of first phase where regional and country priorities are considered, and output excel sheet of second phase where only matching size conditions are considered), v) list of all remaining countries not yet assigned, vi) list of all assigned countries, vii) list of remaining schools that couldn't be assigned due to exceptional constraints (Appendix – Fig.4).

c. *MAL Application Interface*

This interface is to be used by the MAL organizers to do the school-country assignments. These assignments are specified in an output excel sheet and have several criteria. The interface should contain input fields for: i) input excel sheet of all applying schools and their country priorities, ii) input excel sheet of all schools with their corresponding number of students, iii) input excel sheet of all countries with their corresponding number of representatives, iv) location of the output excel sheet of the school-country assignments, v) threshold for schools: input from user specifying the minimum number of extra students needed for the school to be assigned a second country, vi) threshold for countries: input from user specifying the maximum number of vacant places allowed for a country to stay functioning/represented, vii) maximum number of countries per school: input from user specifying the maximum number of countries that can be assigned to a school, viii) maximum number of schools per country: input from user specifying the maximum number of schools a country can be assigned to, ix) list of all assigned countries, x) list of all unassigned countries, xi) list of countries which are not totally filled; they still have vacant places above the countries threshold, and thus, can be assigned to other schools (Appendix – Fig. 5).

B. *Non-Functional Requirements*

The non-functional requirements in our design fall under product and organizational requirements.

a. *Product Requirements*

When designing the software, we should consider the space and time requirements. The execution time of the application

should be in the order of seconds. Regarding the efficiency and performance requirements, it should be able to deliver the maximum amount of information using a certain limited number of inputs. In addition, the software should be reliable in terms of country-school assignments and can have a margin error of 1/100 schools. Another important requirement which we should consider is the usability of the application; it should provide the necessary data through a user-friendly interface that can be easily used by users of different backgrounds. Last but not least, since the data is to be stored in a database system and since all the computations and results are based on this data, we need to make sure that our system complies with the security requirements.

b. Organizational Requirements

The organizational/external requirements include interoperability requirements. Since the software is to be used on different platforms and systems (MUN, GCI MUN and MAL offices), we should ensure that it can operate on any system that has Java Runtime Environment.

IV. PROPOSAL

To solve the school-country assignment problem, we referred to the Transportation and Assignment Algorithms. We first identified the different basic feasible solutions which include North West Corner Rule, Minimum Cost Method, and Penalty Cost Method (Vogel's Approximation Method). Then, we looked at the optimal solutions which include Stepping Stone Method and MODI (Modified Distribution Method) or U-V Method.

After doing several simulations based on the different transportation algorithms, we decided to use the Minimum Cost Method which yielded the lowest costs and best matching of schools-countries [5-11].

The transportation problem (TP) is concerned with shipping a commodity between a set of sources (e.g. manufacturers) and a set of destinations (e.g. retailers/warehouses). Each source has a capacity dictating the amount it supplies. Each destination has a demand dictating the amount it receives (Appendix – Fig.2). In a balanced system, the total number of demand at all destinations is equal to the total amount of supply at all sources. The purpose of this algorithm is to determine the amounts to be shipped between each source and destination with a minimal cost while meeting demand and supply conditions [4].

Table 1: Cost Matrix of a Balanced Transportation Problem

Factories	W1	W2	W3	Supply
F1	6	8	4	14
F2	4	9	3	12
F3	1	2	6	5
Demand	6	10	15	31

Our school-country assignment problem can be mapped to a transportation problem where schools are suppliers and countries are demands (Appendix – Fig.3). The costs for filling the school-country matrix are based on the regional and country preferences provided by the schools.

A. Calculating the Costs – Basic Solution

Since filling the cost matrix is based on the preferences provided by each school, we have to compute the preferences. To do so, and since regions have higher priorities than countries, we used the following method:
"Rank of Region" followed by "Rank of Country/Region"
Example: **51** means it is the region with rank 5 and the country with rank 1 within this region.
Below is the table of preferences/cost mappings.

Table 2: Preference/Cost Mapping

Region Rank	Country Rank/Region
1	1
	2
2	1
	2
3	1
	2
4	1
	2
5	1
	2

Preferences are associated with costs. They are inversely proportional to each other; the higher the preferences, the lower the costs. For instance, the lowest value in this table is **11**, and it refers to region with rank 1 and country with priority 1 within this region → Best preference is mapped to lowest cost. Similarly, the highest value in this table is **52**, and it refers to region with rank 5 and country with priority 2within this region → Worst preference is mapped to highest cost.
For the countries that are not chosen by a school as preferences, they will be assigned very large costs "Infinite" or "Inf" values in the corresponding row of the school.

B. Calculating the Costs – Extensible Solution

This solution is more flexible in terms of setting up and changing the priority of each factor. We can add at any time as many factors as required. It is based on a weighted summation. Each factor is associated a certain cost and weight. The values of the costs are normalized. Currently, we can consider three factors: i) RC: Region Choice, ii) CC: Country Choice, iii) SM: Students Matching Size. The weighted sum is then given by:

$$\sum W_j * Cost(Feature)_j = W(RC)*Cost(RC) + W(CC)*Cost(CC) + W(SM)*Cost(SM)$$

such that $\sum W_j = 1$ and $(W_1, W_2, …, W_n) >= 0$

In this section, we will discuss the Minimum Cost Method and how it was applied to the school-country assignment problem. This being said, we will study the two solutions developed: Excel sheet based solution and database solution.

A.   Minimum Cost Method Implementation

The minimum cost method is a basic feasible solution which takes into account the cost variables of a problem; this makes it more preferred than the North-West Corner Method and other methods. It consists of a cost matrix which has suppliers in its rows and demands in its columns. In our problem, the suppliers are the schools, and the demands are the countries. The last row of the matrix is filled with the value **1**, and the last column of the matrix is also filled with the value **1**. If a school/supplier has a value 1 in its last column, this means that this school has not yet been given any country. If a country/demand has a value of 1 in its last row, this means that the country is not yet assigned to any school. These values are not representing the actual numbers of students of a school and representatives of a country; they are used to flag the fulfillment of a school and the availability of a country. Please note that, what we discussed so far applies to a balanced transportation system. A balanced transportation system is when the total number of suppliers is equal to the total number of demands [13]. If the system was unstable, then we need to add a dummy column or a dummy row (based on which of the total number of suppliers or total number of demands is lower).

a.   From Unbalanced to Balanced System: Dummy Row

If the total number of suppliers is lesser than the total number of demands, in other word, if the total summation of **ones** in the last column is lesser than the total summation of **ones** in the last row, we need to add a dummy row that will have in its last column the difference of these two summations. This way, we recover for the deficiency in the number of suppliers by adding dummy suppliers/dummy schools which can supply as many demands based on previously calculated difference. In our problem, the dummy schools are added because not all countries need to be assigned to schools (most of the cases, the list of countries worldwide exceeds the number of schools who apply); thus, these dummy schools will not be considered.

b.   From Unbalanced to Balanced System: Dummy Column

If the total number of suppliers is greater than the total number of demands, in other word, if the total summation of **ones** in the last column is greater than the total summation of **ones** in the last row, we need to add a dummy column that will have in its last row the difference of these two summations. This way, we recover for the deficiency in the number of demands by adding dummy demands/dummy countries. In our problem, the dummy countries are added when there are no more countries (all countries have been assigned) and still there are schools not assigned any countries yet. These schools are then assigned dummy countries which do not exist. This case is rare; to have more schools than the total number of countries worldwide.

c.   Steps and Procedure

After setting up the structure of the matrix and making sure the transportation system is balanced, we can proceed with the algorithm [12].

The first step to do is to find the cell with the minimum unit cost in the entire matrix. If there is a tie, we break it arbitrarily. In our case, the tie will not be broken arbitrarily, but we will take the first occurrence of the minimum unit cost because time is considered in our design.

The second step is to cross out the column and row which correspond to the minimum cost unit cell after checking if the supply and demand conditions are satisfied. In other words, we need to check whether the school still have students and whether the country to be assigned still have vacant places for representatives. In our design, the school should have in its last column a value greater than or equal to **1**, and the country should have in its last row a value greater than or equal to **1**. If this condition was not met, we will need to restart the process and find the next minimum unit cost cell.

After doing the assignment of the row and column corresponding to the minimum cost unit cell, we need to take them out from the matrix and update their corresponding supply and demand values. To take them out of the matrix, we replace the values of the cells in this column and row by large values, "infinite" values, so that they will not be chosen anymore when looking for the minimum cost unit cell. Then, we update the supply and demand through subtracting **1** from each of them.

This process is repeated until all rows and columns are assigned. To keep track of the assignments, when doing them, we can replace the value of the cell corresponding to the assigned row and assigned column by a certain fixed constant unused value. This way, when we want to get the assignments after we finish the run of the algorithm, we can search for this specific value and match the rows and columns to get the assignments.

d.   Simplified Version of the Minimum Cost Method

The Minimum Cost Algorithm that was discussed in the previous section does not consider the timestamp criteria. It does the assignments after checking the whole matrix for the minimum cost unit; however, in some cases, this will not achieve the correct assignments. Let us consider the following scenario: School X chose a country A with a cost C1 at a timestamp T1; school Y chose the same country B with a cost C2 < C1 at a timestamp T2 > T1. Based on the Minimum Cost Algorithm, Country A should be given to school Y since it has a lesser cost than school X, but if we consider the time condition which has higher priority in our design, Country A **might** be assigned to school Y if school X was assigned another country which it chose with higher priority. The simplified version of the algorithm that is used in our transportation problem is to search row by row for the

minimum cost unit cell, and assign the schools sequentially; this will ensure that a school which comes first has higher probability of being assigned countries from its list of preferred countries than a school which comes last.

B. Excel Sheet Based Solution

This solution is based on providing the input and getting the results through Excel sheets. The desktop application is developed using Eclipse IDE and using Java and JavaFX programming languages. For all types of conferences, we need as input three excel sheets which will be processed differently.

a. MUN and GCI MUN Excel Sheets

The excel sheets needed to run the application as mentioned in section III, part A., include the sheet of applying schools and their country preferences. Please refer to the appendix, Table 3, to see an example of the input excel sheet of all schools applying for the GCI MUN Conference. Once the excel sheets are entered, we read the data and store it in files.

We process the input excel sheet of the schools to get an excel sheet with a list of ten countries listed from the highest to the lowest priority instead of having five regions ranked from 1 to 5 and two countries in each region.

After converting the excel sheet, we apply the Simplified Version of the Minimum Cost Algorithm. We get an output excel sheet with the preliminary school-country assignments. This is the output of the first iteration. If there are schools that were not assigned because we could not find for them countries from their preferred list which have same number of representatives as their number of students, they will be the input of the next iteration. In this new phase, we no longer consider the country preferences of a school, but we only consider the matching size criteria. In this second phase, we use the Minimum Cost Algorithm as well; however, this is the original version (not simplified version) of the algorithm since the time factor (which school applies first) is no longer considered. To create the structure of the new cost matrix, we take the remaining schools that need to be assigned countries and place them in rows; we take the remaining countries (countries that have not been assigned in the first phase) from the total list of countries and place them in columns. To fill the costs inside this matrix, we calculate the difference between each column (number of representatives per country) and each row (number of students per school):

$$\rightarrow \textbf{Difference} = \textbf{Representatives}_{\textbf{Country}} - \textbf{Students}_{\textbf{School}}$$

If this difference is **zero**, we place a 0 in the corresponding cell; if it is positive, we place it itself in the cell; otherwise, we place a very large value in the cell to prevent the assignment of the specified country to the school (since it is not allowed to assign a country which has lesser number of representatives than the school's students).

After running the algorithm and getting the assignments of this second phase, we can provide the users with list of countries that are still not assigned to any school; in addition,

we can provide the list of countries that have been assigned so far; and most importantly, we need to provide the list of schools that did not find any country with number of representatives less than or equal to its number of students. These schools will be assigned manually by the user (exceptions handled by the user's decision).

b. MAL Excel Sheets

The procedure for the MAL is similar to the MUN and GCI MUN, with slight changes:

i) The input excel sheet is already given as a list of ten countries (no region ranking).

ii) We can have several countries assigned to the same school (the maximum number of countries assigned per school is specified by the user); similarly, we can have several schools representing the same country (the maximum number of schools representing a country is specified by the user).

iii) The outputs of the first iteration are the preliminary country-school assignments, list of countries that are assigned but still have vacant places for representatives, list of schools that still have students not represented.

iv) The second phase consists of having recursive calls until all schools are completely assigned if the school threshold specified by the user is 1. For instance, if this threshold was 3, a school will not be assigned another country unless it has more than 3 students not represented. Similarly, we have a threshold for the countries; for example, if the country threshold is 5, then a country having lesser than 5 remaining representatives cannot be assigned to a second school.

v) The iterations in the second phase follow the Minimum Cost Algorithm as well; defining the structure and calculating the differences are done in the same way as MUN and GCI MUN; but, in this case, we fill in all cells the absolute value of the differences.

C. Database Solution

The database solution consists of creating a database that will contain all the data entered by the schools through the school application interface, and the control data entered by the MUN, GCI MUN, and MAL organizers (Appendix – Fig.1). Thus, other than the school interface, we need to have another interface for the database dedicated to the organizers. Through this DB, we can extract the input excel sheets that were provided as inputs previously. This way, the organizers can either run the software using external sheets or run it with a button that extracts directly the data regarding the countries and schools from the database. Using this solution not only allows for efficient and better security, but also it can be used for data mining, statistics and visualization techniques.

VI. TESTING

The testing procedure consists of checking if the functional and non-functional requirements of the system are met or not. To do so, we ran a multiple tests and we provided some of the results in the appendix section.

Table 4 in the appendix provides the school-country assignments of the first phase when running the GCI MUN application. The first two columns of this table represent the name and number of students of a school, and the next two columns represent the name and number of representatives of the country that is assigned to a school; the last column represents the difference between the number of students and number of representatives. In the first phase, the difference in all rows is zero since all schools got countries with exactly matching numbers. However, if we look at table 5 which represents the results of the second phase, we can see that some rows have negative values since in worst case scenarios, schools are assigned to countries which have larger number of representatives than their number of students.

If we look at table 6 which represents some of the school-country assignments of the MAL conference, we can notice that the table has a different structure than tables 4 and 5. The first five columns represent the same entities as before; but, the new added columns represent the additional countries that are assigned for the schools which still have remaining students. The number of these columns differs each and every time we run the software since they are added recursively, based on the need of the schools, and the maximum allowed number of countries per school. We can notice that, starting from column 6, we have the notation "country A/x" which indicates that after assigning country A, the school still has x students to be represented. The value of x is obtained by subtracting the number of representatives of country A from the previous value of remaining students. Please note that, for example, in the fifth column and fifteenth row, we have a value of **-3** which means that country **Ghana** still has three places for it to be represented completely. If the country threshold is lesser than three, we disregard it; otherwise, we can assign Ghana to another school.

VII. CONCLUSION

United Nations (MUN) is an international simulation program in which students, known as delegates, represent countries from all over the world, do research, formulate positions, and conduct debates regarding regional and global issues. Through this academic competition, students get to participate in simulated UN diplomatic exercises, allowing them to acquire many skills including performing information research, public speaking, debating, critical thinking, team work, writing formal memos, and acquiring leadership skills. One of the upcoming MUN program initiatives is titled *Global Classrooms LAU MUN for Middle Schools*, co-organized with the Lebanese American University (LAU), and which is going to take place in New York in March 2016.

The Model United Nations Simulation Software allows helping organizing committee members to organize and process the data needed to launch MUN simulation exercises and conferences. It is used to come up with school-country assignments for different types of conferences: MUN, GCI MUN and MAL conferences, using adapted transportation problem algorithms, namely the Minimum Cost optimization method. Each type of conference (e.g., middle schools, or high schools) has its own rules and requirements, with different interfaces and algorithmic designs. Two solutions for data acquisition have been suggested: External Excel sheets and internal imports from a database. The improvements we are aiming to do on the software include developing a mobile application for schools registration, and analyzing the data in the database over the years through visualization and data mining techniques.

REFERENCES

[1] Marković, M. (2015, October 6). 2016 global classrooms international model united nations conferences. Retrieved from //www.munplanet.com/articles/global-classrooms-lebanese-american-university/2016-global-classrooms-international-model-united-nations-conferences-q-and-a

[2] (n.d.).Retrieved March 20, 2016, from https://en.wikipedia.org/wiki/Model_United_Nations..

[3] Global Classrooms International Model UN Conferences. (2016). Retrieved March 20, 2016, from http://www.gclaumun.org/GCI/gci.html.

[4] Maddah. (2005, December 20). the transportation problem (1). Retrieved from http://staff.aub.edu.lb/~bm05/ENMG500/Set_7_TP_a.pdf.

[5] Shafaat, A., & Goyal, S. K. (1988). Resolution of Degeneracy in Transportation Problems. Palgrave Macmillan Journals, 39(4), 411-413. Retrieved from http://www.jstor.org/stable/2582121.

[6] Hung, M. S., ROM, W. O., & Waren, A. D. (1986). Degeneracy in transportation problems. Discrete Applied Mathematics, 13, 223-237. Retrieved from http://www.sciencedirect.com/science/journal/0166218X/13.

[7] The Assignment Problem and the Hungarian Method. (n.d.). Retrieved March 21, 2016, from http://www.math.harvard.edu/archive/20_spring_05/handouts/assignment_overheads.pdf.

[8] Carnegie mellon university, G.A., A.S., M.D. (2007). The Dynamic Hungarian Algorithm for the Assignment Problem with Changing Costs. Retrieved 21 March, 2016, from https://www.ri.cmu.edu/publication_view.html?pub_id=5805.

[9] The Transportation and Assignment Problems. (2008). *College of Management, NCTU*. Retrieved March 21, 2016, from http://ocw.nctu.edu.tw/upload/classbfs121001503719748.pdf.

[10] Linear Programming Notes VIII – The Transportation Problem - University of California, San Diego. (n.d.). Retrieved March 21, 2016, from http://econweb.ucsd.edu/~jsobel/172aw02/notes8.pdf.

[11] Winston, W. L., Venkataramanan, M., & Goldberg, J. B. (2003). *Introduction to mathematical programming: Operations research* (4th ed., Vol. 1). Place of publication not identified: Duxbury.

[12] Transportation Problems. (n.d.). Retrieved March 21, 2016, from http://personal.maths.surrey.ac.uk/st/J.F/chapter7.pdf.

[13] Balance An Unbalanced Transportation Problem By A Heuristic approach. (n.d.). Retrieved March 21, 2016, from http://www.academia.edu/6732317/Balance_An_Unbalanced_Transportation_Problem_By_A_Heuristic_approach.
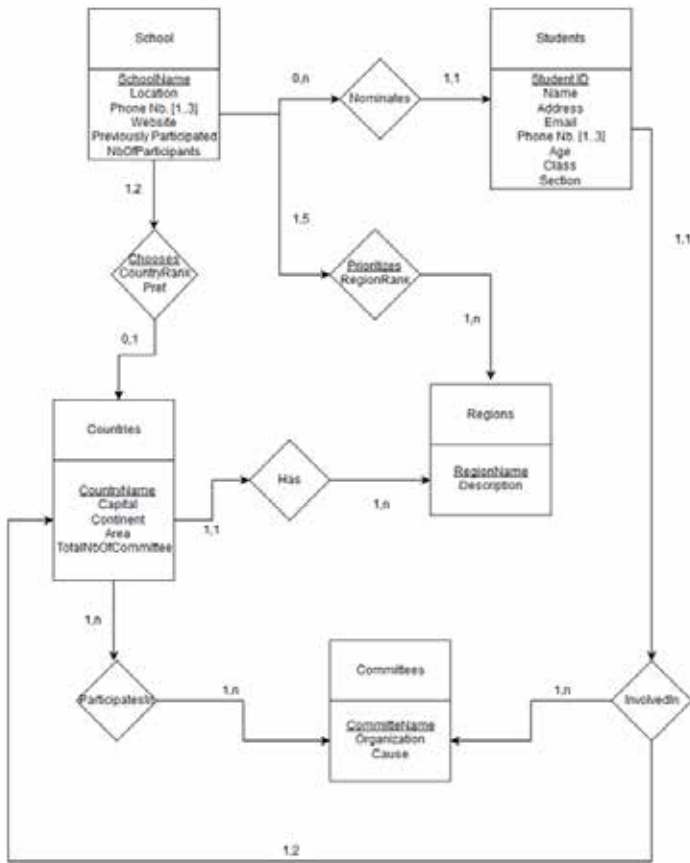
Figure1: Database Conceptual Data Model – Shema Level - Entity Relationship Diagram

Figure 2: TP Network Representation

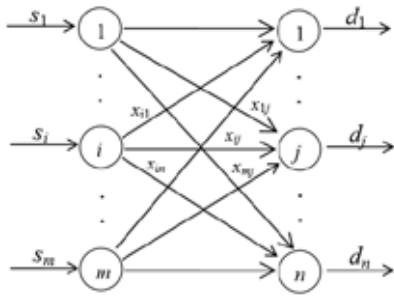

Figure 3: Example of a School-Country Matrix

Figure 4: GCI MUN Application Interface

Figure 5: MAL Application Interface

Table 3: Example of Input Excel Sheet for a GCI MUN Conference

Official School Name	What	Region 1: Africa	Region 2:	Region 3	Region 4:	Region 5:	First Priority in Africa	Second Priority in Africa	First Pric	Secondc	First P	Second	First P	Second	First Pr	Second Priority in West
School A	16	4	2	1	5	3	Kenya	Nigeria	India	China	Russia	Czech R	Brazil	Mexico	Sweder	Germany
School B	12	2	1	5	3	4	Mali	Cape Verde	United Ara	Oman	Russia	Russian	Jamaic	Saint Kit	Malta	San Marino
School C	16	5	2	4	3	1	Kenya	Egypt	Japan	China	Czech	Poland	Brazil	Mexico	Denmar	Germany
School D	16	5	2	4	3	1	South Africa	Egypt	China	Iran (Isla	Russia	Czech R	Brazil	Venezu	United K	Germany
School E	16	5	2	4	3	1	South Africa	Nigeria	Republic c	India	Czech	Croatia	Brazil	Mexico	German	Sweden
School F	12	2	1	5	4	3	Mali	Togo	United Ara	Mongolia	Bosnia	Bosnia a	Saint K	Jamaica	San Mar	Malta
School G	15	4	3	2	1	5	Guinea	Tunisia	Lebanon	Jordan	Ukraine	Romania	Argent	Chile	Luxemb	New Zealand
School H	15	5	3	4	1	2	Gabon	Guinea	Qatar	Lebanoc	Ukraine	Romania	Argent	Chile	Luxemb	New Zealand
School I	14	5	4	1	3	2	Seychelles	Burkina Faso	Thailand	Philippin	Estonia	Slovakia	Costa I	Dominica	Iceland	Iceland
School K	9	4	3	2	1	5	Lesotho	Lesotho	Kyrgyzsta	Kyrgyzs	Lithuan	Lithuania	Nicarag	Nicarag	Russian	Russian Federation
School L	16	4	3	2	5	1	South Africa	Morocco	India	Iran (Isla	Russia	Czech R	Mexico	Colombia	United K	Germany
School M	15	3	1	2	5	4	Tunisia	Rwanda	Saudi Ara	Jordan	Ukraine	Serbia	Argent	Chile	Luxemb	New Zealand
School N	16	5	2	4	3	1	South Africa	Cameroon	China	Japan	Poland	Bulgaria	Brazil	Mexico	German	France
School O	13	2	1	4	5	3	Ethiopia	Libyan Arab Jamahiriya	Kuwait	Indones	Hungar	Hungary	Trinidad	Guatemi	Ireland	Ireland
School P	15	5	4	1	3	2	Mozambique	Tunisia	Qatar	Jordan	Armeni	Ukraine	Chile	Argentin	Luxemb	New Zealand
School Q	12	5	1	2	4	3	Guinea-Bissau	Togo	United Ara	Maldives	Bosnia	Bosnia a	Jamaic	Barbado	Malta	San Marino

Note: Please note that due to privacy requirements, the actual names of schools are not disclosed.

Table 4: Part of the Output Excel Sheet for School-Country Assignments (First Phase) – GCI MUN Conference

Name of the School	# of Students	Country Assigned #1	RepresentativesPerCountry	(-): Deficiency (in Nb. of Country Representatives) (+): Surplus (in Nb. of School Students) 0: Satisfied
SABIS® International School - Adma	16	Russian Federation	16	0
Beirut Evangelical School For Girls And Boys - Rabieh	12	United Arab Emirates	12	0
Collège Des Saints-Cœurs Kfarhbab-Ghazir	16	Denmark	16	0
Collège Notre Dame De Lourdes - Jbeil	16	United Kingdom of Great Britain and Northern Ireland	16	0
College Des Saints Coeurs Ain Najm	16	Germany	16	0
West Hill College	15	Argentina	15	0
Collège De La Sainte Famille Française - Jounieh	14	Estonia	14	0
Al-Manar Modern School	9	Nicaragua	9	0
College Notre Dame De Nazareth	16	Czech Republic	16	0
Collège Protestant Français	15	Saudi Arabia	15	0
Makassed Islamic High School	16	France	16	0
Collège Notre Dame De Louaizé	13	Kuwait	13	0
Central College	15	Armenia	15	0
Collège Saint Joseph Antoura	12	Maldives	12	0
Rawdah High School	16	United States of America	16	0
Bayssour Official Secondary School	11	Singapore	11	0

Table 5: Part of the Output Excel Sheet for School-Country Assignments (Second Phase) – GCI MUN Conference

A	B	C	D	E
Name of the School	# of Students	Country Assigned #1	RepresentativesPerCountry	(-): Deficiency (in Nb. of Country Representatives) (+): Surplus (in Nb. of School Students) 0: Satisfied
Beirut Al Horj Public School	16	Benin	16	0
College Du Sacre Coeur-Gemmayze	16	Botswana	16	0
Collége De La Salle - Kfaryachit, Zgharta	14	Burkina Faso	14	0
Antoine Sisters School - Jamhour	13	Burundi	13	0
The International School Of Choueifat	16	Cameroon	16	0
Komayel Official School	8	Sao Tome and Principe	8	0
Christian Teaching Institute CTI	16	Chad	16	0
Daughters Of Charity School Ashrafieh	16	Congo (Republic of the)	16	0
Ecole Sainte Rita	13	Djibouti	13	0
SSCC Bouchrieh	8	Tajikistan	8	0
Eastwood College Mansourieh	2	Micronesia (Federated States of)	2	0
Lebanese American School Las	12	Cape Verde	12	0
Batroun Official Secondary School	6	Namibia	6	0
Maroun Abboud Official Secondary School	15	Mozambique	15	0
Notre Dame De Moussaitbeh	6	Saint Vincent and the Grenadines	8	-2
Saint Sauveur College	8	Guyana	8	0
International College	16	Côte d'Ivore	16	0
The National Evangelical School - Kab Elias	6	Timor Leste	10	-4
Cedars Cultural School	10	Solomon Islands	10	0
Al Imam Hasan High School	3	Kiribati	3	0

Table 6: Part of the Output Excel Sheet for School-Country Assignments – MAL Conference

A	B	C	D	E	F	G
Name of the School	# of Students	Country Assigned #1	RepresentativesPerCountry	(-): Deficiency (in Nb. of Country Representatives) (+): Surplus (in Nb. of School Students) 0: Satisfied	Country Assigned #2	Country Assigned #3
SABIS® International School - Adma	10	China	10	0		
Wellspring Learning Community	10	India	10	0		
Saint Joseph School-Cornet Chahwan	10	Germany	10	0		
Bahmuzine High School	10	Netherlands	10	0		
Bourj International College	10	Russian Federation	10	0		
Houssam Eddine Hariri High School	10	Estonia	10	0		
Universal College Aley	10	Brazil	10	0		
International Learning Community	10	Ethiopia	10	0		
Tripoli Evangelical School For Boys And Girls	10	United States of America	10	0		
Armenian Evangelical Torossian School	9	Ukraine	9	0		
Beirut Modern School	10	Algeria	10	0		
Antoine Sisters School - Mar Elias Ghazir	10	Australia	10	0		
Brummana High School	100	Japan	90	3	Congo (Republic of the) / 80	Togo / 70
Beirut Evangelical School For Girls And Boys - Rabieh	7	Ghana	10	-3		
Phoenix International School	10	Venezuela (Bolivarian Republic of)	10	0		
College Notre Dame De Lourdes - Jbeil	10	France	10	0		
Jesus & Mary School	10	Spain	10	0		
Daughters Of Charity School Dar En Nour	10	Colombia	10	0		
German International School Beirut	9	Central African Republic	9	0		
College Saint Joseph Miniara	9	Mali	9	0		
Collège Mariste Champville	9	South Africa	9	0		
Students' Paradise Secondary School	9	Egypt	9	0		
College Notre Dame De Balamand	50	Canada	9	41	Cameroon / 31	Nigeria / 21
Adma International School	9	Switzerland	9	0		
College Du Sacre Coeur-Gemmayze	10	Argentina	9	1	New Zealand / 0	

Non-invasive Prenatal Baby Monitoring Techniques

Ralph Deeba, Rashad Saab, Malek Srour
Department of Electrical and Computer Engineering
American University of Beirut
Beirut, Lebanon
rgd06@aub.edu.lb, rms78@mail.aub.edu, mms112@mail.aub.edu

Abstract—**Premature babies are normally born with several health problems and need intensive care during their stay in the hospital. More than 15 million babies are born too early every year in the U.S. alone. These babies stay in Neonatal Intensive Care Units (NICU) where sensors are used to continuously monitor their vital signs parameters such as heart rate, respiratory rate, SpO2 and temperature. Self-evidently, this monitoring is crucial but the main concern is the discomfort and skin irritation that potentially results from the mounted sensors on premature babies. To remedy the problem, this project discusses alternative non-invasive monitoring techniques to replace existing contact sensors. The feasibility of employing these techniques at the American University of Beirut Medical Center (AUBMC) is being investigated. Camera-based techniques used to detect the heart rate proved to be feasible. Heart rate can be obtained by simple processing applied on the photoplethysmogram signals. More precisely, within every heartbeat, the color of skin region varies where the absorption of green color by oxygenated and deoxygenated hemoglobin varies significantly. Thus monitoring the change in intensity of reflected green color from a region of interest on the subject's skin yields in the extraction of heart rate. In order to monitor the respiratory rate of these babies, a camera is used to track their chest movement. The respiratory rate can then be extracted from the rate of variation of the chest volume.**

I. Introduction

When a premature baby is born, s/he is subjected to multiple sensors via latches that measure heart rate, respiratory rate, arterial or central venous pressure, and other critical physiological information. These measurements are essential and inevitable and thus the baby is forced to endure the discomfort and irritation caused by the multiple apparatuses used for monitoring. For example, the tape used in temperature sensing leaves the skin wounded as described in [5]. In addition to suffering, prolonged usage of the ECG electrodes results in possible skin irritation effects [5]. Avoidance of as much sensors and direct contact with the baby is thus favored. The suffering of premature babies raises the voice to search for non-contact alternative techniques to monitor the inevitable medical parameters, as required per patient-case. Several techniques are available in the literature but they are not yet deployed in the medical field as a replacement to existing equipment. Health and lives are not to be subjected to technological "fallacies"; therefore, we aim to be able to get accurate, reliable, repeatable and reproducible measurements of the monitored vital parameters via non-invasive monitoring techniques that agree with the existing direct contact sensors used for measurement. Our project aims to contribute effectively to reduce the suffering and avoid skin injuries for premature babies monitored in a highly controlled environment at AUBMC. In this paper we first introduce video magnification techniques that shed the light on PhotoPlethysmoGraphic (PPG) signals for medical use, alongside a literature review of related work regarding extracting the heart rate and respiratory rate from the observed non-contact signal. Following that we describe our implementation of the system including the image processing techniques and experimental setup. Finally we will present our preliminary results and the methodologies we will use to assess and verify the effectiveness of the system.

II. Previous Work

Video magnification is a concept that came out in 2005 at MIT-CSAIL [6]. Several image-processing techniques allow the magnification of small movements that are not visible to the naked eye. Some examples of these motions are that corresponding to the pumping of blood by the heart or the movement of the chest due to the respiratory cycle. The concept has seen improvements over time and the literature is rich with various methods. There are three present methods with the first using motion detection and trajectory amplification [4], the second using a Taylor series expansion to manipulate the variations in RGB pixel values over time while introducing an amplification factor to amplify the results [1] and finally, the third relies on a phase based magnification rather than the amplitude in an attempt to avoid noise magnification [2].

The variations in the blood volume in a body segment alter the transmission and reflection of light from that segment. These reflected light signals are the PhotoPlethysmoGraphic signals (PPG) that carry relevant information for heart rate extraction. More precisely, the heart rate can be extracted from the time interval separating two consecutive peaks of the PPG [7]. This approach works under perfect conditions where there is no noise from external sources, which is unrealistic assumption in real-world applications. Aliased frequency components due to artificial light flickering were enough to deteriorate the results.

A real-world applicable approach from light modulation is the auto-regressive (AR) model method [7]. In this method, PPG signals are de-trended, band-pass filtered, fitted into an AR model and then cleaned from unwanted signals based on differential comparison against a reference that merely contain the same ?unwanted? signals. Heart rate information is then inferred from the processed PPG signal. The advantages of

this method lie in its robustness to noise [7]. By selecting two regions of interest, the subject and the reference region, it is thus possible to detect noise coming from unwanted frequencies and therefore cancel it from the subject?s region data. This technique also helps in avoiding quantization errors introduced by the Fast Fourier Transform (FFT). However, the use of sliding windows, which introduces a latency of half the window duration (typically 7.5 seconds) in measuring the heart rate.

Similar to the heart rate AR model approach and relying on the fact that the respiratory rate (RR) modulates the amplitude of PPG signals [7], the RR can be determined by processing the PPG signal. The latency of measuring the respiratory rate is half the window duration (typically 15 seconds) [7]. It is noteworthy to mention that the respiratory rate obtained via AR method provides promising results that are more reliable than the EQ-02 belt (EquivitalTMVital Signs Physiological Monitor EQ-02, Hidalgo, Cambridge, UK) [7]. One of the possible methods that could be used for non-invasive respiratory rate detection is based on depth image sensing. As its name indicates, depth image sensing relies on collecting information about the depth of the chest to infer the volume of the lungs and thus differentiate the processes of exhalation and inhalation. Afterwards, counting of the number of breathing cycles, consisting of an exhalation and inhalation, is adequate to determine the respiratory rate. The preceding approach was experimented in [9] where five sets of tests were performed. Thorough testing was conducted while taking into consideration different scenarios to verify the robustness of the method. The results showed an error with a maximum value of 0.533 respirations per minute between the measurements obtained by a spirometer and the values calculated by the algorithm. Although the results show a significantly low error, we are constrained not to expose the babies to any radiations projected by the RGB Depth camera on the body of the subject.

III. HEART RATE IMPLEMENTATION

A. Face Detection

In order to remove added noise and artifacts, there is a need to specify a region of interest. As revealed by the video magnification techniques described in [1], [2] and [4], this ROI is the face. Hence before processing the input from a web-cam, the algorithm has to detect the face and isolate the forehead. For this purpose we apply the Viola-Jones algorithm which uses Haar feature-based cascade classifiers object detection methods to detect the subject?s face in each frame [10]. This is a machine learning based approach where a cascade function is trained from a lot of positive images (images including faces) and negative images (images without faces). After defining the face position within a frame, we slice the detected face in order to extract the forehead which is at a defined position relative to the faces coordinate box provided by the face detection algorithm.

B. Blind Source Separation

The collected RGB values across frames are affected by noise, however they all have the common component of the heart rate frequency. Hence by performing independent component analysis (ICA), we are able to reduce noise and artifacts making it simpler to identify the dominating heart rate frequency. We opted for the Jade ICA implementation since it proved to offer fast and reliable results [8].

C. Extracting the Heart Rate from Independent Components

In order to extract the dominant frequency component we passed the second component of the Jade ICA result through a Fast Fourier Transform (FFT) filter to get the frequency components of the signal. The second component is selected because it almost always contains the best detectable and dominant heart rate frequency. The signal is also subjected to a band-pass filter with cutoff frequencies of 0.83 Hz and 2.66 Hz corresponding to 50-160 BPM to further reduce noise. Following this processing, we select the frequency with the highest power as the heart rate frequency.

IV. RESPIRATORY RATE IMPLEMENTATION

A. Optical Flow Computation

Optical flow depicts the pattern of visible motion of objects, edges and surfaces. Several methods are adopted in the literature to determine the optical flow. Optical flow analysis is based on the assumption that pixel brightness remains constant in a local neighborhood of consecutive frames in the video. Let $I(x, y, t)$ be the Intensity of a pixel at position (x, y) at time t; and that frame pixels advance with velocities v_x and v_y in directions x and y, respectively. Therefore, the displacement of the image after time dt is $v_x dt$ and $v_y dt$:

$$I(x, y, t - dt) = I(x + v_x dt, y + v_y dt, t) \qquad (1)$$

Considering that time is represented by frames, $dt = 1$, and velocity in area surrounding the monitored point is constant, the unknown velocities v_x and v_y are determined by minimizing the following expression

$$\sum_{x, y \in W(t)} \left(I(x, y, t - 1) - I(x + v_x, y + v_y, t) \right)^2 \qquad (2)$$

where $W(t)$ is the window surrounding the monitored point. The solution of this problem as derived in [11] is

$$\boldsymbol{v} = \boldsymbol{C}^{-1} \boldsymbol{d} \qquad (3)$$

where

$$\boldsymbol{C} = \begin{pmatrix} \sum_{x,y \in W(t)} I_x^2(x,y,t) & \sum_{x,y \in W(t)} I_x I_y(x,y,t) \\ \sum_{x,y \in W(t)} I_x I_y(x,y,t) & \sum_{x,y \in W(t)} I_y^2(x,y,t) \end{pmatrix} \qquad (4)$$

and

$$\boldsymbol{d} = - \begin{pmatrix} \sum_{x,y \in W(t)} I_x I_t(x,y,t) \\ \sum_{x,y \in W(t)} I_x I_t(x,y,t) \end{pmatrix} \qquad (5)$$

where I_x and I_y are the partial derivative of I with respect to x and y respectively.

B. Velocity Signal Processing

Velocities, being the signals that are processed, can be obtained from various places of the image. However, noise along with the motion are eminent in these places. To remedy this problem, velocities are obtained from different areas and averaged with respect to a calculated signal to noise ratio. The signal is modeled as

$$x_i(t) = s(t) + n_i(t) \qquad (6)$$

where n_i is the noise signal with variance σ_i^2 and s is the signal to be determined. Following the derivation in [11], the model in (6) is generalized to

$$x_i(t) = A_i s(t) + n_i(t) \qquad (7)$$

where the amplitudes A_i are such that

$$A_i = 2 \max_{k \in <k_1, k_2>} \{|X_{iFFT}(k)|\}/N \qquad (8)$$

where N is number of samples; k_1, k_2 are discrete frequencies corresponding to the frequency band lacking RR in which the noise is predominant.

The noise variance is estimated in the frequency band $L_1 - L_2$ as

$$\sigma_i^2 = \frac{2.k_f}{N.N_{FFT}} \sum_{k=L_1}^{L_2} |X_{iFFT}(k)|^2 \qquad (9)$$

where N_{FFT} is number of FFT points and k_f is the ratio of signal bandwidth to the bandwidth used for noise variance estimation.

C. Algorithm

The optical flow computations were implemented in Matlab where the image sequence was first converted to gray-scale. Then, image gradients were calculated in each frame using the Sobel filter. Afterwards, a sliding window of a tuned size was used to traverse each frame both horizontally and vertically, and for every window, the velocities v_x and v_y are computed. We plotted the derived velocity vectors on the original video sequence to observe them and noise was significant in some instances. We are currently working on implementing the velocity signal processing method ? described in B ? followed by a Band-pass filter to reduce the noise.

V. Experimental Setup

The system used consists of a web-cam (Logitech C920) set at a resolution of 640 x 480 and sitting on a tripod next to the incubator. The camera provides a top view of the baby and is connected to a laptop. Video recording are saved in chunks of five minutes each and saved locally on the laptop. The laptop processes the videos on the spot and saves the computed values in a file including time stamps of each value computed.

VI. Preliminary Results

The obtained results from the heart rate detection implementation was tested on adults prior to the setup at AUBMC. We used a digital automatic blood pressure monitor from Omron as a reference measurement. The proposed method results in an error of at most two beats per minute as a worst case in presence of normal room conditions. This is a promising evaluation as premature babies have thinner skin than adults and thus blood light modulation should be more evident with less noise levels.
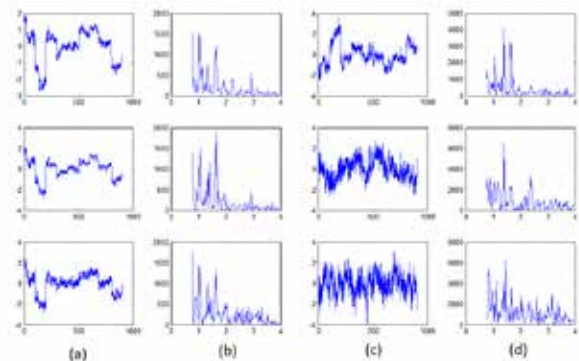
Fig. 1. (a) Raw RGB Data; (b) FFT of RGB signals; (c) Independent Components of the RGB sources; (d) FFT of the Independent Components, the peak here represents the heart rate frequency

VII. Future Work

A. Data Collection

To collect data from the employed Philips monitors, a serial MIB/RS232 card will be used.

B. Data Verification

The recorded videos and collected data will be aligned in order to verify results. Moreover, doctors will be reviewing the videos and identifying abnormal situations in the recordings that will further help in medical research.

VIII. Conclusion

This paper summarizes the importance of non-invasive monitoring techniques and presents some of the available methods, in addition to detailing a proposed implementation of heart rate and respiratory rate detection using blind source separation and optical flow calculations respectively. Preliminary results of the heart rate detection are promising. As well, respiratory rate detection that proved to. We have started collecting video recording of premature babies and looking forward to comparing the results of the proposed non-invasive techniques with the current sensors data extracted from the neonatal intensive care unit data centers. This work aims primarily to minimize pain that premature babies are exposed to and to provide pediatricians at AUBMC additional data that are needed for medical research purposes.

References

[1] H. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand and W. Freeman, *Eulerian video magnification for revealing subtle changes in the world*, vol. 31, no. 4, pp. 1-8, 2012.

[2] N. Wadhwa, M. Rubinstein, F. Durand and W. Freeman, *Phase-based video motion processing*, vol. 32, no. 4, p. 1, 2013.

[3] M. Poh, D. McDuff and R. Picard, *Non-contact, automated cardiac pulse measurements using video imaging and blind source separation*, Opt. Express, vol. 18, no. 10, p. 10762, 2010.

[4] C. Liu, A. Torralba, W. Freeman, F. Durand and E. Adelson, *Motion magnification*, vol. 24, no. 3, p. 519, 2005.

[5] Janani, *Temperature Sensors for Premature Babies - CAST — Center for Advanced Sensor Technology - UMBC*, Cast.umbc.edu, 2015. [Online]. Available: http://cast.umbc.edu/blogs/temperature/.

[6] People.csail.mit.edu, *Video Magnification*, 2015. [Online]. Available: http://people.csail.mit.edu/mrub/vidmag/publications.

[7] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. Clifton and C. Pugh, *Non-contact video-based vital sign monitoring using ambient light and auto-regressive models*, Physiol.Meas., vol. 35, no.5,pp. 807-831, 2014.

[8] A. Hyvrinen and E. Oja, *Independent component analysis: algorithms and applications*, Neural Networks, vol. 13, no. 4-5, pp. 411-430, 2000.

[9] F. Benetazzo, S. Longhi, A. Monteri and A. Freddi, *Respiratory rate detection algorithm based on RGB-D camera: theoretical background and experimental results*, Healthcare Technology Letters, vol. 1, no. 3, pp. 81-86, 2014.

[10] P. Viola and M. Jones, *Rapid object detection using a boosted cascade of simple features*, Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, 2001, pp. I-511-I-518 vol. 1, doi: 10.1109/CVPR.2001.990517

[11] T. Lukac, J. Pucik and L. Chrenko, *Contactless recognition of respiration phases using web camera*, Radioelektronika 24th International Conference, 2014.

RF Power Cube
Electrical and Computer Engineering
American University of Beirut
Ingrid Bakri Kasbah – Rayan Awad – Rena El Hajj
Hamra – Bliss Street – Beirut, Lebanon
imb03@aub.edu.lb – raa132@aub.edu.lb – rka21@aub.edu.lb

I. INTRODUCTION

RF Harvesting is one of the newest innovative techniques to produce DC voltage. Although it produces a low voltage output in addition to harvesting low energy power, RF harvesting is becoming the buzzword of people and will continue to evolve in the future with the advancement of technological changes. Very briefly, it consists of an antenna that harvests RF energy and transforms it to direct current energy through a high efficiency rectifier. There is more than one possible design for an efficient rectifier, in fact a lot of research and work have already been done about the topic, and many circuits with different output efficiencies were implemented. Although it can be implemented on low power devices only, it is a very promising topic that has grabbed many researchers' attention because of its innovative technique that can have considerably high efficiency.

Over the past decade, the dependence on wireless devices has significantly increased. These devices offer many important functions, however they have limitations mainly due to the short lifetime of their rechargeable and disposable batteries. RF Harvesting has been recently the subject of many research topics due to its technique of collecting "free" energy and reusing it, making efficient and continuous recharging possible. Beside what is already mentioned, eliminating the finite battery life and replacing it with an infinite one using the harvester is probably the greatest motivation that each and every one of us would like to have. It gives the possibility to produce direct current power out of thin air by harvesting and converting ambient radio frequency energy. As a matter of fact, radio frequency energy is broadcasted from TV stations, mobile phones, radio stations, Wi-Fi routers etc.…which makes it highly available. Efficiencies of harvesting high radio frequency power have reached 70% and sometimes exceeding that number [1]. Hence, finding better efficient solutions for the world to benefit from and obtaining a battery-less charger for our electronic devices is the priority of our project.

This paper allows the reader to know more about the design of a high efficiency rectenna that harvests radio frequency energy and converts it to direct current energy. The rectenna consists of an antenna that will harvest the desired frequency and a high efficiency rectifier that will convert the RF energy into direct current energy. Our system will harvest energy from 2.4 GHz, which belong to the Wi-Fi band of the IEEE standards IEEE 802.11b.

II. RECTENNA DESIGN

a) Antenna

Different antenna configuration has been reported in [2], [3] and [4]. In [2] a differential electromagnetically coupled Microstrip antenna with improved gain, efficiency and bandwidth is developed where the antenna has an efficiency of 80%, a gain of 8.5 dBi at center frequency and VSWR $\leq$ 2 in frequencies ranging from 870MHz to 1.05 GHz, but this antenna design is only applicable for differential I\Os. A higher gain is obtained in [3] where it presents a broadband 1x 4 quasi-Yagi antenna array with a bandwidth that goes from 1.8 to 2.2 GHz and high gains of 10.9dBi and 13.3dBi at 1.85GHz and 2.15 GHz respectively. This specific design requires a specific optimization of the driver, director, and reflector parameters. A new antenna design is introduced in [4] where 2 antennas were used and fabricated on the surface of a 3D cube that was printed using Inkjet printing to allow multi directional orthogonal harvesting. The antennas in this design are not designed as an antenna array like in [3], the output of each antenna is connected to the input of a rectifier circuit. Patch antennas were used to limit the electromagnetic coupling between the antenna and the electronics circuit board.

b) Rectifier Design

Different rectifier designs were introduced in [5], [3] and [4]. In [5] the rectifier design is a voltage doubler design, made up of two Schottky HSMS-285C diodes, the design gives a maximum efficiency of 24% and output voltage of 1.9 operating at 2.1 GHz, and a maximum efficiency of 18% and output voltage of 1.7 V operating at 2.45 GHz at 10dBm ambient power. Where as in [5] the rectifier design is made up of one Schottky diode Avago HSMS-2852, the measurement showed an efficiency of 40% and output dc voltage of 400mV at power density of 455W/m, in this design the efficiency is relatively higher than in [3] this is due the number of diodes used where less number of diodes

implies less non-linearity and thus more efficiency, as for the voltage it is smaller than in [3] this is due to the voltage doubler design that doubles the output voltage used in [3].

In [4] the design is made up of two voltage doubler circuits (see figure below) and the output is combined by a DC combiner that drives the load, each rectifier circuit is connected to an antenna that is fabricated on the surface of the cube.

III. PRELIMINARY TESTING

In order to make sure that we can obtain high efficiency at low power, two circuits were simulated using ADS. The first design is a tentative design that we simulated which is similar to our design, and the second design is a simulation of a design that is already found in one of the papers that we read in the literature review.

First design:

Two RF harvesters were simulated; each harvester is made up of two (SMS-7630) Schottky diodes from Skyworks, a matching circuit and a resistive load. One of the harvesters is simulated on 2.4GHz while the other is simulated on 5.8 GHz Fig.11 shows the slot antennas' top layers as well as the rectifying circuit diagrams for both targeted frequencies of operation. Fig.1 c, d, g, and h present the rectifying circuits efficiency and reflection coefficient results as predicted by Keysight's ADS. Fig. 12 presents the reflection coefficient analysis and the layout of the 2.4 GHz rectifier circuit.
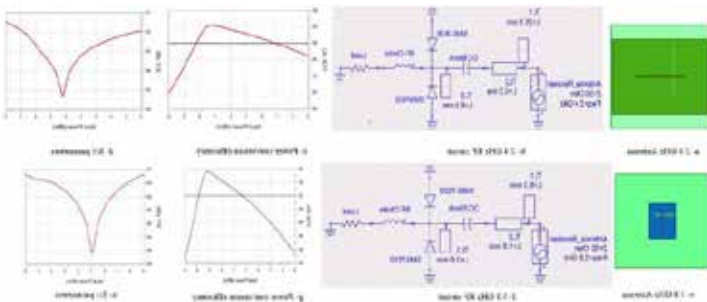
Figure 1: First design simulation

A re-simulation of [6] was done using ADS at a frequency of 900 MHz in order to check if previously implemented circuit results can be achieved.

We built the same circuit presented in [6] but instead of using conjugate input matching we used stubs to match the 50 Ω to Z_{in} of the rectifier circuit which is found using ADS. We used a capacitor as a DC block of 1μF and an inductor as RF choke of 1μH. After simulation the S11 parameter value is less than -16 dB as shown in Fig.2-a, which is a good result since it is below -10 dB. As for the efficiency, its value is slightly higher than 80% at 14 dBm input power as shown in Fig.15-b which is closer to the value in the article where 82% efficiency is obtained at 15.5 dBm input power as shown in Fig.10-b.

When the input power reaches 14 dBm the efficiency starts dropping (Fig.2-b) and the output voltage saturates at 3.3V as presented in Fig.16 since the diode starts to break down.

a- S11 analysis   b- efficiency analysis
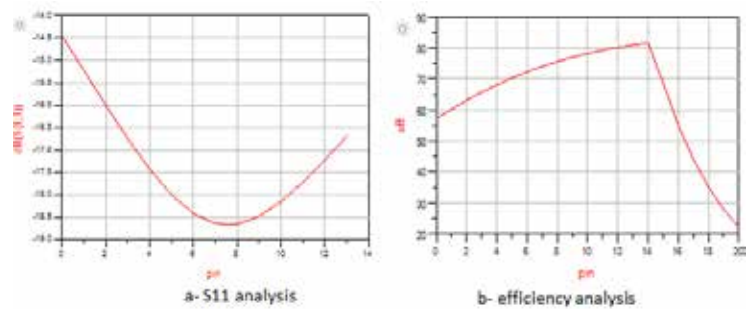
Figure 2: S11 and efficincy analysis of the resimulated circuit at 900 MHz frequency
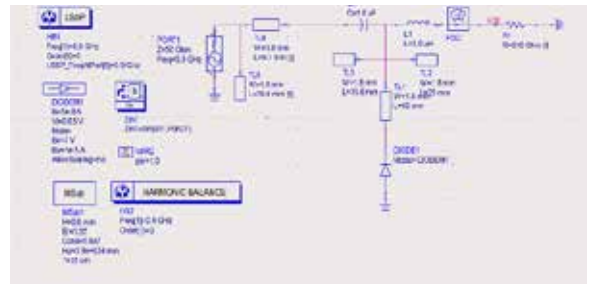
Figure 3:Re-simulated circuit on ADS

IV. OUR FINAL DESIGN

The design consists of four antenna components that are placed on the four sides of a foam based cube. Each antenna is connected to a rectifying circuit. The cube is of maximum size of 10 cm x 10 cm x 10 cm. The antennas are microstrip slot antennas that are built on top of a Rogers Duroid substrate, the Rogers substrate will be integrated with the hardened foam to form the walls of the cube. Four antennas are designed to operate at 2.4 GHz within the IEEE 208.11 b standard.

Each antenna will be connected to an RF harvester by an SMA Cable. The harvester which is of size 4 cm*2 cm is placed at a distance of a quarter wavelength behind the antenna. The RF harvester is composed of Schottky diodes from Skyworks (SMS-7630), a matching circuit and a resistive load. The output of the four harvesters will be combined in parallel and the output will be connected to a capacitor in order to store the dc power.

The total RF-to-DC efficiency of the system should be higher than 50%.

In the following sections, we will present the virtual simulations on ADS and HFSS, as well as the testing and measurements to validate our design.

a) HFSS Simulations

Sketches were developed on HFSS to obtain a more technical view for the schematic. We also drew the schematic of the antenna design although this is not the scope of the paper
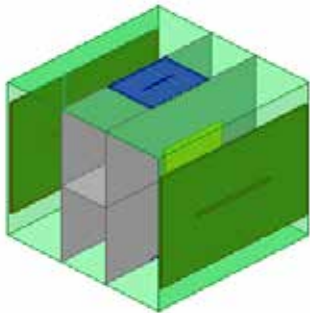
Figure 4 - Designed overall cube view

b) ADS Simulations:

The final design of the rectifier is shown in the figure below, as drawn on ADS:
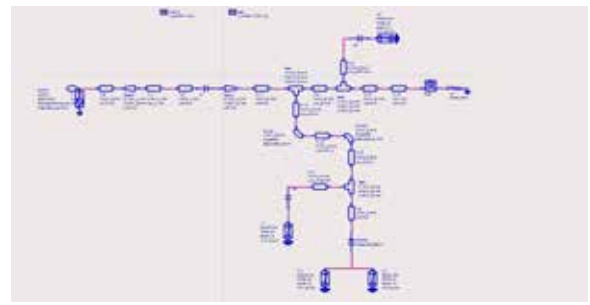
Figure 5: Final Design Schematic

The simulations on ADS resulted in the following:
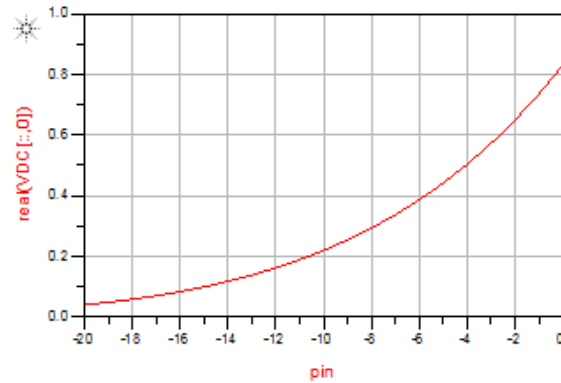
Figure 6: DC Voltage output of the final design
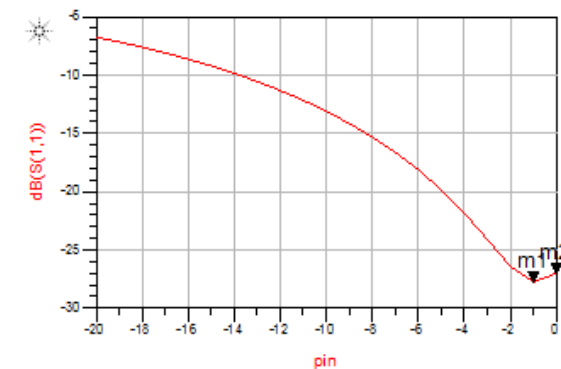
Figure 7: S11 parameters (dB)
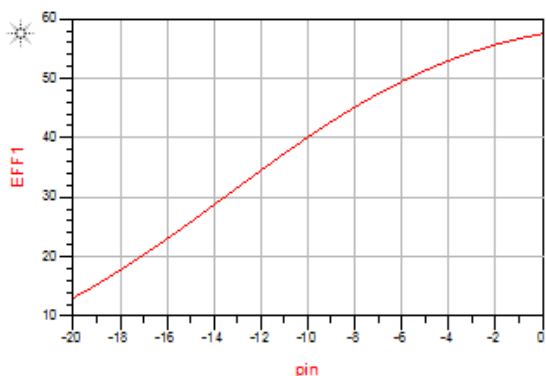
Figure 8 - Efficiency in %

c) Power management testing results:

In order to test the addition in parallel for the power management, a voltage supplier supplied 5V through a 1 kOhm resistor, we obtained a current of 5 mA. We can see the results through the figure below.

Figure 9- 5 mA Current

After setting 2 supply voltages to give 5 V each and placing them in parallel to obtain a 5V voltage after the combination and reducing the resistance to 500 Ohms, we realized that the total current by the circuit was 10 mA, yet the current of each individual circuit is 5 mA hence transferring maximum efficiency of the voltage source.

Finally, we can realize that the solution we wanted to test by dividing the resistor in half and putting the voltage sources in parallel is efficient and ensures a maximum efficiency transfer.

Figure 10- 5 V - 10 mA measurements

d) Rectifying Circuit testing results:

Table 1- Measurement results of efficiency

Power Detector	Power generator	efficiency
-30	-29.6	-
-25	-24.57	24%
-15	-14.64	29%
-10	-9.6	38%
-5	-4.56	42.2%
-3	-2.56	45%
0	0.46	50.06
2	2.48	54%
5	5.48	56.4

The resistance used in these measurements was a 2660 Ohm resistor.

Some of the results above were recorded as pictures

Figure 11- voltage output at 0 dBm and -30dBm

The following graph compares between theoretical value that we got in ADS and the actual values

Figure 12- Efficiency vs detected power graph

measured, we can obviously notice that both values are close to each other.

V. CONCLUSION

An efficiency of 50% or above is achievable with the proposed circuit in this paper. Four similar RF harvesters were designed and built. The outputs of each harvester will be combined by parallel addition, which is a feasible solution that was tested.

VI. REFERENCES

[1] V. Marian, B. Allard., C. Vollaire and J. Verider "Strategy for Microwave Energy Harvesting From Ambient Field or a Feeding Source", IEEE TRANSACTIONS ON POWER ELECTRONICS, VOL. 27, NO. 11, NOVEMBER 2012 4481

[2] M. Arrawatia, M. S. Baghini, , and G. Kumar " Differential Microstrip Antenna
for RF Energy Harvesting ", IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, VOL. 63, NO. 4, APRIL 2015 1581

[3] J. Kimionis*, A. Georgiadist, M. Isakov, H. J. Qi, and M. M. Tentzeris "3DlInkjet-printed Origami Antennas for Multi-direction RF Harvesting" School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0250

[4] H. Sun, Yo. Guo, , M. He, and Z. Zhong " A Dual-Band Rectenna Using Broadband Yagi Antenna Array for Ambient RF Power Harvesting", IEEE ANTENNAS AND WIRELESS PROPAGATION LETTERS, VOL. 12, 2013

[5] E. Khansalee, Y. Zhao, E.Leelarasmee and K. Nuanyai" A Dual-Band Rectifier for RF Energy Harvesting Systems" IEEE vol.14,201

[6] Jiapin Guo; Hongxian Zhang; Xinen Zhu, "Theoretical Analysis of RF-DC Conversion Efficiency for Class-F Rectifiers," in *Microwave Theory and Techniques, IEEE Transactions on* , vol.62, no.4, pp.977-985, April 2014

Smart Headgear: An Instrumented Orthodontic Appliance with a Mobile App

Ahmad Sharaf Aldine, Iman Sonji, Mohamad Shaaban
Department of Electrical and Computer Engineering
American University of Beirut
Beirut, Lebanon
aks13@mail.aub.edu, ims10@mail.aub.edu, mms113@mail.aub.edu

Abstract- **Achieving compliance in medicine is a well-known struggle that every physician probably has to deal with on a daily basis. This is especially true in the field of Orthodontics, where complying with the prescribed treatment gets tough for patients due to the physical inconvenience of most of the treatment methods, especially headgear wear. As a result, the aim of our project is to remedy the problem at hand and provide a means for monitoring pat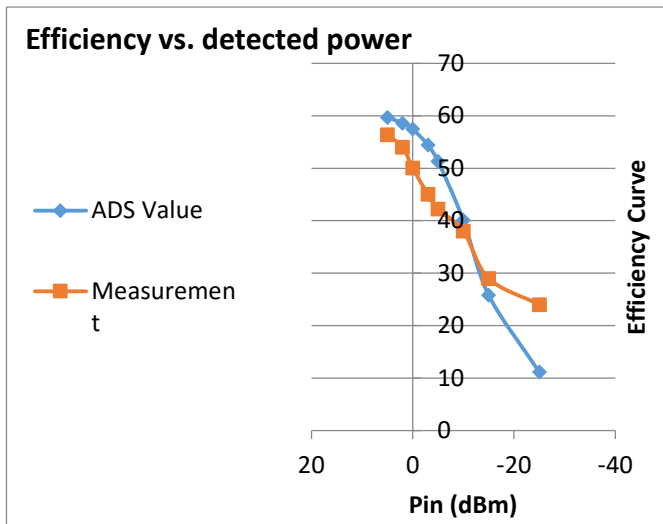ient compliance while promoting it at the same time. Monitoring the compliance is made possible through implementing a headgear timer device which is activated by sensor inputs that serve to make sure that no cheating or tampering is occurring. Enhancing the compliance is done through providing an Android game application that can be accessed during wear time. This headgear device is also connected to the game application via Bluetooth. This game incorporates the received wear time measurements by allocating in-game rewards. It also allows the patients to compete in terms of a leader board that keeps track of the scores. Moreover, this Android application is connected to an online database where the wear time measurements are periodically updated, enabling the orthodontists to monitor them remotely through a web application. This web application would have a list of the orthodontist's patients along with their profile details aside from being able to plot each patient's wear time versus the hours of the day. In this paper, we will present the research we have conducted on the topic of monitoring orthodontic compliance which will be followed by the design process of our device with its different components.**

I. INTRODUCTION

Headgears are orthodontic appliances that help take advantage of the growth of the jaws to correct bite problems in children. Usually, they are required to be worn for 14 hours per day. A study conducted in the Academic Centre for Dentistry Amsterdam in 2005 clearly highlights a certain compliance problem. The study involved 56 patients of mean age 12.89 years and whose headgear wear-time was monitored over a period of 29 days by using a headgear timer device. This device included a thermistor which sensed the patient's neck temperature during the wear time. The results show that the mean wear time after the 29-day period was 5.58 hours per day as opposed to the required 14 hours per day [1]. Normally, there are two types of compliance measures: subjective and objective. Subjective measures such as self-reporting or parent reporting tend to be unreliable and overestimated. Since most measures nowadays tend to be on the subjective side, orthodontists are faced with an additional struggle of not being

able to truly determine the amount of time their patients wear the headgear. Consequently, if this compliance problem is not resolved, the effectiveness of the orthodontic treatment will diminish notably and the treatment duration will increase. In some cases, this can also lead to surgical procedures and dental extractions which are two unfavorable outcomes. As a result, we believe this problem can be solved by combining a group of different sensors to ensure headgear wear and minimize patient tampering. These sensors, monitoring temperature, touch and head position of the patient, activate a timer which only increments if the sensor outputs fall within the acceptable ranges. Timing measurements are kept with their respective time stamps all of which are sent to an Android application. This data is uploaded to the patient's online profile which is remotely accessible for the doctor through a web application. In the mobile application, the time measurement is incorporated into the game so that patients with higher wear time could gain a competitive advantage. Therefore, our comfortably designed and packaged device with its Android application could solve this monitoring and compliance problem faced by orthodontists nowadays.

II. LITERATURE REVIEW

The first timing headgear that was established was by Northcutt and was produced by Aledyne Corporation in 1974 [2]. The timer was activated based on a pressure switch placed on the neck strap. However, this invention was easily circumvented by patients by placing something heavy on the pressure switch to activate the timer. This headgear was then updated by Northcutt in 1975 [3]. He added a pull switch on the pull of the strap beside the pressure switch to activate the timer. This updated headgear was found to increase patient motivation. Although Northcutt mentioned that the timer was 98% accurate, this accuracy is questionable in light of an investigation into the reliability of the timing headgear by Banks and Reed in 1987 [4]. The investigation revealed that 9/13 timers tested were less than 30% accurate. Thus, the mode of operation of this headgear is as follows. There are two switches used to activate the timer. The first one is a pull switch placed on the pull of the straps and the second is a pressure switch placed on the neck strap. Both have to be activated in order to activate the timer.

Another timing headgear was established by Malcolm Savage in 1982 [5]. The first approach for inventing this headgear was to use an E-cell timer based on a spring or a saliva switch. However, it was then canceled since it is easily circumvented by patients. This is because the spring switch can be activated by simply stretching the strap and saliva switch can be activated by salty water. A second approach was followed which is based on a soluble controlled release glass timing disc implanted in the removable appliances. This headgear indicated high levels of accuracy due to the accurate solubility rate measured. However, this timing headgear was proven to be ineffective later on due to issues related to short disc life and poor disc maintenance [6].

In "An Accurate, Inexpensive Headgear Timer", a third timing headgear is mentioned. It was invented by Steven L. Cureton in 1991 [7]. Time measurement was done using a small quartz calendar wrist watch. The timer is activated based on a mini lever switch placed on the spring of the headgear. When the straps of the headgear move forward, a switch is closed and completes the circuit. This headgear recorded very high accuracy. In 1997, the Selcuk timing headgear was developed by Enis Guray and Metin Orhan [8]. Similarly to the timing method mentioned above, this device also consists of a small quartz calendar watch timer that is activated based on power supplied to it. When the power is interrupted, the watch will stop working. However, these headgears are easily circumvented by patients because the timer can be activated simply by stretching the band (achieving a closed circuit) without even wearing it.

After the several drawbacks and the suspicion regarding the reliability of timing headgears mentioned above, P. A Kyriacou, D.P Jones and M.B Gungor invented a new micro-controller based wear time monitor that is implemented on headgear devices in 1997 [9]. The timer was activated based on two sensors. The first one is a temperature sensor that senses the body heat, outputting an extra 10mV per degree Celsius above the reference point of 300mV corresponding to 30ºC. The second is a force sensor which is placed on the headgear band to fit its spring mechanism. It senses the retraction force when the headgear is worn with an error margin of ±0.62N. This error can be determined since the pull up headgear does not cause large variations on the force. Cheating is considered hard in this case since two conditions must be duplicated to activate the timer and the patient was not aware of what attributes are being used. This device reported accurate and reliable results with an error of ± 36 sec. This way it is able to sense the exact preset force. This is not the case in cervical headgears which allow the head to move independently of the strap causing high variations in the force [6].

Later in 1997 at Washington University, Doctors Douglas Ramsey and his associates sought out to correct the flaws with the previous headgear monitoring instruments [10]. They aimed to design an instrument that is durable, not bulky, comfortable, and equipped with a safety breakaway feature. It was still qualified to record the temporal characteristics of the headgear, approximate how much force is applied on the headgear straps, tell whether a patient is cheating in headgear use, and offer accessible feedback to patients, doctors, and parents. Measuring the adjustment strap position is done by placing a small magnet at the end of the adjustment strap. Using Hall sensors to record the change in magnetic field strength, it was possible to estimate the position of the strap. Thus, the strap position allows knowing the amount of spring compression. The tensile force was measured at every strap extension, and then a relationship was described between strap position and force. The force is always varying when worn on a human head due to the effects of head posture, even while sleeping (although the variation decreases). When there was no variation in the force for more than one hour, the headgear monitor would stop registering time worn. The problem with this headgear remains with the risk of patient tampering due to the ability of stretching it around a spherical object. This is because integration of the force variations as a condition for timer activation was not present.

The following table summarizes the work of the mentioned authors in terms of the sensing method used for the monitoring as compared to the method employed in the Smart Headgear.

TABLE I
DIFFERENTIATION FROM LITERATURE

	Force	Temperature	Pressure	Pull	Solubility	Circuit Switch	Position/Motion	Touch
Northcutt			■	■				
Cureton						■		
Savage					■			
Selcuk						■		
Douglas							■	
Kyriacou	■	■						
Smart Headgear		■					■	■

However, in terms of overall implementation, our Smart Headgear system is not restricted to sensing methods that the doctor could later check during checkups. Instead, our project will also use the data coming from the microcontroller as an input to a game application (via Bluetooth). This application aims to incorporate the data into the game dynamics to enhance patient compliance which has not been previously done.

III. DESIGN

A. Hardware

i. Motion Sensor

This sensor option fits into our project in terms of a technique we can use to activate the timer. One sensor would be placed on the side of the headgear to measure the different positions as time changes. This would be useful since the patient's head posture will be highly variable at different instants. That would result in large variations in the position measurement each time the head is tilted or moved in a certain way. The variations are an indication of actual human use, which prove the absence of any patient tampering or falsification of time recordings. Position sensors are also accurate: can measure within short period intervals on the order of seconds.

ii. Temperature Sensor

The sensor in this case should be able to sense the body heat. It would be placed on the inner side of the neck strap. Thus, it would be in contact with the back of the patient's neck. Once implemented, it should control the activation of the timer by the following mechanism: if the sensed temperature on the back of the neck is between 30 and 42 degrees Celsius, the reading is considered acceptable.

iii. Touch Sensor

This sensor employs capacitive sensing which is sensitive to the dielectric difference specific to human tissue. According to research, this difference taken on the surface of the human neck is approximately equal to 0.2 S/m. The sensor would be placed on the inside of the neck strap. Thus, it would be activated upon being contact with the patient's neck. This sensor reduces the threat of tampering when paired with a couple of the aforementioned sensing methods since it ensures that the headgear strap is in contact with the patient.

iv. Controller

In order to integrate the three sensors we have chosen for the monitoring mechanism of our smart headgear, we need a microcontroller of size that matches the size of the provided headgear straps (20x5 cm). We program the microcontroller to increment a wear time variable whenever the outputs of the sensors are acceptable. In this case, we need the position sensor to show constant changes in the deltas of change of the three axes x, y, and z. The temperature sensor should output a reading between 30 and 42 degrees Celsius; and the touch sensor must sense the dielectric difference unique to humans which is 0.2 S/m. A recording is made every second, leading up to a decision concerning the wear status of the headgear.

B. Software

i. Game Application

Our smart headgear is accompanied by a game application in which the monitoring data is stored. The data transmitted from the controller, containing the patient's wear time, is received by this application. After the application receives the time measurements from the controller, it uploads this measurement to the unique patient profile which is available on the online server. This measurement is accessible for the patient so that he/she may keep track of it, while having an indicator of how much more time he/she needs to keep the headgear on. In order to encourage patient compliance, the application contains a game which takes into consideration how much the patient has been wearing the headgear and consequently allocates a maximum number of game attempts which would allow the patient to play the game as long as he/she has lives left. This would give the patient a competitive advantage over the other patients who aren't wearing their headgears as much since he/she would become more accustomed to the game. The game allows each patient to achieve a certain score, where the competition would be leaderboard and score-based.

ii. Server

In order to store patient-specific data which include their wear-time, our smart headgear is also accompanied by an online server whose databases are regularly updated at a rate of once per day. Each patient's Android application uploads their respective time measurements once a day to their online profile. This is necessary because the doctor needs to be able to access this information remotely, rather than monitor his/her patients' compliance only during checkups. The server is also necessary for the backend functionalities and application support.

iii. Doctor's Interface

Each orthodontist is given access to a web application through which he/she would have to login with their respective credentials in order to view their list of patients. Upon logging in, the orthodontist will be able to remotely monitor each patient's wear-time while having the capability of adding new patients to the online database.

C. Connectivity

The controller is connected to the game application via Bluetooth, utilizing IEEE standard 802.15. The application is connected to the online server via WiFi, utilizing IEEE standard 802.11.

Figure 1. Hardware System Interconnection

IV. IMPLEMENTATION AND TESTING

A. Hardware

i. Motion Sensor

The ADXL345 accelerometer is the motion sensor chosen. This sensor is capable of measuring static gravitational acceleration which is suitable for tilt-sensing applications. Placing the sensor on the left side of the strap provides delta changes (along X,Y,Z) in the patient's head posture. Delta changes not equal to zero proves that the headgear is actually being worn by the patient, and not just stretched around a fixed figure. The sensor is also characterized by a high resolution, low power consumption, and small size (with dimensions of 2.2x1.2 cm). All of these factors prove that the ADXL345 is capable of performing the position sensing functionality of our monitoring system.

ii. Temperature Sensor

The temperature sensor that is used in our smart headgear is the LM35DZ. This sensor is known for its extremely small size, very high measurement, and being a contact temperature sensor. Since it satisfied our constraints, we obtained the LM35DZ and tested it using several experiments where the sensor was embedded in the neck strap.

iii. Touch Sensor

A touch pad sensor is used to make sure the headgear strap is touching the patient's head. The touch functionality is performed by a digital capacitive sensor which would give a voltage when it comes in contact with skin. The chosen sensor is also compatible in size, having dimensions of 2.2 x 1.7cm.

iv. Controller and System Interconnection

To compensate for the small size of the headgear straps as previously explained, we chose the Arduino Pro Mini microcontroller (dimensions of 3.3 x 1.7cm). This microcontroller is embedded in the neck strap. Thus, the Arduino is connected to the accelerometer, which is situated on the side of the headgear strap, the LM35DZ and the touchpad sensor, which are both placed on the inner side of the strap as to provide contact with the patient's neck. In order to provide the microcontroller with Bluetooth transmission capabilities, an HS-06 Arduino Bluetooth module is also connected to the system. This way, the Arduino is able to pair to the patient's phone and send wear-time measurements to the game application. Finally, an RTC component is also connected in order to allow for the time stamping of the measured data. This is necessary since it is required to know the patient's headgear wear status at all instants.

Figure 1. Hardware System Interconnection
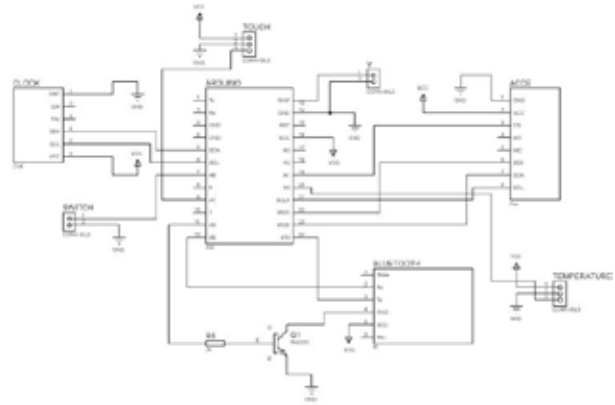
The previous figure explains the system interconnection, including all the hardware components and their connection to the Arduino microcontroller. These implementation of these connections is facilitated with the design of a PCB which was fabricated and used during the testing of the hardware and connectivity.

The Arduino, accelerometer, and RTC component are shown in the figure below after being soldered to the bottom side of the PCB which is embedded in the neck strap. The touch sensor shown will be in contact with the patient's neck.

Figure 2. Bottom side of PCB with components

Figure 3. Top side of PCB with components

The second figure above shows the top side of the PCB along with the soldered components that include the LM35DZ (in contact with the patient's neck) and the HC-06 Bluetooth module along with the necessary terminals for the battery connections.

v. Testing

Each of the sensors was tested individually with the timer to observe whether or not the logic of the microcontroller program is correct. All of the sensors provided desirable results

where the timer would only show an increment in measurement if the measurand of the sensor was suitable i.e. only when temperature fell in appropriate range, the touch was activated by skin, and the position was changing (which indicates motion). To test the reliability of the headgear device, ten trials were carried out. Each of the thirty-minute testing experiments was performed under different conditions. The different conditions were the person who is wearing the headgear during trial and the position of this person (sitting, walking, sleeping or laying down). All of these experimental trials yielded wear-time results of around 1530 seconds as opposed to a true value of 1800 seconds (according to a stop watch). This shows that our timing headgear has a wear-time measurement accuracy of about 85% so far.

B. Software

i. Game Application

An Android version of the application is developed for patients. After the patient receives his/her headgear, his/her doctor will be able to create an account for him/her. Using the provided login credentials, each patient is able to access the application interface. Upon successful login, the patient will have access to the two main functions of the application which are his/her wear-time statistics and the game. The game is designed to promote patient compliance due to the competitive advantage gained the longer the headgear is worn. The game is accessed based on availability of lives. Lives are incremented according a wear-time where each 15 minutes increases lives by one. However, this will motivate the patient to wear the headgear to gain more lives. For our application, the game concept is the following. Each patient will get an avatar once he/she opens the game application for the first time. This dental-themed avatar will be customizable in terms of color. This way each patient would feel more connected to his avatar because he/she played a part in designing it. Once the game starts, the avatar will be in a mode of free-fall. The job of the player is to move the avatar across the screen and land it on rising blocks as to avoid falling in the sea of candy below him. The figure below shows two screenshots of two stages of the game. The first stage is early on when the difficulty level is minimum: the blocks are of normal size and there is no constraint on touching the ceiling. As time progresses, difficulties increase. For example, when the time mark reaches 40, a block of teeth becomes visible on the ceiling (figure. 4) and the game would end if the avatar comes in contact with it. Also, the speed of fall increases with time as seen in the two screenshots. However, there are different power-ups that are implemented in our game in terms of increasing score and block size, for instance.

Figure 4. Later stage of the game

Testing

The game application was installed on 10 different Android devices which included different manufacturers. These were Samsung, HTC, and Sony devices. The phone users were instructed to play the game several times a day so we could survey number of crashes and get feedback from them. Over the course of two weeks, only two application crashes were reported and very good feedback about the game was received. Thus, the results of this testing experiment are very desirable.

ii. Server

The application is linked to an online server in which different databases keep track of patient and doctor information. The function of this server is essential for the project as it links the phone application to the remote web doctor interface. MySQL is used to develop the two server database tables shown in the figure below. The ID is the primary key of each patient in addition to having the Doctor_id as a secondary key which would allow each doctor to access a list of all his registered patients.
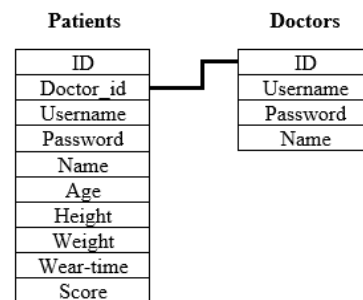
Figure 5. Database Tables

Through different developed PHP scripts, the server was connected to the phone application providing successful login functionality while also linking to the doctor interface. The implemented functions of the server extend to the computation of each patient's allowed game attempts and the support of the application's backend functionalities such as implementing the game's scoreboard and updating the score. The wear-time is used to compute the game lives as previously mentioned. All of these was tested numerous times over the course of days in order to ensure the absence of any bugs or errors in the code.

The results are desirable. The testing of wear-time storage has been completed as well. This was done through a parsing function which extracts the wear-time from the data which is sent from the Android application and uploaded to the server. Fifty attempts of sending different manually-input wear-time measurements from the Android were done, all of which resulted in a correct database update for the patient's wear-time field and returned a correct number of game lives to each patient's application.

iii. Doctor's Interface

In order to allow the orthodontist to remotely access patient information and have the ability to add patients to the database, an online web application was developed using HTML and CSS, and then linked to the server using PHP scripts. Webhosting and domain services were bought in order to put the interface online. It is now functional and available through www.toothytheapp.com. For now, we manually add doctors into the database to provide them with their own login credentials. After the doctor uses the website to login, he is able to perform two functions: retrieving patient statistics and adding new patients into the "Patients" database table. The doctor's welcome page includes a table which displays all of their patients along with their attributes. The doctor is then able to view a bar graph which displays the wear-time measurements of each patient during each day of the last week. Moreover, the doctor is able to download a file

Testing:

Fifty attempts were made to add patients and retrieve data which was manually added to the tables for testing purposes. These attempts were 100% successful. The following is an example of this interface where the orthodontist is Dr. Maria Doughan and the different patients she added are showing after she logs in to the doctor portal.

Figure 6. Information Retrieval Interface

iv. Overall System Functionality

An experiment which seeks to prove the operation of our three project components in integration was carried out. The purpose of this experiment was to prove that the Arduino is capable of sending wear-time measurements along with their day stamps to the Android application. It in return analyzes this data and uploads it to the online server where it is parsed and further analyzed to compute game lives which are sent back to the application and update the patient's file, in addition to setting up the wear-time timeline specific to the patient. The images present in the Appendix show that the data recorded by the microcontroller is concurrent with that received on the server and the resulting bar graph. Measurements of around 1530 seconds were recorded every day. This is correctly received by the server as seen in the bar graph in addition to sending 25 extra game lives back to the application. It is notable to mention that the game-life-computing equation was altered for testing purposes. Thus, the results of this experiment were desirable.

V. CONCLUSION

The results of the testing at this point prove that our smart headgear design is feasible. In this paper, the different functionalities and design details of the three main components of the Smart Headgear project are explained. The full testing of the fully integrated system is also completed, yielding promising results which can be improved in the future. This Smart Headgear will change the reality of patient compliance in orthodontics and positively provide orthodontists with an objective tool capable of accurately measuring their patients' adherence to headgear treatment.

REFERENCES

[1] A. Bos et al., "Comparing subjective and objective measures of headgear compliance," *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 132, pp.801-805, Dec.2007.

[2] M. E. Northcutt, "The timing headgear," *Journal of Clinical Orthodontics*, vol. 8, pp.321-324, Jun.1974.

[3] M. E. Northcutt, "Updating the timing headgear," *Journal of Clinical Orthodontics*, vol. 9, pp.713-717, Nov.1975.

[4] K P. A. Banks, and M. J. Read, "An investigation into the reliability of the timing headgear," *British Journal of Orthodontics*, vol. 14, pp. 263-267, Nov.1987.

[5] M. Savage, "A preliminary report into the development and use of soluble controlled-release glass timing discs implanted into orthodontic appliances," *British Journal of Orthodontics*, vol. 9, pp. 190-193, Oct.1982.

[6] P. A. Kyriacou and D. P. Jones, "Compliance monitor for use with removable orthodontic headgear appliances," *Medical & Biological Engineering & Computing*, vol. 35, 1997, pp.57-60.

[7] S. L. Cureton et al., "An accurate, inexpensive headgear timer," *Journal of Clinical Orthodontics*, vol. 25, pp.749-754, Dec.1991.

[8] E. Guray, and M. Orhan, "Selcuk type headgear-timer," *Journal of Orthodontics and Dentofacial Orthopedics*, vol. 111, pp. 87-92, Jan.1997.

[9] M. B. Gungor et al., *"A new micro-controller based wear-time monitor for use with removable orthodontic appliances,"* Proc. of the 19th Annual International Conference of IEEE, Chicago: IL, 1997, pp. 2419-2421.

[10] D. S. Ramsy et al., "Orthodontics headgear compliance monitor," U.S. Patent 5 980 246, Apr, 3, 1997.

APPENDIX

```
0:48:01 0/0/14 Day of week: Sunday
0:48:01 0/0/14 Sunday,1,34.8,0.008,0.043,-0.008,11,
tempC: 35.13
touch: 1
0.97,-0.97,-0.17
0.00
0.05
0.01
WearTime: 12
0:48:08 0/0/14 Day of week: Sunday
0:48:08 0/0/14 Sunday,1,35.1,0.000,0.051,0.012,12,
tempC: 34.80
touch: 1
0.95,-0.98,-0.16
-0.02
-0.01
0.00
WearTime: 13
0:48:09 0/0/14 Day of week: Sunday
0:48:09 0/0/14 Sunday,1,34.8,-0.020,-0.008,0.004,13,
tempC: 35.13
touch: 1
0.97,-1.02,-0.18
0.02
-0.05
-0.02
WearTime: 14
```

Figure 7. Serial output on Arduino (date stamp error due to component reset and was fixed)

```
0:48:01 0/0/14 Sunday   1    34.8   0.008   0.043   -0.008  11
0:48:08 0/0/14 Sunday   1    35.1   0.000   0.051   0.012   12
0:48:09 0/0/14 Sunday   1    34.8   -0.02   -0.008  0.004   13
```

Figure 8. Concurrent parsed information in patient file online

Figure 9. Wear-time measurements in the last 7 days where 30 minute trials were performed each day

Figure 10. Successful receipt of 25 lives after analyzing received wear-time

Spectrum Detection with Software Defined Radio Tools

Michel Abdelmassih, Christelle Azar, Joe Akiki

Electrical and Computer Engineering Department

American University of Beirut

mna53@mail.aub.edu, cea12@mail.aub.edu, jja29@mail.aub.edu

Abstract- **In this project we attempt to use the USRPs[1] to sense the medium. We acquired data about the signal using the energy detector and the cyclic autocorrelation function. The results are represented in a digital map on LabVIEW. The main purpose of our project would be to provide the telecom operators with areas that have bad coverage either due to bad transmission or the presence of illegal repeaters.**

I. INTRODUCTION

Telecom companies should offer good cellphone coverage to their customers. However, certain areas have bad coverage. This can be due to the lack of cell towers or the presence of illegal repeaters. Illegal repeaters are transmitters that detect a certain signal in the spectrum and retransmit it [1]. Without proper synchronization with other transmitters, these repeaters will cause an interference noise that would corrupt the original signal resulting in a high SIR.

In Lebanon, Touch and Alfa suffer greatly from the consequences of having illegal repeaters. This shows the need for a tool that can analyze the spectrum in the different Lebanese areas and help companies improve their signal coverage.

Using tools and software provided by the university and National Instruments, our project aims to provide a real-time spectrum analysis tool. This tool will integrate maps; users can use the maps to view areas with bad spectrum characteristics, and add new measurements.

II. LITERATURE REVIEW

A. Cognitive Radio

Defining cognitive radios is not the most straightforward thing considering that multiple resources define it differently. However, the definition that we will follow in our project is closest to that of the Federal Communications Commission (FCC) [3]:

"Cognitive radio: A radio or system that senses its operational electromagnetic environment and can dynamically and autonomously adjust its radio operating parameters to modify system operation, such as maximize throughput, mitigate interference, facilitate interoperability, access secondary markets."

In other words, a cognitive radio is a system that can sense the medium and change its parameters to better the transmission and reception of signals.

B. Spectrum Sensing

Spectrum sensing is an essential feature in every cognitive radio. A cognitive radio should be able to sense the medium correctly in order to react accordingly.

Usually in a medium, we speak of Primary and Secondary user. The primary user (PU) is in a way the user that can send signals whenever they have the need to [4]. On the other hand, secondary users (SU) can exploit the spectrum only when the primary users aren't. A SU should not interfere with the PU's signals [4]. In order to avoid interference, SU need to sense the medium correctly. Moreover, the SU should be able to change its parameters according to the currently available spectrum.

It is important to note that most of the licensed spectrum today remains unused in time and frequency [5]. This means that the SU has a good chance of transmitting.

Regarding spectrum sensing, secondary users face two types of errors [5]:
- Type I error: False alarms.
- Type II error: Missed detection.

False alarms happen when an idle channel is detected as busy. The SU does not use the spectrum because it believes that other users are currently using it. This leads to a waste of an opportunity for transmission.

Missed detection happen when a busy channel is detected as unused. The SU tries to transmit while another user is transmitting. This leads to a collision and both users must retransmit.

The different spectrum sensing techniques should try to minimize both Type I and Type II errors. In the following sections, we will discuss 4 main types of spectrum sensing techniques, which are the energy detector, the matched filter, the cyclostationary feature detector and the cooperative sensing.

C. Energy Detector

One famous spectrum sensing technique is the energy detector. It is in fact the most popular technique in literature [6]. In the following paragraphs, we give a small explanation of the energy detector spectrum sensing technique.

If we consider that the primary user sends signal x(t), then the received signal at the SU is either noise or only gain*x(t) + noise [7] as shown in equation (1).

$$y(t) = \begin{cases} w(t) & : H_0 \\ hx(t) + w(t) & : H_1 \end{cases} \quad (1)$$

---

1. USRP: (Universal Software Radio Peripheral) It's a transceiver that accommodates a wide variety of functions.

When the PU is using energy detection, these are the steps to be followed for the received signal [7]:
- The signal is filtered with an ideal band-pass filter
- The above output is squared
- The above output is integrated over a certain time period Tf

After following the above steps, the SU compares the output to a certain threshold λ [6]. When the output is greater than λ, the SU assumes the existence of a signal in the spectrum and does not transmit until it detects an output less than λ. The test statistic of the energy detector can be represented by [5] by equation (2).

$$z = \frac{1}{2n\sigma_0^2}\sum_{k=1}^{n}|y(t)|^2 > \lambda \qquad (2)$$

D. Matched Filter

Another spectrum sensing technique is the Matched Filter. This technique is used when we have knowledge about the signals in the spectrum. If the signal received by the SU is of the form y(n) =hs(n) +w(n) (h is a gain and w(n) is AWGN), the matched filter is the best detector for such AWGN [5].

The test statistic is represented by the following equation (3) [8]. Equation (4) explains more on the components of (3).

$$\sum_{n=0}^{N-1} x(n)z(n) \qquad (3)$$

$$z(n) = 2\sqrt{P_s}\cos(2\pi f_m n).\cos(2\pi f_0 n) \qquad (4)$$

The advantages of this technique at low SNR is the lower number of samples required (1/snr) as compared to the energy detector (1/snr2).

However, this technique can have bad performance in case of [5]:
- Timing and frequency offsets
- Fading
- Delay spread

Moreover, this technique requires knowing the signal structure and it insinuates that every primary user needs a receiver that is dedicated to it [9]. The SU using matched filter can't be used for spectrum sensing without prior knowledge of the characteristics of the primary user.

E. Cyclostationary Feature Detector

Cyclostationary signals are signals that exhibit periodicity in their mean and correlation [5].

Unlike other detection models, this one exploits the correlation between spectral components of the signal [5].

The test statistic for this detector is (equation (5)):
$$S(f;\tau) = \frac{1}{N}\sum_{n=1}^{N} y(n)y(n+\tau)e^{-j2\pi fn} \qquad (5)$$

The cyclostationary feature detector can be implemented very easily using FFT. The advantages of this technique are:
- Doesn't require knowing the noise variance [5].
- Fast, Accurate, Efficient [10].
- Can distinguish between noise and the PU signal [10].

- Can identify the modulation scheme used by the PU [10].
- Has high noise immunity [10].
- Performs well with low SNR [10].
- Can know the type of modulation scheme used by the PU [10].

It is important to note that a fast spectrum sensor is very important: if the PU is not using a certain frequency for a certain time, the SU should know that and have enough time to send data before the PU wants to send data again. Hence if the spectrum sensor is slow, there will be a delay and the PU might start sending again before the SU acknowledges that.

However, this technique can have bad performance in case of:
- Timing and frequency jitters [5].
- RF nonlinearities [5].
- Small number of samples [10].

This technique requires complex calculations, which might take time [10].

F. Cyclostationarity in OFDM

In communication, modulated signals often have a periodic property [11]. Using cyclostationary analysis of an unknown signal, we can determine some of its characteristics. This analysis would help us identify the kind of modulation that the transmitter is using, hence we can determine what these signal are being used for. For example, if we detect an OFDM signal we can safely assume that this is an LTE signal used for mobile Internet. Moreover this analysis would give us further information about the signal; even if the received signal is encrypted we can always tell if it is a useful signal (ie: a receiver with the proper demodulation scheme and key will be able to retrieve the data).

The following paragraphs explain the math behind cyclostationarity.

Consider a signal with a null mean and whose autocorrelation Rxx(t,τ) is defined by equation (6) [12].

$$R_{xx}(t,\tau) = E\{x(t)x^*(t+\tau)\} \qquad (6)$$

If the autocorrelation were periodic for all values of τ, then this signal would be defined as second order cylostationary. The autocorrelation of such a signal can be shown as a Fourier series (equation (7)):

$$R_{xx}(t,\tau) = \sum_{\alpha} R_{xx}^{\alpha}(\tau)e^{j2\pi\alpha t} \qquad (7)$$

In the formula above α represents the cyclic frequency. This sum is evaluated over multiple values of α.

Based on this formula we define cyclic autocorrelation (CAF) as following (equation (8)).

$$R_{xx}^{\alpha}(\tau) = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} R_{xx}(t,\tau)e^{-j2\pi\alpha t}dt \qquad (8)$$

All the computations that we have done so far concern an analogue signal. But in practice, we are working with digital

signals; the USRP samples and quantizes the signals received before applying any modification. Hence the CAF would have a slightly different expression (equation (9)).

$$R_{xx}^{\alpha}(l) = \lim_{N \to \infty} \frac{1}{N} \sum_{m=0}^{N-1} x[m]x^*[m+l]\, e^{-j2\pi\alpha m \Delta m} \quad (9)$$

The offset/lag that was previously represented by τ, is now represented by l. We have a string of discrete value x[m]; N is the length of x[m]; and m is the index of the array of values.

G. Detecting OFDM

While cooperative sensing takes care of the problem of multipath and shadowing [15], we are still in need of ways that can help us determine that the signals we receive, are of any relevance to us. We aim to retrieve two characteristics of any unknown electromagnetic signal:
- If the received signal is a useful signal
- If the received signal is an OFDM signal

To detect weather we have a useful signal or just noise, we have to follow the following formula (equation 10) [16]:

$$\lambda = \sigma_n^2 . \left(1 + \sqrt{\frac{2}{N}} Q^{-1}(P_{fa}) \right) \quad (10)$$

Using the expression of the threshold above, we know when the signal is considered useful. We can see that the threshold is independent of the noise present in the medium; it is only dependent on the number of samples and the probability of false alarm [13].

The second point we need to tackle is detecting if the signal we are using is an OFDM signal. Detecting an OFDM signal is based on retrieving the IFFT periodic part from it. Using the CAF of the received series of samples, we should evaluate CAF for α=0. Once we have the CAF for α=0, we should analyse the graph; if we can detect a periodic repetition of peaks all centered at zero this implies that the message received is an OFDM modulated message signal.

III. Solution Methodology

First, we started by reading a series of articles and papers about cognitive radios, spectrum sensing, energy detectors, and cyclostationary feature detector. These articles provided us with the formulas needed in our project.

In parallel, we started getting acquainted with the USRPs and LabVIEW. We chose LabVIEW as the main software because it's the standard platform for USRPs. Also, any MATLAB code can be integrated in the design using MATLAB script. The first thing we did was to build a basic transmitter-receiver scheme. To go further, we decided to check the accuracy of our receiver by catching radio frequencies and listening to FM channels.

Then we had three alternatives to check existence of a useful signal: energy detector, Eigenvalues [4], and matched filter. Eigenvalues was dismissed as an alternative. As for the matched filter, its implementation requires knowledge about the signal in the spectrum. This is why we chose to implement the energy detector.

Then we did some testing on MATLAB. We built an energy detector's ROC (Receiver Operating Characteristic) and observed how it varies as SNR varies. From the MATLAB testing, we clearly saw the relationship between the variance of the noise (depends on the noise level), the probability of false alarm, and the detection threshold. These results are similar to ones found in publications.

The next step was to implement this code on LabVIEW. We started by building a receiver based on the Energy Detection Scheme. Afterwards, we were required to test the receiver by using the already built transmitter. These models were implemented and tested on USRPs in the lab.

Using the NI user's forum, we downloaded and tested a program that detects the modulation scheme of the received signal using AMC [19]. It indicates if the signal follows a BPSK, QPSK, FSK or MSK modulation.

After checking the functionality of the AMC, we integrated the energy detector into the VI.

Another requirement of our project is to detect OFDM signals. Using NI forums, we found an OFDM transmitter and receiver [20]. To detect OFDM signals, we need to use a cyclic autocorrelation function (CAF) on the receiver. The CAF gives the OFDM signature. We had three different alternatives for implementing the CAF and have implemented one.

The final step of our project has been to integrate maps into LabVIEW. Based on the location, users can view the measurements done in certain areas. They can also add data to the database and display them on the map.

IV. Testing

In the following paragraphs, we describe some of the testing conducted during the development of the project.

A. Cyclic autocorrelation on OFDM

After developing the VIs on LabVIEW, we tested the CAF with the OFDM. The first VI was used to transmit OFDM signals, and the second VI was used to detect the signals. In the second VI, we used a MATLAB script to apply a cyclic autocorrelation function on the received signals. After doing so we got the following results (Figure 1):
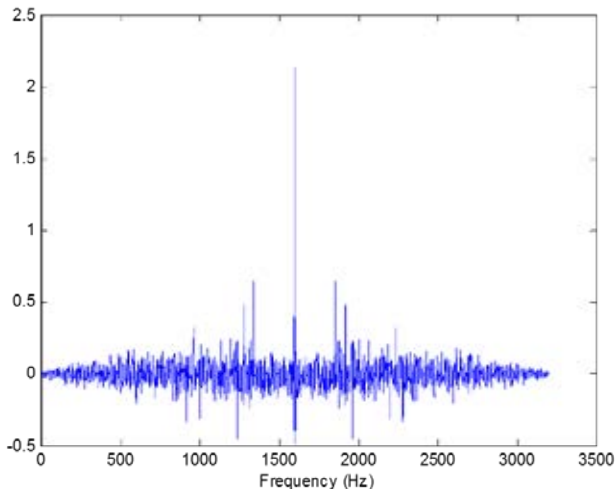
Figure 1: CAF of OFDM with 256 subcarriers

Figure 3: PFA vs Number of samples

B. OFDM with variable transmission power

We wanted to see how our system reacts to the decrease of the transmission power. Since the transmission power is constant in the USRPs, we increased the distance between the Tx and Rx USRPs. As the distance increases the transmission power virtually decreases. We transmitted the signals while separating the USRPs gradually. One of the graphs we got is shown below, for a separation of 1.5 meters (Figure 2):

D. 2G and 3G signals

In a later stage, we worked with real signals: 2G and 3G signals.

For the 2G signals we set the central frequency at 890MHz, we got a power of: $1.65*10^{-9}$ mWatts. Figure 4 represents the Power Spectral density of the 2G signal.
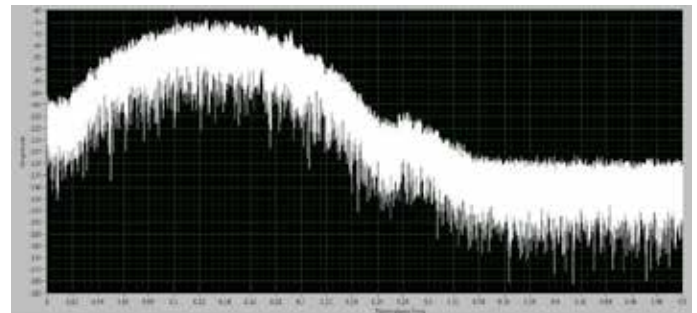
Figure 4: PSD of 2G signal

For the 3G signals we measured it at 1.8GHz, and we got a power of: $3.1*10^{-12}$ mWatts.

Figure 2: PSD of OFDM signal, with 1.5m seperation

C. Probability of False Alarm

In a next step, we conducted tests on the receiver. We plotted the probability of false alarm versus the number of samples (Figure 3).

CONCLUSION

In conclusion, our final product is a VI that senses the spectrum and detects the quality of a signal. It analyses the signal using the energy detector and the cyclic autocorrelation function. All the results are saved in a database that can be used to display the results on an interactive map. In the future, further research can be done regarding the energy detection threshold equation.

REFERENCES

[1]   "Illegal mobile phone signal boosters 'hijacking' other network users," ABC News, Jul-2015. [Online]. Available at: http://www.abc.net.au/news/2015-03-07/mobile-repeaters-disrupting-mobile-phone-signal/6287256. [Accessed: 2015].

[2]   [Online]. Available at: http://www.ni.com/pdf/manuals/375839a.pdf. [Accessed: 2015]

[3]   Federal Communications Commission, "Notice of proposed rule making and order: Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies," ET Docket No. 03-108, Feb. 2005.

[4]   T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications", IEEE Communications Surveys & Tutorials, vol. 11, no. 1, pp. 116-130, 2009.

[5]   A. Wyglinski, M. Nekovee and Y. Hou, Cognitive radio communications and networks. Burlington, MA: Academic Press, 2010.

[6]   H. Urkowitz, "Energy detection of unknown deterministic signals," Proc. IEEE, vol. 55, no. 4, pp. 523–531, Apr. 1967.

[7]   S. Atapattu, C. Tellambura and H. Jiang, "Energy Detection Based Cooperative Spectrum Sensing in Cognitive Radio Networks", IEEE Transactions on Wireless Communications, vol. 10, no. 4, pp. 1232-1241, 2011.

[8]   D. Bhargavi and C. R. Murthy, "Performance comparison of energy, matched-filter and cyclostationarity-based spectrum sensing," 2010 IEEE 11th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC).

[9]   Chen, A. Gibson, and J. Zafar, "Cyclostationary spectrum detection in cognitive radios," IET Seminar on Cognitive Radio and Software Defined Radio: Technologies and Techniques

[10]   A. P.s and M. Jayasheela, "Cyclostationary Feature Detection in Cognitive Radio using Different Modulation Schemes," International Journal of Computer Applications IJCA, pp. 12–16, 2012

[11]   A. Gardner, Cyclostationarity in Communications and Signal Processing, IEEE Press, 1994

[12]   K. PO and J.-ichi TAKADA, "Signal Detection Method based on Cyclostationarity for Cognitive Radio," pp. 2235–2238.

[13]   V. Dandawate, B. Giannakis, "Statistical Tests for Presence of Cyclostationarity" IEEE Transactions on Signal Processing, pp. 2355–2369, No.9, vol. 42, Sept. 1994.

[14]   A. William, H. Herschel, "Digital implementations of spectral correlation analyzers" IEEE Transactions on Signal Processing, pp. 703–719, No.2, vol. 41, Feb. 1993.

[15]   S. Mishra, A. Sahai, and R. Brodersen, "Cooperative Sensing among Cognitive Radios," 2006 IEEE International Conference on Communications.

[16]   Kobeissi, H.; Bazzi, O.; Nasser, Y., "Wavelet denoising in cooperative and NonCooperative spectrum sensing," in Telecommunications (ICT), 2013 20th International Conference on , vol., no., pp.1-5, 6-8 May 2013

[17]   Standards.ieee.org, 'IEEE SA - 802.15.1-2005 - IEEE Standard for Information technology-- Local and metropolitan area networks-- Specific requirements-- Part 15.1a: Wireless Medium Access Control (MAC) and Physical Layer (PHY) specifications for Wireless Personal Area Networks (WPAN)', 2015. [Online]. Available: http://standards.ieee.org/findstds/standard/802.15.1-2005.html. [Accessed: 2015].

[18]   Ieee802.org, 'IEEE 802.22 WRAN WG Website', 2015. [Online]. Available: http://www.ieee802.org/22/. [Accessed: 2015].

[19]   "Community," Automatic Modulation Classification using Cyclic Feature Detection with NI USRP Version History. [Online]. Available at: https://decibel.ni.com/content/docs/doc-23284. [Accessed: 2015].

[20]   "Community," Streaming OFDM Transmitter and Receiver Version History. [Online]. Available at: https://decibel.ni.com/content/docs/doc-34781. [Accessed: 2015].

[21]   Mathworks.com, 'USRP Support from Communications System Toolbox - Hardware Support', 2015. [Online]. Available: http://www.mathworks.com/hardware-support/usrp.html. [Accessed: 2015].

[22]   A. Al-Habashna, O. Dobre, R. Venkatesan and D. Popescu, 'Second-Order Cyclostationarity of Mobile WiMAX and LTE OFDM Signals and Application to Spectrum Awareness in Cognitive Radio Systems', IEEE J. Sel. Top. Signal Process., vol. 6, no. 1, pp. 26-42, 2012.

[23]   "Community," Automatic Modulation Classification using Cyclic Feature Detection with NI USRP Version History. [Online]. Available at: https://decibel.ni.com/content/docs/doc-23284. [Accessed: 2015].

[24] "Community," Using Google Maps in LabVIEW Version History. [Online]. Available at: https://decibel.ni.com/content/docs/doc-12505. [Accessed: Jan-2015]

Statistical Voxel-Based Morphometric Image Processing for Automated Diagnosis of Epileptic Lesions from MRI

Majed El Helou
Department of Electrical and
Computer Engineering
American University of Beirut
Mhe33@mail.aub.edu

Rawan Chanouha
Department of Electrical and
Computer Engineering
American University of Beirut
Rsc08@mail.aub.edu

Lynn Debs
Department of Electrical and
Computer Engineering
American University of Beirut
Lmd15@mail.aub.edu

Abstract- **Major contributing factors to the overall performance of hospitals are the accuracy and efficiency of doctors, which have been increasingly improving with technological advancement. One important challenge in neurology that has been recently receiving particular interest is the diagnosis and 3D localization of epileptic lesion regions in the brain, without which performing surgery on epilepsy patients is hindered. This paper describes NeuroAnalyst, a novel tool designed to assist neurologists in the diagnosis of epilepsy from MRI scans. NeuroAnalyst builds on state of the art morphometric image processing techniques, in particular, research targeting MRI processing and the detection of epileptic lesions. The tool generates post-processed 3D maps of the brain which can be visually assessed by neurologists to diagnose patients, each map emphasizing different important features of the brain that can help with the detection. The main contributions of this work are first to enhance the readability of the generated maps, by highlighting statistically probable abnormal regions across the brain for the diagnosing neurologists. Also, NeuroAnalyst generates its own automated diagnosis by leveraging statistical analysis techniques on the pre-processed MRI scans and studying the 3D geometric relations between possible abnormalities which are challenging false positives. Attributes and inter-relations automatically assessed by the tool were chosen with the collaboration of experienced neurologists from the neurology department of the American University of Beirut Medical Center (AUBMC).**

I. INTRODUCTION

Epilepsy is a neurological disorder due to a disruption of some brain cells' activities [1]. This results in the patients having recurrent seizures. Some conditions are considered mild (these are treated with medications) but others (30-35% of cases) require surgery to be treated. The symptoms vary from a blank stare and twitching arms or legs all the way to losing consciousness [1]. According to a recent study published by the world health organization, about 50 million people in the world have epilepsy [2]. Among these patients, approximately 80% are residents of low and middle income countries [2]. Nearly three fourth of these people do not have access to the necessary treatment for this disease [2]. Particularly, studies conducted in the Eastern Mediterranean have shown that the prevalence rate of epilepsy varied from relatively low rates (about 4.04 per 1000 population) to reach rates as high as "12 or more per 1000 population" [3]. Unfortunately, misdiagnosis of epilepsy is also still a common problem. Indeed, a study conducted on a population of adults has found misdiagnosis rates to be around 23% and 26% [4]. The two principal diagnosis tools for this disease are electroencephalograms (EEGs) and MRIs. The latter is used to detect and delimit the cells geometrically in the brain responsible for the recurrence of seizures so that surgery can be carried out.

In particular, functional MRIs are used to identify the activities which will be affected in case the neurologist decided to remove certain brain cells by performing a surgery. For that reason, the accurate analysis of MRIs is a crucial task. In many cases, neurologists find it challenging to visually perceive what is causing the seizures from MRIs. Abnormal transition from white matter to gray matter, abnormal cortical thickness values and abnormal extension from gray matter into white matter regions [5] are the main characteristics used to detect epileptic regions and these are usually very hard to discern visually. In this paper, we present the developed tool NeuroAnalyst which studies MRIs and creates maps highlighting the characteristics of abnormal regions then detects these regions automatically and thus greatly assists neurologists in the diagnosis process at the pre-surgical stage. All tests used by neurologists are ultimately aimed at helping them obtain a 3D MRI localization of lesion regions which are crucial for surgery.

Running other tests such as SPECT or EEG for the diagnosis requires the occurrence of a clinically observed seizure. Forcing a seizure can not only be dangerous for the patient, but can also lead to the stimulation of wrong parts of the brain and the detection of incorrect regions as abnormal. Lastly, the importance of MRI lesion region diagnosis reflects directly on surgery success rates.

II. RELATED WORK

A. Preprocessing of MRIs

Image normalization: This technique is commonly used by neurologists prior to any attempt to search for abnormalities in MRIs. The algorithm, described in [6], maps the patient's MRI to a well-known coordinate system by using a reference image of the human brain. This operation not only helps in accounting for the various biases in the MRIs due to factors such as the size or volume of the patient's brain but also allows neurologists to work in a well-defined normalized coordinate system instead of dealing with unknown new spaces that vary with patients. This common coordinate system is adopted by all the members of the "brain imaging community" [6]. In addition, normalization facilitates comparing MRIs of different patients as well

as matching a whole set of images to a single image instead of two images to each other for instance, which makes all the voxel-based morphometric computations more accurate. Spatial normalization is achieved by determining the optimal coefficients which minimize the squared differences between the template image and the initial image, and this operation has been thoroughly researched and developed over the years. It is also noteworthy to mention that normalization will not remove the visually intractable abnormalities, on the contrary, it will make them easier to catch by eliminating differences that are not abnormalities (these being normalized following the given normal template of the brain) [8].

Image segmentation: This method generates three images from the original image [7]. The first image captures the gray matter in the brain while the second shows the white matter and the third represents the cerebro-spinal fluid (CSF) present in the brain. Voxels belonging to each zone follow a Gaussian distribution which is used to assign apriori probabilities that make the segmentation process more effective. These apriori probability maps are obtained by studying a large dataset of patient brains, and generating a probability assigned to each voxel, for example the probability that it belongs to gray matter, to white matter or to CSF. A tool used by doctors to normalize and segment MRIs efficiently is SPM (statistical parametric mapping) which we will be discussing in details in the experiments section.

Smoothing: This step is usually performed once the segmented images are generated. Smoothing is used to "morph" or transform sharp differences in gray matter or white matter intensities in the brain from a voxel to its neighbors into a smoother and more probable difference in intensities [8]. For that purpose a Gaussian smoothing 3D filter is used. In short, each map is convolved in 3D matrix convolution with a regular 3D Guassian filter, whose three weight parameters we chose empirically to match a physical smoothing width of 6mm in the brain, an empirically tested reasonable width to avoid any smearing of abnormalities which would make them die out. This smoothing reduces high frequency noise that can come from the measurement equipment and which can lead to more false positives in the diagnosis. Smoothing is a necessary step used prior to most post-processing techniques of MRIs, in particular what is known by morphometry which is discussed in details in the next section.

B. Morphometry

There are three main types of morphometry: Deformation-Based Morphometry (DBM), Tensor-Based Morphometry (TBM) and Voxel Based Morphometry (VBM). However, there is no doubt that the prevalent analysis technique used in the processing of MRI scans is as mentioned earlier the voxel based approach. Voxel based morphometry was first mentioned in 1995, in [9], and since then it became the most used approach for studying the brain and its abnormalities [8]. Voxel based morphometry analysis of MRI can be outlined by the following steps: spatial normalization of the scan, segmentation of the brain and possibly smoothing,

preprocessing of gray matter segments and statistical analysis, all computed on a voxel study basis (i.e. the set of studied objects in the processing is composed of single voxels), to finally evaluate the obtained results [10]. Our proposed technique, building on VBM, will study gray matter segments as well as white matter maps to come up with the diagnosis and maps of the characteristics of abnormalities.

C. Z-score Maps

The three maps obtained through the segmentation of the normalized MRI scan are further processed to obtain three new maps which have shown through empirical testing to be the most suitable for highlighting abnormalities in the brain. These three characteristics maps being generally recognized by neurologists to be the most revealing of abnormalities. The computation of each map and its importance are summarized below.

Extension Image: The extension image highlights abnormal extension zones of the gray matter into the white matter regions compared to the normal database of healthy patients [11]. The map is obtained by applying the following steps, all done voxel by voxel:

a. The gray matter image of the patient generated from the segmentation step is smoothed, as discussed earlier.

b. The average image of the normal dataset is obtained and then subtracted from the image generated in part a.

c. The resultant image of part b is then divided by the standard deviation image (also computed from a normal dataset) to get the extension image.

Junction Image: Zones in which the borders between gray and white matter are not clearly defined appear bright in the junction image and therefore can be discerned as abnormalities [11]. Other regions that fall within the thresholds and that are significantly different from the normal dataset are highlighted as statistically probable abnormalities. The junction image is generated as follows:

a- From the gray matter and the white matter images generated in the preprocessing step, the standard deviations and means are used to compute an upper and a lower threshold as shown below [11]:

$$Lower_{Threshold} = Mean_{GM} + 0.5 \times STD_{GM}$$

$$Upper_{Threshold} = Mean_{wM} - 0.5 \times STD_{wM}$$

If the gray value of a voxel falls between these given thresholds, then it is set to 1, otherwise it is set to 0.

b- Smoothing of the generated binary image is then performed by applying the 3D convolution mentioned earlier. We chose to use specifically a Gaussian kernel to perform the smoothing operation for all the images as MRI scans are generally relatively accurate measurements.

c- The previous described steps are also performed on a normal database and the resultant images are then averaged. The obtained average image is subtracted from the convolved image in part b to account for "the variability of gray-white

matter transition zone in different brain regions" [11]. On the other hand, a "standard deviation image" is also generated from the normal database and is used to get rid of outliers. The image obtained from the subtraction is divided by this standard deviation image. The finally resultant image is called the junction image.

Thickness Image: The thickness image not only highlights regions in which cortical thickness is considered abnormal compared with the normal dataset but also accentuates zones where gray matter tissues are present in the patient's brain but not in the normal dataset subjects. This image is computed by the following steps:

a. The first step is to compute the binary image from the gray matter image generated using segmentation.

b. Next, for each voxel in the gray matter, run-length vectors are generated in all 26 possible directions which reach either a preset boundary (chosen to be 7 because almost all vectors will be shorter, so it is a good way to reduce unnecessary computations) or the nearest voxel were the gray matter boundary is reached. This gives a measure of how far the gray matter extends away from the studied voxel in any given direction.

c. Similarly to the previous two images, the thickness image is obtained by subtracting the mean image of a normal dataset and dividing this resultant image by the standard deviation image of the normal dataset.

III. PROPOSED METHODOLOGY

The methodology adopted can be divided into two main phases. During the first phase, we produce the three statistical maps which were described in the previous section by pre-processing MRI scans provided by the AUBMC using the SPM function toolbox. NeuroAnalyst will then use these images as an input to detect abnormalities by clustering the regions which have high z-scores values and generating scores for each cluster to detect any abnormal region if present. The developed tool is therefore able to perform the following:

1. Format handling
As discussed in the previous section, MRI scans which are the system's primary input, are done according to predefined and agreed on standards. The DICOM standard is followed when it comes to MRI scans, which are thus available in the "dcm" format and stored as a set of 160 2D "dcm" slices of the brain. The software tool therefore has to be able to process this format of images and be able to convert them into a more practical format. We propose to convert the 160 "dcm" slices into a single Disc Image File ("img"). The 160 2D slices will be combined and stored as a single 3D "img" scan. The reason is that the "img" format can be more easily processed for later stages in the proposed tool. The conversion mechanism is summarized in Figure 1.
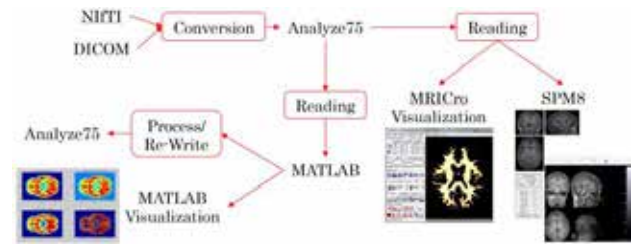
Figure 1: NeuroAnalyst Handling MRI formats

2. Image preprocessing
The next step consists of segmenting the MRI scan which is then normalized by the tool. The segmentation is in essence closely related to the partitioning/edge detection image processing problem aiming at obtaining a separation between brain gray matter (GM), brain white matter (WM) and brain cerebrospinal fluid (CSF). NeuroAnalyst performs the segmentation algorithm used by SPM which generates not only the three images mentioned earlier but also a parametric file which is applied in the normalization of the three generated images as well as the original image, the parametric files holds patient related information obtained when segmenting that are used to improve the normalization.

As discussed earlier, normalization is a necessary step in order to draw meaningful conclusions since a difference in the brain size for instance will be reflected as an abnormality which will potentially mislead the neurologists. Once the patient's brain is re-mapped, its size, brain zones' volumes, approximate shape and distribution of gray and white matter should become consistent with any re-mapped normal brain. Studying inconsistencies in that "standardized" brain scan relative to the expected normal brain scan is the approach for discovering abnormalities. SPM is generally used for the purpose of normalization, it has its own template for normal brains, and can also accept new templates that we create. Therefore, the next step consists of normalizing the patient's MRI scan relative to the normal brain template. NeuroAnalyst applies the normalization algorithm used by SPM since it is a widely used trusted tool in the field.

3. Generation of the z-score maps
The three z-score maps are then generated by NeuroAnalyst adopting the approach proposed in [11] and which was summarized in the Related Work section. This step consists of numerically computing the maps and then each is processed voxel by voxel in order to get their z-score normalization. After that, NeuroAnalyst enhances the visibility of these maps for neurologists by running a different z-score averaging, this time over the 2 dimensional slices that neurologists look at. Each voxel V will have the slice mean μ subtracted from it then it will get divided by the slice's voxel standard deviation σ:

$$V' = \frac{V - \mu}{\sigma}$$

This computation will mathematically always yield two dimensional slices with zero mean and unity variance, and is known for being useful for finding good statistically chosen outlier thresholding values, which is a matter of importance in our application. The generated maps are shown in the Experiments section.

4. Clustering

After the generation of all the statistical maps, a clustering algorithm is applied to group together voxels which satisfy both an 8-connectivity criterion and statistical outlier criteria in order to detect later any potential abnormal cluster. The same dynamic algorithm is used for the three maps and it works as follows:

a. An initial thresholding operation is applied on the map in order to filter out all the voxels having z-scores less than a certain threshold. Because of the slice-wise z-scoring, this thresholding is actually a statistical outlier identification. Also, the thresholds are generated both locally (current 2D slice) and across the 3D brain. All the voxels below the double-threshold are set to 0 at this stage, thus masking them out.

b. Each slice of the resultant image is then scanned bottom up and from left to right over all voxels. Whenever one outlier voxel is found, it is considered part of a cluster and its neighbors are examined. These are shown in the figure below (the currently studied voxel being located at (x, y)). The algorithm creates a toy 2D matrix for each slice, where every entry is either a 0 meaning the voxel is not an outlier, or have value n non-zero where n is its cluster ID.

Figure 2: The detected voxel and its neighbors

Once an outlier voxel is located, three different scenarios can occur:

Case 1: If all the neighbors have 0 values, then the found voxel is a new cluster. We create a new cluster ID in this slice and assign this new cluster ID to that voxel in the toy matrix.

Case 2: If some neighbors have 0 values and others have an old cluster ID that is unique (i.e. all neighbors are in the same cluster), the new voxel is added to the old cluster.

Case 3: If some neighbors have cluster ID1 and others have cluster ID2 and others possibly 0 values, then this voxel is assigned to cluster ID1 (which is the older cluster ID) and the second cluster is merged with the older cluster. After that, if there are IDs with larger values than that of ID2, these IDs are decremented by 1 to simplify the algorithm's output for later stages.
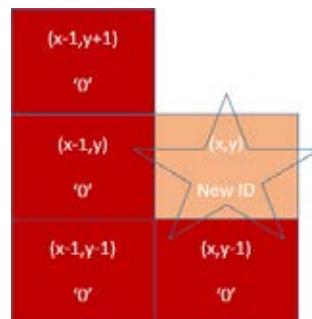
Figure 3: The detected voxel is assigned to a new created cluster
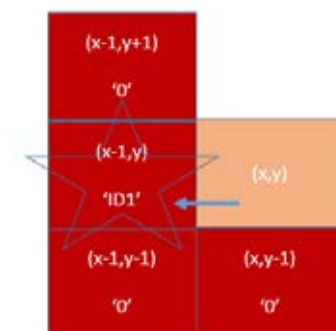
Figure 4: The detected voxel is assigned to an old cluster
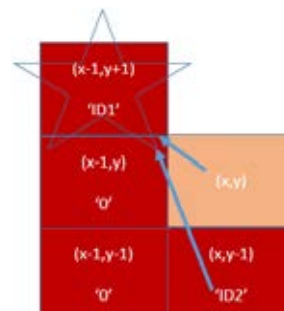
Figure 5: The detected voxel is assigned to an old cluster and two neighboring clusters are merged

It is worthy to note that this algorithm is quite efficient as it requires only a unique pass over the data to solve the entire clustering task.

5. Features extraction

From here onward, the objective is to decide whether a given cluster represents an abnormality or not. The criteria used in the assessment were deduced from the literature as well as our observation of the present abnormality and collaboration with neurologists. The features used are:

- Robustness: From our observation of the abnormality region in our patient's MRI, we noticed that one distinguishing feature between an abnormal cluster and

any false positive is the robustness of the abnormality across slices. In other terms, the abnormality is likely to appear and fade gradually whereas false positive are subject to more sharp changes. This feature is quantified by the number of slices spanned by each cluster.

- Size of the cluster: The size of the cluster is an important factor but not decisive. It is very likely to find false positive clusters that are larger than an abnormal cluster. However, the size of the cluster can still give an important indication and thus the algorithm takes it into account but with the assignment of a reasonable weight to it.

- Intensities of the voxels in the cluster: One for instance can consider the maximum intensity in the cluster and the mean of the intensities. This should give a hint of whether the examined cluster is a false positive.

- Number of other clusters in the same slice: By examining the number of clusters in the same slice, one can deduce whether this cluster models a region that is significantly distinct from others in the slice or if the slice simply has no real outliers thus has many apparent false ones (due to z-scoring on top of a low variance).

- Standard deviation of the intensities: The higher the standard deviation the more likely is the cluster to be abnormal. It is shown empirically that the intensities within an abnormal cluster follow a Gaussian distribution (can be fitted by a bell shaped curve).

Finally, a weighted average of the calculated features is used to derive a score for each cluster present in all the slices across the brain. The algorithm developed for the 3D spanning of 2D clusters works recursively, in a slightly similar fashion to the dynamic clustering algorithm, in order to only construct the span of any given 3D structure once as it moves across the 2D layers. Also, parts of the algorithms run in parallel on capable machines so that all in all the tool can handle the processing of around 8.4 million voxels per given scan, generated by novel MRI scan machines, in the matter of minutes.

IV. EXPERIMENTS AND SAMPLE RESULTS

The experiments were conducted using sample MRI datasets. The tool was used to first normalize and segment the MRI scan and then the provided normal database was employed in the generation of the three z-score maps from the pre-processed image. The normalized MRI as well as segmented grey and white matter are shown in Figures 6 and 7.

It is also worth mentioning that all the results (normalized, segmented and z-score maps) were validated by doctors from the neurology department at AUBMC. In addition to its ability to perform z-scoring across the whole brain (detecting suspicious regions

compared to those across all the brain which is the classical approach proposed in [11]), the NeuroAnalyst tool also performs a slice wise z-scoring to highlight abnormalities by comparing them to regions in the same slice. This modification has shown significant improvement in the extension map by eliminating a significant number of the false positives that were previously present. Figure 4 shows the generated extension image with and without the slice wise z-scoring. Similarly Figure 9 and 10 show the junction and thickness maps, respectively. It can be seen that the abnormal region has a higher intensity relative to the rest of the slice, in the slices that have slice-wise z-scoring, which helps the neurologists in localizing those abnormal regions visually on our maps.

Figure 6: Normalized MRI scan
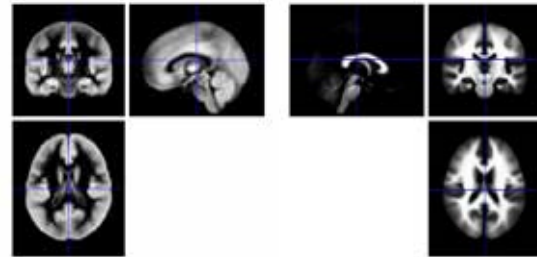
Figure 7: Grey matter segmentation (left) and White matter segmentation (right)
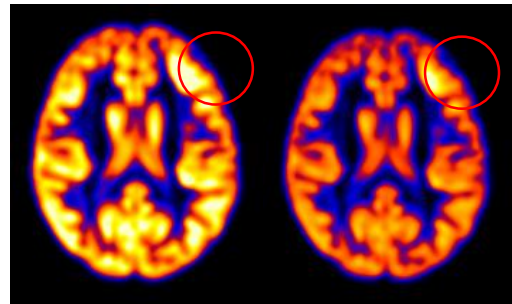
Figure 8: Extension map generated without 2D z-scoring (left) and with 2D z-scoring (right)

The generated z-scores maps are then passed to the clustering algorithm. Scores were computed based on the features extracted from each cluster. Primary results of the clusters' scores are summarized in Table 1. Only the clusters with the three highest scores in each map are reported and the abnormality among these three clusters is highlighted. What can be deduced from the following Table is that the abnormal region is centered around slice 50, because in the extension map the slices between slices 45 and 55 have a relatively (relative to other slices in this map) high score. More importantly, because here we can see they have a large span of 11 slices, while

other false positives with large scores span 3 or 5 slices at most, thus breaking the robustness criterion.
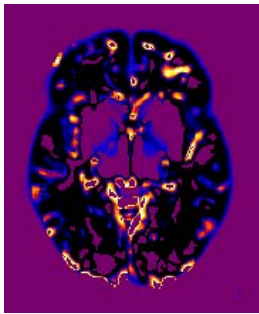
Figure 9: Junction map (without 2D z-scoring)
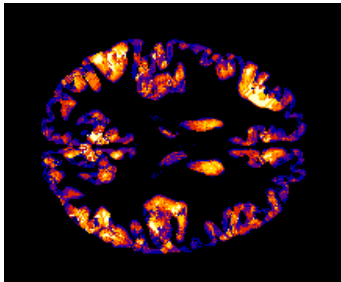
Figure 10: Thickness map (without 2D z-scoring)

Map	Slice span	Score value
Extension	**45-55**	**100**
	105-108	100
	139-145	100
Junction	77-79	19.9986
	50-52	**19.6109**
	6-9	14.2459
Thickness	19-21	33.3333
	25-35	25
	47-53	**5.2632**

Table 1: Primary Results

V. CONCLUSION

In this paper, we introduce a novel tool that aims to assist neurologists in accurately localizing epileptic lesion regions for surgery through MRI scans. The NeuroAnalyst tool takes the patient's MRI as an input and performs statistical analysis on it after the completion of the preprocessing steps. The normalized gray matter and white matter images obtained by this preprocessing are used to generate three statistical maps which give more insight on whether an abnormality is present as well as its potential location in the brain. These maps benchmark the obtained images with a normal database in order to emphasize any potential abnormality that could be depicted through observing the extension of the gray matter into the white matter, the junction between these two tissues as well as the thickness of the gray matter in the brain. The obtained maps are then passed to a clustering algorithm which connects a voxel having a high z-score value with its neighbors also having a z-score exceeding a set threshold

to form clusters. Features are then extracted from each cluster and are employed in the calculation of a cluster score by using weighted averaging. Primary results have shown that cluster scores can indicate not only clusters which could potentially represent a lesion region but also clusters which represent false positives that could mislead neurologists when searching for the abnormality using the generated maps.

Future research will target new features that relate to shapes and statistical variation pattern of intensities over the volume of the detected abnormal regions. This last feature will require the design and fitting of a geometric model for the way intensities vary over the three dimensional volume of abnormalities and requires the availability and study of a data set of epileptic patients. Future research will also involve more work on the design of a full grown machine learning algorithm that will study the extracted features.

REFERENCES

[1] Mayoclinic.org, 'Overview - Epilepsy - Mayo Clinic', 2015. [Online]. Available: http://www.mayoclinic.org/diseases-conditions/epilepsy/home/ovc-20117206.
[2] Who.int, 'WHO | Epilepsy', 2015. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs999/en/.
[3] Who.int, 'Epilepsy in the WHO Eastern Mediterranean Region', 2015[Online].Available: http://applications.emro.who.int/dsaf/dsa1039.pdf?ua=1.
[4] C. Ferrie, 'Preventing misdiagnosis of epilepsy', *Archives of Disease in Childhood*, vol. 91, no. 3, pp. 206-209, 2006.
[5] M. Reiser, H. Hricak and M. Knauth, *MRI in Epilepsy*. Berlin: Horst Urbach - Department of Neuroradiology - University Hospital Freiburg - Germany, 2013.
[6] J. Ashburner and K. Friston, 'Nonlinear spatial normalization using basis functions', *Human Brain Mapping*, vol. 7, no. 4, pp. 254-266, 1999.
[7] J. Ashburner and K. Friston, 'Image Segmentation', *Human Brain Mapping*, vol. 7, no. 4, 1999.
[8] J. Ashburner and K. Friston, 'Morphometry', *Human Brain Mapping*, 2012.
[9] I. Wright, P. McGuire, J. Poline, J. Travere, R. Murray, C. Frith, R. Frackowiak and K. Friston, 'A Voxel-Based Method for the Statistical Analysis of Gray and White Matter Density Applied to Schizophrenia', *NeuroImage*, vol. 2, no. 4, pp. 244-252, 1995.
[10] J. Ashburner and K. Friston, 'Voxel-Based Morphometry—The Methods', *NeuroImage*, vol. 11, no. 6, pp. 805-821, 2000.
[11] M. Reiser, H. Hricak and M. Knauth, *MRI in Epilepsy*. Berlin: Horst Urbach - Department of Neuroradiology - University Hospital Freiburg - Germany, 2013.

Temperature Control System for

Residential Solar Water Heating Systems

Mohammad Ghamlouch   Haytham Dbouk   Rayan AlSayed Ali

Department of Electrical and Computer Engineering

American University of Beirut

Hamra, Bliss Street

mig05@mail.aub.edu

hmd13@mail.aub.edu

raa131@mail.aub.edu

Abstract

The demand for solar water heating systems has increased significantly throughout the world considering that solar energy is a renewable source able to decrease the reliance on scarce resources [1]. However, the over exposure of these solar water heating systems to sun radiations (especially during summer) has resulted in the overheating problem which reduces the unit lifespan, causes premature component failure, and reduces the system's performance. This paper presents a design for a temperature control system that can reduce the overheating of residential solar water heating systems, thus protecting the unit. The system accounts for weather conditions and future household demand. The automated protection method controls any type or model of solar water heating systems, so that it can attain a safe desired temperature without influencing availability. The developed prototype is currently in the early stages of being tested on a real solar water heater.

I.   Introduction

Despite the fact that solar heating systems are considered hi-tech systems, they are still facing serious problems which need further research and effort to be tackled. Overheating is the main problem solar water heating systems experience. Overheating is due to the fact that most solar water heating systems

produce about twice as much heat in the middle of summer in comparison with what they produce in mid-winter. System

overheating can degrade heat transfer fluids, accelerate scaling, cause premature component failure, and reduce system performance. What we look for is a temperature control system that can reduce the overheating of residential solar water heating systems, thus protecting equipment and unit. This overheating protection would be implemented along predicting weather and household's future demand. That is, if the overheating protection control system predicts that tomorrow's weather would be rainy, it would allow the solar water heating system, today, to reach a higher set point temperature, so that hot water would be available for the household on both days. We aim to control any type or model of solar water heating systems so that it can attain a desired temperature degree limited by two set points that may change depending on the prediction of weather and the prediction of demand the temperature control system does [2]. This would assure that the solar system would not overheat to an extent that threatens the equipment, components, and unit. At the same time, the household would be supplied with enough hot water along any day, even if it is not sunny.

For the sake of protecting the solar system's unit against overheating, an automated temperature control system is designed. The system should precisely determine the lower and higher temperature set points so that the solar system would not operate outside the safe temperature range. Whatever the model of the solar system is, our temperature control system is capable of protecting it against overheating without requiring any modification to the solar system or user intervention. Safety is

highly considered in the system design, and it is accounted for in two senses: (1) the system isn't harmful on any user who tries to access it, because it might be necessary to access the system for maintenance and this shouldn't be harmful, and (2) the system is rigid enough so that it does not get damaged easily in the presence of strong winds.

Moreover, the designed temperature control system is considered reliable under different weather conditions and is expected to maximize its performance measure. This is accomplished by protecting the solar system and supplying the household with enough hot water simultaneously. For this purpose, two significant features were designed within this temperature control system, which are: (1) future weather prediction and (2) future household demand prediction. The temperature control system would use predicted future weather conditions and forecast future household demand and account for both of them when assigning the lower and higher temperature set points. Together, the predicted household demand and forecasted future weather conditions are mapped through a fuzzy logic process to their appropriate temperature set points.

II.   Literature review

Solar water heating systems are systems which captures sun radiations energy to heat water for households or companies [1]. Thus, the usage of solar water heater systems replaces the usage of natural gas, fossil fuels, or electricity with green, limitless, and free solar energy [1].
Solar water heater is characterized by their high efficiency; this is due to the issue that system collectors are able to absorb about 60% of the solar radiations which reaches the roof where it is mounted [1].

In addition, since its invention solar water heater system has been used in various parts of the world; where it's domain of usage is rapidly increasing with the improvement of the technology beyond this system and the decrease of its cost [1]. In European Union, 13% of resources and emissions are saved by solar water heater systems [2].
Furthermore, solar water heating systems have various properties which make it preferable over other heating systems. Solar water heating systems are systems which uses renewable energy source (solar radiations) causing reduction of electricity bills, thus taking advantage of new technology aspects with low maintenance costs.

Moreover, the challenge with using solar energy is that the sun is a highly variable power source. Thus overheating of solar water heating system takes place when heat can no longer be dissipated because maximum temperature is reached. Furthermore, overheating is likely to take place when we use wrong system type or tilt to shallow or when lifestyle and load changes that is when you for example go for a vacation or

simply reduces hot water usage [7].
In summer time water in the collector can easily approach boiling point this can cause excessive pressure and can damage the lines connected to water tanks from the collectors. Thus the temperature pressure relieve valve is not enough, since it can be easily damaged after successive overheating accidents and its damage is of big economic and technical cost.

In addition, overheating is a series problem which can cause capacity overload for water tanks (leading to its damage), gas emanation, calcification, lowering the lifespan of the system, and requiring high maintenance costs. All these drawbacks are besides the main safety concern because of very high temperature water.

Hence, several overheat protection methods have been implemented since the overheating problem has been encountered. For instance, drain-back system is a method which imposes differential control with high limit function is actuated when the storage tank sensor reaches the set point. At this point, the solar collector loop pump shuts off causing the water to be drained [3]. Two things make this method rejected: (1) the fact that it needs redesign for active solar water heating systems, and (2) the fact that it is not efficient due to the waste of water it causes when it drains overheated water out of the system.

Steam-back is another existing method where the pump switches off automatically, whenever the water tank reaches its maximum temperature. Once the water in the collectors starts boiling, the collector is fill with steam and all the liquid water is pushed down into the expansion tank [4]. However, this method does not protect the collectors from the very high temperature, and thereby the lifespan of the system will decrease continuously.

Heat dump is another existing method where a package automatically activates when the storage tank exceeds a certain high limit temperature: a circular pump in the heat dump works on pouring the excess heat into the atmosphere using a copper coil as a heat exchanger, allowing cooler water to flow into the collector loop [3]. Nevertheless, this installed package is a secondary loop plumped into the primary collector loop, so few connections to the storage tank are made. Hence, this solution requires some modifications to the existing solar system: that's why we rejected it.

Controller vacation mode is another existing method where a vacation mode that is present only in some controllers monitors the temperature of the storage tank by turning on the pump at night to dump excess heat through the collector [3]. But, this method of protection is well workable only with flat-plate collectors and not with evacuated tubes that have irreversible heat pipe cycles. Hence, this protection method does not work with all types and models of solar water heating systems.

Pump cycling is another existing method where the water temperature is allowed to increase to a pre-determined danger setting to increase convective heat loss [5], after the pump is stopped due to the tank's high temperature. As soon as reaching the danger point, the pump cycles turn on to carry the heat to the tank.

Heat-safe collector is a patented overheat-protection mechanism which allows excess hot air to vent, preventing overheating and the damage it can cause to the collector and fluid [6]. However, this solution causes a big waste of water.

Automatic Solution is another existing method where water with high temperature moves into a finned tube radiator where the dissipation of heat into the outside air happens. After cooling, the water that entered the reservoir is reintroduced to the system [3]. This is similar to an automobile coolant reservoir [3]. Nevertheless, this cooling system will consume too much electricity, diminishing the whole purpose of installing a solar water heating system to reduce reliability on fossil fuel.

Covering the collector is an easy, inexpensive method that is based on covering the solar water heating system manually by a cover, whenever it is overheated [3]. However, this method needs a direct access to the solar system, which is dangerous on specific roofs (especially steeply pitched roofs) and under some weather conditions. Add to this that the solar system is covered manually, hence this protection method is not automated.

Tilting is another existing method where the collector is tilted to latitude plus 15°. The higher the tilting angle is, the lower is the heat production [3]. Nevertheless, tilting differs between active (for tubes) and passive (for both tubes and tank) systems so it is highly dependent on the type or model of the solar water heating system.

Check-Valve bypass is another existing method where a manual valve is placed in a bypass pipe in a way to stop the circulation of water in the collectors whenever the system overheats [3]. However, the valve implemented in this overheat protection method is manual, hence the temperature control system is not automated. Moreover, this method does not protect the collectors from over hot water, it only protects that storage tank.

Pump shut-off at a pre-determined tank high limit temperature is another existing method where the pump shut off and don't push water in the collectors when the temperature reaches a set point below the overheating limit [5]. Nevertheless, this method is not useful if the user didn't use the hot water, so as the pump is off the water is still boiling in the collectors.

Proprietary collector design is another existing method where passive vents open when the collector's temperature hits a specific level, so that ambient air can enter the collector box limiting water heating and cooling the absorber by convection [3]. However, the temperature will drop slowly and this solution doesn't protect the collectors.

Pressure stagnation is another existing method where a higher pressure is in the collector glycol loop, which allows an increase in the water temperature and delays boiling at the same time [5]. Nevertheless, this doesn't work in countries falling on the equator where the day is long and the sunlight is strong.

III. Implementation

Thus upon an innovative temperature control system concept is implemented at three different levels: mechanical, electrical and software. The system was designed to fit "Deema 200" solar water heating system, taken as a sample for testing.

A. Mechanical implementation

This is the level at which the appropriate material needed for the accomplishment of the full system design are chosen. The full system design is built using the following materials:

- Stand: the stand of the temperature control system is totally composed of galvanized steel. This material was chosen based on a research done concerning the manufacturability of the system. The stand is composed of two rods, one of them is used as a shaft that allows the DC motor to control the shade placement, and the other is needed to identify the full covering/ uncovering of the solar water heater through accessing the corresponding limit switches. In addition to these rods, two pipes are used to allow the up/down movement of the shade to cover/uncover the solar water heater. Moreover, a case is also designed for the sake of protecting the shade when the solar water system is uncovered, and protecting the Arduino along with its power and control circuit. The temperature control system is combined and fixed over the solar water heater using screws and nuts.
- Shade: an appropriate shade is used to cover/uncover the solar system and is capable of preventing 80-90% of the sun radiations upon covering the system.
- DC motor: an appropriate DC motor capable of producing the necessary torque needed to spread and fold the fabric upon covering/uncovering of the solar water heater is used.

- Arduino along with its power and control circuits: this is the control unit of the temperature control system. It is composed of an arduino-uno along with its Ethernet shield, photovoltaic cell along with a rechargeable 9v battery, a voltage divider circuit (responsible for continuously feeding the Arduino microprocessor), and an Arduino

control circuit ensuring the feasibility of the temperature control system features (future weather prediction and demand prediction).

B. Electrical implementation

This is the circuit needed to ensure that the shade covers/uncovers the solar collectors, taking into consideration lower and upper temperature thresholds read by the NTC temperature sensor. Thus when temperature falls below the lower temperature threshold, the shade uncovers the collectors, and when temperature hits the higher temperature threshold, the shade covers the collectors. The below figure represents the full electric and control circuit.
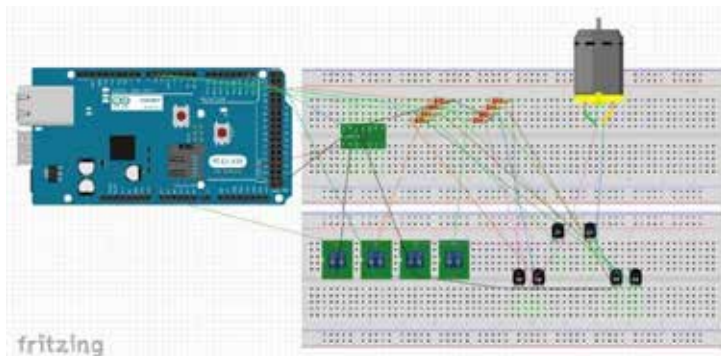

Figure 1: Electric and Control Circuit

This electric and control circuit is composed of the following components:

- 6 Volts PV panel (used as the electric and control circuit's power source).

- 5/6 Voltage divider using resistors (used because the Dc motor needs 6 v while the Arduino needs 5v).

- 4[2N2222 BJT] (used in the buildup of the H-bridge power divider circuit).

- 3 watts' rechargeable battery (used to continuously feed the Dc motor and the Arduino).

- Arduino Ethernet shield/ Arduino Uno (used as the control unit for the temperature control system, where the Arduino Ethernet shield is connected on top of the Arduino Uno, so that the sensors' pins are directly connected to the shield's input/output pins).

- NTC temperature sensor (used for measuring solar water heater temperature).

- 2 push buttons (used to disconnect the DC motor from the power supply when the shade is completely closed/ opened. The two limit switches will set the output D1 or D2 to 0 if pressed).

C. Software implementation

This implementation level is the manufacturing step capable of generating the algorithms required for demand and future weather predictions. Moreover, the algorithms implementation is the part ensuring the automation and independence of the temperature control system due to the serious criterion of "learning" the system will have. Hence the software implementation process is composed of:

- Future weather prediction:

Future weather conditions are classified into possible fuzzy sets, taking into consideration the domain of each variable weather condition. Weather conditions variables are: (1) today's min-weather temperature, classified into three different domains ([-10, 15[, [15, 30[, [30, 45]), (2) today's max-weather temperature, classified into three different domains ([-10, 15[, [15, 30[, [30, 45]), (3) today's weather current state, classified into four discrete values (sunny, rainy, snowy, cloudy), (4) tomorrow's min-weather temperature, classified into three different domains ([-10, 15[, [15, 30[, [30, 45]), (5) tomorrow's max-weather temperature, classified into three different domains ([-10, 15[, [15, 30[, [30, 45]), and (6) tomorrow's weather current state, classified into four discrete values (sunny, rainy, snowy, cloudy). Moreover, for the ease of decision and implementation and to avoid dealing with a huge number of combinations, the processes of setting inference rules is accomplished using layering along with divide and concur techniques. That is, all possible combinations of today's weather conditions will be listed in the form of (today's min-weather temperature, today's max-weather temperature, today's weather current state) triplets and classified into the appropriate group; each group corresponding to one of the three values (low, medium, high) of *today state variable*. The same is done for tomorrow's weather conditions to generate three values (low, medium, high) of *tomorrow state variable*.

- Demand Prediction:

The system counts the number of sudden drops in temperature throughout the day and uses this frequency to update the running average of frequencies at the end of the day. The updated value of the average frequency will be used to predict the household demand on the same day one week later. This will be done through mapping the average frequency value to one of the following three states that describes demand. The three states are: (1) low demand, (2) medium demand, and (3) high demand.

All together, the demand state, today state and tomorrow state are mapped through a fuzzy logic process to the appropriate temperature set points.
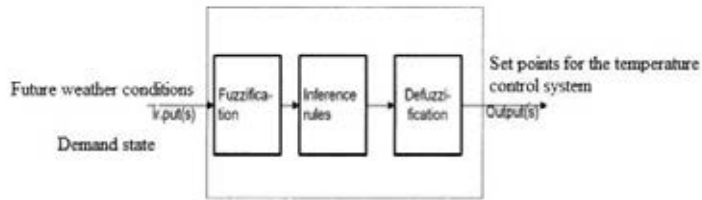
Figure 2: Fuzzy controller composition
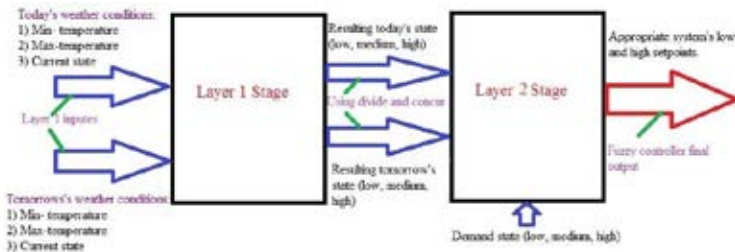
Figure 2: Fuzzy controller layering and divide-concur

Demand state, today state and tomorrow state are the fuzzy algorithm inputs; thus noticing the domains corresponding to today state, tomorrow state and demand conditions inputs along with their classification; all possible combinations of the input values are scattered into groups according to the ones which results in the same output values (the appropriate set points).

In addition, input possible combinations will be listed in the form of (today state, tomorrow state, demand state) triplets and classified into the appropriate group; each group corresponding to certain output.

Where appropriate temperature set points are classified into possible values each corresponding to a group of combinations. Thus within this state a certain type of membership function is developed over the output variables possible ranges resulting in the rules consisting the fuzzy algorithm. Moreover, the defuzzifized output consist of pairs of (lower set point, upper set point) specifications for the opening/closing mechanism observing its dynamic environment.

Thus there exist 3 distinct pairs of resulting output each related to a group of input combinations through one of the fuzzy algorithm rules, these rules are:

Rule 1: If the input combination belongs to group 1, then the corresponding output pair is (70, 80).

Rule 2: If the input combination belongs to group 2, then the corresponding output pair is (50, 60).

Rule 3: If the input combination belongs to group 3, then the corresponding output pair is (40, 50).

IV. Experimental Results

After setting a detailed design schematic and listing all needed materials several levels of testing are required as a prerequisite for establishing final prototype for the temperature control system**.**

A. Fabric testing

To verify the chosen design, we conducted several tests to quantify the fabric quality, efficiency compared to cost, and percentage of sun radiations preventing. Furthermore, 2 tests were conducted to obtain the shading sunlight preventing percentage.

- Test 1:

This test is conducted using LDR (light dependent resistor). First the LDR should be calibrated, where calibration is ensured when the LDR reads 0 Ω in total obscurity.
Testing starts by placing a source of light at a fixed distance from the LDR it indicates 200 Ω; the fabric is then introduced in front of the LDR thus lowering the reading to 85 ohms.

Taking into consideration the linear relationship between resistance and the inverse of light intensity:

$$R = \frac{500}{Lux} \quad (1)$$

where Lux is the light intensity, and R is the LDR read resistance, the percentage of light sunlight passing through the fabric was found to be:

> % of sunlight passing the fabric = R with fabric / R without fabric = 80/200 = 40%. Thus % of sunlight reduction = 60%.

- Test 2:

This test is conducted to assure the exact fabric sunlight preventing percentage. Thus the test uses the previously installed LDR which is connected in series with a 330 kohm resistor, and a 5V power supply.
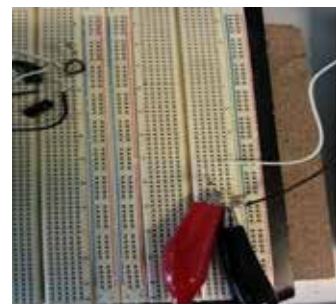
Figure 3: Test 2 circuit

By substituting equation (1) in the voltage divider equation, we arrive at the following relation: $Lux = \frac{\frac{2500}{V_o} - 500}{3.3}$ (2)

Moreover, the light source is placed at a certain fixed distance from the LDR then voltage across the LDR is measured before and after inserting the fabric. Results are:

VLDR1 = VLDR before inserting the fabric=1.05 V.

VLDR2 = VLDR after inserting the fabric= 1.5 V.

Thus, Lux1 = 569.98, Lux2 = 216.45, and hence
% of reduction= 62.02%.

Upon repeating this experiment for 5 iterations, percentages ranged between 57.8 and 62.02%. Thus fabric sunlight preventing is considered to be 60%.

B. Prototype Testing

To accommodate the full system testing the first iteration prototype was designed and built to simulate real application conditions as accurate as possible. Thus several steps where followed to reach the required prototype before the testing, these steps are:

a- Fabric casing (this is series simulation of the required casing for fabric to resist various weather risks)
b- Adding a shaft for the DC motor
c- Cascading the cascaded fabric along with the motor with shaft part
d- Manufacturing the needed tracks for the fabric to rollup/roll down
e- Putting all parts together

After designing the full prototype system testing was conducted, where the material used on the testing process are:

1) The manufactured prototype
2) A 10 v supply (a variable power supply was used to supply the motor with 10 v and approximately 1 A, so that the motor is capable of rolling up/rolling down the fabric to cover/uncover the collectors).

Figure 4: Simulation where fabric is in its case

Figure 5: Simulation where fabric is rolled down

C. NTC Sensor Testing

The temperature and voltage of the NTC temperature sensor were recorded at different instants. The resulting set of data was then plotted on excel so that an explicit relation between temperature and voltage was established.
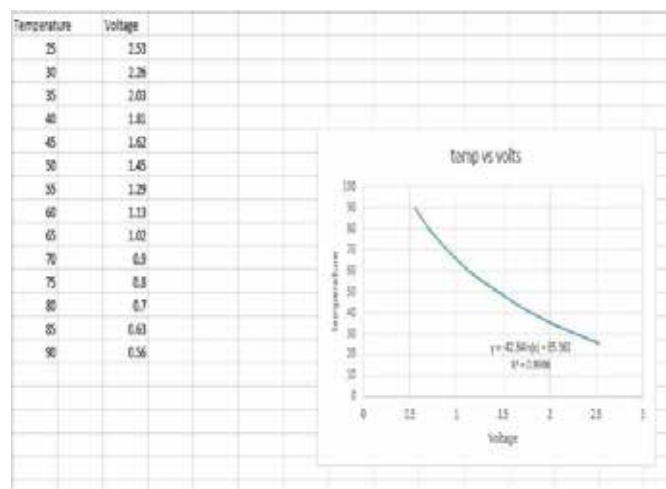
Figure 6: NTC Voltage as a function of temperature

This relationship was found to be:

$$V=-42.84\ln (T) +65.361 \quad (3)$$

V. Conclusion and Future Work

The designed and implemented temperature control system is a 50-100 dollars' system capable of solving the problem that all solar water heating systems suffer from, the overheating problem. This would accomplish the aim of protecting residential solar systems reliably under different weather conditions without the need of accessing them. Once, the temperature control unit is installed, it will automatically protect the system (with accuracy of control ± 5 degree Celsius) even if household members aren't present at home.

What makes this temperature control system important, is that it will apply the overheat protection along weather prediction and near future household demand prediction. This protecting system is always aware of the temperature of the solar water heating system. Hence, once the solar system attains a specific dangerous temperature that threatens the unit, the temperature control system will automatically initiate protection without any user intervention.

Future work in this problem will consider modeling the fuzzy logic algorithm inputs establishing a mathematical characterization capable of identifying the overall system.

Figure 7: Final full system design

VI. Acknowledgements

VII. References

[1] M. Azzouzi, "Control of solar water heater design", Environment and Electrical Engineering (EEEIC),2011 10th International Conference on. Pages:1-5, DOI: 10.1109/EEEIC.2011.5874684.

[2] S. Yousfi, "L'énergie solaire en Algérie", Quotidien national El- Khabar, Algiers, 2007.

[3] C. Marken,"Overcoming Overheating", Issue 142, April 2011

[4] F. Mueller, "Steamback Shows Promise for Solar Water Overheating", August 2011

[5] Solar rating and certification corporation, "Assessment of Overheat & Freeze Protection Mechanisms", February 2013

[6] Enerworks solar heating and cooling, 'Enerworks', [Online]. Available: http://enerworks.com/products/. [Accessed: 12-Sep-2015]

[7] F. Bautista, J. Aguilar, J. Ortega, F. Rivera, M. Gonzalez, M. El-Halwagi, "Optimal design of domestic of water heating solar systems", Clean Technologies and Environmental Policy, March 2015, Volume 17, Issue 3, pp 637-656.

Temporal Entity Normalization in Arabic

Jad El-Kik, Rawane Issa, and Ramzi Karam
Department of Electrical and Computer Engineering
American University of Beirut
Beirut, Lebanon
Email: jwe05@aub.edu.lb, rgi01@aub.edu.lb, rrk16@aub.edu.lb

Abstract— The normalization of temporal entities extracted from text is an important natural language processing problem. Given a temporal text expression, normalization is concerned with projecting the expression onto a timeline and accurately mapping it to a numeric temporal value. This allows for a plethora of applications in information retrieval and machine translation. Some research work to extract and normalize temporal entities for Latin text exists. Research for Arabic is currently limited to temporal entity extraction. This paper describes a new method with supporting tools for normalization of Arabic temporal entities using morphological analysis, expert rules, hidden markov models and temporal fact scoping.

I. INTRODUCTION

In the field of Natural Language Processing (NLP), less work and research has been carried on the Arabic language compared to Latin languages such as English. In fact, the rich morphological nature of Arabic makes it a hard language to process, and requires further processing compared to morphologically poor languages like English. In addition to this, Arabic is an under-resourced language meaning that there is not enough tagged data available and statistical methods alone often fail to produce accurate results due to the lack of necessary data. For these reasons, tools such as ATEEMA [1], developed by the Arabic Natural Language Processing team at AUB, have been recently developed in order to extract entities and relational entities from Arabic documents. A key to these tools is the morphological analysis of Arabic they provide which is necessary due to the rich nature of Arabic morphology. However, these tools are still unable to normalize the extracted temporal entities and give them standard numerical values. Our team proposes to work on solving this problem, thus helping Arabic catch up with other languages in terms of technological advancement.

A. Motivation

Some of the direct applications of temporal entity normalization are the following:

- Automating the process of temporal ordering: ordering of medical events (medication times, symptom dates etc.) ultimately reducing possible medical errors and saving crucial diagnosis times; ordering of historical events in texts and books thus helping scholars in their tasks of building timelines and large databases.
- Providing Arabic users with the option of creating calendar events based on the timestamps found in their emails.

- Augmenting the usage of Arabic across the world by making it a more accessible language and providing it with the tools that Latin languages have.

B. Literature Review

Mani and Wilson [2] represent times following the ISO standard CC:YY:MM:DD:HH:XX:SS (ISO-8601 1997) where CC represents the century, YY the year in that century, MM the month, DD the day, HH the hour, XX the minutes and SS the seconds. One way in which the work presented in this paper is innovative compared to previous work is its concern with temporal expressions that are relative to the speaker or to a certain reference time, like yesterday which could mean very different days depending on when it is said. These context-dependent time expressions are evaluated against a Reference Time which the authors of this paper have chosen to be one of two things: the document creation time or the time currently being talked about in the narrative.

II. ARCHITECTURE AND PROPOSED SOLUTION

The Arabic Temporal Entity Normalization is divided into 3 tasks: Entity extraction (using ATEEMA), entity normalization and scoping. We will briefly describe entity extraction, then carefully describe entity normalization and scoping.

The following is a description of each step of the diagram in Figure 1.

A. Entity Extraction: ATEEMA

The Arabic Temporal Entity Extraction with Morphological Analysis (ATEEMA), is an open source software which allows users to process Arabic texts using knowledge-based morphological analysis. ATEEMA first takes an Arabic text as input and passes it to a morphological analyzer (ATSarf). The analyzer processes the input by considering the stem of the word along with its prefix, suffix and infix. Each of these is tagged with several tags such as a part-of-speech (POS) and a gloss tag. The morphological analyzer also reports categories: TIME, NUM (for numbers) and TIME_PREP. The analyzer then calls ATEEMA back with the current content and the found morphemes (prefix, suffix, infix and stem). The callback then adds the morphemes to a sequence of unresolved morphemes or resolves them and produces a temporal category as a finite state transducer (FST). The FST then recognizes the temporal entity while taking into consideration both morphological variations and ambiguous solutions. The

Fig. 1. Proposed Solution Flow Diagram

في أول كل شهر ,من سنة ٢٠١٤ etc.

Having classified temporal expressions into the following categories, we will now discuss how we are representing the values of time expressions.

FST is capable of detecting temporal expressions by counting the number of possible temporal categories and either going to the states of Time or Maybe Time. If a certain threshold is passed the expression is either discarded as non-temporal or accepted. Heuristic optimizations in ATEEMA include: Ignoring the word العام in the context of analyzing newspaper texts, because of its high level of ambiguity; And using the word أحد as a number, and الأحد as meaning "Sunday".

Temporal Data Structure

In order to conduct the normalization process, we have devised a data structure (presented in Table I) to store and represent an expression's values:

From: +flags : (true?|-inf)	To: +flags: (true ?|+inf)
For/Span	Weekday
frequency link (flag/flag)	

TABLE I

DATASTRUCTURE FIELDS

Entity Normalization

Temporal expressions can be classified into different categories depending on how these expressions are placed onto a timeline:

- **Explicit Expressions**: these expressions include those which can be directly hooked onto a calendar. These expressions only include those which fully specify a date in terms of DD/MM/YYYY and can refer to a specific interval like: من ٢٠١٥-١١-٠٣ حتى ٢٠١٥-١٢-٠٣".
- **Indexical Expressions**: these expressions include words which require references in order to be evaluated. Examples of the expressions might be الأمس, "في الخامس من هذا الشهر , etc. These expressions can also refer to an interval: من اليوم حتى الخامس من الشهر المقبل. The words اليوم and الأمس do not explicitly point to a specific date, but depend on when they have been pronounced.
- **Ambiguous Expressions**: these expressions are those with no clear projection onto a time space and cannot be expressed as fixed points or intervals like in the cases of indexical expressions or explicit expressions. For example: في السنين المقبلة.
- **Semi-determined Expressions**: this final category represents expressions which are only partially expressed in the DD/MM/YYYY format, but do not contain all of the information required. Example: في شهر كانون الثاني

The From-To entry represents temporal expressions which express an interval of time with start and end dates. These dates will be written in a granular system, which we will discuss later on, and will therefore not contain a specific weekday. Such expressions are generally of the following format: ..... من .... إلى .....
If we are dealing with an event that always happens, for example في كل يوم, the + and - infinity flags are set to true along with part of the recurring date, in our example جمعة.
If we are dealing with an expression that will always happen in the future starting from a certain date, then we will fill that starting date in the From entry and set the infinity flag in the To entry. An example of such an occurrence is من اليوم وصاعداً. We will work similarly with events that have always happened up to a date (From = - infinity, To = date in that case).
However, this entry is not exclusively restricted to such spanning events. It can also represent single point events, for example: الإثنين ٥ أب. In such a case the To entrys is true? flag is set to false.
Moreover, these entries can represent an expression which is not very informative but expresses a duration. For example: في فترة زمنية . In such case we only need to set both is true? flags to true.

The Span entry represents an interval of time. This interval can be calculated from the From-To entries if these are expressive enough to allow us to do so (if, for example, both are only set to true without values, then there is no way of computing an interval). This entry can also express the value of intervals of time with are not hooked to specific

dates for example شهر لمدة. These expressions have no From-To entry but are not uninformative.

The Weekday entry represents a specific day of the week. Notice that a day in the form DD/MM/YYYY does not explicitly tell us which day of the week this date will occur on. This observation initially led us to create this entry. We later noticed, while reading articles, that some temporal expressions only contain a weekday, for example: في يوم الجمعة. This entry can also help us represent expressions like صلاة الجمعة, which always occur on specific weekdays.

The Frequency flag and link connect two temporal expressions which are related via a period/frequency relation. In other words, if an event will occur between two dates in a specific frequency, then we will represent this as 2 separate temporal expressions related via the frequency flag: the From/To date and the frequency of the event (either as a Span or as a specific Weekday). For example: أخذ الدواء كل أربعة أيام لمدة شهر.

In order to fill our data structure we first follow a rule-based approach in which we match expressions with hardcoded rules and try to extract as much information as we can, before proceeding to statistical methods if the rule based method did not fill the data structure enough (we will formally explain what enough is in the statistical method section.)

Ontology tree

We will give values to specific words by looking them up in an ontology tree. Based on the relations found in the ontology tree, we will infer possible ways of filling our data structure and feed the result as a feature input to our statistical methods. The Arabic ontology tree is a tree with the meanings of arabic words and the relations of these words with other words (part of, subtype of, and instance of relations) if any relation exists. The ontology tree we are using was provided to us by Dr. Mustapha Jarrar from Birzeit University. Other lexical databases, such as Wordnet, are based on psycholinguists analysis of words. These lexical databases rely on the relation is an instance of which does not represent meanings of words and might lead to an inability to classify certain words. For example if we consider the words جدول and نهر ابراهيم as instances of جدول, does that mean that جدول and نهر ابراهيم share the same relation with نهر? The Arabic Ontology is well founded (logically and philosophically) and has so far proved to be a good representation of words for our purposes. For example, the semantic relations (part of/جزء من and subtype of/جنس من) between words lead to a removal of a large part of the possible ambiguities linked with classifying words. For example: عيد الأضحى becomes an instance of عيد and عيد becomes a subtype of يوم. For these reasons, we have chosen to make use of the Arabic ontology tree.

Granular system

We will call granularity a unit of time: second, minute, hour, day, week, month, and year. Our time values will be represented as functions T of the granularities and more specifically as a sum of the granularities (i.e. $T(a1, a2, a3, a4, a5, a6, a7) = a1 * second + a2 * minute + a3 * hour + a4 * day + a5 * week + a6 * month + a7 * year$). In such a way, any operation we conduct on our time becomes a well-defined operation of refinement (if we are adding precision to our time value), or abstraction (if we are removing precision from our time value). This concept of granularity also allows us to define shifting operations (if we are increasing or decreasing one or more granularities). The shifting operation is effectively shifting our position on the timeline by a given distance on the axis (i.e. the span). (Note that if we have for example $9 * month$ and add 4 to our month granularity, we carry a 1 to our year granularity and set $9 * month$ to $1 * month$.)

Expressions Processing

We can now proceed on to seeing how we will process incoming expressions from ATEEMA. Each expression is represented in the form of a 3-tuple consisting of a sequence of words, $< POS, Stemmer, Category >$ tags for each of these words, and the expressions' value represented in our data structure. We will follow the following algorithm we have devised for normalization:

- For each word in an expression, look for the word in the ontology tree and add the value found in the tree onto the word's feature vector. The value returned by the searchTree method is obtained by looking the word up in the tree using a simple traversal. Since a word can have multiple definitions and the ontology tree is based on definitions and not words, then an input word may have several matches (e.g. "عيد الأضحى" can mean different things). Also, since some of the expressions in the tree are partial matches, an input word may get several of these too (e.g. عيد in itself and عيد الأضحى). For every match, we mark the node as "1" in the ontology feature vector, and all other nodes are marked with "0". For every node marked with 1, all of its predecessors are also marked with 1. Given a definite

order of the nodes in the graph, there will be a definite resulting bit vector of the marks on these nodes.

- Then, using a rule-based method composed of several rules, in which have for example ....من...إلى, we fill our temporal data structure. These rules may depend on the ontology tree results as well as the different tags which were extracted using the morphological analysis. Several solutions may result from this stage of normalization and each will be retained. (Different solutions might also occur if we find different meanings of a word in the tree: الإثنين might for example mean two or Monday).

- Pass each solution to a statistical method (discussed in a later section). The input to the statistical method will thus be each expression with each word's feature vector. The output of the statistical method will be the previous input, along with a set of tags which indicates where to place the word in our temporal data structure, and to which granularity the word belongs, if such a tag holds.

- Then, if the user has decided to use scoping, we suggest a list of words than can be scoped for the user, and let him decide on which words to scope (from the suggested list or elsewhere in the input text). For each of these words, we find a scoping interval and add it to a vector of intervals. Then for each interval in the vector of intervals, we check these intervals for consistencies using collective inference constraints, which we will discuss further in the scoping section.

Rule Based Approach

Each temporal expression, augmented with the information gathered about the words, is matched against a list of rules. The first rule that matches is used to process the expression. Each rule is expressed as a regular expression that tries to match the whole expression, or a subset of it, while capturing different meaningful parts together. Whenever a rule matches, the captured parts of the expression are processed in code in order to get normalized values we fill our temporal data structure with. Some expressions are composed of a combination of other smaller expressions, and therefore instead of repeating the same rules multiple times, we define some rules as a combination of the others. Whenever the general structure is matched, the expression is decomposed into smaller sections and each one of these is matched against some other possible rules, and processed accordingly. For example, the rule for:

من الرابع من شباط إلى الثامن من آذار

is defined in this manner:

(RULE 2) إلى *(RULE)* من

with *(RULE)* and *(RULE 2)* being some of the rules that are compatible with this expressions format, and that match the respective subparts. This abstraction saves us from having to rewrite the same rules multiple times, and allows us to extend the subrules once and have that change propagated to all the larger rules that use it.

Since an expression could match against different rules, the rules are checked according to an order of priority, from the one that gives us the most information to the one that gives us the least. We thus favor the rules that allow us to correctly fill a bigger part of our data structure.

Statistical Method: Ngram HMM tagger

For our purposes, we have decided to stick with a supervised n-gram Hidden Markov Model (HMM) tagger as it is a generative model which requires little structure in the data.

HMMs are temporal probabilistic models (i.e. probabilistic models which depend on observations) in which processes have states which are discrete random variables. A Markov model is a model where a state S_t is independent of all states with indexes less than $t-1$; and where an observation made at time t is generated by a process with a state S_t hidden from the observer. HMM is thus a generative model which tries to generate labels and data from distribution parameters. HMMs' ability to find patterns over time tends to our purpose of analyzing a sequence of words. Our sequence of words becomes a sequence of observations in a text, and each expression is one time step away from its direct predecessors. HMMs can be extended to output a sequence of labels by using the Viterbi algorithm (we will formalize this idea shortly).

N-grams on the other hand are very simply N sequences of words. N-grams are generally used in language models to predict the tag of the last word in an N-sequence of words given its predecessors in the same sequence, and to assign the probability of the occurrence of the whole sequence.
By using n-gram HMM taggers, we are using the tag context. Rather than assuming that the tag of the current word only depends on the previous state, we would in effect be making the probability of the tag of the current word depend on (N-1) previous states.
More formally, given a finite sequence of possible observed words ν and a finite sequence of tags κ we will define an n-gram HMM tagger as:

1) $p(t_N|t_1, ..., t_{N-1})$, the probability of seeing the tag t_N immediately after the (N-1)-gram of tags $(t_1, .., t_{N-1})$ where $t_N ... t_1 \in \kappa$
2) $p(x|t_n)$, the probability of pairing the observation x with the tag t_n, for $x \in \nu$ and $t_n \in \kappa$
3) We will define S to be the set of all sequences of $< x_1, .., x_n, t_1, .. t_n >$ paired words and tags, such that $0 \leq n$ and $N \leq n$, $x_1, ... x_n \in \nu$ and $t_1, .., t_n \in \kappa$

4) The probability of any $< x_1, .., x_n, t_1, ..t_n >$ is then:

$$p(x_1, .., x_n, t_1, ..t_n) = \Pi_{i=1}^{n} p(t_i|t_{i-1}, ..., t_{i-(N-1)}) \Pi_{i=1}^{n} p(x_i|t_i)$$

(1)

In our architecture:

- The set of tags is a bit vector which indicates which part of the data structure a word in an expression belongs to (i.e. FROM, TO etc.) and which granularity a word expresses, if any (i.e. HOUR, DAY, MONTH etc.).
- Each expression is passed to three HMM taggers: a unigram, a bigram and a trigram.
- The resulting tag is thus the argmax over the probabilities of the resulting tags presented by the unigram, bigram and trigram.
- In case several of the three HMM taggers presented the same tags, the resulting probabilities are normalized and summed before the argmax is computed.

Temporal Scoping

We will finally proceed onto scoping. Scoping is a feature implemented in our tool in order to either remove possible ambiguities and inconsistencies in our data structure, or in order to fill our data structure if the temporal expressions could not do so.

The basic design concept of scoping is based on the paper Coupled Temporal Scoping of Relational Facts by P.Talukdar, D. Wijaya and T. Mitchell.

The idea behind scoping is the following: given a large database of documents, one can count the number of occurrences of facts in each document in order to assess their validity across time. Scoping is uses the idea of collective inference over multiple temporal dependent facts. It allows for several useful features: eliminating noise i.e. facts which are not well represented by counts, and being able to proceed with scoping even if certain facts are not widely mentioned in the literature and the counts are very sparse. An example of coupled scoping are the facts President Clinton and Vice President Al Gore which might appear together in the data or during separate time ranges. These are two facts which are aligned and occurred in almost the same time frame. Scoping is done in two steps in our implementation: querying and collective inference. Our words are first queried and counted in a database.Then collective inferences based on constraints allow us to detect a words "activation date". This activation date is based on a threshold we deduced from a sensitivity analysis of our data. Possible straightforward constraints one can use in scoping are:

- Consistency: a given fact can only be true between its beginning and end times. It must also be true at the beginning time, true at the end time, but false at times beginning-1 and end+1.
- Functional: Under a certain relation, no two facts of that relation can be true at the same time. For example facts President Mubarak and President Sisi, are both linked by the relation presidency, but cannot be true at the same time.

- Single Span: Under a relation r, any fact must be true continuously during its active span of time, and must only have a single beginning and end.
- Point: A fact must happen at a single point in time under such a constraint. Its span of time is unit.
- Aligned: two facts under such a constraint must have the exact same temporal span. Containment: if a fact1 is contained within another fact2, it must have its timespan1 contained in timespan2. Succession: if a fact1 happens after another fact2, then its timespan1 must end before timespan2s beginning.
- Mutex: if fact1 and fact2 are mutually exclusive then they cannot happen at the same time, their timespans cant intersect.

III. PRELIMINARY RESULTS

Our main testing corpora consists of articles from two local Lebanese newspapers (An-Nahar, and Al-Mustaqbal). The table that follows (Table II) summarizes the analysis of the data.

	AnNahar		Al-Mustaqbal	
Number of Articles	599		8659	
Expressions by Size	Number	Percentage	Number	Percentage
1 word	2209	35.7	19008	31.3
2 words	1626	26.2	14946	24.7
3 words	1209	19.5	12331	20.3
4 words	551	8.9	6017	9.9
5 words	329	5.3	3941	6.5
6+ word	271	4.4	4389	7.2
Total	6195	100	60632	100

TABLE II

RESULTS OF NORMALIZATION ON NEWSPAPER CORPUS.

We notice that a majority of temporal expressions found in our corpus consist of one or two words, as reported by the ATSarf temporal entity extraction tool. Some of the notable most frequent expressions are words such as "أمس" (yesterday) and "اليوم" (today) which denote a relative date. These kinds of expressions can be captured by the rule system but need to be resolved to an absolute point in time based on a reference point. This is yet to be done at this stage.

In addition to that, we have developed an annotation tool for temporal expressions and their normalized values, as shown in the following figure (Fig. 2).
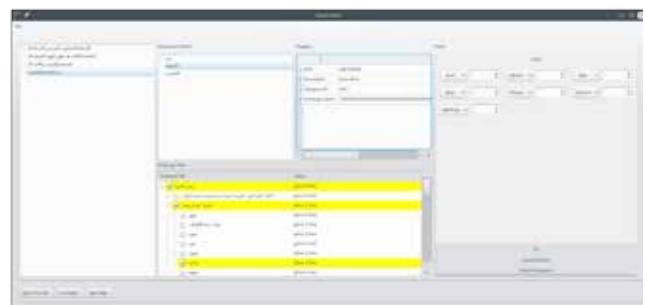

Fig. 2.   Screen capture showing the GUI of the tool

It displays the extracted temporal expressions and displays them in a friendly Graphical User Interface with the following features:

- Show the details of the morphological analysis of each word in the expression in the center-top panels
- Show the temporal extraction analysis results in the left panel and center panels
- Show the ontology analysis and highlight the match visually for inspection by the user in the center-bottom panel
- Allows the user to view the result of the rules and manually annotate each expression with values of every field in the data structure in the right panel

In order to evaluate the effectiveness of the rules and the HMM tagger, we have processed and tagged manually 40 documents from the An Nahar newspapers corpus. In total, we so far have an annotated corpus of 636 temporal expressions that we have used to train and test the HMM tagger. The training set and testing set for the HMM tagger are generated by randomly sampling 90% and 10% of the corpus respectively, and the results are measured on the feature basis (i.e. partial matches of the expression are taken into consideration). The tagger achieves 70% to 80% accuracy depending on the testing set.

IV. Conclusion

Our solution for the normalization of temporal expressions in Arabic into a set of numbers that can be processed, starts by augmenting the extracted temporal expressions with additional information such as matches in the ontology tree, part-of-speech, stem information and others. This information is then fed into a rule-based processor that tries to match and normalize some expressions using regular expressions. Some expressions that are ambiguous, or that do not have specific rules covering them, will not be matched at this stage and will continue through a statistical analysis that employs a Hidden Markov Model approach to try to normalize them. A user friendly graphical application accompanies our program and allows researchers and interested people to test the normalizer on custom files, as well as see the features of the different words, and annotate the data that was extracted. The annotated data is helpful for training the HMM. Our approach to the problem is promising and we hope to further develop rules that cover more cases as well as a more precise statistical model through larger training sets.

Once the expressions are normalized, they can be used in different applications such as the classification of historical documents, the ordering of events, or the automatic event extraction and scheduling from emails.

References

[1] F. Zaraket and J. Makhlouta, "Arabic Temporal Entity EXtraction using Morphological Analysis," International Journal of Computational Linguistics and Applications, vol. 3, no. 1, pp. 121-136, 2012.

[2] I. Mani and G. Wilson, "Robust temporal processing of news," ACL '00 Proceedings of the 38th Annual Meeting on Association for Computational Linguistics, pp. 69-76, 2000.

[3] A. Dehghan, "Temporal ordering of clinical events," arXiv preprint, no. arXiv:1504.03659, 2015.

[4] P. P. Talukdar, D. Wijaya and T. Mitchell, "Coupled temporal scoping of relational facts.," ACM, pp. 73-82, 2012.

[5] F. Schilder and C. Habel, "From Temporal Expressions to Temporal Information: Semantic Tagging of News Messages," Proceedings of the workshop on Temporal and spatial information processing, vol. 13, p. 9, 2001.

[6] D. Llid, R. Berlanga and M. J. Aramburu, "Extracting Temporal References to Assign Document Event-Time Periods," Database and Expert Systems Applications, pp. 62-71, 2001.

[7] M. Stede, S. Haas and U. Kssner, "Tracking and Understanding Temporal Descriptions in Dialogue," Technische Universitt Berlin, Berlin, 1998.

[8] A. Farghaly and K. Shaalan, "Arabic Natural Language Processing: Challenges and Solutions," ACM Transactions on Asian Language Information Processing (TALIP), vol. 8, no. 4, pp. 1-22, 2009.

[9] D. Jurafsky and J. H. Martin, Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition. Upper Saddle River, NJ: Prentice Hall, 2000.

[10] C. Allauzen, M. Mohri and B. Roark, "A General Weighted Grammar Library," AT&T Labs Research, Shannon Laboratory, Florham Park, 2005.

[11] Y. Benajiba, I. Zitouni, M. T. Diab and P. Rosso, "Arabic Named Entity Recognition: Using Features Extracted from Noisy Data.," in ACL 2010, Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, Uppsala, Sweden, 2010.

[12] M. Gross, "Local grammars and their representation by finite automata," In M. Hoey - Data, Description, Discourse, pp. 26-38, 1993.

[13] International Organization for Standardization, "Date and time format - ISO 8601," [Online]. Available: http://www.iso.org/iso/home/standards/iso8601.htm. [Accessed 25 October 2015].

[14] H. Traboulsi, "Arabic named entity extraction: A local grammar-based approach," Computer Science and Information Technology, 2009. IMCSIT '09. International Multiconference on, pp. 139-143, 2009.

[15] A. Alexandrescu and K. Kirchhoff, "Data-Driven Graph Construction for Semi-Supervised Graph-Based Learning in NLP," NAACL HLT, pp. 204-211, 2007.

Uni-Lab: A User Centric Web-Based Application for Laboratory Test Results

Freddy Vartabedian   Ibrahim El Kawam   Jad Tarabay

{fkv00, ike03, jat10}@mail.aub.edu

Abstract—-**In the absence of a centralized health system in many parts of the world, accompanied with the absence of electronic health records that involve multiple care providers (such as hospitals and private clinics), that serve to store a patient's complete medical history, patients usually have access to their laboratory test results in paper form or via emails. Over time, these results tend to get lost or end up in hard to organize chaotic piles; a fact that not only makes temporal tracking of changes in results difficult or even impossible, but also increases future health problems together with their related costs. From here comes the necessity of a single repository for lab test results that would store and organize a patient's data, make it available for him anytime and anywhere and at the same time satisfy the issues of data security, privacy and integrity. While numerous products in the form of applications and services are available, none address laboratory test results in a single application and more importantly, none provide a fully automated user-centric system. In this paper, we address the need for a system that would satisfy the aforementioned criteria. In this work we design an automated Web-based application, Uni-Lab, that will read the user's lab test results in different formats and from different sources and will serve to aggregate and store them. Uni-Lab, with its innovative server-based OCR technology, will completely digitize and categorically archive the user's medical lab results, in addition to providing a comprehensive and temporal visualization of the different results. The application will also provide the user with data access through a desktop, tablet, or a mobile phone, in addition to an optional sharing of data with selected medical care providers.**

I. INTRODUCTION

Generally, in our part of the world, where a centralized health system and electronic health records are uncommon, lab tests results are delivered to patients either in paper form or through emails. Over time and especially if different clinical providers are involved (as opposed to a single hospital as a source), these results tend to get lost or they end up in chaotic piles that become hard to organize, making tracking changes in health issues difficult. However, the huge advances in information technology and the extensive research and development in the area of electronic health records have given rise to various systems and applications under the general name of personal health records (PHR). These are generally user-centric applications or services that enable the user or patient to gather his/her entire medical or health related information and data into one place. According to the National Committee on Vital and Health Statistics (NCVHS), these systems serve to "improve wellness activities, improve understanding of health issues, and increase sense of control over health" for the consumer [1]. In fact, lab test results, which are an integral part of personal health records, fail to serve their purpose in the absence of a temporal tracking of changes and trends; a fact that not only increases the risk of future health problems, but also related costs. Numerous applications for PHRs exist in the market today and most rely on mobile based architectures where data is stored on the mobile device such as My Medical, and Capzule. Others make use of mobile applications but also provide cloud storage of data like Microsoft HealthVault. Smaller and more specific applications also exist such as Blood Test Pro for the visualization of blood test results on mobile phones (A detailed description of these products is provided in the literature review section). However, in our review of the available products, none was found to address the field of lab tests results as a whole. Moreover, most of the products reviewed were found to make use of a mobile system, which doesn't come without security and privacy concerns, especially when dealing with medical results which are to be dealt with high standards of security and privacy. Moreover, most products were found to lack automation and needed the manual entry of data.

Our goal in this project is to develop an automated web-based system, Uni-Lab, which would allow the user to securely and automatically store all his lab tests results in one place. The application will also organize the results in a categorical view, together with a temporal visualization of the data in a comprehensive and user- friendly interface. A graphical view will allow the user to view trends and deviations from averages. In this way, a user can make useful use of his/her lab test results and most importantly, he/she can monitor and keep track of his health from any location and anytime. An additional share option will allow the patient to share lab results with medical care providers of his choice and subject to his authorization.

II. RELATED WORK

A. PHR Designs

User-centric electronic personal health records have recently received substantial popularity and various designs and

systems have been developed by researchers in the field to improve existing functionalities.

In [2] a user or patient-centric PHR system is proposed with a web-based architecture that makes use of open-source tools and cloud computing to address the issue of storing and sharing electronic health records by the patient. The described system is one that can integrate patient records from different health care delivering systems like labs, doctors, hospitals, etc. and from any location. The web application in [2], runs on J2EE platform hosted by JBOSS and it uses a 2 component database system comprising dcm4che server and Microsoft SQL database layer. The authors suggest Amazon EC2 as a possible cloud computing network for development [2]. Data security and privacy is achieved through password protected access. Patients can only access their own records and other CDOs can only access the records that they are given access to by the patient.

In [3], another PHR system (PMHS: Personal Health Monitoring System) that makes use of the cloud and a mobile application for the android system is proposed. It consists of 4 modules: a clinical data collection module, cloud based file manager module, CDA (Clinical Document Architecture) file query module and diagnosis module [3]. For clinical data collection, the system uses an Android based mobile app that serves to collect the patient's medical records and generates CDA files for them. The application places the patient in control of his/her data in that he/she can control the sharing of his/her records. The gathered data is next uploaded to the cloud using systems like Dropbox. The application enables the user to retrieve desired records from the cloud, and convert them into a text file that will be used with the diagnosis module for potential disease alerts. Details of the XML parsing and mapping are provided in [3]. The authors of [4] proposes a patient-centric system of electronic health records(HER)in China that makes use of SD cards for medical data storage and complies with HL7 communication protocols (triggering events, messaging, information sharing, etc.). The system stores and deploys data in 4 different locations: the SD memory card, the hospital data center, the area or city storage center and a data backup center. Of particular interest in the proposed system is the use of an SD memory card that is the property of the patient to store his/her health records. The card in this system also contains a small database application that can run in Windows operating systems. A 2- fingerprint authentication (medical worker and patient) is needed for operating on the data on the card making the stored data tamper-proof [4].

In [5] a hybrid Web-based PHR design, Personal Health Manager (PHM) is described. With PHM, the patient's health records are owned solely by the patient, but the updates can only be performed by authorized medical personnel who are granted access to the data by the patient. The design in [5] uses role-based access with access tokens (over which the patient has control) and Apache Shiro, a tool that prevents users from using pages for which they have no authorization. The basic tools used include J Boss 6 (Java application server), Java 6, MySQL5, Java Enterprise Edition which includes Java Server Faces 2.0 (Java Web development framework), JPA (Java persistence API), Hibernate implementation, Enterprise Java Beans 3 and Apache Shiro

(for security). Security is ensured through X.800 architecture, including SSL and ID/password encryption, in addition to a timestamp in the Message Authentication Code (MAC) allowing only for a 1-time message acceptance.

The authors of [6] describe a system that allows for the integration of a patient's medical records in an already fully functional hospital DB and server with a mobile platform. In this system, a mobile server is designed to manage user ID/PW, access log and authorization to the corresponding hospital server. It receives a retrieval request and retrieves the required document through the hospital server in XML format followed by parsing it and converting it to an HTML document which is then sent to the mobile device for viewing. The mobile server in [6] is constructed with ASP.NET MVC4 and Visual Studio (C#) and the client side with Xcode (Objective-C). The server-client protocol is HTTP and JSON data interchange.

PHR Apps & Products	My Medical	Capzule	My Quest by Quest 360	Blood Test Pro	Blood Test Guide	3In1 Lab Values + Medical reference	Uni-Lab
Cloud Based Storage	✘	✘	✔	✘	✘	✘	✔
Automated Entry	✘	✘	✔	✘	✘	✘	✔
Multi-Platform	✔	✘	✔	✘	✘	✘	✔
Graphical Visualization	✔	✔	✔	✔	✘	✘	✔
Multi-Format Files	✔	✔	✔	✘	✘	✘	✔
Internet Connection	✘	✘	✔	✘	✘	✘	✔
Share	✘	✔	✔	✘	✘	✘	✔
Device	iOS(6.0 & higher) Mac(OSX) Android(2.2 or higher)	iOS 6.0 or later iPhone/iPad/iPod	All	iOS5.0 &6.0 iPhone/iPad	iOS 4.3	Android 2.3 & up	All

Table 1 PHR Products

B. PHR Applications and products

Various PHR applications, as well as apps for blood tests and ranges are available in the market such as HealthVault, MyMedical, Capzule, MyQuest by Quest 360, Zweena, LabCorp, Blood Test Pro, etc. The following table summarizes the significant features offered by some apps that can be considered as samples with respect to their features.

Of particular interest is the feature of automated data entry on the user side which is absent in all reviewed products in Table 1. Most of the existing apps require tedious manual entry of all medical results on the user side. MyQuest by Quest 360 (exclusive for Quest Diagnostic Medical services) and a few others like Zweena and LabCorp are exceptions in that the user's medical data collection, entry and organization are accomplished by their assigned professional staff. The user in such systems can access, view and share his results but cannot edit or change them [7, 8, 9, 10, 11, 12, 13, 14, 15].

Our review of related work and existing apps has led us to conclude that no single design or available application covers

clinical laboratory results in a single and secure Web-based application that spares the user the privacy issues associated with mobile systems; more importantly, none address the issue of automated data entry on the user side. In the absence of a secure application for the automated storing and managing of laboratory test results, the software we develop, Uni-Lab, will provide users with a simple, automated and secure way of storing their lab results and a comprehensive way of viewing their data anytime and from any location. Uni-Lab will also help them replace paper results that are prone to get lost with a permanent and reliable repository for their lab results with a simple click of an upload button and without the tedious task of manual data transfer commonly encountered in existing PHRs. In doing so, Uni-Lab will not only help the patient to easily monitor his/her health, but also understand the values of his/her tests and consequently increase his/her health-awareness.

III. METHODOLOGY

Uni-Lab comprises of 2 major components, the User Side which includes the patient, the laboratory and the medical practitioner, and the Server side. Our goal is to design a Web-based application and develop it fully so that it will store the user's medical results, categorize them, give a dedicated and informative visualization of the results and display the outcome of the results with data accessible through a desktop computer or a mobile phone or whatever can access our Web App online.
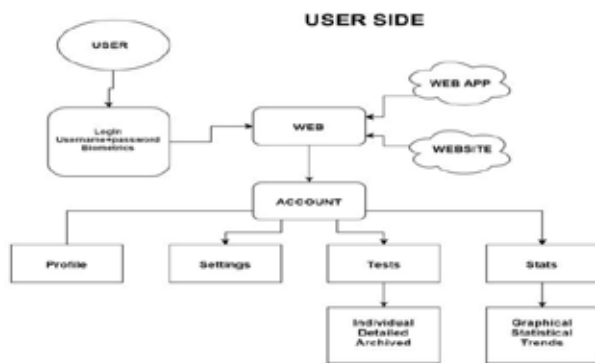

Figure 1 Front-End

The Front-End shown in Figure 1, deals mainly with the components and the method by which the user will interact with our Web App. Uni-Lab mainly relies on the users: patient or lab, providing us with their results by uploading them to our server, and it ensures a secure and private service that enables them to make use of the service with complete confidence. Using a standard encrypted login, in addition to an added layer of security through the use of biometrics, ensures that only the appropriate users can access their own accounts and share them with selected medical practitioners of their choice. The ability to access the Web App using almost all online connected devices makes it convenient for users to use the service at any time and place provided there is an internet connection during login and download of data. Upon logging in successfully the user (patient) will have full access to his/her account and all that entails, including Profile(where the account is personalized),Settings(Where preferences as

well as other information can be edited), Tests(where all the uploaded tests can be found sorted and archived) and Stats(where the user specified analyzed data can be viewed). Results can be uploaded by the patient himself/herself or by the lab. The patient can then share his/her results with the medical practitioner of his choice following a process of authorization.
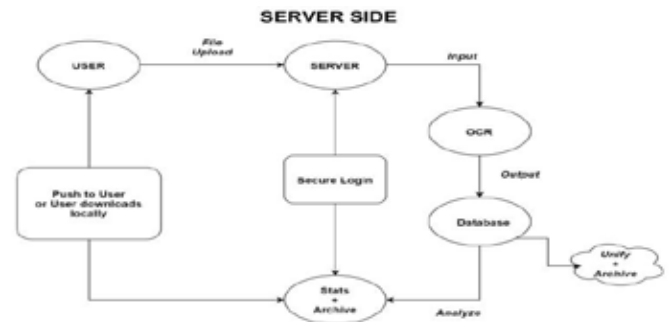

Figure 2 Back-End

The Back-End shown in Figure 2 above deals mainly with all components and actions done by our side whether on the actual server or on the dedicated computer that will be used for performing the OCR and Statistical Analysis. The Server receives all uploaded files by the users then automatically commences OCR parsing of the data followed by entering the data into our databases which is constantly backed- up and archived. This allows the analysis of the stored data to be done at any point in time. The Server itself is secured up to the highest standards applicable to ensure the integrity of our data; only authorized administrators can directly access server data. The Server allows any-time-place access to the data, as well as local download capability.

IV. DESIGN AND IMPLEMENTATION DETAILS

A. General Overview

Figure 3 displays a general overview of our Web-based application. The user side includes the patient, the lab and the medical care provider (doctor). Access to the data might be given to authorized statistical researchers at a later stage in the project.

The users can access the service via the device of their choice which can be smart phones, tablets and PCs over an internet connection. The patient after receiving his/her lab results document uploads it to the server. The document can be of any type including image, PDF, text and html files.
The server at this stage will automatically start processing the uploaded document. It will pass it through the OCR software (Tesseract). This process will convert the document into parsed format. This stage is followed by a parsing process that will extract the information relevant to the test results from the new file and will then send it to the database for storage. The patient can view his stored results anytime and anywhere. The application also enables a graphical visualization of the various parameters of the lab tests over time. The patient can also share his results with a medical care provider (doctor) of his choice via a process of authorization.

As for the lab, its function is to send the results to the patient or upload them directly to the server. The file transfers in the system will be encrypted using AES 256 for ensuring privacy and security, in addition to HTTPS protocol and a two-factor authentication that will include Username/ password and a biometric identification, mainly fingerprint identification. The server will also be protected by a firewall.
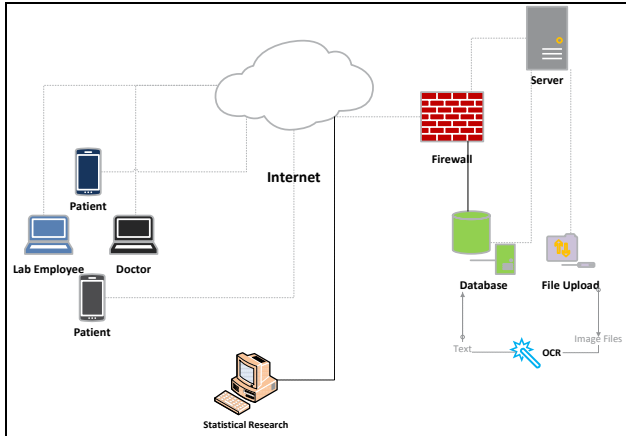

Figure 3 Application General Overview

B. OCR

What distinguishes Uni-Lab is its server side automated OCR technology implemented using Tesseract. In fact, a server-based OCR design, as explained in [16] is superior in quality and accuracy to "standalone" OCR mobile applications. [16] Mentions a text accuracy rate of 99.98% and a format accuracy of 99.8% for multiple pages processing with server-based OCR. As for Tesseract, in [17] an image processing technique is described as a necessary preprocessing step to achieve 97.1% OCR accuracy.

Therefore, for our server side OCR we implemented an image preprocessing step using Image Magic library. We used built-in shell scripts to refine the uploaded image by adjusting the contrast and color balance. Additionally, we removed outlier pixels and shades from the image using *textcleaner* shell script from the same library. The script works on removing any unnecessary pixels that may affect the written text. The output of this step is a clear text image that has a high contrast and a sharp white and black color balance. This would give a better image with the same high quality resolution which will be the input for the next step. The used arguments in the image refining process are shown in Table 2 OCR arguments' best values. We've concluded with a set of values that returned a very acceptable image quality.

Argument name	Description	Value
-g	Grey scale	On
-e	Enhance brightness	Stretch
-f	Filter size	25 px
-o	Noise offset	10 px
-s	Sharpness	4 (scale)
-T	Trim background	On
-p	Pad around border	10 px
-Psm	Image segmentation	tabular

Table 2 OCR arguments' best values

The next step is executing the OCR script from Tesseract library. Since it is an open source, Tesseract accepts training data from the user in order to raise detection accuracy and use user-defined strings in the detection process. We've downloaded shared training data for the OCR from the library website and used it in the process. The input arguments for the OCR would be the image file, language, and the training dataset to be used in detection. The output of this step is a text file containing the converted strings contained in the processed image.

All the pre-mentioned steps are automated on the web server side, where we used WAMP web server to host and execute the scripts using PHP. Image refining and processing scripts are called through the PHP code after uploading the image by the user. The output of image preprocessing will trigger the call for the OCR script which will output the text contained in the image and auto save it on our web-server in a specified directory. All files are saved with a timestamp for later use. The final step would be reading and parsing the output text and save it to our database for analysis.

Adding the results to the database is done through a parser that reads each section of the document and saves it in the results table. The results table will contain each patient's lab-test results to be used in future analysis.

C. Graphical Visualization
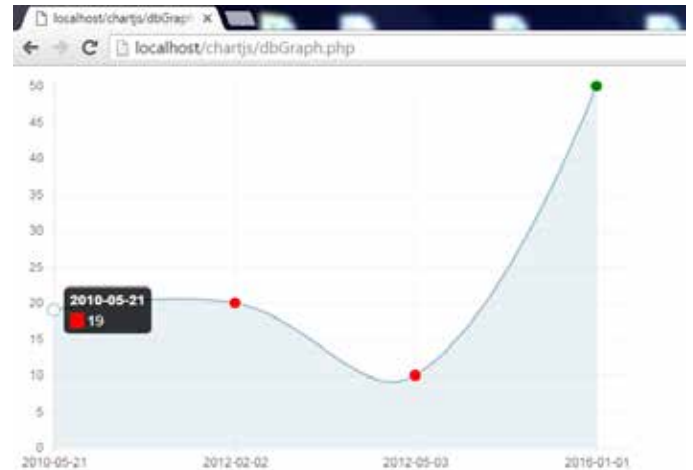
Figure 4 Graph extracted data

Uni-Lab makes use of the open-source library Chart.js to provide patients and medical practitioners with a historical graphical visualization of clinical results. Figure 4 shows a sample graph that uses extracted data from a placeholder DB. The red dots represent abnormal values whereas the green ones indicate results falling within the specified normal range.

D. Tools and Languages

We used HTML, JavaScript and PHP for our web application. HTML and CSS were used for building web pages and forms to upload results and visualize them. Our pages are user friendly and simple where we used CSS templates to add effects and changes to HTML pages. JavaScript is used mainly for client side validation to check user inputs and uploaded files. We used JSON strings and

arrays to transfer data between client and server side because it is light weight and much faster than XML and any other data structure language. Finally, we used MySQL as our database, because it is compatible with PHP and very lightweight on the web server. Additionally, MySQL can handle large files and keep a remarkably low latency in retiring query results. As for the management of the database, it was achieved with phpMyAdmin.

E. Access Control

All users are authenticated using a saved combination of a username and a hashed password. Once the user is registered the username/password combination will be saved in a database table on the web server. For every user we generate a verification code based on the name, current date and time, and email. After the authentication each user can access a subset of authorized pages and forms that corresponds to his profile (i.e. patient or doctor). For example the patient can see a list of available doctors and contact them. On the other hand the patient can't access other patients' information.

V. CONCLUSION

In this paper, we have described Uni-Lab, a user-centric and Web-based application, to provide users with a simple and automated management of clinical laboratory results. The design is based on a server-side automated OCR technology with Tesseract. Uni-Lab offers a friendly and easy to use interface, a secure 2-factor access control, low latency in results queries and a graphical visualization of trends; which are all features that simplify clinical results storage and retrieval. Uni-Lab aims at building a long term clinical lab results storage and history which in turn serve to raise health awareness.

ACKNOWLEDGEMENT

REFERENCES

[1] United States, National Committee on Vital and Health Statistics.; United States, Dept. of Health and Human Services.; National Center for Health Statistics (U.S.), (2005), "Personal health records and personal health record systems; a report recommendation from the National Committee on Vital and Health Statistics" [Online]. Available:http://stacks.cdc.gov/view/cdc/6758

[2] V. Ved, V. Tyagi, A. Agarwal, A. S. Pandya, "Personal Health Record System and Integration Techniques with Various Electronic Medical Record Systems", High Assurance Systems Engineering, IEEE 13th International Symposium, 2011. DOI: 10.1109/HASE.2011.63

[3] Y. T. Song, S. Hong, J. Pak, "Empowering Patients Using Cloud Based Personal Health Record System," Software Engineering,Artificial Intelligence, Networking and parallel/Distributed Computing,16 IEEE/Acis International Conference, 2015. DOI:10.1109/SNPD.2015.7176216

[4] G. Ma, J. Liu, Z. Wei, "The Portable Personal Health Records:Storage on

SD Card and Network,Only for One's Childhood," IEEE International Conference on Electrical and Control Engineering, 2010. DOI: 10.1109/iCECE.2010.1168

[5] J. Israelson , E. C. Cankaya, "A Hybrid Web Based Personal Health Record System Shielded with Comprehensive Security," 45 Hawaii IEEE International Conference on System Science, 2012, pp. 2958-2968. DOI: 10.1109/HICSS.2012.61

[6] Y. A. Kim, S. S. Kim, S. Kang, K. Kim, J. Kim, "Development of Mobile Platform Integrated with Existing Electronic Medical Records", Healthcare Informatics Research by Korean Association of Medical Journal editors, vol. 20, Issue 3, pp. 231-235, July 2014. DOI: 10.4258/hir.2014.20.3.231

[7] "Microsoft Health Vault" [Online]. Available: https://www.healthvault.com/lb/en. (Accessed: October 14, 2015).

[8] "My Medical" [Online]. Available: http://mymedicalapp.com/. (Accessed: October 13, 2015)

[9] "Capzule" [Online]. Available: https://www.capzule.com/. (Accessed: October 14, 2015)

[10] "My Quest by Care 360" [Online]. Available:https://myquest.questdiagnostics.com/web/home. (Accessed: October 19, 2015)

[11] "Blood Test Pro" [Online]. Available: http://www.blood-test-pro.com/. (Accessed: October12, 2015)

[12] "Blood Test Guide-Analyze Your Blood Test Report" [Online]. Available: https://itunes.apple.com/us/app/blood-test-guide-analyze-your/id491681195?mt=8. (Accessed: October 21, 2015)

[13] "3in1 Values + Medical Reference", [Online]. Available: http://www.imedicalapps.com/2013/12/3-1-lab-values-medical-reference-android-app/

[14] "Zweena" [Online]. Available: http://zweenahealth.com/

[15] "LabCorp" [Online]. Available: https://www.labcorp.com/wps/portal/patient/results

[16] T. Mantoro, A. M. Sobri, W. Usino, "Optical Character Recognition (OCR) Performance in Server-Based Mobile Environment," IEEE International Conference on Advanced Computer Science Applications and Technologies, pp. 423-428, 2013. DOI: 10.1109/ACSAT.2013.89

[17] T. Tuna, J. Subhlok, S. Shah, "Indexing and Keyword Search to Ease Navigation in Lecture Videos," IEEE Applied Imagery Pattern Recognition Workshop (AIPR), pp.1-8, 11-13 Oct. 2011. DOI: 10.1109/AIPR.2011.6176364

Vehicles Recognition from Partial Images

Salwan M. Alwan

Electrical and Computer Engineering Department, Rafik Hariri University, Meshref, Lebanon

alwansm@students.rhu.edu.lb

Abstract— **Vehicles recognition from partial images can be used in many security applications. Towards this goal, we present in this paper a vehicles recognition system based on SIFT features and the use of SVM to classify the types of cars. A benchmark of pictures (BMW and Mercedes) was used to explain the benefits of the proposed method.**

1. INTRODUCTION

Most, if not all, of the vehicles used in illegal activities have their paint changed and their license plate replaced by that of another vehicle.

In this paper a system is described which identifies classifies cars from their partial pictures. The suggested system uses Scale-invariant feature transform (SIFT) to extract features from cars and Support vector machine (SVM) to classify the type of cars. A plate recognition software is integrated with this system to compare the collected data with the government database to identify stolen cars or any cars used in a criminal activity.

The outline of the paper is given as follows: Section 2 presents related previous work on vehicles recognition. We present in Section 3 the proposed method of recognition and classification. Section 4 describes the experimental results obtained for the classification between two famous types of cars (Mercedes and BMW). The conclusion and some perspectives of this work are given in Section 5.

2. RELATED WORK

Vehicles recognition is essential and can be used in many security applications such as identification of stolen cars. Despite this, few researches have been conducted on this subject. (Sparta Cheung)[1] recommended that the problem should be to investigate the sample using a volumetric framework in combination with an appearance model. (Ferencz 2008)[2] considered the problem to be addressed as a visual identification and attempted to learn distinguished appearance plots. (Lowe 1999)[3] Used Scale Invariant Feature Transforms (SIFT). However, the approach to identification has been more of a general object categorization. (Mahasin Dimassi)[5] proposed a method that used SIFT, SURF and Hausdorff distance to match cars and classify them.

(Sparta Cheung and Alice Chu)[6] proposes a technique in performing make and model recognition given an image of an unidentified car viewed from an arbitrary angle. (David João Adão dos Santos)[7] recommended a method to identify the vehicle's manufacturer and model based on the external shape and the light shape.

In this paper we aim to contribute by presenting a vehicle recognition method that uses SIFT-based features and SVM to classify the type of cars.

3. PROPOSED METHOD

The system is trained using pictures which are already labeled. Both, pictures from manufacturer's catalogs as well real pictures taken by ordinary people, are processed by image processing software to extract features. Those features are tested using Kruskal-Wallis (KW) test to choose significant features then those significant features are used by SVM to learn how to identify the vehicle from a picture as shown in Figure 1.

3.1 Feature Extraction

In this part, several features where extracted using SIFT. For the descriptor vector computation, several methods exist such as the scale-invariant feature
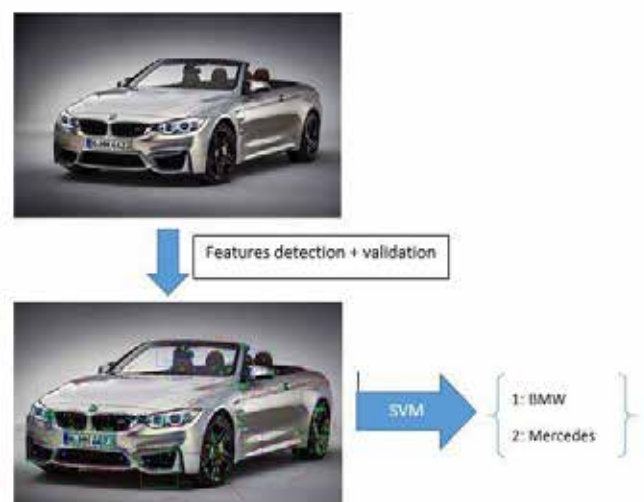

Figure 1 General scheme of the proposed method

transform [3], shape contexts [8], and speed up robust features (SURF) [9].In our study, we have used the SIFT algorithm since a comparison study presented by

Mikolajczyk and Schmid [10] shows that SIFT is better than the other methods. SIFT algorithm has been also successfully used in biometric recognition for different modalities such as the veins [11], face [12], fingerprint [13], iris [14] as well as in 3D facial recognition [15]. SIFT algorithm consists of four major stages: 1) scale space extrema detection, 2) key point localization, 3) orientation assignment, and 4) key point descriptor. In the first stage, potential interest points are identified, using a difference-of-Gaussian function, that are invariant to scale and orientation. In the second stage, candidate key points are localized to sub-pixel accuracy and eliminated if found to be unstable. The third stage identifies the dominant orientations for each key point based on its local image patch. The key point descriptor in the final stage is created by sampling the magnitudes and orientations of the image gradients in a neighborhood of each key point and building smoothed orientation histograms that contain the important aspect of the neighborhood. Each local descriptor is composed of a 4×4 array (histogram). For each coordinate of this array, an eight-orientation vector is associated. A 128-elements 8×(4×4))vector is then built for each key point. In other words, each image im is described by a set of invariant features X(im) ={ ki = (si, sci, xi, yi)| i = 1:N(im)} where si is the 128-elements SIFT invariant descriptor computed near the key point ki, (xi,yi) is the position of ki in the original image im, sci is the scale, and N(im) is the number of detected key points in image im. The features extracted are invariant to image scaling and rotation and partially invariant to change in illumination and 3D camera view point. From these features, six criteria are collected (see Table 1):

- Scale: value is the sub level adjusted scale value
- Size: is the size of the feature in pixels on the original image
- Edge flag: is zero if the feature is classified as an edge
- Edge orientation: is the angle made by the edge through the feature point
- Scale space curvature: is a rough confidence measure of feature prominence
- Descriptor vector

For each criterion, we have calculated the mean, median, variance, standard deviation, and the interquartile range. Except the descriptor vector, we have calculated the DC coefficient of the matrix Ms, with N(im) rows and 128 columns, related to SIFT invariant descriptor for si, i = 1:N(im) where N(im) is the number of detected key points for image im.

3.2 Kruskal-Wallis Test

After extracting features from car images we use the KW test which will determine which features are significant as shown in Figure 2. It is a non-parametric (distribution free) test, which is used to decide whether K independent samples are from the same population. In other words, it is used to test two hypothesis given by equation (1) the null hypothesis H0 assumes that samples originate from the same population (i.e., equal population means) against the alternative hypothesis H1 which assumes that there is a statistically significant difference between at least two of the subgroups

$$(1) \quad \begin{cases} H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \\ H_1: \mu_i \neq \mu_j \, \exists(i,j) \, where \, i \neq j \end{cases}$$

The KW test statistic H is given by equation (2), and the p-value is calculated using a χ2 distribution with k−1 degrees of freedom. The decision criterion to choose the appropriate hypothesis is given in equation (3).

$$(2) \quad H = \frac{12}{N(N+1)} \sum_{i=1}^{g} n_i r_i^2 - 3(N+1)$$

Where n_i is the number of observations in group i, r_{ij} is the rank of observation j from group i and N is the total number of observations across all groups.

(3)

$$r_i^2 = \frac{\sum_{j=1}^{n_i} r_{ij}}{ni} \, and \, \bar{r} = \frac{1}{2}(N+1)$$

$$\begin{cases} p - value \geq 0.05 & accept \, H_0 \\ otherwise & reject \, H_0 \end{cases}$$

Figure 2 scheme of Kruskal-Wallis test
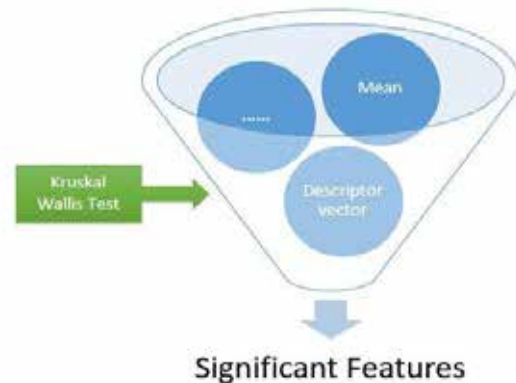
3.3 SVM

In order to classify the input image, we use SVM. From all existing classification schemes, a SVM-based technique has been selected due to high classification rates obtained in previous works [16] and to their high generalization abilities. SVMs have been proposed by Vapnik [17] and are based on the structural risk minimization principle from statistical learning

theory. SVMs express predictions in terms of a linear combination of kernel functions centered on a subset of the training data, known as support vectors (SV).

Suppose we have a training set $\{x_i,y_i\}$ where xi is the training pattern and y_i the label. For problems with two classes, with the classes $y_i \in \{-1,1\}$, a support vector machine [17] implements the following algorithm. First, the training points $\{x_i\}$ are projected into a space H (of possibly infinite dimension) by means of a function ϕ ($\cdot$).The second step is to find an optimal decision hyperplane in this space. The criterion for optimality will be defined shortly. Note that for the same training set, different transformations ϕ ($\cdot$) may lead to different decision functions.

A transformation is achieved in an implicit manner using a kernel K ($\cdot$,) and consequently the decision function can be defined as:

$$(8) \quad f(x) = \langle w, \phi(x) \rangle + b = \sum_{i=1}^{l} \alpha_i^* y_i K(x_i, x) + b$$

with $\alpha_i^* \in R$. The values w and b are the parameters defining the linear decision hyperplane. In SVMs, the optimality criterion to maximize is the margin, that is to say, the distance between the hyperplane and the nearest point ϕ (x_i) of the training set. The α_i^* which optimize this criterion are obtained by solving the following problem:

$$(9) \quad \begin{cases} max_{\alpha_i} \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{ij=1}^{l} \alpha_i \alpha_j \ \gamma_i K(x_i, x_j \gamma_j) \\ with\ constraints, \\ 0 \le \alpha_i \le C \\ \sum_{i=1}^{l} \alpha_i \gamma_i = 0 \end{cases}$$

where C is a penalization coefficient for data points located in or beyond the margin and provides a compromise between their numbers and the width of the margin. In this paper, we have used WEKA [18] to train SVM.

4. EXPERIMANTAL RESULTS

The goal of the proposed method is to recognize the types of cars from partial images. In the following experiment, we will use binary classification to predict if the input car is BMW or Mercedes.

In Section 4.1, we present the experimental protocol. In Section 4.2, we present the results of KW test followed by the SVM classification results in Section 4.3.

4.1 Protocol

A benchmark of 150 images of different Mercedes cars (Figure 3) and 150 images of different BMW (Figure 4) were used in this study to clarify the benefits of the presented method.

4.2 Kruskal-Wallis Test results

After extracting features, we used the KW test which will return for us significant features to be used as input to the SVM classifier. According to the explanation presented in Section 3.2, the interpretation of results

Figure 3 Samples of Mercedes car images

Figure 4 Samples of BMW car images

should be done according to the p-values. Each feature that has a p-value less than 0.05 would be significant. Table 1 presents the obtained *p-value* for each extracted feature. According to this table, all features are significant except the mean of size of feature on image, edge orientation and curvature.

In addition, the interquartile range of size of feature on image and edge orientation is not important. Furthermore, the median of edge orientation and curvature are not significant. Those features are not significant because their p-*value* is larger than *0.05*. Furthermore, Table 1 shows also that the DC coefficient is important to be used as input to the SVM (With a p-*value* = 0).

4.3 SVM Classification results

In this part, the significant features (*p-values* less than *0.05*) that were obtain from the KW test are used as input to SVM to classify the type of cars either BMW or Mercedes.

Using SVM, we have obtained results of 80% correctly classified instances using 10-fold cross-validation. Those results are promising due to the small set of data.

	Scale	Size of feature	Edge flag	Edge orientation	Curvature	Descriptor vector
DC coefficient						0
Mean	0	0.539	0	0.989	0.064	
Variance	0	0	0	0.01	0	
Standard	0	0	0	0.001	0	
Interquartile	0.035	0.687	0	0.177	0	
Median	0	0	0.03	0.435	0.704	

Table 1 *p-values of features*

5. CONCLUSION AND PERSPECTIVE

Car make and model recognition is an unfamiliar field in machine vision, but some progress have been made in several areas that can help further studies. Our work in achieving vehicles recognition from partial images technique extracts SIFT-based features and uses the SVM classifier for identifying cars' type. The tests conducted on 300 images (BMW and Mercedes) have shown promising classification results that would pave the way for our vision to reach the future researchers.

For the perspectives of this work, our aim is to build a significant benchmark that serves as the building block of future researches for extending our work to cover simultaneously different types of cars by using a multi-class SVM. Furthermore, combining this work with license plate recognition software will help in security applications such as identification of stolen or engaged-in-criminal-activity cars.

6. REFERENCES

[1] R. Farrell, O. Oza, N. Zhang, V. Morariu, T. Darrell, and L. Davis, "Birdlets: Subordinate categorization using volumetric primitives and pose-normalized appearance," in Computer Vision, 2011. ICCV 2011. IEEE 13th International Conference on, Nov. 2011.

[2] A. Ferencz, E. Learned-Miller, and J. Malik, "Learning to locate informative features for visual identification," International Journal of Computer Vision, vol. 77, pp. 3–24, 2008.

[3] D. Lowe, "Distinctive image features from scale-invariant keypoints", IJCV, 2(60):91–110, 2004.

[4] Learning, Modeling, and Classification of Vehicle Track Patterns from Live Video Brendan Tran Morris, Student Member, IEEE, and Mohan Manubhai Trivedi, Senior Member, IEEE

[5] Mahasin Dimassi, Partial Image recognition, Masters in Communications and Electronics Engineering, Beirut Arab University, 2014

[6] [1]Sparta Cheung, [2]Alice Chu, Make and Model Recognition of Cars, [1]Department of Electrical and Computer Engineering, [2]Department of Computer Science, Winter 2008

[7] David João Adão dos Santos, Automatic Vehicle Recognition System, Master in Electrical and Computer Engineering, September 2008

[8] S Belongie, J Malik, J Puzicha, inInternational Conference on Computer Vision. Matching shapes (IEEE, 2001), pp. 454–461

[9] H Bay, A Ess, T Tuytelaars, L VanGool, Speeded-up robust features (SURF). Comput. Vis. Image Underst. 110,346–359(2008)

[10] K Mikolajczyk, C Schmid, A performance evaluation of local descriptors. IEEE Trans. Pattern Anal. Mach. Intell. 27,1615–1630(2005)

[11] P-O Ladoux, C Rosenberger, B Dorizzi, in the 3rd IAPR/IEEE International Conference on Biometrics (ICB'09). Palm vein verification system based on SIFT matching (Springer-Verlag Berlin, Heidelberg, 2009), pp.1290–1298

[12] DR Kisku, A Rattani, E Grosso, M Tistarelli, in 5th IEEE International Workshop on Automatic Identification Advanced Technologies (AUTOID'07). Face identification by SIFT-based complete graph topology (IEEE Alghero, 2007), pp.63–68

[13] U Park, S Pankanti, AK Jain, in Proc. SPIE 6944, Biometric Technology for Human Identification. Finger print verification using SIFT features (IEEE Orlando, Florida, 2008)

[14] F Alonso-Fernandez, P Tome-Gonzalez, V Ruiz-Albacete, J Ortega-Garcia, in IEEE Proc. Intl. Conf. on Biometrics, Identity and Security (BIDS). Iris recognition based on SIFT features (IEEE, 2009), pp.1–8

[15] S Berretti, AD Bimbo, P Pala, BB Amor, M Daoudi, in Proceedings of the 20th International Conference on Pattern Recognition (ICPR). A set of selected SIFT features for 3D facial expression recognition (IEEE Istanbul, 2010), pp.4125–4128

[16] G Lebrun, C Charrier, O Lezoray, C Meurie, H Cardot, in the 11th International Conference on Computer Analysis of Images and Pattern (CAIP). Fast pixel classification by SVM using vector quantization, tabu search and hybrid color space (Springer Heidelberg, 2005), pp. 685–692

[17] VVapnik, The Nature of Statistical Learning Theory. (Springer, New York, 1995)

[18] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten (2009); The WEKA Data Mining Software: An Update; SIGKDD Explorations, Volume 11, Issue 1

DEPARTMENT OF MECHANICAL ENGINEERING

CONTENT

Application on Periodic Structures Under Seismic Load Using Recursive Finite Element Approach

Reem Yassine and Faten Salman

Mechanical Engineering Department, School of Engineering
Lebanese International University
Beirut, Lebanon

Abstract— Frequency response function is a significant dimension for structural analysis. For an accurate result of this function, large number of cells is to be used, which will lead to a high amount of computational time. Moreover, complex periodic structures also require high amount of computational time costs for the relative study. Recursive method is a finite element method used for computing the frequency response function, independent of the number of cells. It is based on eliminating the common degrees of freedom between a cell and it's succeeding one, for the nodes that are under no load. By means of products and inverses, the ultimate dynamic stiffness matrix will have a particular dimension, depending on the boundary and interior degrees of freedom. Computing the frequency response function will consume low amount of time costs, governed by high frequencies. The presented method is to be applied on periodic structures under seismic load.

Keywords—finite element analysis; seismic load; vibrations; periodic structure; recursive method

I. INTRODUCTION

Recently, most studies were centered on modelling periodic structures under seismic loading. For that, getting the frequency response function, which is a major dimension used for modelling, is the objective of such studies. For having an accurate result, large number of cells is to be considered. This will lead to large amount of computations, which will increase the computational time.

There are several approaches investigated to find the frequency response for structures. Starting with the spectral finite element method (SFEM), it uses general uniform structures with complex cross-sections, where all boundary conditions are to be considered. Using finite element method (FEM), the displacement in the cross section is studied. The variation along axis of symmetry is expressed by means of ordinary differential equation. The dynamic stiffness matrix is established by applying the virtual work method. SFEM uses merely one method, but it requires long computational time. Various scientists used spectral finite element approach on different applications. It was mostly used by Finnveden [1, 2]. Nilsson [3] applied SFEM on plates and shells. While, for excited structures under high frequencies, a competent approach is used which is dynamic stiffness matrix (DSM) [4]. This method divides the member under study into distinct element. Then, for finding the dynamic stiffness matrix, the relationship between the nodal displacement and forces of element is defined. A solution with higher accuracy is obtained with less computational time. A relationship between the nodal displacement and forces of the element is calculated to derive the dynamic stiffness matrix. Plane wave and interaction of fluid structure were examined by Birgersson [5-7]. Gry and Gravic [8, 9] used similar approaches for finding the wave propagation on rails, where relative dispersion relations were found. Bartoli and Marzani [10, 11] analyzed similar techniques, where they calculated the dispersion relations for damped structures of arbitrary cross-sections, and with symmetric elements. The deformation on each point on the cross-section was calculated by Dong and Aalami [12, 13] using finite element analysis. These methods aimed to find the frequency response using number of operations proportional to that of cells.

Duhamel [14, 15] interpreted the frequency response using recursive method (RM) on waveguide structures, such as beams, plates and tyres. Beams were examined as a 2 cell element. Increasing the number of cells lead to an increase in the accuracy of the results. Plates under excitation in the mid x and y position were also studied. After meshing the simply supported plate, the mass and stiffness matrices drawn from ABAQUS will be placed in a MATLAB program to find the frequency response, using recursive method.

Demonstration of the efficiency of the recursive method is the intent of this paper for periodic structures that cannot be designed as waveguides. The study of large structures with complex geometries will be facilitated by using recursive method. It is a method used for computing the global dynamic stiffness matrix by means of products and inverses of matrices, which have the same dimension as the dynamic stiffness matrix of a single cell. The structure is modeled and meshed using conventional finite element software (ANSYS). Then, the mass stiffness and damping matrices found will be post processed in a MATLAB program for computing the frequency response function. Nodes that are under no load or not under study will be omitted using this method. Hence, internal degrees of freedom between the adjacent cells will be recursively removed. Trusses and frames are structures characterized by simplicity of geometry, and have in their structural composition inability to guide waves along their longitudinal axis. Cranes and buildings were modeled as trusses and frames, respectively, under various loading.

Periodic structures were examined in this paper. For the first application, 1-D bar structures are assembled in a crane. When it is under excitation it will be dealt with as a truss. Seismic loading was also examined on cranes. On the second hand,

application of recursive method will be on 2-D frame, where a harmonic displacement is loaded at its base (model of seismic load). Frame is considered as a 2-D periodic structure. In this procedure, an accurate response will be found with few computations compared to the number of computations used in finite element analysis.

II. RECURSIVE METHOD

A. Introduction

Recursive method is a computational method that calculates the frequency response function of structures under loading. Finding the frequency response using other methods consume much higher computational time. This method facilitates identifying the global dynamic stiffness matrix, without spending much time for the relative calculations. The structure is modeled and assembled by conventional finite element analysis, resulting with mass, stiffness, and damping matrices for a one cell structure.

The Dynamic stiffness matrix of the section studied, is built after post processing the obtained matrices using MATLAB program. Using the recursive method, the global dynamic stiffness matrix is found. This matrix is obtained by products and inverses of matrices with the same dimensions as the dynamic stiffness matrix of one cell.

The structures studied, truss and frame consists of N identical cells. The response is denoted by u. Following is a demonstration for the recursive method, numerically [14].

B. One cell element

To start with contemplating one cell element under a time dependent loading $e^{-i\omega t}$, the discrete dynamic equation of this cell is given by:

$$(K + i\omega C - \omega^2 M)q = f \tag{1}$$

The angular frequency is denoted by ω. Stiffness, damping, and mass matrices are denoted by K, C and M, respectively. The displacement vector is q and the loading vector is f.

Then the dynamic stiffness matrix is as follows:

$$\widetilde{D} = K + i\omega C - \omega^2 M \tag{2}$$

After separating the degrees of freedom into boundary (B) and interior (I), the dynamic equation will be given by:

$$\begin{bmatrix} \widetilde{D}_{BB} & \widetilde{D}_{BI} \\ \widetilde{D}_{IB} & \widetilde{D}_{II} \end{bmatrix} \begin{bmatrix} q_B \\ q_I \end{bmatrix} = \begin{bmatrix} f_B \\ 0 \end{bmatrix} \tag{3}$$

No external force exists on the internal degrees of freedom, then the second row in equation (3) can be eliminated. This will result in a relation between q_I and q_B.

$$\widetilde{D}_{IB}q_B + \widetilde{D}_{II}q_I = 0$$
$$q_I = -\widetilde{D}_{II}^{-1}\widetilde{D}_{IB}q_B \tag{4}$$

Then the first row of equation (3) will be as shown in equation (5).

$$f_B = (\widetilde{D}_{BB} - \widetilde{D}_{BI}\widetilde{D}_{II}^{-1}\widetilde{D}_{IB})q_B \tag{5}$$

This shows that, only the boundary conditions are to be considered in the study.

Boundary conditions will also be separated into left (L) and right (R) for each node.

After decomposing into right and left, equation (5) will be:

$$\begin{bmatrix} f_L \\ f_R \end{bmatrix} = \begin{bmatrix} D_{LL}^{(1)} & D_{LR}^{(1)} \\ D_{RL}^{(1)} & D_{RR}^{(1)} \end{bmatrix} \begin{bmatrix} q_L \\ q_R \end{bmatrix} = D^{(1)} \begin{bmatrix} q_L \\ q_R \end{bmatrix} \tag{6}$$

$D^{(1)}$ is the dynamic stiffness matrix of one cell. It is a symmetric matrix since it consists of K, C and M matrices which are symmetric.

C. Assembly of two cells

Consider two consecutive cells 1 and 2 shown in "Fig. 1".
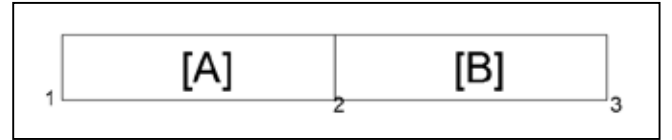
Figure 1. Two cell element

The dynamic stiffness matrices are A and B respectively. The dynamic stiffness matrix is denoted by $D^{(2)}$ it relates the degrees of freedom between left and right side. The structure's dynamic stiffness matrix is calculated by

$$\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} A_{LL} & A_{LR} & 0 \\ A_{RL} & A_{RR} + B_{LL} & B_{LR} \\ 0 & B_{RL} & B_{RR} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} \tag{7}$$

The interior node is under no load, $f_2 = 0$
The second row in equation (7) will give a relation between q_1, q_2 and q_3.

$$q_2 = -(A_{RR} + B_{LL})^{-1}(A_{RL}q_1 + B_{LR}q_3) \tag{8}$$

Then, for a two cell structure the global dynamic stiffness matrix is given by:

$$\begin{bmatrix} f_1 \\ f_3 \end{bmatrix} = \begin{bmatrix} D_{LL}^{(2)} & D_{LR}^{(2)} \\ D_{RL}^{(2)} & D_{RR}^{(2)} \end{bmatrix} \begin{bmatrix} q_1 \\ q_3 \end{bmatrix} = D^{(2)} \begin{bmatrix} q_1 \\ q_3 \end{bmatrix} \tag{9}$$

$D^{(2)}$ is the dynamic stiffness matrix of the two cell structure.

$$D^{(2)} = \begin{bmatrix} A_{LL} - (A_{RR} + B_{LL})^{-1}A_{RL} & -A_{LR}(A_{RR} + B_{LL})^{-1}B_{LR} \\ -B_{RL}(A_{RR} + B_{LL})^{-1}A_{RL} & B_{RR} - B_{RL}(A_{RR} + B_{LL})^{-1}B_{RL} \end{bmatrix}$$

This matrix is symmetric, since it consists A and B matrices which are symmetric. Thus equation (9) can be written in the form:

$$D^{(2)} = \{A, B\} \qquad (10)$$

D. General case

For studying a general case, consider a structure consisting of N identical cells, where the internal nodes are under no load. The total dynamic stiffness matrix is denoted by $D^{(N)}$. The structure is periodic and isotropic.

For the structure made up of identical cells, recursively remove the internal degrees of freedom which are under no load. Thus the element which is made of two cells will be considered as one. The current one cell element will be assembled with an identical cell element. Doing the same procedure for the whole structure, the resultant will be a one cell element that includes all nodes under.

To avoid any limitation in the number of cells, the number of operations will be $log_2 N$. An illustration for the whole process is shown in "Fig. 2".

Steps	Number of elements	Demonstration
1	2 cells	
2	4 cells	
3	8 cells	
⋮	⋮	⋮
logNlog2	N cells	

Figure 2. Procedure of removing cells

After computing the frequency response function conventionally (using ANSYS), the response will be compared to that determined from the recursive method. RM is defined by a self-written MATLAB code where it takes the K, M and C matrices of a one cell from ANSYS for the purpose of assembling it reaching N cells, recursively. Functions were used for demonstrating the work.

Large number of computations were saved by using recursive method. This method is an economic and accurate approach, for computing the dynamic stiffness matrix of structure consisting of large number of cells.

E. Truss

Trusses are element structures that are made up of more than one bar element. Cranes are dealt with as trusses under harmonic vibrations. The repeated cell in the truss is shown in "Fig. 3". The vertical bars tolerate larger load, for that, those bars have larger cross-sectional area than the cross-sectional area of horizontal bars and inclined ones. Each element is aligned with respect to the global coordinate system.
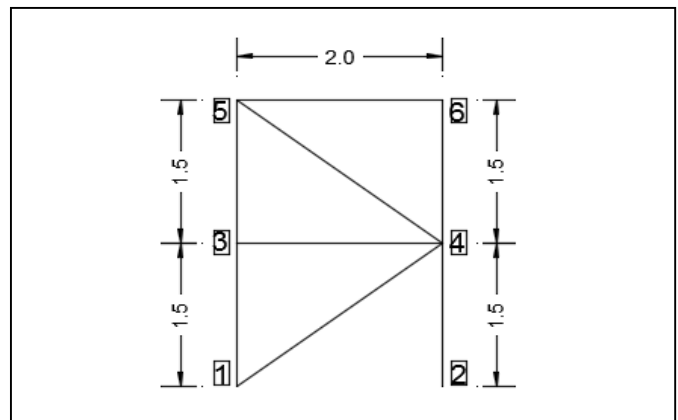
Figure 3. Repeated cell of the truss structure (dimensions are in meters)

The discrete dynamic stiffness matrix is:

$$[\bar{K}^e + i\omega\bar{C}^e - \omega^2\bar{M}^e][\bar{\Delta}^e] = [\bar{F}^e] \qquad (11)$$

The angular frequency is denoted by ω, and $i = \sqrt{-1}$.
Local stiffness, damping, and mass matrices are denoted by $\bar{K}^e$, $\bar{C}^e$ and $\bar{M}^e$ respectively. The local displacement vector is $\bar{\Delta}^e$. The local loading vector is $\bar{F}^e$.
The bar element which have constant cross-sectional area and length l, is assembled in truss, $\bar{K}^e$ and $\bar{M}^e$ matrices are:

$$\bar{K}^e = \frac{EA}{l}\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (12)$$

$$\bar{M}^e = \frac{\rho A l}{6}\begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix} \qquad (13)$$

$\bar{C}^e$ matrix is established using hysteretic damping where the dynamic stiffness matrix is:

$$D^e = (1 + 0.01 * i)K^e - \omega^2 M^e \qquad (15)$$

The dimensions of the crane are found from a real life application; a free standing tower crane with a fixed foundation. The cells are repeated for 16 times, thus the total length taking the foundation length into consideration is 53.65 m. "Table. I" shows the characteristics of the studied steel bar.

Using recursive method, it is capable to find the dynamic stiffness matrix without building a global assembly. At the node studied, the frequency response function is calculated under a range of driving frequencies.

TABLE I. CHARACTERISTCS OF A BAR

Young's modulus of elasticity	$E = 200$ GPa
Density	$\rho = 7800 \frac{\text{kg}}{\text{m}^3}$
Larger cross-sectional area	$A1 = 0.001175$ m^2
Smaller cross-sectional area	$A2 = 2.91 * 10^{-4}$ m^2

F. Truss under forced vibration at the last node

A model of free standing crane which is considered as periodic structure is shown in "Fig. 4". Last nodes are exposed to forced vibrations where $F = F_0 e^{-i\omega t}$ and $F_0 = 1$ N.

In the global system, nodes one and two are fixed, this will remove the degrees of freedom at those nodes. Using a relationship between the cell and the successive one, the matrix is built up.

The results found using RM and conventional finite element program (ANSYS) for determining the frequency response function are presented in "Fig. 5".

The time ratio between the two methods is approximately 1:18, shown in "Table. II".
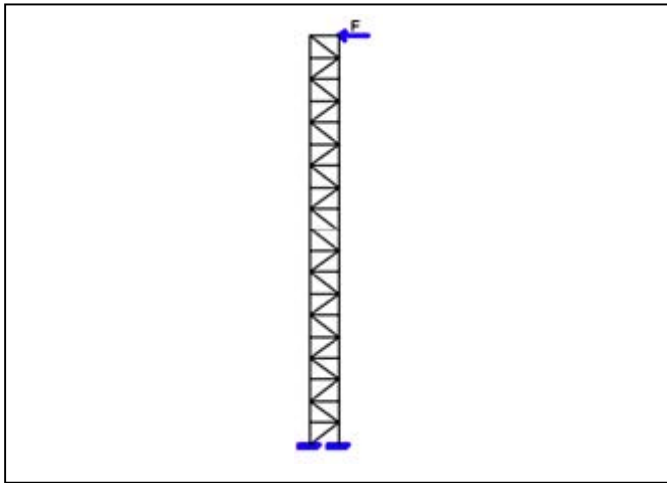
Figure 4. Overview on the crane structure
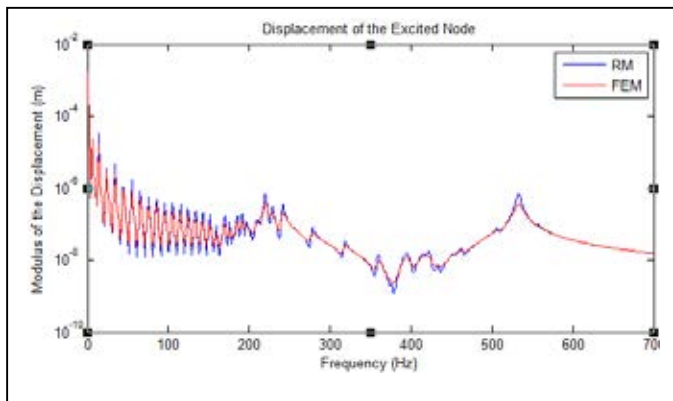
Figure 5. Displacement at exited node (logarithmic scale)

TABLE II. TIME ELAPSED AND TIME RATIO FOR BOTH METHODS

Method	Elapsed time (sec)	Time ratio
RM	0.638	$\dfrac{t_{RM}}{t_{FEM}} \cong \dfrac{1}{18}$
FEM	11.43	

G. Truss under siesmic load

For the periodic structure, the first two nodes are exposed to seismic load which is modeled as a harmonic displacement of 0.01 m on the base of the structure.

The results found using recursive method and conventional finite element program (ANSYS) for determining the frequency response function are presented in "Fig. 6".

The time ratio calculated between two methods is approximately 1:22, as in "Table. III".

Figure 6. Displacement of the last node (logarithmic scale)

TABLE III. TIME ELAPSED AND TIME RATIO FOR BOTH METHODS

Method	Elapsed time (sec)	Time ratio
RM	0.505	$\dfrac{t_{RM}}{t_{FEM}} \cong \dfrac{1}{22}$
FEM	11.08	

H. Frame under seismic load

Frames are structures that have a combination of beams which resist loads. Such structures are modelled to overcome large moments developed due to the applied loading.

Building under seismic load is modelled as a frame. The cross sectional dimensions of repeated cell used is illustrated in "Fig. 7", and the repeated cell is shown in "Fig. 8", respectively. The vertical and horizontal beams have the same cross-sectional area.

Frame analysis for local element is similar to the truss element. $\bar{K}^e$ and M^e matrices are found from ANSYS program $\bar{C}^e$ matrix is established using hysteretic damping as mentioned for the truss. The beam used is made up of steel which have modulus of elasticity and density similar to that of the bar element. The nodes at the base are under a harmonic displacement, where the displacement is 0.01 m.

The results found using recursive method and conventional finite element program (ANSYS) for determining the frequency response function are presented in "Fig. 9".

The time ratio calculated between two methods is approximately 1:31.6, as in "Table. IV".

Figure 7. Dimensions of a beam (in meters)

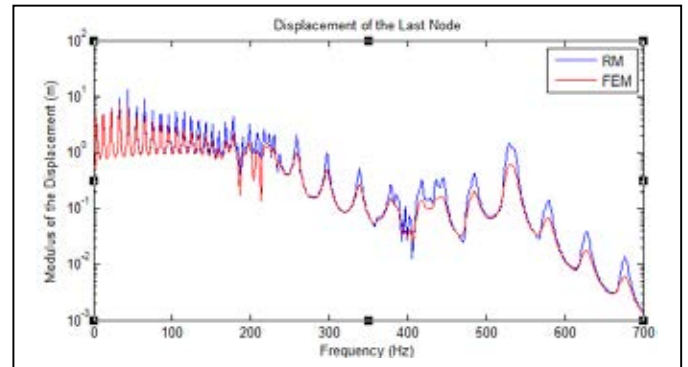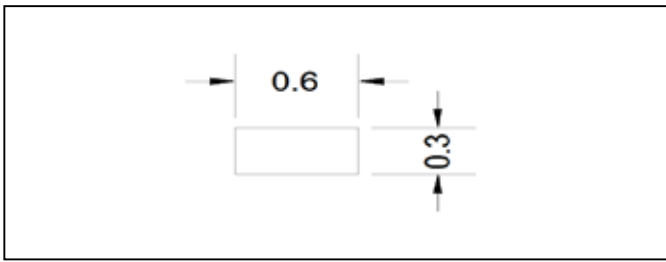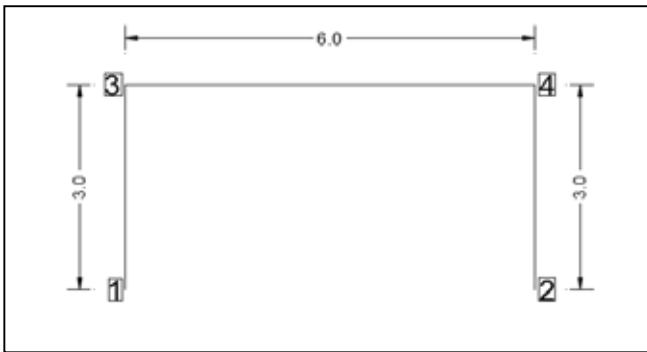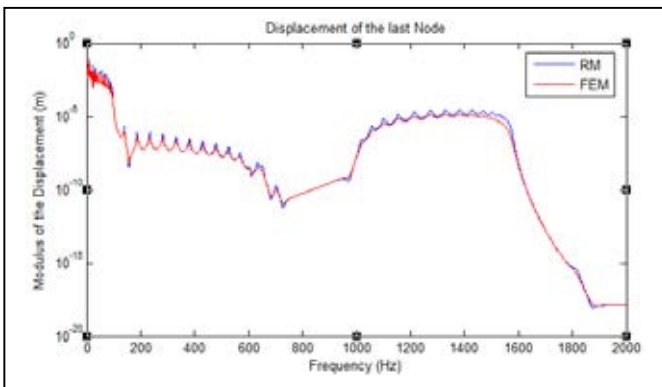Figure 8. Repeated cell for the frame structure (dimensions are in meters)

Figure 9. Displacement of the last node (logarithmic scale)

TABLE IV. TIME ELAPSED AND TIME RATIO FOR BOTH METHODS

Method	Elapsed time (sec)	Time ratio
RM	2.078	$\dfrac{t_{RM}}{t_{FEM}} \cong \dfrac{1}{31.6}$
FEM	65.8	

III. CONCLUSION

Recursive method is an approach used for calculating the frequency response function of periodic structures. For finding an accurate result, high number of cells must be considered, which will lead to large number of computations. This method provides an accurate solution and it decreases the time consumption. Also, recursive method is efficient for calculating the frequency response function for complex structures. This method showed high accuracy level at high frequencies. Its accuracy can be increased by changing the type of element from linear to quadratic, therefore vibrations at internal nodes may be computed. In general this method is applicable and effective when the structure is not under large number of forces. When the forces are exerted on every node, the method will not be feasible for calculating the frequency response function, since the advantage of removing the internal nodes recursively will be lost. The percentage of error for the truss under forced vibrations was calculated to be 11.783%, and for the truss under seismic load was 18.92%, while that for the frame was 12.544%. This difference is due to the inverse of the dynamic stiffness matrix, which will leads to a slight numerical difference between finite element method and recursive method.

IV. ACKNOWLEDGMENT

REFERENCES

[1] S. Finnveden, "Finite element techniques for the evaluation of energy flow parameters", Proceedings of the Novem, Lyon (keynote paper), 2000.

[2] S. Finnevden, "Evaluation of modal density and group velocity by a finite element method", Journal of Sound and Vibration 273 (2004) 51-75.

[3] C. M. Nilsson, "Waveguide finite elements for thin-waled structures", Licentiate Thesis, KTH, Stockholm, 2002.

[4] C. P. Yu, J. M. Roesset, "Dynamic stiffness matrices for linear members with distributed mass", Chaoyang University of Technology, Texas A&M University, Wufeng, Taichung 413, Taiwan, College Station, TX 77843, USA.

[5] F. Birgersson, "Prediction of random vibration using spectral methods", PhD Thesis, KTH, TRITA-AVE, Stockholm, 2003. p. 30.

[6] F. Birgersson, S, Finnveden, C. M. Nilsson, "A spectral super element for modelling of plate vibration", Part1: general theory. Journal of Sound and Vibration 28 (2005) 297-314.

[7] F. Birgersson, S. Finnveden, "A apectral super element for modelling of plate vibration", Part 2: turbulence excitation, Journal of Sound and Vibraton 287 (2005) 315-328.

[8] L. Gry, "Dynamic modelling of railway track based on wave propagation", Journal of Sound and Vibration 195 (1996) 477-505.

[9] L. Gavric, "Computation of propagative waves in free rail using finite element technique", Journal of Sound and Vibraton 185 (3) (1995) 531-543.

[10] I. Bartoli, A. Marzani, F. Lanza di Scalea, E. Viola, "Modelling wave propagation in damed waveguide of arbitrary cross-section", Journal of Sound and Vibraton 295 (3-5) (2006) 685-707.

[11] A. Marzani, E. Viola, I. Bartoli, F. Lanza di Scalea, P. Rizzo, "A semi-analytical finite element formulation for modelling stress wave propagation in axisymmertic damped waveguides", Journal of Sound and Vibraton 318 (2008) 488-505.

[12] S. B. Dong, R. B. Nelson, "On natural vibrations and waves in laminated orthotropic plates", Journal of Applied Mechanics 39 (3) (1972) 739-745.

[13] B. Aalami, "Waves in prismatic guides of arbitrary cross section", Journal of Applied Mechanics 40 (1973) 1067-1072.

[14] D. Duhamel, "A recursive approach for the finite element computation of waveguides", Journal of Sound and Vibration 323 (2009) 163-172.

[15] D. Duhamel, S. Erlicher, H. Hai Nguyen, "A recursive finite element method for comuting tyre vibrations", EJCM-20/2011. Dynamics of materals, structures and systems.

Damage Detection in Aerospace Structures Using Ultrasonic Waves

Ibrahim Joudi, Mahdi Bzeih, Michael Khayyat, Zachariah El-Hajj
Department of Mechanical Engineering
American University of Beirut
P.O.Box 11-0236 Riad El-Solh / Beirut 1107 2020
ikj02@mail.aub.edu, msb19@mail.aub.edu, mmk68@mail.aub.edu, zae17@mail.aub.edu

Abstract-**This paper presents the development of software elements for a structural health monitoring system, utilizing a network of piezoelectric sensors mounted on a sandwiched carbon epoxy structure. Normalization of collected data is achieved via a simple frequency response function, and damage classification via a supervised linear support vector machine reliant upon pre-identified training data. What follows is the validation and optimization of the generated model, and the authors express further venues for expanded performance.**

I. INTRODUCTION

Over the turn of the century, there has been a vast increase in the use of composite structures for high performance aerospace applications, for which conventional testing typically requires partial disassembly and lengthy inspection. Such structures are responsive to guided waves, as material imperfections and minor deformations give rise to changes in the transmitted waveform, which could enable semi-synchronous in-place examination. However, the developed responses are too complex for human assessment, making software essential for any in-depth study. Full automation could result in novel applications, as it could function as the basis of an in-flight damage response system active during nominal operation.

The main goal of this project is to build a structural health monitoring (SHM) system for said applications, capable of detecting whether damage is present in the structure, and locate the damage if present. Emphasis is placed mostly on developing the necessary software, as the component hardware is already well developed.

In this paper, we detail the experimental testbed and software in section 2, the results of said setup in section 3, and the conclusion in section 4. Section 5 features venues for further development.

II. EXPERIMENTAL SETUP

A. Equipment

The specimen to be tested is a sandwiched carbon epoxy plate, with waves transmitted and collected by a square array of 12 piezoelectric transducers (PZT's). As the plate is nominally healthy, damage effects are simulated by placement of small masses at desired locations, with increasing weight simulating greater damage.

The SHM system hence described utilizes the following cutting-edge instruments: an NI PXIe-2593 Multiplexer, NI PXIe-5122 Digitizer, and an NI PXIe-5422 Arbitrary Waveform Generator. The waveform generator is used to provide input waves, which are all sinusoids of varying frequency. All PZT's are connected to the multiplexer, as the multiplexer determines which acts as input (the others all being designated as output) at the beginning of each test cycle. The digitizer then transforms this collected data into a computer-readable for m.

B. Software

Our software consists of three elements operating in sequence – a LabVIEW data collection code, a MATLAB frequency response function, and a two-stage MATLAB support vector machine (SVM).

The LabVIEW code governs the operation of the above equipment, and can if necessary supplant the waveform generator by providing signals of its own. Its chief function is controlling the multiplexer and organizing the collected data into files readable by the next two elements.

The frequency response function is used for normalizing the data, rendering them into comparable sets. A Fast Fourier transform is used to move from the time domain to the frequency domain. From this, a vector is created detailing amplitude of transfer function (ratio of the output to the input) at a distinct set of frequencies. Which frequencies and how many are utilized can be varied for optimized study.

The SVM is responsible for data classification, through which damage is classified. In this particular example, the SVM is a supervised single class device, which is to say firstly that it must be 'trained' beforehand with pre-classified data in order to operate, and secondly that it can only distinguish two distinct damage states from one another. It can, for instance, distinguish healthy sites from damaged ones, or mildly damaged from highly damaged sites, but cannot do both within the same classifier. Once the two damage states have been identified and given appropriate labels (1 or -1), the classifier produces a decision value for each data set; if positive, it is considered part of the type 1 set, and if negative, it is considered part of the other set. This is done through the following:

Assume we have the following graph with x's and o's. We want to draw a straight line that separates the 2 examples such that it can have the largest width from the 2 examples:
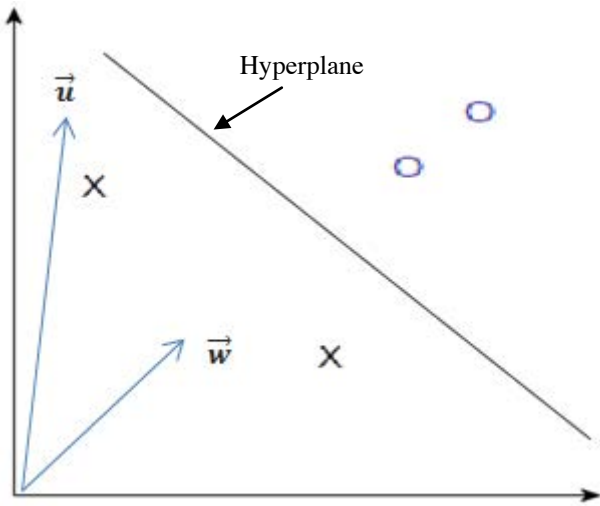
Figure 1

We are going to assume a vector $\vec{w}$ and that vector (of any length)is constrained to be perpendicular to the hyperplane. In addition assume that we have an unknown that we would like to know on which side of the gap it falls.

And we do that by projecting the vector $\vec{u}$ on $\vec{w}$. Therefore we get the distance of that unknown in the perpendicular direction giving us where this unknown resides. Representing this in an equation:

$$\vec{w} \cdot \vec{u} \geq b$$

Where b represents some constant, thus will account for any change in the sign.

$$\vec{w} \cdot \vec{u} + b \geq 0 \qquad (1)$$

Let's call this the decision rule and this would represent the unknown falling on the o's side. The problem lies in what the value of b is and which $\vec{w}$ to use since there are many vectors perpendicular to the hyperplane with several lengths for a particular w. Therefore we need to provide some constraints to know what the following b and w's are.

$$\vec{w} \cdot \vec{o} + b \geq 1$$

$$\vec{w} \cdot \vec{x} + b \leq -1$$

We introduce an additional variable y_i, to simplify the math, such that:
$y_i = +1$ if the sample is o
$y_i = -1$ if the sample is x

Therefore:

$$y_i(\vec{w} \cdot \vec{o_i}) + b \geq 1$$

$$y_i(\vec{w} \cdot \vec{x_i}) + b \geq 1$$

We can see that both equations are the same so let's represent any sample, o or x, by q.

$$y_i(\vec{w} \cdot \vec{q_i}) + b \geq 1$$

We'll add one more constraint:

$$y_i(\vec{w} \cdot \vec{q_i}) + b - 1 = 0 \qquad (2)$$

And this equation would represent the parts of the sample that lie on the maximum distance from the hyperplane.
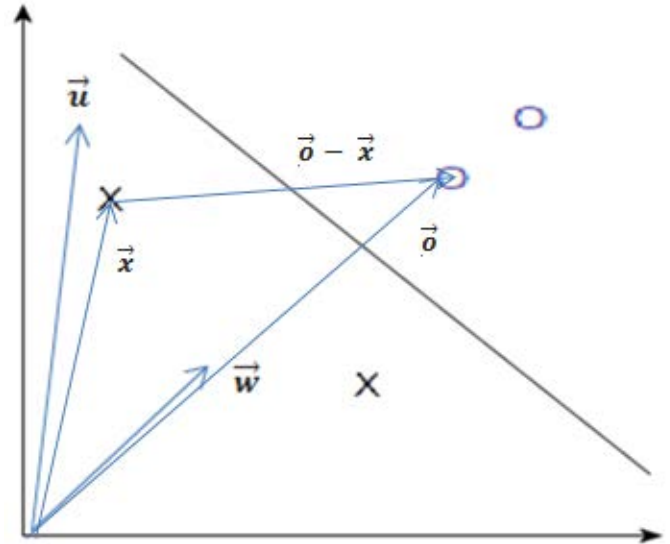
Figure 2

To find the width of the gap we calculate:

$$\text{width} = (\vec{o} - \vec{x}) \cdot (\frac{\vec{w}}{||\vec{w}||}) = \frac{2}{||\vec{w}||}.$$

So, to maximize the width we need to minimize $||\vec{w}||$ which in turn means we can minimize $\frac{1}{2} || \vec{w}||^2$.

Since this mathematical problem is a constrained problem, thus an efficient and most commonly used method to solve it is to construct the Lagrange function which will transform the constrained problem into an unconstrained problem.

Lagrange function:

$$L(w,\lambda) = \frac{1}{2} || \vec{w}||^2 - \sum_i \lambda_i[y_i(\vec{w} \cdot \vec{q_i} + b) - 1] \, (3)$$

After manipulating the equation based on the previously mentioned equations **(1)** and **(2)**, **(3)** becomes:

$$L(q,\lambda) = \sum_i \lambda_i - \frac{1}{2}\sum_i \sum_j \lambda_i \lambda_j y_i y_j \vec{q}_i \cdot \vec{q}_j$$

Which also tells us that this decision rule depends on the dot products of $\vec{q}$ and $\vec{u}$.

III. RESULTS

A. Equipment

Four damage cases were to be studied - healthy, extremely mild (individual micrometer-depth cracks), intermediate (fractions of a millimeter crack) and heavy (large cracks sufficient to cause shearing or separation of layers) – hereafter

identified as 1, 2, 3 and 5. For initial tuning, comparison was made between healthy and heavy damage cases, as it was presumed these would be most readily distinguishable. 162 samples are chosen, with one-third of them randomly designated as training data (pre-labeled) and the remainder as testing (to be classified).
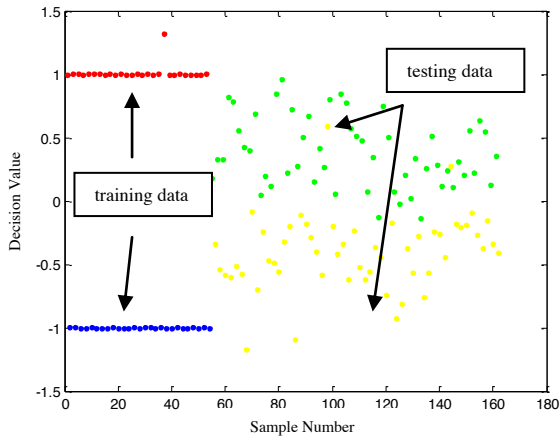
(a)
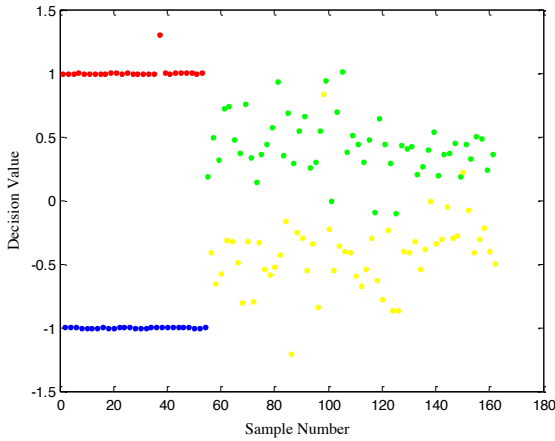
(b)
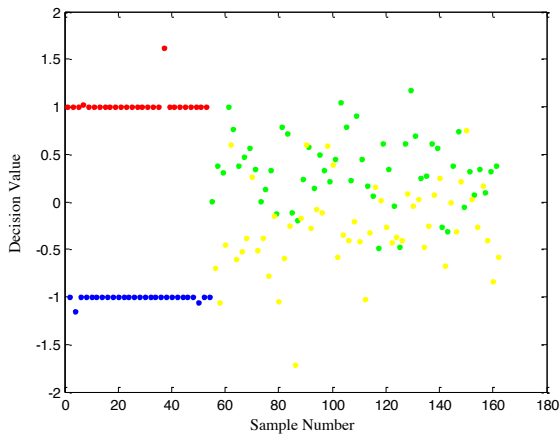
(c)

Figure 3. Plot of SVM classification results based on number of features: (a) 2500, (b) 1250, (c) 500. Red/blue indicate healthy/unhealthy training data, green/yellow indicated healthy/unhealthy testing data.

The training data is sorted perfectly, as is to be expected, seeing as it is the basis of the model. However, classifier accuracy with regards to the testing data varies depending on the number of features, i.e. number of data points, which in our case would be 15000 due to the high sampling frequency of the waveform generator and the digitizer being used. In the first case, all features are used – all 2500 – followed by 1250 and 500, with corresponding accuracies of 95.37%, 95.37% and 77.78%.

As this yields the largest accuracy, what remains is repeated with 1250 features. This analysis is repeated for varying combinations of damage cases.

TABLE I
SVM ACCURACY FOR DAMAGE CASE COMPARISON

Case	1 vs. 2	1 vs. 3	1 vs. 5	2 vs. 3	2 vs. 5	3 vs. 5
Accuracy	50.93%	87.04%	95.37%	87.96%	95.37%	99.07%

IV. CONCLUSION

As can be noticed from the plots in Fig.3, dispersion of decision values increases as the number of features is reduced. This suggests that the majority of features are not changing meaningfully, creating similarities between sampled data points that drive down the decision value. These features are not necessary for classification and can be done without. However, removing too many features results in a loss of those which are changing, and though decision values increase in absolute magnitude as points are more easily distinguished, accuracy decreases as the characteristic pattern becomes missing.

From Table 1, it is most notable that the classifier is all but incapable of distinguishing case 1 from case 2, and that comparison of case 1 or case 2 vs. a particular damage case yields similar results. As damage case 2 describes microscopic damage which does not notably weaken the structure, this apparent similarity is not a critical issue, but as greater damage tends to originate from such sites the ability to distinguish them would nevertheless prove useful. It is noteworthy that higher damage states are more readily distinguished from one another than from the healthy plate, which suggests different modes of damage. This agrees with our understanding of the physical phenomena behind them.

V. FURTHER DEVELOPMENT

While all components are satisfactory as they are, further testing could undoubtedly result in greater optimization. The most effort can be put in the classifier, and an obvious roadmap is already in place. First of these steps is to develop a multi-class version of the developed classifier. This may be a one-vs.-all scheme, which is essentially a sequence of single class classifiers each separating one category from the others, or an all-vs.-all scheme, in which all data points are plotted and clustered and hence separated simultaneously. The large accuracies demonstrated in Table 1 are testament to its viability. Second is to modify the single-class SVM to be

unsupervised, so that it does not require pre-training. Third would be to combine the above two cases to produce a qualified unsupervised multi-class SVM. Finally, it may pay to see if accuracy varies with shuffling of the input data, particularly in the case of the unsupervised classifiers.

A fourth stage is still necessary in order to localize the damage. This could be achieved by plotting the transducers and producing equations for the lines of response and noting which of them intersect, then seeing where two intersecting lines both show damage. In our case, the fixed geometry of the piece makes this a simple task, but a generalized form could be made in which the shape of the PZT array is an input, from which the lines of response are determined by the software rather than given. The possibility for improved precision by comparing back-and-forth signals along the same line of response (between cases where a is input and b is output, and where b is input and a is input) should also be put under consideration.

Unsupervised classification algorithm is in the final stages where the famous k-means algorithm is being implemented. K-means algorithm is a partitional clustering algorithm where it partitions the given data into k-clusters and the data will be categorized based on the location of the centroid of the data present in each cluster.

Currently, one of our objectives is trying to integrate LibSVM within our LabVIEW program to automate the process with minimum number of stages and softwares being used.

It is noteworthy that our whole setup including the LabVIEW program and the MATLAB code, could be implemented on a higher scale system, i.e. using n number of sensors in real life applications such as damage detection in aircraft wings and in fuel piping systems.

REFERENCES

[1] Agarwal, Sushant and Mira Mitra. "Lamb Wave Based Automatic Damage Detection Using Matching Pursuit and Machine Learning." Smart Materials and Structures (2014): 085012.

[2] D. C. Barker, "Circuit waveforms and frequency response: Nyquist plot,"MINNIE and HSpice for Analogue Circuit Simulation, vol. 1, no. 2, pp. 103–138, 2011.

[3] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene Selection for Cancer Classification using [7] Liu, Bing, et al. "2-D Defect Profile Reconstruction from Ultrasonic Guided Wave Signals Based on QGA-Kernelized ELM." Neurocomputing (2014): 217-223.

[4] J. Greben, "On the physical relevance of the discrete Fourier transform,"Applied Mathematical Modelling, vol. 15, no. 11, pp. 657–660, 1991.

[5] C. Hwang, M.-J. Lu, and L. S. Shieh, "Improved FFT-based numerical inversion of Laplace transforms via fast Hartley transform algorithm,"Computers & Mathematics with Applications, vol. 22, no. 1, pp. 13–24, 1991.

[6] J.-T. Kim, Y.-S. Ryu, H.-M. Cho, and N. Stubbs, "Damage identification in beam-type structures: frequency-based method vs mode-shape-based method," Engineering Structures, pp. 57–67, 2003.

[7] N. Maia, J. Silva, E. Almas, and R. Sampaio, "Damage Detection In Structures: From Mode Shape To Frequency Response Function Methods," Mechanical Systems and Signal Processing, vol. 17, no. 3, pp. 489–498, 2003.

[8] Oukhellou, L, et al. "Combined Use of Sensor Data and Structural Knowledge Processed by Bayesian Network: Application to a Railway Diagnosis Aid Scheme." Transportation Research Part C (2008): 755-767.

[9] R. Sampaio, N. Maia, and J. Silva, "Damage Detection Using The Frequency-Response-Function Curvature Method," Journal of Sound and Vibration, vol. 226, no. 5, pp. 1029–1042, 1999.

[10] S. Sharma, "Application of Support Vector Machines for Damage Detection in Structures," 2008.

[11] I. Steinwart and A. Christmann, Support vector machines. New York: Springer, 2008.

[12] Z. Su and L. Ye, Identification of damage using Lamb waves. Berlin: Springer, 2009.

[13] Vishnuvardhan, J, C V Krishnamurthy and Krishnan Balasubramaniam. "Genetic Algorithm Based Reconstruction of the Elastic Moduli of Orthotropic Plates Using an Ultrasonic Guided Wave Single-Transmitter-Multiple-Receiver SHM Array." Smart Materials and Structures (2007): 1639-1650.

[14] K. Worden and A. J. Lane, "Damage identification using support vector machines," 2000.

Design of a Computational Model for Magnetic Targeted Drug Delivery in Human Diseased Blood Vessels

Adnan Arnaout, Fouad Barakat, Ribal Hachem, Tarek Jaber
Mechanical Engineering Department
American University of Beirut
Beirut, Lebanon
aja28@mail.aub.edu, fab18@mail.aub.edu, ryh06@mail.aub.edu, ttj00@mail.aub.edu

Abstract- **Cardiovascular diseases are among the leading causes of human death worldwide. These include highly lethal conditions such as coronary artery and peripheral vascular diseases, as well as cases such as bulges and stenosis inducing different complications in the organism. Such cases in the arterial walls affect blood flow, which may lead to the damage of various organs and regions in the organism. Cardiovascular diseases were traditionally treated using drug delivery methods that involved pills and injections, chemical activation, controlled release and drug entrapment using proteins and polymers. However, these methods have several major side effects on the patient, such as overdose. These negative effects can be reduced by treating the disease using targeted drug delivery. These methods include magnetic targeted drug delivery, microrobot injection and gene therapy. This paper focuses on the analytical and numerical investigation of the magnetic targeted drug delivery in a human diseased arterial system. The principle of this technique is based on allowing flow retention and enabling the drug to be deposited in the desired tumor site for a longer duration. In this study, a 1D analytical model is developed to describe the hemodynamics in an artery under the influence of an external magnetic field. The model is used to simulate the effect of several parameters on flow distribution. These parameters include the volume fraction of the magnetic nanoparticles, the artery diameter and the magnetic field intensity. Furthermore, a 3D computational model will be developed to address the hemodynamics for a patient-specific arterial model obtained from CT scan data.**

I. Introduction

Around 17 million deaths per year are approximated by the World Health Organization (WHO) with cardiovascular related causes. In addition, the projection to 2030 is around 24 million deaths due to cardiovascular diseases, maintaining its position as the most fatal disease around the globe [4].

Among these cardiovascular diseases are stenosis and aneurysm. Stenosis induces narrowing of the blood vessels, causing the obstruction of blood flow. For instance, stenosis in the arteries can block the oxygen transport to different organs and tissues, which can cause serious damage within the human body. Another cardiovascular disease is aneurysm, which is an arterial condition causing the weakening of the arterial wall, thus allowing the artery to bulge which may cause the blood vessels to rupture. Thus, it is important to develop an appropriate and suitable mathematical model and numerical simulations to better understand and develop the appropriate knowledge to treat these cardiovascular diseases.

Current drug delivery systems, such as oral and intravenous drug delivery methods present numerous limits whether in methods of delivery or on desired impact. Drug delivery to diseased sites in the organism may fail or can be largely attenuated because of the delivery method. Also, current drug delivery techniques may present negative side effects on individuals and cause other sources of complications in the body. That's why it is crucial to seek novel and non-invasive drug delivery systems, such as magnetic targeted drug delivery. Magnetic targeted drug delivery is done by injecting drug carriers such as magnetic nanoparticles into the artery in which the blood flows to the targeted infected cells. An external magnet is then used in order to magnetize the magnetic particles. Using magnetic targeted drug delivery methods would decrease the body's exposure to toxic and harmful drugs, such as the side effects of chemotherapy, by reducing the dosage of drugs taken into the blood. Several studies were conducted to investigate the flow of a ferrofluid under an externally applied magnetic field. Andersson et al. analyzed the flow of a ferrofluid driven by a stretchable sheet, under the influence of an external magnetic field [1]. In their study, the fluid is considered to be viscous and non-conductive, while the flow is assumed to be in 2D. It is found that the magnetic field applied decelerates the flow of the ferrofluid. On the other hand, Wang et al. modelled a 3D flow of a ferrofluid inside a blood vessel for magnetic drug delivery, where the velocity is seen to decrease with increasing magnetic field intensity, and the pressure drop is seen to decrease near the vessel wall [2]. Another study was conducted by Weng et al. where the effect of natural blood particles and artificial nanoparticles on the flow of a ferrofluid under an applied magnetic field were examined. It is found that the increase in the volume fraction of artificial particles in the presence of a magnetic field in the direction of fluid flow leads to an increase in the magnetic body force, hence an increase in the flow axial velocity. As for the natural particles, an increase in their volume fraction leads to an increase in the fluid effective viscosity hence a decrease in the flow axial velocity. The results also reflect the effect of increasing the magnitude of magnetic field along the flow where velocity is seen to increase as the magnitude increases [3].

In this study, a 1D analytical model as well as a 3D computational model are built to analyse the hydrodynamics of ferrofluids in human diseased blood vessels under the influence

of an external magnetic field generated by a magnetic dipole. A CT scan of a human artery will be used to simulate the problem numerically in order to replicate the natural conditions more accurately. For the same reason, the pressure at the outlet will be taken as variable, which was assumed to be constant in other studies. This study will contribute to the enhancement of future magnetic targeted drug delivery systems which carry an immense potential in disease treatment and clinical applications.

II. PROBLEM APPROACH

The flow problem in 1D and 3D analysis will be approached using the following fundamental equations:
Continuity equation in cylindrical coordinates [3]:

$$\frac{1}{r}\frac{\partial(ru_r)}{\partial r} + \frac{1}{r}\frac{\partial u_\emptyset}{\partial \emptyset} + \frac{\partial u_z}{\partial z} = 0 \ (1)$$

Momentum equation (Navier Stokes):
$$\frac{dp}{dz} - \mu_0\sigma_m\varphi m_l\frac{dh}{dz} = \mu_{eff}\left(\frac{1}{r}\frac{d}{dr}\left(r\frac{du_z}{dr}\right)\right) (2)$$

Where m is the magnetization, "h" the internal magnetic field, μ_{eff}: the effective viscosity that considers the effects due to the nanostructures and $\mu_0\sigma_m\varphi m_l\frac{dh}{dz}$ is the magnetic force per unit volume.

The flow is considered incompressible, fully developed and steady state.

III. METHODS

A. 1D analytical method:
   Boundary conditions:
Static pressure inlet=89.44486mmHg
Static pressure outlet=89.06982mmHg
Velocity profile:
$$u_z(r) = \frac{1}{4.\mu_{eff}} \times \left(\frac{P_2-P_1}{L}\right) \times (1-AB) \times \left(r^2 - \frac{D^2}{4}\right) (3)$$
Where $A = \mu_0 \times \sigma_m \times \varphi \times m_l$, $B = \frac{-\Delta H}{\mu_0 \times \sigma_m \times m_l}$, φ is the volume fraction of artificial particles, μ_0 is the magnetic permeability of vacuum, σ_m is the ratio of the volume occupied by the magnetic phase to the total volume of artificial particles and m_l is the saturation magnetization of magnetic solid material [3].

ΔH is a unitless parameter describing the magnetic field gradient.

$$\Delta H = -\mu_0 \times \sigma_m \times m_l\frac{\frac{dh}{dz}}{\frac{dp}{dz}} (4)$$

Pressure profile: $P(z) = \frac{P_2-P_1}{L} \times z + P_1 (5)$
Volume flow rate:
$$Q = \frac{1}{4.\mu_{eff}} \times \left(\frac{P_2-P_1}{L}\right) \times (1-AB) \times 2 \times \pi \times \frac{-D^4}{64} (6)$$

B. 1D numerical method:
   Case 1 boundary conditions:
Total pressure inlet=90mmHg
Vascular resistance outlet range= $\frac{\Delta P}{Q} = 29873$Pa.s/m^3 to 16627Pa.s/m^3 where ΔP is the pressure drop across the arterial section and Q is the theoretical flow rate in the artery. The vascular resistance accounts for the vessel branch system at the outlet of the artery.
   Case 2 boundary conditions:
Static pressure inlet=90mmHg
Vascular resistance outlet range=29873Pa.s/m^3 to 16627Pa.s/m^3

C. 3D numerical method:
   Ferrohydrodynamics (FHD) relies on the assumption that the magnetic force plays the largest role, and that the biomagnetic fluid (here, the blood) is not electrically conductive. For a magnetic source with high magnetic field gradient, this approach will be followed. For this, the following formulation for "h" will be used:
$h = -\nabla\emptyset$ (7)
"h" is an irrotational vector that represents the gradient of the scalar field. [1]
$\nabla.\mu_0 h = 0 \rightarrow \nabla^2\emptyset = 0$ (8)
This formulation follows Laplace's equation.
After solving the above Laplace equation with the corresponding boundary conditions we get the following solution:
The following formulas will be used for the parameters in the simulation while assuming that the flow is axisymetric:
Based on [1], the following equations we obtained:
Magnetic scalar potential:

(9)

$$\emptyset = \frac{\alpha}{2\pi}\left(\frac{x}{x^2+\left(y+d\right)^2}\right) \quad \text{Magnetic field "h":}$$

$$h_y = -\frac{\partial\emptyset}{\partial y} = \frac{\gamma}{2\pi}\left(\frac{2x\left(y+d\right)}{\left(x^2+\left(y+d\right)^2\right)^2}\right) \quad (10)$$

$$h_x = -\frac{\partial\emptyset}{\partial x} = \frac{\gamma}{2.\pi}\left(\frac{x^2-\left(y+d\right)^2}{\left(x^2+\left(y+d\right)^2\right)^2}\right) \quad (11)$$

IV. RESULTS

IV.I. 1D analytical results

Figure 1. Volume flow rate versus ΔH while varying phi for P_{total} at inlet and resistance at outlet
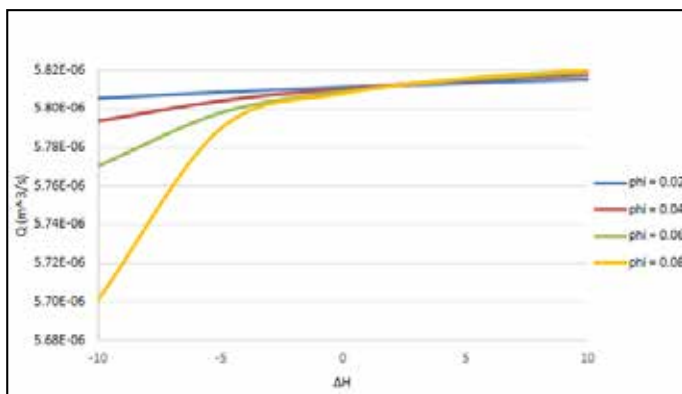
Figure 2. Volume flow rate versus ΔH while varying phi for P_{static} at inlet and resistance at outlet

For a negative ΔH, the flow rate decreases dramatically as the volume fraction of the magnetic nanoparticles (phi) increases. For positive ΔH, the increase of phi has no significant effect on the flow rate.
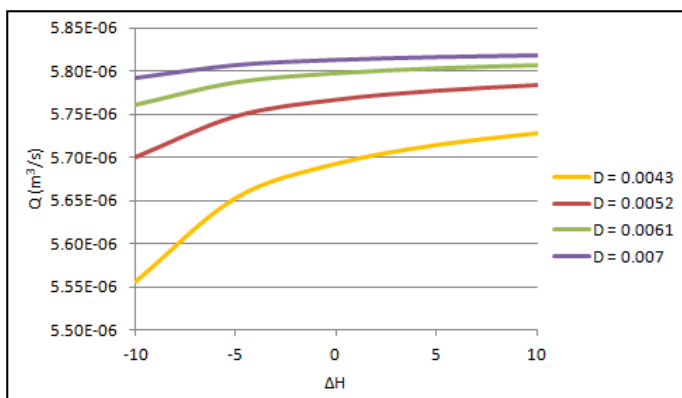
Figure 3. Volume flow rate versus ΔH while varying vessel diameter for P_{total} at inlet and resistance at outlet
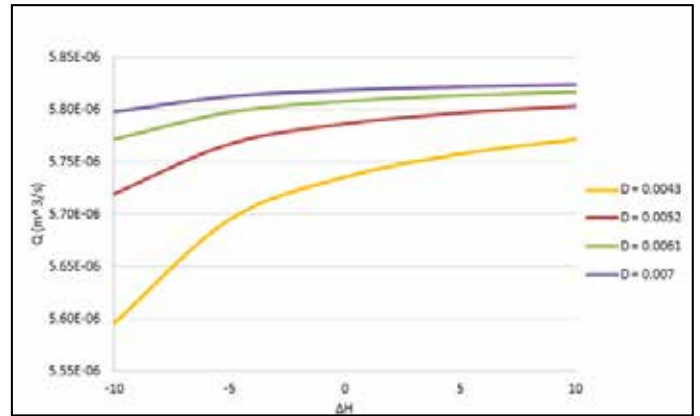
Figure 4. Volume flow rate versus ΔH while varying vessel diameter for P_{static} at inlet and resistance at outlet

As the artery diameter increases, the flow rate shifts upward.
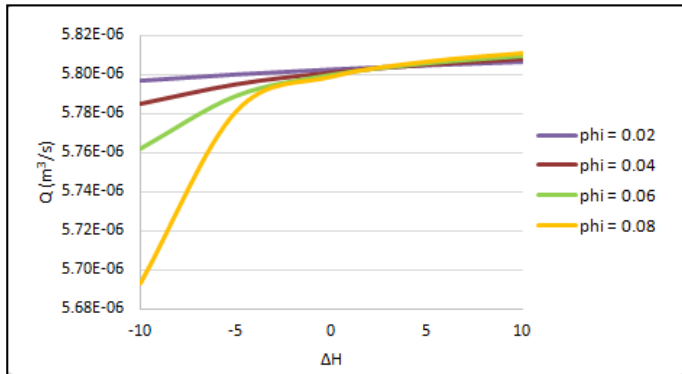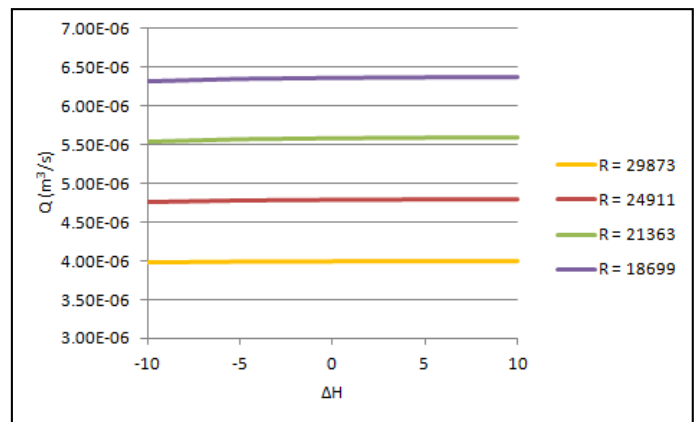

Figure 5. Volume flow rate versus ΔH while varying outlet resistance for P_{total} at inlet and resistance at outlet
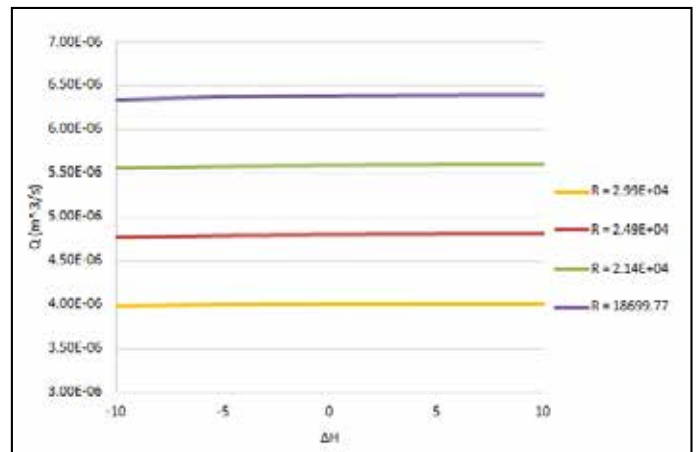
Figure 6. Volume flow rate versus ΔH while varying outlet resistance for P_{static} at inlet and resistance at outlet

As the outlet resistance increases, the flow shifts downward.

Figure 7. Volume flow rate versus ΔH while varying phi for P_{static} at inlet and outlet

For negative ΔH, the flow rate decreases as phi increases, while for positive ΔH, the flow rate increases as phi increases.

Figure 8. Volume flow rate versus ΔH while varying vessel diameter for P_{static} at inlet and outlet

As the artery diameter increases, the flow rate shifts upward.

Figure 9. Magnetic force versus ΔH while varying phi

For negative ΔH, the magnetic force decreases dramatically as phi increases, while for positive ΔH, phi has no significant effect on the magnetic force.

V. DISCUSSION

As ΔH increases in the negative direction, the magnetic force opposing the ferrofluid flow increases, causing its retention, hence decreasing it. This method can be applied to target a diseased site inside a blood vessel and accurately release the nanoparticles encapsulating the drug to treat the condition.

On the other hand, the increase of volume fraction of magnetic nanoparticles while applying the magnetic field in the positive direction has little effect on the magnetic force, hence a non-significant effect on the flow rate.

Finally, assuming a constant static pressure at inlet and outlet, as the work done in previous articles, results in a linear behavior of the flow rate when varying ΔH. However, when the resistance at the outlet of the artery, due to the distribution of blood in different branches after exiting the arterial section, was accounted for, the flow rate behaves non-linearly when ΔH is varied, which better mimics cases in the natural environment.

REFERENCES

[1] H. Andersson and O. Valnes, "Flow of a heated ferrofluid over a stretching sheet in the presence of a magnetic dipole," *Acta Mechanica*, vol. 128, pp. 39-47, 1998.

[2] S. Wang, H. Liu, and W. Xu, "Hydrodynamic modelling and CFD simulation of ferrofluids flow in magnetic targeting drug delivery," *International Journal of Computational Fluid Dynamics*, vol. 22, pp. 659-667, 2008.

[3] H. C. Weng, "Hydrodynamic Modeling of Targeted Magnetic-Particle Delivery in a Blood Vessel," *Journal of biomechanical engineering,* vol. 135, p. 034504, 2013.

[4] World Health Organisation, Cardiovascular diseases (CVDs), 2013. http://www.who.int/mediacentre/factsheets/fs317/en/index.html. [Accessed: Oct. 15, 2013].

APPENDIX

Parameters	Value
L (m)	0.05
P_{in} (Pa)	$15*10^3$
P_{out} (Pa)	$13.3*10^3$
μ_0 (H.m^{-1})	$4*pi*10^{-7}$
σ_m	0.64
m_l	$4.78*10^5$
D (m)	0.006
μ_f (mPa.s)	1.19
ε	1.05
ε_m	1.08

$$\mu = \mu_f \times \left(1 + \frac{5}{2}\varepsilon(phi + phic)\right)$$

$$k = \frac{3}{2} \times \mu \times phi$$
$$\mu_{MV} = \varepsilon_m \times \sin(phi)^2$$
$$\mu_{eff} = \mu + k \times \mu_{MV}$$

Where μ is the shear viscosity, μ_f is the shear viscosity of carrier liquid, ε is the nonmagnetoviscosity correction factor, phi is the volume fraction of natural particles, phi$_c$ is the volume fraction of natural particles, k is the vortex viscosity, μ_{MV} is the magnetoviscosity, ε_m is the magnetoviscosity correction factor, and μ_{eff} is the effective viscosity [3].
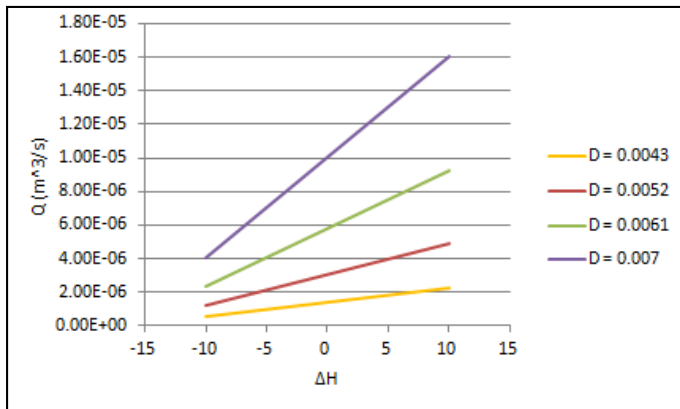
Effectiveness of a Dual Dynamic Vibration Absorber in Attenuating Pathological Tremor

Sarah Gebai

Graduate student, Mechanical Department, School of Engineering

Lebanese International University

Beirut, Lebanon

31130357@students.liu.edu.lb

Abstract-**Neurologically disordered people, suffering from resting tremor, are known to have Parkinson Disease [1]. Medications used to decrease involuntary antagonistic muscles contraction can threaten the life of a Parkinson's patient [2]. However, mechanical vibration absorbers can be used as an alternative treatment [3]. In this paper, the human hand is modeled dynamically as three-degree-of-freedom dissipative system to reflect the biodynamic response of Parkinson's pathological tremor at musculoskeletal level. Shoulder, elbow, biceps brachii and wrist muscles are modeled to produce movements at the proximal joints. One-degree-of-freedom frictionless joints are modeled to permit the flexion-extension angular displacements in horizontal plane. Involuntary tremor is modeled as a dual harmonic excitation due to shoulder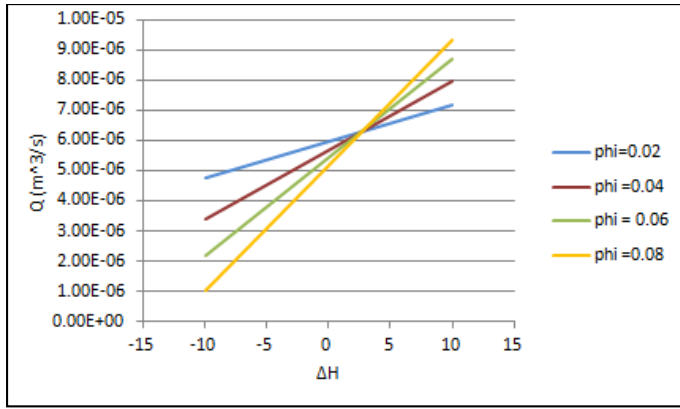 and elbow muscles activation at the resonance frequencies. The performance of the dynamic vibration absorber (DVA) is studied when attached to the forearm of the hand. The response of the system with a single DVA is compared to that with dual DVA which is tuned at both excitation frequencies. The equations of motion are derived for the non-Lagrangian modeled systems. The response is determined using the complex transfer function of the dynamically coupled system. The absorbers are designed to have long period of operation. Each absorber is sized and appropriate materials are selected. The choice of the best efficient absorber is analyzed in terms of its capability in reducing the amplitude of resting tremor. As future work, an experimental rig will be fabricated to hold the actual hand and permit the flexion angular displacement at the proximal joints. Also, an experiment can be conducted on a real case to validate the semi-analytical approach and to probe any potential "real-life" side effects of the system.**

I. INTRODUCTION

Tremor is considered as the most common faced abnormal involuntary movement disorder [4] and the source of functional disability. Movement disorders cause patients with pathological tremors to have significant uncontrollable hand tremor movement [5]. PD is a multi-system neurodegenerative disorder caused by the lack in the level of dopamine (>80% of dopamine in the brain). PD tremor is related to resting and postural tremors which are in range of 3-7 Hz and 5-12 Hz [6], respectively. Essential tremor is a bilateral of kinetic and posture tremor in range of 4-12 Hz [7]. The negative correlation between amplitude and frequency justifies the significant movement in hand of PD patient when the person is at the rest condition.

Tremor in hand of PD patient makes them suffer while performing their daily tasks and feel embarrassed to face other people. Motor symptoms can be improved by dopaminergic replacement therapy, but PD tremor may not respond to dopamine and may be dopa resistant.

To reduce tremor caused by problems with supplying certain commands to the muscles, vibration absorber can be placed on the muscles to counter act the vibration. The DVA is a passive vibration controller added as a secondary system to reduce the steady state vibrational motion of the structure at a particular undesired frequency.

Hashemi *et al*. [8] proposed a two degree of freedom (DOF) biodynamic model of the upper limb. Hand was modeled in horizontal plane as two rigid segments to describe the flexion-extension planar motion at elbow and shoulder joints with sinusoidal excitation. He designed a 1 DOF passive TVA modeled as pendulum that was able to suppress rest tremor at elbow and shoulder of PD patient. Rahnavard *et al*. [9] used a similar hand model but with a randomly excited system. He designed an optimal 1 DOF pendulum absorber using the H_2 optimization method which caused a very considerable reduction in tremor at shoulder and elbow joints. Taheri *et al*. [10] designed a pneumatic cylinder actuator with high power density. It was able to control pressures in cylinder chambers to attenuated involuntary motion of a 4 DOF motion of human hand (wrist joint flexion-extension and abduction-adduction).

Human hand is excited by muscular activity over a range of driving frequencies and this may confirm the need of multi-vibration absorber. Igusa [11] studied the multiple mass dampers tuned within a frequency range. He found that it is more effective than a tuned mass damper with same total mass. Brennan [12] demonstrated the use of a parallel multiple TVA tuned at slightly different frequencies which results in an improved broadband device.

In this paper, the flexion-extension motion of the palm will be considered within the human hand modeling. Single and dual DVA will be designed to suppress the pathological tremor of human hand. A ready to be used absorber will be sized and suitable materials will be selected respecting absorber constraints and limitations.

II. SYSTEM'S DESIGN

A. Hand Design

Human hand is modeled in the horizontal plane as a 3 DOF rigid segments to describe the biodynamic behavior of the upper arm, forearm and the palm. The dimensions of the modeled hand segments and the used 1 DOF frictionless joints

are shown in "Fig. 1". The geometric parameters are calculated based on the experimentally determined values of the right hand segments as demonstrated by Drillis *et al*. [13]:

$$l_1 = 36.4\,cm, l_2 = 29.9\,cm, l_4 = 20.3\,cm$$
$$a_1 = 0.427\,l_1\,m, a_2 = 0.417\,l_2\,m, a_4 = 0.361 l_3\,m \qquad (1)$$
$$D_1 = 1088.0\,kg/m^3, D_2 = 1108.6\,kg/m^3, D_4 = 1112.6\,kg/m^3$$

Calculated parameters are shown in "Fig. 1", in addition to:

$$m_1 = 2.070\,kg, m_2 = 1.160\,kg, m_3 = 0.540\,kg$$
$$I_1 = 0.0228\,kg.m^2, I_2 = 0.0082\,kg.m^2, I_4 = 0.0012 kg.m^2 \qquad (2)$$
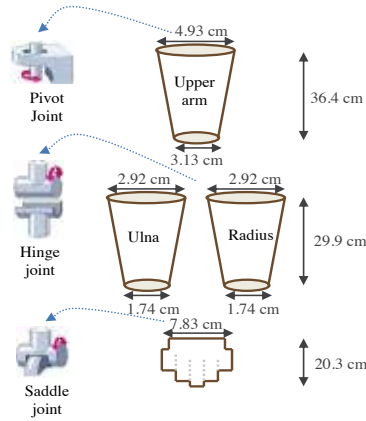
Figure 1. Modeled geometry of the upper arm, forearm and the palm and their connecting joints

The hand is modeled in the horizontal plane to describe the flexion motion at each joint. The mass is concentrated at the centroid of each segment. The kinematic and dynamics are determined using the effective hand modeled linkages when the absorber is attached to the forearm "Fig. 2".
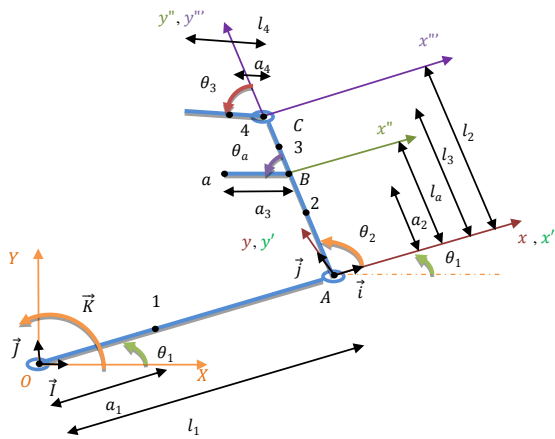
Figure 2. The controlled hand system

Four muscles are modeled as spring-damper systems to produce movements and are assumed to be regulated independently. Muscles considered are: The shoulder, elbow and wrist single joint muscles and the biceps brachii double joint muscle which pairs at shoulder and elbow. Their stiffness

and damping coefficients are assumed to be proportional by a constant Table I.

TABLE I
DESIGNED PARAMETERS OF THE MUSCLES

Muscle	Shoulder	Elbow	Biceps	Wrist
k (N.m/rd)	180	70	40	10
c (N.m.s/rd)	$0.002k_1$	$0.002k_2$	$0.002k_3$	$0.001k_4$

B. Equations of Motion

Motion will start at an instant of stable equilibrium: $\theta_1 = 0°, \theta_2 = 90°\,and\,\theta_a = 0°$ with zero initial velocities.

At this instant, the kinematics are determined using Coriolis' theorem for the system (with a dual DVA) formed by five frames rotating with respect to the global coordinate system.

Then, equations of motion for the non-Lagrangian system can be derived using the Rayleigh dissipation function to form the equations of motion for the dissipative systems:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i} - \frac{\partial U}{\partial \dot{q}_i}\right) - \frac{\partial T}{\partial q_i} + \frac{\partial U}{\partial q_i} = F_i + F_{c_i}, i = \{1,2,3,a_1,a_2\}$$

$$F_{c_i} = -\frac{\partial R}{\partial \dot{q}_i}\,with\,R_i = \frac{1}{2}c\dot{q}_i^2 \qquad (3)$$

Then, the generalized equation of motion has the form:

$$[M]\{\ddot{\theta}\} + [C]\{\dot{\theta}\} + [K]\{\theta\} = \{f\} \qquad (4)$$

Where,
$$M = \begin{bmatrix} M_{11} & M_{12} & M_{13} & M_{14} & M_{15} \\ M_{21} & M_{22} & M_{23} & M_{24} & M_{25} \\ M_{31} & M_{32} & M_{33} & M_{34} & M_{35} \\ M_{41} & M_{42} & M_{43} & M_{44} & M_{45} \\ M_{51} & M_{52} & M_{53} & M_{54} & M_{55} \end{bmatrix} \qquad (5)$$

$$M_{11} = \left(I_1 + m_1 a_1^2\right) + \left(I_2 + m_2 a_2^2\right) + m_2 l_1^2 + m_3\left(l_1^2 + l_3^2\right)$$
$$+ m_4\left(l_1^2 + l_2^2 + a_4^2 + 2l_2 a_4\right)$$
$$+ m_{a_1}\left(l_1^2 + l_{a_1}^2 + a_3^2 + 2l_{a_1} a_3\right) + m_{a_2}\left(l_1^2 + l_{a_2}^2 + a_5^2 + 2l_{a_2} a_5\right)$$

$$M_{12} = \left(I_2 + m_2 a_2^2\right) + m_3 l_3^2 + m_4\left(l_2^2 + a_4^2 + 2l_2 a_4\right)$$
$$+ m_{a_1}\left(l_{a_1}^2 + a_3^2 + 2l_{a_1} a_3\right) + m_{a_2}\left(l_{a_2}^2 + a_5^2 + 2l_{a_2} a_5\right)$$

$$M_{13} = m_4\left(a_4^2 + l_2 a_4\right), \; M_{14} = m_{a_1}\left(a_3^2 + l_{a_1} a_3\right)$$

$$M_{15} = m_{a_2}\left(a_5^2 + l_{a_2} a_5\right), \; M_{21} = M_{12}, \; M_{22} = M_{12}$$

$$M_{23} = M_{13}, \; M_{24} = M_{14}, \; M_{25} = M_{15}, \; M_{31} = M_{13}$$

$$M_{32} = M_{23}, \; M_{33} = I_4 + m_4 a_4^2, \; M_{34} = 0, \; M_{35} = 0$$

$$M_{41} = M_{14}, \; M_{42} = M_{24}, \; M_{43} = M_{34}, \; M_{44} = m_{a_1} a_3^2$$

$$M_{45} = 0, \; M_{51} = M_{15}, \; M_{52} = M_{25}, \; M_{53} = M_{35}$$

$$M_{54} = M_{45}, \; M_{55} = m_{a_2} a_5^2$$

$$K = \begin{bmatrix} k_1 + k_3 & k_3 & 0 & 0 & 0 \\ k_3 & k_2 + k_3 & 0 & 0 & 0 \\ 0 & 0 & k_4 & 0 & 0 \\ 0 & 0 & 0 & k_{a_1} & 0 \\ 0 & 0 & 0 & 0 & k_{a_2} \end{bmatrix} \quad (6)$$

$$C = \begin{bmatrix} c_1 + c_3 & c_3 & 0 & 0 & 0 \\ c_3 & c_2 + c_3 & 0 & 0 & 0 \\ 0 & 0 & c_4 & 0 & 0 \\ 0 & 0 & 0 & c_{a_1} & 0 \\ 0 & 0 & 0 & 0 & c_{a_2} \end{bmatrix} \quad (7)$$

$$f = \{f_1 \quad f_2 \quad 0 \quad 0 \quad 0\}^T \quad (8)$$

Input moments to the model are considered as muscular activity due to shoulder and elbow muscles activation to see their effect on the motion transmitted to the palm. Same equations are derived for the primary system and systems controlled by single absorbers, but with some simplifications.

The natural frequencies of the primary system are obtained to be in range of pathological tremor:

$$\omega_{n_1} = 3.564Hz, \omega_{n_2} = 5.296Hz \, and \, \omega_{n_3} = 12.496Hz$$

C. Absorber Design

The absorber is modeled as stainless steel alloy cantilevered beam with a copper mass attached along its length "Fig. 3". Two absorbers are designed to be attached on the forearm.

The dimensions of 'absorber 1' and 'absorber 2' systems are shown in Table II. The effect of using each absorber alone is compared to the response of the hand when combining both absorbers together attached at the same position on the forearm to form the 'dual DVA'. The absorbers are considered with an additional 52 g respresenting the mass of the controller device.
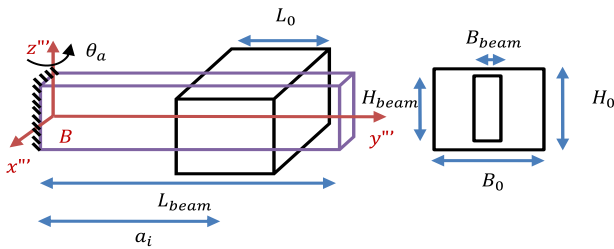
Figure 3. The dynamic vibration absorber (DVA)

TABLE II
DIMENSIONS OF THE DESIGNED ABSORBER

Dimensions		L (cm)	H (cm)	B (cm)
Absorber 1	Beam 1	9	2.4	0.03
	Attached mass 1	2.24	2.3	2.3
Absorber 2	Beam 2	7.5	1.8	0.03
	Attached mass 2	2.4	2	2

The natural frequency of the absorber's system is approximated using Dunkerley's formulation which gives a lower band approximation:

$$\frac{1}{\omega_a^2} \cong \frac{1}{\omega_{beam}^2} + \frac{1}{\omega_{m_0}^2} \quad (9)$$

Where, $$\omega_{beam} = 3.5160 \sqrt{\frac{E_{beam}I_{beam}}{m_{beam}L_{beam}^3}} \quad (10)$$

$$\omega_{m_0} = \sqrt{\frac{6E_{beam}I_{beam}}{m_0 a_i^2 (3L_{beam} - a_i)}}, \, i = \{3,5\} \quad (11)$$

$$E_{beam} = 189.6GPa, \rho_{beam} = 7800 \, kg/m^3 \, and \, \rho_{m_0} = 8900 \, kg/m^3$$

Hence, $$a_3 = 7.76cm \, and \, a_5 = 5.31cm \quad (12)$$

III. RESULTS

The behavior of the system at the proximal joints can be determined by solving the system's equation of motion. The absorber's parameters are determined depending on the tuning condition. Graphs will show the flexion angle before and after adding the absorbers in the frequency and time domains. Then, the period of operation of the designed absorber is calculated.

A. Response

The response is determined using the complex transfer function of the system:

$$H(\omega) = \{(-\omega^2[M] + [K]) + j\omega[C]\}^{-1} \quad (13)$$

The input forces due to shoulder and elbow muscles activation are assumed to be equal. Both are assumed to behave as sinusoidal functions to reflect the rhythmic motion of hand tremor and are driven at the resonance of the primary system. Muscles can operate under a range of driving frequencies, but only the critical frequencies in range of resting tremor will be considered.

$$f_k = F_{k_1} \cos(\omega_1 t) + F_{k_2} \cos(\omega_2 t), \, k = \{1,2\}$$
$$F_{k_1} = F_{k_2} = 1 \, Nm \, and \, \omega_m = \omega_{n_m}, \, m = \{1,2\} \quad (14)$$

The response depends on the Receptance function having the form:

$$\alpha_k = \begin{Bmatrix} \dfrac{A_{1_{ik}} + B_{1_{ik}}}{A_2 + jB_2} \\ \vdots \\ \dfrac{A_{1_{nk}} + B_{1_{nk}}}{A_2 + jB_2} \end{Bmatrix}, \, i = \{1,\dots,5\} \quad (15)$$

The response is the sum of the responses due to each excitation moment. Then, the response in frequency domain

can be determined to analyze the behavior of the system over a range of driving frequencies.

$$\Theta = \sum_{k=1}^{2}\sum_{m=1}^{2}\alpha_k F_{k_m} \qquad (16)$$

As well as the time domain response, in order to analyze its behavior at the specified excitation frequencies.

$$\theta_{ik} = |\Theta_{ik}| e^{j(\omega_m t - \varphi)}$$

$$|\Theta_{ik}| = F_{k_m}\sqrt{\frac{A_{1_{ik}}^2 + B_{1_{ik}}^2}{A_2^2 + B_2^2}} \quad and \quad \varphi = \tan^{-1}\left(\frac{B_{1_{ik}}}{A_{1_{ik}}}\right) \qquad (17)$$

The response serves in designing the unknown absorber's parameters. Each absorber is tuned from the root of the real part in the numerator of the chosen response, and then the stiffness of the absorber can be calculated. The damping coefficient of each absorber is assumed to be proportional to its stiffness coefficient by a constant.

Although there is no damper added to the absorber, a very little damping will be provided from beam's material. However, the damping coefficient must be increased to maintain low amplitudes of oscillation in the absorber and to avoid impact with the forearm. So, the beams material can be coated by damping material to increase its value to:

$$c_{a_1} = 0.05 k_{a_1} \, and \, c_{a_2} = 0.05 k_{a_2} \qquad (18)$$

'Absorber 1' is designed to be tuned at the palm's response 'θ_{31}' with the first resonance frequency 'ω_{n1}', due to shoulder muscle activation 'F_{11}'.

$$A_{1_{31}} = -\omega^6\left(M_{21}M_{32}M_{44} - M_{22}M_{32}M_{44}\right) + \omega^4\left(-M_{32}M_{44}k_2 + M_{21}M_{32}k_a - M_{22}M_{32}k_a - M_{32}c_2 c_a\right) + \omega^2\left(M_{32}k_2 k_a\right) \qquad (19)$$

'Absorber 2' is designed to be tuned at the palm's response 'θ_{32}' with the second resonance frequency 'ω_{n2}', due to elbow muscle activation 'F_{22}'.

$$A_{1_{32}} = -\omega^6\left(M_{11}M_{32}M_{44} - M_{12}M_{32}M_{44}\right) - \omega^4\left(M_{32}M_{44}k_1 + M_{11}M_{32}k_a - M_{12}M_{32}k_a + M_{32}c_1 c_a\right) + \omega^2\left(M_{32}k_1 k_a\right) \qquad (20)$$

Solving (23) and (24), then all absorbers parameters are determined (see Table III) to satisfy the tuning conditions:

$$A_{1_{31}} = 0 \, and \, A_{1_{32}} = 0 \qquad (21)$$

TABLE III
PARAMETERS OF THE DESIGNED ABSORBER

Parameters	m_a (g)	k_a (N.m/rd)	c_a (N.m.s/rd)
Absorber 1	105.239	0.3351	$0.05 k_a$
Absorber 2	85.044	0.2388	$0.05 k_a$

Both absorbers can have their equivalent conventional mass-linear spring and damper system shown in "Fig. 4".

These pendulum absorbers can be attached to the forearm in the form of their equivalent models with the corresponding parameters using determined (22) and are shown in Table IV.

$$k_{a_{1x}} \cong \frac{k_{a_1}}{a_3^2} \, and \, k_{a_{2x}} \cong \frac{k_{a_2}}{a_5^2} \qquad (22)$$

TABLE IV
PARAMETERS OF THE EQUIVALENT LINEAR ABSORBERS

Parameters	m_a (kg)	k_{ax} (N/m)	c_{ax} (N.s/m)
Absorber 1	105.239	55.55	$0.05 k_{ax}$
Absorber 2	85.044	84.47	$0.05 k_{ax}$

B. Numerical Simulations

Results are obtained using MATLAB in order to show the behavior of the system in the time and frequency domains. The palm is the main organ in the hand which is responsible of most of human's activity. So, graphs will show the behavior of the palm due to shoulder and elbow muscles activation when the absorbers are tuned at the palm's response. The time and frequency domain responses of the palm are shown for the uncontrolled system and compared to that of the system controlled by different passive absorbers "Fig. 5". The performance of 'absorber 1' and 'absorber 2', each is tuned to different driving frequency, will be compared to that of the 'dual DVA' which is a combination of both absorbers under the same controller device. The performance will be analyzed in terms of absorbers capability in reducing tremor amplitude.

$$\% \, \text{Reduction} = \frac{\Theta_{\text{uncontrolled}} - \Theta_{\text{controlled}}}{\Theta_{\text{uncontrolled}}} \times 100 \qquad (23)$$

In "Fig. 5a", the frequency domain response of the palm is shown. Absorber's high damping coefficient has damped some peaks which result from tuning the absorber. Absorber's damping coefficient can be decreased to show the suppression in the 1st two peaks of the primary system near zero with the shifted peaks in the controlled system. However, this will cause the extremely high amplitudes of vibration of the absorber which will results in its damage. The natural frequencies of the controlled systems are determined as:

- For the 4 DOF system with 'Absorber 1':

$$\omega_{n_1} = 3.489 Hz, \omega_{n_2} = 3.734 Hz, \omega_{n_3} = 5.253 Hz, \omega_{n_4} = 13.256 Hz$$
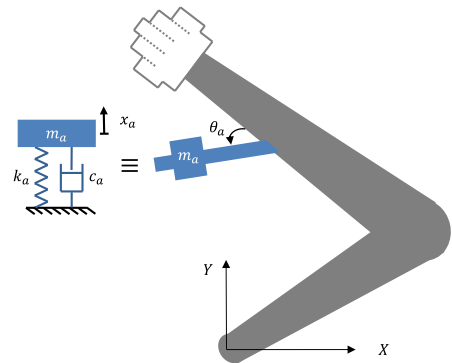
Figure 4. Equivalent model of the pendulum absorber

- For the 4 DOF system with 'Absorber 2':

$$\omega_{n_1} = 3.491 Hz, \omega_{n_2} = 5.155 Hz, \omega_{n_3} = 5.219 Hz$$
$$and \; \omega_{n_4} = 12.367 Hz$$

- For the 5 DOF system with the 'Dual DVA':

$$\omega_{n_1} = 3.400 Hz, \omega_{n_2} = 3.737 Hz, \omega_{n_3} = 5.145 Hz,$$
$$\omega_{n_4} = 5.162 Hz \; and \; \omega_{n_5} = 12.321 Hz$$

Table V shows the percentage of reduction in tremor magnitude at the peak of each controlled system with respect to the uncontrolled system. All modeled absorbers had caused tremor reduction. 'Absorber 1' was more efficient at the 1st peak of 3.489Hz with 37.6% reduction, which is near its tuning frequency. The 'dual DVA' was more efficient than 'absorber 1' at the 1st peak. It causes 40.9% reduction at 3.491Hz which is near the tuning frequency of 'absorber 1'. The 'dual DVA' was effective also at the 2nd peak with 29.8% reduction at 5.162Hz. 'Absorber 2' causes a comparable reduction in tremor magnitude at the three peaks with maximum reduction of 14.9% at the 3rd peak of 12.367Hz. The 'dual DVA' was more efficient than 'absorber 2' at the 3rd peak. It causes 32.6% reduction at 12.321 Hz.

TABLE V
PERCENTAGE OF REDUCTION IN FREQUENCY DOMAIN

% Reduction	1st Peak	2nd Peak	3rd Peak
Absorber 1	37.6	13.2	19.5
Absorber 2	14.7	12.8	14.9
Dual DVA	40.9	29.8	32.6

In "Fig. 5b", the time domain response of the palm is shown when the system is excited at the 1st two resonance frequencies. The % reduction ranges due to attaching each absorber are shown in table VI. The 'dual DVA' causes the highest reduction in range of 33.1 – 50.6%.

(a)

(b)
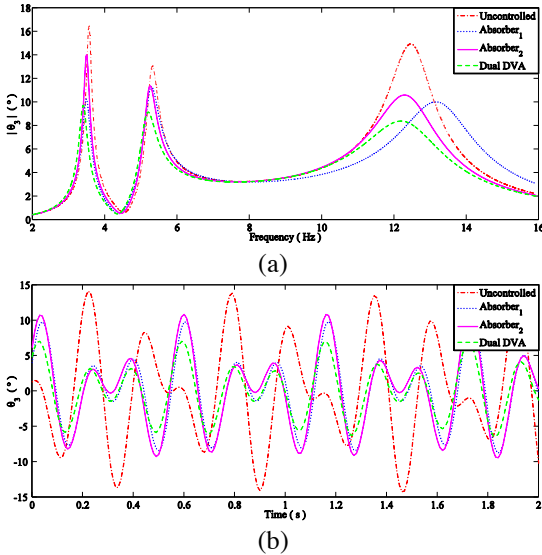
Figure 5. Response at wrist joint in (a) frequency domain and (b) time domain

TABLE VI
PERCENTAGE OF REDUCTION IN TIME DOMAIN

	Absorber 1	Absorber 2	Dual DVA
% Reduction	12.1 – 31.6	13.81 – 25.1	33.1 – 50.6

C. Absorber's Life Time

The period of operation for the designed absorber can be calculated to know its estimated number of cycles before its failure. "Fig. 6" shows the behavior of 'absorber 1' and 'absorber 2' when attached separately to the primary system and their behavior when attached together as a 'dual DVA'.

The behavior of the absorbers in time domain is shown in "Fig. 6a and b". The absolute maximum and minimum tremor amplitude for each absorber alone and their combination are summarized in Table VII.

TABLE VII
MAXIMUM AND MINIMUM FLEXION ANGLE OF THE ABSORBERS

Flexion Angle	Absorber 1	Absorber 2	Dual DVA	
			Absorber 1	Absorber 2
Θ_{min}	6.6⁰	5.3⁰	3.6⁰	2.1⁰
Θ_{max}	20.5⁰	16.9⁰	14.2⁰	10.2⁰

The cold rolled – stainless steel alloy (type 301) is selected as beams material of both absorbers (S_{ut} = 1379 MPa and S_y = 1138MPa) and is assumed to have 90% reliability. Then, the safety factors against failure by fatigue and yielding of the beam can be calculated according to the following equations:

$$\sigma = k_f \frac{M_a \frac{B_0}{2}}{I_{beam}}, \quad where \; M_a = k_a \Theta_a \qquad (24)$$

$$S_e' = 0.5 S_{ut} \quad (S_{ut} < 1400 \, MPa) \qquad (25)$$

$$S_e = C_{load} C_{surf} C_{size} C_{temp.} C_{relib.} S_e' \qquad (26)$$

(a)

(b)
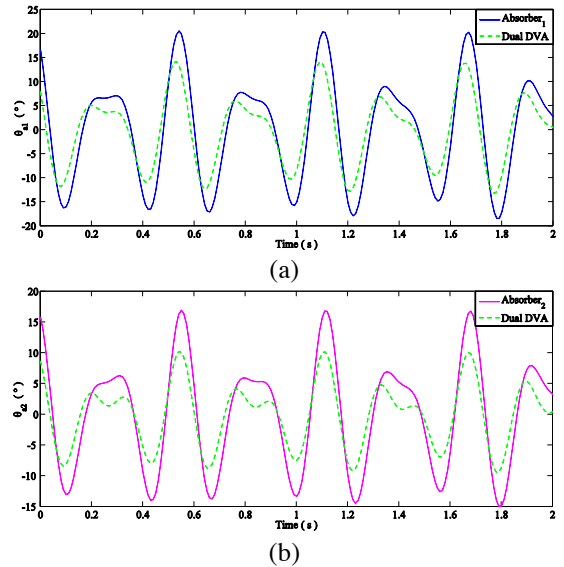
Figure 6. Response in time domain : (a) 'absorber 1' and (b) absorber 2'

$$N_f = \frac{S_e S_{ut}}{\sigma'_{alt}.S_{ut} + \sigma'_{mean} S_e} \quad \text{and} \quad N_y = \frac{S_y}{\sigma_{max}} \qquad (27)$$

The obtained safety factors for each of 'absorber 1', 'absorber 2' and their combination in the 'dual DVA' are shown in Table VIII. All the absorbers are safe against failure by fatigue ($N_f > 1$), so they are designed to have infinite life time. The beams are also safe against yielding ($N_y > 1$).

TABLE VIII

FATIGUE AND YIELDING SAFETY FACTOR OF ABSORBERS BEAMS

Safety Factor	Absorber 1	Absorber 2	Dual DVA	
			Absorber 1	Absorber 2
N_f	2.30	2.70	3.17	3.97
N_y	3.41	4.36	4.93	7.22

IV. DISCUSSION

In a previous study, Hashemi et al. [8] has not considered the motion at the wrist joint. The system was excited at elbow joint with single excitation frequency at 2.24 Hz neighborhood the resonance frequency. The absorbers natural frequency was 2.755 Hz which is higher than the excitation frequency. The tremor flexion motion amplitude of the excited system was reduced due to attaching the absorber from 3^0 to 0.5^0 at shoulder joint and from 4^0 to 2.2^0 at elbow joint. The modeled hand was fabricated for a comparison with the numerical study. Experimental and theoretical results were qualitatively similar.

Rahnavard et al. [9] used the same 2 DOF hand model used by Hashemi. However, they designed an optimum single DOF dynamic absorber to suppress rest tremor at the joints. The absorber was able to reduce more than 98% and 80% of the flexion motion at elbow joint at the first and second natural frequencies of the primary system in the frequency domain. Tremor is modeled to have random nature with two types instead of sinusoidal one. In the first type, the absorber was able to reduce 60% and 39% of flexion motion at shoulder and elbow joints and in the second type 33% and 50% reduction is revealed in the time domain

In our study, a 3 DOF system is designed by considering flexion motion of the palm. Single and dual DVAs are compared to select the best efficient absorber. The dual DVA shows the highest performance which reduces 33.1 – 50.6% of flexion motion at the palm.

V. CONCLUSION

Human hand was modeled as 3 DOF system to describe the flexion angular motion at the proximal joints in the horizontal plane. The modeled system reflects the biodynamic response of patients suffering from pathological tremor. The system was excited at resonance using two harmonic functions due to each of shoulder and elbow muscles activation. Mechanical treatment using tuned vibration absorbers was suggested to help the neurologically disordered people.

The effectiveness of the single and the dual DVAs were compared in terms of their capability in reducing the involuntary tremor amplitude. Two single DVAs were used and each is tuned to one of the excitation frequencies. Each absorber alone was able to reduce tremor amplitude at the proximal joints and especially at the wrist joint. Both absorbers are combined together under one device to form the dual DVA tuned at both excitation frequencies. The dual DVA was more efficient than both absorbers in attenuation the tremor. It causes 33.1 – 50.6% reduction in the flexion motion of the primary system. The absorbers were designed respecting their geometric limitations and constraints and by considering patients comfort in holding this absorber. Absorber's device was considered to allow absorbers vibration in flexion direction only. These absorbers are expected to operate for an infinite life without yielding.

As future work, an experimental rig will be fabricated to hold the actual hand and permit the flexion angular displacement at the proximal joints. Also, an experiment can be conducted on a real case to validate the semi-analytical approach and to probe any potential "real-life" side effects of the system.

REFERENCES

[1] Deuschl, G., Bain, P., & Brin, M. (1998). Consensus statement of the Movement Disorder Society on tremor. Movement Disorders, 13(S3), 2-23.

[2] Morrison, S., Kerr, G., & Silburn, P. (2008). Bilateral tremor relations in Parkinson's disease: Effects of mechanical coupling and medication. Parkinsonism & related disorders, 14(4), 298-308.

[3] Rahnavard, M., Hashemi, M., Farahmand, F., & Dizaji, A. F. (2014). Designing a hand rest tremor dynamic vibration absorber using H2 optimization method. *Journal of Mechanical Science and Technology*, 28(5), 1609-1614.

[4] Miwa, H. (2007). Rodent models of tremor. *The Cerebellum*, 6(1), 66-72.

[5] Kazi, S., As' Arry, A., Zain, M. M., Mailah, M., & Hussein, M. (2010). Experimental implementation of smart glove incorporating piezoelectric actuator for hand tremor control. *WSEAS Transactions on Systems and Control*, 5(6), 443-453.

[6] Morrison, S., Kerr, G., & Silburn, P. (2008). Bilateral tremor relations in Parkinson's disease: Effects of mechanical coupling and medication. *Parkinsonism & related disorders*, 14(4), 298-308.

[7] Louis, E. D. (2001). Essential tremor. *New England Journal of Medicine*, 345(12), 887-891.

[8] Hashemi, S. M., Golnaraghi, M. F., & Patla, A. E. (2004). Tuned vibration absorber for suppression of rest tremor in Parkinson's disease. *Medical and Biological Engineering and Computing*, 42(1), 61-70.

[9] Rahnavard, M., Hashemi, M., Farahmand, F., & Dizaji, A. F. (2014). Designing a hand rest tremor dynamic vibration absorber using H 2 optimization method. *Journal of Mechanical Science and Technology*, 28(5), 1609-1614.

[10] Taheri, B. (2013). Real-Time Pathological Tremor Identification and Suppression in Human Arm via Active Orthotic Devices..

[11] M Igusa, T., & Xu, K. (1994). Vibration control using multiple tuned mass dampers. *Journal of sound and vibration*, 175(4), 491-503.

[12] Brennan, M. J. (1997). Characteristics of a wideband vibration neutralizer.*Noise Control Engineering Journal*, 45(5), 201-207.

[13] Drillis, R., Contini, R., & Maurice Bluestein, M. (1966). Body segment parameters 1. Artificial Limbs, 44.

List of Symbols

a_1, a_2, a_4 — Distance between centroid of the upper arm, forearm and the palm to its corresponding proximal joint (m)

a_3, a_5 — Position from the attached mass to the fixed joint of 'absorber 1' and 'absorber 2' (m)

B_{beam}, B_0 — Base of absorber's cantilevered beam and the attached mass (m)

C — Damping coefficient matrix of the system (N.m.s/rd)

$C_{load}, C_{reliab}, C_{size}, C_{surf}, C_{temp}$ — Load, reliability, size, surface and temperature correction factors

c_a, c_1, c_2, c_3, c_4 — Bending damping coefficient of absorber's beam and shoulder, elbow, biceps brachii and wrist muscles (N.m.s/rd)

$c_{a_{1x}}, c_{a_{2x}}$ — Linear damping coefficients of 'absorber 1' and 'absorber 2' (N.s/m)

D_1, D_2, D_4 — Density of the upper arm, forearm and the palm (kg/m³)

E_{beam} — Modulus of elasticity of absorber's beam material (GPa)

F_i, F_{c_i} — Generalized conservative and frictional moments (N.m)

F_1, F_2 — Input moment at shoulder and wrist joints (N.m)

f — Vector of bending moment functions (N.m)

g — Gravitational acceleration (m/s²)

H_{beam}, H_0 — Height of absorber's cantilevered beam and the attached mass (m)

$H(\omega)$ — Complex transfer function (rd/N.m)

I_{beam} — Area moment of inertia of the beam (m⁴)

I_1, I_2, I_4 — Mass moment of inertia of the upper arm, forearm and the palm (kg.m²)

K — Stiffness coefficient matric of the system (N.m/rd)

k_a, k_1, k_2, k_3, k_4 — Bending damping coefficient of absorber's beam and shoulder, elbow, biceps brachii and wrist muscles (N.m/s)

$k_{a_{1x}}, k_{a_{2x}}$ — Linear stiffness coefficients of 'absorber 1' and 'absorber 2' (N.m/s)

k_f — Fatigue stress concentration factor

L_{beam}, L_0 — Length of absorber's cantilevered beam and the attached mass (m)

l_a, l_3 — Distance between absorber's joint and controller device to the elbow joint (m)

l_1, l_2, l_4 — Distance from the location of concentrated masses of the upper arm, forearm and the palm to its corresponding proximal joint (m)

M — Mass matrix of the system (kg.m²)

m_{a_1}, m_{a_2} — Effective proof mass of 'absorber 1' and 'absorber 2' (kg)

m_0 — Mass attached to absorber's beam (kg)

m_1, m_2, m_3, m_4 — Masses of the upper arm, forearm, controller device and the palm (kg)

M_a — Bending reaction moment on absorber's beam (N.m)

N_f, N_y — Fatigue and yielding safety factors

$q_i, \dot{q}_i$ — Generalized coordinates of angular displacement and velocity (rd) and (rd/s)

R — Raleigh dissipation function (J)

S_e, S'_e — Corrected and uncorrected endurance limit of the absorber's beam (MPa)

S_{ut}, S_y — Ultimate and yielding strength of absorber's beam (MPa)

s — Proportional constant relating stiffness and damping coefficients

T, U — Kinetic and potential energy of the system (J)

α_j — Receptance transfer function (rd/N.m)

$\theta, \dot{\theta}, \ddot{\theta}$ — Angular displacement, velocity and acceleration functions (rd), (rd/s) and (rd/s²)

Θ — Angular displacement magnitude (rd)

$\Theta_{controlled}, \Theta_{uncontrolled}$ — Angular displacement magnitude of the controlled and uncontrolled systems (rd)

ω — Driving frequency (rd/s)

ω_a, ω_{beam} — Fundamental frequency of absorber's system and its beam alone (rd/s)

ω_{m_0} — Natural frequency of the attached mass (rd)

ω_n — Natural frequency of the system (rd/s)

φ — Phase angle of the response (⁰)

σ — Bending stress (MPa)

ρ_{beam}, ρ_0 — Density of the cantilevered beam of the absorber and the attached mass (kg/m³).

Evaluation of the Performance of Two Smoothers within a Multi-Grid Framework

Sobhi Takkoush

Department of Mechanical Engineering
Faculty of Engineering and Architecture
American University of Beirut
P.O. Box 11-0236, Beirut, Lebanon
smt09@mail.aub.edu

Abstract- Incomplete LU decomposition with no fill-in ILU(0) has been used as a standard smoother with algebraic multigrid solvers in many applications. With recent developments, new techniques have emerged following the sparse approximate inverse approach. Methods like AINV have been implemented as preconditioners but rarely as smoothers. In this paper, the ILU (0) and AINV will be implemented as smoothers in both scalar and block versions within an algebraic multigrid solver. A comparative assessment of the performance of these techniques as smoothers in an algebraic multigrid solver will be performed in the context of a finite volume discretization method using open source CFD toolbox OpenFOAM. Two test cases with high aspect ratio are used to compare the two smoothers. Residual convergence rates, number of iterations as well as CPU time are used to evaluate the performance. For segregated flow solver, AINV and ILU(0) have same convergence rate, number of iterations and CPU time with slight difference; however, ILU(0) outperforms AINV in every aspect for coupled flow solver.

I. INTRODUCTION

In 3D fluid flow problems, the variables to be solved are mainly the velocity fields in x,y and z directions as well as the pressure field by discretizing and solving the momentum and continuity equations. The solution procedure can be divided into segregated and coupled. In segregated solver, the continuity and momentum equations are decoupled and solved sequentially [1]. The individual momentum equations are solved for each velocity component then continuity equation is solved for pressure. Although segregated solvers require less memory, they have low convergence rate and need more iterations. In contrast, the coupled or block algorithms have high rate of convergence on the expense of high memory since all velocity components and pressure fields are solved at the same time or simultaneously where all discretized equations are solved in one system [2] .

System of algebraic equation at centroid of one cell in segregated:

$$a_C \emptyset_C + \sum_{NB} a_F \emptyset_F = b_C \qquad (1)$$

While for coupled system [3], the system of algebraic equations at centroid of one cell would be:

$$
\begin{bmatrix}
a_C^{uu} & a_C^{uv} & a_C^{uw} & a_C^{up} \\
a_C^{vu} & a_C^{vv} & a_C^{vw} & a_C^{vp} \\
a_C^{wu} & a_C^{wv} & a_C^{ww} & a_C^{wp} \\
a_C^{pu} & a_C^{pv} & a_C^{pw} & a_C^{pp}
\end{bmatrix}
\cdot
\begin{bmatrix}
u_C \\ v_C \\ w_C \\ p_C
\end{bmatrix}
+
$$

$$
\sum_{NB}
\begin{bmatrix}
a_{NB}^{uu} & a_{NB}^{uv} & a_{NB}^{uw} & a_{NB}^{up} \\
a_{NB}^{vu} & a_{NB}^{vv} & a_{NB}^{vw} & a_{NB}^{vp} \\
a_{NB}^{wu} & a_{NB}^{wv} & a_{NB}^{ww} & a_{NB}^{wp} \\
a_{NB}^{pu} & a_{NB}^{pv} & a_{NB}^{pw} & a_{NB}^{pp}
\end{bmatrix}
\cdot
\begin{bmatrix}
u_{NB} \\ v_{NB} \\ w_{NB} \\ p_{NB}
\end{bmatrix}
=
\begin{bmatrix}
b_C^u \\ b_C^v \\ b_C^w \\ b_C^p
\end{bmatrix}
$$

$$(2)$$

In either approaches (segregated or coupled), we need a robust and fast iterative linear solver. While requiring low storage, the rate of convergence of standard iterative methods (Gauss-Seidel, Jacobi, and SOR) drastically deteriorates as the mesh is refined since the size of the algebraic system increases [4]. This problem is more critical in highly compressible systems such as turbo machinery, supersonic flow, anisotropic diffusion and others where an ill-conditioned and weakly diagonally dominant matrix arises which leads to large Eigen values and breakdown in convergence rate of the solution [5]. This has constituted a severe limitation for iterative solvers. The development of Multigrid methods remedied this weakness.

Multigrid Methods were basically developed for solving elliptic problems. These methods were independently introduced by Fedorenko [6] and Poussin [7], and then it gained popularity with the work of Settari and Azziz [8], and later with the theoretical work of Brandt [9]. These methods became widely accepted and used in computational fluid dynamics where the Poison equation, which is considered an elliptic type, arises from the discretization of pressure equation in the solution of Navier Stokes equation [10-14].

The rate of convergence is related to the function of error frequency. Since iterative solvers are capable of removing oscillatory or high frequency errors but not the smooth components of the error, solving on successively coarser grid errors that are smooth on a fine grid will look more oscillatory on a coarser one and thus are easier to be removed by the

iterative solver. This is why these iterative solution methods are denoted by smoothers in the context of multigrid methods.

The purpose of multigrid methods is to improve the convergence rate of iterative solvers by transferring the remaining low frequency errors (smooth errors) from a fine grid to a coarser grid where the low frequency error will appear as high frequency; making it easier for the smoother to detect the error and eliminate it, thus increasing the rate of convergence [4,5,15]. This is done through a hierarchy of coarse grids as shown in Figure 1 by restricting the smooth error to coarser grids where it is smoothed until the coarsest grid is reached. Then this correction is prolonged back to fine grids also with application of smoothers, until reaching the original finest grid.
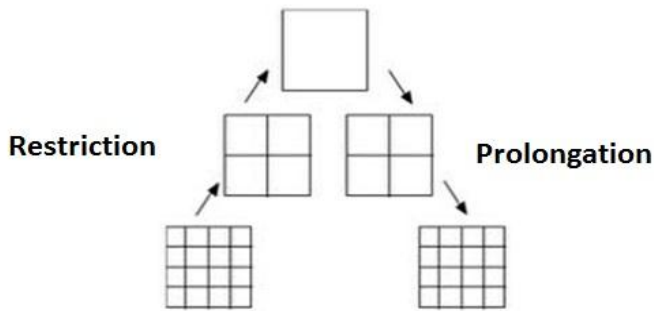
Figure 1. A schematic grid systems hierarchy with MG approach [16]

Traditionally the ILU(0) [17,18] (Incomplete LU decomposition with no fill-in) iterative methods have been used as smoother in the context of algebraic multigrid methods. AINV [19-24] Sparse Approximate Inverse iterative methods have never been used for that purpose.

The objective of this paper is to improve the performance of iterative solvers with Multigrid by implementing AINV as smoother in an algebraic multigrid environment, an acceleration framework, in the context of the Finite Volume Method (FVM), and to evaluate and compare its performance to the ILU(0) Smoother, in terms of CPU time by solving a number of three-dimensional fluid flow problems using an open source software environment OpenFOAM.

The two algorithms are first presented in scalar version then in block (coupled) version. This is followed by a comparison of the performance of the two smoothers by solving two test-case problems. Finally the convergence rate and CPU time are compared and discussed.

II. THE ILU(0) SMOOTHER

Incomplete factorization methods were introduced by Buleev [21] and Oliphant [22, 25]. The first who introduced the term preconditioning was Evans [26] who also considered the use of sparse LU factors as preconditioner. These methods were of particular interest in the field of oil reservoir simulation [27, 28]. The major breakthrough was in mid 1970s by Meijerink and van der Vorst [29] who established, recognized and proved the existence of incomplete factorization for M-matirces and showed that preconditioning Conjugate Gradient, by using Incomplete factorizations, can result in efficient results. In their paper, they also introduced the ILU (0) which is the most basic form of ILU preconditioner. This article had a very important role in capturing the attention of researchers about the importance of this preconditioning technique.

A standard LU factorization of matrix A will lead to A=LU where L is lower matrix and U is upper matrix, however this has very high computation and storage cost. A more simple approach is an incomplete factorization of coefficient matrix A which will yield A=LU+r where r is the residual of factorization. The algorithm of ILU (0) by Van der Vosrt [29-31] will be described. In ILU (0), factorization can be performed using Gaussian elimination but the LU factors have the same nonzero patterns as in matrix A, so any new non-zero element arising in the process is dropped if it appears at location where zero element appear in A. The algorithm used is the one in OpenFOAM is shown in algorithm 1.

Algorithm 1
ILU (0)

- Take Coefficient matrix A as input
- For k=1,2,…n:
- Compute Residual $\quad r^{(k)} = b - Ax^{(k)}$
- For i=k+1,…n:
- Compute l values $a_{ik}a_{kk}^{-1}$
- Solve $y^{(k)} = L^{-1}r^{(k)}$ by forward substitution
- For j=k+1,…n:
- Compute u values $a_{ij} = a_{ij} - a_{ik}a_{kj}$
- Solve $d^{(k)} = U^{-1}y^{(k)}$ by backward substitution
- Update solution Solve $\quad x^{(k+1)} = x^{(k)} + d^{(k)}$
- Repeat until $\left\| r^{(k)} \right\|_2 \le \varepsilon$ where ε is tolerance

III. THE AINV SMOOTHER

Inner-product version of AINV (Approximate Inverse) method had been developed by Benzi and Tuma [20]. Given sparse matrix $A \in \mathbb{R}^{nxn}$, this method is based on incomplete inverse factorizations which mean incomplete factorizations of A^{-1}. If the factorization of A=LDU where L is unit lower triangular matrix, D is diagonal matrix, and U is unit upper triangular matrix, then $A^{-1} = U^{-1}D^{-1}L^{-1} = ZD^{-1}W^T$ where $Z = U^{-1}$ and $W = L^{-T}$ are unit upper triangular matrices. AINV is based on an algorithm which computes two sets of vectors $\{z_i\}_{i=1}^n$, $\{w_i\}_{i=1}^n$, which are A-biconjugate such that $w_i^T A z_j = 0$ iff $i \ne j$ with small entries dropped to preserve sparsity.

Since the coefficient matrix is sparse (contains too many zeros), the inner product of the Benzi & Tuma version can be computationally expensive. These inner products are often zero. To avoid this, in this paper we shift to outer-product form developed by Bridson and Tang [24]. Their algorithm has same concept as that of Benzi but with some difference in the order of the algorithm loops. Let consider the vectors l and u the j'th column of row LD and DU respectively.

We have implemented the AINV smoother of Bridson and Tang in OpenFOAM framework with below algorithm.

Algorithm 2
The Outer Product Version of AINV

- Take as input the Coefficient Matrix A and drop Tolerance δ

- Set W & Z as Identity matrix I

- For i=1,2,…n :

- Compute $l = AZ_i$

- Compute $u = W_i^T A$

- Compute $D_{ii} = uZ_i$ or $D_{ii} = W_i^T l$

- For j > i

- Orthogonalize z & w by subtracting multiple columns of z from ith column of z (same for w) :

$$Z_j = Z_j - drop(\frac{u_j}{D_{ii}} Z_i, \delta) ;$$

$$W_j = W_j - drop(\frac{l_j}{D_{ii}} W_i, \delta)$$

- Apply dropping to preserve sparsity pattern with magnitude below δ are dropped

- End for

- End for

- Set Z= $[z_1, z_2….z_n]$ and W= $[w_1, w_2….w_n]$ that are calculated in above steps

- Compute $x \approx M^{-1}b = ZD^{-1}W^T b$

IV. BLOCK METHODS

The block version is almost the same as segregated; however, the main differences are that scalar variables become tensors and division would be replaced by an inverse as shown below.

Algorithm 3
Block ILU(0)

- For k=1,2,…n:

- Compute Residual $R^{(k)} = b - Ax^{(k)}$

- For i=k+1,…n:

- Compute l values $A_{ik}A_{kk}^{-1}$

- Solve $y^{(k)} = L^{-1}r^{(k)}$ by forward substitution

- For j=k+1,…n:

- Compute u values $A_{ij} = A_{ij} - A_{ik}A_{kj}$

- Solve $d^{(k)} = U^{-1}y^{(k)}$ by backward substitution

- Update solution Solve $x^{(k+1)} = x^{(k)} + d^{(k)}$

- Repeat until $\|r^{(k)}\|_2 \le \varepsilon$ where ε is tolerance

For block version of AINV, with tensors L and U as the j'th block column of row LD and DU respectively. The algorithm would become as follow.

Algorithm 4
Block AINV

- Take as input the Coefficient Matrix A and drop Tolerance δ

- Set W & Z as Identity matrix I

- For i=1,2,…n :

- Compute block $L = AZ_i$

- Compute block $U = W_i^T A$

- Compute block $D_{ii} = UZ_i$ **or** $D_{ii} = W_i^T L$

- For j > i

- Orthogonalize Z & W by subtracting multiple block columns of z from ith column of z (same for w) :

$$Z_j = Z_j - Z_i (D_{ii}^{-1} U_j)$$

$$W_j = W_j - W_i (L_j D_{ii}^{-1})$$

- Apply dropping to preserve sparsity pattern with norm of Z_j or W_j below δ are dropped

- End for

- End for

- Set Z= $[z_1, z_2….z_n]$ and W= $[w_1, w_2….w_n]$ that are calculated in above steps

- Compute $x \approx M^{-1}b = ZD^{-1}W^T b$

In addition to scalar version, we have implemented block version in OpenFOAM to be compared witl Block IU(0).

Note that in this paper, we didn't apply the dropping techniques strategy; however, we computed the approximate inverse with sparsity pattern same as that of coefficient matrix A. In this case, while comparing performance of ILU0 and AINV, both smoothers would have same sparsity pattern to that of A.

V. TEST CASES

Two test cases will be used to evaluate and compare the performance of each smoother. The test cases setup and run were conducted on OpenFOAM using segregated and coupled fluid flow solver.

A. Flow over a Stator Blade (Compressible-Segregated)

The first test case is based on the experimental setup of Hylton et al. [32]. In their study, they investigated two aero-thermodynamic linear cascades. The NASA-C3X cascade experiment, made up of three linear cascade vanes, was chosen as representative of the first stator stage of a gas turbine.

The shape of the blade of C3X in the original setup is prismatic which allow us to use periodic and symmetry boundary conditions. Furthermore, a reduction is applied to quasi 3D of the domain, shown in Figure 2, in which the 2D section is discretized with only one cell in the span direction of the blade.
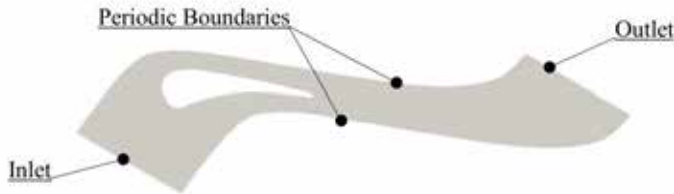
Figure 2. Computational domain and boundary conditions [33]

Table I summarizes the reference boundary conditions which are based on a particular operating point reported by Hylton et al. (code no 4422, run no 112), with an isentropic exit Mach number of 0.9.

Table I
Stator Vane Data [33]

Inlet Total Pressure	P_0	321800	Pa
Inlet Total Temperature	T_0	783	K
Dissipation Length	L_d	0.001	m
Turbulence intensity	Tu	4%	[-]
Outlet Static Pressure	P_{out}	192500	Pa

As shown in Figure 3, the computational grid is made of multi-blocks O-type grid, resulting in a mesh with size of 14,500 hexahedral elements. In addition, Figure 3 also shows that the grid fully resolves the boundary layer close to the walls of the blade and the wake. This test case, despite of its moderate mesh size, is still very demanding due to the very high anisotropic mesh used which yields elements of aspect ratios reaching a maximum of 30,000, and resulting in a very stiff system of equations [33].
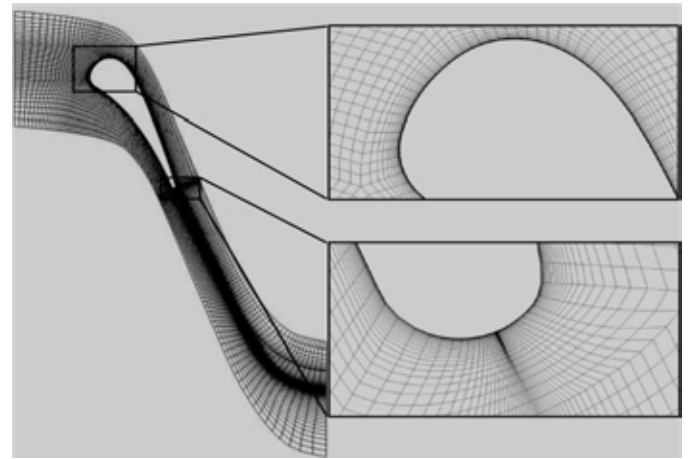
Figure 3. Computational grid for the flow over a blade problem[33]

B. 90° Pipe Bend (Incompressible-Coupled)

The second test case is based on an interesting field, which is turbulent flow through curved pipes and channels, has been investigated theoretically and experimentally for decades [34, 35]. In addition, with high computing power, numerical investigations have been developed recently to include unsteady techniques [36].

Turbulent flow through pipe bend is of great importance since it tackles important features in turbulence modeling. Some of these features are geometry induced pressure gradients, longitudinal streamline curvatures and many others [37].

The test case in interest is a 90 pipe bend with circular cross section as shown in Figure 4. The boundary conditions are of one inlet and one outlet with no slip condition on the walls and symmetry on the bottom wall for computational efficiency.
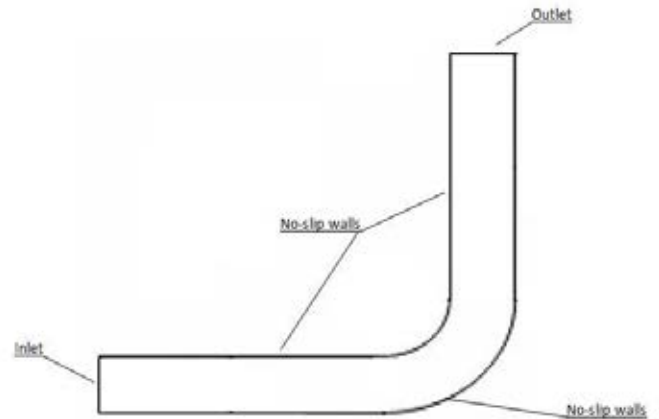
Figure 4. Computational domain and boundary conditions

Table II summarizes the initial and boundary conditions of this test case.

Inlet Pressure	P_{inlet}	Zero Gradient	*Pa*
Inlet Velocity	V_{inlet}	1.67	*m/sec*
Outlet Pressure	P_{outlet}	10000	*Pa*
Outlet Velocity	V_{outlet}	Zero Gradient	*m/sec*

As shown in Figure 5, the computational grid is made of a mesh with size of 1.5 million hexahedral elements. This test case is also of high importance due to high anisotropic mesh which results in elements of very high aspect ratio reaching maximum of 478174 which is considered very large and will definitely yields a stiff system of equations with an ill-conditioned matrix.
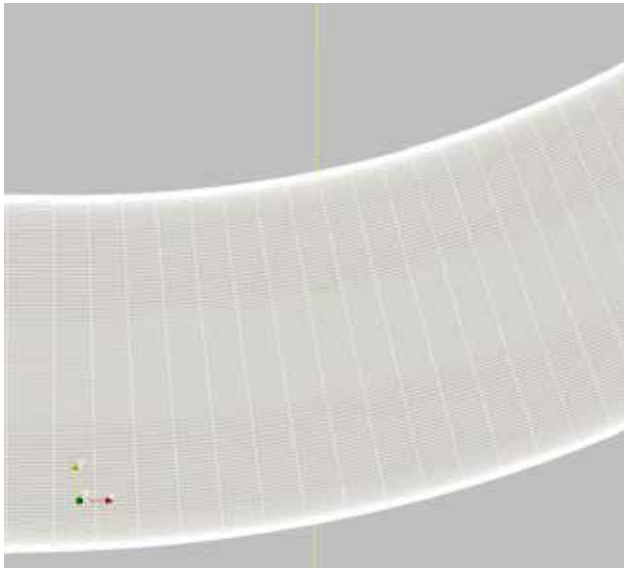
Figure 5. Top View of Mesh on the Pipe Symmetry Plane

VI. RESULTS AND DISCUSSION

We run the two test cases on OpenFOAM using segregated solver for first test case and coupled one for the second. For each test case, we compare the efficiency and robustness of each smoother by showing the residual convergence rate for each smoothers and calculating how much the residual is decreased at each time step, the number of iterations needed by each smoother to achieve this decrease. In addition we will present the CPU convergence time needed to achieve convergence. We focus on the pressure equation since it is of an elliptic type and is critical for the solution and convergence of whole system.

A. *Flow over a Stator Blade (Compressible-Segregated)*

Figure 6 shows the convergence rate residual with each smoother versus time. It can be noticed the high agreement between ILU (0) and AINV in the convergence rate and time/iterations needed. In the next graphs, the Residual of pressure before and after applying each smoother is presented. In Figure 7, it can be seen how the residual is consistently decreasing with time. After applying the ILU(0) smoother, the residual is locally decreasing with an average of 27 times the initial residual with standard deviation of 5. This shows how ILU (0) is stable. By comparing this residual decrease to that of AINV in Figure 8, the pressure residual is decreasing with a mean value of 25 times with standard deviation of 4. This shows how AINV is quite effective as ILU (0) in reaching convergence over time with nearly same robustness and stability.
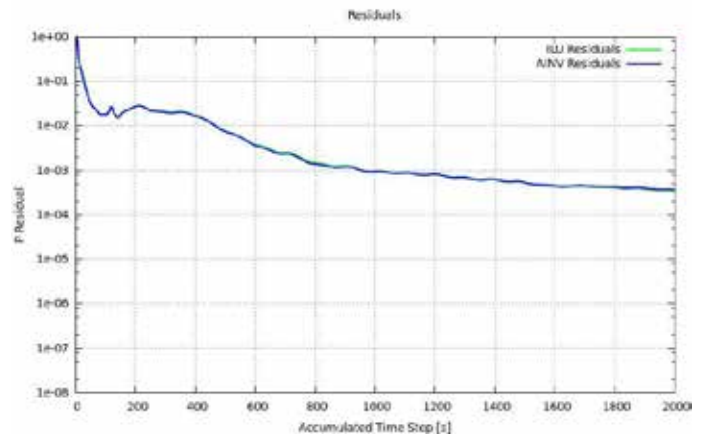
Figure 6. Residual vs Time for Smoothers
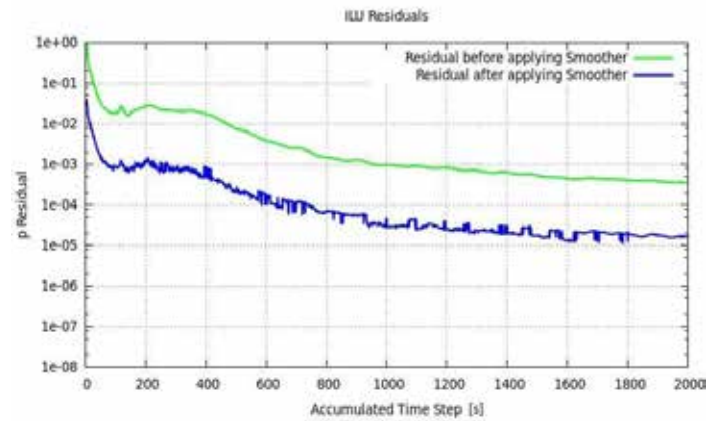
Figure 7. Initial & Final Residual using ILU(0)

AINV Residuals
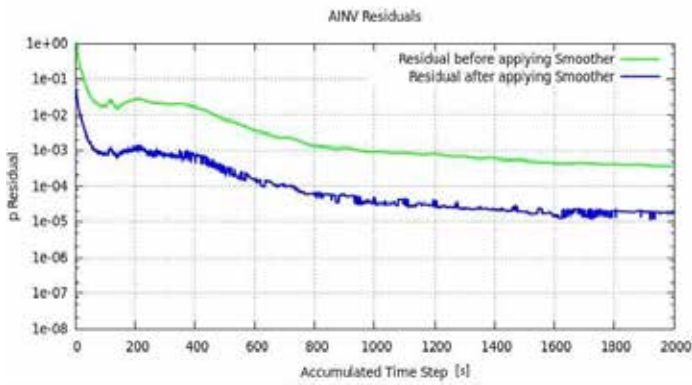
Figure 8. Initial & Final Residual using AINV

In addition to how much the residual is decreasing with time, it is also important to monitor the number of iterations needed by each smoother to reach the final residual. As shown in Figure 9, ILU (0) and AINV start with high number of iterations in each time step then decrease to almost reach a steady number of iterations. ILU(0) requires minimum 4 iterations and maximum of 10 with a mean value of 5 iterations in each time step to lower the residual by order of 27 as mentioned earlier, while the AINV smoother requires a little more to reach minimum of 5 and maximum of 14 iterations with mean value of 6 iterations. This clearly shows how robust these two smoothers are with performance of AINV almost reaching that of ILU (0).

Linear Solver Iterations
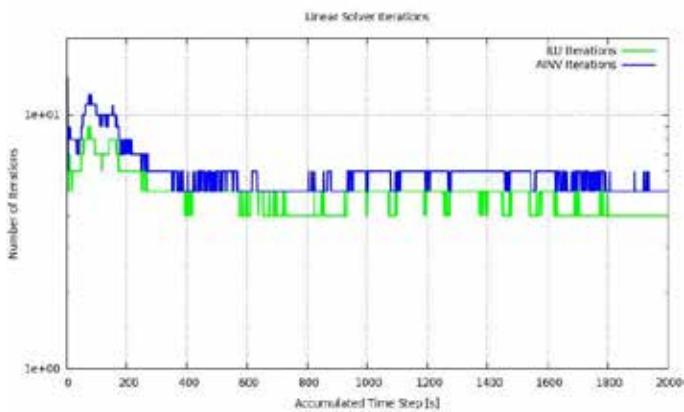
Figure 9. Number of Iterations vs Time Step of Smoothers

Table III summarizes the residual decrease after applying each smoother, number of iterations needed as well as CPU time. It is clearly shown how the performance of the two smoothers in this test case, which has high aspect ratio and causes an ill-conditioned matrix, is nearly identical with slight difference.

Table III
Smoothers Comparison

Smoother	Avg. of Res Decrease	Standard Dev. of Res Decrease	Avg. # of iterations	CPU Convergence Time (sec)
ILU (0)	27	5	5	216.96
AINV	25	4	6	238.03
Error (%)	7.4	-	-	9.71

B. *90° Pipe Bend (Incompressible-Coupled)*

Figure 10 shows the residual convergence rate of pressure versus time for both smoothers. It can be noticed how Block ILU (0) is consistently converging and decreasing the residual without any fluctuations; whereas, Block AINV causes the residual to decrease at first with some fluctuation then after some time steps the residual increases causing the solution to diverge.
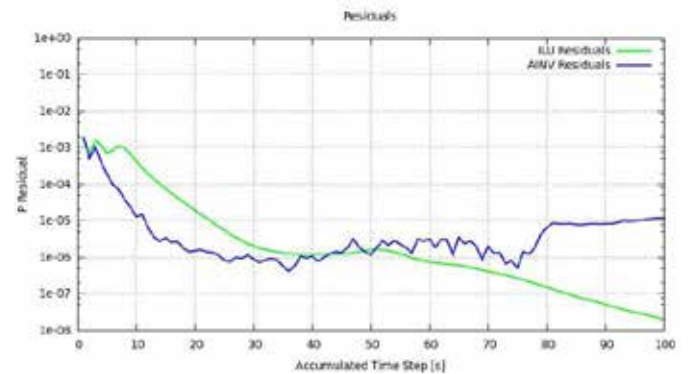
Residuals

Figure 10. Residual vs Time for Block Smoothers

By calculating the number of iterations each smoother required at each time step do decrease the residual by a certain tolerance, Figure 11 shows how the Block ILU (0) almost keep the same number of iterations in each time step with minimum of 3 and maximum of 5 iterations with mean value of 4 iterations, while Block AINV requires a lot more in each time step with minimum of 3 and maximum of 17 iterations with mean value of 7 iterations with a lot of fluctuations. This clearly reflects the divergence that occurred with Block AINV Smoother.

Residuals
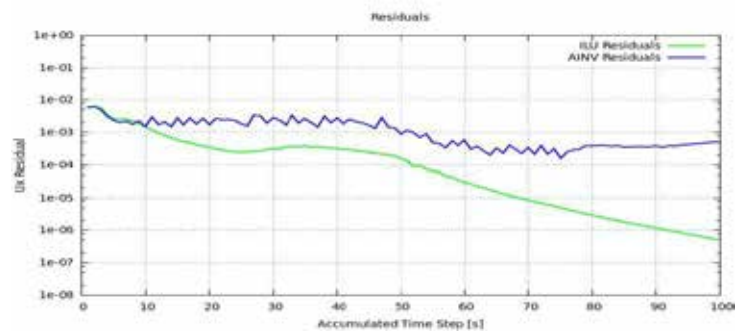
Figure 11. Number of Iterations vs Time Step of Smoothers

Table IV summarizes the number of iterations needed by each smoother as well as CPU time. It is clearly shown how the Block ILU(0) smoother in this test case outperforms Block AINV requiring less iterations and CPU time and being more robust, while Block AINV diverges requiring more iterations and more CPU time.

Table IV

Smoothers Comparison

Smoother	Avg. # of iterations	CPU Convergence Time (sec)
ILU (0)	4	2530
AINV	7	4500
Error (%)	-	77.86

VII. CONCLUSION

In this paper two algebraic Multigrid smoothers were presented and compared. AINV in scalar and block version was implemented in OpenFOAM and compared to already existing ILU (0) smoother.

The two smoothers were tested for two demanding test cases with turbulent flow having anisotropic mesh resulting in elements with very high aspect ratio and yielding a stiff system of equations.

Results clearly show that AINV has same performance as ILU (0) in segregated solver, whereas it does not perform well compared to ILU (0) in coupled solver. This may due to many inverse operations required by Block AINV. In addition no drop tolerance was used. The implementation of drop tolerance might result in a better smoother.

Improvements adopted for future work include drop tolerance implementation for AINV thus increasing its robustness.

ACKNOWLEDGMENT

Nomenclature

a coefficients in the discretized equation

b   general source term

p pressure

u,v,w velocity components in x-,y- and z- directions, receptively

Greek Symbols

ϕ general dependent variable

δ drop tolerance

Subscripts

C   refers to central/diagonal cell point

f   refers to a control volume face

F   refers to neighboring cell point

NB   refers to neighbors cells

REFERENCES

[1] M. Darwish et al., "A Coupled Finite Volume Solver for the Simulation of Disperse Multiphase Flows", presented at the Fifth European Conference on Computational Fluid Dynamics ECCOMAS CFD 2010, Lisbon, Portugal, 2010.

[2] U. Falk and M. Schafer, "A Fully Coupled Finite Volume Solver for the Solution of Incompressible Flows on Locally Refined Non-Matching Block-Structured Grids", presented at the Sixth International Conference on Adaptive Modeling and Simulation ADMOS, 2013.

[3] L. Mangnai et al., "Development of a Novel Fully Coupled Solver in OpenFOAM: Steady-State Incompressible Turbulent Flows", *Numerical Heat Transfer, Part B*, vol. 66, pp. 1-20, 2014.

[4] W. Hackbusch, *Multi-Grid Methods and Applications*, Berlin, Germany: Springer, 1985.

[5] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, New York: Springer, 1994.

[6] P. Fedorenko, "A relaxation method for solving elliptic difference equations", *USSR Computational Mathematics and Mathematical Physics 1*, pp. 1092–1096, 1962.

[7] F.V. Poussin, "An accelerated relaxation algorithm for iterative solution of elliptic equations", *SIAM Journal of Numerical Analysis 5*, pp. 340-351, 1968.

[8] A. Settari and K. Aziz, "A Generalization of the Additive Correction Methods for the Iterative Solution of Matrix Equations", *SIAM Journal of Numerical Analysis*, vol.10, no. 3, pp. 506–521, 1973

[9] A. Brandt, "Multi-Level Adaptive Solutions to Boundary-Value Problems", *Mathematics of Computations*, vol. 31, no. 138, pp. 333–390, 1977.

[10] A. Brandt and N. Dinar, "Multigrid Solutions to flow problems", *Numerical Methods for Partial Differential Equations*, Part S., pp., 43-147, Academic Press, New York, 1979.

[11] C.P. Thompson and P. Lezeau."Application of the Full Approximation Storage Method to the Numerical Simulation of Two-Dimensional Steady Incompressible Viscous Multiphase Flows", *International Journal for Numerical Methods in Fluids*, vol. 28, no. 8, pp. 1217-1239, 1998.

[12] M.C. Thompson and J.H. Ferziger, "An Adaptive Multigrid Technique for the Incompressible Navier-Stokes Equations", *Journal of Computational Physics*, vol., 82, pp. 94-121, 1989.

[13] P. Rubini et al., "Multigrid Acceleration of Three Dimensional Turbulent Variable Density Flows", *Numerical Heat transfer* ,part B, vol. 22, pp. 163-177, 1992.

[14] S. Zeng and P. Wesseling, "Multigrid Solution of the Incompressible Navier-Stokes in Generalized Coordinates", *SIAM Journal of Numerical Analysis*, vol. 31, p. 1764-1784, 1994

[15] P. Wesseling, *An Introduction to Multigrid Methods*, Chichester, UK: Wiley, 1992.

[16] Moukalled, F., Mangani, L., & Darwish, M. *The Finite Volume Method in Computational Fluid Dynamics*. Springer, 2015.

[17] R. Suda, *Large scale circuit analysis by preconditioned relaxation methods*, Keio University, Japan, 1994.

[18] A. C. Van Duin, "Scalable parallel preconditioning with the sparse approximate inverse of triangular matrices", *SIAM Journal on Matrix Analysis and Applications*, vol. 20, pp. 987-1006, 1999.

[19] M. Benzi *et al.,* "A sparse approximate inverse preconditioner for the conjugate gradient method", *SIAM Journal on Scientific Computing*, vol. 17, pp. 1135-1149, 1996.

[20] M. Benzi and M. Tuma, "A sparse approximate inverse preconditioner for nonsymmetric linear systems", *SIAM Journal on Scientific Computing*, vol. 19, no. 3, pp. 968-994, 1998.

[21] N. I. Buleev, "A numerical method for the solution of two-dimensional and three dimensional equations of diffusion", *Math. Sb, 51*, pp. 227-238, 1960.

[22] T. A. Oliphant, "An implicit numerical method for solving two-dimensional time-dependent diffusion problems", *Quarterly of Applied Mathematics*, vol. 19, pp. 221-229, 1961.

[23] M. Benzi *et al*., "Stabilized and block approximate inverse preconditioners for problems in solid and structural mechanics", *Computer Methods in Applied Mechanics and Engineering*, vol. 190, pp. 6533-6554, 2001.

[24] R. Bridson and W.P. Tang, "Refining an approximate inverse", *J. Comput. Appl. Math.*, vol. 123, pp. 293-306, 2000.

[25] T. A. Oliphant, "An extrapolation process for solving linear systems", *Quarterly of Applied Mathematics*, vol. 20, pp. 257-267, 1962.

[26] D. J. Evans, "The Use of Preconditioning in Iterative Methods for Solving Linear Equations with Symmetric Positive Definite Matrices", *Journal of the Institute of Mathematics and its Applications*, vol. 4, pp. 295-314, 1968.

[27] H. Stone, *Iterative Solution of Implicit Approximations of Multidimensional* Partial Differential. Equations", *SIAM Journal on Numerical Analysis*, vol. 5, pp. 530-558, 1968.

[28] T. Dupont et al., "An approximate factorization procedure for solving self-adjoint elliptic difference equations", *SIAM Journal on Numerical Analysis*, vol. 5, pp. 559-573, 1968.

[29] J. A. Meijerink and H. A. van der Vorst, "An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix", *Mathematics of Computation*, vol. 31, pp. 148-162, 1977.

[30] H. A. van der Vorst, "Iterative solution methods for certain sparse linear systems with a nonsymmetric matrix arising from PDE Problems", *Journal of Computational Physics*, vol. 44, pp. 1-19, 1981.

[31] R. Beauwens and L. Quenon, "Existence criteria for partial matrix factorizations in iterative methods", *SIAM Journal on Numerical Analysis*, vol. 13, pp. 615-643, 1976.

[32] L.D. Hylton et al., *Analytical and experimental evaluation of the heat transfer distribution over the surfaces of turbine vanes*, Tech Rep. 182133, NASA, Lewis Research Center, Cleveland Ohio.

[33] L. Mangnai et al., "Development of a Pressure-Based Coupled CFD Solver for Turbulent and Compressible Flows in Turbomachinery Applications", in *Proc. Of ASME TURBO EXPO 2014: Power for Land, Sea & Air*, Dusseldorf, Germany, June 16-20, 2014.

[34] S.A. Berger et al., "Flow in Curved Pipes", *Annual Review of Fluid Mechanics*, vol. 15, pp. 461-512, 1983.

[35] F.T. Smith, "Fluid Flow into a Curved Pipe", in*: Proceedings of the Royal Society A,* vol. 351, pp. 71-87.

[36] F. Rutten et al., "Large-eddy simulation of low frequency oscillations of the Dean vortices in turbulent pipe bend flows", *Physics of Fluids*, vol. 17, 2005.

[37] R. Rohrig et al., "Comparative computational study of turbulent flow in a 90° pipe elbow", *International Journal of Heat and Fluid Flow,* vol. 55, pp. 120-131, 2015.

Integrated Sustainable Power Bank

Rema Daher
Roy Deeb
Walaa El Safadi
Yazan Fanous
Department of Mechanical Engineering
American University of Beirut
Bliss Street, Beirut, Lebanon
Rgd05@mail.aub.edu
Ymf02@mail.aub.edu
Rjd06@mail.aub.edu
Wge02@mail.aub.edu

Abstract- **When the United Nations set a decision to form a "Decade of Sustainable Energy (2014-2024)", most technologies contemplated having sustainability factors in their respective final products. Some considered recycling, others adopted renewable energy systems. In this paper, a synergy of multiple sustainability factors is proposed such that all gymnasiums with green certifications can have a net-zero plug load. The Integrated Sustainable Power Bank (ISPB) is a machine combines recycling, solar and human generated electricity and physical sustainability. The life cycle analysis of the ISPB proved the optimization of the energy sources, possessing the least amount of energy possible used by its components, and an efficient long run economic advantage.**

I. INTRODUCTION

The energy crisis with no debate is a worldwide issue that is specifically exemplified in Lebanon. Lebanese citizens suffer daily from the pollution produced due to the energy sector, as well as the scarcity of energy availability throughout the whole country causing frequent and long power cuts. The current energy distribution system is highly inefficient putting multiple constraints on the economic growth of the country. To repair the damage, new and more sustainable energy sources should be discovered and implemented. These include alternative solutions such as wind energy, solar energy, hydropower, or even human mechanical energy. In this machine, the aim is to build an Integrated Sustainable Power Bank (ISPB) which constitutes an energy storage bank from both human mechanical and solar power.

The energy generation market is monopolized and owned by the government with one provider, Electricité du Liban (EDL). However, EDL supplied the energy needs for only 60% of the population in 2008; this number is estimated to have decreased drastically in the past 7 years. New additions to increase their capacity stopped after the implementation of two combined cycle plants in the 1990's [5]. Due to this deficient capacity,

most regions experience multiple lengthy power cuts during the day. Thus a market of off-grid network emerged and supplied around 33 to 38% of the demand in 2008 and this number was estimated to go up to 60% in 2015 [5].

Energy Statistics conducted by the International Energy Agency (IEA) indicate an astonishing 74% of greenhouse gas (GHG) emissions in Lebanon associated with the energy sector. A major part is due to the diesel generators populating major cities; the average residential area in Beirut is subjected to 2.5 times more carcinogenic pollutants when the generators are in use. After consistent exposure, they cause cancer, genetic mutations and influence the physiological development of the residents. Off grid private generator market is not regulated and the diesel sources are of very poor quality. This is not only costing citizens their health but it is also more expensive to pay for operating diesel generators than to get the supply from EDL [6].

The effective solution requires taking sustainable measures in houses and offices and implementing secondary power sources that use renewable energy. This investment will help reduce emissions by decreasing demand of unsustainable sources including diesel generators. Constraints on economic growth would decline leaving more liquidity for EDL to invest in upgrading to a more efficient power distribution system. Integrated Sustainable Power Bank (ISPB) aims to satisfy the above. ISPB is to be employed in gyms where mechanical energy from gym bikes is utilized. It also integrates recycling bottle caps that are made of higher grade plastics than the exterior of the bottle. REVA has implemented several reverse vending machines for recycling plastic bottles. However, the caps have no proper recycling system inside AUB facilities.

II. Specifications and Deliverables

A. Specifications

ISPB incorporates two renewable resources of power to provide an alternative source of energy. Solar energy and mechanically human produced energy is utilized to decrease the plug loads in the location in which the machine is placed. In addition to the aforementioned, elements of recycling and physical sustainability are added to increase the green value of this machine. The purpose is to generate a green energy source that is tangible to users and is a step towards shifting the energy dependence towards carbon free sources.

The first energy source is produced by cycling on a gym bike. The ISPB incorporates two bikes that charge its power bank. The power generated by the end user is then used to charge electronic gadgets. Since human power would not be enough to provide energy for multiple full phone charges, a solar panel is added to the design. The machine then allows the users to charge their electronic gadgets after recycling a set amount of bottle caps. This bottle cap collection mechanism is highly sought for as an alternative to the usual processing of bottle caps. Bottles going into recycling are managed in machines that separate them from the bottle caps and the end result is the disposal of the caps. The plastic of the bottles is then recycled and reused [1]. Therefore, a simple and cheap machine that can collect the bottle caps along with a reward for the final user will increase bottle cap recycling in a significant manner.

B. Deliverables

Students are required to use knowledge gained through their training in mechanical energy and integrate different tracks to develop a machine and build a prototype of the final design. The project work is inspired by the "Integrated Process" suggested by LEED certifications entailing an iterative design and constant charrettes amongst the team members. The project work is listed as follows:

1) Energy modeling of the solar panel to study the influence of different environmental factors. Those include humidity, temperature, shading effects and angle of tilt.

2) Stress analysis to help with design decisions and maintenance predictions based on the resulted fatigue.

3) A study of the environmental factors to optimize the efficiency of power generation.

4) Material Selection based on life cycle analysis (LCA). LCA entails analyzing the material endurance starting at the manufacturing level and throughout the life of the machine until it is destroyed, recycled or reused.

5) Business Plan

C. Business Plan

The cost of production per machine for the ISPB will be close to $1070. Once a production line is in place, the initial cost of intermediate goods will be much less than the cost proposed by the initial estimation; the cost would decline to about $900 to $950. Assuming the Gross Marginal Profit percentage is 20%, the machine will cost $1284. Payback period for the end user will be 5.499-6 years taking into consideration the electricity tariff set by the Lebanese Electricity Company (EDL).

III. Technical and Non-technical Constraints

A. Technical Constraints

Technical limitations facing the design of the ISPB are listed as follows:

1) Efficiency limitations: the relatively low energy outputs of bikes and solar panels are limiting factors for the power generation expected from the machine.

2) Size: the power bank should be sized appropriately to satisfy both indoor and outdoor applications.

3) Power Consumption: Power generated by the machine shall compensate more than at least 6 full charges for regular phones

4) Heat Generation: The heat generated by the machine and its user should not exceed the cooling loads of the existing air conditioning.

5) Noise Reduction: the noise level generated by the machine should be acceptable under human comfort standards for indoor environments.

B. Non-technical Constraints

1) Environmental: this work aims towards a carbon free energy production process to help decrease GHG emissions.

2) Economical: the power bank should be sized appropriately to satisfy both indoor and outdoor applications.

3) Social: the machine should be user friendly and accessible to all classes of society.

4) Heat Generation: The heat generated by the machine and its user should not exceed the cooling loads of the existing air conditioning.

5) Noise Reduction: the noise level generated by the machine should be acceptable under standards of human comfort in indoor environments.

IV. Applicable Standards

ASHRAE 55: Human thermal comfort standards for indoor air quality acceptable by 80% of the occupants. This guideline will be used for the analysis of heat generated by the machine and its effects on the cooling load of the space.

American Society of Testing Materials (ASTM): These standards will be used for analysis including emissivity, irradiance, solar reflectance and recyclability.

ENERGY STAR: Hardware parts used should meet the requirement of using 30% to 75% less electricity than standard equipment.

Forest Stewardship Council (FSC): The wood used is FSC certified ensuring a trusted source that use recycled wood.

V. Methodology

The research, design and building process of the machine is divided into several subprojects. These include an energy study of the ISPB, bottle cap recognition system, solar panel analysis, material selection including all electrical parts and a mechanical design analysis for a mechanism used to change the tilt angle of the panel. A bidding process is then implemented to agree on design decisions and to integrate the work of all subprojects.

A. Bottle cap recognition system

Bottle caps are being disposed through the conventional recycling process of bottles. Instead they can be recycled and utilized as material for other products such as cars, garden rakes, storage containers, reusable shopping bags, yarn, ropes, brooms and even more bottle caps [7]. Coca Cola has attempted this project with its "hello happiness" machine that collects bottle caps in return for time on a phone booth. The technology used in this machine was, however, not published. Several designs were therefore studied and are listed as follows:

1) Metal detection: Mini Gadgets MD-RING Metal Detector Ring is a low cost sensor. This device is able to detect metals and could be used as the main or as a secondary recognition mechanism [2].

2) Plastic detection by pattern recognition: A Light Emitting Diode (LED) light is placed inside the bottle cap processing chamber of the ISPB to illuminate the space. Plastic type identification number is usually grooved into the caps inside a triangular symbol. Number 2 classifies the only kind of recyclable caps [3]. A high definition camera is used to capture a screenshot which then gets processed using pattern recognition software.

3) Plastic detection by shape recognition: a high definition camera takes two pictures of the inserted bottle cap from the front and the side view. These pictures are processed using software to detect the shape of the circular front and the rectangular side of the bottle cap.

4) Hardness/Friction constant test: a hardness detection sensor is used to identify the friction constant of the inserted bottle cap. This value is then compared with a database of different material hardness and categorized as recyclable plastic according to a set range of desired values. "Digital Shore Tire Durometer Plastic Tester Meter 0~100HC" is a high cost device that measures the hardness of plastics [4].

The pattern recognition system is the most economical and uses minimum hardware. After processing and recognition, a LED light turns on signifying the acceptance of the bottle cap. It then passes through a purely mechanical system constituting of a solenoid valve. The solenoid valve under a command from the recognition software sends the bottle cap down a pathway that leads to a collection tank. An ultrasonic HC-SR04 range detection sensor is mounted on the inner side of the tank to measure the level of bottle caps reached. A signal is sent to turn on a LED indicator to notify when the tank is full and ready for pick up. The system gives the option of using a power socket when a certain amount of bottle caps is inserted or when a certain amount of calories is burnt on the bikes. The calories counter uses an optical encoder that measures the rotational speed of the bike and accordingly an estimated value of the calories burnt is calculated after a set amount of time.

B. Solar Panel Efficiency

The main parameter determining the overall performance of the solar photovoltaic (PV) panel is the efficiency. The solar panel efficiency is generally expressed in the following equation:

$$\eta = P_m / (G \times A). \tag{1}$$

η: Efficiency of the solar panel
P_m: Maximum power output from the panel
G: Irradiance of the surface of the panel
A: Surface area of the panel

The maximum power of the panel is strictly affected by the change in temperature, which will be a major factor of significance in the design of the ISPB.

$$P_{m, variable} = (1/P_m) \times (dPm /dT). \tag{2}$$

T: operating temperature

(2) can be rewritten as

$$P_{m,variable} = (1/V_{oc}) \times (dV_{oc}/dT) + (1/FF) \times (dFF/dT) + (1/I_{sc}) \times (dI_{sc}/dT)$$

V_{oc}: Open circuit voltage
I_{sc}: Short circuit current
FF: Fill Factor, which is the useful energy that can be gathered from the maximum power output

Variables FF, I_{sc} and V_{oc} are all temperature dependent functions that affect the maximum power generated by the solar panel. All three parameters are directly proportional to temperature; however, a decrease in temperature mostly affects FF and V_{oc} only. The performance of the solar panel of the ISPB is evaluated under different combinations of environmental conditions and geometries. The environmental conditions studied include the following:

1) Location: indoors versus outdoors installations
2) Seasonal changes
3) Presence and absence of Air Conditioning
4) Shading effects

Different design geometries were analyzed in reference to figures 3, 4 and 5 in the appendix as follows:

1) Solar panel support as a triangular solid.

2) Solar panel support as a hollow triangular solid with both ends exposed to the surrounding atmosphere such that an air flow is present under the solar panel providing a cooling effect. This geometry is preferable for indoor applications.

3) Solar panel support as a hollow triangular solid with both ends closed. This design prevents rain water from entering the system interior and is preferable for outdoor applications.

"On average (as a general "rule of thumb") modern photovoltaic (PV) solar panels will produce 8-10 watts per square foot of solar panel area. For example, a roof area of 20 feet by 10 feet is 200 square-feet (20 ft. x 10 ft.). This would produce, roughly, 9 watts per sq.-ft., or 200 sq.-ft. x 9 watts/sq.-ft. = 1,800 watts (1.8 kW) of electric power. One kilowatt-hour (1 kWh) means an energy source supplies 1,000 watts (1 kW) of energy for one hour. Generally, a solar energy system will provide output for about 5 hours per day. So, if you have a 1.8 kW system size and it produces for 5 hours a day, 365 days a year: This solar energy system will produce 3,285 kWh in a year (1.8 kW x 5 hours x 365 days)." [12]
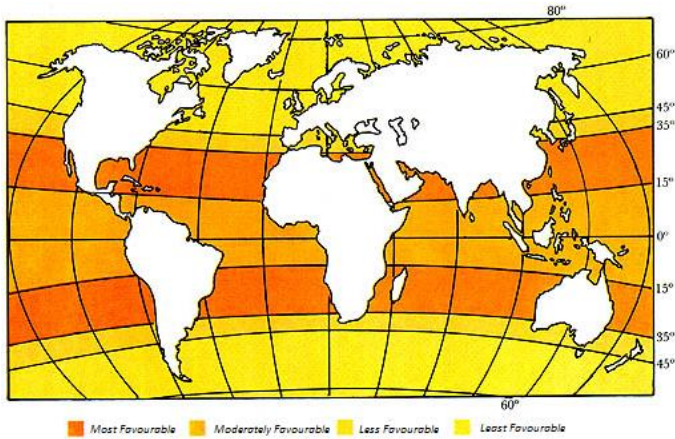
Figure 1. Solar energy intensity based on geographical location

Lebanon is at the border of the most favorable belt of angles having a latitude angle φ of 33.2° based on figure 1. The most favorable belt represents the regions in which the geographical location of the country has the best conditions for setting up solar panels, where more than 90% of the solar radiation comes to the solar panels are direct radiation, simply because the regions on this belt have fewer clouds than regions in other belts. In fact, these regions experience over 3000 hours of sunshine per year [8].

If the incident solar radiation $S_{incident}$ perpendicular to the sun is exactly the radiation that reaches the solar module (panel) S_{module}, the maximum power possible is generated in the solar panel. The following equation relates S_{module} with $S_{incident}$ based on the elevation angle of the sun α and the angle of tilt of the module from the horizontal β.

$$S_{module} = S_{incident} \times sin\ (\alpha + \beta). \quad (4)$$

The energy received by the panel is optimized when $\alpha+\beta=90°$. In other words, tilting the module in an angle from the vertical that matches α is the maximum possible radiation received by the module. To determine α, the following equation is used:

$$\alpha = 90 - \emptyset + \delta. \quad (5)$$

where the Earth declination angle δ is represented by:

$$\delta = 23.45 \times sin\ (\ 360/365 \times (284+d)\ ). \quad (6)$$

The tilt angle measured from axis of rotation of Earth changes on a daily basis. The day number d is thus essential in this equation as it will eventually alter the amount of radiation reaching the module [9]. Changing the angle of tilt constantly is a time consuming and effort intensive process. Most rooftop solar panels have a fixed tilt angle hence losing system efficiency due to the decreased radiation meeting the module. The ISPB compensates using a rotary mechanism that allows variation of the angle of tilt of the module. The mechanism is manually handled, and the angle at which the module is set varies with every season of the year. Winter starts on the 21st of December and ends on the 21st of March. In other words, it starts on the 355th day of the year and ends on the 81st. The latitude angle φ is 33.2° as mentioned earlier. (5) and (6) are solved using these values resulting in a declination angle range of [-23.4°, 0°). Therefore, the sun angle α has a range of [33.35°, 56.8°]. The Solar Electricity Handbook 2015 Edition states that the best tilt angle for winter is 32° (close to 31.55°) from the vertical [10]. The angle of tilt of the module β is then 58°. Other calculations are made by means of the same logic and the following table is constructed that summarizes the setup the panel.

TABLE I
DECLINATION, TILT AND SUN ELEVATION ANGLES

Season	Day Number	Declination	Sun Angle α	Angle of Tilt β
Winter	355-81	[-23.4, 0)	[33.35, 56.8)	[58,34)
Spring	81-173	[0, 23.45)	[56.8, 80.25)	[34,10)
Summer	173-265	[23.45, -0.61)	[80.25, 56.19)	[10,34)
Autumn	265-355	[-0.61, -23.4)	[56.19, 33.35)	[34,58)

Figures 3, 4 and 5 in the appendix show the behavior of the direct solar radiation H on the panel based on the three different angles of tilt of the panel 58°, 34° and 10° consecutively given the latitude angle of Lebanon. The energy output generated by the solar panel is given by the following equation [11]:

$$E = E_p \times H \times PR \quad (7)$$

E = Energy in kWh

E_p = Energy peak in kWp

H = Annual average solar radiation on tilted panels (shadings not included); refer to figures 3, 4 and 5 in appendix.

PR = Performance ratio, coefficient for losses (range between 0.5 and 0.9, default value = 0.75)

The potential solar panel chosen for the ISPB has dimensions 1×1.6 meters, maximum power of 270 Wpeak. Taking the lowest *H*, the machine will solely give us energy output from the solar module equal to 1.215 kWh per day. This will be equivalent to 121.5 Watts per hour (given that the time span of the sun, from sunrise to sunset, is 10 hours), which is enough to charge 7 full charges for a single phone with power rating of 12 Wh.

C. Human Powered Energy Generation

A stationary bicycle generator is studied by the Canadian Center of Science and Education, where the motion of the front wheel is connected to a flywheel and rotates coils of wires in a magnetic field. The DC output current could be either channeled to a battery bank or converted to AC voltage for purpose of operating a higher range of electrical appliances. Tests done on this schematic presented a work output of 75W for an average one-hour workout. On a speed of 25kph output can reach up to 200W for a limited time. Devices that can use this energy draw little current such as radios, phones, laptops and high efficiency fluorescent lightings. A 25W fluorescent bulb can light up to 8 hours on 200W power. On this basis it is deduced that the most efficient use of this technology is in fitness clubs and gyms where more than one exercise machine is used and their power output could be stored in one common power bank.
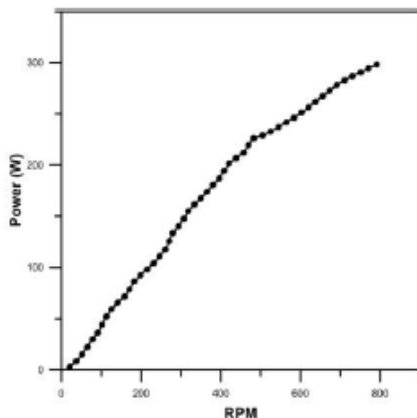
Figure 2. Power vs. RPM of MNS Power Generator

The efficiency generation of the human powered bike generators will be studied with reference to figure 2 above. The MNS Power Dynamo Generator has a maximum output of 300W power and can be mounted to any rotating shaft. The generator costs 300,000 LBP and its average Amp Hours during one hour of use is 0.1kWh. A typical operating time for each bike is around 5 hours per day and according to the cost of electricity in Lebanon, the two gym bikes will save around $54.5.

There seems to be no short term benefit for energy produced mechanically by human power, but their energy generation efficiency can go up to 94%. Thus the integration of such generators with other sources of energy is beneficial and helps decrease the plug load in the space where it is installed.

VI. Budget and Payback Period

A. Budget Listing

1) Solar panel+ Battery bank + invertor: $800
2) Dynamo: $200
3) Sockets: $2 per socket.
4) Recycling tank: could be free if made in manufacturing lab using scrap wood.
5) Outer cover of power bank: (cast iron) 0.15 cents
6) Valve opening mechanism: $10.
7) Web Camera: $6
8) LED light: $1
9) Ultrasonic sensor hc-sr04: $1
10) Interface data acquisition board (Arduino or raspberry pie): $40.

Total=$1070.15

B. Payback Period

- Dollars saved using solar power: ((0.27*0.75*11.5*92*320)/1500) + 2*((0.27*0.75*9*91*320)/1500) + ((0.27*0.75*6*91*320)/1500)
- Dollars saved using dynamo= 54.5
- Payback period=1070/ (solar + bike)

Peak rate tariff, according to EDL, is 320 LL/kWh. According to the previous calculations, the module of the ISPB will save around $140.05 in one year, while the two gym bikes will save around $54.5. Having a cost of $1070.15, we can conclude that the payback period of the machine is 5.50 years.[1]

VII. CONCLUSION

The synergy of several sources of energy in the ISPB, along with the benefits of recycling proved to be a beneficial invention that should be considered in the market. LEED certified gyms are highly encouraged to have views to the outdoors to increase productivity of

---

[1] Analysis done neglects costs of transportation and operation

the occupants. The presence of the ISBP in such buildings is a significant improvement to the sustainable performance of the gym in which it exists. The decrease in plug load, the possibility of recycling, and most importantly the brief and rapid payback period are impressive assets to any gym seeking green certifications.

ACKNOWLEDGMENT

NOMENCLATURE

η	Efficiency of the solar panel
P_m	Maximum power output from the panel
G	Irradiance of the surface of the panel
A	Surface area of the panel
T	Operating temperature
V_{oc}	Open circuit voltage
I_{sc}	Short circuit current
FF	Fill Factor, which is the useful energy that can be gathered from the maximum power output
$S_{incident}$	Incident solar radiation perpendicular to the sun
S_{module}	Solar radiation reaching the solar module
α	Elevation angle of the sun
β	Angle of tilt of the module from the horizontal
δ	Earth declination angle
d	The day number
$\emptyset$	Latitude angle
E	Energy in kWh
E_p	Energy peak in kWp
H	Annual average solar radiation on tilted panels (shadings not included)
PR	Performance ratio, coefficient for losses

REFERENCES

[1] "PET Recycling Machine that solved the problem of removing pet bottle caps, rings, labels almost 100%", YouTube, 2014. [Online]. Available: https://goo.gl/tR1BG3 [Accessed: 06- Nov- 2015].

[2] "Mini Gadgets MD-RING Metal Detector Ring", Google.com, 2011. [Online]. Available: https://goo.gl/vTxcuq. [Accessed: 14- Nov- 2015].

[3] Plastic Resin Codes, 1st ed. 2015.

[4] "Digital Shore Tire Durometer C Rubber Plastic Hardness Tester Meter 0~100HC - Google Search", Goo.gl, 2011. [Online]. Available: https://goo.gl/ZBw5Vg. [Accessed: 07- Nov- 2015].

[5] The World Bank Group ,"Republic of Lebanon, Electricity Sector - Public Expenditure Review", Sustainable Development Department, Middle East and North Africa Region, Report No. 41421-LB, 2008. [Accessed:25-Nov-2015]

[6] Effect of distributed electric power generation on household exposure to airborne carcinogen, AUB, Alan Shehadeh, 2008. [Accessed: 26-Nov-2015]

[7] recycleyourplastics.org, 'what plastics can become', [online]. Available at: https://www.recycleyourplastics.org/consumers/kids-recycling/plastics-can-become/

[8] Pveducation.org, 'Solar Radiation on a Tilted Surface | PVEducation', 2015. [Online]. Available: http://www.pveducation.org/pvcdrom/properties-of-sunlight/solar-radiation-on-tilted-surface. [Accessed:20-Nov-2015].

[9] Pveducation.org, 'Solar Radiation on a Tilted Surface | PVEducation', 2015. [Online]. Available: http://www.pveducation.org/pvcdrom/properties-of-sunlight/solar-radiation-on-tilted-surface. [Accessed: 08- Nov-2015].

[10] Pveducation.org, 'Solar Radiation on a Tilted Surface | PVEducation', 2015. [Online]. Available: http://www.pveducation.org/pvcdrom/properties-of-sunlight/solar-radiation-on-tilted-surface. [Accessed: 20- Nov-2015].

[11] WholesaleSolar.com, 'SolarWorld SW270 Silver Mono Protect Solar Panel - Wholesale Solar', 2015. [Online]. Available: http://www.wholesalesolar.com/1922268/solarworld/solar-panels/solarworld-sw270-silver-mono-protect-solar-panel. [Accessed: 01- Dec- 2015].

[12] Solar Power Returns|Solar Calculations| benefits from installing solar panels", Solar-estimate.org, 2016. [Online]. Available: http://www.solar-estimate.org/?page=solar-calculations. [Accessed: 20- Mar- 2016].
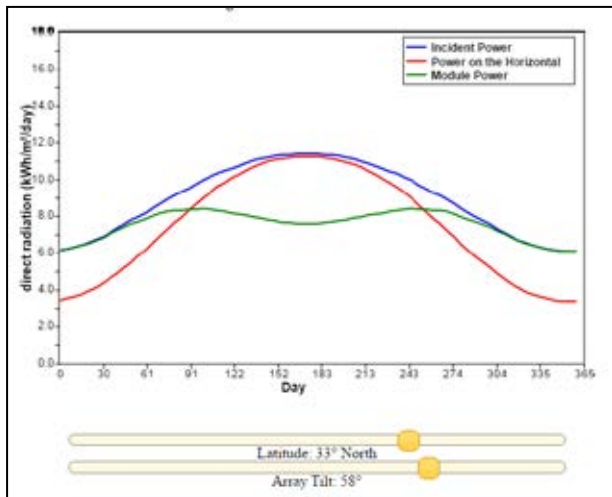
Figure 3. Module power in winter with β = 58º
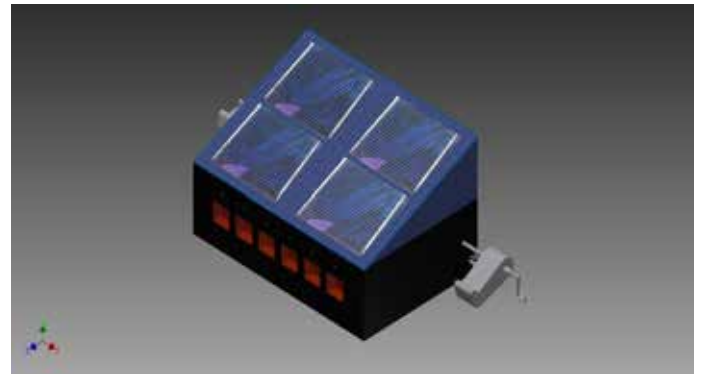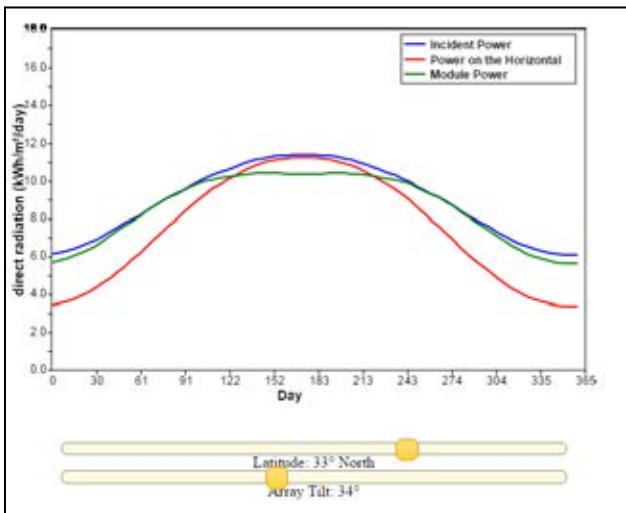

Figure 6. CAD design of the ISPB (top view)
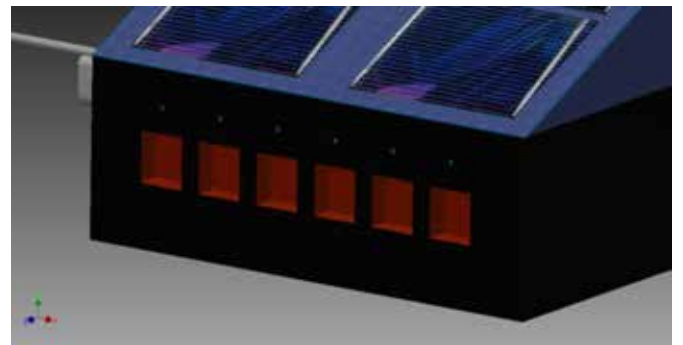

Figure 4. Module power in spring and autumn with β=34º
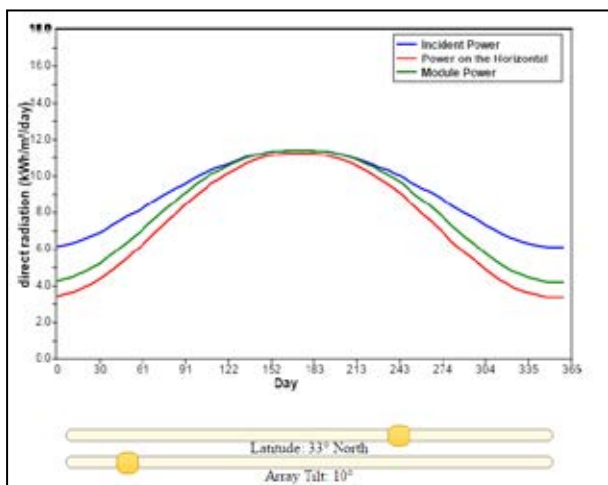

Figure 7. CAD design of the ISPB (frontal view)
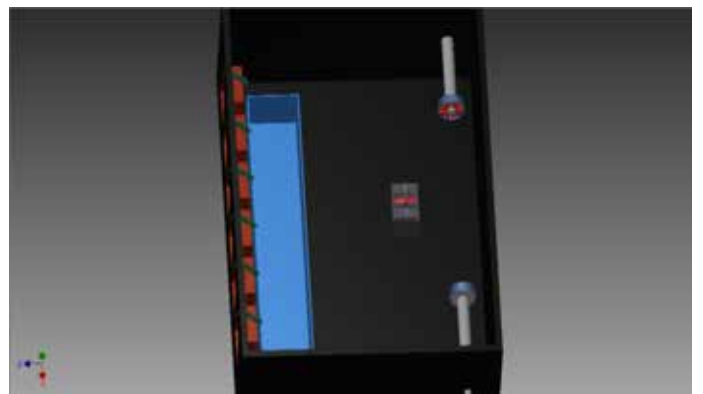

Figure 5. Module power in the summer with β=10º

Figure 8. CAD design of the ISPB (components present inside the machine)

Inverse Lagrangian Particle Tracking in Stochastic Flow Fields

Ali Ayoub, Ali Saab, Nabil Ramlawi, and Wael Hajj Ali
Department of Mechanical Engineering
American University of Beirut
Beirut, Lebanon
Emails: asa67@aub.edu.lb, aks14@aub.edu.lb, nsr07@aub.edu.lb, wrh01@aub.edu.lb

Abstract— In this work, efficient methods that can be used to predict the Lagrangian motion of passive tracers in stochastic flow fields were developed. These methods were further used for the inverse Lagrangian prediction problem in which estimation of the trajectories within a turbulent oceanic flow, given final positions and description of the flow field, is addressed. Of particular interest is to define probabilistic inverse maps that can lead to the estimation of the source location which is critical in a variety of environmental and safety-related applications, such as oil spill sources or other search operations. However, in such applications, one is challenged by huge amounts of data due to the sheer size of the simulations, as well as the need to assess the impact of model and data uncertainties. This work presents an approach using probability-weighted piecewise particle trajectories in which the domain is binned and the tracing process is splitted up over the individual assimilation cycles. A forward model was first developed in which a passive tracer is seeded at some locations in the Red Sea and the model was marched forward in time in order to predict the trajectory and final position of the tracer in a stochastic flow field. This model was further extended to account for cases where several tracers are emitted at the same location with different emission times, as well as tracers emitted at different locations. Finally, a backward in time trajectory model is presented in which probabilistic inverse maps leading to estimations of the source location are defined.

I. INTRODUCTION

Following the tragic Deepwater Horizon oil rig explosion on April 20, 2010, oil gushed from the bottom of the Gulf of Mexico at a depth of 1500 m for 87 days before it was capped. The gushing Mississippi Canyon (MC) 252 oil lead to a buoyant plume spread over a surface area of about 17,725 km^2 [1]. In such a case as well as in many source location search operations, the temporal evolution of the spatial extent of the tracer, and a measure of the error of that estimate, will be needed for the natural resource damage assessment.

This objective may be addressed using model-based estimates or satellite imaging. However, in the case of oil source location estimation, these methods lead to considerable differences and they both exhibit significant variation in the temporal sequence of oil location maps. This is mainly due to the fact that in-situ observations are limited by large search domain and observing error, while high resolution images of the upper few millimeters of the ocean surface can be obscured by clouds [2]. As a result, efficient and accurate methods to determine the source, given the final location, are required.

One approach can be implemented using the dynamics and the velocity fields in the current search area. This approach would try to estimate the source by relying on the given velocity fields. However, forecasts of currents and surface flows are themselves sources of uncertainties due to the inaccuracies in the measurement as well as forecast models parameterisations. Consequently, tracking oil in such a stochastic flow field becomes difficult. Many tools from uncertainty quantification and data assimilation can be used to address these needs.

In this work, we tackle this problem by working towards defining probabilistic inverse maps for the case of a passive tracer. The structure of this paper is as follows: in section II we define what a passive tracer is. Section III gives a brief description of the velocity field data set used throughout this work. In section IV, we present the forward problem in which the final position and trajectory are seeked given initial position and velocity field. Section V introduces the techniques that were used to reduce the computational cost associated with the sheer amount of data. In section VI, we provide the methodology followed to address the backward problem. Section VII presents the results obtained for both the forward and backward problems. Finally, in section VIII, we provide concluding remarks and in section IX we discuss possible extensions of the presented work.

II. PASSIVE TRACER REPRESENTATION

A passive tracer representation of particles involves eliminating any interaction between the tracer and its surrounding and results in the simple case of a tracer being driven by its surrounding flow with no inertia. It is of particular importance since it allows extracting information about the deep circulation which is not readily available from examination of the velocity or density fields. As a result, we will use passive tracer representation throughout all our computational simulations.

III. VELOCITY FIELD DESCRIPTION

Throughout our work, we will be only considering passive tracers in the Red Sea. The flow field data set used in this work comes from a simulation for MITgcm (MIT general circulation model) with 10 assimilation cycles over a one month period

(June 2011) and a spatial resolution of around 4 km.
Over each assimilation cycle, 50 ensemble members are re-sampled, i.e. at each time step, there are 50 velocity realizations at a certain point in the domain. As discussed before, the spread in these realizations comes from the uncertainties associated with such models as well as errors in data used for assimilation. As a result, the given velocity field becomes a stochastic flow field.

IV. FORWARD PROBLEM

In this section, we will start first by formulating our forward problem. Then, the methodology followed throughout our work will be presented along with the computational cost associated with it.

A. Formulation

The forward problem can be formulated as follows:
Given the source location, emission time, and the velocity field, it is required to find the final location of the tracer.

B. Methodology

Since we are dealing with a passive tracer, the velocity of the particle will be that of the surrounding flow. The problem will be initialized by specifying a seed at a certain position within the domain. Then, for each member of the velocity field, we will be marching in time using 4th order Runge-Kutta (RK) scheme interpolating in time and space to the specified location within the grid. In other words, since we're using a Lagrangian transport advection scheme, marching forward in time requires solving equation (1) below at each time step as shown.

$$\frac{d\boldsymbol{r}}{dt} = \boldsymbol{u}(\boldsymbol{r}, t) \tag{1}$$

where

$$\boldsymbol{u}(\boldsymbol{r}, t) = \boldsymbol{u}_{fluid}(\boldsymbol{r}, t)$$

is a 3D velocity field. Hence:

$$\boldsymbol{r}(t_2) - \boldsymbol{r}(t_1) = \int_{t_1}^{t_2} \boldsymbol{u}(\boldsymbol{r}, t) \, dt \tag{2}$$

And the integral is carried out using 4th order RK. The time step used in our simulations was 3 days over one month period.

C. Computational Cost

The methodology presented before needs to be applied for every member of the velocity field. As a result, for an ensemble with n members and m assimilation cycles, the number of all possible combinations is n^m. The Red Sea data used throughout this work has 50 members with 10 assimilation cycles. This results in 50^{10} combinations which leads to a high computational cost [6]. The need to develop efficient methods is hence evident. These are discussed in the following section.

V. PROBABILITY MAP COMPUTATION

Instead of tracing quadrillions of individual particles, and since a single a possible path is not of particular interest, one approach could be implemented using probability-weighted bins by which a probability is assigned to each bin according to the number of particles that reached it from the different ensemble runs. In the next time step, only a single particle, weighted by the corresponding bin probability, is traced. Although individual paths are lost during the process, this approach- presented by Hoellt et al in [6]- leads to a probability map that captures the uncertainties associated with the velocity field.

Specifically, computing the pathline segments is done as follows:
Given the flow field data set (figure 1(a)), a particle is first injected at the user-defined seed point (the blue dot in figure 1(b)), then each member of the flow field is used to march the particle forward for one time step (figure 1(c),(d)&(e)). Then, the domain is binned (figure 1(f)) and the number of particles within a bin is used to associate a weight with the specified bin (figure 1(g)). The last segment of a pathline of one cycle does not necessarily end in the center of a bin, consequently, the final position of each pathline is adjusted to the center of the correponding bin to avoid gaps when connecting segments of two cycles (figure 1(h)). The weights of all pathlines are finally summed to get a probability map of the particle's position at the specified time step (figure 1(i)) [6].

VI. BACKWARD PROBLEM

In this section, we will start first by formulating our problem. Then, the methodology followed throughout our work with the backward problem.

A. Formulation

The backward problem can be formulated as follows:
Given the final position of the tracer and the velocity field, it is required to determine the source location. However, since going backward in time is by itself an ill-posed problem, and adding to this the uncertainties in the velocity field, a probabilistic inverse map will be computed leading to estimates of the source location.

B. Methodology

In order to tackle the backward problem, we followed the approach presented hereafter:

Since we're dealing with passive tracers advected by the surrounding field, diffusion, and the associated irreversibility with it, is not present. Hence,in order to march backward in time, it is enough to invert the velocity field at the tracer's position as shown in figure 2. In fact, as presented in the forward problem formulation, solving the backward problem requires also solving the Lagrangian advection equation:

(a) Sample velocity field

(b) Inject particle

(c) Integrate particle trajectory

(d) Obtain particle trajectories for 5 members

(e) Final particle positions in 1 cycle for 5 members

(f) Bin the domain

(g) Count particles per bin and use as weight

(h) Render pathlines with their weight

(i) Sum the weights of all pathlines
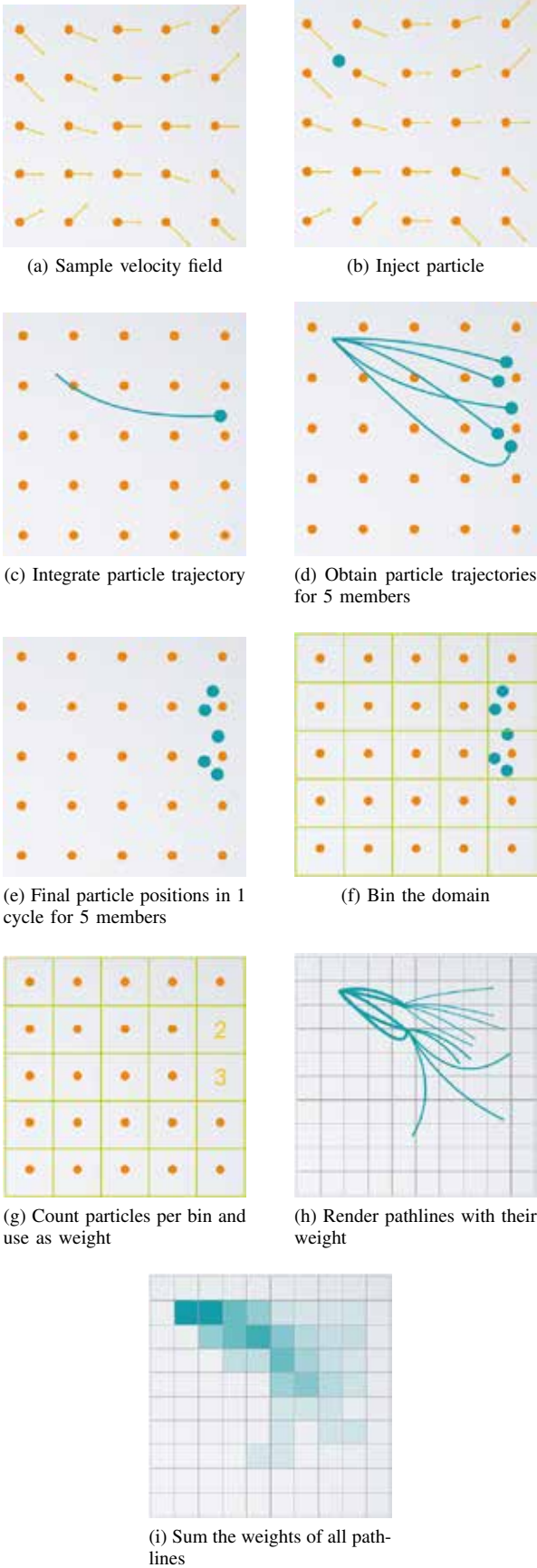
Fig. 1: Probability map computation using spatial binning technique
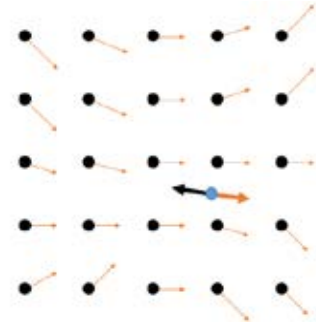
Fig. 2: Inverting the velocity field
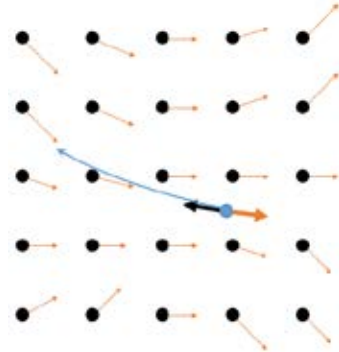
Fig. 3: Marching backward in time using approach 1

$$\frac{d\boldsymbol{r}}{dt} = \boldsymbol{u}(\boldsymbol{r}, t) \tag{3}$$

$$\boldsymbol{r}(t_2) - \boldsymbol{r}(t_1) = \int_{t_1}^{t_2} \boldsymbol{u}(\boldsymbol{r}, t) \; dt \tag{4}$$

But now, we are given the position at t_2, and we seek finding the position at t_1, therefore:

$$\boldsymbol{r}(t_1) = \boldsymbol{r}(t_2) - \int_{t_1}^{t_2} \boldsymbol{u}(\boldsymbol{r}, t) \; dt \tag{5}$$

Which is equivalent to:

$$\boldsymbol{r}(t_1) = \boldsymbol{r}(t_2) + \int_{t_1}^{t_2} -\boldsymbol{u}(\boldsymbol{r}, t) \; dt \tag{6}$$

Thus, running the backward problem is nothing but running the forward problem with inverting the velocity fields and starting from the final time. This approach shown in figure 3 is repeated for each member of the ensemble data and makes use of the binning probability technique presented before.

VII. RESULTS

After implementing the methodologies presented above numerically in Fortran, the results obtained for both the forward and backward problems are presented hereafter.
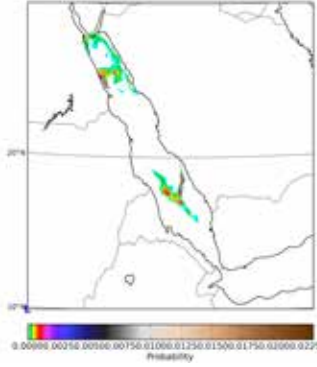
Fig. 5: Several particles with different locations after 10 time steps
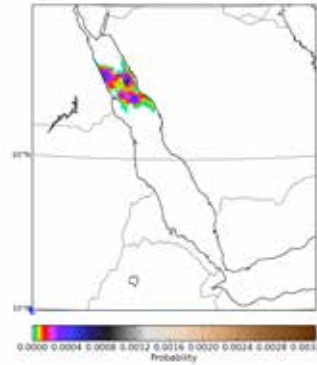
Fig. 6: Several particles with different locations after 10 time steps

A. Forward Problem Results

In what follows, results of the forward problem are presented for 2 cases that vary depending on the number of tracers, their location, and their emission time used during the simulation.

1) Case of a single tracer: Results for a single tracer are presented in figure 4 for different time steps.

2) Case of several tracers at different source locations: Corresponding results are presented in figure 5.

B. Backward Problem Results

In what follows, results of the backward problem are presented in figure 6. It was run from the location where the forward problem lead to the highest probability. Comparing the results, we see that the backward problem has detected the source location with high probability.

VIII. CONCLUSION

The prior history and trajectory of a tracer or a spill might be critical in environmental and safety related applications. This paper presented an approach that can be used for inverse Lagrangian particle tracking in which the source location is sought given the final position and the velocity fields. The approach presented overcomes the computational cost associated with such problems due to the uncertainties associated with measured and simulated quantities. Specifically, instead of tracking each path alone, a probability map is computed that can be used to predict both the forward and backward motion of particles in an oceanic field.

IX. FUTURE WORK

A. Part 1

We are currently working on applying our advection algorithm on the case of normally distributed velocity ensembles at every position and compare the results with the following mathematical reasoning:

For a given particle's position r_i and given time t_j

$$\frac{d\boldsymbol{r}}{dt} = \boldsymbol{u}(\boldsymbol{r_i}, t_j) \tag{7}$$

where

$$\boldsymbol{u}(\boldsymbol{r_i}, t_j) = \boldsymbol{u}_{fluid}(\boldsymbol{r_i}, t_j)$$

is a 3D velocity field that is normally distributed with mean and standard deviation vectors:

$$\boldsymbol{\mu}(\boldsymbol{r_i}, t_j) = (\mu_x(\boldsymbol{r_i}, t_j), \mu_y(\boldsymbol{r_i}, t_j), \mu_z(\boldsymbol{r_i}, t_j))$$

and

$$\boldsymbol{\sigma}(\boldsymbol{r_i}, t_j) = (\sigma_x(\boldsymbol{r_i}, t_j), \sigma_y(\boldsymbol{r_i}, t_j), \sigma_z(\boldsymbol{r_i}, t_j)) respectively.$$
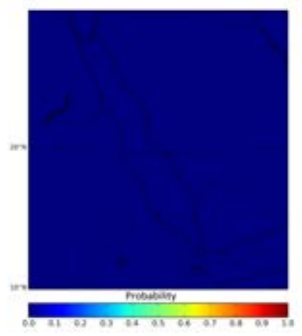
(a) t=0;inject particle
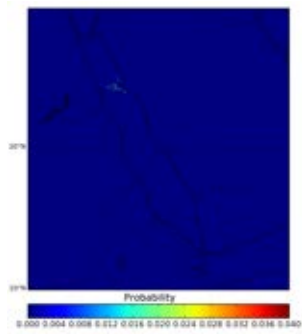
(b) t=1 (after 3 days)
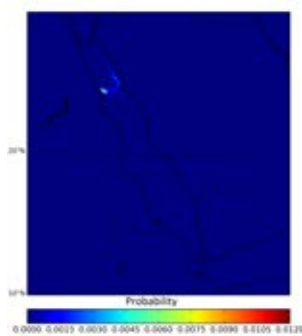
(c) t=2 (after 6 days)
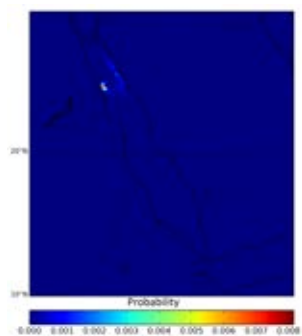
(d) t=3 (after 9 days)
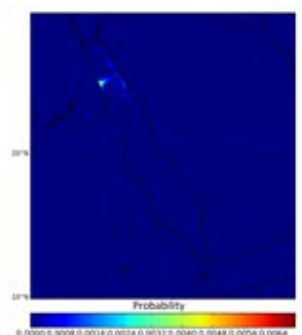
(e) t=4 (after 12 days)
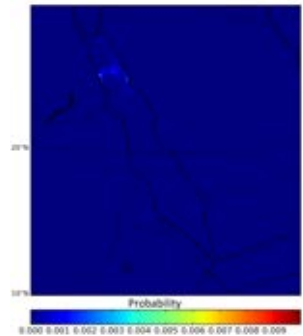
(a) t=5 (after 15 days)
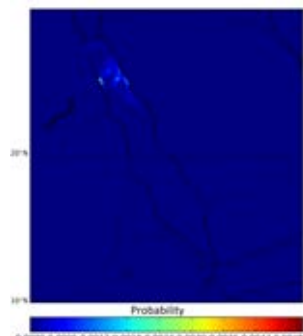
(b) t=6 (after 18 days)
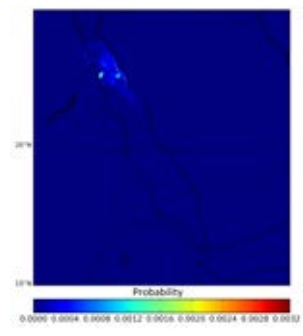
(c) t=7 (after 21 days)
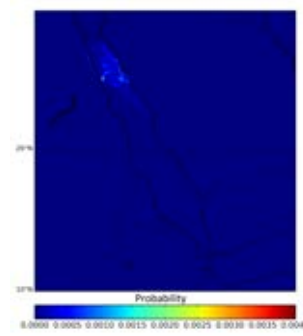
(d) t=8 (after 24 days)

Fig. 4: Forward model sample results for different time steps

Hence

$$r(t_f) - r(t_j) = \int\limits_{t_j}^{t_f} u(r_i, t_j) \, dt \qquad (8)$$

equivalently

$$r(t_f) = u(r_i, t_j) \times (t_f - t_j) + r(t_j) \qquad (9)$$

This $r(t_f)$ will also be normally distributed since it is a linear transformation of $u(r_i, t_j)$. Calculating the first and second moments of $r(t_f)$, we find that it has the following mean and standard deviation vectors:

$$\mu(r_i, t_j) \times (t_f - t_j) + r(t_j)$$

and

$$\sigma(r_i, t_j) \times (t_f - t_j)$$

respectively.

B. Part 2

As previously mentioned, we are dealing with passive tracers in our work, and there is no need to account for the physical and chemical processes taking place with the surrounding environment. However, for future work, a more realistic representation of an oil spill would be by incorporating some non-conservative processes such as coalescence and breakup, bio-degradation, emulsification, evaporation, and sedimentation.

Diffusion had shown to have no important effect in this specific work. Yet, for a different case, when working with another type of particles rather than oil, or when studying particles motion in the atmosphere for example, the diffusion there will be much more important and must be accounted for. The approach of "Inverting the velocity field" presented before would be no longer valid once diffusion (or any other non conservative process) is introduced. Two approaches were thought of to help solving the problem of diffusing particles.

1) Running instances of the forward problem: Given the final position at time t_2, several instances of the forward model need to be run at time t_1. The issue of where to seed the forward model becomes important as seeding everywhere (or at least on a finely resolved grid) "upstream", and identifying the few particles that arrive at the final location, would be computationally expensive. One solution can be introduced by limiting the search region to a rather smaller domain by using a spatial scale obtained from the magnitude of the velocity at time t_2, and the time step, i.e.:

$$R_{searchregion} = \|\mathbf{V}\| \, \triangle t \qquad (10)$$

This method might still require high computational cost when the search regions are themselves large.

2) A combination approach: As seen before, the approach of "Inverting the velocity field" cannot be implemented once nonconservative processes are introduced. In addition to that, the approach of "running instances of the forward problem" might require high computational cost. A good idea would be then to combine both approaches, i.e by inverting the mean velocity of the ensemble, the particle can be marched backward in time, and then a search region is constructed at the detected position within which multiple instances of the forward problem are launched, and those leading to the final location are combined.

REFERENCES

[1] M. McNutt, R. Camili, G. Guthrie, P. Hsieh, V. Labson, B. Lehr, D. Maclay, A. Ratzel, ans M. Sogge, Assessment of flow rate estimates of the Deepwater Horizon/Maconda Well Oil Spill Flow Rate Technical Group report to the National Incident Command, Interagency Solutions Group,(2011).
[2] C. Brekke, A. Solberg, Oil spill detection by satellite remote sensing, Remote Sensing of Environment, Volume 95, Issue 1, (2005).
[3] J. Lubchenco, M. McNutt, B. Lehr, M. Sogg, M. Miller, S. Hammond, and W. Conner, Deepwater Horizon/BP Oil Budget: what happened to the oil? National Oceanic and Atmospheric Administration Report. Silver Spring, MD, (2010).
[4] D.K. Scholz, J.H. Kucklick, R. Pond, A.H. Walker, and A. Bostrom, Fate of Spilled Oil in Marine Waters American Petroleum Institute Publication, 4691, (1999).
[5] Hazen et. al, Deep-sea oil plume enriches indigenous oil-degrading Bacteria Science 330 (6001), (2010).
[6] T. Hollt, M. Hadwiger, O. Knio, I. Hoteit, Probability Maps for the Visualization of Assimilation Ensembles Flow Data (2015).

Multi-Dimensional Model of the Human Arterial System

Israa Issa | Nicole Al Hanna | Andrew Debbas
Mechanical Engineering Department
American University of Beirut
Riad El-Solh | Beirut 1107 2020
isi00@mail.aub.edu | nia16@mail.aub.edu | acd01@mail.aub.edu

Abstract- Advanced state-of-the-art comprehensive predictive techniques applied on patient-specific data have been recently becoming inevitable for clinical applications. In specific, the use of Computational Fluid Dynamics (CFD) methodologies combined with patient specific data (CTA images) allows computing in-vivo arterial hemodynamics for patient follow up and diagnostic purposes. However, such modeling techniques still incur several burdens, starting from accurate reconstruction of patient specific arterial system and heart complicated dynamics, besides the challenges of replicating actual in-vivo patient dependent physiologic conditions (pressure and flow waveforms). This research aims at developing an advanced 3D numerical tool to model patient specific vasculature system which could be later adopted in different research investigations to mimic real patient hemodynamic distribution, understand various vascular diseases prognosis, and aid clinicians in decision-making and patients' following up. The construction of the modeling process starts by first acquiring CT scan data of a specific patient's human arterial system. The associated DICOM files were then converted into 3D geometric model that was then discretized into small control volumes. However, this 3D model is disconnected from the patient heart and the downstream arterial system. Therefore, to account for these upstream and downstream effects, lumped models were integrated with the generated 3D physiologic domain. These integrated models represent the heart at the inlet section of the truncated domain and the downstream arterial system at the outlet. The heart was assigned an electric circuit representing the function of the left ventricle while the outlet sections of the 3D physiologic domain were modelled by "Three Element Windkessel" electric analog circuits. To predict the blood hemodynamics (blood pressure and velocity fields), the set of Reynolds-averaged Navier-Stokes equations were applied. The finite volume method (FVM) was adopted to solve the algebraic discrete form of the conservation equations using ANSYS Fluent. In this research, a major contribution has been made to couple the lumped heart model to the 3D domain inlet section via "one type of boundary condition", rather than shifting between two types of boundary condition formulations to model the systolic and diastolic cardiac phases as been done in the literature.

Besides the 3D computational analysis, a reduced mathematical model was developed on simple human arterial tube models and solutions of pressure and waveforms were generated using Matlab code simulated under different boundary condition formulations. This step was very essential in formulating the problem, understanding the physics and the impact of each parameter, and designing a smooth numerical coupling between the patient model (artery), upstream (heart model) and downstream (3-element WK) models, before implementing the concept in the full 3D CFD model.

Predictions of cardiac properties were presented in terms of pressure-volume (PV) loops of the left ventricle, ventricular and aortic pressure, and lumen pressure of the arterial system. The obtained outcomes of the developed numerical models were realistic and compares well with what's being observed clinically.

Keywords: Human Vascular System, CFD, Finite Volume Method, Lumped Model, Windkessel

I. INTRODUCTION AND LITERATURE SURVEY

The Human Cardio-Vascular system is a vital system and understanding its complex nature and functionality can have extremely profound implications in medical research. Blood hemodynamic properties can be found using in-vivo techniques, but that is relatively expensive, hard to perform, and undoable for some parts of the body. Therefore, with the fast growing numerical advancements came the idea of building a comprehensive model which mimics the inherited physical phenomena that occur in human vascular systems. Computational modeling of cardiovascular systems has been used in studies related to diseased arteries, enhancements of medical devices [1][2], and planning of vascular surgeries (from therapeutic decision-making to outcome assessments [3][4][5]). On this basis, several researches have worked on developing computational-based models to be used as an algorithmic tool- the ultimate goal being to attain a certain level of predictive accuracy, given real patient-specific inputs. After rigorous dissection of the work provided in the open literature, it was found that the mathematical modeling of the human vasculature system was based on combining three modeling approaches: lumped, tube, and physiologic modeling.

Before discussing those approaches, an understanding of the basic functioning of the human heart is needed as a prerequisite for the terminology used in this paper. The heart consists of four main chambers: the left and right atriums, and the left and right ventricles (Fig. 1). Newly oxygenated blood coming from the lungs enters the heart into the left atrium. During the phase called diastole, the mitral valve, which lies between the left ventricle and left atrium, opens to make way for the blood into the ventricle. Then comes the diastole, during which the aortic valve allows the flow out of the heart. This valve separates the left ventricle from the aorta, which is the largest artery. From the aorta, all arterial branches emerge and extend to feed the whole body with oxygenated blood. It is important to note that whenever one of the valves mentioned is open, the other is closed. Finally, the cardiac cycle drives the blood to circulate in a closed loop and a pulse-like form. [6]

The information provided above may not be a detailed explanation on all the chambers and valves, but it is enough to understand the methodology of this research.
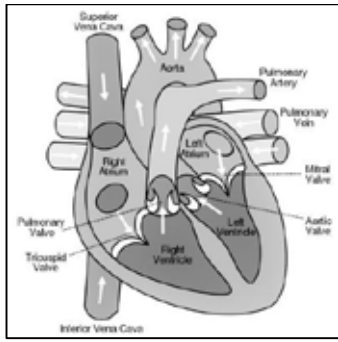
Figure 1. Labeled Model of the Human Heart [7]

The literature presents the lumped approach as a common method of modeling the arterial system. Segers et al. [8] have assessed the systemic and pulmonary hemodynamics following a lumped parameter approach to account for the interaction between the heart ventricles and the arteries. In their paper, the 4 element Windkessel model (4WM) was used to represent the system. Many other researchers adopted the simpler 2 element Windkessel model (2WM) to estimate aortic variables [9][10][11]. The difference between 2, 3 and 4 element circuits is the number of electric components utilized.
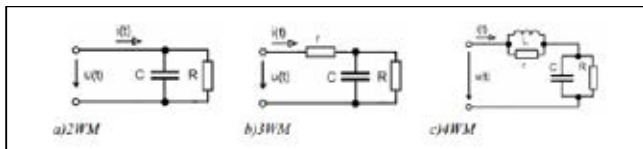
Figure 2. (a) 2-element Windkessel Model (b) 3-element Windkessel Model (c) 4-elements Windkessel Model [12]

In all of them, the vein compliance and peripheral resistance are represented by a capacitor and resistor, respectively. However, in the 3 and 4 element models, an extra resistance accounting for the opening and closing of the valves in the heart is used, and in the 4 element model, an inductor represents blood inertia. [12] Clearly, the more the components, the more accurate and the closer the model is to the real biologic system.

Other models have been developed such as the comprehensive Guyton model [13] which is a complex and detailed representation of most of the main circulatory branching, from the aorta to the veins. This, however, requires much more advanced studies since only through a thorough analysis of each branch would the model be reliable. Therefore, the Windkessel circuits, which represent the whole system by a few elements and yield accurate results, are enough for this work.

Not many studies use more than one approach to model the arterial system. One of the few is that of Kim et al. [14] who investigated the heart-arterial interaction by coupling a lumped parameter heart circuit and a 3D finite element physiologic model. In their study, they consider switching between two different boundary conditions at the inlet section, depending on the position of the valves. They adopted a weak pressure boundary condition when the aortic valve is open, and a zero velocity boundary condition when the mitral valve is open (the aortic valve is closed). At the outlet, a pressure boundary condition is used. In this work, we take it as a challenge to use only one type of boundary condition on each side of the 3D model. We find out the key is to use an outward mass flow rate on the outlet. This enables us to represent the inlet with only a pressure function, and therefore simplify the problem and avoid divergence and numerical fluctuations. In what follows, the research design method, numerical outcomes, and final recommendations and conclusions will be presented.

II. RESEARCH DESIGN METHOD

A. *Physical Model (Physiologic Model) Reconstruction and Numerical Grid Generation*

CT scan data were first acquired for a specific patient from the American University of Beirut Medical Center (AUBMC). The DICOM files were then entered into Mimics software to segment the model and keep the relevant features of the model. After then, a 3D reconstructed model was generated, forming the basis for the numerical computations (Fig. 3a). A total of 6,100,588 control elements (Fig. 3b) were then generated to numerically model the 3D patient specific physiologic domain.
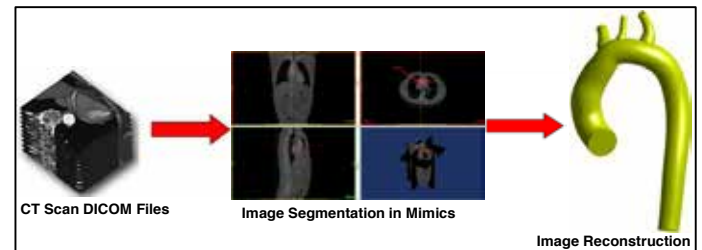
(a)
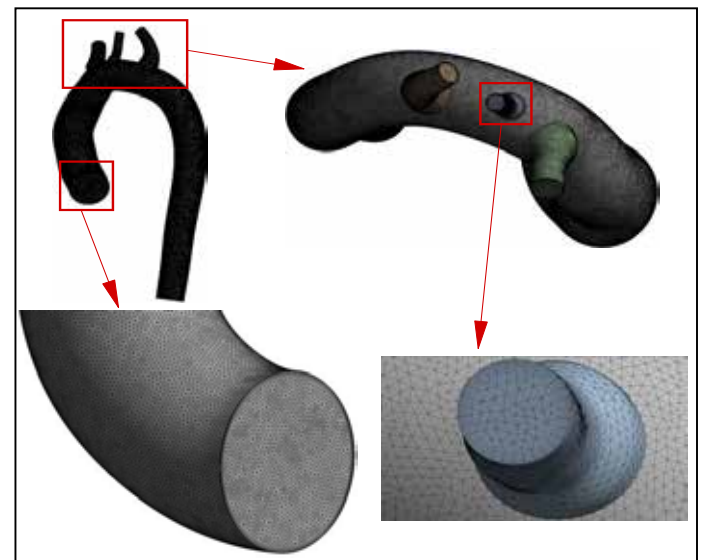
(b)

Figure 3 (a) Image Segmentation/Reconstruction; (b) Numerical Grid Generation.

B. Mathematical Modeling

Two different modeling approaches were formulated in this research work: **In the first part,** a full 3D numerical model based on computational fluid dynamics (CFD) methodologies coupled with lumped heart at inlet and 3-element Windkessel Models at outlets was formulated. **In the second part,** a reduced 1D model of an arterial system coupled with lumped heart and 3-element Windkessel Models was formulated. In what follows, the description of each approach will be mathematically and physically discussed.

1) 3D-Dimensional Finite Volume Based Aortic Model coupled with lumped heart and 3-element Windkessel

1-a) Three Dimensional Finite Volume Simulations of Blood Flow

Blood was assumed to be an incompressible Newtonian fluid with density and dynamic viscosity of values of 1050 kg/m^3 and 0.003 Pa.s, respectively. Hemodynamic properties of blood (i.e. pressure and velocity fields) are computed by solving the set of Navier-Stokes equations (continuity and momentum) given by:

$$\frac{\partial \rho}{\partial t} + \nabla.(\rho \mathbf{v}) = 0 \qquad (1)$$

$$\frac{\partial (\rho \mathbf{v})}{\partial t} + \nabla.(\rho \mathbf{v} \mathbf{v}) = -\nabla p + \nabla.(\boldsymbol{\tau}) + \mathbf{F_b} \qquad (2)$$

$$\boldsymbol{\tau} = \mu \{\nabla \mathbf{v} + (\nabla \mathbf{v})^T\} \qquad (3)$$

1-b) Lumped Parameter Heart Model

To induce the effects of the heart which is not fully modeled as done for the patient-specific arterial physiologic model, a lumped parameter heart model was coupled to the inlet boundary (A) of the aortic model as shown in Fig. 4. In this study, the left atrium and ventricle were modeled constituting of the left atrial pressure (atrium was modeled as one lumped element), mitral valve, atrio-ventricular resistance (representing viscous effects of blood flow) and inductance (representing inertial effects of blood flow), aortic valve, ventriculo-arterial valvular resistance and inductance, and left ventricular pressure (modeled also as lumped element). A time-varying Elastance function was considered to model the left ventricular pressure. The heart physiologic function is divided into two main phases, systole and diastole. During systole, the left ventricular pressure rises and when it reaches above the aortic pressure, the aortic valve opens and enables interactions between the left ventricle chamber and arterial system. (Fig. 5a). Accordingly, the aortic pressure and flow become solution due to the interactions between the lumped parameter heart model and the three dimensional finite volume model of the aorta. Moreover, the aortic flow is coupled to the ventricular volume during which the left ventricle is ejecting

blood to the aorta. On the other hand, during diastolic phase, the aortic valve closes (Fig. 5a), and hence the lumped heart model and the three dimensional aortic model are decoupled. Therefore, the solution of the ventricular pressure and its volume are predicted independently of the three dimensional physiologic aortic model and are solely function of the heart lumped model parameters. As the left ventricle continue to further relax during diastole, its pressure decreases and falls below the left atrial pressure upon which the mitral open (Fig. 5b). and ejects blood to the left ventricle. In this case, the left ventricle pressure, volume, and left atrial flow are determined solely by the interactions between the left ventricle and left atrium. The mitral valve closes (Fig. 5b) after and the cycle repeats again.

When the aortic valve is open, the lumped parameter heart model is given by:

$$P_{aorta}(t) = E(t)[V_v(t) - V_0] - R_{V-art}Q(t) - L_{V-art}\frac{dQ(t)}{dt} \qquad (4)$$

When the mitral valve is open, the lumped parameter heart model is given by:

$$P_{LA} = E(t)[V_v(t) - V_0] + R_{A-V}Q_{LA}(t) + L_{A-V}\frac{dQ_{LA}(t)}{dt} \qquad (5)$$

The ventricular volume and pressure are determined by the following equations:

$$\frac{dV_v(t)}{dt} = Q_{LA}(t) - Q(t) \qquad (6)$$

$$P_v(t) = E(t)[V_v(t) - V_0] \qquad (7)$$

P_{aorta}: the aortic pressure, [mmHg]
$P_v(t)$: the ventricular pressure, [mmHg]
P_{LA} : the left atrium pressure, [mmHg]
$E(t)$: the Elastance function of the left ventricle, [mmHg/cm^3]
$V_v(t)$: the ventricular volume, [cm^3]
Q: the aortic volume flow rate, [cm^3/s]
Q_{LA}: the left atrial volume flow rate, [cm^3/s]

$R_{V-art}, R_{A-V}, L_{V-art}$, and L_{A-V} : the aortic and mitral valves resistances and inductances, respectively

The time-varying Elastance model shown Fig. 4a is based on the work performed in an isolated heart prepared by Suga and Sagawa in the early 1970s, where they analyzed cardiac mechanics in the pressure-volume plane [15][16]. They recorded hemodynamic data measured both from isovolumic (non-ejecting) and ejecting cardiac contractions and found that the 'isochronic' points (i.e., points recorded after a given time from onset of contraction) in the different contractions are located on a single line, characterized by its slope and intercept with the volume axis. The slope is the ratio of the increase in pressure associated with an increase in ventricular volume, and hence has the dimension of stiffness or Elastance, $E(t)$.

Since the 3D physiologic domain (arterial system) is truncated at specific locations, the effects from the separated downstream physiologic model was accounted for via 3-elements WK models (Fig. 4a), governed by the following differential equation:

$$\left(1 + \frac{Z}{R}\right)Q(t) + CZ\frac{dQ(t)}{dt} = \frac{P(t)}{R} + C\frac{dP(t)}{dt} \tag{8}$$

2) Reduced 1D Arterial Model coupled with lumped heart and 3-element Windkessel Models

Before conducting the full 3D simulation using the finite volume method in fluent solver, the patient vasculature model was simplified to a simple arterial tube model (Fig. 4b) while the lumped models including the heart and the 3-element Windkessel (WK) models were maintained. The reduced model was extensively used to investigate the effect of various patient related physiological parameters on blood hemodynamics. This step was essential in formulating the problem, understanding the physics behind it, the impact of each parameter, and ensuring a smooth numerical coupling between the patient model (artery) and the upstream (heart model) and downstream (3-element WK) model. In this mathematical formulation, hemodynamic properties of blood (i.e. pressure and velocity fields) are computed by solving the set of Navier-Stokes equations (continuity and momentum) given in cylindrical coordinates as follows:

$$\frac{du_z(r)}{dz} = 0 \tag{9}$$

$$-\frac{dP(z)}{dz} + \mu\left(\frac{1}{r}\frac{d}{dr}\left(r\frac{du_z(r)}{dr}\right)\right) = 0 \tag{10}$$

The differential equation (10) is subjected to the following boundary conditions:

- $no - slip$ conditions @ walls
- $\frac{du_z}{dr}\Big]_{r=0} = 0$

u_z is the axial velocity (m/s), P is the pressure (Pa), and μ is the dynamic viscosity (Pa.s).

The solution of the velocity field is given by:

$$u_z(r) = \frac{1}{4\mu}\frac{\partial P}{\partial z}(r^2 - R^2) \tag{11}$$

To obtain the pressure distribution, apply the continuity constraint (eq. 9) on eq. 11,

$$\frac{\partial}{\partial Z}(u_z) = 0 \xrightarrow{yields} \frac{1}{4\mu}\frac{\partial^2 P}{\partial z^2}(r^2 - R^2) = 0 \tag{12}$$

The general set of solutions of eq. (12) is given by

$$P(Z) = C_3 Z + C_4 \tag{13}$$

The volume flow rate is then computed as follows

$$Q = \iint \overrightarrow{v_z(r)} . \vec{n} dS \tag{14}$$

Rearranging the obtained equations, the final solution of the flow in the artery could be written as:

$$P_1(t) = P_2(t) + \frac{128\mu L}{\pi D^4} Q(t) \tag{15}$$

Where: $P_1(t)$ and $P_2(t)$ are the inlet and outlet static pressures of the arterial model, respectively and $Q(t)$ is the blood volume flow rate.
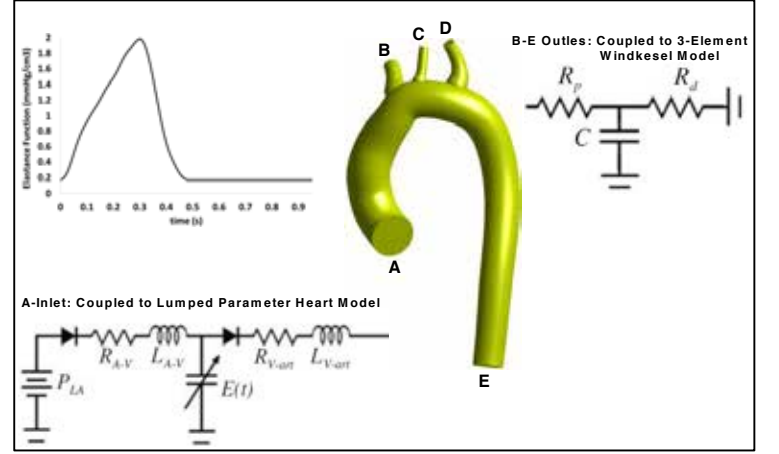
(a)
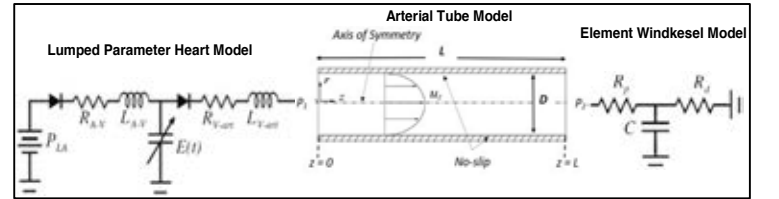
(b)

Figure 4. (a) 3D physiologic patient aortic model coupled to a heart and 3-element Windkessel models; [14] (b) 1D arterial model coupled to a heart and 3-element Windkessel models. [14]

(a)   (b)
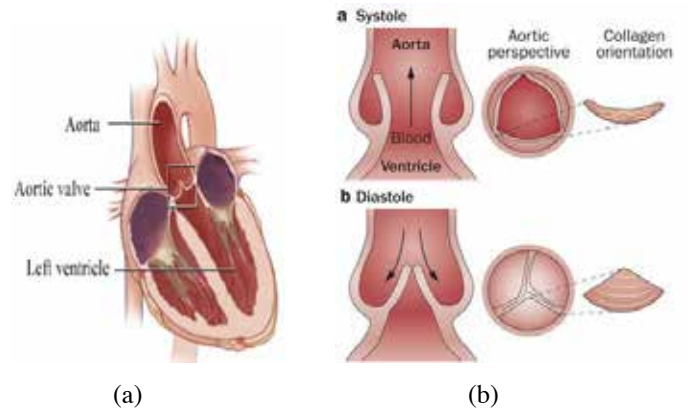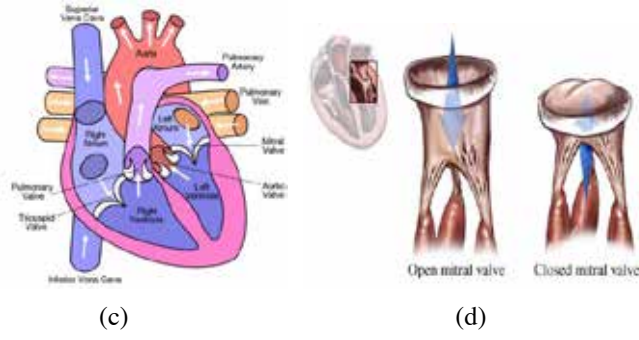
Figure 5. (a) Aortic Valve [19]; (b) Structure-function relationship in aortic valve [17] (c) Diagram of the human heart showing the chambers, valves and great vessels- by Wapcaplet in Sodiopodi [7]; (d) Normal Function of the Mitral Valve [18];

III. NUMERICAL FORMULATION AND COUPLING

Using the finite volume method (FVM), the governing partial differential equations (Equations 1, 2, and 3) are integrated over each cell of the numerical domain to form the discrete model upon which a set of algebraic equations are generated and iteratively solved over each element. The final discrete model governing φ at every location in the computational domain becomes:

$$a_P \varphi_P = \sum_{F=NB(P)} a_F^{CD} \varphi_F + b_P \tag{16}$$

Where: P is an arbitrary location of a control element and φ is a general variable that could be either pressure or velocity.

On the other hand, the forward finite difference scheme was used to evaluate the derivative terms appearing in equations. (4, 5, 6, and 8). The resulting numerical model was obtained,

$$\left(1 + \frac{Z}{R}\right) Q^{(t+\Delta t)} + CZ \frac{Q^{(t+\Delta t)} - Q^{(t)}}{\Delta t}$$
$$= \frac{P^{(t+\Delta t)}}{R} + C \frac{P^{(t+\Delta t)} - P^{(t)}}{\Delta t} \tag{17}$$

$$P_{aorta}{}^{(t+\Delta t)} = E^{(t+\Delta t)} [V_v{}^{(t+\Delta t)} - V_0] - R_{V-art} Q^{(t+\Delta t)}$$
$$- L_{V-art} \frac{Q^{(t+\Delta t)} - Q^{(t)}}{\Delta t} \tag{18}$$

$$P_{LA} = E^{(t+\Delta t)} [V_v{}^{(t+\Delta t)} - V_0] + R_{A-V} Q_{LA}{}^{(t+\Delta t)}$$
$$+ L_{A-V} \frac{Q_{LA}{}^{(t+\Delta t)} - Q_{LA}{}^{(t)}}{\Delta t} \tag{19}$$

$$\frac{V_v{}^{(t+\Delta t)} - V_v{}^{(t)}}{\Delta t} = Q_{LA}{}^{(t+\Delta t)} - Q^{(t+\Delta t)} \tag{20}$$

$$P_{aorta}{}^{(t+\Delta t)} = P_2{}^{(t+\Delta t)} + \frac{128\mu L}{\pi D^4} Q^{(t+\Delta t)} \tag{21}$$

$$P_v{}^{(t+\Delta t)} = E^{(t+\Delta t)} [V_v{}^{(t+\Delta t)} - V_0] \tag{22}$$

Coupling between the lumped heart, physiologic arterial system, and WK models was done via an automatic-based numerical algorithm. The major contribution in this work has been made in the coupling of the lumped heart model to the 3D domain inlet section using "one type of boundary conditions" rather than shifting between two types of boundary conditions (Dirichlet and Neumann) which is extensively used in literature. In literature when the aortic valve is closed, a zero velocity Dirichlet condition is assigned and when the aortic valve opens the boundary condition is switched back to Neumann pressure inlet boundary condition [13]. However, in previous works it has been shown how the choice of boundary conditions can change significantly the velocity and pressure fields [13]. To ensure accurate results and to reduce implementation complexities, the pressure inlet was modelled as the inlet boundary condition throughout the simulation instead of shifting between pressure inlet during systole and velocity inlet during diastole as previously done.

Moreover, to avoid fluctuations and divergence and to make the solution prediction stable, an under relaxation factor ("alfa") was introduced in the numerical algorithm. This guarantees the way to the solution to be smooth and stable, however, it decreases a bit the convergence rate. It's recommended whenever using iterative method to control the rate at which variables are changing during iterations, it's also worth mentioning that this under-relaxation factor changes the path to the solution only and not the discretization [19].

Fig. 6 depicts a portion of a developed Matlab code integrated to simulate numerically the problem. Further, generic C++ codes for serial and parallel processing were built to implement the model represented by equations. (17-22) and coupling it to Fluent solver. The Semi Implicit Method (the iterative method) for Pressure Linked Equations (SIMPLE) algorithm was adopted to resolve the pressure-velocity coupling.

Fig. 6b and 6c show the velocity field of the reduced 1-D model (see Fig. 6b) coupled with lumped heart and 3-element Windkessel models with specifications of 3.5 mm diameter and resistance of 120000. Comparison of the two graphs indicate that both simulations match perfectly for different diameters as well as for different resistances. Also, the pressure field simulated on both CFD and Matlab exhibits similar patterns for different diameters and different resistances.

```
%Iterative Procedure
convergence_criterion = 1e-6;
modes = mod(time,T);
for itime=1:length(time)-1

    delta_t(itime) = time(itime+1) - time(itime);

    if (modes(itime+1) >= 0) && (modes(itime+1)<tao))
        Q(itime+1) = 0;
        Q_LA(itime+1) = 0;
        Vv(itime+1) = Vv(itime);
        P1s(itime+1) = ((C/delta_t(itime))*P1s(itime))/((1/R + C/delta_t(itime)));
        P2s(itime+1) = P1s(itime+1);

    elseif ((modes(itime+1) >= tao) && (modes(itime+1)<tao))
        Q_LA(itime+1) = 0;
        %Initial Guess
        Q(itime+1) = 0;
        Vv(itime+1) = 110e-3;
        P1s(itime+1) = 90*133.3224;
        P2s(itime+1) = 90*133.3224;
        for iterations=1:1900000
            iterations(itime) = iterations;
            Q_previous = Q(itime+1);
            Vv_previous = Vv(itime+1);
            P1s_previous = P1s(itime+1);
            P2s_previous = P2s(itime+1);
            %check if initial guess is satisfied by computing the errors
            error1(itime) = abs((1 + Z/R + C*Z/delta_t(itime))*Q(itime+1) - ((1/R +
C/delta_t(itime))*P2s(itime+1) - (C/delta_t(itime))*P2s(itime) + (C*Z/delta_t(itime))*Q(itime)));
            error2(itime) = abs((P1s(itime+1) - (E(itime+1)*(Vv(itime+1) - Vo) - (Rv_art +
Lv_art/delta_t(itime))*Q(itime+1) + (Lv_art/delta_t(itime))*Q(itime)));
            error3(itime) = abs(Vv(itime+1) - (delta_t(itime)*(Q_LA(itime+1) - Q(itime+1)) +
Vv(itime)));
            error4(itime) = abs(P2s(itime+1) - (P1s(itime+1) -
(8*mu*L/(pi*(D/2)^4))*Q(itime+1)));
            errors = [error1(itime) error2(itime) error3(itime) error4(itime)];
            Max_Error(itime) = max(errors);
            if Max_Error(itime) < convergence_criterion
                break
            end
            %Update Variables (in this case they are: P1s, P2s and Q, and Vv)
            Q(itime+1) = ((1/R + C/delta_t(itime))*P2s(itime+1) - (C/delta_t(itime))*P2s(itime) +
(C*Z/delta_t(itime))*Q(itime))/(1 + Z/R + C*Z/delta_t(itime));
            P1s(itime+1) = (E(itime+1)*(Vv(itime+1) - Vo) + (Rv_art +
Lv_art/delta_t(itime))*Q(itime+1) + (Lv_art/delta_t(itime))*Q(itime));
            Vv(itime+1) = (delta_t(itime)*(Q_LA(itime+1) - Q(itime+1)) + Vv(itime));
            P2s(itime+1) = (P1s(itime+1) - (8*mu*L/(pi*(D/2)^4))*Q(itime+1));

            P2s(itime+1) = alfa*(P2s(itime+1)) + (1 - alfa)*P2s_previous;
            Q(itime+1) = alfa*(Q(itime+1)) + (1 - alfa)*Q_previous;
            Vv(itime+1) = alfa*(Vv(itime+1)) + (1 - alfa)*Vv_previous;
            P1s(itime+1) = alfa*(P1s(itime+1)) + (1 - alfa)*P1s_previous;
        end
```

Figure 6a. Portion of the integrated numerical algorithm in Matlab code.

Fig 6b. Velocity field of the reduced 1D arterial model coupled with lumped heart and 3-element WK Models simulated on CFD.
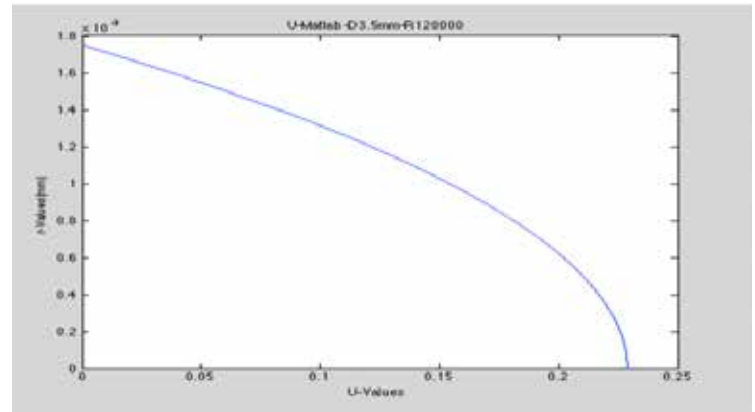
Fig 6c. Velocity field of the reduced 1D Arterial Model coupled with lumped heart and 3-element WK Models on Matlab.

IV. RESULTS AND DISCUSSIONS

The parameter values of the lumped parameter heart and 3-element Windkessel model given in equations. (4, 5, and 8) were selected based on measurements from the literature [20] [21].

Results of the cardiac properties were reported for rest conditions. Solutions were obtained using a 6,100,588 elements (control volumes) mesh with a time step size of 1e-3 sec. The 3D computer based model was simulated using parallel processing on a 128 core high performance computer (HPC) machine at AUB (CFD Labs). A total of minimum 8 cardiac cycles were required until the final cardiac properties and hemodynamics yielded periodic solutions. Fig. 7 depicts the simulated pressure-volume (PV)-loop of the left ventricle simulated over 12 cycles. In the results shown in this study, the time dependent resistance and inertial effects during the closing of the aortic and mitral valves were not accounted for, and this has led to modification of the PV-loop shape during the isovolumic contraction and relaxation phases. However, this issue could be easily resolved by varying the valvular resistance during closure phases. Ongoing research is being performed to formulate a general time dependent function of the mitral and aortic valves to capture the ventricular pressure and volume changes during the isovolumic contraction and relaxation phases. Note also that PV-loops could have different shapes other than a perfect rectangular shape in some cases, for instance when valvular diseases exist. Fig. 7b shows the pressure-volume loop of the left ventricle for a single cycle [14]. To account for the closing of the aortic valve during diastole, there is a shift between Neumann boundary conditions to Dirichlet boundary conditions (zero velocity) [14]. This could be the main reason why there is some discrepancy between the two figures (Fig 7a and 7b).
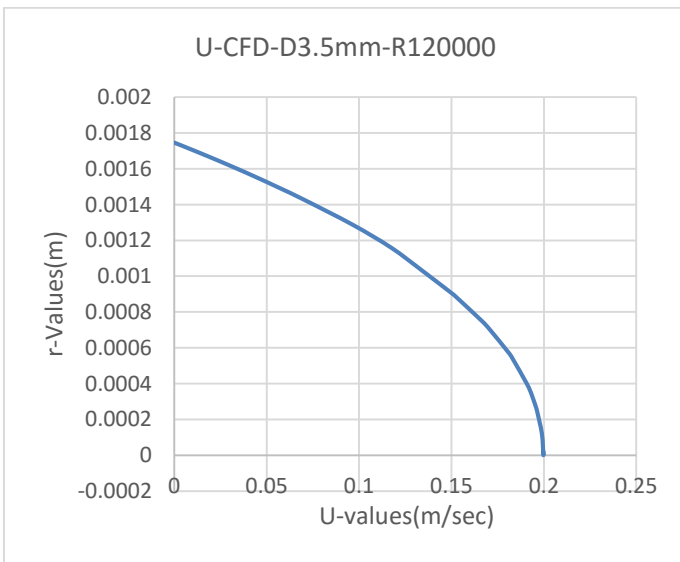
Fig. 8 shows the aortic and ventricular pressure during 1 cardiac cycle. This behavior matches well with what is

clinically observed, where the left ventricular pressure increases slightly above the aortic one during cardiac ejection to systemic circulation, and then it drops below it during diastolic phase. This simulation gives a computed left ventricular pressure ranging from 18 to 125 mmHg and a Ventricular pressure ranging from 62 to 125 mmHg (see Fig. 8a). Fig. 8b shows the pressure and velocity waveforms of the left ventricle and aorta extracted from a real physiologic patient. By comparing the red and black graphs of 6b with that of 6a, these two graphs show very similar results to what we got during our simulations (Fig. 8a), where both graphs (8a and 8b) exhibit a very similar pattern over the whole cycle (0.92 sec). Fig. 8c on the other hand shows computed pressure waveforms of the inlets and outlets for thoracic aorta at rest and during exercise [14]. By comparing the thoracic aorta pressure curve at rest (Fig. 8c) with the simulated aortic pressure (Fig. 8a), the two graphs show similar trends, with little to no discrepancies.

Fig. 9 shows the predicted lumen pressure at different flow times over 1 cardiac cycle in order to illustrate the varying flow features during different phases of the heart mainly: systole, diastole and mid-diastole. and it also agrees well with what is expected (compared to other literature work and physiologic data).

Figure 8a. Simulated Ventricular and Aortic Pressure after 12 cycles.
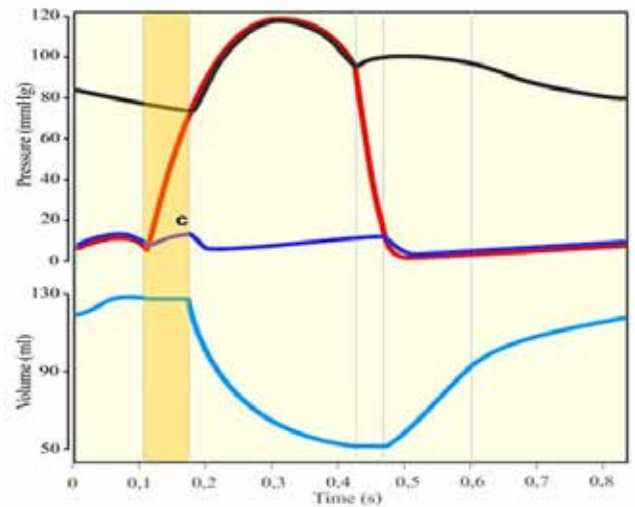
Fig 8b. Pressure and Volumetric changes from a real physiologic data during one cardiac cycle [20].

Figure 7a. Simulated PV-loop

Figure 7b. PV- loop of the left ventricle for a single cardiac cycle, [14]
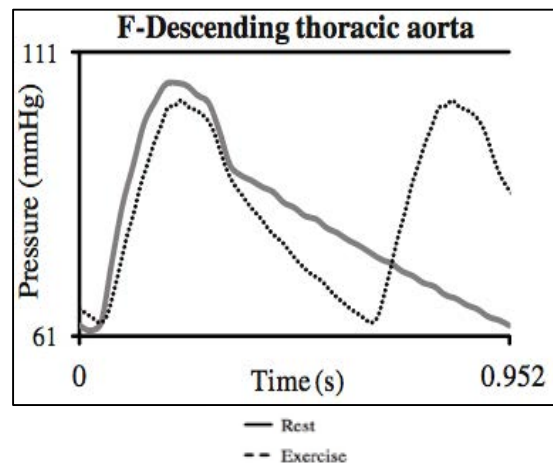
Fig 8c. Computed pressure waveform at the inlet and outlet boundaries of thoracic aorta at rest and during exercise [14].
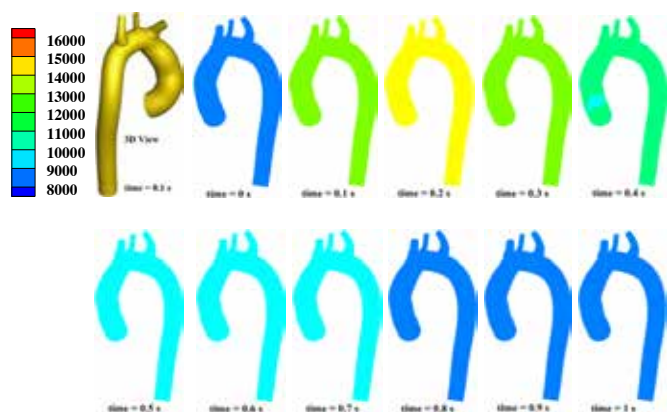
Figure 9. Predicted lumen pressure (scale shown in Pascal. 1mmHg = 133.32 Pa) at different flow time over 1 cardiac cycle.

V. CLINICAL TRANSLATION

The purpose of this project is to build a comprehensive integrated lumped-3D computational model that will serve as a research tool for several future clinical and biomedical engineering applications. This model relates several cardiac parameters of specific patient that allows predicting different hemodynamic properties including but not limited to pressure, velocity distribution and recirculation, wall shear stress, etc.… Moreover, it allows understanding the change in aortic flow and blood pressure due to the interaction with the cardiac cycle and vice versa. This computational-based tool could be integrated in different clinical applications to study blood hemodynamics in diseased arterial segments (valvular diseases, plaque prognosis, heart defects, and peripheral artery diseases), enhance the design of medical devices, planning of vascular surgeries and in therapeutic decision-making, prediction of disease prognosis in human vascular system, design of artificial medical devices to sustain failing hearts, grafting procedures and diagnostic procedures.

VI. CONCLUSIONS

An advanced integrated lumped-3D numerical tool was developed in this study to model patient specific vasculature system. The interaction between the 3D model and the upstream and downstream domains is challenging as it arises complicated modeling issues at the boundaries. In this research, a major contribution has been made to couple the lumped heart model to the 3D domain inlet section via "one type of boundary condition formulation", rather than shifting between two types of boundary condition formulations to model the systolic and diastolic cardiac phases as been done in the literature. This project is the first time to be researched in AUB, and the promising outcomes and the resulting numerical platform will serve as an inevitable tool for future research

investigations by AUB engineering and medical community. This numerical platform could be later translated into practical clinical applications that could be adopted in different research investigations to mimic real patient hemodynamic distribution, understand various vascular diseases prognosis, and aid clinicians in decision-making and patients' following up. This project is the first time to be researched in AUB, and the resulting numerical platform will serve as an inevitable powerful tool for future research and clinical investigations by AUB engineering and medical community.

REFERENCES

[1] Li, Z., and C. Kleinstreuer. Blood flow and structure interactions in a stented abdominal aortic aneurysm model.Med. Eng. Phys. 27(5):369–382, 2005.

[2] Stuhne, G. R., and D. A. Steinman. Finite-element modeling of the hemodynamics of stented aneurysms. J. Biomech. Eng. 126(3):382–387, 2004.

[3] Migliavacca, F., R. Balossino, G. Pennati, G. Dubini, T. H. Hsia, M. R. de Leval, and E. L. Bove. Multiscale modelling in biofluidynamics: application to reconstructive pediatric cardiac surgery. J. Biomech. 39(6):1010–1020, 2006.

[4] Soerensen, D. D., K. Pekkan, D. de Zelicourt, S. Sharma, K. Kanter, M. Fogel, and A. P. Yoganathan. Introduction of a new optimized total cavopulmonary connection. Ann. Thorac. Surg. 83(6):2182–2190, 2007.

[5] Taylor, C. A., M. T. Draney, J. P. Ku, D. Parker, B. N. Steele, K. Wang, and C. K. Zarins. Predictive medicine: computational techniques in therapeutic decision-making. Comput. Aid. Surg. 4(5):231–247, 1999.

[6] Catanho, Marianne, Mridu Sinha, and Varsha Vijayan. Model Of Aortic Blood Flow Using The Windkessel Effect. 2012. Web. 17 Apr. 2016.

[7] R. I. Ionasec, I. Voigt, B. Georgescu, Y. Wang, H. Houle, F. Vega-Higuera, . Comaniciu, D.. Patient-Specific Modeling and Quantification of the Aortic and Mitral Valves From 4-D Cardiac CT and TEE. IEEE Transactions on Medical Imaging IEEE, Trans. Med. Imaging, 29(9), 1636-1651,2010

[8] Segers, P., N. Stergiopulos, N. Westerhof, P. Wouters, P. Kolh, and P. Verdonck. Systemic and pulmonary hemodynamics assessed with a lumped-parameter heartarterial interaction model. J. Eng. Math. 47(3):185–199,2003.

[9] Burattini R, Natalucci S: Complex and frequency-dependent compliance of viscoelastic Windkessel resolves contradictions in elastic Windkessel. Med Eng. Phys 1998, 20: 502–514. 10.1016/S1350-4533(98)00055.

[10] Burattini R, Campbell KB: Physiological relevance of uniform elastic tube-models to infer descending aortic wave reflection: a problem of identifiability. Ann Biomed Eng 2000, 28: 512–523.

[11] John LR: Forward electrical transmission line model of the human arterial system. Med Biol Eng. Compute 2004, 42: 312–321. 10.1007/BF02344705.

[12] CTU,. SYSTEM ANALYZE OF THE WINDKESSEL MODELS. Prague: N.p., 2016. Web. 17 Apr. 2016.

[13] Guyton AC, Coleman TG, Granger HJ: Circulation: overall regulation. Annu Rev Physiol 1972, 34: 13–46. 10.1146/annurev.ph.34.030172.000305.

[14] H.J.Kim, I.E.Vigon-Clementel, C.A. Figueroa, F.LaDisa, E.Jansen, J.A Feinstein and C.A. Taylor, "On Coupling a Lumped Parameter Heart Model and a Three-Dimensional Finite Element Aorta Model", Annals of Biomedical Engineering, Vol.37, No.11,pp.2153-2169, Nov. 2009.

[15] H. Suga, K. Sagawa and A. A. Shoukas, Load independence of the instantaneous pressure-volume ratio of the canine left ventricle and effects of epinephrine and heart rate on the ratio. Circ. Res. 32 (1973) 314-322.

[16] H. Suga and K. Sagawa, Instantaneous pressure-volume relationships and their ratio in the excised, supported canine left ventricle. Circ. Res. 35 (1974) 117–126.

[17] W.D. Merryman, J.D. Hutcheson, E. Aikawa, Potential drug targets for calcific aorta valve disease. Nat Rev Cardiol Nature Review Cardiology, 11(4), 218-231, 2014.

[18] Staff Healthwise, K.Pai, MD, FACC- Cardiology, Electrophysiologic, J. Gabica, MD-Family Medicine, Retrieved from http://www.nucleusinc.com, Feb. 2015.

[19] S.R. Mathur, J.Y. Murthy, Numerical Methods in Heat, Mass, and Momentum Transfer ME 608, School of Mechanical Engineering Purdue University, Spring 2012.

[20] Ottesen, J. T., M. S. Olufsen, and J. K. Larsen. Applied Mathematical Models in Human Physiology. SIAM Monographs on Mathematical Modeling and Computation. Philadelphia: SIAM, 2004.

[21] Segers, P., N. Stergiopulos, and N. Westerhof. Relation of effective arterial elastance to arterial system properties. Am. J. Physiol. Heart Circ. Physiol. 282(3):H1041–H1046, 2002.

[22] N.A. Cardiac Cycle, Retrieved from: https://mefanet.lfp.cuni.cz/download.php?fid=31, N.D.

Numerical and Experimental Investigation of Heat Transfer in Ovens

Hussein Daher- Mohamad Fleifel- Amr Gali- Bassel Jannoun
Mechanical Engineering Department
American University of Beirut
Bliss Street, Beirut, Lebanon
baj03@mail.aub.edu- hyd04@mail.aub.edu- mhf14@mail.aub.edu- amg15@mail.aub.edu

Abstract Summary:

This paper discusses the validation of a Computational Fluid Dynamics (CFD) model to calculate the heat transfer in an industrial diesel forced-convection oven. We will investigate the performance of the oven while varying the operating conditions, with and without bread in the baking chamber, to reach the most reliable CFD model portraying the actual oven. The Simulations and Results are still in progress and will be discussed when finished.

I. INTRODUCTION

Our project is a well-synchronized combination of Fluid Dynamics, Thermodynamics, Heat Transfer, and Design of Thermal Systems. The plan is to discuss the validation of a Computational Fluid Dynamics (CFD) model to calculate the heat transfer in an industrial gas forced-convection oven.

The heat oven we aim to study is not functioning at its maximum capacity. Through our study, we aim to provide KARAKI, by using a CFD model, the conditions needed to minimize diesel fuel input, maximize the heating process, and reduce the time needed to bake bread across the baking chamber whether placed at the center or at the borders of the baking chamber

We intend to guarantee that the technical specifications of our final product are met. These specifications include: The heat is homogenously spread across the baking chamber reaching bread to all locations equally, reaching the optimum conditions needed to maintain bread moisture and quality, reducing the power requirement.

II. LITERATURE REVIEW

The study aims on performing a numerical and experimental CFD 3D model of an industrial forced convection bread baking oven. The 3D model is coupled with an appropriate turbulence model, taking into account bread moisture, quality, and focusing on the optimization and uniformity of the heating process. The oven under study is composed of three main parts, a burner, a heat exchanger, and a baking chamber. Our focus lies on the CFD modeling of the heat exchanger and the baking chamber.

Computational fluid dynamics is tremendously advantageous for predicting temperature and velocity fields in an oven.

Lately, it is evident from the numerous publications of the last decade that the use of CFD in food processing studies has become a growing movement, particularly in industrial ovens.

In 2000, Verboven et al. [1] discussed isothermal air flow in a forced convection oven using the application and authentication of CFD modeling. Three years later, Therdthai et al. [2] performed a 3D CFD analysis coupled with a moving grid approach to examine temperature outlines and flow profiles in a continuous industrial baking process. The 2D CFD model was later on used with the sliding mesh technique to study the oven chamber of a continuous industrial bread baking oven [3]. In 2009, Williamson et al. [4] presented the analysis and optimization of the design of an innovative gas fired radiative burner for industrial tunnel ovens using CFD simulation to calculate radiation heat fluxes applied on the exterior crust of the food. Recently a CFD model of a bakery pilot oven was designed with three heat transfer mechanisms measured and combined with turbulent flow [5]. The results were compared with experimental values, reporting promising equivalence for all three models.

Heating processes adopted in industrial ovens include the bake and broil cycle. The broil cycle denotes cooking with heat generated only from the top heater while the bake cycle denotes cooking with heat generated only from the bottom heater.

Mistry et al. [6] uses CFD methodology to study transient natural convection heat transfer in an electric oven under two different cooking cycles, the Bake and the Broil. Heat loss contribution analysis was performed to understand the contribution of heat loss through the walls, vent, and glass door of the oven. It was concluded that the maximum uncertainty between the gas temperature and the thermocouple reading is 9 %. After a thermal flow field analysis was executed, it was observed that the thermal field is similar for both laminar and

turbulent flows, thus a laminar approach was used. The core results state that the Broil Cycle attains non uniform temperature distribution inside the cavity while the Bake Cycle reaches uniform temperature distribution inside the cavity. Additionally, the dominant mode of heat transfer in a Broil cycle is surface to surface radiation between the heater and oven walls and between the walls themselves, while that of a Bake cycle is radiation heat transfer between the heater and oven walls combined with convective heat transfer. Thus in a Bake cycle, efficient convective heat transfer leads to temperature uniformity in the baking chamber. Our study focuses on the experimental and numerical modeling of a forced convection industrial oven. Rek et al. [7] states that there are three different methods of baking that can be classified as classic, fan, and grill. In the classic method, an electric heater is placed on both the floor and ceiling and the heat transfer is due to natural convection, conduction and radiation. In the fan method the air entering the chamber is heated before entering and then it's circulated in the chamber by a fan, so the mechanism of heat transfer is forced convection. The third method is grill where an electric heater is placed on the ceiling and it's heated to a very high temperature so the main mechanism of heat transfer is radiation. The oven under study relies on the fan method and the main mechanism of heat transfer is forced convection. It's important to note that we will be concerned with the steady state condition, thus the equations will be simplified since all the time derivatives are zero. The oven is considered to be static i.e. the bread enters the oven in a tray and remains in its position until the baking process ends. Litovchenko [8] discusses the computer simulation of equipment in the food industry including heat exchangers, pumps and ovens. Software Flow Vision was used to design and simulate the movement of liquids and gases in devices consisting of four steps. These steps are the development of 3D geometry in the CAD programs, the implementation of boundary conditions (input, output, wall, etc.), the assignment of rheological properties (density, viscosity, emissivity), and situating the driving force (velocity, pressure drop, convection, elevation).

Rek et al. [9] notes in their evaluation of emissivity that the emissivity of food products (in this case biscuits) is not well known, so different emissivities were associated to biscuits and the corresponding surface temperature variations were recorded from which it was noticed that biscuit surface emissivity is not of significant effect on their analysis. In a different study, hot air is used as the heating medium and results for an empty oven as well as for an oven loaded with bread. It was concluded that the quantitative aspect of the model was limited due to the approximations in model design which include: uniform boundary conditions, geometrical approximations, variable flow field in oven cavity, K-ε model wall function estimates, and limitations of computational grid

refinement [10].

Also discussed is the experimental and computational fluid dynamics (CFD) analysis of thermal air flow distribution in a 3-zone forced convection bread-baking oven. The oven is controlled at different constant temperatures within each zone and a CFD model is validated against experimental data. CFD results prove that selection of the flow model as well as realistic boundary conditions will result in accurate temperature predictions [6].

To ensure bread quality, moisture content should be taken into consideration throughout the CFD simulations. Mondal [11] studies the quality of bread and crust formation in relation to water and how this can be controlled using spraying of water at prefixed intervals. It uses CFD modelling for better understanding of the process of bread baking (simulating both the mass and heat transfer from the bread). In the baking process there is a process of simultaneous exchange of heat, liquid water, and water between the product and its surrounding's in the baking chamber. The understanding of this exchange is very important for the optimization of the baking process and will complement the design and implementation of advanced control systems for industrial ovens.

Knowing that processing is the time needed to cook the food, bread is baked when its center reaches a temperature of 90-95 °C (Cauvain, 2003) noting that the middle section of the bread requires the highest time to reach this temperature. The crust develops at a temperature of 110 °C. It's also important to note that the moisture rate is very high in the bread at the beginning of the baking process (when surface temperature increase sharply) and tends to decrease as the baking process progresses. After all, temperature distribution plays a major role in the quality and the effective bread baking processes.

Khatir [12] uses CFD methodology and an optimization framework to analyze the optimization of 3D ovens of different sizes and flow conditions for desired features such as temperature uniformity. Temperature uniformity and velocity across the width of an industrial bread-baking oven are fundamental to ensure even heat circulation to each piece of bread. Air flows in the oven are analyzed using steady state Reynolds Averaged Navier-Stokes (RANS) equations for 3D flow. RANS and turbulence transport and energy equations are solved on ANSYS FLUENT 13.0. Grid independency analysis is performed along with CFD to obtain values of σT (Temperature Function) that are used to undergo temperature uniformity assessments. This paper does not incorporate moisture content and volume change, only temperature distribution. Findings indicate that the greater the value of ΔT the lower the optimum core bread temperature and the more time it takes to heat the bread. Thus, the optimum ΔT desired between the top of the bread and the baking chamber is 10 °C.

Moreover, Boulet [5] considers all the mechanisms of heat transfer (convection, conduction, and radiation). Major factors of interest are the air flow, oven load, humidity and baking time. The heat absorption of the product is of prime importance. It will be very high at the beginning when there are high temperature differences between the bread (that is at room temperature) and the fluid in the baking chamber and will decrease with time.

III. METHODOLOGY

After completing extensive literature review regarding the relevant topics related to this project. We visited KARAKI to acquire the required dimensions for the heat exchanger and the baking oven as well as the fluid properties (attached detailed test report in appendix).

1. Mass flow rate of flu gas (kg/s) 0.0020 kg/s
2. Inlet Temperature of the flu gas 800°C
3. Blower Speed (rpm: rotation per minute) 1400 rpm
4. Outlet Temperature of the flu gas 470°C
5. Material of the heat exchanger AISI 310s
6. Material of the oven AISI 430
7. Flu gas material 6%O_2, 7%CO_2, 5%H_2O, 78%N_2, 1%NOx, 2% SOx, and 1% Soot.

The first step of the CFD analysis is drawing the oven on SOLIDWORKS. A heat exchanger and a baking chamber make up the oven. The heat exchanger is made of an inlet, two collectors, twelve tubes, an outlet, and thirty-seven fins on the surfaces of the heat exchanger.

SOLIDWORKS Assembly (void drawing):
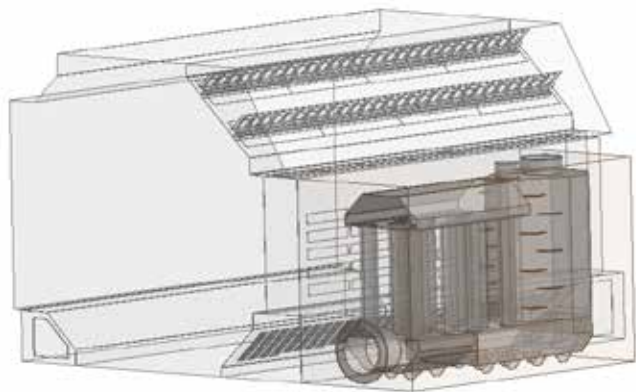
After completing the SOLIDWORKS drawing, the oven was imported to ANSYS FLUENT 16.2. The SOLIDWORKS drawing was made of void faces, thus the oven was then filled with flu gas (fluid) and the boundaries of the oven were covered with solid.

The next step of the CFD analysis is meshing the oven. The oven was meshed as one complete unit with inlet and outlet faces for the heat exchanger and inlet and outlet faces for the baking chamber.

The next step is inserting the flu gas, bread, and oven material properties in the setup and choosing the features (velocity variations, pressure variations, temperature variations, etc.) to be evaluated.

Simulations will be made with and without bread in the oven to get a better understanding of the oven's performance.

Simulations & Results are still in progress.

Heat exchanger solid:
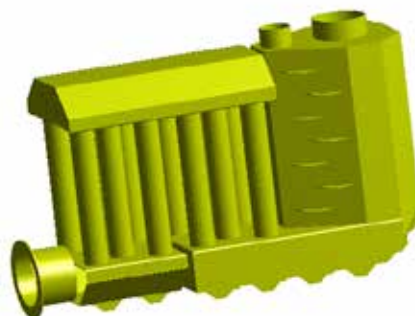
Heat exchanger flue gas:
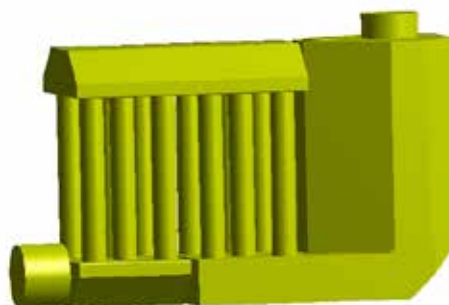
Full Assembly:
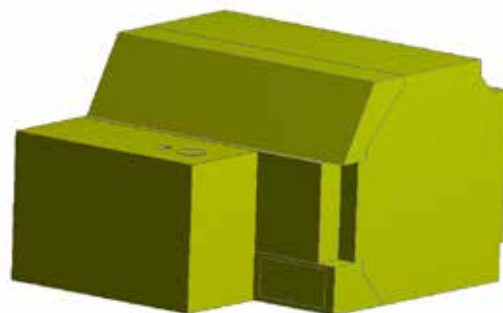
In the preprocessing stage we worked on understanding the physics and theory behind the problem at hand and understanding the objective constraints, assumptions to be used, the governing equations, and the conservation equations (mass conservation, momentum conservation, energy conservation, and radiation intensity conservation).

Also to better understand the finite volume theory that lie at the base of ANSYS Fluent we worked a simplified version of the code using MATLAB.

A significant part of the project deals with meshing the 3D computational model. This operation involves subdividing the domain to numerous elements and entering the necessary boundary conditions, fluid parameters and physical properties underlying the problem such as the flow rates and fluid characteristics of both fume gases and the air entering the baking chamber.

Then there is the processing stage (which is in progress) once the meshing process is complete. This process is repeated iteratively until convergence and until the required accuracy is achieved (this process is time and computationally consuming).

Finally, there is the post-processing stage (to be completed) where the data generated by the CFD software is compared with experimental data to evaluate the validity of the CFD model. In this stage the experimental data collection is of prime importance for validation where the temperature, humidity and other factors must be measured at certain specific locations in the oven.

IV. Experimental Process

The experimental investigation of the heat transfer in the oven represents all the temperature profile at the inlet and outlet at different circumstances. The measurements have been recorded with the assistance of KARAKI Company; they provided us with the required equipment for the experiment as well as the appropriate setup. The measurements include the oven internal and external temperatures at consecutive 2 mins time intervals. In the attached appendix there are full details of the temperature at each time in addition to all the details and the operating conditions of the Oven and the Burner.

In addition we have analyzed the oven performance by measuring the energy rate at the input and output, and the baking energy rate to get the efficiency of the baking oven and the overall production capacity. Moreover we have measured the fuel consumption, fuel density and flow rate at different stages of the baking process.

V. RESULTS

Simulations & Results are still in progress and will be presented in the poster.

VI. REFERENCES

1) Verboven, P., Scheerlinck, N., De Baerdemaeker, J., & Nicolaï, B. M. (2000). Computational fluid dynamics modelling and validation of the isothermal airflow in a forced convection oven. Journal of Food Engineering, 43(1), 41-53.

2) Therdthai, N., Zhou, W., & Adamczak, T. (2004). Three-dimensional CFD modelling and simulation of the temperature profiles and airflow patterns during a continuous industrial baking process. Journal of Food Engineering, 65(4), 599-608.

3) Wong, S. Y., Zhou, W., & Hua, J. (2007). CFD modeling of an industrial continuous bread-baking process involving U-movement. Journal of Food Engineering, 78(3), 888-896.

4) Williamson, M. E., & Wilson, D. I. (2009). Development of an improved heating system for industrial tunnel baking ovens. Journal of Food Engineering, 91(1), 64-71.

5) Boulet, M., Marcos, B., Dostie, M., & Moresoli, C. (2010). CFD modeling of heat transfer and flow field in a bakery pilot oven. Journal of food engineering, 97(3), 393-402.

6) Mistry, H., Dey, S., Bishnoi, P., & Castillo, J. L. (2006). Modeling of transient natural convection heat transfer in electric ovens. Applied thermal engineering, 26(17), 2448-2456.

7) Rek, Z., Rudolf, M., & Zun, I. (2012). Application of CFD simulation in the development of a new generation heating oven. Strojniški vestnik-Journal of Mechanical Engineering, 58(2), 134-144.

8) Litovchenko, I. (2013). The study of the baking ovens by computer simulation. Acta Universitatis Cibiniensis. Series E: Food Technology, 17(2), 107-114.

9) Verboven, P., Scheerlinck, N., De Baerdemaeker, J., & Nicolaï, B. M. (2000). Computational fluid dynamics modelling and validation of the temperature distribution in a forced convection oven. Journal of food engineering, 43(2), 61-73.

10) Khatir, Z., Paton, J., Thompson, H., Kapur, N., Toropov, V., Lawes, M., & Kirk, D. (2012). Computational fluid dynamics (CFD) investigation of air flow and temperature distribution in a small scale bread-baking oven. Applied Energy, 89(1), 89-96

11) Mondal, A., & Datta, A. K. (2010). Two-dimensional CFD modeling and simulation of crustless bread baking process. Journal of food engineering, 99(2), 166-174.

12) KHATIR, Z., Thompson, H., & Kapur, N. (n.d.). THE APPLICATION OF COMPUTATIONAL FLUID DYNAMICS (CFD) AND OVEN DESIGN OPTIMISATION IN THE BRITISH BREAD-BAKING INDUSTRY. 8th International Conference on CFD in Oil & Gas, Metallurgical and Process Industries.

13) Ozden, E., & Tari, I. (2010). Shell side CFD analysis of a small shell-and-tube heat exchanger. Energy Conversion and Management, 51(5), 1004-1014.

14) Zhang, M., Meng, F., & Geng, Z. (2015). CFD simulation on shell-and-tube heat exchangers with small-angle helical baffles. Frontiers of Chemical Science and Engineering, 1-11.

15) Chhanwal, N., Anishaparvin, A., Indrani, D., Raghavarao, K. S. M. S., & Anandharamakrishnan, C. (2010). Computational fluid dynamics (CFD) modeling of an electrical heating oven for bread-baking process. Journal of food engineering, 100(3), 452-460.

16) Xia, B., & Sun, D. W. (2002). Applications of computational fluid dynamics (CFD) in the food industry: a review. Computers and electronics in agriculture, 34(1), 5-24.

17) Codes and Standards | ICC public ACCESS. (n.d.). Retrieved November 5, 2015.

VII. APPENDIX

Oven Test

Manufacturer:	KARAKI (logo)	**"KARAKI"**
Address:		**Amrousieh-Choweifat (Industrial Area)** **Lebanon**
Model:		**CUTE Oven**
Type:	**Oil Fired Oven** — X	**Electrical Oven** —
Serial Number:		
Site:	"KARAKI" Factory (Choueifat)	**Date:** 26-Feb-16

1- Oven

CUTE OVEN	OIL FIRED
Proper thermal insulation	rockwool: 80 mm thickness 70 Kg/m3 density
Pre-heating starting tempearture	30°C
Pre-Heating tempearture setting	270°C
Pre-Heating tempearture difference	240°C
Pre-Heating time	20 minutes

2- Burner

Brand	RIELLO 40
Type	431T1
Model	G3
Thermal power - output	19 - 35 KW - 1.6 - 3 Kg/h
Fuel	Gas oil, max. viscosity at 20°C: 6 mm2/s (1.5 °E)
Electrical supply	Single phase, 230 ± 10% ~ 50HZ
Motor	Run current 0.7 A - 2850 rpm - 298 rad/s
Capacitor	4 µF
Ignition transformer	Secondary 8 KV - 16 mA
pump	Pressure 7 - 15 bar
Absorbed electrical power	0.115 KW
Flu gas exhaust section cm2	113 cm2
Combustion gas Velocity m/s	0.78 m/s
Flow m3/h	32 m3/h
Fuel injector capacity	0.65 gallon/hour
Fuel injector pressure	14 bars maximun
Fuel Consumption for pre heating	842 g
Fuel specific energy	48 MJ/Kg
Fuel Energy Consumption	40 MJ ≡ 11.2 KWh
(combustion analyser test results)	6 % O2
Exhaust flu temperature	470°C
Flame temperature	800°C

3-Pre-Heating Temperature gradient

Staring Time: 15:56 burner ON	Ambient temperature	20°C	Notes
Time	Oven internal temperature	Oven external temperature	
15:56	30°C	21.2°C → 24.2°C	
15:58	44°C	21.2°C → 24.2°C	
16:00	70°C	21.2°C → 24.2°C	
16:02	101°C	21.2°C → 24.2°C	
16:04	131°C	21.8°C → 25.0°C	The door glazing temperature increases from 21.2°C to 56°C during testing
16:06	159°C	21.8°C → 25.0°C	
16:08	186°C	21.8°C → 25.2°C	
16:10	210°C	21.8°C → 25.2°C	
16:12	232°C	21.9°C → 25.6°C	
16:14	253°C	21.9°C → 25.6°C	
16:16	270°C	21.9°C → 25.8°C	Burner OFF
16:18	277°C	22.0°C → 26.0°C	Oven Door opening and insertion of 7'500 g of dough
16:20	270°C	22.0°C → 26.0°C	
16:22	266°C	21.2°C → 24.2°C	Burner ON

4-Dough composition:

5'000 g Flour
2'500 g Water
Total: 7'500 g Dough
Dough were divided into 150 loafs
Each loaf weight is 50 g

5-Oven baking test

Ambient temperature	20 °C	Notes
Staring Time @ 15:56	Oven door closed	Burner ON
Time	Oven Internal temperature	
17:03	265 °C	
17:06	271 °C	Burner OFF
17:06	Time Interval Setting: **10 minutes**	
17:11	Exhaust fan ON	
17:16	Oven door opening and trolley pulling	
Fuel consumption for baking	175 g	
Fuel specific energy	48 MJ/Kg	
Fuel Energy Consumption	8.4 MJ ≡ 2.3 KWh	
Dough inserted weight	7'500 gram	
Ten Steel plates total weight	17'800 gram	
Fuel Energy Consumption per KG of dough	307 watt-hour/Kg	

Time Interval	2 hours

Oven Test

6-Idle Energy	Idle energy fuel consumption average (gram/h)	540	
	Fuel specific energy	48 MJ/Kg	
	Fuel Energy Consumption per hour	26 MJ ≡ 7.2 KWh	

7-Summary of CUTE Rack Oven performance:	Measured Energy Input Rate KW	22 KW	
	Measured Energy Input Rate BTU/h	75'134 BTU	
	Pre-heat Temperature	270 ºC	
	Pre-heat Time	20 minutes	
	Pre-heat Energy	40 MJ ≡ 11.2 KWh ≡ 38'216 BTU	
	Baking Energy Rate @ 270 ºC (KW)	13.8 KW	
	Baking Energy Rate @ 270 ºC (BTU/h)	47'130 BTU/h	
	Idle Energy Rate @ 270 ºC (KW)	7.2 KW	
	Idle Energy Rate @ 270 ºC (BTU/h)	24'590 BTU/h	
	Baking Energy Efficiency (%)	67%	Taking into consideration that a pre-heat occurs each two hours
	Production Capacity (Kg/h)	45 Kg/hour	
	Production Capacity (lb/h)	99 lb/hour	

8-Mass flow rate of flue gas Kg/s:		Baking process	For pre-heating process	Procedure
	Fuel combustion flue gas (m3/liter)	10.4	10.4	
	Fuel consumption average per hour (gram/h)	1,296	2,526	Fuel consumption (supposing that a pre-heating occurs each 2 hours): pre-heating: 842 g fuel /20 min Baking energy: 175 g fuel / 10 min ==> fuel consumption g/h = (842 +175*10) / 2 hours = 1'296 g/h
	Fuel density (gram/liter)	832	832	
	Fuel consumption average per hour (liters/h)	1.56	3.04	measured
	Volumetric flow rate of Flue gas m3/h	16.2	31.58	Calculated
	Density of flue gas @ 470 ºC (Kg/m3)	0.435		
	Mass flow rate of flu gas Kg/h	7.05		Calculated
	Mass flow rate of flu gas Kg/s	0.0020		Calculated

9-Blower fan air speed and volume:	Blower fan dimensions L x W (cm)	18.5 x 15.5 cm
	Blower fan section cm2	287 cm2
	Blower fan air speed m/s	18 m/s
	Blower fan volumetric flow m3/h	1'733 m3/h
	RPM	1390
	Power (KW)	0.75

10-Flu gas material (combustion analyser test results)	6 % O2	
	7 % CO2	
	0 % CO	
	5 % H2O	These figures are partially measured and assumed
	78 % N2	
	1 % Nox	
	2 % SOx	
	1 % Soots	

11- Instruments Used during test:	
Description	Reference
Infra-red thermometer	3M
Combustion Analyzer	UEI - C75
Energy check meter	VOLTCRAFT 3000
Digital Thermometer -50ºC to +300ºC	TP3001

Standards	ASTM F2093: Standard test method for performance of rack ovens General principles for design - Risk assessment and risk reduction

Prototype of a Real Life 3D Printed Hand

Diana El Hajj[1], Julie El Jurdi[2], and Magdalena Assaad[3]
[1,2,3]Department of Mechanical Engineering
American University of Beirut
Beirut, Lebanon
[1]die03@mail.aub.edu, [2]jfe09@mail.aub.edu, [3]mma163@mail.aub.edu

Abstract-**3D printing allows the production of three dimensional solid objects including various prototypes, end user products, biomaterials, tissues, and organs. Implementing 3D printing in medical applications is currently a challenge that promises to highly impact the future of medicine. The aim of this project is to integrate 3D printing of bones, cell culturing and artificial muscle fabrication to produce a prototype of a human hand. The fabricated hand is to be composed of three main components resembling the three main constituents of a hand: the bones, muscles and skin. The prototype is the first step towards a fully 3D printed functional one that allows amputees to regain their lost limb with an identical and natural one.**

I. INTRODUCTION

3D printing has provided a new platform where it is possible to transform virtual designs into physical 3D objects. These objects, whether prototypes or functional parts, have been integrated within many scientific studies [1]. All areas of medicine have also adopted this new platform to enhance quality of life of patients. The different uses of this technology can be seen in cardiology [2], blood vessels engineering [3], and prosthetics [1]. Another interesting and useful application is the 3D printing of bones. For example , the technology was applied to print a physical 3D temporal bone model from a computed tomography CT image [4].

In this project, we aim to use this technology to produce a real-life duplicate of the hand of an amputee. In order to achieve this goal, a CT image of a hand will be first obtained, mirrored and segmented using an appropriate CAD software to prepare it for 3D printing. After printing the bones and connecting them using cartilage-like material, fabricated electromagnetic muscles will be attached to their periphery to facilitate their movement. Before the final phase is initiated, the skeletal hand will be covered with an elastomer to take the shape of a real-life hand. Finally, 3D skin cell culturing will be performed in order to produce multiple skin layers that attach to the chosen elastomer.

The use and complexity of 3D printing in medical applications has been expanding with the development of possible 3D printing applications. For instance, integrating cell cultures with 3D printed scaffolds to closely mimic real-life bone function and structure have been done before [5]. However, due to the limited available facilities and materials, we had to resort to 3D printing of bones using non-biocompatible. The use and complexity of 3D printing in medical applications has been expanding with the development

of possible 3D printing applications. For instance, integrating cell cultures with 3D printed scaffolds to closely mimic real-life bone function and structure have been done before [5]. However, due to the limited available facilities and materials, we had to resort to 3D printing of bones using non-biocompatible material in our prototype. In what comes next, the methods used to produce the 3D prototype are explained in detail.

II. MEHODOLOGY

To construct the 3D printed hand, the process was divided into three phases, discussed below in order: (1) bones (and cartilage), (2) muscles, and (3) skin (and mold).

A. BONES

The production of the printed bones begins with a computerized tomography (CT) scan of a hand. As advised by two radiologists at the American University of Beirut Medical Center, the optimal CT image is obtained when the hand is in a flat position. Moreover, previous scans for medical purposes were all also taken in that position [6, 7]. The CT scan was taken at the American University of Beirut Medical Center with a resolution of 0.625 mm. Even though this resolution is not ideal for our application, the radiologist reconstructed it with a resolution of 0.4 mm, which was enough to capture the cartilage between the bones (seen as empty gaps between the different phalanges in figure 1). This scan provides an accurate imaging of the needed tissues and allows us to transfer this image to a biomedical CAD software. In this case, MIMICS was used for the analysis, segmentation and surface finishing of the bones. A preview of the segmentation of the bones is shown in figure 1.

The model is then exported to a stereolithography (.stl) file to be printed. The advantage of STL is that it facilitates the geometry by reducing it to its initial components. However, the object loses some of its resolution because triangles are used to represent the geometry [8]. This disadvantage can be fixed by secondary image manipulation in order to create well-defined contours [9]. After software processing is done, the image is transferred to a 3D printer which prints the bones using a thermoplastic polymer: acrylonitrile butadiene styrene (ABS). Previous use of ABS as a bone replacement material showed satisfactory results based on the feedback of surgeons on its anatomic characteristics [9]. The properties of the ABS composition that will be used are shown in Table 1. The 3D

printer is ProJet 3510 SD with a resolution of 375x375x790 DPI (xyz), layers of 32μm thickness, and an accuracy of 0.025-0.05 mm.

The bones are then separated and cleaned from the wax produced by the 3D printer. To avoid contamination and in preparation for the next step, the bones are rubbed with isopropyl alcohol. When the bones are ready, specific grade of silicone (is used as articular cartilage to reduce the friction at joints between the phalanges of the fingers, and simultaneously adheres them appropriately. To make sure that the bones are stationary during the process in order to avoid any cracks in the connection, a fixture is used to keep them in place while the silicone cures. Multiple alternative materials and grades of silicone have been tested, and the optimal adhesive/cartilage-like grade of silicone was used.

Property	Condition	Specification
Density	ASTM D4164	1.02
Tensile strength	ASTM D638	42.4
Tensile modulus	ASTM D638	1463
Elongation at break	ASTM D638	6.83
Flexural strength	ASTM D638	49
Heat distortion temperature	ASTM D648 at 0.45MPa	56

Table 1: Properties of the ABS plastic used in 3D printing

Figure 1: CAD image of the hand obtained from a CT scan

Figure 2: ABS printed bones

B. MUSCLES

When the skeletal hand is obtained, the fabricated muscles must be inserted and connected to facilitate the movement of the fingers. Two different methods are tested for this purpose: a mechanical approach and a magnetic approach. The idea of using electricity as a motion initiator is inspired by the fact that the human body already possesses the ability to generate a certain voltage. However, this voltage requires amplification in order to match the power sources (6-12V) required for the mechanical and electromagnetic muscles, which is out of the scope of this project.

The mechanical approach includes the use of servo-motors attached to each finger through nylon fishing lines and controlled through an Arduino. The advantage of using a servomotor is the fact that the position of the shaft is known and regulated using an encoder that supplies feedback [10], which is suitable for mimicking the selective grip of the hand. A servo motor should be capable of holding a given load (its maximal power should be larger than the peak power required for the task), satisfying the volume and weight criteria [11].

For this project, the team opted for an E-flite super sub-micro s60 servo, due to its light weight (6.0 g) and small size. This servo requires a voltage of 5V, which means that once connected to the body, an amplifier should be integrated in order to increase the human body's voltage (10-100 mV) to 5V. This servo has a torque of 0.9 Kg-cm.

Figure 3: An E-flite super sub-micro S60 drvo motor [12]

The servo is connected to both a power supply (the black wire connected to the negative pole and the red wire to the positive pole) and a data acquisition device (the NI myDaq by connecting the orange wire to pin 5 of the DIO pins) while control is done using a LabView program.

Once the program runs, the servo motor shaft goes back to position 0, and when the knob is turned, the shaft turns along. The maximum turn is reached once the knob is fully turned and the shaft has made a 180 degree angle.

This method ensures precise and controlled movement and a good torque for the future use of the fingers.

The second magnetic method offers the advantage of remote actuation. It involves attaching the bones to a cobalt rod (also using nylon fishing lines) centered in an electromagnet formed by winding an enamel coated copper wire around an iron rod, such that when a current passes through the coil, it attracts the cobalt rod causing the fingers to bend. An iron core serves to magnify the magnetic field up to one hundred times more than a coil without a core, and several experiments were done prior to choosing iron as the core material, including steel and lead. Before fixing the muscles in place, any moving wire or rod is inserted in a tube in order to prevent any interference from the added layers that hinders motion. For optimal performance, the power delivered should be available in high currents but low voltage reducing the magnet's resistance [13]. Studies show that the most favorable coil configuration is that

of a solenoid electromagnet, as it can operate at larger distances than other configurations [14]. After a series of experiments, two final configurations were tested taking into account the size limitations: 1,500 turns of a 1mm coil wound around a 10cm long iron rod (diameter=6mm), and 7,000 turns of a 0.1mm coil wound around a 13cm long iron rod (diameter=6mm). The coil with the larger diameter showed slightly better results and is recommended for use. As for the power source, different voltages and currents were tested to ensure a good range and strength of the electromagnet (knowing that increasing the current while decreasing the voltage increases the efficiency of the electromagnet), and the most suitable values were a voltage of 5.8V and a current of 0.97A that resulted in a range of 1.5cm. Adding a neodymium magnet to the cobalt rod increases the range to 4cm and is highly recommended. The force required to bend the finger was also measured by gradually adding weights and observing the bending pattern of the finger and the mold. The results showed, using Newton's Second Law, that a force of 5N is enough for a good grip.

Figure 4: Two types of electromagnets fabricated- 0.1mm diameter to the left and 1mm diameter to the right

The magnetic field is calculated using the following equation:

$$B= \mu NI/2\pi r$$

where B is the magnetic field, N is the number of turns of the coil, r is the radial distance and μ is the permeability of free space. Thus, as implied from the formula, through increasing the number of turns and the current, and decreasing the radial distance of the turns, it is possible to obtain a very strong magnetic field that is capable of carrying very high loads.

The nylon fishing lines serve as tendon replacements in structure and function. Two different configurations were tested, depending on the attachment of the fishing line to the fingers, specifically the phalanges. A compromise between the available space and the attainable degrees of freedom of the fingers had to be made. On one hand, providing the maximum degrees of freedom for the movement of the fingers, or attaching a nylon line to each phalange, secures natural-like movement. However, the feasibility of attaching the lines on all the phalanges is aesthetically complex and mechanically difficult. On the other hand, attaching one line to each finger consumes less space, but provides only one degree of freedom.

Three methods of attachment between the bones and the tendon-like lines were also tested. The first method requires drilling the bones to insert the lines in the hole, which could lead to weakening of the bones. The second method involves inserting the lines into the silicone cartilage before it dries; however, this could hinder motion. The third and most feasible method is to simply tie the lines around the phalanges. Since knots made by nylon lines are known to be unstable and loose, silicone will be added to secure stability and adhesion.

C. SKIN

After the muscles are fully connected and protected, the structure must be covered with a mold that replaces the tissues of the hand (e.g. fat). In order for the mold to adopt the shape of a hand, a cast will be formed using alginate. Alginate powder is thoroughly mixed with water with a ratio of 4:1 (water weight to alginate powder weight); this ratio is usually indicated on the package. The hand is wetted and immediately placed in the mixture maintaining the required open and flat position of the fingers. After around 10 minutes, the alginate solidifies and the hand is removed carefully, revealing the exact shape the mold needs to take. Next, the skeletal hand along with the attached muscles are inserted into the cast and precisely laid down, after which silicone is poured in to fill in the remaining volume. Silicone is left to dry guaranteeing satisfactory mimicry of the texture, dimensions and volume of the hand. Another layer composed of an FDA approved elastomer (medical rubber) is added to the mold which prevents any cell leakage from the final layer to penetrate through the lower layers.

As for the final layer, skin fibroblasts are cultured on different meshes and in various methods to determine the most reliable combination. Previous studies have shown that a PLGA (poly(DL-lactic-co-glycolic-acid)) mesh is an efficient scaffold for the culturing of fibroblasts; however, combining the mesh with collagen results in a higher success and proliferation rate [15]. The two types of meshes tested are prolene polypropylene meshes and PLGA meshes. Polypropylene meshes were tested due to their availability, low cost, and similarity to PLGA meshes. Since the available meshes have a large pore size,As for the final layer, skin fibroblasts are cultured on different meshes and in various methods to determine the most reliable combination. Previous studies have shown that a PLGA (poly(DL-lactic-co-glycolic-acid)) mesh is an efficient scaffold for the culturing of fibroblasts; however, combining the mesh with collagen results in a higher success and proliferation rate [14]. The two types of

meshes tested are prolene polypropylene meshes and PLGA meshes. Polypropylene meshes were tested due to their availability, low cost, and similarity to PLGA meshes. Since the available meshes have a large pore size, which allows the cells to seep through, the meshes will either be placed on top of each other so that the pores become relatively blocked, or the cells will be mixed with collagen prior to being added to the mesh. Finally, the cultured skin is sewn onto the hand, providing a final protective and aesthetic layer.

III. RESULTS AND DISCUSSION

After evaluating the alternative approaches mentioned in the methodology section, the most suitable approach for each layer will now be discussed in detail.

The printed ABS bones are to be placed on a fixture made from a modeling compound that holds the bones in place. Silicone is then added in between them and slightly extended to their periphery to increase adherence. The grade of silicone was chosen due to its flexibility as well as its relative adherence strength, knowing that silicone is not the best adhesive material.

As for the optimal muscle configuration, although servomotors provide a more precise motion and higher stability; the existence of a rotating shaft in the hand is very inconvenient (due to produced noises and vibrations) and sensitive. Thus, the servomotor needs to be engulfed in a protective box which requires space and defies the purpose of the hand being as close to reality as possible. On the other hand, using the electromagnet provides the hand with the ability to carry high loads through increasing the number of turns of the coil. In addition, it requires less space and provides smooth movements. The challenge resides in increasing the effective range of the magnetic field while maintaining a reasonable size in order to be able to operate it at larger distances; in other words, to increase the distance of the magnetic pull and thus finger bending. Solutions include increasing the diameter of the coil and changing the electromagnet's geometry. The decision between the two methods is made depending on the plausibility of the challenge presented above. It is worth noting that currently, electroactive polymers are under research and are great candidates for artificial muscle in the future.

A different grade of silicone, along with an alginate cast, will be used as a mold since it does not adhere to the other layers, preventing interference. The mold also possesses texture and flexibility properties similar to that of a real life hand. Finally, the decision regarding which method and mesh to use for skin cell culturing will depend on the results of the experiments which are still under the process. Maintaining the skin requires keeping the hand in an incubator, but since this is only a prototype, this issue will not be addressed. However, in real life, after the hand is fabricated and covered with the skin layer, its attachment to an arm will allow the culture to grow and become one with the person's own skin.

IV. CONCLUSION

Although the end product is a prototype, it will still serve as one of the starting points for wholly 3D printed hands in the future. Merging three different fields of biomedical engineering accentuates both the novelty of the project and the challenges faced. The fully functional system, comprised of 3D printed and synthetic musculoskeletal components protected with a layer of cultured skin, leaves room for possible advancement and development. For instance, multiple breakthroughs have provided the ability to integrate 3D printing with living cells.

In short, a hand comprised of 3D printed bones, muscles, and skin can be fully biocompatible and functional. The potential progression of the project is based on the development of 3D printing and its applications, particularly in the biomedical field.

REFERENCES

[1] A. Bandyopadhyay, S. Bose and S. Das, "3D printing of biomaterials," *MRS Bull.*, vol. 40, no. 02, pp. 108-115, 2015.

[2] S. Mishra, "Application of 3D printing in medicine," *Indian Heart Journal*, vol. 68, no. 1, pp. 108-109, 2016.

[3] C. Pinnock, E. Meier, N. Joshi, B. Wu and M. Lam, "Customizable engineered blood vessels using 3D printed inserts," Methods, 2015.

[4] J. Cohen and S. A. Reyes."Creation of a 3D printed temporal bone model from clinical CT data," *Am. J. Otolaryngol*. 36(5), pp. 619-624. 2015.

[5] C. Piard, Y. Chen and J. Fisher, "Cell-Laden 3D Printed Scaffolds for Bone Tissue Engineering," *Clinic Rev Bone Miner Metab,* vol. 13, no. 4, pp. 245-255, 2015.

[6] S. Nasab and M. Pipelzadeh, "Osteoid Osteoma of Proximal Phalanx of the Index Finger of the Right Hand," *Open Journal of Medical Imaging,* vol. 1, no. 2, 2011, p. 50-52.

[7] M. Inam et al., "Myositis Ossificans of Hand," *Gomal Journal of Medical Sciences,* vol. 8, no. 1, 2010.

[8] Y. Azem, P. Ikonomov, and P. Fleming, "Application of 3D Printing for Human Bone Replacement," *ispe-usa.com,* 2014.

[9] S. Mowry*et al.*, "A Novel Temporal Bone Simulation Model Using 3D Printing Techniques," *Otology and Neurotology,* vol. 36, no. 9, 2015, p 1562-1565.

[10] "Patent US4808898 - Gripper Assembly for Robot Arm". *Google Books*, 2016.

[11] H.J. Van De Straete et al. "Servo Motor Selection Criterion for Mechatronic Applications". *IEEE/ASME Transactions on Mechatronics*, vol. 3, pp. 43-50, 1998.

[12] Servo, E-flite. "E-Flite 6.0 Gram Super Sub-Micro S60 Servo [EFLRS60] | Electronics". Heliproz.com. N.p., 2016. Web. 17 Apr. 2016.

[13] F. Biter, "The Design of Powerful Electromagnets Part IV. The New Magnet Laboratory at M. I. T," *AIP Scitation*. 2016.

[14] Sudhoff, Scott D. *Power Magnetic Devices*.

[15] G. Chen, T. Sato, H. Ohgushi, T. Ushida, T. Tateishi and J. Tanaka, "Culturing of skin fibroblasts in a thin PLGA collagen hybrid mesh." *Biomaterials*, vol. 26, no. 15, pp. 2559-2566, 2005.

Smart Trash Can

Jamil Ballout
Gabriel el Hakim
Ezzedine Hashisho
Georges Bekhaazi
Mechanical Engineering Department
American University of Beirut
Jmb26@mail.aub.edu
geb07@mail.aub.edu
gme16@mail.aub.edu
emh09@mail.aub.edu

I. EXECUTIVE SUMMARY

The operation and cost of waste sorting at the sorting plants is a major problem with respect to environment and cost. This report describes and evaluates a possible solution using auto-waste- sorter for the home use. The report is intended to provide a background about the waste problem based on surveys and statistical data and how our project, the auto-waste-sorter, is going to contribute to the solution, the design of this machine is going to be discussed in details later in this report.

Many studies attempting to solve this problem showed that sorting at home would result in a much more efficient recycling process. Numerous attempts at initiating an "at home" recycling program have been made by Sukleen. They have proposed distributing recycling bins and trash bags every month to households and collecting them frequently, but these attempts were not successful due to the unresponsiveness of the Lebanese people.

Although the waste is sorted at the sorting plant after the garbage is collected in huge mount, this method is inefficient, because it only sorts a few percentage of the recyclable material. With the use of the auto-waste sorter, this will help sorting material at home before reaching the dumpster.

.

II. INTRODUCTION

On the global scale hundreds of thousands of tons of garbage are produced every day, for example in Beirut and Mount Lebanon solely produce 3500 ton/day. With the increasing awareness of environmental impact of burying such amount of garbage underground, the world has turned in the early 1990's to the recycling, to reduce the impact on the environment and to benefit economically from the recyclable material. But the problem was how to sort these materials? At first it started using huge machines and labors to sort, then later households started sorting at home, thus the cities started providing the households with a recycling bin, this was only in the advanced countries like United States and Europe. However in the third world this doesn't exist, for example in Lebanon households doesn't sort at home, and the garbage collector (Sukleen) sorts only 10% and compost only 10 % of the amount collected thus the remaining 80% goes to the landfills. Even in the advanced countries, not all garbage is sorted due to the huge variety of waste and the time consumption of the process.

A. Motivation

With this current situation the only solution is to start sorting at home where the mixing of garbage starts, however who guaranties the accuracy of this process? For example if we provide the households with a three recycle bins of each material, do we guarantee that they will sort at home? According to the sorting plant manager at Sukleen, he said "we tried to provide 3 bins to the households to sort at home, but they were irresponsive". In addition, people are lazy in nature, they do not care about this issue unless a governmental regulation imposes sorting at home, and if imposed it is very hard to control. Therefore the only solution is an Auto-Waste- Sorter that does the job accurately at home.

III. DESIRED NEEDS

With this solution, sorting at the plant is eliminated and the cost of garbage collection is reduced. In addition, it provides new energy sources and, without doubt, saves resources due to efficient recycling, which leads to a cleaner and healthier environment.

IV. PROPOSED SOLUTION/METHODOLOGY

The main challenge in our project is to find a way to identify all different kinds of trash. Therefore, step one would be coming up with a solution to the identification task.

This solution is based on the theory of natural frequency. In fact, each sound wave is a result of an object's vibration. Vibration could be due to any disturbance of an object; in our

case, it is the collision of the object with the bottom of the trashcan. When this object vibrates, it tends to do so at a particular frequency know as the natural frequency of the object, which is mainly dependent on the intrinsic properties of the object in question and slightly affected by a variety of factors such as the shape, defects, temperature and humidity. If the vibration of the object is significant, meaning that its collision was powerful, the amplitude of this natural frequency will be high enough to be detected.

V. PROGRESS DESCRIPTION

Our design is based on a translational system. We choose this type of system because it is very easy to manufacture. The body material is Plexiglas. We will use only a glue to join 2 parts and this will make our design easy to manufacture. A microphone sensor place on the plate will identify the 3 main waste materials:

The object is dropped to the plate at constant altitude. Based on material selection, the plate will be lifted upward to the corresponding material compartment. Power is transmitted to the plate by a motor through a pulley-belt system. Next, the squeegee will drop the object through the corresponding material compartment and it will remove any sticky material from the plate. This mechanical design would be common to the main design and the one alternative design.

Our preliminary electrical system design that is used to identify the objects' material consists of:

1-directional microphone: is used to detect object's material from its impulsive sound.

2- Microprocessor: MYDAQ would be our processor.

3-Programming: Labview is used for programming
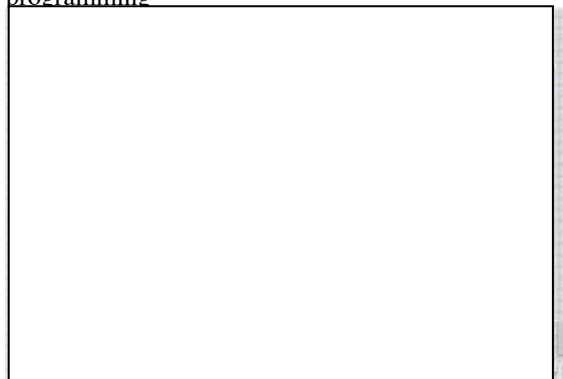
Fig. 1 (above) (Showing the resulting frequencies and the total bandwidth produced due to collision)

Fig.2(showing the developed program)

A. Design Alternative

It would be the same of the preliminary design, but we will add two more sensors in order to minimize the error as much as possible, these sensors are: 1- Inductive sensor to detect metals 2- capacitive to detect water containing objects (organic).

A. Preliminary Implementation

After completing the mechanical design, we will start building our model to test its functionality and feasibility. Based on this preliminary implementation we would be able to identify the errors and modify our design accordingly. In addition, in this phase we will test the functionality of each electrical component especially sensors and motors, because the response time of these components are of supreme importance for our project, because we are intending to make our system as fast as possible.

A. Preliminary Testing

This phase was the most important one; because we were able to decide which system we have to use in order to make our project as efficient as possible. At first we have tested the functionality of the inductive sensor to check its reliability in identifying the metallic objects and we have found that it is 100% reliable. Second we have tested the capacitive sensor, we have found that this sensor detects any water containing object like fruits, but this sensor did not meet our expectations to differentiate between organics, plastics and metals, because the plastic and metallic containers sometimes might contain water that are left over after the consumption. Therefore this setup is not reliable enough to perform the job.

After obtaining the result explained above we have switched our design for object's material identification to use microphone, MYDAQ and Labview Program. We have conducted several experiments on the microphone setup, and we came up with very promising results. The experiments showed that the microphone setup is able to identify the objects' material thrown with an error of approximately 15%. The results are shown in the table below, noting that we will conduct more experiments on the microphone setup to improve the error through programming and tuning. The table below

shows some experimental data with an approximate of 15% based on experiments done:

Type	Max dB	Min dB
Metals	-48.68	-67.16
Plastics	-68	-74
Organics	-75.26	-85.46

VII. BUDGET

Components	Quantity	Cost
Capacitive Sensor	1	$50
Inductive Sensor	1	$50
Microphones	1	$100
Reed Switch Sensor	1	$5
Stepper Motor	2	$20
Container	1	$20
Electronic components	….	$50
Power Supply (Transformer)	1	$10
Belt and Pulley Setup		$50
Gears	3	$30
Machine Elements	….	$50
Seals	….	$5
Testing	….	$200
Total		**$640**

These quantities are based on a tentative design.

VI. CONCLUSION

This invention is an opportunity for us to welcome a better world, where the environment and the resources are kept safe and available. It brings more back more benefits, which definitely compensate for the cost. For future progresses, coupling more sensors will yield a better, and much more effective integrated system.

The Development of a 3D Low Computational Cost Non-Invasive Tool to Predict Disease Functional Significance in Coronary Arteries

Iyad Fayssal, *PhD Candidate*

Mechanical Engineering Department, American University of Beirut, Riad El-Solh / Beirut 1107 2020
iaf04@mail.aub.edu

Abstract- Recently, non-invasive predictive techniques for diagnosing ischemia in diseased coronary arteries have shown to be promising, however not yet fully formulated and validated. Few studies based on coupling lumped-parameter coronary models of the downstream vascular bed with patient specific 3D CT scan data have been recently reported to evaluate coronary computed tomographic angiography-FFR. Though these studies "claiming" to be "non-invasive", however, on the contrary are not. Such models, though comprehensive, require invasive measurements of patients' time-dependent pressure and blood flow waveforms to tune the lumped-elements of the coupled downstream coronary impedance. Hence, these approaches can't be translated into future clinical practices. In this research, the problem has been viewed from a different perspective, not reported on previously. A new boundary condition has been formulated and coupled to a 3D computational model to account for the myocardial blood flow demand by a downstream vascular bed. This boundary condition relates different patient dependent parameters which can be non-invasively extracted for each patient. The current formulation allows a "purely" non-invasive diagnosis of ischemia for suspected patients with coronary artery disease under steady state conditions rather than conducting full transient simulations, while preserving the same level of accuracy. This further allows computing the level of ischemia with minimum time and thus translating it into future clinical practices. The solution methodology starts by constructing the geometry of the diseased artery via image processing techniques. The physical domain is then decomposed into small elements forming the basis of the numerical platform. While previous studies have adopted allometric scaling laws to quantify total coronary flow and myocardial resistance, an effort was put in this study to develop a new comprehensive physical-based approach to quantify the arterial-venous impedance at the outlet boundary of the truncated diseased vessel. This starts by quantifying the total baseline coronary hypothetical artery flow to the left myocardium following different approaches. The specific coronary artery flow for each artery is then computed based on geometric characteristics of the target vessel to be simulated. Then, the arterial-venous impedance of the modeled vessel is computed based on the central pressure, taking into account the reduction of arterial-venous impedance during maximal hyperemia. The set of conservation equations with the developed boundary condition formulation are then solved numerically via a collocated pressure-based finite volume method following a segregated approach. To validate the designed numerical method, available experimental data for dog patients were considered. Predictions of normalized blood volume flow rates in healthy and diseased left circumflex arteries were compared with 240 experiments performed on 12, 17 to 40kg, black Labrador dogs simulated under rest and maximal hyperemic conditions. Similar behavior was obtained between the numerical and experimental results (computed R^2 values for a selected single dog model were 0.9290 and 0.9504 for rest and hyperemic conditions respectively). For all dog models, third order polynomials were initially fitted to the numerical normalized mean flow predictions for each simulated single dog, with R^2 values ranged from 0.9376 to 0.9850 and 0.9804 to 0.9847 for rest and hyperemic conditions respectively. A single average correlation was then obtained for each of the rest and hyperemic simulated cases in order to compare with those generated by the study of Gould et al. The computed R^2 values of the obtained average numerical based correlations were 0.8809 and 0.9733 for the rest and hyperemic conditions respectively. Extensive investigations were then done on the validated dog models and constructed idealized and actual human single and branch arterial models by varying various physiologic parameters. A criterion was then developed to identify patients prone to higher levels of ischemia. As a summary, the future significant value of the designed method lie in its potential to (1) replace the traditional experimental method which is based on intrusive process and (2) be done at low computational cost. The numerical tests performed on dog patients and human arterial models have proven viability and were promising towards translating the developed method into future clinical, research and educational applications.

Keywords: ischemia, coronary arteries, non-invasive, finite volume method, impedance.

I. INTRODUCTION

Coronary heart disease is considered one of the leading causes for human death worldwide. In 2004, heart syndromes claimed approximately 7.2 million lives, which represented 12.2% of the world deceased population, with about 82 % of these deaths occurring in low and middle-income countries [1]. A widespread heart disease is the coronary artery disease (CAD), which results from an insufficient supply of oxygen to cardiac tissue beds. CAD is typically caused by the excessive accumulation of atheromatous plaques and fatty deposits within certain regions of the arteries leading to a restriction of blood flow. Several techniques have been existing for detecting CAD. The most widely adopted one is the invasive coronary catheterization (Cath) procedure that provides 2-D images of coronary arteries showing diseased segments and lesions from which a geometric estimate of the obstruction (% stenosis diameter) is obtained. However, Cardiac Cath, the backbone of diagnosing CAD for decades, suffers a certain degree of operator subjectivity in estimating the extent of the disease and its physiologic importance. More pertinently, the percent stenosis may often misestimate the true functional dependent myocardial perfusion effect of a certain lesion and hence the need for intervention. Computed Tomographic Angiography (CTA) is another adopted technique that clinicians use in assessing non-invasively the severity of CAD relying only on geometric characterization of images without discerning whether a stenosis causes ischemia or not. Accordingly, relying only on anatomic measures captured via angiography or

coronary computed tomographic angiography (CCTA) does not reflect the in-vivo functional hemodynamic significance of the stenosis [2]. Such difficulties in characterizing the possibility for myocardial ischemia relying only on CTA or 2-D Cath images alone suggested the need to look for additional physiological parameters that aid in detecting the hemodynamic significance of CAD. In-vivo measurements of fractional flow reserve (FFR), an indicatory index developed by Pijls and De Bruyne [3], was shown to correlate with the disease physiologic importance. This is the solely physiological parameter used to determine whether the stenosis impedes oxygen delivery to the heart muscle (ischemia) by measuring the ratio of the mean distal to mean proximal pressure across the stenosed region. Diseased coronary arteries with FFR values ⩽0.8 (FFR = 0.8 means 80% of the blood flow that should be crossing a healthy artery is crossing the stenosed artery) are considered prone to ischemia and require percutaneous coronary intervention (PCI). This technique used in coronary catheterization is considered as a measure of lesion-specific ischemia, providing the fact that the diagnostic performance of such a parameter still unexplored deeply. Moreover, this measuring technique is costly, time consuming and incurs complicated procedures as it requires invasive intervention through a costly pressure-sensor guide-wire in the diseased artery and administration of a vasodilator for inducing hyperemia. Therefore, this invasive technique is still not widely adopted in clinical practice (as reported in the recent US CathPCI Registry [4], less than 10% of stenosed arteries were evaluated by FFR). Because of that, non-invasive comprehensive predictive techniques involving advanced numerical methods have recently gained attention. In specific, the use of Computational Fluid Dynamics (CFD) techniques combined with patient specific data has shown promising, "yet unresolved", outcomes in predicting in vivo hemodynamics for future diagnostic purposes [5, 6]. With the state-of-the-art developments in computer technology, such numerical techniques are expected to become essential tools in future clinical applications by providing an incremental value to stand-alone coronary imaging on one hand and in-vivo FFR on the other. Recently, coronary computed tomographic angiography–FFR (CCTA-FFR) has been developed as a non-invasive tool for functional assessment of the coronary tree. This technique relies on applying CFD coupled with lumped models to capture the systemic circulation and coronary microcirculation [7, 8]. In the study presented by Taylor et al. [8], the implemented types of boundary conditions including the lumped-parameter heart model, the Windkessel model, and the lumped-parameter coronary models are patient dependent and requires invasive measurements of flow properties (time-dependent pressure and blood flow waveforms) to tune the lumped elements presented in their boundary condition formulation. This is totally comprehensive in terms of the physical definition of the problem; however, "for designing a generic non-invasive tool for ischemia prediction", such approach could not be clinically adopted as it requires pre-knowledge of the transient pressure and flow waveforms in a healthy situation for the specific patient.

Similar to the study of Taylor et al. [8], Min et al. [9] investigated the diagnostic accuracy of noninvasive FFRCT on 238 patients for the purpose of detecting ischemia-causing stenosis. In their study, FFR-CT was computed from acquired CTA images and compared to a reference FFR measured invasively. Calculations were performed for the full left and right arterial trees, thus increasing the computational solution cost. Similar studies are found in [10-12]. Morris et al. [13] developed a computer model to predict myocardial fractional flow reserve for nineteen patients with stable coronary artery disease. In addition to the implementation of the lumped Windkessel model, transient simulations were performed to compare with measured FFR. Besides the limitation of the patient dependent lumped model, the adoption of a transient analysis resulted again in a high computational cost. Papafaklis et al. [14] presented an approach for virtual functional assessment index (vFAI) of coronary stenoses, using three-dimensional quantitative coronary angiography (3D-QCA) and blood flow simulation. However, their approach requires two simulations to be performed at two different imposed values of the volume flow rate to compute the unknown coefficients of pressure loss due to viscous friction and due to flow separation. In addition, their suggested method was limited to a medium-sized population and requires several computational steps for assessment. Moreover, their method doesn't provide a direct assessment of ischemia level, but rather, pre-specified uniform flow values were used; thus the mass and functional capacity of the distal myocardial bed which affect coronary flow were not incorporated into their model.

In this research, the problem has been viewed from a different perspective, not reported on previously. A new boundary condition has been formulated and coupled to a 3D computational model to account for the myocardial blood flow demand by a downstream vascular bed. This boundary condition relates different patient dependent parameters which can be non-invasively extracted for each patient. The current formulation allows a "purely" non-invasive diagnosis of ischemia for suspected patients with coronary artery disease under steady state conditions rather than conducting full transient simulations, while preserving the same level of accuracy. This further allows computing the level of ischemia with minimum time and thus translating it into future clinical practices.

II. Methodology

A. Physical Model Construction

In an actual situation, the solution process starts by constructing the geometry of the artery using image processing techniques. In this study, simulations were performed on real patient models and idealized geometries of arterial segments (healthy and diseased branch and single arterial models are considered) constructed using actual "physiological" dimensions of human coronary arteries (Fig. 1). Though arteries with the highest degree of stenosis severity (the ones with Maximum %DS > 80 %) is rarely encountered in clinical applications, they were used to show the theoretical capability

of the developed model in predicting extreme situations. Branch arterial models are also considered, constructed from human anatomical physiological dimensions.

(a)

(b)
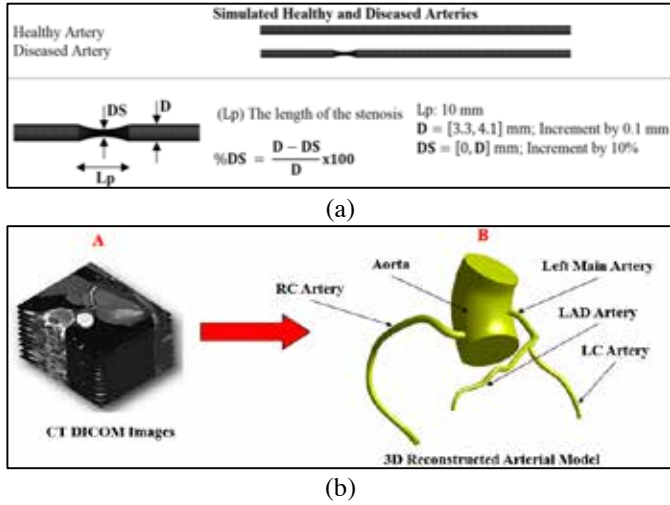
Figure 1. (a) Geometries of the simulated idealized healthy (hypothetical) and diseased arteries; (b) Reconstructed geometries of real patient coronary arteries.

B. Governing Equations

To predict the level of ischemia in a diseased artery, blood hemodynamic properties (i.e. pressure and velocity fields) are first computed by solving the Reynolds-averaged Navier-Stokes equations (continuity and momentum) given by

$$\frac{\partial \rho}{\partial t} + \nabla . (\rho \mathbf{v}) = 0 \tag{1}$$

$$\frac{\partial (\rho \mathbf{v})}{\partial t} + \nabla . (\rho \mathbf{v} \mathbf{v}) = -\nabla p + \nabla . (\mathbf{\tau} + \mathbf{\tau^R}) + \mathbf{F_b} \tag{2}$$

Where $\mathbf{v}, p, \rho$, and $\mathbf{F_b}$ represent the velocity vector, pressure, density, and body force per unit volume, respectively. $\mathbf{\tau}$ is the deviatoric stress tensor, which, for a Newtonian fluid is expressed as

$$\mathbf{\tau} = \mu\{\nabla \mathbf{v} + (\nabla \mathbf{v})^T\} - \frac{2}{3}\mu(\nabla . \mathbf{v})\mathbf{I} \tag{3}$$

$\mathbf{\tau^R}$ is the Reynolds stress tensor which within the Boussinesq hypothesis can be written as

$$\mathbf{\tau^R} = \mu_T\{\nabla \mathbf{v} + (\nabla \mathbf{v})^T\} - \frac{2}{3}\mu_T(\rho k + \nabla . \mathbf{v})\mathbf{I} \tag{4}$$

Where μ_T is the turbulent viscosity computed using the shear stress transport turbulence model of Menter (i.e. $k - \omega\ SST$).

C. Materials and Boundary Conditions

Results reported in this paper are generated by modeling blood as an incompressible Newtonian fluid with density of 1050 kg/m^3 and dynamic viscosity of 0.003 Pa.s. An objective of this work is to estimate the level of ischemia by performing computations in the diseased artery only without including other arterial connections. This should be reflected by the solution process and starts by truncating the domain of interest at specific locations and separating it from the upstream zone and downstream vascular beds. To account for downstream effects, a new boundary condition physics-based model is developed

and coupled to the three-dimensional domain at the outlet boundary. Even though the hemodynamics of human vasculature is a 3D complex and time dependent phenomenon, however, the level of ischemia in a suspected patient could be estimated from steady state conditions (further computational-based evidences will be supported in other section of this manuscript). The interpretation of this goes to the fact that the total blood volume flow rate supplied to a vascular bed during a specific duration (T) is equal to the integral of the instantaneous blood flow rate supplied to the vascular bed over a period T. In other words, is the vascular bed receiving the required blood volume flow rate to avoid ischemia? This is independent of weather this amount is received as an average value or continuously over a period T. This concept is mathematically expressed by the following relation:

$$\bar{Q} = \frac{1}{T} \int_{t=0}^{t=T} Q(t)dt \tag{5}$$

The lumped model relating the static pressure and mean blood volume flow rate crossing the outlet boundary is then expressed by

$$P_{static,outlet} - P_v = R_{a-v}\bar{Q} \tag{6}$$

Where R_{a-v} represents the arterial-venous resistance of the truncated downstream vascular bed of the modeled vessel and P_v is the venous pressure.

Quantification of total left ventricular myocardium coronary flow ($Q_{T,h}$), specific coronary artery flow ($\bar{Q}$), and arterial-venous resistance (R_{a-v})

While previous studies have adopted allometric scaling laws to quantify total coronary flow and myocardial resistance, an effort was put in this study to develop a new comprehensive physical-based approach to quantify the arterial-venous resistance, R_{a-v}, at the outlet boundary of the truncated diseased vessel. This starts by quantifying the total baseline coronary hypothetical artery flow to the left myocardium, ($Q_{T,h}$). The specific coronary artery flow ($\bar{Q}$) for each artery is then computed based on geometric properties of the target vessel to be simulated. Then, according to the patient central pressure, the corresponding arterial-venous resistance, R_{a-v}, of the modeled vessel is computed, taking into account the reduction due to maximal hyperemic conditions. Several mathematical models were developed to quantify the arterial-venous resistance at maximal hyperemia given as,

$$R_{a-v} = \left[P_{central}/F_f(a_1X_1, a_2X_2, \ldots, a_{15}X_{15})\right] \times \frac{1}{hf} \tag{7}$$

hf is the hyperemic factor, considered generally to be between 0.2 and 0.3 [15]). F_f are functions developed to quantify the specific coronary blood flow demand. X_i are patient related parameters, including for instance left ventricular geometric properties, blood hemoglobin level (Hgb), patient's weight and height, etc… a_i could take a value of 0 or 1 depending on the mathematical formulation of the function to be used. The inlet section of the domain is modeled using a total pressure formulation, with the static pressure computed as

$$p_{static} = p_{total} - \frac{1}{2}\rho\mathbf{v}.\mathbf{v} \qquad (8)$$

Coupling these prescribed inflow and outflow conditions to the 3D domain, allows characterizing the myocardial perfusion of the modeled vessel from a single simulation. In the context of this paper, this boundary condition formulation will be sometimes referred to "scenario #0".

D. Numerical Solution Procedure

The set of conservation equations (Eqs. 1-4) subject to the boundary conditions (Eqs. 6-8) discussed above is solved numerically via a collocated pressure-based finite volume method following a segregated approach. Solutions are obtained by subdividing the computational domain into a finite number of elements, each associated with a grid point placed at its centroid. The partial differential equations are first integrated over each element and with the aid of the divergence theorem, the volume integral of the convection and diffusion fluxes are transformed into surface integrals along the element faces. Then, fluxes are approximated and transformed into algebraic relations using interpolation profiles. The diffusion flux is discretized using a linear profile, while the second order upwind scheme is adopted for the discretization of the convection flux and is applied in the context of the NVSF methodology [16] and implemented using the deferred correction approach of Khosla and Rubin [17]. For the transient simulations, the unsteady term is discretized using a two-step Adam-Bashforth method leading to a second order implicit transient scheme. The integral value of the source term over an element is evaluated by assuming the source at the cell center to be equal to the mean value over the whole element. The resulting system of algebraic equations is then solved using an iterative algebraic solver [18, 19]. To evaluate the pressure field, a pressure correction is defined and a pressure correction equation is derived by combining the momentum and continuity equations as in the SIMPLE procedure of Patankar [19]. A collocated grid is used and checkerboard pressure and velocity fields are suppressed through the use of the Rhie-Chow interpolation technique [18, 21]. The model at the outflow boundary (i.e. scenario #0) is carefully implemented to damp spurious numerical reflections into the three-dimensional computational domain and reduce the calculation time needed to achieve a stable converged solution. Numerical experimentations indicated that under relaxing the pressure at outlet greatly enhances numerical stability and accelerates convergence.

III. NUMERICAL MODEL VALIDATION

Predictions of normalized blood volume flow rates in healthy and diseased left circumflex arteries were obtained based on 240 experiments performed on 12, 17 to 40kg, black Labrador dogs from the experimental study of Gould et al. [22] simulated under rest and hyperemic conditions. Fig. 2 demonstrates the normalized mean flow predictions for a single dog mapped over the measured values of Gould et al. [22]. For the hyperemic flow predictions, results were multiplied by the hyperemic factor corresponding to increase after Hypaque injection. Similar

behavior was obtained between the numerical and experimental results (computed R^2 values were 0.9290 and 0.9504 for rest and hyperemic conditions respectively) where the normal hyperemic response become blunted when the %DS reached 35 percent narrowing and decreased markedly before the resting flow was affected (changes in resting mean flow from the hypothetical value was reported when %DS became more than 80%).
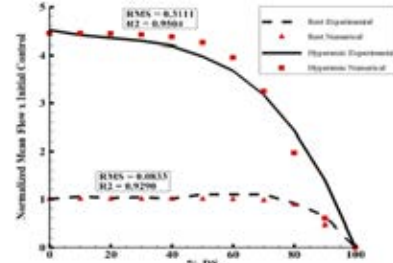
Figure 2. Predicted normalized mean flow for a single dog mapped over the measured data of Gould et al. [22]. For hyperemic flow, results were multiplied by the hyperemic factor as shown in figure.

To validate the situation for all dogs, third order polynomials were initially fitted to the numerical normalized mean flow predictions for each simulated single dog for rest and hyperemic conditions (Fig. 3), with R^2 values ranged from 0.9376 to 0.9850 and 0.9804 to 0.9847 for rest and hyperemic conditions respectively. A single average correlation was then obtained for each of the rest and hyperemic simulated cases in order to compare with those generated by the study of Gould et al. [22]. The computed R^2 values of the obtained average numerical based correlations were 0.8809 and 0.9733 for the rest and hyperemic conditions respectively.
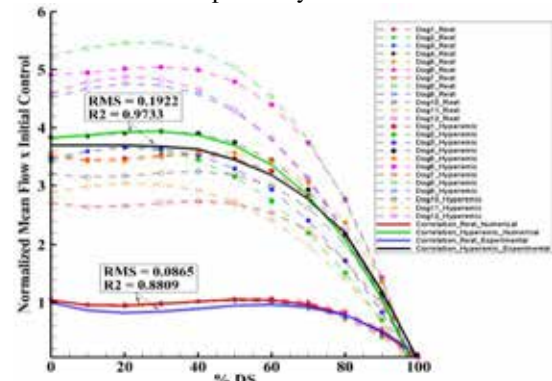
Figure 3. Comparison between the predicted average normalized mean flow and experimental ones for all dogs under rest and hyperemic conditions (Refer to solid lines in figure).

IV. NUMERICAL OUTCOMES

The capability of the developed numerical methodology to predict ischemia was first tested on idealized single healthy and stenosed arterial segments and branch models with diseased arterial segment (Numerical simulations were also conducted on actual patient coronary arteries that will be presented in later section of this paper). The criterion used to assess the performance of the model was based on comparing the ratios P_d/P_a and Q_s/Q_h, where P_a is the proximal pressure to the stenosis, P_d is the distal pressure, Q_s is the volumetric flow rate

in the diseased segment, and Q_h is the volumetric flow rate in the hypothetical (healthy) situation. Scenario #0 of boundary condition is applied and the models were primarily simulated under steady state conditions for rest and hyperemic conditions.

A. Idealized Healthy and Stenosed Arterial Segments

The left anterior descending (LAD) artery is the most commonly occluded of the coronary arteries. It provides the major blood supply to the interventricular septum. For that purpose, idealized healthy (hypothetical) LAD and stenosed arterial segments with plaque profiles of increasing stenosis severity were simulated under rest and hyperemic conditions and over a mean pressure range from 70 to 130 mmHg. The physical dimensions of the simulated arterial geometries are shown in Fig 1. The arterial-venous resistance R_{a-v} was set at 120000 and 26664.4 dynes.s/cm^5 to induce the rest and hyperemic responses respectively. The ratio of blood volumetric flow rate in a stenosed artery to that of a hypothetical healthy arterial segment (Q_s/Q_h) was compared with the ratio of distal pressure (P_d) to proximal one (P_a) in the stenosed artery (P_d/P_a) for rest and hyperemic conditions. An excellent agreement was obtained (Fig. 4). Thus, the mathematical model and implemented type of boundary conditions have revealed reliability in predicting ischemia in diseased arteries. This confirms well with the experimental measurements obtained by Pijls and De Bruyne [3] and the capability of the developed numerical methodology and the implemented inflow and outflow boundary conditions in estimating the level of ischemia in diseased coronary arteries. However, it should be noted that the equality ($Q_s/Q_h = P_d/P_a$) was satisfied due to the fact that pressure drop in a healthy (hypothetical) situation is negligible (A mathematical proof of this issue was developed in this study based on fluid dynamics theory, and extensive simulations were done to provide physical evidences; however, for compactness it will not be presented in this manuscript). The advantage of the numerical (non-invasive) model will be shed light on in a later section of this manuscript. It will be shown in the following section (Effect of Boundary Conditions Formulation) that these outcomes doesn't always hold on when applying different types of boundary conditions formulations.
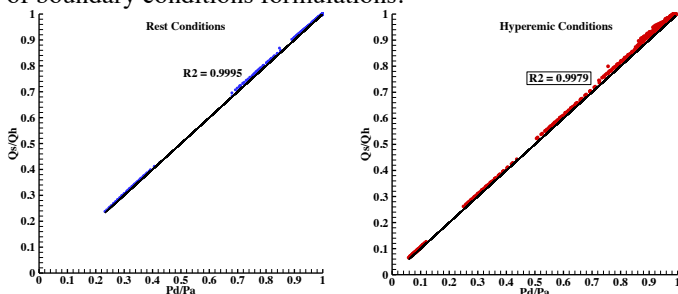
Figure 4. Graph shows the correlation between the predicted P_d/P_a and Q_s/Q_h ratios for all simulated cases.

B. Effect of Boundary Condition Formulation

In this section, a light is shed on the effect of boundary conditions formulation on solution prediction. The main purpose of this section is to provide computational-based evidences on the importance of choosing a correct boundary condition

formulation that accurately capture in-vivo real situations. Two different scenarios of boundary conditions formulations that are generally adopted in most theoretical investigations of blood hemodynamics were investigated and applied on single healthy and stenosed arterial segments. For compactness, only a brief description of each type of boundary condition and resulting predictions will be presented.

Scenario #1: Fixed pressure outlet combined with total pressure formulation at inlet

In this scenario, a fixed outlet pressure is adopted combined with a total pressure formulation at the inlet sections of the simulated domains. Two idealized single arterial segments of respective lumen diameters of 3.3 and 4.1 mm and %DS from 0 to 90 % were simulated. The total pressure at the inlet sections of the simulated arteries was fixed at 12,100 Pa while the outlet sections were modeled with a fixed pressure outlet with values of 12,000, 9,000, and 6000 Pa (These values were chosen based on the known physiologic range that the mean pressure could fall within for all possible situations of healthy and/or diseased coronary arteries). Results depicted in Fig. 5 show discrepancies between the ratios P_d/P_a and Q_s/Q_h as compared to the outcomes obtained from scenario #0. This boundary condition formulation (Scenario #1) was shown to overestimate the level of ischemia in diseased arteries (only the level of ischemia in diseased arteries with %DS < 12% were well estimated by this scenario for rest and hyperemic conditions). The predictions obtained by this scenario do not account for the downstream effects induced from the myocardium resistance which always force the pressure to decrease in association with blood volume flow rate reduction. Therefore, "apparently", this scenario could not be adopted as a "direct assessment tool" for ischemia in human coronary arteries, though it's possible that this scenario could inherit a non-linear virtual assessor of ischemia, left questioned for future possible investigations.
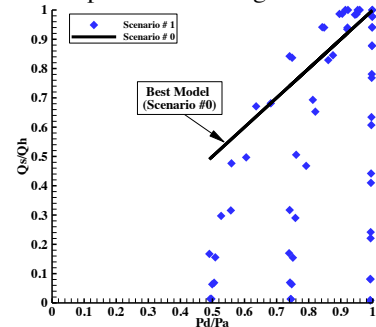
Figure 5. Q_s/Q_h vs P_d/P_a when applying scenario #1 of boundary condition formulation.

Scenario #2: Pre-specified mass flow rate and static pressure at inlet combined with a zero diffusion flux outflow model

In this scenario, pre-specified mass flow rate and static pressure were adopted to model the inlet section of the domain combined with a zero diffusion flux model considered at the simulated domains' outlet sections. Two idealized single arterial segments of respective lumen diameters of 3.3 and 4.1 mm and %DS from 0 to 90 % were simulated. For the arteries of

lumen diameter of 3.3 mm, pre-specified mass flow rate values of 0.001150804 and 0.004830245 kg/s were set at the inlet with static pressure fixed at 100 mmHg. These mass flow rate values were obtained from the rest and hyperemic conditions for a healthy case using Scenario #0 and were used in the current scenario as "pre-specified" values to simulate the healthy and diseased cases. Similar setup was considered for the other artereis. Since mass flow rates were pre-specified, then same mass flow rate will cross the healthy and diseased arteries. Accordingly, the ratios Q_s/Q_h will remain fixed and could not be compared with P_d/P_a. The predicted ratios of P_d/P_a obtained by scenarios #0 and #2 were then considered for comparison. Under rest conditions, Scenario #2 was capable of predicting ischemia level for diseased arteries of %DS $\leqslant$ 70% and 80% for arteries of 3.3 mm and 4.1 mm lumen diameters, respectively (Fig. 6a), while it fails to accomplish that for the other cases. On the other hand, under hyperemic conditions, the performance of this scenario decreased compared to the outcomes obtained from the rest conditions. As shown in Fig. 6b, this scenario was capable of predicting ischemia level for diseased arteries of %DS $\leqslant$ 40% and 50% for arteries of 3.3 mm and 4.1 mm lumen diameters, respectively, while it over predicted it for the other cases. The reason why Scenario #2 acts much more better than Scenario #1 in most of the simulated cases is that a fully developed flow assumption at the outlet section of the domain is valid and thus inherits the physical condition of Scenario #0; however, one disadvantage of this scenario is that it doesn't allow to evaluate the loss in blood volume flow rate (ischemia level) due to the use of pre-specified mass flow rate values at the inlet section of the domain and over predicts ischemia.
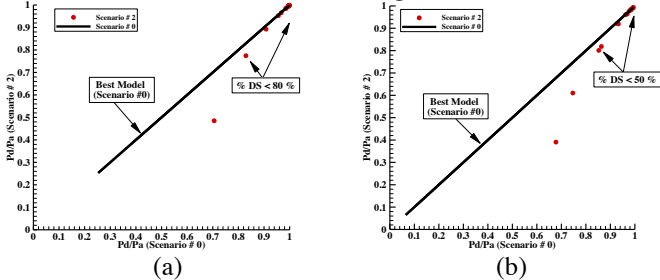
Figure 6. Q_s/Q_h vs P_d/P_a when applying scenario #2 of boundary condition formulation. (a) Rest Conditions; (b) Hyperemic Conditions.

C. Branched Arterial Models

Even though one of the goals in this paper is to isolate the diseased arterial segment and simulate it separately, numerical investigations were also performed on healthy and stenosed branch arterial models to validate the performance of the developed methodology in capturing the level of ischemia in interconnected arterial branches. Fig. 7 shows the predicted pressure color code map for healthy and stenosed branch models respectively simulated under rest and hyperemic conditions. Predictions of blood volume flow rate crossing the branch arterial segments were mapped over the pressure contours. The developed method has proven high performance in estimating the cardiac perfusion and its capability to predict hemodynamic characteristics. Under rest conditions, the pressure and blood

volume flow rate distribution were barely affected in the stenosed branch (Fig. 7b) compared to the hypothetical (healthy) branch model (Fig. 7a), with the ratios Q_s/Q_h and P_d/P_a were matched ($Q_s/Q_h = P_d/P_a = 0.984$). On the other hand, blood volume flow rate has dropped by 10% from the healthy situation in the stenosed segment (left artery) of the arterial branch when simulated under hyperemic conditions accompanied by a difference in pre and post stenotic pressures, with the ratio of distal (P_d) to proximal (P_a) pressure in the diseased segment ($P_d/P_a = 82.11 / 92.76 = 0.8852$) matches well the ratio of blood volume flow rate in that segment of the branch to that of the healthy situation ($Q_s/Q_h = 188.2905 / 209.1747 \cong 0.9$).
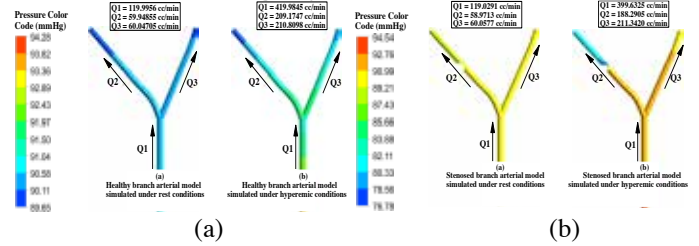
Figure 7. (a) Pressure map and predicted blood flow rates for a healthy branch; (b) Pressure map and predicted blood flow rates for a stenosed branch.

V. STRATEGIES FOR FUTURE CLINICAL APPLICATIONS

Reducing computational cost in biomedical applications while preserving the accuracy of the solution prediction is a challenging issue and indispensable for future clinical applications. However, the success in achieving this target greatly reflects the adoption of the method and translating it into practical clinical applications. Two effective strategies were developed in this study for the sake of reducing computational run time for solution prediction. The first strategy is based on simulating the diseased artery under specified mean conditions (i.e. performing steady state analysis) rather than conducting a full transient simulation. The second strategy is based on isolating the diseased arterial segment under investigation from its branch model and simulating it separately while retrieving the accuracy of results (i.e. level of ischemia) as if it was connected to its main branch model.

A. From Full Transient to Mean Condition Simulation

The first strategy to reducing computational run time cost is shifting towards simulating the arterial segment under mean conditions rather than undergoing detailed transient simulations. In order to prove the validity of the concept, transient and steady state (steady state = under mean conditions) simulations were performed for healthy and diseased arterial segments under rest and hyperemic conditions. Time dependent inflow and outflow boundary conditions were applied to simulate the transient hemodynamics, while computed averaged inflow and outflow boundary conditions were adopted for the steady state simulations. A newly time dependent impedance function was introduced to model the effects of the downstream vascular bed on the truncated domain. This impedance function is considered a characteristic function of the blood flow demand required by the vascular bed and thus is used for all simulated healthy and stenosed arterial models under rest and hyperemic conditions. The pressure distribution and blood volume

flow rate in a diseased artery will be predicted for same imposed demand conditions and their corresponding values will depend on the stenosis severity. On the other hand, for steady simulations at mean conditions, the time dependent impedance function is replaced by a fixed arterial-venous resistance (R_{a-v}) and averaged values of the hypothetical time dependent pressure and blood volume flow rate are considered based on the mean value theorem of integrals. The arterial-venous resistance, R_{a-v}, and total inlet pressure were then evaluated based on the computed mean pressure and volume flow rate. The ratio P_d/P_a and FFR (P_d/P_a = FFR at hyperemic conditions) computed from the transient simulation was compared to that obtained from a steady state simulation. Predictions of P_d/P_a from transient and steady state mean conditions were conformed for rest and hyperemic conditions. Extensive simulations were performed on large number of arteries, however, for compactness, results for one selected diseased LAD artery will be reported. Under rest conditions, the predicted ratio P_d/P_a for the simulated diseased artery was 0.9554 when performing a transient simulation, while it was 0.957 under steady state mean conditions (Fig. 8 and 9). Under hyperemic conditions, the predicted ratio P_d/P_a was 0.868 and 0.865 under transient and steady state simulations, respectively. The required total number of iterations to simulate a hyperemic case of 1 period (T = 0.5s) was 50,000 (not considering here the transient effects that were deleted) while it took less than 14000 iterations to simulate the model under mean conditions.
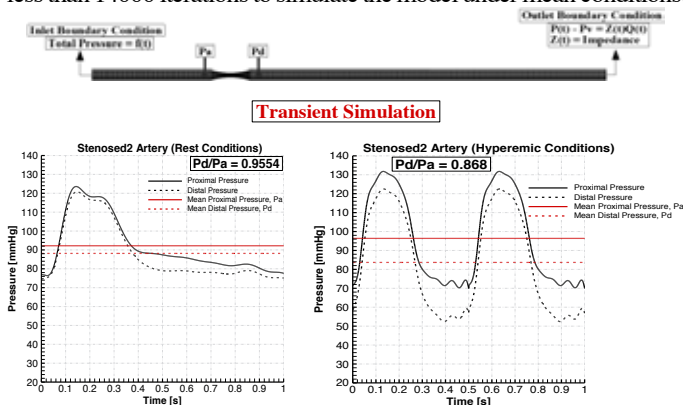
Figure 8. Variations of predicted proximal (Pa) and distal (Pd) pressures during rest and hyperemic conditions for a stenosed artery computed via transient simulation.
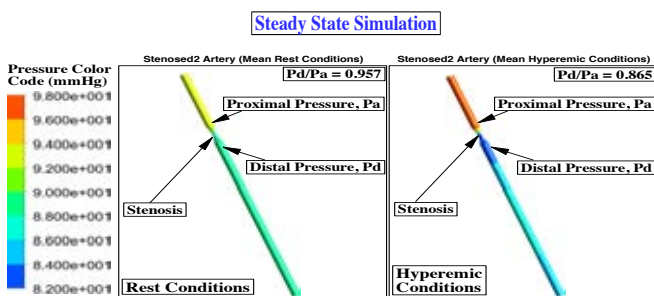
Figure 9. Pressure map (mmHg) showing predicted proximal (P_a) and distal (P_d) pressures during rest and hyperemic conditions for a diseased artery computed via steady state simulation.

In what have been presented in this section, the average value of the pressure that was used in the steady state simulations was computed based on the mean value theorem of integrals. In this study, several approaches were also evaluated to compute an average value of the inlet pressure by knowing few parameters that could be non-invasively retrieved for each suspected patient. 7 mathematical relations based on the mean arterial pressure (MAP) were used to compute the average inlet pressure (Fig. 10) and impedance at the outlet of the 3D truncated domain (These were referred to as, MAP #1, MAP #2, … MAP # 7, in the context of this manuscript). Extensive simulations steady state simulations were conducted and the performance of each mathematical relation was assessed by comparing the level of ischemia with the original one predicted by the full transient simulation; for compactness only one set of results is presented. Results in Fig. 11 are presented for a diseased artery of 57.1 % DS. The predictions obtained by all mathematical relations were reliable when comparing FFR values with the ones obtained by transient simulations under rest and hyperemic conditions. For rest conditions, all P_d/P_a ratios simulated in steady state environment were found to match well with that of the transient simulation with maximum relative error of 0.733% obtained by MAP #7. Under hyperemic conditions, MAP formulations # 5 and #7 have the highest relative error of 1.17 and 1.744% respectively, while all other formulations resulted in a maximum relative error of 0.69 % by MAP #4. According to the above findings, the established outcomes demonstrate promising significance towards adopting this strategy in practical clinical applications.
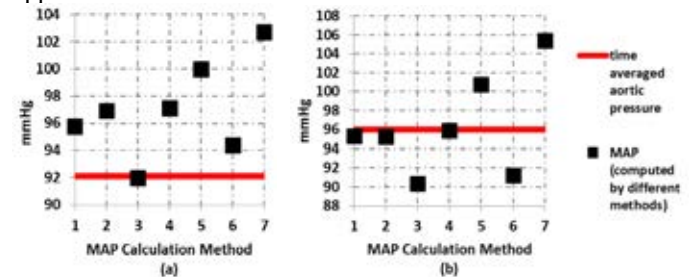
Figure 10. Computed average pressure by different MAP relations for (a) rest Conditions and (b) hyperemic conditions.
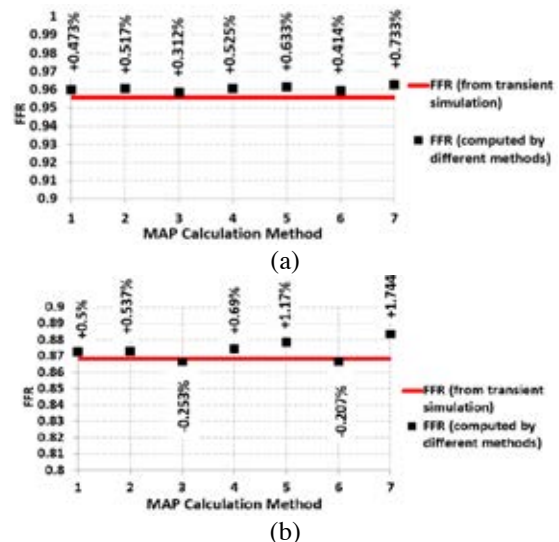
Figure 11. Predicted FFR by differen MAP relations for (a) rest Conditions and (b) hyperemic conditions.

B. Isolating the Diseased Arterial Segment

The second strategy to reduce the computational run time cost is based on isolating the arterial segment under investigation from its branch model (Fig. 12) and simulating it separately while retrieving the accuracy of the targeted outcome (ischemia level or FFR) as if it was connected to its original branch model.
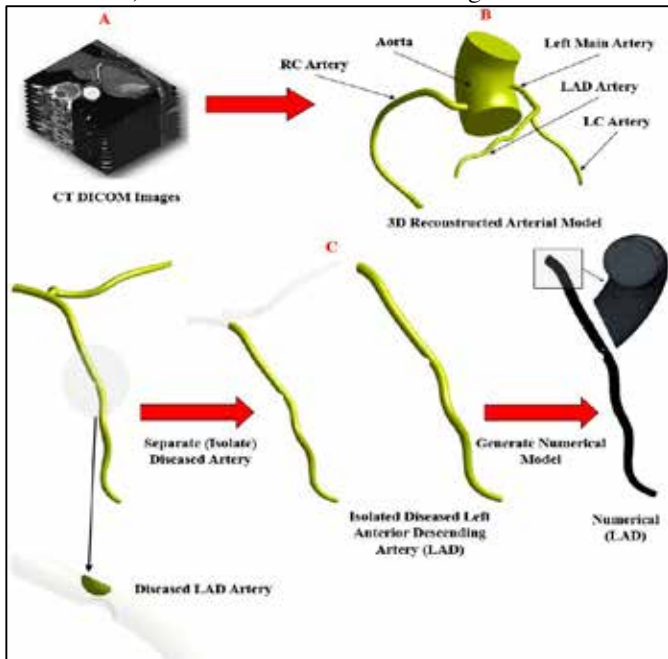
Figure 12. Arterial Isolation Strategy.

Though the intention is to lower computational cost, however, it was found that the ratio P_d/P_a of a certain lesion depends on whether only one segment is diseased in a branch or is simultaneously found with another diseased segment. Thus, this requires isolating the diseased segment from its original branch model and simulating it separately to quantify its level of ischemia and its effects on the downstream vascular bed. Further scientific evidence is depicted in Fig. 13. As shown in the figure, same geometric lesion found in a downstream artery of a branch model has, apparently, a different disease functional significance when found simultaneously with another upstream geometric lesion. For the single downstream diseased segment in the branch, the predicted FFR (0.91) value correspond to the actual one and is causing a loss in cardiac perfusion by 10%. On the other hand, when this diseased segment is found at the same time with another upstream diseased artery, the predicted FFR value of the downstream stenosed artery is increased to 0.965. This reading correspond to the apparent FFR and not the actual (real) one and it appears that this lesion is less significant; hence could not be relied on for decision making. It should be also noted that relying on the pressure gradient of each lesion will allow only "qualitatively" assessing which lesion is more significant than the other, but one cannot "quantitatively" guess the actual FFR for each lesion, unless simulated independently to remove the effects of multiple-lesions interference. This is among the driving factors to isolate the diseased artery and simulating it separately.
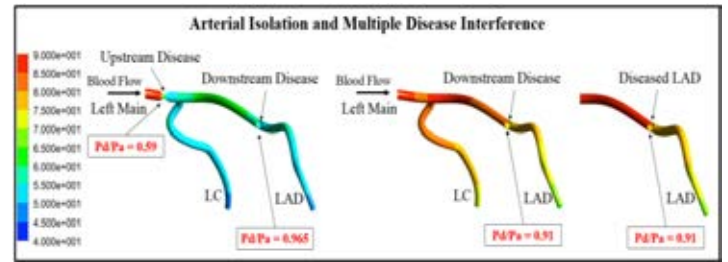
Figure 13. Same geometric lesion found in a downstream artery of a branch model has, apparently, a different disease functional significance when found simultaneously with another upstream geometric lesion.

VI. CRITERION TO IDENTIFY PATIENTS PRONE TO HIGHER LEVEL OF ISCHEMIA

An investigation was performed to determine the criterion that identify patients prone to higher levels of ischemia using the results of the previously validated dog models in this study. Based on the achieved outcomes presented in this manuscript, predictions have shown, for fixed downstream vascular impedance and inlet total pressure, arteries with minimum lumen diameter assumed minimum FFR. On the other hand, for a fixed lumen diameter and inlet total pressure, arteries with lower downstream vascular impedance achieved higher ischemia levels. Moreover, arteries with same lumen diameter and fixed downstream vascular impedance are prone to higher levels of ischemia when simulated under higher total inlet pressure conditions. Fig. 14 depicts the product of the LC diameter (D) and computed downstream myocardial impedance (R) for the 12 dogs simulated under rest and hyperemic conditions. Predictions of Q_s/Q_h ratios obtained under rest and hyperemic conditions for different inlet total pressure values (P = 90, 120, and 150 mmHg) were mapped over the product R*D in the same figure. Based on the inherited observations, Dog #2 has the minimum product of R*D while Dog #5 has the maximum one. The computed Q_s/Q_h ratios for Dog #2 for different %DS were the lowest with respect to that obtained for the other dog patients for rest and hyperemic conditions. For instance, the predicted Q_s/Q_h ratios for Dog #2 were 0.9858 and 0.876 when simulated under mean total inlet pressure of 90 mmHg and %DS of 60% for rest and hyperemic conditions respectively. On the other hand, the computed Q_s/Q_h ratios for Dog #5 were the highest among all dog patients for all disease geometric severity and rest and hyperemic conditions. Further, for same dog patient, the predicted ratio decreased with the increase of total inlet pressure (Fig. 14). In addition, the ratio Q_s/Q_h was shown to be more sensitive to the changes of the imposed value of the total inlet pressure for dog patients with the lowest R*D product (for example Dogs #2, 8, 9, and 12), while it was less sensitive for dog patients with high R*D product (for example Dogs #5 and 7). According to the aforementioned analysis, patients with minimum R*D/P factor possess the least Q_s/Q_h ratios and hence are prone to higher levels of ischemia. This criterion to identify patients prone to higher levels of ischemia was shown to be independent of the patient's body mass, arterial diameter (when considered solely),

and rest and hyperemic myocardial blood flow demand (when considered solely). For instance, Dogs # 2 and 9 possess almost same body mass, however, different ischemia levels (Q_s/Q_h = 0.849 and 0.888 for Dogs #2 and 9 respectively, simulated under hyperemic conditions at inlet total pressure of 120 mmHg) as depicted in Fig. 15. Moreover, Dogs # 2 and 12 have almost same hyperemic myocardial blood flow demand and LC diameter, however, possess different ischemia level (Fig. 15).

(a)

(b)
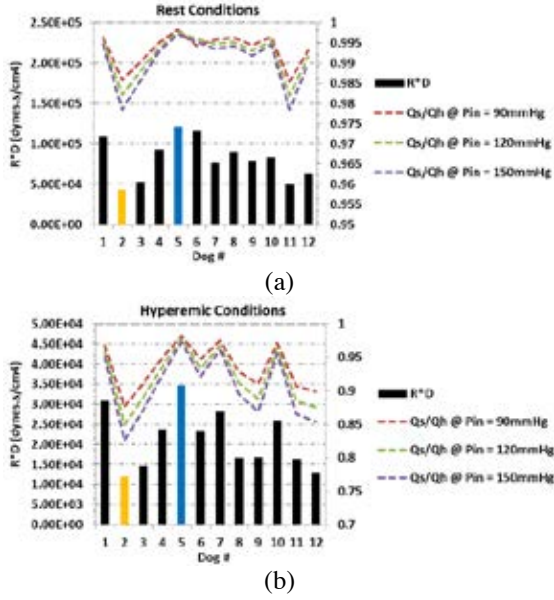
Figure 14. The product of the LC diameter (D) and computed downstream myocardial impedance (R) for the 12 dogs simulated under rest and hyperemic conditions. The figure also shows the variation of Q_s/Q_h ratio for all dog models.
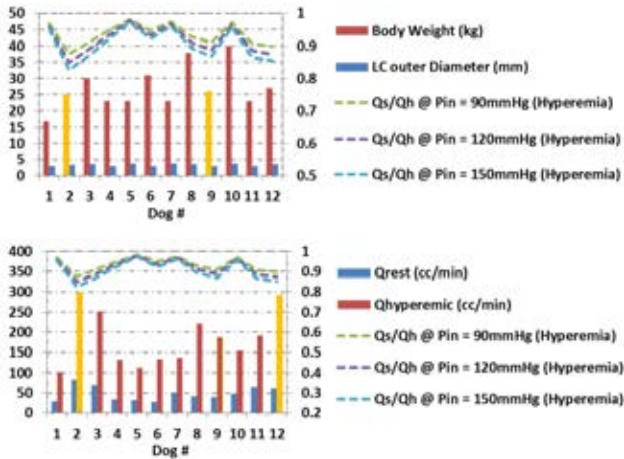
Figure 15. (a) Body weight, LC diameter and variation of Q_s/Q_h ratio for all dog models; (b) rest and hyperemic myocardial blood flow demand and variation of Q_s/Q_h ratio for all dog models.

VII. FUTURE PROSPECTS

An analytical model was developed to show a reason of the extent level of ischemia dependency in a diseased segment on the imposed inlet total pressure. The model relates blood volume flow rate to the total pressure at inlet section of the modeled domain and all other geometric and physiologic parameters. The model was developed for idealized healthy and stenosed arterial

segments of uniform cross sectional area and then generalized for an arterial segment of non-uniform cross sectional area, by applying the conservation laws of continuity and momentum. Extensive mathematical derivations were performed to generate the model prototype; for compactness issues, the details of which and test cases will not be presented in this manuscript. The final outcome of the derived model is expressed by:

$$\frac{\partial Q}{\partial P_{Total,inlet}} = \frac{1}{\sqrt{c_1 + c_2 c_3}} \tag{9}$$

$$c_1 = \left(-\frac{\mu}{\pi}\left[\sum_{i=2}^{i=n} F_i''\right] + R_{a-v}\right)^2 \tag{10}$$

$$c_2 = -4\left[\frac{4\rho}{\pi^2}\left[\sum_{i=2}^{i=n} F_i'\right] + \frac{1}{2}\rho\left(\frac{1}{A_{inlet}}\right)^2\right] \tag{11}$$

$$c_3 = \left[P_v + \mu\frac{\partial v_Z}{\partial Z}\Big|_{inlet} - P_{Total,inlet}\right] \tag{12}$$

F_i' and F_i'' are characteristic non-linear functions depending on geometries of arterial segment and stenosis; R_{a-v} is the arterial venous resistance; P_v is the venous pressure at the vascular bed; $\mu\frac{\partial v_Z}{\partial Z}\Big|_{inlet}$ is the normal stress at the inlet section of the modeled artery; $P_{Total,inlet}$ is the total pressure imposed at the inlet section of the simulated artery; ρ and μ are the blood density and viscosity respectively; A_{inlet} is the inlet artery cross sectional area.

VI. CONCLUSIONS

A new boundary condition has been formulated and coupled to a 3D computational model to account for the myocardial blood flow demand by a downstream vascular bed. The model relates different patient dependent parameters which can be non-invasively extracted for each patient. The current formulation allows a "purely" non-invasive diagnosis of ischemia for suspected patients with coronary artery disease under steady state conditions rather than conducting full transient simulations, while preserving the same level of accuracy. The designed numerical method was validated with available experimental data for dog patients. The future significant value of the designed method lies in its potential to (1) replace the traditional experimental method which is based on intrusive process and (2) be done at low computational cost. The numerical tests performed on dog patients and human arterial models have proven viability and were promising towards translating the developed method into clinical, research and educational applications. The aforementioned findings can significantly impact the applicability of this method and the clinical practice at large.

REFERENCES

[1] WHO Media Center, editor. The 10 leading causes of death by broad income group (2004). World Health Organization, October 2008.

[2] Melikian N, et al., 2010, "Fractional flow reserve and myocardial perfusion imaging in patients with angiographic multivessel coronary artery disease," JACC Cardiovasc Interv., 3(3), pp. 307–314.

[3] Pijls N.H., De Bruyne B., Peels K., 1996, "Measurement of fractional flow reserve to assess the functional severity of coronary-artery stenosis," N Engl J Med., 334, pp. 1703-1708.

[4] Dehmer GJ et al., 2012, "A contemporary view of diagnostic car¬diac catheterization and percutaneous coronary intervention in the United States: a report from the CathPCI Registry of the National Cardiovascular Data Registry, 2010 through June 2011," J Am Coll Cardiol.,60(20), pp. 2017-2031.

[5] Meijboom WB et al., 2008, "Comprehensive assessment of coronary artery stenoses: computed tomography coronary angiography versus conventional coronary angiography and correlation with fractional flow reserve in patients with stable angina," J Am Coll Cardiol., 52(8), pp. 636-463.

[6] Yong AS, Ng AC, Brieger D, Lowe HC, Ng MK, Kritharides L., 2011, "Three-dimensional and two-dimensional quantitative coronary angiography, and their prediction of reduced fractional flow reserve," Eur Heart J.,32(3), pp. 345-353.

[7] Kim HJ, Vignon-Clementel IE, Coogan JS, Figueroa CA, Jansen KE, Taylor CA, 2010, "Patient-specific modeling of blood flow and pressure in human coronary arteries," Ann Biomed Eng., 38(10), pp. 3195–3209.

[8] Taylor CA, Fonte TA, Min JK., 2013, "Computational fluid dynamics applied to cardiac computed tomography for noninvasive quantification of fractional flow reserve: scientific basis," J Am Coll Cardiol., 61(22), pp. 2233-2241.

[9] Min JK et al., 2011, "Rationale and design of the DeFACTO (Determination of Fractional Flow Reserve by Anatomic Computed Tomographic AngiOgraphy) study", J of Cardiovasc Comput Tomogr, 5(5), pp. 301–309.

[10] Koo BK et al., "Diagnosis of ischemia-causing coronary stenoses by noninvasive fractional flow reserve computed from coronary computed tomographic angiograms. Results from the prospective multicenter DISCOVER-FLOW (Diagnosis of Ischemia-Causing Stenoses Obtained Via Noninvasive Fractional Flow Reserve) study," J Am Coll Cardiol. 2011 Nov 1;58(19):1989-97

[11] Min JK et al., Diagnostic accuracy of fractional flow reserve from anatomic CT angiography. JAMA 2012;308:1237–1245.

[12] Norgaard, B. L. et al. Diagnostic performance of noninvasive fractional flow reserve derived from coronary computed tomography angiography in suspected coronary artery disease: the NXT trial (analysis of coronary blood flow using CT angiography: next steps). J. Am. Coll. Cardiol. 63, 1145–1155 (2014).

[13] Morris PD, Ryan D, Morton AC, Lycett R, Lawford PV, Hose DR, Gunn JP., 2013, "Virtual fractional flow reserve from coronary angiography: modeling the significance of coronary lesions: results from the VIRTU-1 (VIRTUal Fractional Flow Reserve From Coronary Angiography) study," JACC Cardiovasc Interv., 6(2), pp. 149-157.

[14] Papafaklis MI et al., 2014, "Fast virtual functional assessment of intermediate coronary lesions using routine angiographic data and blood flow simulation in humans: comparison with pressure wire – fractional flow reserve," EuroIntervention, 10(5), pp. 574-583.

[15] Wilson RF, Wyche K, Christensen BV, Zimmer S, Laxson DD. Effects of adenosine on human coronary arterial circulation. Circulation 1990;82: 1595-606.

[16] M. Darwish and F. Moukalled, "Normalized Variable and Space Formulation Methodology for High-Resolution Schemes," Numerical Heat Transfer, Part B, vol. 26, pp. 79-96, 1994.

[17] P. K. Khosla and S. G. Rubin, "A Diagonally Dominant Second-Order Accurate Implicit Scheme," Computers and Fluids, vol. 2, pp. 207-209, 1974.

[18] F. Moukalled, L. Managani, and M. Darwish, The Finite Volume Method in Computational Fluid Dynamics: An Advanced Introduction with OpenFOAM® and Matalab, Springer, Switzerland, 2015.

[19] S.V. Patankar, Numerical Heat Transfer and Fluid Flow, New York, Hemisphere Publishing Corporation, 1980.

[20] F. Moukalled and M. Darwish, "Pressure Based Algorithms for Single and Multifluid Flow," Handbook of Numerical Heat Transfer, second edition, W. J. Minkowycz, E. M. Sparrow, and J. Y. Murthy (eds.), pp. 325-367, Wiley, 2006..

[21] C.M. Rhie and W.L. Chow, "Numerical Study of the Turbulent Flow Past an Airfoil with Trailing Edge Separation," AIAA Journal, vol. 21, no. 11, pp. 1525-1532, 1983.

[22] Gould, K. L., K. Lipscomb, and G. W. Hamilton. Physiologic basis for assessing critical coronary stenosis. Instantaneous flow response and regional distribution during coronary hyperemia as measures of coronary flow reserve. Am. J. Cardiol. 33(1):87–94, 1974.

Towards a Safer Design of Helmets - Finite Element and Experimental Assessment

Sarah Siblini, Bilal Wehbi, Sari Kassar, Omar Abro
Department of Mechanical Engineering
American University of Beirut
Riad El Solh, Beirut, Lebanon
fyphelmet@gmail.com

Abstract- **Motorcycle helmets are required to protect the vulnerable head and maximize shock absorption upon impact. This research presents a new and plausible bio-inspired design affined to the foam liner material and structure in helmets. The proposed liner design is inspired from animal horn microstructure and tubule arrangement. An innovative drop-testing apparatus is presented with a spring-winch system for experimental testing. The aim is to validate the new design by meeting the ECE 22.05 standard for motorbike helmets using peak acceleration and HIC criteria, as well as confirming experimental results with finite-element simulations.**

Keywords- **foam-liner, helmet, shock absorption, brain trauma, acceleration curve, bio-inspired, drop test, finite element analysis, Head Injury Criterion (HIC)**

I. INTRODUCTION

The human head is profoundly vulnerable to multi-directional impacts and injuries mainly triggered by motorcycle accidents. Motorcycle helmets are considered the best preventive method for head injuries to date [1]. While helmets provide remarkable mitigation of head injuries, it proves crucial to assess the energy-absorbing capacities of foam liners in commercial helmets. Helmets typically consist of two parts: the outer shell and inner foam liner, such that the padding foam in contact with the wearer's head is not considered [2]. The shell is usually composed of an outer thermoplastic, acrylonitrile butadiene styrene (ABS), polycarbonate shell, or composite material such as fiberglass and Kevlar. Polymer foam liners are most commonly used for the inner liner such as expanded polystyrene (EPS) and expanded polypropylene (EPP) ranging from 30-40mm in thickness, and absorb the most impact energy (30-50%) [3], [4]. EPS, which is a closed-cell low weight foam, is used most commonly as the foam liner in helmets, and absorbs impact energy by permanent crushing and deformation [4]. This deems it impractical for multi-impact applications. That is why a new design is proposed to accommodate the multi-shock nature of head-to-helmet-to-ground collisions. Such a design would take structural features of cattle horn and equine hoof structures to improve crashworthiness.

Understanding the amount of protection required for foam liners calls for understanding mechanisms of head injury and tolerance levels allowed in the head. Head injury is defined as "temporary or permanent damage to one or more of the head

components from a blow or accident"[5]. It includes four main components: brain Injury (BI), skull fracture, neck injury, and scalp damage. Brain injury takes precedence over other types of fractures/injuries based on statistics by Otte et al. and H H Hurt Jr et al. The first mechanism of BI is concussion due to linear acceleration, which produces pressure gradients in the brain and accounts for diffuse or focal traumatic brain injury. Angular acceleration and acute subdural hematoma (ASDH) is the second mechanism that occurs due to stretching and tearing of veins leading to a clot between the surface of the brain and Dura Matter. Thereby the liner design is to mitigate the effects of linear and rotational acceleration of the head upon impact.

The first aim is to test experimentally a commercially available motorcycle helmet used by the Lebanese civil defense to check for peak linear acceleration shown in Figure 1.

Figure 1. Motorcycle helmet sample used by civil defense individuals in Lebanon

New material and bio-inspired structural designs would then be investigated using finite element simulations to come up with a proposed liner design. This is backed with reference to the impressive behavior of horns when subject to radial loading, whose structure can be mimicked to reduce rotational effects upon impact [6]. The final product is implementing the liner design in commercially available motorcycle helmets and comparing the energy absorbing capacities.

II. Materials and Structures

A. Bio-inspired Design

Cattle horns, rhinoceros horns, and equine hooves are known to offer great shock absorbance for applications involving linear and rotational impact. The horn structure, much like that of the motorcycle helmet is composed of two layers: an outer porous "shell" of α-keratin, and inner soft tissue. Keratin is a fibrous, structural protein found in hair, nails, claws, and hooves [7]. Horns appear to be arranged in a lamellar structure (2–5 μm keratin lamellae in thickness) of keratin sheets (1-2mm in thickness). The lamellar structures are organized in the radial direction with inter-laminar tubules (40-500 μm in diameter) dispersed between the lamellae, extending along the length of the horn as hollow tubules as it grows. The horn structure is shown in Figure 2 and is reproduced from J. McKittrick et al. [7].
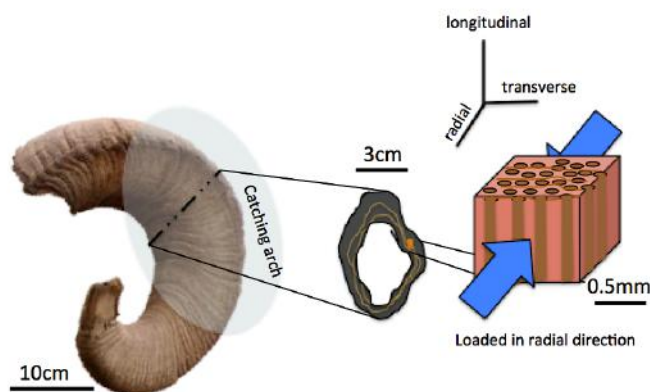
Figure 2. Oriented keratin filament structure interspersed in a protein-matrix of a sheep horn reproduced from J.McKittrick et al.

Tombolato et al. show that sheep horns exhibit gradient porosity in the horn structure, varying from 8-12% at the distal regions to about 0% at the interior surfaces, similar to what is observed in hooves. They concluded in their study that the compression in the radial direction has the lowest elastic modulus and yield strength in comparison with transverse and longitudinal directions, allowing for more compression in the tubules and more energy absorbance, owing to the highest toughness. The radial configuration that allows for most energy absorbance in compression is shown in Figure 2.

Accordingly, the proposed bio-inspired design consists of choosing solid material foam and designing it in a sheet slices of varying porosity from 0% near the head to about 10% near the outer shell oriented in the radial direction of the helmet as shown in Figure 3.
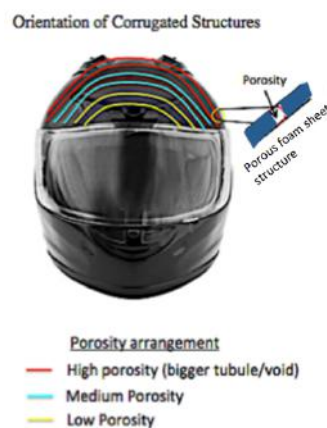
Figure 3. Proposed Design Schematic

B. Foam Liner Material

EPS is used frequently as the liner material in helmets due to its lightweight and high-energy absorbance capability as it gets permanently crushed, acting like a damper. EPS also comes in various densities. Cui et al. worked on a multi-layered foam liner of different densities for equestrian helmets [8]. Four layers were considered and different combinations of densities of EPS liner (25, 50, 64, and 80kg/m^3) were tested against the commercially available EPS of density 64kg/m^3 to compare energy absorbance. The final proposed design based on test results consisted of manufacturing the foams in decreasing density outwards. However, one disadvantage is that stresses may be localized at the layer boundaries due to the variation in material properties between the different densities, which may lead to crack propagation at these interfaces unless particular care is taken during manufacturing.

Other material have been used in foam liner as well such as polyurethane or polyvinylchloride (PVC) foam [9]. Most recently, Micro agglomerated cork (MAC) has been suggested as a substitute material for multiple impacts due to its spring-like viscoelastic nature. MAC has an advantage over natural cork as it consists granules (0.5 to 2mm) with binding agents leading to more uniform properties and more ease in machining. A composite of EPS and cork (20-40%) MAC in a linearly packed structure or parallel structure with MAC cylinders in EPS (24-29%) showed best properties in shock absorption [10]. One disadvantage of MAC is that it is denser than EPS, but adding a porous structure may help reduce the weight.

III. Methodology

A. Apparatus

Our testing apparatus is based on drop, spring-loaded mechanism. Although testing standards require a drop tower 3m in length to reach speeds of 7.5m/s upon impact of the head and helmet, the distance of the drop has been decreased for ease of manufacture and transportation, replacing it with a spring to maintain the required speed upon impact. The mechanism consists of a heavy bottom base where a flat anvil of diameter 14cm is welded to it, an upper ceiling, four guiding

columns (3cm diameter), a drop plate, spring, and pulling winch and release mechanism. To minimize the weight of the drop plate where the head form and helmet will be attached, it is convenient to design it in a shape to reduce volume shown in Figure 4. The pulling mechanism is a manual ratchet-pulley mechanism that can withstand 2 tons of pulling with release clamps.
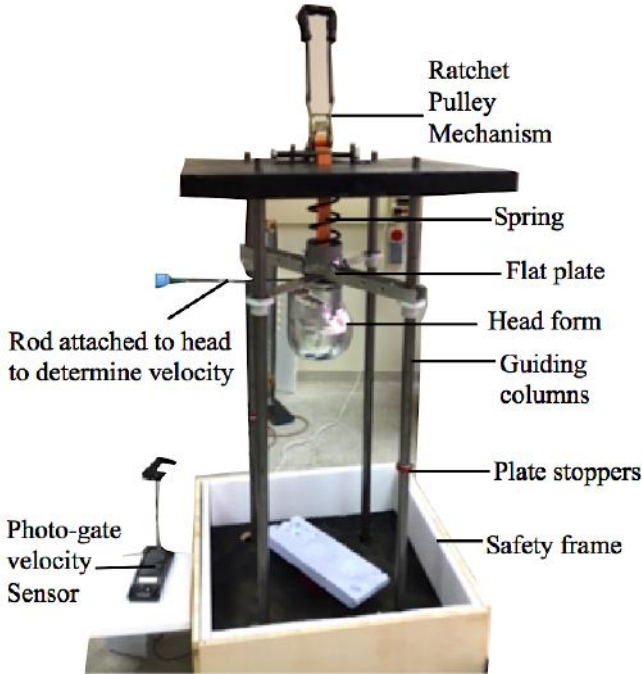
Figure 4. Drop Testing Apparatus without the head form

B. Experimental Procedure

The experimental assessment will be conducted based on the United Nations Economic Commission for Europe Standard for motorcycle helmets ECE 22.05 requiring an impact speed of 7.5m/s and peak linear acceleration less than 275g (gravitational force) and head impact criterion (HIC) below 2400. The HIC is the maximum value obtained from the linear acceleration curve upon impact taking into account duration of impact (10-15ms) as shown in equation (1) [11].

$$HIC = [(\frac{1}{t_2 - t_1}) \int_{t1}^{t2} a(t) \, dt \,]^{2.5} (t_2 - t_1) \qquad (1)$$

Where a(t) is linear acceleration (g) and t_1 and t_2 are the bounds of total impact duration.

The ECE 22.05 standard also requires a tri-axial accelerometer withstanding a maximum acceleration of 2000g and mass below 50g. For our application, Slam Stick X from Midé would be used that can measure up to 500g, weighs 40g, and has a sampling rate of 10KHz (up to 20KHz). A photo-gate velocity sensor is used near the anvil to validate that the impact velocity reaches 7.5m/s. Furthermore, a metallic "ISO-J" head-form as per the standard, and considered the near average sized

head, is manufactured at the American University of Beirut (AUB) workshops through CNC machining. For the purpose of acceleration measurement and in order to meet head injury tolerance levels, the accelerometer is placed at the center-of-gravity of the head-form.

C. Spring Requirements for impact

Upon testing, we aim to achieve an impact speed of 7-10 m/s. Our apparatus will have a stiff spring mounted on the top surface so that the desired impact speed is achieved without dropping the head-from a large height.
The dynamics for the process can be described in the following two equations (2) and (3):[12]

$$\left(\frac{1}{2}\right)(m.V^2) = m.g.(h+x) + \left(\frac{1}{2}\right)(k.x^2) \qquad (2)$$

The required stiffness can be easily calculated by rearranging the equation:

$$k = \frac{m.V^2 - 2m.g.(h+x)}{x^2} \qquad (3)$$

Where m is the mass of the helmet, head-form, and plate combined, V is the impact velocity, h is the drop height, k is the spring stiffness, and x is the compressed spring length.

Standards allow a maximum head-form weight of 4.84 kg. In our case, the combined weight of the fat plate and head form is found to be 6 kg. The added weight is to make the experiment closer to real life situations where the mass of the body is taken into account. For experimental purposes however, the head-form would separate from the flat plate upon impact thereby decreasing the mass upon impact. The drop height h is set to be 1.1m and a constraint is imposed on the compressed spring displacement to be at most 10 cm.
If the desired velocities are 7 m/s and 10 m/s, we calculate the stiffness k to be 15.2 kN/m and 45.8 kN/m.
And spring having stiffness 30 kN/m is made suitable for the apparatus by adjusting the compressed length. In this case, the compressed length is to be 7 cm and 12 cm to reach 7 m/s and 10 m/s impact speed.
Our spring specifications:
- Wire diameter (d): 1cm
- Outer diameter (D): 9.3 cm
- Number of turns (N): 4.5 turns
- Material: Iron (G=80 GPa)
- Free length : 25 cm

Equation (4) is used to calculate the stiffness of the available spring to be about 30KN/m [12].

$$k = \frac{G.d^4}{8.N.D^3} = 27.6 \, kN/m \qquad (4)$$

For stage III: Densification

$$\sigma = \frac{\sigma_y}{D}\left(\frac{\epsilon_D}{\epsilon_D - \epsilon}\right)^m + \frac{P_0 \epsilon}{1 - \epsilon - R} \quad (7)$$

where D and m are constants typically equal to 2.3 and 1 respectively and ϵ_D is densification strain expressed by Gibson and Ashby as: $\epsilon_D = 0.9(1 - 1.4R + 0.4R^3)$

Note that some models take $\frac{P_0 \epsilon}{1 - \epsilon - R}$ to be negligible

MAC can also be modeled using the "hyper foam" model in ABAQUS library as done by Fernandes et al.[13].

Preliminary testing was also conducted on the police helmet that was impacted at a speed of 7.3 m/s at the crown region. The peak acceleration was recorded at 265g <275g deeming it to be safe, however upon taking a second impact, the accelerometer saturated above 500g. The crushing of the EPS foam at the crown region can be seen in Figure 8.

Figure 8. EPS liner crushed (right) at crown region after impact

V. FUTURE DIRECTIONS

Conducting a second impact test on the police helmet in a different helmet orientation where EPS is not densified to obtain the peak acceleration curve and deduce HIC is require to deduce safety of the police helmet. Furthermore, 3D scanning of a helmet would be conducted to introduce it to ABAQUS along with the head form to impact test it against the proposed configurations in Figure 9.
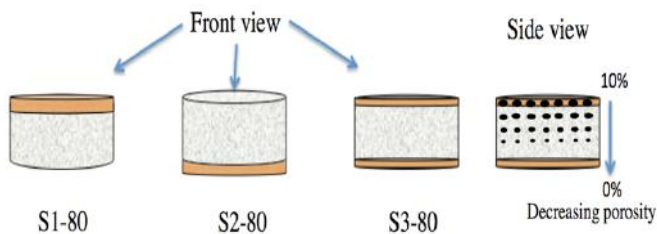
Figure 9. Proposed Series configurations for FEA (80% EPS in white and 20% MAC in orange)

REFERENCES

[1] M. C. Tsai and D. Hemenway, "Effect of the mandatory helmet law in Taiwan," Inj Prev, vol. 5, no. 4, pp. 290–291, 1999.

[2] F. M. Shuaeib, a. M. S. Hamouda, S. V. Wong, R. S. R. Umar, and M. M. H. M. Ahmed, "A new motorcycle helmet liner material: The finite element simulation and design of experiment optimization," Mater. Des., vol. 28, no. 1, pp. 182–195, 2007.

[3] P. K. Pinnoji, P. Mahajan, N. Bourdet, C. Deck, and R. Willinger, "Impact dynamics of metal foam shells for motorcycle helmets: Experiments & numerical modeling," Int. J. Impact Eng., vol. 37, no. 3, pp. 274–284, 2010.

[4] V. Kostopoulos, Y. . Markopoulos, G. Giannopoulos, and D. . Vlachos, "Finite element analysis of impact damage response of composite motorcycle safety helmets," Compos. Part B Eng., vol. 33, no. 2, pp. 99–107, 2002.

[5] F. M. Shuaeib, A. M. S. Hamouda, R. S. Radin Umar, M. M. Hamdan, and M. S. J. Hashmi, "Motorcycle helmet - Part I. Biomechanics and computational issues," J. Mater. Process. Technol., vol. 123, no. 3, pp. 406–421, 2002.

[6] L. Tombolato, E. E. Novitskaya, P. Y. Chen, F. A. Sheppard, and J. McKittrick, "Microstructure, elastic properties and deformation mechanisms of horn keratin," Acta Biomater., vol. 6, no. 2, pp. 319–330, 2010.

[7] J. McKittrick, P. Y. Chen, S. G. Bodde, W. Yang, E. E. Novitskaya, and M. A. Meyers, "The structure, functions, and mechanical properties of keratin," JOM, vol. 64, no. 4, pp. 449–468, 2012.

[8] L. Cui, M. A. Forero Rueda, and M. D. Gilchrist, "Optimisation of energy absorbing liner for equestrian helmets. Part II: Functionally graded foam liner," Mater. Des., vol. 30, no. 9, pp. 3414–3419, 2009.

[9] L. Chang, G. Chang, J. Huang, S. Huang, D. Liu, and C. Chang, "Finite element analysis of the effect of motorcycle helmet materials against impact velocity," J. Chinese Inst. Eng., vol. 26, no. 6, pp. 835–843, 2003.

[10] R. M. Coelho, R. J. Alves de Sousa, F. A. O. Fernandes, and F. Teixeira-Dias, "New composite liners for energy absorption purposes," Mater. Des., vol. 43, pp. 384–392, 2013.

[11] L.-T. Chang, C.-H. Chang, and G.-L. Chang, "Fit Effect of Motorcycle Helmet. A Finite Element Modeling.," JSME International Journal Series A, vol. 44, no. 1. pp. 185–192, 2001.

[12] R. Budynas and K. Nisbett, Shigley's Mechanical Engineering Design, SI Version. 2009.

[13] F. A. O. Fernandes, R. J. S. Pascoal, and R. J. Alves de Sousa, "Modelling impact response of agglomerated cork," Mater. Des., vol. 58, pp. 499–507, 2014.

Appendix A: Reference ISO J Head form
Extracted from ECE 22.05 Manual

REFERENCE HEADFORMS
(shape, dimensions above reference plane)
Dimensions in millimetres
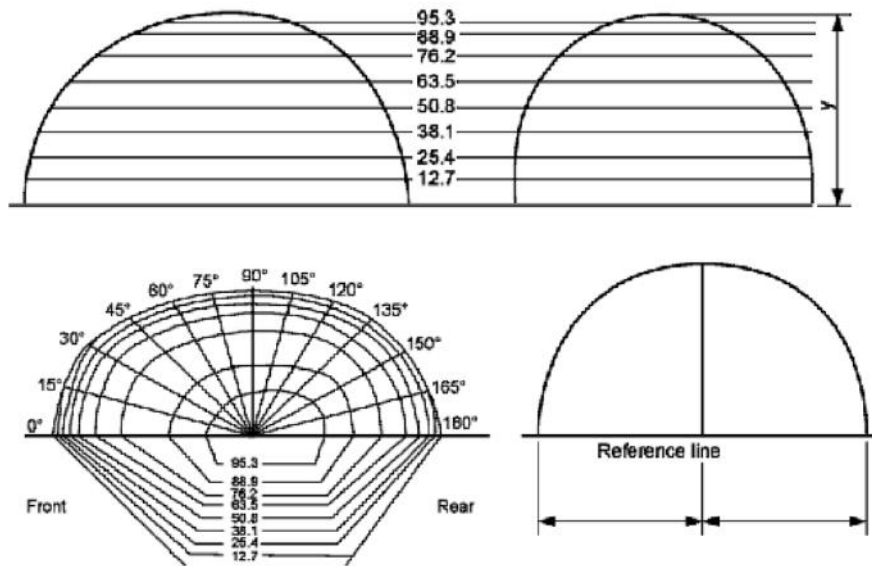
J (Dimensions in millimetres)													
Height above reference line	0° Front	15°	30°	45°	60°	75°	90°	105°	120°	135°	150°	165°	180° Rear
0	100.8	98.8	96.3	88.1	82.0	79.5	79.2	82.0	85.9	91.7	96.8	100.1	100.8
12.7	99.6	98.0	95.8	88.1	82.0	79.5	79.2	82.0	85.9	91.7	96.8	100.1	100.8
25.4	96.8	95.8	94.5	88.1	82.0	79.5	79.2	82.0	85.9	91.7	96.5	98.3	98.8
38.1	93.7	92.7	91.9	86.1	80.0	77.2	77.7	80.0	83.8	89.4	94.5	95.8	96.0
50.8	89.2	88.6	87.9	82.0	76.2	73.9	74.4	77.0	80.5	85.9	90.4	90.9	90.9
63.5	81.5	80.8	81.0	75.9	70.6	68.1	68.3	71.1	71.4	79.5	83.8	84.1	84.1
76.2	69.3	69.1	69.3	65.3	61.2	58.9	59.2	61.7	65.0	69.3	73.2	73.4	73.4
88.9	47.2	47.5	48.0	46.2	44.4	43.7	44.2	46.2	50.0	54.1	58.2	58.4	58.4
95.3	32.8	32.8	33.3	32.5	32.0	32.3	33.0	35.1	38.1	42.2	46.5	47.2	47.2
Dimension: Y : 102.4 mm - Head circumference : 570 mm													

Dimension Z'
Reference plane

Reference plane

G

Dimension Z

Side view
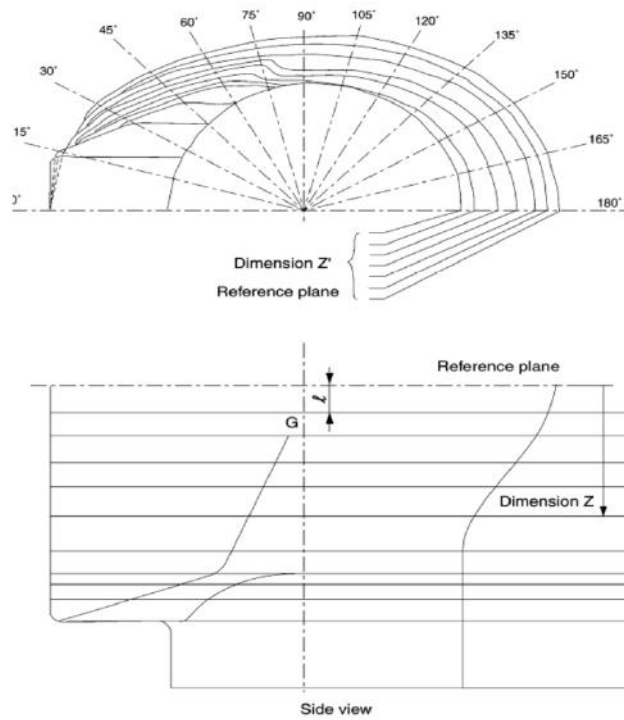
J													
Dimension Z'	0°	15°	30°	45°	60°	75°	90°	105°	120°	135°	150°	165°	180° Rear
0	101.0	99.5	95.5	88.5	82.5	79.5	79.5	82.0	86.0	92.0	97.0	100.5	101.0
-12.7	101.0	99.5	94.5	85.0	78.0	75.5	75.5	78.0	82.0	88.0	93.0	96.5	97.0
-22.7	101.0	100.5	94.5	84.5	76.0	72.0	70.5	73.5	77.0	83.0	88.0	91.5	92.0
-35.0	101.0	102.0	92.5	81.5	74.0	71.0	64.0	66.5	70.5	76.0	81.0	84.0	84.5
-45.0	101.0	102.0	90.0	79.0	72.0	68.5	61.5	63.0	66.0	70.5	74.5	77.0	76.5
-60.0	101.0	102.0	88.0	76.5	69.0	61.5	59.0	59.5	61.0	64.5	67.5	68.5	67.0
-75.0	101.0	102.0	86.0	74.0	67.0	60.0	57.5	58.5	59.0	60.0	60.5	61.5	62.0
-85.0	101.0	102.0	84.0	71.5	66.0	58.0	57.5	58.5	59.0	60.0	60.5	61.5	62.0
-90.0	101.0	102.0	81.5	69.0	56.5	57.0	57.5	58.5	59.0	60.0	60.5	61.5	62.0
-96.5	101.0	102.0	79.5	54.5	56.5	57.0	57.5	58.5	59.0	60.0	60.5	61.5	62.0
-106.0	101.0	105.0	54.0	54.5	56.5	57.0	57.5	58.5	59.0	60.0	60.5	61.5	62.0
-136.0	53.5	53.5	54.0	54.5	56.5	57.0	57.5	58.5	59.0	60.0	60.5	61.5	62.0
Dimension 1 : 12.7 mm - Head circumference : 570 mm													

INDUSTRIAL ENGINEERING AND MANAGEMENT

CONTENT

Classification and Analysis of Notices under the 1999 FIDIC's General Conditions of the Construction Contract

Salam G. Khalife
Industrial Engineering and Management
American University of Beirut
Beirut- Lebanon
sgk19@mail.aub.edu

Abstract-**The rights, obligations, and responsibilities of the parties to a construction contract are described in the conditions of contract through both the general and particular conditions. These provisions need to be well comprehended and administered by contract administrators, working with each of the parties. Notice provisions are viewed as main contractual provisions, as these, in general, preserve the right of knowledge by either party about events arising during the construction phase and serve to highlight the following or consequent actions in response to such events. Under certain clauses, notices are sometimes regulated by time bars for their timely issuance; one well-known example is the notice of intention to submit a claim for an additional compensation or an extension of time upon encountering certain events during the course of construction. Yet, if the notice is not served within the regulated period set by the contract conditions, the contractor, under many jurisdictions, loses his right for such a claim.**

Contract administrators encounter various events that necessitate the issuance of notices, all in accordance with contract stipulations. The problem is that contract administrators often deal with notices in a negligent way, thereby exposing the contract to major risks. This is because of the lack of material that explains the notice provisions and their importance. The objective of this paper is to look into the wide range of notices and study their relevant provisions. For that purpose, the globally used 1999 standard contract conditions, issued by the International Federation of Consulting Engineers (FIDIC), are examined, filtered and analyzed for the purpose of giving full information and guidelines about notices.

I. INTRODUCTION

The increase in size and complexity of construction projects and the fact that, the budget agreed on, the completion time and the scope of work tends to change, are main reasons that make the construction contract one of the complex contracts to be administered [1]. In recent years, employers are increasingly forcing the improvement of management procedures on construction sites [2] to maintain successful completion of projects and to avoid or resolve claims and other events that affect the work progress. Consequently, the contract document has been the initiative step towards constituting efficient project management procedures and fair distribution of project

risks for a better relation and less adversarial approach between parties to construction project [2].

Construction contracts include general and specific conditions in which contractual and legal obligations are imposed on parties to agreement. These conditions describe the rights, duties and responsibilities of the parties throughout the construction phase. The standard forms of contracts are being highly adhered recently, seeking the balanced conditions that they are meant to establish. Yet the crucial duty is held upon managing these contractual conditions and provisions included. Administering the contract and its terms is equally important for the success of the project. This is where the role of contract administrator appeals.

Today, multiple parties constitute the team required to follow up work on site and its relevant requirements under the contract. The contract administration team is required to have sufficient knowledge of the contract document. They evaluate any event on site that might require a certain act to be established under the contract conditions. The provisions included in the contract conditions describe how to address these events. Therefore, full compliance with these provisions is important to maintain and ensure project's execution and completion is in conformance with quality, time and budget requirements.

Notice provisions form an integral part of the contractual requirements under the contract. They describe the form of communication between the owner, engineer and contractor, and how and when knowledge shall be communicated about different situations [3]. Notices are directly related to the procedural measures of the construction project, in which they cover the performance, methods and practices required when there are submittals, payments, testing, variations, defects, and unforeseen events or actions that may lead to claims for extension of time and/or money.

Therefore, ultimately notice provisions, along with other provisions described in the construction contract, shall be referred to almost on daily basis in the construction phase; Contract administrators along with the efficient cooperation with the staff on site shall be well comprehended with the

circumstances that warrants the need to issue notices to preserve the contractual obligations and protect all parties' rights for knowledge, judgement, and action.

It is of great importance that parties know about the highly beneficial aspects of issuing notices and especially notices that are condition precedents under certain clauses. These notices serve all parties interests and protect them as long as they are properly complied with [4].

II. PROBLEM STATEMENT

Going through previous researches, notice provisions have been tackled mainly from a limited perspective as in connection to contractor's claim notice. Notice provisions related to claims are discussed within published literature since early 80s [3]. This is because failure to issue the notice of a claim within the time stipulated in the contract is being vulnerable to dismissal of the claim [5].

Yet still, the material covering these requirements is vague or incomplete, and notice provisions are simply presented superficially [3]. Therefore, the question is in what other situations during the construction phase are notices required? What are the different provisions and time periods that governs the issuance of these notices? And what can be said about the notices and their mechanism?

III. OBJECTIVE AND METHODOLOGY

The objective of this paper is to (1) create a complete image of notices found under the 1999 FIDIC general conditions, (2) analyze the different types of notices and what is common between them, (3) help unveil the implied thinking or rationale governing how time barring has been decided on, and (4) present guidelines about notices that help contract administrators fulfill their duties under the contracts. This is done according to a five-step methodology:

1. Reviewing the FIDIC clauses to filter out the clauses and provisions that prescribe the need for issuing notices;

2. Deducing descriptors to help classify the notices called for under the filtered clauses;

3. Documenting the filtered notice provisions in a classification matrix using the deduced descriptors;

4. Analyzing the classification matrix in depth in order to reveal the underlying needs for such notice provisions; and

5. Drawing conclusions as to the importance of satisfying notices requirements and providing guidelines to serve in betterment of contract administrators' work.

IV. NOTICES UNDER THE FIDIC 1999 CLAUSES

A. FIDIC clauses and filtered notice

The twenty clauses, in the general conditions of the 1999 FIDIC construction contract, are examined for the purpose of relating which of these clauses call out for notices within their sub-clauses. Table I shows the results of filtering out the sub-clauses.

Table 1 Number of notices filtered out from the 20 clauses of the 1999 Clauses

Clause 1	General provisions	5
Clause 2	The Employer	4
Clause 3	The Engineer	3
Clause 4	The Contractor	14
Clause 5	Nominated Subcontractors	2
Clause 6	Staff and Labour	0
Clause 7	Plant, Material and Workmanship	7
Clause 8	Commencement, Delays and Suspension	12
Clause 9	Tests on Completion	4
Clause 10	Employer's Taking Over	5
Clause 11	Defects Liability	6
Clause 12	Measurement and Evaluation	3
Clause 13	Variations and Adjustments	2
Clause 14	Contract Price and Payment	1
Clause 15	Termination by Employer	5
Clause 16	Suspension and termination by Contractor	3
Clause 17	Risk and Responsibility	3
Clause 18	Insurance	5
Clause 19	Force Majeure	5
Clause 20	Claims, Dispute and Arbitration	2
TOTAL Number of notices		91

As shown in the above table all clauses call out for issuing a number of notices except for clause 6 [Staff and Labour]. This reveals that notices are connected to all related aspects of the construction phase and process. The highest number of notices called for are in reference to Clause 4 [The Contractor] where 14 notices are requested under this clause. It is followed by Clause 8 for the commencement, delays ad suspension of works.

The 91 notices found need to be classified and analyzed for the purpose of understanding all their aspects.

B. Descriptors used to help with the analysis

Two types of descriptors have been used for the purpose of understanding related information about notices and analyzing them afterwards: (1) The expressed descriptors; which are extracted and explicitly stated within the sub-clauses of the FIDIC general conditions. These include:

1. Context
2. Party issuing and party receiving notice,
3. Notice prerequisite(s),
4. Time bar,
5. Subsequent action, party responsible for the subsequent action and its time bar; and
6. Notice description.

The second type (2) is the deduced descriptors and they include:

1. Type of notice,
2. Works' relevant aspects,
3. Notice/Event sequence, and
4. Structured or non-structured notices.

Type of notice describes the nature of the notice and context related to it, whether it is only an informing notice, attention notice or even a causation/alarm notice stating that there is a cause that might lead to a claim or a progress issue and so on. 14 different types of notices were found each expressing the language included in the notice.

The second deduced descriptor is the work relevant aspect. It is used to classify notices in relation to which type of work it is correlated to: is it design related, site work related, administrative, financial, etc. Also 14 work sections were found connected to notices.

The third descriptor is the sequence between the notice and the event related to it. The notice could either precede the event, follows it, or be related to two events which is a dual-event notice. Dual-event notice has to be either:

- Preceded by an event which triggers issuing this notice and followed by a time period for the notice event to be effective, e.g.: Termination Notice (Sub-Clause 16.2); or
- Preceded by an event which triggers issuing this notice and followed by a time barred event described in the notice itself.
  e.g.: Failure to Remedy Defects (Sub-Clause 11.4): Contractor fails to remedy defects→ Employer fixes a date for the defect to be remedied by giving notice → Contractor is expected to remedy defect within the time bar stipulated in the notice.

The forth descriptor is whether the notice is structured or non-structured. This helps in identifying the notices that are related to a chain of events, notices, or time bars and therefore are called structured. The situations in which the notice is specified as structured are as follows:

- It is preceded by another notice and/or a regulated period,
  e.g.: Termination Notice (Sub-Clause 16.2);
- It sets itself a regulated period for the required subsequent action,
  e.g.: Delayed Drawings or Instructions (Sub-Clause 1.9); or
- It alters itself the regulated period already set for the subsequent action,
  e.g.: Contractor's Claim (Sub-Clause 20.1).

C. Classification Matrix

The classification matrix was constructed and all relevant information were plugged in. Each row in the matrix is a separate notice described under its corresponding sub-clause. The columns are the descriptors selected to help reveal common information about notices; and using the filter mode in the excel spreadsheet where the matrix is formulated,

statistical results are produced in which the analysis is based upon.

The following figures Fig.1 to Fig.4 show the four deduced descriptors and their count according to the frequency of repetition within the sub-clauses.
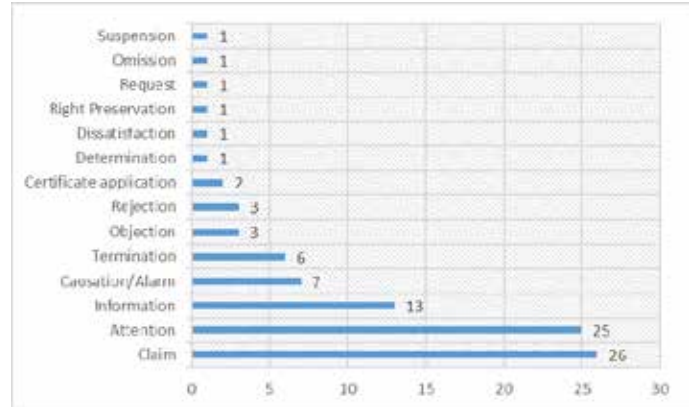
Figure 1 Notices' types and their count
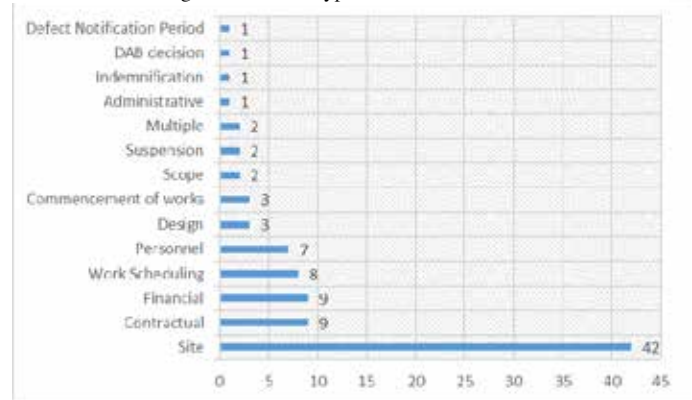
Figure 2 Notices' work relevant aspect and their count
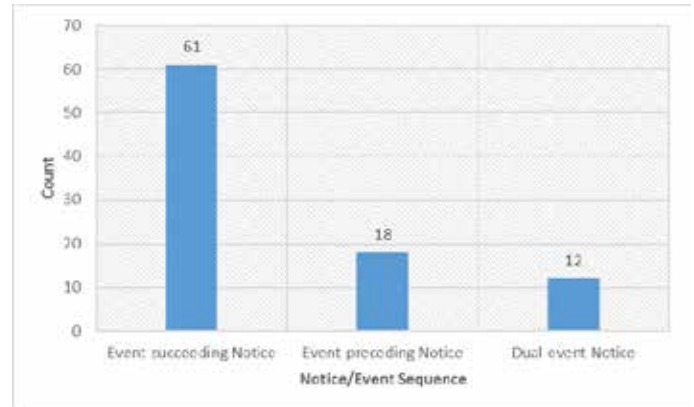
Figure 3 Notice/Event sequence and their count
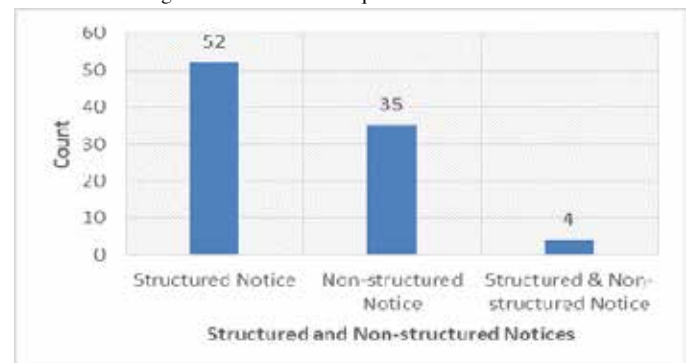
Figure 4 Number of structured and non-structured notices

V. ANALYSIS OF THE NOTICES MATRIX

In reference to the descriptors and the classification matrix, a detailed analysis is carried out in order to reveal the importance of notices and what are some rationale behind them.

Party issuing Notice

Out of the ninety-one notices extracted from the FIDIC sub-clauses, the results showed that:
- 29 are related to the employer
- 56 are related to the contractor
- 28 are related to the engineer

This proves that notices are not only a burden on the contractor as often deduced; instead, this shows that under balanced conditions of the contract, the responsibilities and duties are distributed among parties. The high number of notices issued by the contractor though, shall be explained in respect that his team is the party executing the work and therefore most events are to be encountered under his knowledge.

Time Periods for Notices

As for the time periods stipulated under the sub-clauses, we analyzed two types of descriptions to these periods, one qualitative numerical and the second is qualitative non-numerical. Sixty-one notices have a time period stipulated as follows:
- 22 notices' time bars are non-numerical: those include 9 stating that the notice shall be given promptly and the other 13 are to be issued as soon as practicable. From the analysis of the FIDIC contractual language, 'promptly' is concluded to be the shortest time duration possible; and 'as soon as practicable' is less than 28 days.
- 39 notices are regulated with time periods which range from 24hrs-period to 42-days period. Whenever notices are time regulated with numerical values, these should be well comprehended and 'noticed' by contract administrators.

Time periods set for notices under the contract help organize the construction works and parties' actions. The majority of time regulated notices are found to be set for profound situations, for instance: program rejection notice, notice of force majeure, notice of disagreeing the records for works measured, right to access, taking-over the works, etc. these represent important milestones in the project timeline and therefore regulating their issuance represents a vital norm in project management.

Type of Notices

Classification of notices according to their type helped in giving a clear idea in what sense each party shall deal with the notice. So when notice is referred to as an attention notice, which is the case of the majority of notices other than claim notices, then it is related to a certain event that need to be carefully dealt with as it would affect the progress of work and/or has a negative impact on the project as a whole. Examples on these:

- Under sub-clause 1.8 [*Care and Supply of Documents*]: "*If a Party becomes aware of an error or defect of a technical nature in a document which was prepared for use in executing the Works, the Party shall promptly give notice to the other Party of such error or defect.*" Defect in such documents has a major effect on the project, and therefore failure by any of the parties whom became aware of this error or defect will have an undesirable effect on the execution of the works.
- Under sub-clause 3.4 [*Replacement of the Engineer*]: "*If the Employer intends to replace the Engineer, the Employer shall, not less than 42 days before the intended date of replacement, give notice to the Contractor of the name, address and relevant experience of the intended replacement Engineer.*" This is another attention notice with a regulated period of 42 days, where the employer gives attention to the contractor through a notice, that he intends to replace the engineer.

In addition to attention notices, claim notices, which are very popular in the construction industry, need to have more focus on. The spread sheet matrix (a sample of the matrix is followed in the appendix) explains more about these notices which are a major issue as they often lead to disputes.

Works relevant aspect

The aspect of works described in fig. 2 reveal that notices cover all aspects related to the project, from relation to financial issues, site works, contractual issues and so on. Parties shall know about the aspects related to notices in order to specify on what level these notices shall be dealt with and who is concerned to know about them, for instance: design team, staff on site, accountants, etc. The highest value encountered for works aspect is site related notices. This implies that contract administrators have an important role in connecting with parties found on site daily to follow up the works and evaluate the situations with respect to issuing notices.

Notice/event sequence

This descriptor helped in identifying the relation between the notice and the event triggering its issuance. Therefore, the contractor, engineer, and employer may want to know in which circumstances the notice shall precede the event or corresponding action; in which events notice follows a certain action; and in what sense sometimes the notice is connected to dual events.

Structured and Non-Structured notices

Another main specification for the notices is whether they are structured or not. Some notices have a sequence of events relating to them and therefore are at a level of sophistication and need to be understood well. One important example is the

notice of termination which in case it is not well met, there is a breach of contract.

VI. GENERAL GUIDELINES

According to the above classification matrix and analysis, we can deduce the following guidelines and recommendations:

1. All parties included within the construction project, weather the owner, engineer, or contractor, shall be aware of the events that necessitates issuing notices according to the contract document. Therefore, before going through with the works, the contract document and conditions shall be analyzed as per the suggested form of matrix so that each party will know its obligations, rights, and duties.

2. For the contractor, the causation/alarm notices that lead to claims under sub-clause 20.1 should be of great importance, and their corresponding provisions shall be well comprehended. These include the following sub-clauses:
   - "*1.9 [Delayed Drawings or Instructions]*"
   - "*2.1[Right of Access to The Site]*"
   - "*4.7 [Setting Out]*"
   - "*4.12 [Unforeseeable Physical Conditions]*"
   - "*4.24 [Fossils]*"
   - "*7.4 [Testing]*"
   - "*8.4 [Extension of Time for Completion]*"
   - "*8.9 [Consequences of Suspension]*"
   - "*10.2 [Taking Over of Parts of the Work]*"
   - "*10.3 [Interference with Tests On Completion]*"
   - "*13.7 [Adjustments for Changes in Legislation]*"
   - "*16.1[Contractor's Entitlement to Suspend Work]*"
   - "*17.4 [Consequences of Employer's Risks]*"
   - "*18.1 [General Requirements for Insurances]*"
   - "*19.4 [Consequences of Force Majeure]*"

3. For the Engineer, other than his main role as to determinations under sub-clause 3.5, he shall be responsible for notices of commencing, inspecting, accepting, and measuring of works in addition to following up with works' execution and scheduling. Interim payment certificate and its corresponding notice of not issuing it is also a duty for the Engineer.

4. As for the Employer, from the analysis of the matrix, his role in issuing notices himself is related to the following: financial arrangements, replacement of the engineer, and notice of termination. Notices related to remedying defects may be issued by the employer or on his behalf. Finally, the notices that could be issued by the employer or engineer are claim notices related to sub-clause 2.5 [Employer's Claims] and one notice for the authorized personnel notice.

VII. CONCLUSION

The information revealed from this study are helpful for all parties included in the projects from the contract management level to the daily staff on site where events are likely to occur and someone has to be informed about it. Notices as discussed above, cover the project's works throughout its different stages, from commencement date and until the contract close-out. Notices are an important contractual and procedural tool for a successful project completion where parties communicate together and share their intensions, do their obligations, and preserve their rights under the contract. In conclusion, we hope this research provide guidance to practitioners and improve the way contracts are being dealt with towards a profound understanding of their requirements.

VIII. FURTHER STUDIES

From the importance of the information revealed in the structured notices, ongoing efforts are being done to produce figures for these notices. This will help enrich this study further and help contract administrators visualize the series of events and time periods related to notices.

REFERENCES

[1] J. Adriaanse, Construction Contract Law, New York: Palgrave Macmillan, 2007.

[2] B. W. Totterdill, FIDIC Users' Guide: A Practical Guide to the 1999 Red and Yellow Books, London: Thomas Telford Publishing, 2006.

[3] H. R. Thomas and R. D. Ellis, Interpreting Construction Contracts; Fundamental Principles for Contractos, Project Managers, and Contract Administrators, United States of America: American Society of Civil Engineers, 2007.

[4] M. Harris, "The Importance and Value of Notice Provisions in Construction Contract," Long International, Inc., 2015.

[5] R. J. Miletsky, "Failure to Comply with Contract Notice Provisions Places your Claims at Risk," *Contractor's Business Management Report,* pp. Issue 01-8, 2001.

Sub-clause	Clause Title	Context/Reason	Notice Action	Notice Count	Party Issuing Notice	Party Receiving Notice	Prerequisite(s)	Event(s)	Notice Time Bar
1						General Provisions			
1.3	Communications	Whenever these Conditions provide for the giving or issuing of approvals, certificates, consents, determinations, notices and requests, these communications shall be: (a) in writing and delivered by hand (against receipt), sent by mail or courier, or transmitted using any of the agreed systems of electronic transmission as stated in the Appendix to Tender; and (b) delivered, sent or transmitted to the address for the recipient's communications as stated in the Appendix to Tender. However,…	…, if the recipient gives **notice** of another address, communications shall thereafter be delivered accordingly.	1	Party	Other Party	–	Change in communication address becomes effective.	Not specified
1.8	Care and Supply of Documents	If a Party becomes aware of an error or defect of a technical nature in a document which was prepared for use in executing the Works, …	…, the Party shall promptly give **notice** to the other Party of such error or defect.	2	Party	other Party	–	Party becoming aware of the error or defect.	Promptly
1.9	Delayed Drawings or Instructions	Whenever the Works are likely to be delayed or disrupted if any necessary drawing or instruction is not issued to the Contractor **within a particular time, which shall be reasonable**, …	…, the Contractor shall give a **notice** to the Engineer. The **notice** shall include details of the necessary drawing or instruction, details of why & **by when** it should be issued, and details of the nature and amount of the delay or disruption likely to be suffered if it is late.	3	Contractor	Engineer	–	Contractor becoming aware of the disruption or expected delay.	Not specified
		If the Contractor suffers delay and/or incurs Cost as a result of failure of the Engineer to issue the **notified** drawing or instruction within a time which is reasonable and is specified in the **notice** with supporting details, …	…, the Contractor shall give a **further notice** to the Engineer and shall be entitled subject to Sub-Clause 20.1 [*Contractor's Claims*] to: (a) an extension of time for any such delay, if completion is or will be delayed, under Sub-Clause 8.4 [Extension of Time for Completion], and (b) payment of any such Cost plus reasonable profit, which shall be included in the Contract Price.	4	Contractor	Engineer	Notice of 'Needed Additional Drawings or Instruction	Failure by the Engineer to issue the notified drawings or instruction within the reasonable time specified in the notice.	As soon as practicable and ≤ 28 days
1.14	Joint and Several Liability	If the Contractor constitutes (under applicable Laws) a joint venture, consortium or other unincorporated grouping of two or more persons, …	…, these persons shall **notify** the Employer of their leader who shall have authority to bind the Contractor and each of these persons.	5	Contractor	Employer	–	Constituting the joint venture.	Not specified
2						The Employer			
2.1	Right of Access to the Site	The Employer shall give the Contractor right of access to, and possession of, all parts of the Site **within the time** (or times) stated in the Appendix to Tender. If no such time is stated in the Appendix to Tender, the Employer shall give the Contractor right of access to, and possession of, the Site **within such times as may be required** to enable the Contractor to proceed in accordance with the programme submitted under Sub-Clause 8.3 [*Programme*]. If the Contractor suffers delay and/or incurs Cost as a result of this failure to give any such right or possession within such time, …	…, the Contractor shall give **notice** to the Engineer and shall be entitled subject to Sub-Clause 20.1 [*Contractor's Claims*] to: (a) an extension of time for any such delay, if completion is or will be delayed, under Sub-Clause 8.4 [Extension of Time for Completion], and (b) payment of any such Cost plus reasonable profit, which shall be included in the Contract Price.	6	Contractor	Si	–	Failure by the Employer to give the Contractor right of access to all parts of the site.	As soon as practicable and ≤ 28 days
2.4	Employer's Financial Arrangements	[The Contractor requires to know about the Employer's financial arrangements possibly in connection with delayed payments by Employer].	[The Contractor /notifies the Employer of a/ request to provide] reasonable evidence that financial arrangements have been made and are being maintained.	7	Contractor	Employer	–	Delayed payment or other financially related event has taken place.	Not specified
		If the Employer intends to make any material change to his financial arrangements, …	…, the Employer shall give **notice** to the Contractor with detailed particulars.	8	Employer	Contractor	–	The Employer makes changes in financial arrangements.	Not specified

Subsequent Action	Party Responsible for Subsequent Action	Time Bar for the Subsequent Action	Notice Description	Type of Notice	Works' Relevant Aspects	Notice/Event Sequence	Structured or Non-structured Notice	Notes on Notice-Event Sequence and Structured Notices
Communications shall thereafter be delivered accordingly.	All Parties	Not specified	Notice of Change of Communication Address	Information Notice	Administrative	Event-preceding Notice	Non-structured Notice	Notice is issued after which communications shall be delivered to the new informed address.
Concerned Party is expected to react to the notice (implied).	Party	–	Notice of a Technical Error or Defect	Attention Notice	Design	Event-succeeding Notice	Non-structured Notice	Party becomes aware of the defect then gives notice.
The Engineer is expected to react within the time bar set in the Contractor's Notice.	Engineer	Notice sets Time Bar (Nested)	Notice of Needed Additional Drawings or Instruction	Causation/Alarm Notice	Design	Event-succeeding Notice	Structured Notice	Notice is issued after the awareness of the Contractor of a delayed drawing or instruction which might affect the works. Notice is structured relative to the time stipulated in the notice itself for the subsequent action.
Within 42 days after the Contractor became aware (or should have become aware) of the event or circumstance giving rise to the claim, or within such other period as may be proposed by the Contractor and approved by the Engineer, the Contractor shall send to the Engineer a fully detailed claim which includes full supporting particulars of the basis of the claim and of the extension of time and/or additional payment claimed.	Contractor	Period remaining out of the 42 days	Incurred Delay/Cost Notice	Claim Notice	Design	Event-succeeding Notice	Structured Notice	Notice is issued after the failure of the Engineer to act within the time stipulated in the first-tier notice and therefore delay has been reached. Notice is structured since it is preceded by a notice and time period and its issuance alters the regulated period for the subsequent action; i.e. the remaining duration for the particulars to be submitted.
The Leader shall have authority to bind the Contractor and each of these persons.	Leader	Not specified	Notice of a Joint Venture Personnel	Information Notice	Personnel	Event-succeeding Notice	Non-structured Notice	Notice is issued following the event of constituting a joint venture or consortium by the Contractor.
The Contractor shall send to the Engineer a fully detailed claim which includes full supporting particulars of the basis of the claim and of the extension of time and/or additional payment claimed. However, if and to the extent that the Employer's failure was caused by any error or delay by the Contractor, including an error in, or delay in the submission of, any of the Contractor's Documents, the Contractor shall not be entitled to such extension of time, Cost or profit.	Contractor	Period remaining out of the 42 days	Incurred Delay/Cost Notice	Claim Notice	Site Accessibility	Event-succeeding Notice	Structured Notice	Notice is issued after the failure of the Employer to give any such right or possession within such time as stated in Appendix or related to proceed in accordance with the programme. Notice is structured since it alters the regulated period set for the subsequent action (particulars submission).
The Employer shall submit, within 28 days after receiving any request from the Contractor, reasonable evidence that financial arrangements have been made and are being maintained which will enable the Employer to pay the Contract Price (as estimated at that time)in accordance with Clause 14 [*Contract Price and Payment*].	Employer	≤ 28 days	Request for Evidence on Employer's Financial Arrangements	Request Notice	Financial	Event-succeeding Notice	Structured Notice	Notice is issued as a result of delayed payment or at any time the Contractor needs to check on the Employer's financial arrangements. Notice is structured as it sets a time bar for the subsequent action. This request/notice may be issued after the elapsing of the 56 days duration required under Sub-Clause 14.7 where the Employer shall pay to the Contractor the amount certified in the Interim Payment Certificate.
The Contractor is updated on the Employer's financial arrangements.	Contractor	–	Notice of Material Change in Financial Arrangements	Attention Notice	Financial	Event-preceding Notice	Non-structured Notice	The notice shall precede the event of making changes to financial arrangements.

Sub-clause	Clause Title	Context/Reason	Notice Action	Notice Count	Party Issuing Notice	Party Receiving Notice	Prerequisite(s)	Event(s)	Notice Time Bar
2.5	Employer's Claims	If the Employer considers himself entitled to any payment under any Clause of these Conditions or **otherwise in connection with the Contract**, and/or to any extension of the Defects Notification Period, …	…, the Employer or the Engineer shall give **notice** and particulars to the Contractor. However, notice is <u>not required</u> for payments due under Sub-Clause 4.19 [*Electricity, Water and Gas*], under Sub-Clause 4.20 [*Employer's Equipment and Free-Issue Material*], or for other services requested by the Contractor. The notice shall be given as soon as practicable after the Employer became aware of the event or circumstances given rise to the claim. A notice relating to any extension of the Defects Notification Period shall be given before the expiry of such period. The particulars shall **specify the Clause** or other basis of the claim, and shall include substantiation of the amount and/or extension to which the Employer considers himself to be entitled in connection with the Contract.	9	Employer or Engineer	Contractor	–	Events described in the Clauses relating to this Sub-Clause has taken place or otherwise in connection with the Contract.	As soon as practicable
3					The Engineer				
3.4	Replacement of the Engineer	If the Employer intends to replace the Engineer, …	…, the Employer shall, not less than 42 days before the intended date of replacement, give **notice** to the Contractor of the name, address and relevant experience of the intended replacement Engineer.	10	Employer	Contractor	–	Replacement of the Engineer being effective.	≥ 42 days
		The employer shall not replace the Engineer with a person, …	…, against whom the Contractor raises reasonable objection by **notice** to the Employer, with supporting particulars.	11	Contractor	Employer	Engineer's Replacement notice	Event preceding: Engineer's Replacement notice. Event succeeding: reaction to the Employer's notice by the Contractor and replacement being either hindered or effective.	Not specified but implied to be issued prior to the intended replacement becoming effective
3.5	Determinations	Whenever these conditions provide that the Engineer shall proceed in accordance with this Sub-Clause 3.5 to agree or determine any matter, the Engineer shall consult with each Party in an endeavour to reach agreement. If agreement is not achieved, the Engineer shall make a fair determination in accordance with the contract, taking due regard of all relevant circumstances.	The Engineer shall give **notice** to both Parties of each agreement or determination, with supporting particulars.	12	Engineer	Employer & Contractor	Notices in reference to Sub-Clauses 20.1, 2.5 and termination notices under 15.2,16.2,and 19.6.	The Conditions provide that the Engineer shall proceed in accordance with this Sub-Clause 3.5.	Hybrid: _Nested stipulation of 42 days for determination in response to Sub-Clause 20.1. _ No stipulation in response to all other concerned Sub-Clauses.
4					The Contractor				
4.1	Contractor's General Obligations	The Contractor shall, whenever required by the Engineer, submit details of the arrangements and methods which the Contractor proposes to adopt for the execution of the Works.	No significant alteration to these arrangements and methods shall be made without this having previously been **notified** to the Engineer.	13	Contractor	Engineer	–	Alteration to any arrangements and methods for the execution of works is to take place.	Not specified
4.3	Contractor's Representative	If the Contractor's Representative is to be temporarily absent from the Site during the execution of the Works, …	…, a suitable replacement person shall be appointed, subject to the Engineer's prior consent, and the Engineer shall be **notified** accordingly.	14	Contractor	Engineer	–	Absence of the Contractor's representative from Site.	Not specified
		The Contractor's Representative may delegate any powers, functions and authority to any competent person, and may at any time revoke the delegation.	Any delegation or revocation shall not take effect until the Engineer has received prior **notice** signed by the Contractor's representative, naming the person and specifying the powers, functions and authority being delegated or revoked.	15	Contractor	Engineer	–	Delegation or revocation of any powers, functions and authority to any competent person.	Not specified
4.4	Subcontractors	The Contractor shall be responsible for the acts or defaults of any Subcontractor, his agents or employees, as if they were the acts or defaults of the Contractor. Unless otherwise stated in the Particular conditions: …	… (c) the Contractor shall give the Engineer not less than 28 days' **notice** of the intended date of the commencement of each Subcontractor's work, …	16	Contractor	Engineer	–	Commencement of Subcontractor's work.	≥ 28 days
			… and [the Contractor shall give the Engineer not less than 28 days' **notice** of the intended date] of the commencement of such work on the Site;	17	Contractor	Engineer	–	Commencement of Subcontractor's work on Site.	≥ 28 days

Subsequent Action	Party Responsible for Subsequent Action	Time Bar for the Subsequent Action	Notice Description	Type of Notice	Works' Relevant Aspects	Notice/Event Sequence	Structured or Non-structured Notice	Notes on Notice-Event Sequence and Structured Notices
The Engineer shall then proceed in accordance with Sub-Clause 3.5 [*Determinations*] to agree or determine (i) the amount (if any) which the Employer is entitled to be paid by the Contractor, and/or (ii) the extension (if any) of the Defects Notification Period in accordance with Sub-Clause 11.3 [*Extension of Defects Notification Period*]. This amount may be included as a deduction in the Contract Price and Payment Certificates. The Employer shall only be entitles to set off against or make any deduction from an amount certified in a Payment Certificate, or to otherwise claim against the Contractor, in accordance with this Sub-Clause.	Engineer	Not specified	Employer's Claim Notice	Claim Notice	Multiple	Event-succeeding Notice	Structured/Non-structured Notice	It is an Event-succeeding Notice since it issued after the occurrence of events described in the Clauses relating to th Sub-Clause or otherwise in connection with the Contract. This notice might be either Structured, it is preceded by a notice, or Non-structured.
The Contractor may raise an objection (as explicitly inferred).	Contractor	Not specified but implied to be issued prior to the intended replacement becoming effective.	Engineer's Replacement Notice	Attention Notice	Personnel	Event-preceding Notice	Structured Notice	The notice shall precede the replaceme of the Engineer. Notice is structured as its time bar itse represents the time within which objection by the Contractor is to be rais prior to the Engineer's replacement becoming effective.
The replacement of the Engineer by Employer may be hindered. Employer is expected to react to this notice.	Employer	_	Contractor's Objection to Engineer's Replacement Notice	Objection Notice	Personnel	Dual-event Notice	Structured Notice	This is a Dual Notice since it is preceded the intension of the Employer to repla the Engineer and is followed by anothe event which is the reaction by the Employer and the replacement becomi effective. Notice is structured since when it is issued, it has the effect of splitting the time bar inherent to the notice preced it, therefore setting forth the time peri remaining for possible reaction by the other party (Employer) before the Engineer's replacement becoming effective.
Each party shall give effect to each agreement or determination unless and until revised under Clause 20 [Claims, Disputes and Arbitration].	Party	Not specified	Determination Notice	Determination Notice	Contractual	Event-succeeding Notice	Structured/Non-structured Notice	Following the different events that ref to determination under this Sub-Clause this notice might be either Structured a is preceded by a notice, or Non-structur since some of the cases no notice is required prior to issuing the determination notice (Figure #3a).
For Engineer's information and implied consent.	Engineer**	_	Alteration in Execution Methods Notice	Attention Notice	Site - Execution of Works	Event-preceding Notice	Non-structured Notice	Notice shall precede the event of alteri any arrangements and methods for th execution of works.
According to the Engineer's consent, the replacement shall take place.	Engineer	_	Contractor's Representative Replacement Notice	Attention Notice	Personnel	Event-preceding Notice	Non-structured Notice	Notice shall be issued for obtaining th Engineer's consent prior to effecting an replacement.
For Engineer's information and implied consent.	Engineer**	_	Contractor's Representative Power Delegation/ Revocation Notice	Information Notice	Personnel	Event-preceding Notice	Non-structured Notice	Notice shall be received prior to the delegation or revocation of powers becoming effective.
For Engineer's information.	Engineer*	_	Subcontractors' Subcontract Work Commencement Notice	Information Notice	Commencement of Works	Event-preceding Notice	Non-structured Notice	Not less than 28 days prior to commencement of Subcontractors' wo the Contractor shall inform the Engine by issuing a notice (preceding).
For Engineer's information.	Engineer*	_	Subcontractors' On-site Work Commencement Notice	Information Notice	Commencement of Works	Event-preceding Notice	Non-structured Notice	Not less than 28 days prior to commencement of Subcontractors' wo on site, the Contractor shall inform th Engineer by issuing a notice (preceding

Continuous (s,S) Inventory Policy with Non-Stationary Stochastic Demand

Ibrahim Elshar

Industrial Engineering and Management

Faculty of Engineering and Architecture

American Univeristy of Beirut

Email: ije04@mail.aub.edu

Abstract—**This work utilizes Markovian Processes to model the stochastic and time-dependent performance measures of inventory systems. We present an approach to identify and capture the key characteristics of the demand process by fitting an approximate Markovian distribution to the demand process. Inventory replenishment policies are investigated for the (s_t, S_t) continuous case for a finite horizon model. A Markovian representation of the system is presented along with the differential equations that model the behavior of the system over time. We present an algorithmic approach to numerically solve the corresponding differential equations to calculate the performance measures of the system over time. We then propose an optimization heuristic to compute the dynamic time dependent reorder and order up-to levels (s_t, S_t). Finally, we test our findings using a numerical example based on real inventory data.**

I. Introduction

Most recent studies on inventory management focused on incorporating more realistic assumptions in their inventory models in order to reduce the gap between theory and practice. A realistic assumption is related to the demand and its non-stationary (time varying) vs. stationary (steady) nature. Inventory models that take into consideration the non-stationary nature of the demand is rather limited in the literature especially when compared to the body of work involving stationary demand. In practice, most real world applications of inventory management involve non-stationary demands, [21]. To the best of our knowledge, no one has studied the non-stationary demand by approximating it with a phase-type process. In this paper, our analytical model is similar to the one used by Nasr and Maddah [16] but instead of using a stationary Markov Modulated Poisson Process (MMPP) correlated demand we first use a general time-dependent phase-type process and then a two-level mixture of Erlangs of common order phase-type process to serve as an approximation of the non-stationary demand process. The 2-MECO phase-type distribution is adopted from Nasr & Taaffe [17] which is utilized to fit the $Ph_t/M_t/s/c$ time-dependent departure process. We implement the fitting approach presented in [17] to fit the time-dependent demand process to a 2-MECO. An efficient heuristic is then proposed to compute the optimal reorder and order up-to levels (s_t, S_t) at each period.

II. Literature Review

In this section, we review the literature related to our subject. We divide this section into two subsections. In section II-A, we review the Markovian modulated demand literature; and in section II-B we review the literature about non-stationary demand.

A. Markovian Modulated Demand Literature

More realistic inventory models that consider demand as a random variable that is dependent on environmental factors other than time was given in 1960 by Karlin and Fabens [12]. The authors used a Markovian demand model that is unlike most classical inventory models depends on the state-of-the-world in each period. This takes into consideration the effects of the randomly changing environmental factors like the fluctuating economic and the uncertain market conditions on the demand which in turn affects the cost functions. The demand is modulated as a random variable having a distribution function that depends on the demand state in each period. Iglehart and Karlin [10] used a discrete Markov processes to model dependent demand with no setup costs and managed to prove the optimality of a state-dependent base-stock policy. Song and Zipkin [22] show that for a continuous review, state-dependent, Markov-modulated, Poisson distributed demand having linear costs, fixed order cost and backlogging in inventory, the optimality of a (s,S)-type policy. Zipkin [25], Aviv & Federgruen [2] and Kapuscinski & Tayur [11] investigated cyclical demand models. Beyer & Sethi [3], Sethi & Cheng [19] proved the optimality of (s,S)-type policy for a generalized Markovian demand distribution and a periodic review inventory model. Cheng & Sethi [7] made an extension to their results by considering lost sales for unsatisfied demand. All proofs were given assuming a zero lead time distribution. Chen & Song [6] studied the optimal policies for multistage inventory problems with non-stationary Markov-modulated Poisson demand process at which the demand process is governed by a discrete time Markov chain. An effective algorithm for the determination of the optimal base-stock is provided. Abhyankar & Graves [1] studied a two-stage serial supply chain with a two-state Markov-modulated Poisson demand and were able to develop an optimization model to determine where it is best to hedge inventory. Nasr & Maddah [16] utilized a Markov Modulated Poisson Process (MMPP) to model stochastic demand that is dependent on the state of the environment and has a fixed lead time. They were able to compute and provide an efficient optimization heuristic

to derive the dynamic changing (s_n, S_n) policy for a single item, continuous and infinite horizon inventory problems. They studied and quantified the impact of autocorrelation of the MMPP demand-count process that is causing the variability in the demand. Their results show that when the demand is highly correlated the dynamically changing (s_n, S_n) policy significantly outperforms the common static heuristically computed replenishment (s, S) policies.

B. Non-Stationary Demand Literature

The literature on stochastic time-dependent demand in the context of inventory systems is scarce in comparison to stationary demand. Non-stationary demand can be due to seasonality or can be represented by a time varying function. Though the demand pattern may be stationary, the fact that the demand is stochastic makes the future demand unknown. However, most studies on the dynamic lot-sizing assume that future demand is known, [9]. At the present time as the product life cycles are becoming shorter and a large variety of products are being introduced to the market and thus affecting customers order stability, the demands faced tend to be more non-stationary. The reason behind the small number of papers that deal with non-stationary demand relative to the large number of papers that deal with stationary demand is that non-stationary demand models are hard to compute because of their irregular structure. The literature on the non-stationary demand can be categorized in to two categories: papers that focus on the optimization and papers that focus on the performance evaluation of the proposed policies. Even though the (s,S) policy has been shown optimal for inventory problems with stationary and non-stationary demand, [14], [13], [18], not much work has been done for computing non-stationary (s,S) policies. The inventory problem with non-stationary demand was discussed by Silver [20], Bookbinder & Tan [5] and Tarim & Kingsman [23] by using alternative policies such as the (R_n, S_n) policy. Silver [20] considered the stochastic time-varying demand by assuming normally distributed forecast errors. A procedure involving sequential optimization is used then to find the periods in which to place orders, the number of periods of the horizon that has to be included in the next order and the size of the order, with the uncertainty of demands in these periods are given. This research is extended by Bookbinder & Tan [5] by considering a rolling horizon with updated demand information. Tarim & Kingsman [23] improved Bookbinder & Tan [5] heuristic by making further improvements. Other heuristics used to compute near optimal (s,S) parameters, is by Bollapragada & Morton [4]. Bollapragada & Morton [4] were able to compute optimal (s_n, S_n) levels for a single stage inventory problem facing a general non-stationary demand with proportional backorder and holding costs. They proposed a new myopic heuristic that is based on approximating part of the initial non-stationary problem with a stationary one and involves computing the static (s, S) policy replenishment values for different demand parameters from which the dynamic replenishment (s_t, S_t) policy of the non-stationary problem for period t is approximated by restricting the state space of the inventory position at the beginning of each time period to integer values. The main development behind this research is computing the upper and lower bounds of the optimal policy efficiently. Ettl et al. [8] worked on minimizing the total expected inventory capital and on approximating the replenishment lead-times in a multistage inventory system. They modeled non-stationary demand by making the assumption that the horizon is made up of a set of stationary phases which form a rolling-horizon and then finding the optimal policy for each phase. Finally, it can be seen that the amount literature taking non-stationary demand into consideration is increasing. Tunc et al. [24] studied the cost of using stationary inventory policies when demand is non-stationary. They took the (s,S) policy as a frame of reference, and they compared the optimal non-stationary (s,S) policy with the best possible stationary (s,S) policy in terms of cost performance. They showed that the cost of neglecting the non-stationarity of demand is significantly high for the majority of cases. There numerical study reveals that, the magnitude of the sub-optimality of stationary policies depends heavily on the variation of the demand pattern, i.e. the non-stationarity of demand, among other factors, such as, the stochasticity of demand, and cost parameters.

III. MODEL

The objective of this study is to model the non-stationary stochastic demand process and to determine the optimal reorder and order up-to levels (s_t, S_t).

A. Demand Process as Phase-Type:(Time Dependent)

In this section, we present an algorithm for computing the non-stationary demand distribution over the lead time by deriving the demand-count distribution conditional on the phase-type process being in a given state at the beginning of the lead time, and the related conditional moments at time t. This algorithm is based on numerically solving a set of differential equations (Kolmogorov Forward Equations, KFEs). First, we define our demand model using the general phase-type process then we introduce the two-level mixture of Erlangs of common order (2-MECO) as defined by Nasr & Taaffe [17] in their paper about fitting the $Ph_t/M_t/s/c$ time-dependent departure process for use in tandem queueing networks. The 2-MECO demand model will be used later for our numerical example.

1) The General Ph_t Demand-Count Process: In general, we define the phase-type process consisting of m transient states $S_T = \{1, 2, \ldots, m\}$ and a single absorbing state $S_A = \{m + 1\}$. We write the infinitesimal generator matrix $\mathbf{Q(t)}$ at time $t \geq 0$ and the probability row vector $\boldsymbol{\alpha(t)}$ of the phase-type process as

$$\mathbf{Q(t)}_{(m+1,m+1)} = \begin{bmatrix} \boldsymbol{\lambda(t)} & \boldsymbol{\mu(t)} \\ {\scriptstyle (m \times m)} & {\scriptstyle (m \times 1)} \\ \mathbf{0} & 0 \\ {\scriptstyle (1 \times m)} & \end{bmatrix}$$

Where $\boldsymbol{\lambda(t)}$ is a $(m \times m)$ sub-matrix containing the transition rates at time $t \geq 0$ describing the transitions between

the transient states. where $\lambda_{i,i}(t) = -\lambda_i(t)$ and $\lambda_i(t) = \sum_{\substack{j=1 \\ j \neq i}}^{m} \lambda_{i,j}(t) + \mu_i(t)$ for all $i = 1, \ldots, m$.

The $(m \times 1)$ vector $\boldsymbol{\mu(t)}$ contains the transition rates $\mu_i(t)$ for $i = 1, \ldots, m$ and $t \geq 0$ from the transient states to the absorbing state. It is evident that the values of $\lambda_{i,i}(t)$, $\lambda_{i,j}(t)$ and $\mu_i(t)$ are dependent on the time $t \geq 0$ from which we intend to exploit the time-dependency in our phase-type model. The row vector $\mathbf{0}$ is a vector consisting entirely of zeros since no transitions are allowed from the absorbing state to the transient states. The last element of the matrix is 0 which represents the transition rate out of the absorbing state. The row vector $\boldsymbol{\alpha(t)}$ contains the initial probabilities that the embedded CTMC starts initially at a time t in transient state $i = 1, \ldots, m$ or directly starts in the absorbing state $m+1$. We set $\alpha_{m+1}(t) = 0$ so the sum $\sum_{i=1}^{m} \alpha_i(t) = 1$ for all $t \geq 0$.

We first write the KFE of finding the general Ph_t process at state $\{A(t) = n\}$ at time $t \geq 0$ and $n = 1, \ldots, m$.

$$P'(A(t) = n) = \sum_{w=1}^{m} \Big(\lambda_{w,n}(t) + \alpha_n(t)\,\mu_w(t) \Big) P(A(t) = w) \quad (1)$$

To calculate the probability $P(A(t) = n)$ of finding the general Ph_t process at state $\{A(t) = n\}$ for $n = 1, \ldots, m$ at time $t \geq 0$, we solve the differential equation in (1) with the following initial condition, $P(A(0) = n) = \alpha_n(0)$.

Let $\{D_0(t) = k : 0 \leq t \leq L\}$ be the demand-count over the time interval $[0, L]$ and $\{D_{t-L}(\tau) = k : t - L > 0, \tau > 0\}$ be the demand-count over the time interval $[t - L, t - L + \tau]$ where $k = 0, 1, 2, 3, \ldots, \infty$ and L is the fixed Lead time. We augment the demand-count process with the state of the Ph_t process, $\{A(t), t \geq 0\}$, and we define the state space of the demand-count and arrival-phase at time t for $0 \leq t \leq L$ and $t - L + \tau$ for $t - L > 0$ by $\{D_0(t) = k, A(t) = i\}$ and $\{D_{t-L}(\tau) = k, A(t - L + \tau) = i\}$ respectively. The KFEs of the augmented state space are presented below where the derivative is with respect to t in equation (eq.) (2) and τ in eq. (3).

For $0 \leq t < L$, $k = 0, 1, 2, \ldots, \infty$ and all $i = 1, 2, ..., m$,

$$P'(D_0(t) = k, A(t) = i) =$$
$$\sum_{w=1}^{m} \lambda_{w,i}(t)\,P(D_0(t) = k, A(t) = w) \quad (2)$$
$$+ \sum_{w=1}^{m} \alpha_i(t)\,\mu_w(t)\,P(D_0(t) = k - 1, A(t) = w)\,I_{(k>0)}.$$

For $t \geq L$, $k = 0, 1, 2, \ldots, \infty$ and all $i = 1, 2, ..., m$, let $t_s = t - L$,

$$P'(D_{t_s}(\tau) = k, A(t_s + \tau) = i) =$$
$$\sum_{w=1}^{m} \lambda_{w,i}(t_s + \tau)P(D_{t_s}(\tau) = k, A(t_s + \tau) = w) \quad (3)$$
$$+ \sum_{w=1}^{m} \alpha_i(t_s + \tau)\,\mu_w(t_s + \tau)$$
$$\times P(D_{t_s}(\tau) = k - 1, A(t_s + \tau) = w)\,I_{(k>0)}$$

where $I_{(k>0)}$ is an indicator function such that,

$$I_{(k>0)} = \begin{cases} 1 & \text{when } k > 0, \\ 0 & \text{otherwise.} \end{cases}$$

2) The 2-MECO Demand-Count Process: The motivation behind adopting the 2-MECO phase-type process lies in its flexibility to match high and low variability processes over any time interval $[t, t + \tau]$, [17]. A count process $D_t(\tau)$ is said to have a high/low variability over an interval $[t, t + \tau]$ if $Var(D_t(\tau))(> / <)E[D_t(\tau)]$, where $Var(D_t(\tau))$ is the variance of $D_t(\tau)$. Now we present the 2-MECO process and its properties as defined by Nasr & Taaffe [17]. The 2-MECO phase-type process consists of two Erlang branches. The number of states in level 1 and 2 is m_1 and m_2, respectively. Thus, we have a total of $m = m_1 + m_2$ transient states $S_T = \{1, 2, \ldots, m_1 + m_2\}$ and a single absorbing state. α and $1 - \alpha$ are the probabilities of starting at phase 1 and $m_1 + 1$, respectively. The transition rates between the states of the 1st and 2nd Erlang branch is $m_1\lambda_1$ and $m_2\lambda_2$, respectively. We write the infinitesimal generator matrix $\mathbf{Q}(t)$ at time $t \geq 0$ and the probability row vector $\boldsymbol{\alpha(t)}$ of the 2-MECO phase-type process as

$$\mathbf{Q(t)} = \begin{bmatrix} \mathbf{Q_1(t)} & 0 \\ {\scriptstyle (m_1 \times m_1)} & \\ 0 & \mathbf{Q_2(t)} \\ & {\scriptstyle (m_2 \times m_2)} \end{bmatrix}$$

$$\boldsymbol{\alpha(t)} = [\alpha(t), 0, \ldots, 0, (1 - \alpha(t)), 0, \ldots, 0]$$

Where matrix $\boldsymbol{Q_i(t)}$ for $i = 1, 2$ represents the infinitesimal matrix of the ith Erlang branch. Now, we write the KFEs of finding the 2-MECO Ph_t process at state $\{A(t) = n\}$ at time $t \geq 0$, for $n = 1, \ldots, m_1$,

$$P'(A(t) = n) = -m_1\lambda_1(t)\,P(A(t) = n)$$
$$+ m_1\lambda_1(t)\,P(A(t) = n - 1)(1 - I_{(n=1)})$$
$$+ \alpha(t)\Big(m_1\lambda_1(t)\,P(A(t) = m_1) \quad (4)$$
$$+ m_2\lambda_2(t)P(A(t) = m_1 + m_2)\Big)I_{(n=1)},$$

and for $n = m_1 + 1, \ldots, m_1 + m_2$,

$$P'(A(t) = n) = -m_2\lambda_2(t)\,P(A(t) = n)$$
$$+ m_2\lambda_2(t)\,P(A(t) = n - 1)(1 - I_{(n=m+1)})$$
$$+ (1 - \alpha(t))\Big(m_1\lambda_1(t)\,P(A(t) = m_1) \quad (5)$$
$$+ m_2\lambda_2(t)\,P(A(t) = m_1 + m_2)\Big)I_{n=(m_1+1)}$$

where $I_{(i=j)}$ is an indicator function such that $\forall\, i, j$,

$$I_{(i=j)} = \begin{cases} 1 & \text{when } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

To calculate the probability $P(A(t) = n)$ of finding the 2-MECO Ph_t process at state $\{A(t) = n\}$ for $n = 1, \ldots, m$, where $m = m_1 + m_2$ at time $t \geq 0$, we solve the differential equation in (4) and (5) with the following initial condition, $P(A(0) = n) = \alpha_n(0)$.

The 2-MECO demand-count KFEs of the augmented state space are presented below where the derivative is with respect to t in equations (6) and (7), and τ in (8) and (9). For $0 \leq t < L$, $k = 0, 1, 2, \ldots, \infty$, and $i = 1, \ldots, m_1$,

$$
\begin{aligned}
P'(D_0(t) &= k, A(t) = i) = \\
&- m_1\lambda_1(t)\, P(D_0(t) = k, A(t) = i) \\
&+ m_1\lambda_1(t)\, P(D_0(t) = k, A(t) = i - 1) \times (1 - I_{(i=1)}) \\
&+ \alpha(t)\Big(m_1\lambda_1(t)\, P(D_0(t) = k - 1, A(t) = m_1) \\
&+ m_2\lambda_2(t)\, P(D_0(t) = k - 1, A(t) = m_1 + m_2)\Big) \times I_{(k>0)}.
\end{aligned}
\tag{6}
$$

For $0 \leq t < L$, $k = 0, 1, 2, \ldots, \infty$, and $i = m_1 + 1, \ldots, m_1 + m_2$,

$$
\begin{aligned}
P'(D_0(t) &= k, A(t) = i) = \\
&- m_2\lambda_2(t)\, P(D_0(t) = k, A(t) = i) \\
&+ m_2\lambda_2(t)\, P(D_0(t) = k, A(t) = i - 1) \times (1 - I_{(i=m_1+1)}) \\
&+ (1 - \alpha(t))\Big(m_1\lambda_1(t)\, P(D_0(t) = k - 1, A(t) = m_1) \\
&+ m_2\lambda_2(t)\, P(D_0(t) = k - 1, A(t) = m_1 + m_2)\Big) \times I_{(k>0)}.
\end{aligned}
\tag{7}
$$

For $t \geq L$, $k = 0, 1, 2, \ldots, \infty$, $t_s = t - L$, and $i = 1, ..., m_1$,

$$
\begin{aligned}
P'(D_{t_s}(\tau) &= k, A(t_s + \tau) = i) = \\
&- m_1\lambda_1(t_s + \tau)\, P(D_{t_s}(\tau) = k, A(t_s + \tau) = i) \\
&+ m_1\lambda_1(t_s + \tau)\, P(D_{t_s}(\tau) = k, A(t_s + \tau) = i - 1) \times (1 - I_{(i=1)}) \\
&+ \alpha(t_s + \tau)\Big(m_1\lambda_1(t_s + \tau)\, P(D_{t_s}(\tau) = k - 1, A(t_s + \tau) = m_1) \\
&+ m_2\lambda_2(t_s + \tau)\, P(D_{t_s}(\tau) = k - 1, A(t_s + \tau) = m_1 + m_2)\Big) \times I_{(k>0)}.
\end{aligned}
\tag{8}
$$

For $t \geq L$, $k = 0, 1, 2, \ldots, \infty$, $t_s = t - L$, and $i = m_1 + 1, ..., m_1 + m_2$,

$$
\begin{aligned}
P'(D_{t_s}(\tau) &= k, A(t_s + \tau) = i) = \\
&- m_2\lambda_2(t_s + \tau)\, P(D_{t_s}(\tau) = k, A(t_s + \tau) = i) \\
&+ m_2\lambda_2(t_s + \tau)\, P(D_{t_s}(\tau) = k, A(t_s + \tau) = i - 1)(1 - I_{(i=m_1+1)}) \\
&+ (1 - \alpha(t_s + \tau))\Big(m_1\lambda_1(t_s + \tau)\, P(D_{t_s}(\tau) = k - 1, A(t_s + \tau) = m_1) \\
&+ m_2\lambda_2(t_s + \tau)\, P(D_{t_s}(\tau) = k - 1, A(t_s + \tau) = m_1 + m_2)\Big) I_{(k>0)}.
\end{aligned}
\tag{9}
$$

The probability distribution of the Conditional Demand-Count (CDC) process can be calculated from the KFEs in (2) and (3) for the general Ph_t distribution and from the KFEs in (6),(7),(8) and (9) for the 2-MECO distribution by conditioning on the state of the Ph_t process at time $t \in [0, T]$.

After which we can find the distribution of the demand-count process by applying the total probability theorem. Thus, in order to find the probability $P(D_0(t) = k)$ of having a demand count equal to $k = 0, 1, 2, \ldots, \infty$ over the time interval $0 \leq t \leq L$, we first find the probabilities $P(D_0(t) = k | A(0) = n)$ of having a demand equal to $k = 0, 1, 2, \ldots, \infty$ over the interval $0 \leq t \leq L$, conditioned on being at all possible phases at time $t = 0$, i.e., for all $n = 1, \ldots, m$. Note that in the general Ph_t distribution m in the number of transient states which is also the case in the 2-MECO distribution where $m = m_1 + m_2$. Hence, $P(D_0(t) = k)$ will be equal to the sum $\sum_{n=1}^{m} \Big(P(D_0(t) = k | A(0) = n) \times P(A(0) = n) \Big)$.

Similarly, to find the probability $P(D_{t_s}(L) = k)$, with $t_s = t - L$, of having a demand count equal to $k = 0, 1, 2, \ldots, \infty$ over the time interval $[t_s, t]$ such that $t_s \geq 0$, we first find the probabilities $P(D_{t_s}(L) = k | A(t_s) = n)$ of having a demand equal to $k = 0, 1, 2, \ldots, \infty$ over the interval $[t_s, t]$ conditioned on being at all possible phases at the beginning of the lead time t_s, i.e., for all $n = 1, \ldots, m$. Hence, $P(D_{t_s}(L) = k) = \sum_{n=1}^{m} \Big(P(D_{t_s}(L) = k | A(t_s) = n) \times P(A(t_s) = n) \Big)$.

To facilitate the computation later on in our model, specifically when computing the net inventory probabilities at time $t \geq 0$ at which we need to compute probabilities for all the possible conditional demand-count values over the lead time, we limit the demand-count to an upper limit d_{max}. The upper bound d_{max} is chosen such that it satisfies $P(X > d_{max}) = \epsilon$ where X is a Poisson random variable having a mean $\lambda = \max(\lambda_{i,j}(t))$ for all $i = 1, \ldots, m$ and $t \in [0, T]$, where ϵ is a sufficiently small number. Note that for the 2-MECO, $\lambda_{i,j}(t)$ is either $m_1\lambda_1(t)$ or $m_2\lambda_2(t)$. It is worth to mention that this approach for selecting d_{max} preserves the independence of the CDC and Demand-count with the reorder policy.

Accordingly, we establish an algorithm to compute the probability distribution of the demand count process for a time interval $[0, T]$.

3) Demand-Count Distribution Algorithm: For $0 \leq t < L$ and $\forall\, k \in \{0, 1, \cdots, d_{max}\}$,

Step 1. Set $n = 1$ and $i = 1$.

Step 2. Numerically solve equations (2), (6) and (7) from 0 to t, $t < L$ with the following initial condition, $P(D_0(0) = 0, A(0) = n) = P(A(0) = n)$, to obtain $P(D_0(t) = k, A(t) = i | A(0) = n)$.

Step 3. If $i < m$, set $i = i + 1$ and go to Step 2.

Step 4. Set $P(D_0(t) = k | A(0) = n)$

$$= \sum_{i=1}^{m} P(D_0(t) = k, A(t) = i | A(0) = n)$$

Step 5. If $n < m$, set $n = n + 1$ and $i = 1$ and go to Step 2.

Step 6. Set $P(D_0(t) = k)$

$$= \sum_{n=1}^{m} \Big(P(D_0(t) = k | A(0) = n) \times P(A(0) = n) \Big).$$

For $t \geq L$ and $\forall\, k \in \{0, 1, \ldots, d_{max}\}$,

Step 1. Set $n = 1$ and $i = 1$.

Step 2. Numerically solve equations (3),(8) and (9) from $t_s = t - L$ to t, $t \geq L$ with the following initial condition, $\mathrm{P}(D_{t_s}(0) = 0, A(t_s) = n) = \mathrm{P}(A(t_s) = n)$, to obtain $\mathrm{P}(D_{t_s}(L) = k, A(t) = i | A(t_s) = n)$.

Step 3. If $i < m$, set $i = i + 1$ and go to Step 2.

Step 4. Set $\mathrm{P}(D_{t_s}(L) = k | A(t_s) = n)$
$$= \sum_{i=1}^{m} \mathrm{P}(D_{t_s}(L) = k, A(t) = i | A(t_s) = n).$$

Step 5. If $n < m$, set $n = n + 1$ and $i = 1$ and go to Step 2.

Step 6. Set $\mathrm{P}(D_{t_s}(L) = k) =$
$$\sum_{n=1}^{m} \Big(\mathrm{P}(D_{t_s}(L) = k | A(t_s) = n) {\times} \mathrm{P}(A(t_s) = n) \Big).$$

Next, we discuss finding the moments of the demand over the lead time. The rth moment of the demand-count distribution over the interval $[0, t]$ where $t < L$ can be found by

$$\mathrm{E}[D_0^r(t)] = \sum_{k=1}^{d_{max}} k^r \, \mathrm{P}(D_0(t) = k) \tag{10}$$

and over an interval $[t_s, t]$ for $t_s = t - L$, $t_s \geq 0$ by

$$\mathrm{E}[D_{t_s}^r(L)] = \sum_{k=1}^{d_{max}} k^r \, \mathrm{P}(D_{t_s}(L) = k). \tag{11}$$

In the next section, we will present the analytical model of the inventory characteristics for a given $(\bar{s}, \bar{S})$ replenishment policy.

B. Inventory Position

The replenishment policy at time t is defined by the reorder level $s(t)$ and the order up-to level $S(t)$. For time $t = n\,\delta$, we represent the reorder policy $(s(t), S(t))$ by (s_n, S_n) for $n = 1, 2, \cdots, N$, where $\delta = \frac{\mathrm{T}}{N}$ such that T = Horizon Length and N = number of periods.

For a given dynamic reorder policy, $(\bar{s}, \bar{S})$ where $\bar{s} = \{s_1, s_2, \ldots, s_N\}$ and $\bar{S} = \{S_1, S_2, \ldots, S_N\}$, the state of the dynamic inventory position at time $t \in [0, T]$ is represented by the value of the inventory position and the state of the phase-type process at time $t \in [0, T]$.

The inventory position at time $t \in [0, T]$ is $\mathrm{IP}(t) \in \{\ell, \ldots, u\}$, where ℓ is the lowest achievable inventory position level such that $\ell = \min_{t \in [0, \mathrm{T}]} (s_t) + 1$ and u is the inventory position upper limit such that $u = \max_{t \in [0, \mathrm{T}]} (S_t)$. We augment the states of the Ph_t demand count process $A(t) \in \{1, 2, \cdots, m\}$ with the inventory position level $\mathrm{IP}(t)$ at time $t \geq 0$.

Let $P_{i,n}(t) = P(\mathrm{IP}(t) = i, A(t) = n)$ for $i = \ell, \ldots, u$ and $n = 1, \ldots, m$, be the resulting Markovian process state probabilities. Recall that m is the total number of transient states in the general Ph_t distribution and the 2-MECO distribution where $m = m_1 + m_2$. Similar to the demand-count process the inventory position level can be defined by a general and

a 2-MECO Ph_t processes. The resulting inventory position Kolmogorov Forward Equations (IP-KFEs) for the general Ph_t and 2-MECO processes are presented in equations (12) to (19).

1) The General Ph_t IP Process KFEs: For $t \geq 0$, $n = 1, \ldots, m$ and $i > s(t)$,

$$
\begin{aligned}
\mathrm{P}'_{i,n}(t) = &\sum_{w=1}^{m} \lambda_{w,n}(t) \mathrm{P}_{i,w}(t) \\
&+ \alpha_n(t) \Big(\sum_{w=1}^{m} \mu_w(t) \mathrm{P}_{i+1,w}(t) \Big)(1 - \mathrm{I}_{(i=u)}) \\
&+ \alpha_n(t) \Big(\sum_{w=1}^{m} \sum_{q=\ell}^{s(t)+1} \mu_n(t) \mathrm{P}_{q,n}(t) \Big) \mathrm{I}_{(i=S(t))}.
\end{aligned}
\tag{12}
$$

For $t \geq 0$, $n = 1, \ldots, m$ and $i \leq s(t)$,

$$\mathrm{P}'_{i,n}(t) = \sum_{w=1}^{m} \lambda_{w,n}(t) \, \mathrm{P}_{i,w}(t). \tag{13}$$

2) The 2-MECO Ph_t IP Process KFEs: For $t \geq 0$, $n = 2, \ldots, m_1$ and $\ell \leq i \leq u$,

$$\mathrm{P}'_{i,n}(t) = -m_1\,\lambda_1(t)\,\mathrm{P}_{i,n}(t) + m_1\,\lambda_1(t)\,\mathrm{P}_{i,n-1}(t). \tag{14}$$

For $t \geq 0$, $n = m_1 + 2, \ldots, m_1 + m_2$ and $\ell \leq i \leq u$,

$$\mathrm{P}'_{i,n}(t) = -m_2\,\lambda_2(t)\,\mathrm{P}_{i,n}(t) + m_2\,\lambda_2(t)\,\mathrm{P}_{i,n-1}(t). \tag{15}$$

For $t \geq 0$, $n = 1$ and $i > s(t)$,

$$
\begin{aligned}
\mathrm{P}'_{i,1}(t) = &-m_1\,\lambda_1(t)\,\mathrm{P}_{i,1}(t) + \alpha(t) \Big(m_1\,\lambda_1(t)\,\mathrm{P}_{i+1,m_1}(t) \\
&+ m_2\,\lambda_2(t)\,\mathrm{P}_{i+1,m_1+m_2}(t) \Big)(1 - \mathrm{I}_{(i=u)}) \\
&+ \alpha(t) \Big(\sum_{w=\ell}^{s(t)+1} (m_1\,\lambda_1(t)\,\mathrm{P}_{w,m_1}(t) \\
&+ m_2\,\lambda_2(t)\,\mathrm{P}_{w,m_1+m_2}(t)) \Big) \mathrm{I}_{(i=S(t))}.
\end{aligned}
\tag{16}
$$

For $t \geq 0$, $n = m_1 + 1$ and $i > s(t)$,

$$
\begin{aligned}
\mathrm{P}'_{i,m_1+1}(t) = &-m_2\,\lambda_2(t)\,\mathrm{P}_{i,m_1+1}(t) \\
&+ (1 - \alpha(t)) \Big(m_1\,\lambda_1(t)\,\mathrm{P}_{i+1,m_1}(t) \\
&+ m_2\,\lambda_2(t)\,\mathrm{P}_{i+1,m_1+m_2}(t) \Big)(1 - \mathrm{I}_{(i=u)}) \\
&+ (1 - \alpha(t)) \Big(\sum_{w=\ell}^{s(t)+1} (m_1\,\lambda_1(t)\,\mathrm{P}_{w,m_1}(t) \\
&+ m_2\,\lambda_2(t)\,\mathrm{P}_{w,m_1+m_2}(t)) \Big) \mathrm{I}_{(i=S(t))}.
\end{aligned}
\tag{17}
$$

For $t \geq 0$, $n = 1$ and $i \leq s(t)$,

$$\mathrm{P}'_{i,1}(t) = -m_1\,\lambda_1(t)\,\mathrm{P}_{i,1}(t). \tag{18}$$

For $t \geq 0$, $n = m_1 + 1$ and $i \leq s(t)$,

$$\mathrm{P}'_{i,m_1+1}(t) = -m_2\,\lambda_2(t)\,\mathrm{P}_{i,m_1+1}(t). \tag{19}$$

where, $I_{i=n} = \begin{cases} 1 & \text{when } i = n, \\ 0 & \text{otherwise.} \end{cases}$

The KFEs in (12) and (13) for the general Ph_t distribution and the KFEs in (14) to (19) for the 2-MECO Ph_t distribution allow us to compute the joint probability distribution of the inventory position and the Ph_t state at time $t \geq 0$. After which the distribution of the inventory position at time $t \geq 0$ can be found by the sum of the joint probability distribution for all phases. Thus, in order to find the probability $P(IP(t) = i)$ of having an inventory position equal to $i = \ell, \ldots, u$ at time $t \geq 0$, we first find the probabilities $P(IP(t) = i, A(t) = n)$ of having an inventory position equal to $i = \ell, \ldots, u$ and being in phase $n = 1, \ldots, m$ at time $t \geq 0$. Then, the $P(IP(t) = i)$ will be equal to the sum $\sum_{n=1}^{m} P(IP(t) = i, A(t) = n)$. Finally, the rth moment of the inventory position at time $t \geq 0$ can be found by

$$E[IP^r(t)] = \sum_{i=\ell}^{u} i^r \, P(IP(t) = i) \qquad (20)$$

Next, we will find the net inventory position probabilities and its moments.

C. Net Inventory Position

Now, we present the equations which allow us to find the distribution of the net inventory position at time $t \geq 0$. As thoroughly discussed in [15], it can be seen that

$$NI(t) = IP(0) - D_0(t) \quad \text{for } 0 \leq t \leq L, \qquad (21)$$

$$NI(t) = IP(t - L) - D_{t-L}(L) \quad \text{for } t > L. \qquad (22)$$

Which allows us to write,
for $0 \leq t \leq L$ and $\ell - d_{max} \leq i \leq u$,

$$P(NI(t) = i) = \sum_{y=\ell}^{u} P(IP(0) = y, D_0(t) = y - i), \quad (23)$$

for $t > L$ and $\ell - d_{max} \leq i \leq u$,

$$P(NI(t) = i) = \sum_{y=\ell}^{u} P(IP(t - L) = y, D_{t-L}(L) = y - i). \qquad (24)$$

By conditioning equations (23) and (24) on the Ph_t system state $n = \{1, \ldots, m\}$ we obtain: for $0 \leq t \leq L$ and $\ell - d_{max} \leq i \leq u$,

$$P(NI(t) = i | A(0) = n) = \sum_{y=\ell}^{u} \Big(P(IP(0) = y | A(0) = n) \\ \times P(D_0(t) = y - i | A(0) = n) \Big), \qquad (25)$$

for $t > L$ and $\ell - d_{max} \leq i \leq u$,

$$P(NI(t) = i | A(t - L) = n) = \\ \sum_{y=\ell}^{u} \Big(P(IP(t - L) = y | A(t - L) = n) \\ \times P(D_{t-L}(L) = y - i | A(t - L) = n) \Big). \qquad (26)$$

Which can be written as,

for $0 \leq t \leq L$ and $\ell - d_{max} \leq i \leq u$,

$$P(NI(t) = i, A(0) = n) = \sum_{y=\ell}^{u} \Big(P(IP(0) = y, A(0) = n) \\ \times P(D_0(t) = y - i | A(0) = n) \Big), \qquad (27)$$

for $t > L$ and $\ell - d_{max} \leq i \leq u$,

$$P(NI(t) = i, A(t - L) = n) = \\ \sum_{y=\ell}^{u} \Big(P(IP(t - L) = y, A(t - L) = n) \\ \times P(D_{t-L}(L) = y - i | A(t - L) = n) \Big). \qquad (28)$$

Accordingly, we establish the following equations of the net inventory probability distribution,

for $0 \leq t \leq L$ and $\ell - d_{max} \leq i \leq u$,

$$P(NI(t) = i) = \sum_{n=1}^{m} P(NI(t) = i, A(0) = n) \qquad (29)$$

for $t > L$ and $\ell - d_{max} \leq i \leq u$,

$$P(NI(t) = i) = \sum_{n=1}^{m} P(NI(t) = i, A(t - L) = n) \qquad (30)$$

Finally, the rth moment of the net inventory position at time $t \geq 0$ can be found by

$$E[NI^r(t)] = \sum_{i=\ell-d_{max}}^{u} i^r \, P(NI(t) = i) \qquad (31)$$

D. On-hand Inventory

The on-hand inventory $I(t)$ is the inventory on the shelf that is available for sale or use at a particular time t. Now, we present the equations which allow us to find the distribution of the on-hand inventory at time $t \geq 0$. Based on equations (21) and (22), we can write
for $0 \leq t \leq L$

$$I(t) = NI(t)^+ = \Big(IP(0) - D_0(t) \Big)^+, \qquad (32)$$

for $t > L$,

$$I(t) = NI(t)^+ = \Big(IP(t - L) - D_{t-L}(L) \Big)^+. \qquad (33)$$

Hence, for $t \geq 0$, $i = 1, \ldots, u$,

$$P(I(t) = i) = P(NI(t) = i) \qquad (34)$$

and for $t \geq 0$, $i = 0$

$$P(I(t) = 0) = \sum_{j=\ell-d_{max}}^{0} P(NI(t) = j) \qquad (35)$$

Therefore, the rth moment of the on-hand inventory at time $t \geq 0$ can be found by

$$E[I^r(t)] = \sum_{i=0}^{u} i^r \, P(I(t) = i) \qquad (36)$$

E. Backorders

The backorders also called backlogs, denoted $B(t)$, is the inventory on the shelf that is already committed due to unmet demands from previous periods. In this section, we present the equations which allow us to find the distribution of the backorders at time $t \geq 0$. Based on equations (21) and (22), we can write

$$B(t) = NI(t)^- = (IP(0) - D_0(t))^- = \\ (D_0(t) - IP(0))^+ \quad \text{for } 0 \leq t \leq L, \qquad (37)$$

$$B(t) = NI(t)^- = (IP(t-L) - D_{t-L}(L))^- \\ = (D_{t-L}(L) - IP(t-L))^+ \quad \text{for } t > L. \qquad (38)$$

Hence, for $t \geq 0$, $i = 1, \ldots, (d_{max} - \ell)$,

$$P(B(t) = i) = P(NI(t) = -i) \qquad (39)$$

and for $t \geq 0$, $i = 0$,

$$P(B(t) = i) = \sum_{j=0}^{u} P(NI(t) = j) \qquad (40)$$

Therefore, the rth moment of the backorders at time $t \geq 0$ can be found by

$$E[B^r(t)] = \sum_{i=0}^{\ell-d_{max}} (i)^r \, P(B(t) = i) \qquad (41)$$

F. Number of Orders

In this section, we find the expected number of orders placed in an interval of length $[0, t]$. Let $R(t)$ be the number of orders placed up-to time $t \geq 0$. Let $\Gamma_{r,i,n}(t) = P(R(t) = r, IP(t) = i, A(t) = n)$ be the probability of having $R(t) = r$, $IP(t) = i$ and $A(t) = n$ at time $t \geq 0$ for $r \geq 0$, $i = \ell, \ldots, u$ and $n = 1, \ldots, m$. Let ω be the fixed ordering cost. We present the below KFEs where the derivative is with respect to t.

1) The General Ph_t Number of Orders KFEs: For $t \geq 0$, $n = 1, \ldots, m$, $s(t) < i \leq u$, and $r > 0$,

$$\Gamma'_{r,i,n}(t) = \sum_{w=1}^{m} \lambda_{w,n}(t) \Gamma_{r,i,w}(t) \\ + \alpha_n(t) \Big(\sum_{w=1}^{m} \mu_w(t) \, \Gamma_{r,i+1,w}(t) \Big) (1 - I_{i=u}) \qquad (42) \\ + \alpha_n(t) \Big(\sum_{w=1}^{m} \sum_{q=\ell}^{s(t)+1} \mu_n(t) \, \Gamma_{r-1,q,n}(t) \Big) I_{i=S(t)} I_{(r>0)}$$

and for $n = 1, \ldots, m$, $\ell \leq i \leq s(t)$ and $r \geq 0$,

$$\Gamma'_{r,i,n}(t) = \sum_{w=1}^{m} \lambda_{w,n}(t) \Gamma_{r,i,w}(t). \qquad (43)$$

To compute the probabilities $\Gamma_{r,i,n}$ we can numerically solve the differential equations in (42) and (43) for the general Ph_t distribution. But since $r = 0, 1, \ldots, \infty$ this approach requires setting an upper bound (r_{max}) on the number of orders such that $P(R(T) > r_{max})$ is sufficiently small. An alternate method for the computation of the expected number of orders is to derive a finite set of moment differential equations to calculate the moments. Next, we write the general Ph_t distributions differential equations for the first two moments of the reorder count process.

2) The General Ph_t Expected Number of Orders KFEs: For $n = 1, \ldots, m$, $s(t) < i \leq u$ and $r > 0$,

$$E'[R^k(t); i, n] = \sum_{r=0}^{\infty} r^k \, \Gamma'_{r,i,n}(t) \\ = \sum_{v=1}^{m} \lambda_{v,n}(t) \, E[R^k(t); i, v] \\ + \alpha_n(t) \Big(\sum_{v=1}^{m} \mu_v(t) \, E[R^k(t); i+1, v] \Big) I_{(i \neq u)} \qquad (44) \\ + \alpha_n(t) \Big(\sum_{v=1}^{m} \sum_{j=\ell}^{s(t)+1} \mu_v(t) \Big(\sum_{z=1}^{k} \binom{k}{z} E[R^z(t); j, v] \\ + P_{j,v}(t) \Big) \Big) I_{(i=S(t))}.$$

For $n = 1, \ldots, m$, $\ell \leq i \leq s(t)$ and $r \geq 0$,

$$E'[R^k(t); i, n] = \sum_{v=1}^{m} \lambda_{v,n}(t) \, E[R^k(t); i, v]. \qquad (45)$$

G. Inventory Total Cost

In this section, we find the expected total cost for a given reorder policy $(\bar{s}, \bar{S})$. The expected total cost includes the expected holding, back-ordering and number of orders costs.

$$\Phi(\bar{s}, \bar{S}) = h \, E[I(t)] + b \, E[B(t)] + \omega E[R(t)] \qquad (46)$$

The below algorithm summarizes the computation of the expected cost for a given reorder policy $(\bar{s}, \bar{S})$.

1) Total Expected Cost Algorithm: For $t \geq 0$; $t_s = t - L$; $d = 0, 1, \ldots, d_{max}$; $i = \ell, \ldots, u$ and $n = 1, \ldots, m$

Step 1. Follow the demand-count distribution algorithm in §III-A in order to find $P(D_0(t) = d | A(0) = n)$ and $P(D_{t_s}(L) = d | A(t_s) = n)$

Step 2. Follow the IP distribution algorithm in §III-B in order to find $P(IP(t) = i, A(t) = n)$.

Step 3. Use equations (27) and (28) to find $P(NI(t) = i, A(0) = n)$ and $P(NI(t) = i, A(t_s) = n)$ and then equations (29) and (30) to compute $P(NI(t) = i)$.

Step 4. Use eqs. (34) & (35) to compute $P(I(t) = i)$.

Set $E[I(t)] = \sum_{i=0}^{u} i \, P(I(t) = i)$.

Step 5. Use eqs. (39) & (40) to compute $P(B(t) = i)$.

Set $E[B(t)] = \displaystyle\sum_{i=0}^{d_{max}-\ell} i\, P(B(t) = i)$.

Step 6. Solve equations (44), (45) for $k = 1$ to find $\mathrm{E}[R(t); i, n]$

Set $\mathrm{E}[R(t)] = \displaystyle\sum_{i=\ell}^{u} \sum_{n=1}^{m} \mathrm{E}[R(t); i, n]$.

Step 7. Set $\Phi(\bar{s}, \bar{S}) = h\, E[I(t)] + b\, E[B(t)] + \omega\, E[R(t)]$.

IV. FUTURE WORK

In this section we list the next steps to be taken in our research.

1) Proposing a heuristic policy to optimize the expected cost. This will most probably be done via line search.

2) Providing a numerical example from real-world data for which we will fit the 2-MECO Ph_t distribution with the data's first and second moments. Using this numerical example, we will show and demonstrate by numbers and figures all the results of our analytical model.

ACKNOWLEDGMENT

REFERENCES

[1] Hari S Abhyankar and Stephen C Graves. Creating an inventory hedge for Markov-modulated Poisson demand: An application and model. *Manufacturing & Service Operations Management*, 3(4):306–320, 2001.

[2] Yossi Aviv and Awi Federgruen. Stochastic inventory models with limited production capacity and periodically varying parameters. *Probability in the Engineering and Informational Sciences*, 11(01):107–135, 1997.

[3] Dirk Beyer and Suresh P Sethi. Average cost optimality in inventory models with markovian demands. *Journal of Optimization Theory and Applications*, 92(3):497–526, 1997.

[4] Srinivas Bollapragada and Thomas E Morton. A simple heuristic for computing nonstationary (s, S) policies. *Operations Research*, 47(4):576–584, 1999.

[5] James H Bookbinder and Jin-Yan Tan. Strategies for the probabilistic lot-sizing problem with service-level constraints. *Management Science*, 34(9):1096–1108, 1988.

[6] Fangruo Chen and Jing-Sheng Song. Optimal policies for multiechelon inventory problems with Markov-modulated demand. *Operations Research*, 49(2):226–234, 2001.

[7] Feng Cheng and Suresh P Sethi. Optimality of state-dependent (s, S) policies in inventory models with Markov-modulated demand and lost sales. *Production and Operations Management*, 8(2):183–192, 1999.

[8] Markus Ettl, Gerald E Feigin, Grace Y Lin, and David D Yao. A supply network model with base-stock control and service requirements. *Operations Research*, 48(2):216–232, 2000.

[9] Chandandeep S Grewal, ST Enns, and Paul Rogers. Dynamic reorder point replenishment strategies for a capacitated supply chain with seasonal demand. *Computers & Industrial Engineering*, 80:97–110, 2015.

[10] Donald L Iglehart and Samuel Karlin. *Optimal policy for dynamic inventory process with nonstationary stochastic demands*, volume 8. Chapter, 1962.

[11] Roman Kapuściński and Sridhar Tayur. A capacitated production-inventory model with periodic demand. *Operations Research*, 46(6):899–911, 1998.

[12] S Karlin and A Fabens. The (s, S) inventory model under Markovian demand process. *Mathematical Methods in the Social Sciences*, pages 159–175, 1960.

[13] Samuel Karlin. Dynamic inventory policy with varying stochastic demands. *Management Science*, 6(3):231–258, 1960.

[14] Samuel Karlin and Herbert Scarf. Inventory models and related stochastic processes. *Studies in the mathematical theory of inventory and production*, (1):319, 1958.

[15] BS Maddah, MY Jaber, and NE Abboud. Periodic review (s, s) inventory model with permissible delay in payments. *Journal of the operational research society*, 55(2):147–159, 2004.

[16] Walid W Nasr and Bacel Maddah. Continuous (s, S) policy with MMPP Correlated Demand. *European Journal of Operational Research*, 2015.

[17] Walid W Nasr and Michael R Taaffe. Fitting the Pht/Mt/s/c Time-Dependent Departure Process for Use in Tandem Queueing Networks. *INFORMS Journal on Computing*, 25(4):758–773, 2012.

[18] HERBERT Scarf. The optimality of (s, S) policies in the dynamic inventory problem. Mathematical Methods in the Social Science, KJ Arrow, S. Karlin, P. Suppes, eds, 1960.

[19] Suresh P Sethi and Feng Cheng. Optimality of (s, S) policies in inventory models with Markovian demand. *Operations Research*, 45(6):931–939, 1997.

[20] Edward Silver. Inventory control under a probabilistic time-varying, demand pattern. *Aiie Transactions*, 10(4):371–379, 1978.

[21] Edward A Silver. Inventory management: An overview, Canadian publications, practical applications and suggestions for future research. *INFOR: Information Systems and Operational Research*, 46(1):15–28, 2008.

[22] Jing-Sheng Song and Paul Zipkin. Inventory control in a fluctuating demand environment. *Operations Research*, 41(2):351–370, 1993.

[23] S Armagan Tarim and Brian G Kingsman. Modelling and computing (Rn, Sn) policies for inventory systems with non-stationary stochastic demand. *European Journal of Operational Research*, 174(1):581–599, 2006.

[24] Huseyin Tunc, Onur A Kilic, S Armagan Tarim, and Burak Eksioglu. The cost of using stationary inventory policies when demand is non-stationary. *Omega*, 39(4):410–415, 2011.

[25] Paul Zipkin. Critical number policies for inventory models with periodic data. *Management Science*, 35(1):71–80, 1989.

Economic order quantity with deviation in the order received

Anthony Khoury-Hanna

Engineering Management Graduate Program, Faculty of Engineering and Architecture, American University of Beirut.

I. **Abstract**

The fundamental assumption of an economic order quantity (EOQ) model is that 100% of items in an ordered lot are received. This paper extends the traditional EOQ model by accounting for order deviation, characterized by a random fraction of items deviating from the ordered amount. This random deviation can lead to over-ordering (receiving more than what is ordered) or to under-ordering (the opposite situation).This has application in several industries where counting the amount of received items is not possible, immediately upon the order receipt. Orders of construction material are examples of this. Simple expressions of the optimal order quantity and the expected cost function per unit time are formulated. Sensitivity analysis shows that out optimal order quantity can be smaller or larger that of the classical EOQ depending on the variability of the order deviation.
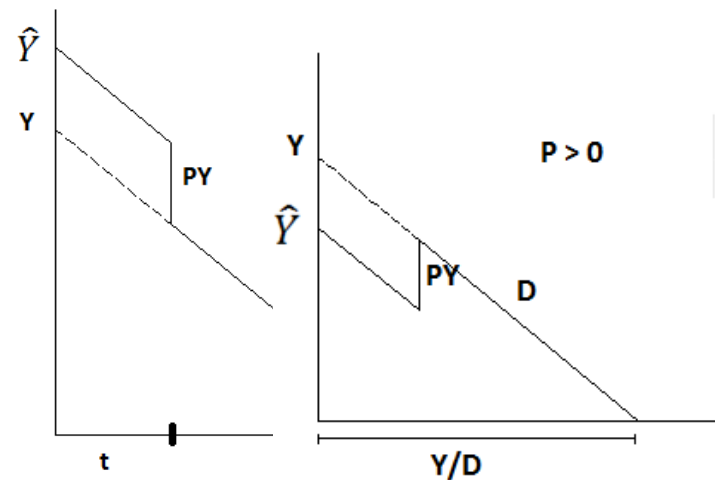
Figure 1 : Inventory level over time , P<0

Figure 2 : Inventory level over time , P>0

II. Introduction

Completing project on time and on cost designed is the target goal that all of project managers focuses on, no matter how large or small this project is. The aim to complete all the project tasks remains the same by using project management, but the question that should be asked is why almost all projects usually exceed their scheduled cost . Many reasons causes projects exceed their schedule like inaccurate quantities, relying just on software, not having a proper description for the project requirements, and one important reason not accounting holding cost of random yield order deviation. The new model leads to simple expressions for the expected profit per unit time and the optimal order quantity. Consider the case where a lot of size y is delivered instantaneously with a purchasing price of c per unit and an ordering cost of K.

It is assumed that each lot received contains percentage deviation, P, with a known probability density function, f (P). A 100% percent counting process of the lot is conducted at a rate of x units per unit time; items deviated are sent back or ordered back on time expected.The behavior of the inventory level is illustrated in Fig. 1-2 , PY is the number of deviated items from inventory and t is the total counting time of $\hat{Y}$ units received per cycle. d is the counting cost per unit and h the holding cost per unit per unit time. Items received are not the same quantity ordered and considers Shortages are no allowed, items are returned or backordered as a single batch by the end of the t period counting process.

Y : order size

$\hat{Y}$: *quantity received*

$\hat{Y} = Y(1 - P)$

D: Demand rate

C : unit variable cost

K : fixed cost of placing order

P : % of order deviation

x : counting validation rate(units/unit time)

* d : unit counting cost*

T : cycle length $= \dfrac{Y}{D}$

* t : counting/validation time* $= \dfrac{\hat{Y}}{x}$

P is a randon variable with pdf f(P) ;

* a<0 , b>0 , $a \leq P \leq b$*

TC = ordering cost + counting / validation cost + Holding Cost

III. Mathematical results

$$TC = K + CY + d\hat{Y} + h(\frac{YT}{2} - PYt)$$

$$ETCU = K\frac{D}{Y} + CD + dDE(1-P) + h(\frac{Y}{2} - \frac{DYE(P(1-P))}{x})$$

$$\frac{\partial C_u(y)}{\partial Y} = \frac{-KD}{Y^2} + 0.5h - E(P(1-P))\frac{Dh}{x} = 0$$

$$Y^* = \sqrt{\frac{2KD}{(h(1-2E(P(1-P))\frac{D}{x})}} \quad ;$$

$$h' = h(1 - 2E(P(1-P))\frac{D}{x})$$

$$ETCU = KD\sqrt{\frac{h'}{2KD}} + CD + dDE(1-P) +$$

$$\frac{h}{2}\sqrt{\frac{2KD}{h'}}\left(1 - 2E(P(1-2P)) \times \frac{D}{x}\right)$$

$$ETCU = CD + dDE(1-P) + \sqrt{2KDh'}$$

IV. Numerical results

$$Y^* = \sqrt{\frac{2KD}{h'}}$$

$$ETCU = CD + dDE(1-P) + \sqrt{2KDh'}$$

P(a ; b)   : a<0 , b>0

P is a random variable with pdf f(P)

P ~ U (-0.1 , 0.1)

$$F(P) = \begin{cases} 5; & -0.1 < P < 0.1 \\ 0 & OTHERWISE \end{cases}$$

- $E(P(1-P)) = 5 \times \left[\frac{P^2}{2} - \frac{P^3}{3}\right]_{-0.1}^{0.1} = -0.00333$

- $E(1-P) = 5 \times \left[p - \frac{p^2}{2}\right]_{-0.1}^{0.1} = 1$

$$Y^* = \sqrt{\frac{2KD}{h(1.00666\frac{D}{x})}} \qquad ;h' = h(1.00666)\frac{D}{x}$$

$$ETCU = CD + dDE(1-P) + \sqrt{2KDh'}$$

Note that to avoid shortages the condition must be satisfied, that is

$$P < 1 - D\frac{Y}{x} \rightarrow P < 1 - 55000 * \frac{1}{185000}$$
$$\rightarrow \quad P < 0.7$$

V. Economic order quantity with divided holding cost

In fact , paying Y to get $\hat{Y}$ must be profitable or loss so $\hat{Y} - Y$ returned or reordered items are not covered by the total holding cost with same assumptions as previous case. Over/under shooting items hold part of the total holding cost . Can be applied for construction sites .

$$H = h_c + h_s = 0.8h + 0.2h$$

$$ETCU = K\frac{D}{Y} + CD + \int_a^0 dD(1-p)\,f_p d_p +$$
$$\int_0^b dD(1-p)\,f_p d_p + \int_a^0 \left(\frac{hY}{2} - \frac{h_s\,p(1-p)DY}{x}\right)f_p d_p +$$
$$\int_0^b \left(\frac{hY}{2} - \frac{h_s\,p(1-p)DY}{x}\right)f_p d_p$$

$$ETCU = K\frac{D}{Y} + CD + \int_a^b dD(1-p)\,f_p d_p +$$
$$\int_a^b \left(\frac{hY}{2} - \frac{h_s\ p(1-p)DY}{x}\right) f_p d_p$$

$$ETCU = K\frac{D}{Y} + CD + dDE(1-P) + h\frac{Y}{2} -$$
$$\frac{h_s\ DYE(P(1-P))}{x}$$

$$\frac{\partial C_u(y)}{\partial Y} = \frac{-KD}{Y^2} + 0.5h - E(P(1-P))\frac{Dh_s}{x}$$
$$Y^* = \sqrt{\frac{2KD}{\left(h - 2E(P(1-P))\frac{h_s\ D}{x}\right)}} \qquad ;$$

$$h' = h - 2E\big(P(1-P)\big)\frac{h_s\ D}{x} = h(1 - 2*0.2*$$
$$E\big(P(1-P)\big)\frac{D}{x})$$

$$ETCU = KD\sqrt{\frac{h'}{2KD}} + CD + dDE(1-P) +$$
$$\frac{h}{2}\sqrt{\frac{2KD}{h'}}\left(1 - 0.2*2*E\big(P(1-2P)\big) \times \frac{D}{x}\right)$$
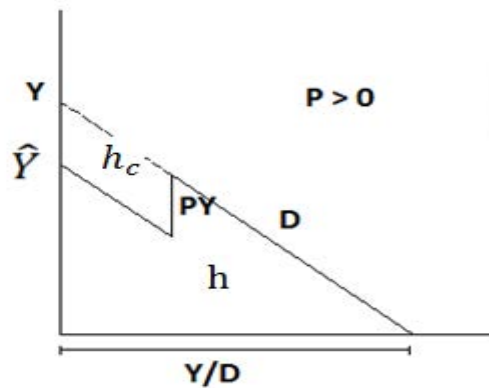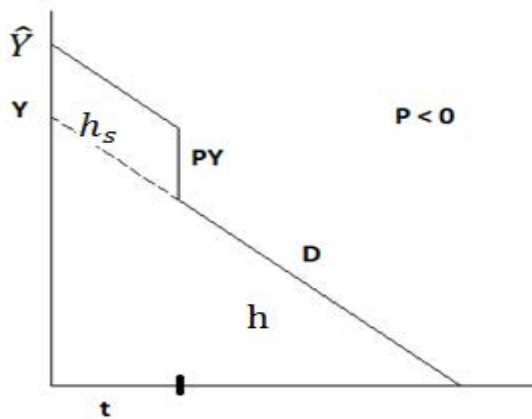
$$ETCU = CD + dDE(1-P) + \sqrt{2KDh'}$$

VI. Numerical results

$$Y^* = \sqrt{\frac{2KD}{h'}} \quad ;$$

$$ETCU = CD + dDE(1 - P) + \sqrt{2KDh'}$$

$P(a\ ;\ b)\quad :\ a{<}0\ ,\ b{>}0$

$P \sim U\ (\ -0.1\ ,\ 0.1\ )$

$$F(P) = \begin{cases} 5\ ; & -0.1 < P < 0.1 \\ 0 & OTHERWISE \end{cases}$$

- $E(P(1 - P)) = 5 \times \left[\frac{P^2}{2} - \frac{P^3}{3}\right]_{-0.1}^{0.1} = -0.00333$

- $E(1 - P) = 5 \times \left[p - \frac{p^2}{2}\right]_{-0.1}^{0.1} = 1$

$$h' = h(1 - 2 * 0.2 * E(P(1 - P))\frac{D}{x})$$

$$Y^* = \sqrt{\frac{2KD}{(h - 2E(P(1-P))\frac{h_s\ D}{x})}} \quad ;$$

$$h' = h - 2E(P(1 - P))\frac{h_s\ D}{x} = h(1 - 2 * 0.2 * E(P(1 - P))\frac{D}{x})$$

$$ETCU = KD\sqrt{\frac{h'}{2KD}} + CD + dDE(1 - P) + \frac{h}{2}\sqrt{\frac{2KD}{h}}\left(1 - 0.2 * 2 * E(P(1 - 2P)) \times \frac{D}{x}\right)$$

$$ETCU = CD + dDE(1 - P) + \sqrt{2KDh'}$$

$$P < 1 - D\frac{Y}{x} \; \rightarrow \; P < 1 - 55000 * \frac{1}{185000}$$
$$\rightarrow \quad P < 0.7$$

VII. Sensitivity analysis:

This analysis shows variable intervals of (a ; b) with optimal order quantity and optimal order cost can be smaller or larger that of the classical EOQ .

1. Expected Optimal ordering analysis :

$$\sqrt{\frac{2KD}{h - 2E(P(1-P))\frac{h_s\ D}{x}}} < \sqrt{\frac{2KD}{h}}$$

$$\frac{2KD}{h - 2E(P(1-P))\frac{h_s\ D}{x}} < \frac{2KD}{h}$$

$$h - 2E(P(1 - P))\frac{h_s\ D}{x} > h$$

$$E(P(1 - P)) < 0$$

$$b^2\left(\frac{1}{2} - \frac{b}{3}\right) < a^2\left(\frac{1}{2} - \frac{a}{3}\right)$$

$$b < -a \ and \ b > a$$

2. Expected Total optimal cost :

$$CD + dDE(1 - P) + \sqrt{2KDh'} \;<\; CD + dD + \sqrt{2KDh}$$

- $E(1 - P) < 1$

$$\frac{1}{b-a}\left(b - \frac{b^2}{2} - a + \frac{a^2}{2}\right) < 1$$

$$\left(\frac{(a-b)(a+b)}{2(b-a)}\right) < 0$$

$$a + b > 0 \qquad\qquad b > -a$$

- $\sqrt{2KDh'} \;<\; \sqrt{2KDh}$

$$h' \;<\; h$$

$$\left(1 - 2 * 0.2 * E\big(P(1 - P)\big)\frac{D}{x}\right) < 1$$

$$E\big(P(1 - P)\big) > 0$$

$$b^2\left(\frac{1}{2} - \frac{b}{3}\right) > a^2\left(\frac{1}{2} - \frac{a}{3}\right)$$

$$b > -a \text{ and } b > a$$

3. Variability of the order deviation

- *"$b \le -a$ ": ordered less than EOQ , receiving more than ordered , optimal total cost of ordering quantity higher than optimal EOQ total cost .*

- *"$b > -a$" : ordered more than EOQ , receiving less than ordered , optimal total cost of ordering quantity less than EOQ total cost.*

- *" b = -a " :* $\quad \dfrac{1}{b-a}\left(\dfrac{b^2}{2} - \dfrac{b^3}{3} - \dfrac{a^2}{2} + \dfrac{a^3}{3}\right) = \dfrac{-b^2}{3} < 0$

$$\rightarrow \quad E\big(P(1 - P)\big) < 0$$

$$\rightarrow \quad \sqrt{\frac{2KD}{h - 2E(P(1-P))\frac{h_s\,D}{x}}} \;<\; \sqrt{\frac{2KD}{h}}$$

References :

Salameh, M.K., Jaber, M.Y., 2000. Economic production quantity model for items with imperfect quality. International Journal of Production Economics 64 (1–3), 59–64.

VIII.   Summary and Conclusion

In this paper , while analyzing the random yield factor , we can see the increase and decrease of total cost and ordered quantity compared to the EOQ model . This is accurate expressions to facilitate project managers and markets work when ordering.

Electromyography Analysis of Maximum Voluntary Contraction Techniques for Trunk, Shoulder, and Neck Muscles

Alif Saba
Industrial Engineering and Management
American University of Beirut
P.O Box: 11-0236, Riad El Solh
Ams98@mail.aub.edu

Abstract- **The objective of this research is to compare literature-based MVC techniques for six muscles in the neck, shoulder, and low back regions in healthy subjects. The six muscles chosen were: the thoracic erector spinae, the lumbar erector spinae, the latissimus dorsi, the posterior deltoid, the upper trapezius, and the sternocleidomastoid. These muscles were chosen based on the amount of attention they receive in ergonomic research. EMG activities were measured while 15 healthy participants performed specific MVC techniques for each muscle chosen.**

The results indicated that the lumbar and thoracic subdivisions of the erector spinae muscle can be maximally activated by four similar MVC technique. Furthermore, it was recommended to use the "Prone Extension" test or the "Chest Supported ROW" test to normalize EMG signals from the latissimus dorsi. The "Shoulder Abduction in Slight Extension" test or the "Transvers Abduction" test were recommended as the MVC technique for the posterior deltoid. The results showed that the levels of EMG signal generated in the upper trapezius by the "Abduction 125" test, the" Elevation and Abduction 90" test, and the "Abduction 90" test were not significantly different. Finally, the "Anterolateral Flexion" was found to be the optimal MVC technique for the sternocleidomastoid muscle.

I. INTRODUCTION

Work-Related Musculoskeletal Disorders (WMSD's) are the costliest work-related diseases [1], with their rates representing 34% of all work-related injuries and illnesses in the US [2], 36% of all lost time claims, and 41% of all costs between 1994 and 2002 [3]. These rates are even more alarming in the UK with musculoskeletal disorders accounting for 42% of all work-related diseases [4].

Electromyography signals (EMG) can be a helpful tool in ergonomic research as it can provide ergonomists with an understanding of muscles' function and reactions. However, these signals can be misleading when used without normalization. The need to normalize EMG signals comes from a number of physiological and technical reasons.

Therefore, EMG normalization techniques are a major focus point in ergonomic literature.

Normalizing EMG signals by using signals from maximum voluntary contractions (MVC) remains the most appropriate, reliable, and common technique [5]. The major problem with this technique remains the lack of consensus on the appropriate technique to elicit the maximum voluntary contraction from a certain muscle [6].

Therefore, the literature presents multiple recommendations for eliciting the MVC of any certain muscle. Nonetheless, as far as known, no study has been able to describe or compare all the tests able to maximally activate any of the eight muscles investigated in this paper.

Statistics show that the low back and upper limbs are the most affected body parts by WMSD's. Therefore, the choice of the six muscles investigated in this paper was restricted to muscles from the low back, shoulder, and neck since these areas are highly affected by WMSD's and have been the primary focus of ergonomic research.

The objective of this paper is to compare literature-based MVC techniques for six muscles in the neck, shoulder and low back regions. The muscles chosen were: the lumbar erector spinae, the thoracic erector spinae, the latissimus dorsi the posterior deltoid, the sternocleidomastoid, and the upper trapezius.

II. METHODOLOGY

A. Selection of Muscles

As explained before, the selection of muscles to be investigated in this study was restricted to muscles from the low back, shoulder, and neck areas. Since low-back musculoskeletal disorders constitute 36% of all work related injuries [2] three muscles were chosen from the low-back area whereas two muscles were chosen from the shoulder area and one muscle from the neck area.

The selection of the muscles was dependent on the amount of attention each muscle receives in ergonomic research. Ergonomic studies using sEMG to investigate WMSD's in the three pre-

selected body regions were reviewed and the muscles they examined were listed. Eventually, the muscles with the most citations were selected for further investigation. Thus, the final list of muscles chosen is as follows:

1- Lumbar Erector Spinae (Low back)
2- Thoracic Erector Spinae (Low back)
3- Latissimus Dorsi (Low back)
4- Posterior Deltoid (Shoulder)
5- Upper Trapezius (Shoulder)
6- Sternocleidomastoid (Neck)

Figure 1 illustrates the location of the aforementioned six muscles.
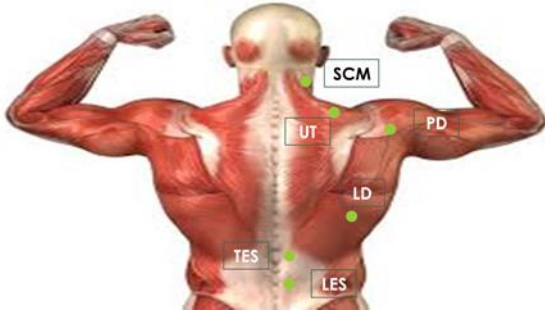
Figure 1 Electrode locations for the six muscles

B. Participants and Apparatus

For time purposes, the experiment was divided into two sections. Each section contained three different muscles. The First section contained the Upper Trapezius (UT), Sternocleidomastoid (SCM), and the Posterior Deltoid (PD). The second section contained the Lumbar Erector Spinae (LES), the Thoracic Erector Spinae (TES), and the Latissimus Dorsi (LD).

Thirty healthy male subjects, fifteen for each section, with no history of back pain volunteered to perform the MVC techniques for each of the two sections. The Physical Activity Readiness Questionnaire (PAR-Q, British Columbia Ministry of Health) was used to screen participants for cardiac and other health problems, such as dizziness, chest pain, or heart trouble [7].

The Tringo wireless EMG system was used to record electrical activity from the chosen muscles. Six rectangular (37mm x 26mm x 15mm, 14g) Ag/AgCL sensors were used. The Tringo system does not require the use of a reference electrode.

C. Experimental Task

A thorough literature review was conducted to identify the most commonly used MVC techniques for each of the six muscles. Four to five MVC techniques that were not directly compared to each other in previous studies were selected. The selected MVC techniques were:

A. Lumbar/Thoracic Erector Spinae
   a. Trunk Holding with a starting position of 60°
   b. Resisted Upper Trunk Extension in a prone position
   c. Resisted Upper Trunk Extension in a standing position
   d. Resisted Lower Trunk Extension
   e. The Arch Test
B. Latissimus Dorsi
   a. Chest Supported Row Contraction
   b. Internal Rotation 90°
   c. Internal Rotation 0°
   d. Lat Pull Down
   e. Prone Extension
C. Posterior Deltoid
   a. Prone Elevation
   b. Abduction 0°
   c. Empty Can
   d. Transverse Abduction
   e. Shoulder abduction in slight extension
D. Upper Trapezius
   a. Abduction 90° or Empty Can
   b. Abduction 125°
   c. Flexion 125°
   d. Shoulder Elevation
   e. Shoulder Elevation and Abduction.
E. Sternocleidomastoid
   a. Resisted neck flexion
   b. Resisted neck lateral bending
   c. Resisted neck rotation
   d. Resisted neck anterolateral flexion

Since it is not possible to explain all the techniques, one technique, resisted neck anterolateral flexion, would be used as an example to clarify the experimental task. This technique is performed in a supine position with the arms abducted at 90° and externally rotated, and the elbow flexed at 90°. The head is rotated to the opposite side and the participants will try to raise their head against resistance applied against the temporal region of the head in an obliquely posterior direction.

D. EMG Preparations

Subsequent to the warm-up session, preparations were made to ready the participants for EMG data acquisition. Any hair on the skin at the electrodes sites was removed. The same areas were cleaned with alcohol. The purpose of cleaning the skin is to get rid of dead skin cells, dirt, and sweat. Then the EMG surface electrodes were attached to the muscles of interest at the following positions:

• Upper Trapezius: electrode was placed 2 cm lateral to the midpoint of the lead line between the spinous process of C7 and the posterolateral border of the acromion [8].

- Posterior Deltoid: In the prone position with the head turned to the right side, the shoulder abducted to 90°, the elbow flexed to 90°, and the thumb pointing upwards, the electrode was placed 2 cm below the lateral border of the scapular spine in an oblique angle towards the arm (i.e. parallel to muscle fibers) [9].
- Lumbar Erector Spinae: electrode was placed 3 cm lateral to the L3 spinous process [10].
- Thoracic Erector Spinae: electrode was placed approximately 5 cm lateral to the T9 spinous process [10].
- Latissimus Dorsi: electrode was placed approximately 4 cm below the inferior tip of the scapula and midway between the spine and lateral edge of the torso [9].
- Sternocleidomastoid: electrode was placed at the lower 1/3rd of the line connecting the sternal notch and the mastoid process [9].

E. Statistical Analysis

For each muscle, the Kruskal-Wallis test was used to assess the effect of the MVC techniques on EMG activity. For all significant effects, post hoc analyses, in the form of the Mann-Whitney test was performed to determine the source(s) of the significant effect(s). The null and alternate hypotheses for the Mann-Whitney test in all muscles were as follows:

H0: Median "Test X" = Median "Test Y"
H1: Median "Test X" $\neq$ Median "Test Y"

The significance level (α) was set at 5%. Statistical significance was based on calculated p-values.

III. RESULTS AND STATISTICAL ANALYSIS

The average age, height, and weight of the participants are provided in Table 7.

TABLE 1
DEMOGRAPHIC DATA OF PARTICIPANTS

Muscles	Age (years)	Weight (kg)	Height (cm)
PD SCM UT	20.7	72.13	176.06
LES TES LD	20.79	75.57	179.36

The P-value of all Kruskal-Wallis tests was (P=0). Therefore, a post-hoc analysis, in the form of a series of Mann-Whitney tests, was performed in all muscles to identify the source(s) of the statistical differences.

A. Lumbar Erector Spinae

The results showed four cases of statistical differences between tests. The medians of the four statistically different pairs of tests were compared to identify the appropriate MVC techniques.

The EMG signal generated by the "Trunk Holding with a starting position of 60°" test was statistically lower than those generated by all other tests. However, there was not enough evidence to suggest any statistical differences between the EMG signals generated by the four other tests. The "Arch" test produced the lowest value of standard deviation.

B. Thoracic Erector Spinae

The results showed four cases of statistical differences between tests. The medians of the four statistically different pairs of tests were compared to identify the appropriate MVC techniques.

The EMG signal generated by the "Trunk Holding with a starting position of 60°" test was statistically lower than those generated by all other tests. However, there was not enough evidence to suggest any statistical differences between the EMG signals generated by the four other tests. The "Arch" test produced the lowest value of standard deviation.

C. Latissimus Dorsi

The results showed eight cases of statistical differences between tests. The medians of the eight statistically different pairs of tests were compared to identify the appropriate MVC techniques.

The EMG signal generated by the "Chest Supported ROW" test had the highest numerical median. However, there was not enough evidence to suggest any statistical difference between the EMG signals generated by the "Prone Extension" test and the "Chest Supported Row" test. The EMG signals generated by the aforementioned two tests were statistically higher than all other tests. The "Internal Rotation 90" test had the lowest numerical median and had an EMG signal that is statistically lower than those generated by the "Internal Rotation 0" and the "Lat Pull Down" tests.

D. Posterior Deltoid

The results showed seven cases of statistical differences between tests. The medians of the seven statistically different pairs of tests were compared to identify the appropriate MVC techniques.

The results indicate that the "Transvers Abduction" test had the highest numerical median. However, there was not enough evidence to suggest a statistical difference between the "Transvers Abduction" test and the "Shoulder Abduction in Slight Extension" test. The "Empty Can" test had the lowest

numerical but was not statistically different than the "Abduction 0" test. The EMG signal generated by the "Prone Elevation" was statistically higher than those generated by the "Empty Can" and "Abduction 0" tests.

E. Upper Trapezius

The results showed five cases of statistical differences between tests. The medians of the five statistically different pairs of tests were compared to identify the appropriate MVC techniques.

The results indicate that the "Abduction 125" test had the highest numerical median. However, there was not enough evidence to suggest statistical differences between the "Abduction 125" test and the "Abduction 90" and the "Elevation and Abduction 90" tests. EMG signal generated by the "Elevation" test is lower than that generated by all other tests. The "Elevation" test had the lowest numerical median and was statistically lower than all other tests.

F. Sternocleidomastoid

The results showed five cases of statistical differences between tests. The medians of the five statistically different pairs of tests were compared to identify the appropriate MVC techniques.

The EMG signal generated by the "Anterolateral Flexion" test was statistically higher than that generated by all other tests. While the "Lateral Bending" test had an EMG signal statistically lower than all other tests.

IV. DISCUSSION AND CONCLUSION

A. Lumbar and Thoracic Erector Spinae

The literature regarding the MVC technique for lumbar and thoracic erector spinae did not present any evidence to suggest that both subdivisions of the erector spinae are maximally activated by the same technique. However, the results of this study indicate that both sections of the erector spinae muscle are maximally activated by the same set of tests. The arch test generated the highest numerical median of EMG signal in both lumbar and thoracic erector spinae. However, there was not enough evidence to suggest a statistical difference between the arch test and the "Trunk Extension Prone" test, the "Trunk Extension Standing" test, and the "Lower Trunk Extension" test. These results, contradict those of Plamondon et al. (1999) wich indicated that the trunk holding exercise with a starting position of 60° provides a more adequate MVC technique.

Additionally, the arch test showed the least value of standard deviation in both subdivisions of the erector spinae. In the lumbar erector spinae, the "Trunk Extension Standing" test showed the second least value of standard deviation. While the "Lower Trunk Extension" test showed the second least standard deviation in the thoracic erector spinae.

Therefore, either one of the following tests can be used as a MVC technique for both the lumbar and thoracic erector spinae:
- The Arch Test
- Trunk Extension Standing
- Lower Trunk Extension
- Trunk Extension Prone

B. Latissimus Dorsi

The results of this research show that the "Chest Supported ROW" test and the "Prone Extension" test generated levels of EMG signals that were not statistically different from each other. The "Prone Extension" test showed a lesser value of standard deviation than the "Chest Supported ROW" test. However, the EMG signals generated from the aforementioned two tests were statistically higher than those generated by the "Internal Rotation 90" test, "Internal Rotation 0" test, and the "Lat Pull Down" test. The "Lat Pull Down" test showed no statistical difference from the "Internal Rotation 0" test.

In conclusion, this research recommends using the "Prone Extension" test or the "Chest Supported ROW" test for normalizing EMG signals from the Latissimus Dorsi.

C. Posterior Deltoid

The literature presented five possible MVC techniques for the posterior deltoid. Three of these five techniques were drawn from a electromyographic study (Boettcher et al., 2008), while the other two were muscle strength testing techniques (Kendall et al., 1993; Perroto et al., 1980). None of the three techniques recommended by Boettcher et al. (2008) proved to be an appropriate MVC technique. This obviously indicates that muscle strength testing techniques, which are not properly investigated in electromyography studies, are possible appropriate normalization techniques.

The "Transvers Abduction" test did not show any statistical difference from the "Shoulder Abduction in Slight Extension" test but had a higher value of standard deviation. This indicates that the "Shoulder Abduction in Slight Extension" test is a more preferable MVC technique since it generates more consistent EMG signals between trials.

In conclusion, this research recommends using the "Shoulder Abduction in Slight Extension" test or the "Transvers Abduction" test for normalizing EMG signals from the posterior deltoid.

D. Upper Trapezius

Arm abduction, arm flexion, and shoulder elevation techniques are the most dominant MVC techniques in the literature. The results indicated that the elevation test generates statistically lower EMG signals than all other tests which contradict the findings of Jensen et al. (1995) who reported that arm flexion techniques were not as effective in activating the upper trapezius

as techniques with arm abduction or shoulder elevation. Therefore, it is clearly inefficient to use elevation based techniques to normalize EMG signals from the upper trapezius. Furthermore, there was not enough evidence to suggest a significant difference between the two abduction tests involved in the study (Abduction 90 and Abduction 125) and the "Flexion 125" test.

However, there was a significant difference when comparing the "Flexion 125" with the "Elevation and Abduction 90" test, which is a "combination test" devised by the authors of this study and not drawn from the literature. Thus, there is evidence to suggest that techniques that are based on a combination of the main movements of the upper trapezius might be more efficient MVC techniques than those currently used in the literature.

There was no statistical difference between the "Elevation and Abduction 90" test, the "Abduction 125" test, and the "Abduction 90" test. The "Abduction 90" test had the least value of standard deviation, followed by the "Abduction 125" test and the "Elevation and Abduction 90" respectively. This means that the "Abduction 90" test generates more consistent MVCs between trials.

Therefore, either one of the following tests can be used to normalize EMG signals from the upper trapezius:
- Abduction 90
- Abduction 125
- Elevation and Abduction 90

The "Abduction 125" test and the "Elevation and Abduction 90" test were not previously described in the literature but are developed by the research team. This shows that previous ergonomic research has not covered all possible MVC techniques and indicates the need to further explore such techniques in future research.

E. Sternocleidomastoid

Electromyography studies have usually used neck flexion, neck bending, or neck rotation to normalize EMG signals from the sternocleidomastoid. However, the results clearly indicate that the EMG signal generated by the "Anterolateral Flexion" test is statistically higher than that generated by all those tests. However, the "Anterolateral Flexion" test had also the highest value for standard deviation. The "Anterolateral Flexion" test is a test used for muscle strength testing purposes (Kendall et al, 1993) and is rarely used in electromyographic studies as a MVC technique.

In conclusion, this research recommends using the "Anterolateral Flexion" test for normalizing EMG signals from the sternocleidomastoid, since it was numerically the largest.

Table 2 summarizes the recommendations for the six muscles.

Muscle	Recommended MVC technique(s)
Lumbar Erector Spinae	• The Arch Test • Trunk Extension Standing • Trunk Extension Prone • Lower Trunk Extension
Thoracic Erector Spinae	• The Arch Test • Lower Trunk Extension • Trunk Extension Standing • Trunk Extension Prone
Latissimus Dorsi	• Prone Extension • Chest Supported ROW
Posterior Deltoid	• Shoulder Abduction in Slight Extension • Transvers Abduction
Upper Trapezius	• Abduction 90 • Abduction 125 • Elevation and Abduction 90
Sternocleidomastoid	• Anterolateral Flexion

RESEARCH LIMITATIONS AND FUTURE WORK

Participant for this study were recruited through flyers distributed in the campus of the American University of Beirut. Therefore, the population of the study is mainly young male students between the ages of 18 and 25. Thus, caution must be taken before generalizing the results of this study.

In three out of the six muscles investigated, the posterior deltoid, the upper trapezius, and the sternocleidomastoid, the results showed that the appropriate MVC technique was a muscle strength testing technique. Therefore, future research should look into muscle strength testing techniques as possible MVC techniques for data normalization purposes.

Additionally, the results of this study indicated that a test combining shoulder arm abduction and shoulder elevation generated an EMG signal in the upper trapezius that showed no statistical difference to those generated by the appropriate MVC technique, the "Abduction 125". Thus, further research may test EMG signals generated in the upper trapezius during tests that have a combination of movements.

REFERENCES

[1] Punnett, L., & Wegman, D. H. (2004). Work-related musculoskeletal disorders: the epidemiologic evidence and the debate. Journal of electromyography and kinesiology, 14(1), 13-23.

[2] BLS. (2012). NONFATAL OCCUPATIONAL INJURIES AND ILLNESSES REQUIRING DAYS AWAY FROMWORK, 2012.

[3] Silverstein, B., Adams, D., & Kalat, J. (2007). Work-related musculoskeletal disorders of the neck, back, and upper extremity in Washington State.

[4] Health and Safety Executive. (2014). Musculoskeletal Disorders in Great Britain.

[5] Burden, A. (2010). How should we normalize electromyograms obtained from healthy participants? What we have learned from over 25years of research. Journal of electromyography and kinesiology, 20(6), 1023-1035.

[6] Chopp, J. N., Fischer, S. L., & Dickerson, C. R. (2010b). On the feasibility of obtaining multiple muscular maximal voluntary excitation levels from test exertions: A shoulder example. Journal of electromyography and kinesiology, 20(5), 896-902.

[7] Hoeger, W. W., Turner, L. W., & Hafen, B. Q. (2001). Wellness: Guidelines for a healthy lifestyle: Brooks/Cole Publishing Company.

[8] McLean, L., Chislett, M., Keith, M., Murphy, M., & Walton, P. (2003). The effect of head position, electrode site, movement and smoothing window in the determination of a reliable maximum voluntary activation of the upper trapezius muscle. Journal of electromyography and kinesiology, 13(2), 169-180.

[9] Cram, J., Kasman, G., & Holtz, J. (1998). Introduction to surface electromyography. Gaithersburg, Marland: Aspen publishers Inc.

[10] McGill, S. M. (1991). Electromyographic activity of the abdominal and low back musculature during the generation of isometric and dynamic axial trunk torque: implications for lumbar mechanics. Journal of orthopaedic research, 9(1), 91-103.

[7] Hoeger, W. W., Turner, L. W., & Hafen, B. Q. (2001). Wellness: Guidelines for a healthy lifestyle: Brooks/Cole Publishing Company.

Generation of Random Variates for PHDs, MAPs and BMAPs

Farah Ghizzawi

Department of Industrial Engineering and Management
American University of Beirut
Bliss Street
P.O. Box 11-0236
Beirut, Lebanon
fng02@mail.aub.edu

Abstract- *We propose the simulation of three arrival processes, the Phase-Type Process (PH), the Markovian Arrival Process (MAP) and the Batch Markovian Arrival Process (BMAP). Although PHDs, MAPs and BMAPs do not belong to the distributions or stochastic processes that are commonly used in simulation tools, it is usually straightforward to integrate them into simulation software by drawing on the underlying Markov chain which governs the activity of the process. We will further study at an alternative approach to generate such numbers whereby the start and end states of an arrival epoch are utilized to set a cumulative distribution for the inter-arrival time which in turn can then be randomly sampled to generate the latter. We will carry out: (1) the simulation of the underlying Markov chain such that arrivals and their corresponding arrival times are recorded and alternatively (2) the discretization of the cumulative distribution function indicated by the start and end states of arrival epochs and random sampling of the latter to produce random inter-arrival times. We compare both approaches, (1) and (2), in terms of accuracy and efficiency. This research aims at expanding the study and practice of special point processes such as PHDs, MAPs and BMAPs via simulation.*

I. INTRODUCTION

In this research, we propose the simulation of three stochastic processes, the Phase-Type Distribution (PH), the Markovian Arrival Process (MAP) and the Batch Markovian Arrival Process (BMAP) such that inter-arrival times are randomly generated. We carry out: (1) the simulation of the underlying Markov chain such that arrivals and their corresponding arrival times are recorded and alternatively (2) the discretization of the cumulative distribution function indicated by the start and end states of arrival epochs and random sampling of the latter to produce random inter-arrival times. We present Approach (2) as a powerful, traceable and equally accurate simulation technique for arrival processes, as well as a generic model which can be further deployed and manipulated to flexibly simulate diverse stochastic processes on the non-negative axis.

We will first simulate the processes under study by mimicking the internal transition activity in the underlying Markov chain, such that arrivals are registered by their occurrence times. The simulation of the underlying Markov chain is based on the Stochastic Simulation Algorithm, yet tailored to adequately capture the specific features of the Markov chains lending their activity to each of the aforementioned arrival processes. We will refer to this approach as Approach (1): Simulation of Underlying Markov Chain.

We further attempt at providing an alternative algorithm to simulate these processes. We try to mimic the inversion method popularly used for all distributions that have a definite and invertible cumulative distribution functions. Since this is not the case for arrival processes understudy, as the cumulative distribution function is either impossible to invert or both non-invertible and indefinite, we attempt at developing an approximate inversion method that falls under the concept of discretization of continuous processes. We will refer to the alternative simulation algorithm by Approach (2): Approximate Inversion Method. Approach (2) first dictates specifying the cumulative distribution function of the inter-arrival time and assessing whether it is invertible or otherwise. We then further utilize this function in creating an inter-arrival time/cumulative probability database which then can be randomly sampled to generate inter-arrival times. The algorithms behind Approaches (1) and (2) were constructed using JavaScript supported on Net Beans IDE platform.

II. OVERVIEW OF PHDs, MAPs AND BMAPs

The analysis of man-made systems such as computer systems, communication networks, manufacturing processes, logistics networks, to mention only a few, is commonly done through the utilization of discrete-event models [1] or simulation [2]. A key and prerequisite stage of developing simulation models is input modeling. Under input modeling, a stochastic model is constructed to capture the key features of an input process from which statistical and real-time measurements are available.

Historically, more often than not, it was assumed that inter-event times are independent and identically distributed and accordingly a best-fitting distribution is selected from a given set of distributions to stochastically model the input data in simulation models. However, sometimes the usual set of available distributions is not flexible enough to capture measured behavior and/or the assumption of independence and identicality of events doesn't hold because the events are rather correlated as it is the case for many real-life systems.

In this context, Markovian processes could be utilized to model many real-life systems, namely PHDs, MAPs and BMAPs. PHDs are known to be very flexible and allow the approximation of a wide variety of distributions on the positive axis [3]. Nevertheless, the use of PHDs was in the past was

mainly restricted to a few of its subclasses, such as the Erlang or hyper-exponential distributions [4]. The approximation of a general distribution by a phase-type distribution is a complex non-linear optimization problem, which has only recently been developed into comprehensive computational algorithms, of which only a few are available in today's popular modeling software [5].

The modeling power of Markovian processes lies in the fact that they can largely reflect the correlation between inter-event times, especially MAPs and their generalized counterpart, the BMAPs [6]. For instance, the analysis of single-server queues with MAP input or service completions is developed and based on well-known matrix analytic techniques primarily led by Neuts [7]. Yet, parameter fitting for MAPs is more complex than that of PHDs [8], and hence is not richly discussed in the literature. Only recently have several algorithms for generating MAPs from measured data become available [9, 10]. Similar to PHDs, MAPs can be integrated in simulation models but again this approach is not really supported by available simulation tools because the first approaches describing the integration of MAPs in simulation models have been published only recently [11, 12, 13].

The inter-arrival time for PHDs, MAPs and BMAPs is a matrix-exponential random variable. The generation of matrix-exponential random variables hasn't been comprehensively studied in the literature. Yet, one can simply generate the latter by utilizing the transition mechanism in the associated Markov chain as we propose in the first simulation approach. However, this can become extensive for higher order Markovian point processes and/or processes with small transition probabilities.

The BMAP is a stochastic point process that generalizes the MAP by allowing correlated arrival batches as opposed to single-unit arrivals, in addition to dependent and correlated inter-arrival times [14]. The origins of the BMAP can be traced back to the development of the versatile Markovian point process (VMPP) by Neuts. Neuts' primary objective was to extend the standard Poisson process to account for more complex customer arrival processes in queuing models. Neuts founded the VMPP on the notion of the phase-type distribution, playing a major role in the advancement of the use of the PHD in queuing theory. Lucantoni further extended the original definition of the VMPP to define the Markovian arrival process (MAP) [14].

III. THE PHASE-TYPE DISTRIBUTION

In a continuous-time Markov Chain (CTMC), a state can either be transient or absorbing. If the long-term probability of absorption in a state is less than 1, then the state is transient; and otherwise absorbing, if the probability converges to 1 in the long run. If every state in a chain is either transient or absorbing, then the Markov chain is called an absorbing Markov chain. A particular example of absorbing Markov chains is the single-absorbing state Markov chain, which lends its transition activity to the underlying Markov chain of a PHD.

If a random variable X has a phase-type distribution, then it has the representation $X \sim (\alpha, D_0)$. The vector α describes the state probabilities as to where the chain is to arbitrarily start. Whenever the single absorbing state is hit, the chain restarts itself according to α.

The transition dynamics within the CTMC underlying a PHD can be described in terms of their respective rates, and hence via D_0 and d_1.

D_0 is an $n \times n$ non-singular and invertible matrix describing the transitions among the transient states, such that:

$$D_0(i, j) = \begin{cases} -\lambda_i & \text{if } i = j \\ \lambda_{ij} & \text{if } i \neq j \end{cases}, \text{for } \forall i \text{ and } j, i, j = 1, 2, \ldots, n$$

d_1 is an $n \times 1$ vector matrix describing the transitions from the transient states to the single absorbing state, such that:
$$d_1(i, n + 1) = \lambda_{i(n+1)}, \text{for } \forall i, i = 1, 2, \ldots, n$$

The identity that $D_0 1 + d_1 = 0$ holds, where 1 and 0 are n-vectors of ones and zeros respectively. In the long-term, the probability of absorption is 1 and the matrix $(-D_0)^{-1}$ is the fundamental matrix of the absorbing CTMC, such that $(-D_0)^{-1}(i, j)$ is the expected time spent in state j before absorption given that the chain started in state i. It is worth noting that d_1 is redundant and can be implicitly derived from D_0.

Given the preceding description of single-absorbing state Markov chains, Neuts derived the concept of the phase-type distribution (PHD). In this sense, the phase-type distribution is the distribution of the lifetime of the single-absorbing state Markov chain $\{X(t)\}_{t \geq 0}^{\infty}$; i.e. the time to enter the absorbing state from the set of transient states. For a PHD, the transient states are called phases and the order of the distribution is defined as the number of transient states.

The cumulative distribution function of X is as given by Equation (1):
$$F(t) = 1 - \alpha e^{D_0 t} 1 \tag{1}$$
The kth moment of X is given by Equation (2):
$$E[X^k] = k! \, \alpha (-D_0)^{-k} 1 \tag{2}$$

IV. MARKOVIAN ARRIVAL PROCESSES: MAPS AND BMAPS

After a PH renewal occurs, the underlying Markov chain is immediately restarted based on the initial probability vector α which predicts the starting state of the chain succeeding a renewal. However, there are many real-life applications, especially in the telecommunications industry, in which strong correlation exists between subsequent inter-event intervals [15], and so it is only natural that the concept of dependence of subsequent intervals is introduced. Accordingly, the phase distribution becomes dependent on the last phase visited and upon which a renewal was triggered.

In other words, each state in the underlying Markov chain can behave as an absorbing state and cause an arrival whenever hit by an observed transition. This can be visualized as a CTMC in which two changing variables are registered, the phase and the level of the system. A hidden transition changes the phase of the

systemas the arrival epoch jumps from one state to another. On the other hand, an observed transition changes not only the phase of the system, but also the level of the system as it causes the occurrence of an arrival epoch. This Markov chain describes the Markovian Arrival Process, shortly referred to as a MAP. This process is heavily used to describe a variety of arrival processes in today's queuing models.

Nonetheless, in today's computer and telecommunication networks, it is very common for multiple jobs to be sent to the server simultaneously, and hence jobs arrive in batches [15]. Therefore, MAPs which describe single arrivals can be further generalized to account for arrival batches of different sizes in the Batch Markovian Arrival Process, commonly referred to as the BMAP. A MAP then becomes a subclass of the BMAP, where correlated single arrival epochs are allowed, and the PHD is a subclass of the MAP whereby single and independent arrival epochs occur.

MAPs generate a single type of events, and so this can be extended by allowing K different batch sizes resulting in a BMAP defined as $(D_0, D_1, D_2, ..., D_K)$, where matrices D_k are non-negative and associated with those arrivals of batch size k.

Let $J \equiv \{J(t): t \geq 0\}$ be an irreducible, continuous-time Markov chain with state space $S = \{1, 2, .., n\}$, where n is a finite and positive integer. Suppose J has just entered state i, such that $i \in S$. The process spends an exponentially distributed amount of time in state i with rate λ_i. The process then transitions to state j, and the transition could be hidden or observed with batch size k. A hidden transition occurs with a probability $P_0(i, j)$ where $i \neq j$. An observed transition of batch size k occurs with a probability $P_k(i, j)$ for $k \geq 1$ where i could be equal to j. Hence Equation (3) is valid.

$$\sum_{k=1}^{\infty} \sum_{j=1}^{n} P_k(i,j) + \sum_{j=1, i \neq j}^{n} P_0(i,j) = 1 \; for \; \forall i, i \in S \quad \textbf{(3)}$$

The transition rates of the aforementioned jumps can be formulated as follows:

$$D_0(i,j) = \begin{cases} -\lambda_i & for \; i = j \\ \lambda_i P_0(i,j) & for \; i \neq j \end{cases} \&$$
$$D_k(i,j) = \lambda_i P_k(i,j) \; for \; \forall i, j \in S, k \geq 1$$

Therefore, D_0 contains the transition rates of J for which no arrivals occur, while D_k for all $k \geq 1$ contains the rates of the observed transitions of batch size k. If $N(t)$ denotes the total number of arrivals up to time t, then the joint process $\{N(t), J(t): t \geq 0\}$ is a BMAP, or a MAP if K is 1.

Inter-arrival times are defined as the accumulation of the sojourn times in all of the states visited prior to the occurrence of the arrival epochs. Knowing that the phase distribution is sensitive to the state at which an arrival epoch occurs, the very initial state of the underlying Markov chain can be predicted using the stationary or steady-state probabilities defined by ϕ, such that an arbitrary arrival epoch begins in any state i with a probability ϕ_i [16]. `

Let X_n be the state of the underlying Markov process $J(t)$ at the time of the nth event occurrence, and T_n the time between the $(n-1)^{th}$ and n^{th} event occurrences, then $\{X_n, T_n\}_{n=1}^{\infty}$ is a Markov renewal process. In particular $\{X_n\}_{n=1}^{\infty}$ is a Markov chain whose transition probability matrix β is given by Equation (4):

$$\beta = (-D_0)^{-1} \sum_{k=1}^{K} D_k \quad \textbf{(4)}$$

The stationary probability row vector can be computed by solving equations (5) and (6).

$$\phi\beta = \phi \quad \textbf{(5)}$$
$$\phi 1 = 1 \quad \textbf{(6)}$$

The mth moment of the inter-arrival time X can be calculated using (7) for both MAPs and BMAPs:

$$E[X^m] = m! \, \phi(-D_0)^{-m} 1, m = 1, 2 \ldots \quad \textbf{(7)}$$

In our research, we will deploy ϕ to initiate the Markov chain under Approaches (1) and (2). Under Approach (1), the Markov chain is triggered using ϕ, and then progressed according to D_0 and D_k up until the maximum number of arrival epochs is achieved. Under Approach (2), ϕ is used to start up the process, then it is restarted according to the absorption status of preceding event according to β such that the correlation and dependence among the inter-event times is not neglected.

V. APPROACH (1): SIMULATION OF UNDERLYING MARKOV CHAIN|

Utilizing the continuous-time Markov chains (CTMCs) underlying PHDs, MAPs and BMAPs, we stochastically simulate the transition dynamics in the chain, registering arrivals and their occurrence times, in addition to their associated batch sizes in the case of the BMAP.

A. The Batch Markovian Arrival Process (BMAP)

The Markov chain is first randomly initiated in any state by randomly sampling ϕ. Given that the chain is at state i, it will spend there an exponential amount of time. With a probability $P_0(i, j)$, the subsequent phase of the chain is j, yet the level remains constant. With a probability of $P_k(i, j)$, not only is the subsequent phase j, but also the level of the process increases by k, marking the occurrence of an arrival epoch of size k. The procedure conducted to record arrivals straightforwardly relies on the sampling of the discrete transition probabilities. The inter-arrival time of an epoch is the aggregation of the exponential sojourn times in all states hit by hidden transitions.

B. The Markovian Arrival Process (MAP)

Since the MAP is a subclass, or a special case of the BMAP, the procedure explained earlier applies; however, the maximum batch size allowed would be 1.

C. The Phase-Type Distribution

For every arrival epoch, the underlying Markov chain is simulated until the single absorbing state is hit. The chain is first initiated by randomly sampling α, and then progressed according to D_0. All transient states visited by an epoch prior to absorption are tracked such that the inter-arrival time is

calculated as the sum of the exponential sojourn times in those states.

VI. APPROACH (2): APPROXIMATE INVERSION METHOD

The execution procedure of Approach (2) is comprised of an initialization and setup procedure in addition to an iterative simulation procedure. In the initialization and setup stage, the cumulative distribution function is identified given the start state and absorption status (state and batch size) of the inter-arrival time and then discretized by setting an appropriate bound on the error and then formulating a time/cumulative probability database to be further utilized in the simulation.

In the simulation procedure, given the start and end states of an arrival epoch, the corresponding cumulative probabilities are randomly sampled such that the appropriate cumulative probability is located using the binary search algorithm and finally linear interpolation is used to estimate the corresponding inter-arrival time.

A. The Batch Markovian Arrival Process (BMAP)

The objective of this approach is to set up the row vectors given by (8) and (9); i.e. the objective is to discretize the phase-type distribution. The vector given by (8) is a vector of increasing time values, such that successive pairs differ by an increment δ. Consequently, (9) is the evaluation of the cumulative distribution function corresponding to the time values in (8).

$$t = [t_o \; t_1 \dots t_{max}]$$

Such that, $t_i = t_{i-1} + \delta$ where $t_0 = 0, \delta = \frac{E[T]}{\varepsilon}$, **(8)**
$i = 1, 2 \dots max$

$$F_{ij}^k(t) = \left[F_{ij}^k(t_0) \; F_{ij}^k(t_1) \dots F_{ij}^k(t_{max}) \right] \quad \textbf{(9)}$$

1. Initialization and Setup

We denote the following:

→ $\boldsymbol{\beta} = (-\boldsymbol{D_0})^{-1}[\boldsymbol{D_1}|\boldsymbol{D_2}|\dots|\boldsymbol{D_K}]$ such that $\beta_{i(j-kn)}$ is the probability that an arrival epoch will get absorbed in state $(j - kn)$ with batch size k given starting in state i.

→ $\boldsymbol{Q}$, infinitesimal generator matrix $\boldsymbol{Q} = \begin{bmatrix} D_0 & D_1 & \dots & D_K \\ 0 & 0 & \dots & 0 \end{bmatrix}$

Given the start and end states, as well as the batch size of an arrival epoch, **(9)** can be evaluated as follows:

$$e^{tQ} = \begin{bmatrix} - & \boldsymbol{R^1}(t) & \dots & \boldsymbol{R^K}(t) \\ 0 & 0 & & 0 \end{bmatrix} \quad \textbf{(10)}$$
$$P(X_m(t) = j, J_m = k, t_m \le t | X_m(0) = i) = R_{ii}^k(t)$$

$$F_{ij}^k(t)$$
$$= \left[\frac{R_{ij}^k(t_0)}{\beta_{i(j+kn)}} \; \frac{R_{ij}^k(t_1)}{\beta_{i(j+kn)}} \dots \frac{R_{ij}^k(t_{max})}{\beta_{i(j+kn)}} \right] for \; \forall \; i,j,k \quad \textbf{(11)}$$

We end up with $(n \times n \times K)$ vectors of (11) corresponding to the times vector (8), to be utilized in the simulation procedure of Approach (2).

2. Simulation

For the very first arrival epoch, the starting state is randomly predicted by sampling ϕ. For subsequent arrival epochs, the starting state of one arrival is the ending state of the one preceding it, highlighting the correlation among the events.

The absorption status of an arrival epoch is predicted by sampling the vector $\beta(iStart, i)$ for all $i = 1, 2, \dots n \times K$, such that iStart is the starting state of the epoch. State i is not the true end state of the arrival epoch m, it rather reflects both the absorbing state as well as the associated batch size, and so it is manipulated to derive iEnd and k according to the following:

$$iEnd = i - kn, where \; k = floor(i/n).$$

According to iStart, iEnd and k, the corresponding cumulative probability vector is extracted from the database and further utilized in the random sampling to approximate the associated inter-arrival time.

B. The Markovian Arrival Process (MAP)

Similar to Approach (1), the execution procedure of Approach (2) for a MAP is the same as that of the BMAP, but reduced such that all entries associated with batch sizes greater than 1 are eliminated.

C. The Phase-Type Distribution

The cumulative distribution function of the inter-arrival time associated with the phase-type process is given in (1). Yet, it is mathematically impossible to invert (1), and hence the approximate nature of Approach (2). The objective of this approach is to set up the row vectors given by (8) and (12).

$$F(t) = [F(t_0) \; F(t_1) \dots F(t_{max})] \quad \textbf{(12)}$$

1. Initialization and Setup

Knowing that $F_X(t)$ represents the cumulative distribution function of $X \sim PH(\alpha, D_0)$, then $F_X(t)$ is by definition continuous and strictly increasing from zero to 1 for all t, such that $t \in [0, \infty)$. It can be explicitly proven that $F_X(t)$ is differentiable over its domain of definition $t \in [0, \infty)$, then it can be expanded about any point a, such that $a \in [0, \infty)$ using Taylor's theorem as given by (13).

$$F_X(t) = F_X(a) + F_X^{(1)}(a)(t - a)$$
$$+ \frac{F_X^{(2)}(\tau)}{2!}(t - a)^2, \quad \textbf{(13)}$$
$$a \le \tau \le t$$

Using linear approximation, (13) is reduced and the error term is designated by (14).

$$\varepsilon(t, a) = \left| \frac{F_X^{(2)}(\tau)}{2!} \right| (t - a)^2 = \frac{\varepsilon(\tau)}{2}(t - a)^2 \quad \textbf{(14)}$$

Accordingly, the smaller the difference between t and a, the error term given by (14) converges to zero, implying that linear approximation given by (15) is a good estimate of $F_X(t)$.

$$F_X(t) = F_X(a) + F_X^{(1)}(a)(t - a) \quad \textbf{(15)}$$

However, under Approach (2), $F_X(t)$ is known for all time values given in (8). The cumulative distribution function values are sampled to estimate the corresponding inter-event time. For

every arrival epoch m, $F_X(t_m) = r, r \sim U(0,1)$ and the corresponding inter-event time t_m is located in an interval $[t_i, t_{i+1}]$, such linear approximation or alternatively linear interpolation can be utilized to estimate t_m.

The main issue is then to determine increment size δ which ensures that the error term doesn't exceed a certain maximum as will be demonstrated hereafter.

$F_X^{(2)}(t)$, the second derivative of (1) is computed for all time values in (8), and $max\{F_X^{(2)}(t)\}$ is utilized to determine the size of δ as follows:

$$F_X^{(2)}(t) = \left[F_X^{(2)}(t_0) \ F_X^{(2)}(t_1) \dots F_X^{(2)}(t_{max}) \right]$$

$$max\{\varepsilon(\tau)\}_{\tau \in [t_0, t_{max}]} = max\left\{F_X^{(2)}(t)\right\}_{t \in [t_0, t_{max}]} \tag{16}$$

$$max\{\varepsilon(t,a)\}_{\tau \in [t_0, t_{max}]} = \frac{max\{\varepsilon(\tau)\}_{\tau \in [t_0, t_{max}]}}{2} \delta^2 \tag{17}$$

We start with an initial estimate of $\varepsilon_{initial}$, we have that $\delta_{initial} = E[X]/\wp_{initial}$, and then perform the following simple check to assess the accuracy of $\varepsilon_{initial}$ and modify according (18).

$$\delta = \begin{cases} \delta_{initial} & if \ 1/\wp_{initial} \geq max\{\varepsilon(t,a)\}_{\tau \in [t_0, t_{max}]} \\ \sqrt{\dfrac{2}{\wp_{initial} \times max\{\varepsilon(t,a)\}_{\tau \in [t_0, t_{max}]}}}, otherwise \end{cases} \tag{18}$$

After determining the adequate spacing of (8) using (18), we reconfigure vectors (8) and (12) accordingly, setting up by that the database for the simulation procedure.

2. Simulation

The simulation procedure is similar to that of the MAP and the BMAP, such that (12) is randomly sampled for every arrival and the corresponding inter-arrival time is estimated via linear interpolation.

It is worth mentioning that Approach (2) was inspired by the work done by Brown, Place and Liefvoort on the generation of matrix exponential random varieties [17]. Analogous to what has been presented later in the context of the algorithm of Approach (2), the authors suggest: (i) the generation of a uniform random number "r" on the interval [0, 1], (ii) setting a maximum decay value for which the cumulative distribution function can be computed with confidence, (iii) if "r" is less than the decay value, then the bisection method is used to locate "r" in a vector of increasing cumulative distribution function values and consequently evaluate the corresponding time value utilizing the decay value and an exponential tail.

VII. APPLICATIONS OF PHDs, MAPs AND BMAPs: RESULTS AND DISCUSSION

A. Randomly Populated Examples

To eliminate bias to either simulation approach, we first assess the efficiency of Approach (2) versus Approach (1) by applying both on the same set of examples. Random PHD, MAP and BMAP examples were generated such the corresponding required parameters were randomly populated.

The output is composed of the average inter-event time, estimates of the variance and skewness (in addition to lag-1 autocorrelation for MAPs and BMAPs) and the duration of simulation procedure for Approaches (1) and (2), as well as the duration of the setup procedure for Approach (2). Randomly populated PHD examples of orders 1 through 10 were run and replicated ten times each. MAPs of orders 2 through 10 and BMAP of orders 2 through 8, each with batch sizes 2 through 5 were also simulated.

1. Randomly Populated MAPs and BMAPs

In terms of the accuracy of estimates, both approaches result in equally accurate estimates of the mean inter-arrival time. However, the simulation of the Markov chain captures the variability of the underlying process and the skewness of the distribution of the inter-arrival time more accurately than Approach (2). Nonetheless, Approach (2) is more accurate in reflecting the correlation among the arrival events. It is worth noting that true values of the aforementioned estimates always fell within the 95th confidence interval of their estimates.

The duration of the setup time of Approach (2) increases as the order and/or maximum batch size of the process increases due to the increase in the complexity of the matrix operations involved. Additionally, the simulation using Approach (1) is almost two times faster than that using Approach (2).

Simulating a fully populated MAP or a BMAP via utilizing the underlying Markov chain is found to be faster because any instantaneous transition between any two states could mark an arrival; the transitional dynamics lend this simulation approach its superior efficiency and accuracy as compared to Approach (2). However, if the probability of arrival; i.e. absorption in one state, given starting in any state were to be reduced, we presume that Approach (2) would be more appropriate. Instead of fully populating D_1, arrival can be restricted to and/or from a specific number of states, which is effectively the case in many real-life applications of MAPs and BMAPs.

Unlike any MAP example whose D_1 is fully and randomly populated, Approach (2) is much more efficient for MAPs and BMAPs with restrictions on marked transitions, such as reducing their probabilities or rates relative to the total departure rate from the origin state and/or limiting marked transitions to a reduced set of specific states. Not only is the simulation time less, but also the overall speed of the execution process under Approach (2) is greater.

2. Randomly Populated PHDs

While the setup time of Approach (2) increases with the order of the PHD, the simulation time varies only slightly, such that no particular trend can be realized for the variation of the simulation time as a function of the order. Generally, the duration of the simulation procedure under Approach (2) fluctuates closely about its average and regardless of the order of the PHD. For Approach (1), the duration of the execution procedure; i.e. the simulation process, increases with the order of the underlying Markov chain. Similarly for Approach (2), the execution time increases with the order of the underlying

Markov chain as the size of the sub-generator matrix grows and the required matrix operations become more extensive. However, given that the simulation time varies only slightly with the order, the increase in the execution time is mostly due to the increase in the setup time, such that the sensitivity to the order of the underlying Markov chain is shifted to the setup stage under Approach (2) as opposed to the case of Approach (1).

B. Effect of Changing Variability of Underlying Process

A sought after property of any simulation approach is the flexibility to accurately reflect the variability of a process. We chose the balanced two-level mixture of Erlangs, which is a subclass of PHDs also referred to as the hyper-Erlang, to analyze the sensitivity of Approach (2) namely to the change in the variability of the process. We applied both approaches to simulate a specific case of a balanced two-level mixture of Erlangs of order 9. The mean is held constant at, yet the variance was changed over 9 variants of the same example. It was deduced that decreasing the variability of the underlying process decreases the duration of the setup procedure of Approach (2) and improves the accuracy of the estimates under both approaches.

VIII. CONCLUSION

In this research, we proposed the simulation of the Phase-Type Process (PH), the Markovian Arrival Process (MAP) and the Batch Markovian Arrival Process (BMAP) such that inter-arrival times are randomly generated via an approximate inversion method, referred to as Approach (2). We used the simulation of the underlying Markov chain as a reference approach (Approach 1) to which our proposal was compared and thusly assessed. The power of Approach (2) lies in its generality, such that given the parameters of any distribution; a database of time-values vs. CDF-values can be compiled and used to perform simulation of corresponding distributions and/or stochastic processes. Results suggest that the Approach (2) is capable of performing as accurately and efficiently as its counterpart. Yet, they also show that simulating the underlying Markov chain is generally faster for fully-populated MAPs and BMAPs. It is also worth noting that the simulation process under Approach (2) is completely trackable such that the number of random numbers utilized in the procedure is a constant and deterministic quantity sensitive to the number of events simulated, unlike the case under Approach (1), whereby this quantity is stochastic and cannot be specifically estimated. This renders the feat of variance reduction more attainable under the proposed approximate inversion method.

ACKNOWLEDGMENT

REFERENCES

[1] W.J. Stewart "Probability, Markov Chains, and Simulation", Princeton University Press, Princeton (2009).

[2] A.M. Law and W.D. Kelton "Simulation Modeling and Analysis", 3rd Edition, McGraw Hill, Boston (2000).

[3] C.A. O'Cinneide "Phase-Type Distributions: Open Problems and a Few Properties" (1999), pp. 764-779

[4] C.H. Sauer and K.M. Candy "Computer Systems Performance Modeling", Prentice Hall, Englewood Cliffs (1981).

[5] P. Buchholz, J. Kriege and I. Felko "Input Modeling with Phase-Type Distribution and Markov Models, Theory and Applications", Springer Cham Heidelberg New York Dordrecht London (2014), pp. 1-27, pp. 105-114

[6] M.F. Neuts "A Versatile Markovian Point Processes", Journal of Applied Probability Vol: 16, (1979), pp. 764-779

[7] M.F. Neuts, "Matrix-Geometric Solutions in Stochastic Models", John Hopkins University Press, Baltimore (1981)

[8] M. Telek and A. Horváth "Markovian modeling of real data traffic: Heuristic phase type and MAP fitting of heavy tailed and fractal like samples", Proceedings of the Performance 2002. Lecture Notes in Computer Science, Vol. 2459, pp. 405–434. Springer (2002)

[9] H. Okamura., T. Dohi, and K.S. Trivedi "Markovian arrival process parameter estimation with group data", IEEE/ACM Transportation Networks, (2009), pp. 1326–1339

[10] M. Telek and G. Horváth. "A minima lrepresentation of Markov arrival processes and a moments matching method", Perform (2007)

[11] F. Bause, P. Buchholz, and J. Kriege "A Comparison of Markovian Arrival Processes and ARMA/ARTA Processes for the Modeling of Correlated Input Processes", Proceedings of the Winter Simulation Conference (2009)

[12] S. Heckmüller and B.E. Wolanger "Using load transformations for the specification of arrival processes in simulation and analysis" (2009)

[13] C. Brickner, D. Indrawan, Williams, D., and S.R. Chakravarthy, "Simulation of a stochastic model for a service system", Proceedings of the Winter Simulation Conference (2010), pp. 1636–1647

[14] J.D. Cordeiro and J.P. Kharoufeh "Batch Markovian Arrival Processes (BMAP)", University of Pittsburgh, Directorate of Force Management Policy Headquarters, United States Air Force (2010)

[15] B.U. Narayan "An Introduction to Queuing Theory: Modeling and Analysis in Applications, 2nd Edition (2015)", Chapter 10, Markovian Arrival Processes

[16] B.U. Narayan "An Introduction to Queuing Theory: Modeling and Analysis in Applications", 2nd Edition, (2015)

[17] E. Brown, J. Place J. and A.V. Liefvoort "Generating Matrix Exponential Random Variates", University of Missouri-Kansas City, Simulation Councils 70:4, pp. 224-230 (1998)
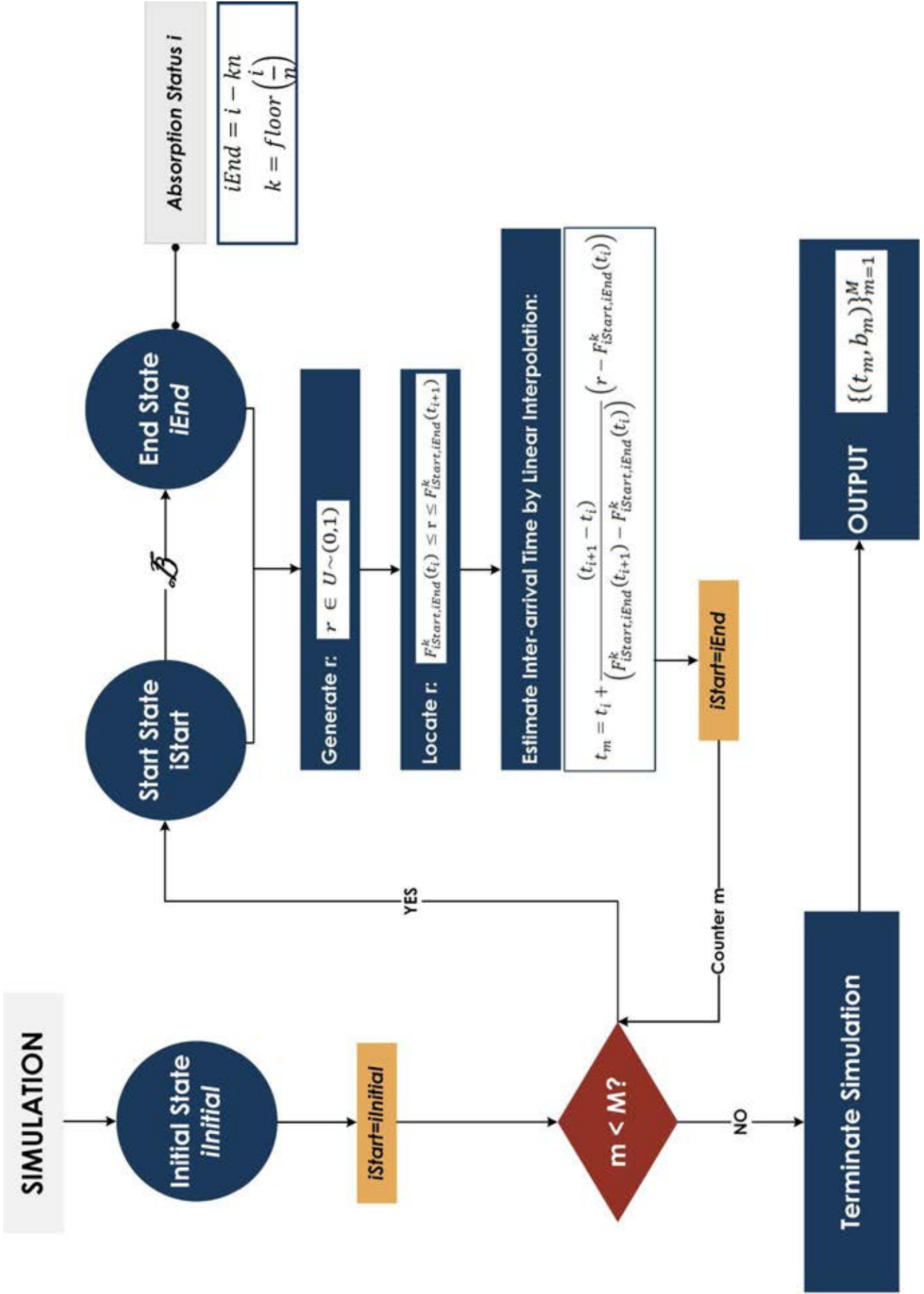
Figure 1 Simulation procedure of Approach (2) for BMAPs and MAPs
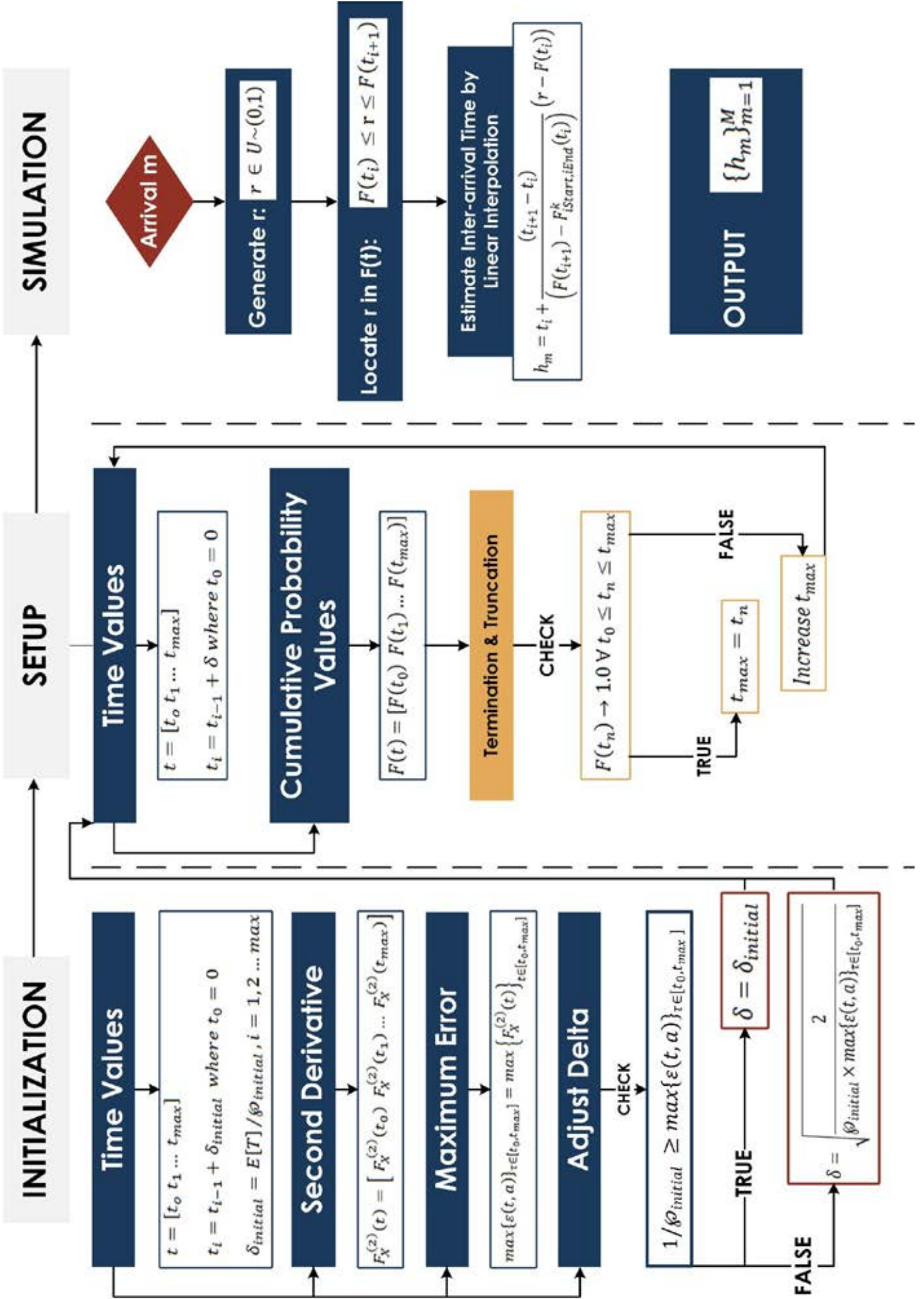
Figure 2 Execution procedure of Approach (2) for PHDs

"Le Blé Croissant" ARENA Simulation

Boutros,K. ; Dawaf,T. & Saad,E.
Department of Industrial and Mechanical Engineering
Lebanese American University
Byblos, Lebanon
E-Mail: kevin.boutros@lau.edu, thomas.dawaf@lau.edu, emilio.saad@lau.edu

I. ABSTRACT

This paper addresses the simulation of *Le Blé Croissant*'s croissant production line. *Le Blé Croissant* is a bakery which produces several baked goods, of which the most produced and highly demanded good is the croissant. *Le Blé Croissant* was visited on regular days to model and record every process of the croissant-making by hand, to come up with the respective probability distributions. The production line was then simulated using Arena, representing the resources accurately, which made it possible to view the queue, number of entities processed, time, and more statistical data. There are no obvious problems in the process, yet the simulation reveals much more about what could not be seen by the naked eye. The analysis of the data allowed the discovery of problems in the production line and thus lead to creating adequate alternatives to optimize results by at least minimizing time and increasing output for the same input. Moreover, all the data collected can be found and formed into distribution functions.

II. INTRODUCTION

At the *Le Blé Croissant* bakery, a series of processes form raw material of flour, butter and other basic dough ingredients into a final output product, the croissant. First, a worker would compress butter blocks of approximately 1.4 kilograms each, to form rectangular sheets of 1 cm of thickness. These sheets are then used inside the dough of approximately 5 kilograms in a process called lamination. There are approximately 100 sheets produced every day. The dough is then sent to the freezers to cool the butter down, so that it will not melt. After the freezing process, one worker would start laminating the dough, piece by piece, and book folding it so that it would have 32 layers of butter. Then, he puts the dough on an automatic laminator, which will spread the dough into a long sheet of 3.8 mm thickness, approximately 3 meters in length, and 50 cm in width. This sheet is cut into around 75 triangles weighing 120 grams when filled with chocolate, cheese, or thyme. These pieces are then rolled into the croissant which then placed on trays onto a trolley. The process of cutting the dough and placing them on a trolley would take around 2 to 3 minutes. When the trolley is full, it is taken to fermentation which

would take 2 hours 45 minutes. After the croissant has gained its shape, it is sent to bake in an oven on high temperature for 14 minutes. The croissants are then sorted into the three flavors and packed according to customer orders. After packing, the orders leave the bakery to the specific customer at the desired time.

III. OBJECTIVES OF STUDY

Knowing the information about the process above, the following study objectives were put:

- Total output when all butter and dough are exhausted
- Entity waiting times in queues
- Resource utilizations

When the above points are studied, weaknesses will be targeted and alternative solutions will be looked for. An alternate model will be created containing the solutions. The results of the improved model will be compared to those of the original model.

IV. LITERATURE REVIEW

In the bakery business, simulation is a famous tool used to model and run the production process using software such as ARENA, WITNESS, and FLEXISM. Hosseinpour and Hajihosseini (2009) [1] list specific issues simulation addresses in production systems:

- Evaluation of the effect of a new piece of equipment on an existing manufacturing system
- Throughput analysis
- Bottleneck analysis
- Times parts spend in queues
- Queue sizes
- Utilization of equipment or personnel

(These points were taken into consideration in the result analysis of the simulation, to reveal the key problems at *Le Blé Croissant* that need to be addressed). In Germany, Hecker et

al.(2010) [2] modeled and simulated a small bakery's production line that produces several kinds of products using ARENA. The bakery comprised of a mixer, dough divider, forming machine, dough retarder, dough proofer, rack oven and tunnel oven. They collected data for one shift of 6 hours and 10 minutes, in order to formulate the model and run it. After succeeding in simulating their model, the team got the numbers they needed to optimize the system. It turns out that by rescheduling both production plans of products and bench worker breaks, they were able to cut down on time and cost wastes, to increase efficiency.

Banks (1998) [3] lists several rules for determining whether or not the technique is appropriate and helpful:
1. The problem can be solved using commonsense analysis
2. The problem can be solved analytically
3. It easier to change or perform direct experiments on the real system
4. The cost of the simulation exceeds possible savings
5. Proper resources are not available for the project
6. There isn't enough time for the model results to be useful
7. There are no data – not even estimates
8. The model cannot be verified or validated
9. Project expectations cannot be met

At *Le Blé Croissant*, identifying problems by just looking around was challenging. Moreover, management wanted to explore other approaches to increase efficiency and reduce cost, including replacing machines. Therefore, the use of simulation was adequate to what management wanted and led to satisfactory results.

As for the data requirements, Kibira and McLean (2002) [4] include the following as essentials:
- Product design, bill of materials, assembly constraints
- Assembly process specification including operation
- Sequences, process times, and testing requirements
- Expected product demand
- Failure and repair data for production line systems
- System configuration and layout options
- Unit cost of labor, tooling, and materials

Indeed, all of the above were accounted for in *Le Blé Croissant*'s study, which lead to accurate and reliable results.

However, the review of literature revealed scarcity of sources when it comes to bakery simulation. The *Le Blé Croissant* simulation is one of the first bakery simulation projects in the Middle East.

V. MODEL AND ASSUMPTIONS

Butter slices arrive to the butter compressing station in a batch of 100. It is assumed that the butter has been already sliced before the process starts. These slices are compressed by Worker 1 using a butter compressing machine, according to a Lognormal distribution of 20+LOGN(4.71,5.75) seconds, in order to have the butter easily spread on dough further on. Each of the compressed butter is then put in a container shared by three workstations, each with one worker carrying out the same process, meaning that the three workstations share the same queue. It is assumed that the dough is put at each of the mentioned workstations before the croissant-making process begins. In order for the lamination stations to have the same queue, a Worker Set is created having workers, each with different expressions, since every worker processes the dough differently. Each of the three workers now takes a unit of compressed butter and spreads it on a unit of dough. The three workers operating the Lamination station are Worker 2, Worker 3, and Worker 1 who has compressed all the 100 pieces of butter in the first process before joining Worker 2 and Worker 3 in the Lamination process. When a unit of dough has been spread with butter, a new worker, Worker 4, transports the latter to a freezer with a maximum capacity of 100 units of dough. The freezing process follows a uniform distribution of 60+UNIF(20,30) minutes. Meanwhile, all workers take a break equal to the freezing time. When the freezing process is over, the dough are removed one by one in the same order as they arrived and are taken to the book folding station where Worker 1 awaits. There, the dough are folded by Worker 1 following a Weibull distribution 23.5+WEIB(11.2,1.71) seconds. The dough is then put on an adjacent automated lamination machine which laminates the dough into a larger piece in a constant 2 minutes. When finished, the bigger piece of dough is lifted to a large table where Workers 2,3,4, and 5 cut the dough into small pieces, fill them with the corresponding filling, and roll them into the desired croissant shape. Worker 1 joins the rest in the Cutting, Filling, Rolling process when finished with the Book Folding process. The cutting, filling and rolling of an individual croissant is a very fast process, so it is assumed that all these actions act as one process following a Poisson distribution of POIS(145) seconds. The difference in time between workers is too small to record, so it is assumed that they all operate with the same distribution.

Note that the processes of book folding and automatic lamination share the same queue. To do so, the book folding process is a Seize Delay, and the automated lamination is a Delay Release. The cut, filled and rolled croissants are then put

on trays of 12. The number of croissant trays is not always the same and follows a Beta distribution of 5.35+1.65*BETA(0.822,0.922)-1 trays. These trays are batched into batches of 32 and put onto trolleys which are taken to the Fermentation process. The croissants are left to ferment in fermentation rooms with regulated temperatures for a constant 165 minutes. When done fermenting, Worker 5 takes the croissants, still on the trolleys, to an oven with a maximum capacity of 9 trolleys, which carries out the Baking process. Note that the Leave/Enter method was used in several places to seize a worker for transportation, which is in fact what happens. Worker 4 takes the laminated dough from the lamination process to the freezing process. Later, Worker 1 takes dough from the freezer to his book folding station to process them. Finally, Worker 5 takes the trolleys of fermented croissants to the oven for baking. Between other processes, routing is almost instantaneous and does not require a Leave/Enter method because of the negligible impact. After baking for a constant 1 hour, the beautifully smelling croissants exit the system to be packed and sent to hungry customers. The following flowchart represents the ARENA process map of the original model:

Figure 1. *Le Blé Croissant* Original Model

VI. SOURCE OF DATA AND DATA COLLECTION TECHNIQUES

After observing the process, simulation requires data gathering in order to use the acquired data to build the model. The data gathering technique was to go to the factory and record the output of the processes. It is crucial that every step is recorded in order to get more accurate results. Moreover, the number of data points must be large enough to form a probability distribution which is feasible (has a chi-squared test p-value greater than 0.1). However, some processes were summarized into one. For the Cutting, Filling, Rolling process, the three steps were combined because it is almost impossible to record five workers working fast on these steps in order. First of all, parts were assigned on the job since the team could not stay together due to many simultaneous processes which needed to be recorded. The first started at 8:00AM recording the butter compression process, recording 45/95 repetitions, each varying between 12 and 20 seconds. While the compressed butter was being sent to the Lamination process, which is 2 to 5 seconds away (almost instantaneous since they are sent by batches), two members were recording each worker's time on the lamination machines. There were only two workers at first before the first worker, who was already on the butter compression machine, moved in. One worker placed dough on the laminating machine (also instantaneously since there is a worker who removes them while the butter is being compressed) so that the butter would be added to it in the Book Folding process, to get 4 layers of butter in between the layers of dough. This process varies for each worker, as 34 data points were collected for the first, 35 data points for the second and only 20 for the third since he was previously on the butter compression and did not have enough dough to get the same number of data as the others. After inputting these values into the input analyzer, the resulting p-values were less than 0.1, so more information was gathered on another day and fitted into another distribution, which resulted in the approximate distribution for each worker. After being laminated, each piece of dough was put in the freezer. The freezing time is a uniform 20 to 30 minutes after the final dough has entered the freezer. Rather than being recorded, the freezing time was provided by management because of time constraints, but in both days the freezing time was 23 minutes and 27 minutes. Then, one worker removes the dough, one by one, from the freezer so that another worker book folds it again to create 32 layers. The latest worker places the dough on an automated laminator, to spread the dough into a 3 meter sheet, which is then placed on a table to be cut, filled and rolled by five workers. For the Book Folding process, 40 intervals were recorded, ranging between 20 to 40 seconds depending on the
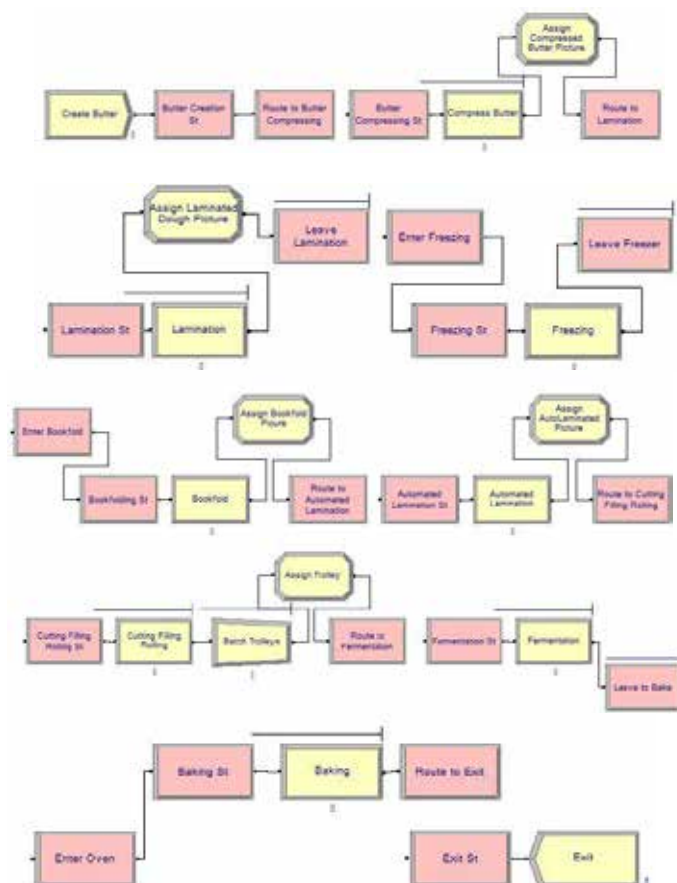
dough. The automated laminator takes a constant two minutes to laminate the dough. The worker in charge of book folding does not prepare more than one piece of dough at a time while waiting for the lamination to be done. After the automatic lamination, the dough is placed on a table where one worker slices it into triangles (around 70), other workers separate the triangles, one worker fills them with the flavor and then rolls them into shape. All of this process is recorded as one, which takes only 2 to 3 minutes. The interval times of the sub-processes are very small to record. This whole process forms around 65 to 80 croissants, depending on how much waste is cut out of the dough. Batches of 12 are placed on trays and batches of 32 trays are placed on trolleys. The problem faced while placing 70 entities after the Separate was "more than 150 entities". So, a distribution for the trays was created, which also generated the same error. After analyzing the data thoroughly, it was revealed that one trolley takes approximately 6 to 7 pieces of dough to be completed. A distribution was then created according to the number of dough instead of croissants on the trolley. After each trolley is filled, it is sent to a fermentation process around 3 minutes away. This process is a constant 2 hours 45 minutes for each trolley where only 9 trolleys can be placed at a time. The final croissants, which do not fill the trolley, are sent to freezing and are packaged as frozen, a process not of interest in this specific simulation. After fermentation, each trolley enters an oven, which can only have 2 trolleys at a time, and undergoes a baking process of a constant 14 minutes before being sent to packaging and shipping. In the meantime, the workers, who have a low utilization in the results, would go and prepare the butter and dough for the next day to be used, another process which is not of interest. Since the last 3 to 5 hour process was not taken into consideration in the simulation, the whole 12 hour process was around 9 hours on ARENA, which is valid.

VII. INPUT AND OUTPUT

Since each process flows differently at *Le Blé Croissant*, data had to be collected in order to give the processes their respective characteristic probability distributions. These distributions vary between Lognormal, Triangular, Weibull, Uniform, Poisson and Beta. It is essential that a Normal distribution is not used, since there is a slight chance a negative number is generated, which is impossible for a time variable.

The following table summarizes the input ranges and output distribution of each process:

Process Name	Range of Input (seconds)	Output
Butter Compression	[12,20]	20 + LOGN(4.71,5.75)
Lamination I (Three workers)	[80,120];[105,180];[60,100]	TRIA(68.5, 108, 119); 57.5 + 22 * BETA(0.88, 0.88); LOGN(84.2, 7.89)
Freezing (After last entering)	[1200,1800]	60 + UNIF(20,30)
Book Folding (Lamination II)	[20,40]	23.5 + WEIB(11.2,1.71)
Automated Lamination	[120]	[120]
Cutting, Filling, and Rolling	[120,180]	POIS(145)
Route to Fermentation	[180,300]	3 + 3.39 * BETA(1.91, 1.85)
Fermentation	[9900]	[9900]
Baking	[840]	[840]

Table1. Process Input and Output

VIII. RESULTS AND ANALYSIS

After running the complete simulation, the first thing realized is that the simulation ends at around 9.1 hours, which is in fact how much time the croissant-making process is taking at *Le Blé Croissant*. The average number of trolleys that have exited the system is 16, meaning that the total number of croissants at the end of the shift is 6144 croissants, which was validated with the company log books. The frozen dough is not facing any major queuing issues, except for leaving the freezer to the book folding station where, on average, the waiting time is around 1.7 hours, with 3.46 hours being the maximum waiting time. This is logical since the book folding, automatic lamination and cutting, filling and rolling processes all share the same queue, meaning that rather than having the dough wait outside the freezer to be processed, it stays in the freezer until the one before it is done being processed. This prevents the dough from melting and getting ruined. However, a waiting time of 1.7 hours for this production line is not practical. The dough is being processed one at a time rather than having a bulk of dough running through the process at the same time. This large waiting time can be easily dealt with in order to increase production and decrease lost time in the process (Practical approaches are covered in the Recommendations and Suggestions). Another issue to address is the average waiting time on the compress butter station. A batch of 90 to 100 pieces of butter arrive to the butter compressing station to be compressed. On average, each piece is waiting for around 11 minutes, with a maximum waiting time of 21 minutes. The butter compressing machine is of regular size and requires one worker. Rather than having the butter pieces cut before hand

and enter one by one to be compressed, a larger machine is required to cut down valuable time. As for the remaining processes, it was found that they are satisfactory in terms of resource utilization as well as waiting times in queue relative to the outcome of one unit. For example, at the lamination process, even though the dough is waiting an average of 17 minutes to be processed, the dough already being processed is directly put in the freezer and being used after the freezing time, giving the following one-queued processes a hefty load of work. So, no big harm is done to the process. However, the butter compressing station is being charged with a large load of work instantaneously and needs to process the butter pieces one by one so that they could be spread on a dough. It is crucial that the butter compressing does not take too much time so that the dough does not melt and , at the same time, have Worker 1, who operates the butter compressor, help workers 2 and 3 with the lamination process as soon as possible. Moving on to resource utilization, it was realized that there are remarkable differences in those of the workers. In the croissant-making process, Worker 1 is working around 60% of the time, Workers 2 and 3 around 21% of the time, and Workers 4 and 5 around 15% to 16% of the time. It is crucial to remember that *Le Blé Croissant* produces various baked goods other than the croissant, meaning that when Workers 4 and 5 seem to have very low utilizations, they are in fact working on another production line when they are not needed for the croissant-making process. When Workers 1, 2 and 3 are laminating the dough, Workers 4 and 5 are busy processing other entities in other production lines such as the donut line, which is not the scope of the simulation. Furthermore, since the simulation comprises of a bakery's croissant production line, there are several processes that require larger amounts of time, such as freezing, fermenting and baking. During these processes, the workers are idle in the simulation, making their utilizations drop. Yet, in fact, the workers are carrying out different tasks on different production lines, meaning that their general utilization is not low. Of all the workers, Worker 1 has the highest utilization in the croissant-making process. This is because he is actually the chief worker with the most experience and is in charge of more tasks than the other workers. Worker 1 also has a higher salary, justifying his high utilization and presence at almost every process. As for routing, since the processes at *Le Blé Croissant* are close to each other, entities are not taking too much time to reach their destination. Transfer times are almost instantaneous and do not take valuable time.

Now that the numbers have been taken care of, issues linked to the results that are seen only when visiting *Le Blé Croissant* can be highlighted. First of all, as previously remarked, the butter compressor is given so much to do instantaneously, where its output is also needed as fast as possible. Furthermore, the waiting time in queue for leaving the freezer is quite high, where it can be reduced in several ways. Moreover, the workers are not well organized, going from station to station and pitching in on numerous processes. What *Le Blé Croissant* needs is automating the production line and increasing machine capacities, which in turn increases production and decreases wasted time.

After highlighting what can be done to improve the system, *Le Blé Croissant* was theoretically upgraded to a semi-automated production line, consisting of a shock freezer, automated cutting, filling and rolling machine, fermentation tunnel, freezing room, and integrated butter feeder and dough preparation machine. Clearly, the new system combines the dough preparation process, which was not of interest in the simulation because of the constant times, and the croissant-making process in one process, which only takes an additional 2 hours to complete. In the improved system, around 15000 croissant pieces exit the system, almost 2.5 times the original output. Furthermore, all workers have very close utilizations that are on average 25%, meaning that the system is much more automatic now. Introducing new automatic machines with larger capacities increased output remarkably as well as cut down lost time in the process. The following flowchart represents the ARENA process map of the revised model:
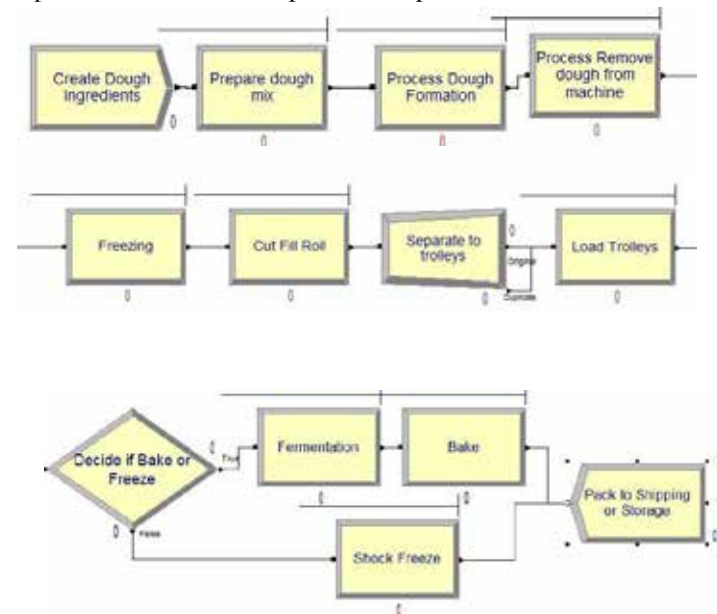
Figure 2. *Le Blé Croissant* Revised Model

IX. RECOMMENDATIONS AND SUGGESTIONS

After meeting management, it was revealed that a budget of around 150,000$ was allocated to upgrade the bakery over 3 years. Given such a budget, a thorough research of bakery equipment was made, to come up with the following recommended production line for *Le Blé Croissant:*

- A butter feeder: This machine reduces the process time of cutting and compressing butter by 1 hour. The whole 25 kg butter block enters the feeder and exits as butter sheets. The price of one butter feeder is approximately $8,000.
- An automatic cutting and filling line: Replacing the Cutting, Filling, Rolling process, the automatic cutting and filling line produces around 4000 pieces (10 trolleys) per hour, using only 3 workers. The price of this line is approximately $100,000 new and $60,000 used, given *Le Blé Croissant* production capacity.
- A shock freezer with a capacity of 1400 pieces per hour: This type of freezer reduces the number of baked entities and stops the yeast from being activated in an unused time. The price of one freezer is approximately $30,000.
- A tunnel proofing system for the fermentation process: Trolleys push one another on their way to the oven, which takes 2 hours and 30 minutes (the same as the fermentation time). The capacity of one tunnel is 10 trolleys. The system is also more energy and space efficient. The price of one tunnel is approximately $15,000.
- A stock freezing room for the shock freezer's output: The room has a capacity of 40,000 pieces, ready to be baked or shipped out as frozen products. The price of one freezing room is approximately $20,000.

Because of the freezing room, *Le Blé Croissant* has two options: shipping out frozen goods or baking the goods at any desired time. Therefore, a new oven is not needed. This freezer room also solves the problem of large orders at an unexpected time.

X. SUMMARY

The simulation of *Le Blé Croissant*'s croissant-making process revealed so much valuable information that can help improve their production line. The model, created on Rockwell's ARENA, consists of the croissant-making process, assuming that the dough is already made and present at the corresponding workstations. Every process and resource behavior was recorded, to come up with respective probability distributions to characterize processing times. After running the simulation, it was revealed that the resulting numbers are logical and reflect the real life process accurately. Having analyzed the results, it was discovered that entities are taking a large amount of time to leave the freezer. Moreover, the butter compressor resource is being handed a large amount of butter at the same time, where it can only process them one at a time, making the waiting time per entity relatively high. Since the general manager wanted to automate the system with an adequate budget at hand, adding a shock freezer, butter feeder, fermentation tunnel, automatic cutter and filler and stock freezing room would be ideal for *Le Blé Croissant*. Another model was then created with all of the above improved resources, and was later simulated. It was revealed that the output was around 2.5 times more than the original model, with reduced worker utilizations and other more efficient aspects. As a conclusion, the new system is in fact very efficient and essential in order to increase output and cut down costs.

REFERENCES

[1]. F. Hosseinpour and H. Hajihosseini, "Importance of Simulation in Manufacturing," *World Academy of Science, Engineering and Technology*, vol. 51, pp. 285–288, 2009.

[2]. F. Hecker, W. Hussein, and T. Becker, "Analysis and optimization of a bakery production line using ARENA," *International Journal of Simulation Modelling*, vol. 9, no. 4, pp. 208–216, Dec. 2010.

[3]. J. Banks, "Principles of Simulation", *Handbook of Simulation*, New York: Wiley, 1998, pp. 3-30.

[4]. D. Kibira, et al., "Virtual Reality Simulation of a Mechanical Production Assembly Line", *Proc. 2002 Winter Simulation Conf*.

Lean Management Impact on Lebanon Government Sector Performance

Hadi Ramzi Abou Fakher, Dr Ramzi Fayad
Industrial Engineering & Engineering Management
Beirut Arab University
Debieh
Hadi.abufakher@hotmail.com, r.fayad@bau.edu.lb

Abstract-**Lean management has started effectively at the late of the 20th century at Toyota production systems, this lean thinking had become a huge area of research in the past 25 years and started to be adopted in the private and public service sectors.**

The Lebanese deteriorated record of servicing citizens has led to civil unrest. This instability has played a major role in the call for improvement within the government public sectors. The government has to do major alterations and improvements to their work culture and structure while maintaining relatively low operational expenses. Lean management was proven to be the optimal solution for governments sectors in order to empower their employees and increase productivity while cutting costs in their departments.

In this paper, we will discuss lean management thinking as an approach that focuses on operational aspects. Lean management could be translated in Lebanon public sector in an operational system that maximizes value added, reduces operation costs and eliminates waste in all process through the value stream. We will shed light on the challenges of applying lean management in Lebanon public sector. A step by step model will be created to facilitate applying lean management in government departments.

I. Introduction

Lean management in government departments is one of the most advanced management techniques applied today. It focuses on creating citizens' satisfaction and as well as creating value through waste reduction. Lean management focuses on value and cost reduction. The effectiveness and efficiency of lean applications has been proven in many successful cases worldwide such as Toyota, GM, Nissan and etc [7]

This basic principle was transferred to many other domains and is applicable outside production sector. Early 21st century lean management transferred into service systems [4]. Several government agencies and departments, such as the U.S. Environmental Protection Agency and the States of Iowa, Minnesota and Washington and etc, are using Lean management to improve the quality, transparency, efficiency and speed of government processes. By referring to the lean government center, we can say that lean management is today applied at many governments with magnificent effect on their countries.

In January of 2010, Connecticut's Governor, M. Jodi Rell, announced that her state would be continuing and expanding their effort to improve state government processes through lean methods. Agencies across the state have changed the way they do business by adopting lean. Lean management was introduced in the Swedish Migrant Board's in 2009, upgraded management practices, and introduced continues-improvement techniques. At the end of 2010, average application-processing times had fallen to half. [3]. The G8 nations have introduced lean management in all their departments to get out of the fiscal and political trap, they have to improve their performance in the public sector, doing more and better with less. They were able to increase their public sector productivity by 1.5 percent annually, this small improvement were able to generate benefits worth $1 trillion annually, up to 2.5 percent of their GDP.

Through analyzing various case studies about countries and agencies applied Lean management, we realize that today it is the best time to apply lean management in Lebanon government departments, in order to improve government performance and service delivery efficiency while cutting costs.

II. Lean Management

Lean is a model that focuses on the customer- the person at very end of the line, who gets services we are delivering- and empowers employees [1]. It's about creating value in order to reduce cost, improves efficiency, productivity, quality, customer satisfaction and morale. It is about the process not the people.

A lean support continues improvement cycle. It's about incremental improvements now instead of 100% later or never. It's not about eliminating or devaluing what people do. It's about how we add things in a way that we remove waste so that the end customer gets the most value. When we remove waste we don't necessary throw things away, we take resources from a place they don't add value to another place they add value. The goal of lean is to end up with process that accomplishes what is needed with least amount of waste.

Lean is a proven, systematic approach for eliminating or minimizing waste that results in the delivery of services at the lowest possible cost. Lean is every process, every system, and every employee in the government departments [6].

III. Waste in Office

Waste is any non-value added activity, which adds problems and blocks the flow of value. Office waste is hard to see it, but easier to feel it. There are four types of office waste: information waste, process waste, physical environment waste, people waste.

A. Information Waste
Any redundant input and output of data is considered waste in an office, information waste is visible in terms of incomplete information systems, the manual checking of data, data inputs that is never used, reentering data, converting formats forms, unnecessary data, unavailable data, missing unknown data, unclear or incorrect data, data safety issues, data discrepancies.

B. Process Waste
Process waste is any defect, scrap, rework, and work around, inspecting and double checking, approvals of unnecessary administrator, variable flow in a process, high inventory rate of data, waiting time, over processing.

C. Physical Environment Waste
Physical environment is any travel from one office to another, organizing a messy meeting room before meeting, going out of your work station to office supply person for stationary.

D. People Waste
People waste is any unclear role of an employee, lack of training, multitasking, worker task interruption, underutilization of talents, hierarchy mistakes, lack of strategic focus, handoffs.

IV. Flow in a Process

There are typically two types of flow, batch processing and single item flow. Batch processing is waiting a certain number of documents for processing. This type of processing will lead for a bottleneck. While single item flow or lean flow is processing each document at a time, through lean flow service can move quickly and continuously through the system.

V. Challenges of Transformation

Many challenges might face the government transformation towards lean management in Lebanon, such as political, economical, legal, social and cultural challenges. With the right tools and practices challenges can be overcome [2][5].

A. Political Challenges
The regime in Lebanon and the politicians has lost people trust and legitimacy, for that they have to start changing their attitude and irresponsibility towards citizens. For that reason new government must show the Lebanese people that it is here for them and willing to change the system. Lean management application requires political commitment and long term vision of the political leadership of the country.

B. Administrative Challenges
Since independence, the authoritarian political regime in Lebanon created a highly centralized public bureaucracy. Citizens faced lack of quality in public sectors, time-consuming processes and waiting time in queues, due to the centralized structure of administration. Moreover this system is operated by mediocre dissatisfied public employees, who lack of efficient training, organization and customer service delivery, corrupted to self interests. This combination of administration irresponsibility, centralized system and mediocre employees will create a resistive environment for applying any corrective and improving systems that could affect their self interest.

C. Economical Challenges

The application of lean management requires an annual budget to cover lean management expenses. Its application necessitates employees training, IT involvement, and reengineering design. Unfortunately, due to the low government performance over these years, Lebanon is facing an economical crisis with public debt that reached 70 billion dollars. For that reason, ministries should afford its expenses from reducing and minimizing the operations costs within.

D. Social and Cultural Challenges

The degrading reputation that Lebanese have toward the public sector service provided as well as the lack of respect towards the citizen besides demanding bribes to get the work done. This way of work has been going on for a long while and it would require a long period of time for people to change their view and opinions about the service provided by the government. The application of lean management need trust relationship between the citizens and the public organizations.

VI. Impact of Applying Lean Management

Lean reduces the non-value added work, increase quality in process, maximizes value added, and reduces operational costs, reduce wait time and queue, and eliminates waste in all process through the value stream. Increase the efficiency of the service system inside the government departments.

Lean management enhances simplicity and transparency and visualization of processes. Improve decision making system through changing managerial attitude which change from report decision system into site decision system. Lean management is centered on making obvious what adds value by reducing everything else. In order to emphasize on the different aspects of the service line and what could be considered essential in minimizing the expenses.

Lean management optimize the services provided within Lebanon public sector and end up with a process that accomplish what is needed with the least amount of waste. Moreover it creates a lean culture in government departments that empowers employees and increase their satisfaction and morale. This work will lead to an increase in citizens' satisfaction and motivates them to pay their taxes; unfortunately, this reflects positively on the Lebanese.

VII. Methodology

To improve the service efficiency, we devised a new methodology of applying lean management in a service sector. This methodology is detailed in the flow chart in appendix 2.

Moreover, relating to the implementation of Lean tools and techniques, it is important to understand current practice and processes from concurrent studies. Assess the pros and cons and try to benefit from the feedback provided.

The questionnaire which would be filled in with citizens will be analyzed using Statistical Analysis Package in order to correlate the factors to be studied; the questionnaire is detailed in a table in appendix 1.

The methods that will be used for data collection are semi structured interview and observations. After data are obtained from the interviews, the data will be analyzed in order to recommend appropriate Lean tools and techniques.

VIII. Lean Tools and Techniques

Visual Management, 5S, Value Stream Mapping, Process Mapping, Fishbone Diagram, Red Tag Techniques

IX. Conclusion

This paper has shed light on the challenges and the opportunities of applying lean management in Lebanon public service sector and to convert Lebanon ineffective government public service performance into a new effective one to fulfill citizens' needs and aspirations. In this paper we defined waste in office, the flow of process, and the impact of applying lean management in Lebanon government public sector. We detailed a methodology of applying lean management in Lebanon public sector in a flow chart.

This project is part of ongoing research at the university and would be submitted toward my master thesis. For the near future, the work planned is as follows: a questionnaire would be filled in by customer being serviced at this public governmental department. After that, the data obtained from the questionnaire would be loaded into SPSS and analyzed statistically in order to verify our hypothesis. After that Lean tools and techniques will be recommended appropriately.

AS well as simulating the current process in the considered public sector.

REFERRENCES

[1]   Perumal Puvanasvaran, Robert Kerk Swee Tian, Suresh A/L Vasu, Lean environmental management integration system for sustainability of ISO 14001:2004 standard implementation, Journal of Industrial Engineering and Management, 2014, vol. 7, no. 5

[2]   Scorsone, E.A New Development: What are the Challenges in Transferring Lean Thinking to Government? Public Money & Management, 2008, vol. 28, no. 1, pp. 61-64

[3]   Miniam Gebre, Petter Hallman, Mark Minukas and Beca O,briem, Transforming government performance through lean management, McKinsey center for government, December 2012, pp. 3-16

[4]   Nina Bahatia & John Drew, Applying lean production to public sector, McKinsey Quarterly, June 2006, pp. 42-45

[5]   Aly, Wael Omran. "Lean Production Role in Improving Public Service Performance in Egypt: Challenges and Opportunities." Journal of Public Administration and Governance 4, 2014, no.2, pp. 90-105

[6]   Francois Bouvard, Robert Carsouw, Eric Labae, Alastair Levy, Lenny Mendonca, Jaana Remes, Charles Roxoburgh, Samantha Test, Better for less: Improving public sector performance on a tight budget, McKinsey center for government, July 2011

[7]   Minh, Nguyen Dang, and Nguyen Thi Van Ha. " Made in Vietnam" Lean Management Model for Sustainable Development of Vietnamese

Enterprises." *Procedia CIRP* 40 2016, pp. 603-608.

جدول إحصائي للوقت الذي يمضيه المواطن لانهاء معاملته

ملاحظات	رضى المواطن	عدد الزيارات	وقت الخروج	وقت الدخول	نوع المعاملة	المواطن

ملاحظاتعامّة--- ---------------------------------- ------------------------------

التاريخ واليوم:----------------------------
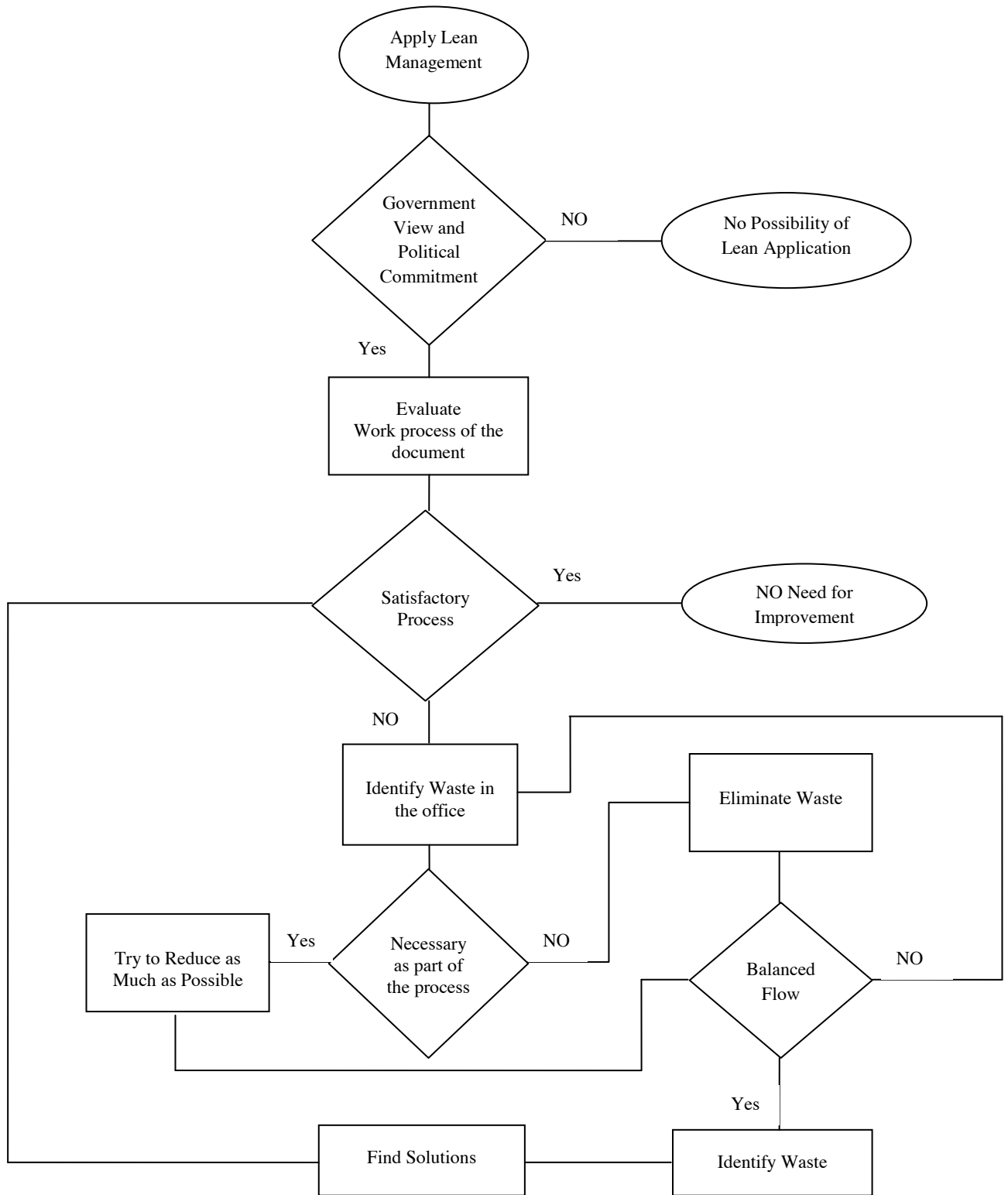
LEARNING CURVE THEORY: THE CASE OF THE CONSTRUCTION INDUSTRY

Daoud Kiomjian
Engineering Management Program
American University of Beirut
Beirut, Lebanon
Email: dak18@mail.aub.edu

Abstract— **This paper reviews the literature on learning curve theory, and presents a suite of learning curve models. The models are then evaluated from the perspective of the construction industry, by assessing their applicability and relevance to construction activities. Due to the complex nature of construction activities, and the significant variability in characteristics between different construction projects, there is still no consensus in the literature regarding the model that provides the best fit and predictive capability for construction productivity data. Accordingly, this paper presents a new recursive learning curve model, which was formulated to accommodate for the significant mechanization existing in the modern construction industry, and for the forgetting effects that are caused by interruptions. Since the proposed model is recursive, it puts more emphasis on recent data. When tested on four case studies, the model forecasted cumulative average completion times with an error of less than 3% for the four cases..**

I. INTRODUCTION

"Practice makes perfect" is a proverb that exists in almost every culture and language. From experience, the time and effort required to perform any task decreases as the number of repetitions increases [1]. This decrease is caused by a phenomenon called the "learning effect", which is represented mathematically by the "learning curve". The learning curve concept originated in the airplane manufacturing industry, and was first formalized by T.P. Wright in 1936 [2]. Subsequently, applications of the learning effect spread throughout other sectors within the realm of manufacturing [3]: the automotive industry [4], the chemical industry [5], photovoltaic production [6], and semiconductors [7, 8, 9, 10, 11, 12]. As witnessed in these fields, understanding the dynamics of the learning effect through the use of the learning curve supports planning. The construction industry is no different and also stands to gain significantly when applying the learning curve [13].

The first reference to learning in the construction industry is in the 1965 Economic Commission for Europe report entitled *Effect of Repetition on Building Operations and Processes on Site* [14]. Since this initial report, the learning effect has been applied to various site activities such as prestressed concrete piles [15], formwork [16], rebars [17] and caisson construction [18]. The learning effect also applies to the various stages of construction projects: design [19], bidding [20], planning [21] and even claim management [1, 15].

Despite the importance and the relevance of the learning effect in the construction industry, the literature addressing this matter is still limited [22, 23]. There is no consensus regarding a model that provides good fit to historical data, while also having acceptable predictive capabilities. A model with good fit is a model that can match historical data, while a model with good predictive capability is one that can use historical data in order to predict future data. The predictive capabilities of a model are particularly important in the construction industry as proper planning and resource allocation requires the prediction of future costs and durations. A contractor will not benefit from a model that perfectly matches the rebar fixing times in past projects, while predicting negative completion times when extrapolated to future data.

II. LEARNING IN CONSTRUCTION

As in any other industry, learning in the construction industry is not instantaneous and happens gradually. Generally, the learning process is divided into two stages: the operational learning stage and the routine acquiring stage [14, 24, 25]. It is during the first stage that the worker, acquires the basic skills and becomes familiar with the task. For example, during this stage workers, tasked with the installation of drainage pipes, learn the functionality of the excavating machines and become familiar with the soil in which they are working. During the second stage, the performance of the workers is boosted as shortcuts are discovered. It is during this stage that the workers know how the equipment will react to the various soil types and start organizing the workplace in a manner that permits optimal performance. However, these two stages are overlapping since the workers can start optimizing performance while also becoming familiar with the equipment.

When measuring learning as a reduction in time to complete a task or cycle, the curve will be steepest following the *startup point* – the time to complete the first cycle. Following the startup point, the curve will begin to level beyond the *standard production point* or the point where the minimum time or maximum reduction in time to complete the task is achieved [15]. The major decrease in the cost or time, when compared to the initial cost or time, happens at the routine acquiring stage. After this stage no further improvement is witnessed and the performance reaches its steady state.

Given this two-point understanding of the learning process, multiple learning curve models have entered the literature. Figure 1 highlights the main classes of learning models and shows how they differ in their mathematical form; the boxes in dark gray are the models reviewed in this paper. This selection was based on our desire to highlight models of practical use within the field of construction – both multivariate and stochastic models, while useful, are not practical as they require significant data and expertise to calibrate properly.

The oldest learning curve model is the *Wright Model*, $y = Ax^{-n}$ [26]. In this model, y is the cumulative average, cost per unit, A is the cost of the first unit, x is the repetition or repetition cycle number, and n, a value between zero and one, is the slope of the logarithmic curve [24]. The learning rate, L, can be derived from the slope of the logarithmic form by using $L = 2^{-n}$. This learning rate information is of particular interest to a construction manager as it can be used to benchmark performance. Learning rates for various trades, that can be used as a starting point or benchmark, are listed in Hijazi et al. [27] and Gottlieb and Haugbølle [25]. The Wright Model remains the most widely used learning curve model due to its simplicity [4, 28] and its ability to provide acceptable precision while having a simple mathematical structure [29].

Despite its popularity, one shortcoming of the Wright Model in construction is that it ignores the workers' previous experience. The *Stanford-B Model*, $y = A(x + B)^{-n}$, developed by the United States' Department of Defense, uses B to represent the number of experience units and shift the curve downwards [30, 31]. If $B = 0$, then this model reduces to the basic Wright Model [30, 25].

However, both of these models assume a continuous decrease in the cost, which is not realistic since there is a minimum non-zero time required to complete every task. The *Plateau Model*, originally introduced by Baloff [4], $y = C + Ax^{-n}$ solves this problem through the use of C, representing the steady state performance of the worker; all other terms are as introduced previously [32, 26]. Knecht [33] also recognized that learning does not continue *ad infinitum* and proposed a model with a changing learning rate, $y = Ax^{-n+1}/(1 - n)$, where the value of n, as before, falls in the interval $(0, 1)$.

While the plateau and Knecht [33] models accommodate for a limit on learning, construction workers are also limited by the technology in use on-site. For example, the learning curve for manual excavation using a shovel is limited by the maximum labor performance. Whereas, excavation done by a mechanical excavator limits the operator's learning by the specifications and capabilities of the excavator. Therefore, when the operation is less mechanized and highly manual the time or cost is more prone to decrease with learning [34]. The *DeJong Model*, $y = A[M + (1 - M)x^{-n}]$, developed in 1957 includes an incompressibility factor, M, that ranges from 0 to 1 [27]. Thus, if the process is completely mechanized, $M = 1$, no improvement will arise with more repetitions [30].

The *S-curve model*, $y = A[M + (1 - M)(x + B)^n]$, devel-oped following World War Two combines both mechanization and experience [30]. In this model, M is the incompressibility or mechanization factor and B represents the acceptable units of previous experience. Zhang et al. [21] developed a similar variation of the S-Curve Model which accommodates for the effects of experience, steady state labor productivity and mechanization. This model, termed the *Improved Learning Curve Model*, is represented by $y = AM + C_0 + [A(1 - M) - C_0](x + B)^{-n}$, where C_0 is the standard time needed to complete the product under optimal conditions with perfect labor; all other terms are as before.

Most planners and estimation engineers assume flat learning rates and calculate costs. Of course, learning is rarely so steady. As such, the Polynomial Models – quadratic and cubic – may exhibit better fit. Both the *Quadratic Model*, $y = A + \beta_0 x + \beta_1 x^2$, where A is the cost of the first unit, β_0 is the initial slope and β_1 is the quadratic factor [35] and the *Cubic Model*, $y = A + \beta_0 x + \beta_1 x^2 + \beta_2 x^3$, where β_2 is the cubic factor, require one to assume or calculate the coefficients. Unlike the learning rate and the cost of the first unit, these parameters have no direct practical meaning and thus may be difficult to estimate or justify. Moreover, these models do not have any limiting parameters allowing for negative estimates of the costs. Therefore, they are unsuitable for using past data to extrapolate and predict future performance as is necessary when preparing bids in construction.

The exponential model was recommended by the Norwegian Building Research Institute as a means to improve predictive capabilities of the model [14]. The most basic *Exponential Model* is that proposed by Knecht [33], $y = Ax^{-n}e^{cx}$, where c is a constant. Just as the Wright model has variations, so too does the exponential model. The *Three-Parameter Exponential Model*, $y = [k(1 - e^{-(x+p)/r})]^{-1}$, includes k as a maximum performance parameter expressed as the number of units per operation time, p as the previous experience parameter expressed in units of time, and r as the learning rate given in units of time; x in this model is the number of units of operation time. By taking the inverse on the right hand side, y becomes the cumulative average time to complete a unit after the passage of x units of time. The *Two-Parameter Exponential Model* is identical to the three-parameter model save for the exclusion of the term p. The value of the exponential models is their ability to estimate and forecast data over long production runs where the bound on learning is encountered. The limitation of the exponential models is that they are best applied to simple tasks such as fixing steel, manual excavation, cable pulling, installing wiring devices and backfilling [36].

A final set of models, the hyperbolic models, was originally designed to capture the effect of learning within compound measures of performance [20]. These models have since been adopted to capture the number of units that can be produced within x units of time [26]. In this paper, the models are manipulated further, by taking the inverse, to represent the cumulative average time required to produce a unit after x units of time have already been

invested. As such the *Two-Parameter Hyperbolic Model*, $y = [k(x/(x + r))]^{-1}$, and the *Three-Parameter Hyperbolic Model*, $y = [k((x + p)/(x + p + r))]^{-1}$, allow for the inclusion of a maximum performance parameter, k, a measure of previous experience, p, and the learning rate, r. The hyperbolic model can be used for novel and complex tasks [26]. Moreover, by adjusting the learning rate appropriately, one can estimate costs after long breaks or model the performance of the workers during crash periods when they are fatigued [37].

All of the models presented here assume that learning is measured as a cumulative average time to complete a task after a given number of repetitions or amount of time (as in the exponential and hyperbolic models) and that a decrease in that cumulative average time depends on only one variable – the number of repetitions. However, task time reduction within the construction industry likely depends on multiple site related factors. Unlike manufacturing, the layout in the location where a specific task is performed can easily vary from cycle to cycle despite the repetitive nature of the task itself.

A. Proposed Model

One strategy to capture site-specific information in the model is through the use of recursion. The need for a recursive model was noted by Adler and Clark [38] who recognized that experience or repetition alone was not enough to fully explain gains and losses in productivity among workers. By including a term relating the time required for the last or previous item to that of the current or next item, we can capture intrinsic changes in the learning process.

Given these considerations, we propose a recursive model represented by $y_n = y_{n-1}2^{-(r_{n-1}+\varepsilon)} + b_0(y_{n-1}H)^{-a} + M$, where y_n is the cumulative average cost or time required to produce the nth unit, y_{n-1} is the cumulative average cost of the previous unit or the last unit before an interruption (if applicable), r_{n-1} is the learning rate associated with the $n - 1^{th}$ item, ε is an additive factor that updates the learning rate between repetitions and prevents the productivity from reaching infinity, b_0 is a binary parameter that allows for the exclusion (0) or inclusion (1) of the experience gained prior to an interruption, H is the length of the interruption, a is a forgetting factor and, M captures the steady state performance, likely influenced by mechanization.

IV. FIT VERSUS PREDICTABILITY

The majority of the existing literature focuses on analyzing the goodness of fit of models rather than their ability to predict future performance [39]. The most widely accepted metric for the goodness of fit is Pearson's coefficient of determination, R^2 [24]. However, when using the learning curve in planning activities it is equally important to analyze a model's capability to predict future performance. The works of Everett and Farghal [40] and Farghal and Everett [39] are among the major contributors in this domain. The method suggested by these works requires splitting the existing set of data exactly in the middle and then applying the fitted equation to predict the

future values in the second set of data. The predictability of the model is then evaluated as an absolute percentage difference between the predicted and actual costs.

In order to study the capabilities of the models presented here, in terms of both fit and predictive capability, we use benchmark data from four different projects cited in the literature. The first project requires estimating the effect of learning in the construction of each floor in a 40 floor building in China [21]. The second project models learning across 19 cycles associated with the construction of tunnel formwork [39]. The third project studies learning across the cementing of 10 floors within a comparatively smaller housing project in Poland [14]. The fourth project examines learning across 26 cycles associated with installing formwork for floors in an office building [16]. These projects were selected given their representative nature in terms of size, type, and era of construction; with older projects representing cases with potentially less automation.

In reality, the model parameters could be obtained either by an optimization process that minimizes the least sum of squares or by using expert opinion. In this paper, the relevant parameters for all models were derived via an optimization process over the first half of the data. The specific models resulting from the optimization of the parameters are shown in Table **??**. Only one representative model was studied from each of the categories: the Wright model was selected from the Wright model and variations, the quadratic model was selected from the polynomial models, the basic Exponential Model from the exponential models, the Two-Parameter Hyperbolic Model from the hyperbolic models, and the proposed model as the only recursive model. The models excluded include those requiring the estimation of experience, mechanization or performance bounds as reasonable estimates for these terms were unknowable from the data provided in the case studies.

The results of both fitting the models to the first half of the data and using the fitted models to predict the second half of the data are summarized in Table II. The fit is stated as the R^2 for which a value closer to 1 represents a better fit; the predictability was measured using the Mean Absolute Percent Error (MAPE) for which a value closer to 0% indicates a better result. Table II also provides the p-value for a paired t-test between the model predicted cumulative average times per unit and the actual observed cumulative average times. In these tests, the null hypothesis is that the difference between the predicted and actual data is zero; a p-value greater than 0.05 supports this null hypothesis.

The results in Table II indicate that the proposed model exhibits a level of fit comparable to all other models. This is to be expected as the parameters were set with the goal of optimizing the fit across the first half of the data in all cases and for all models. The ability of the proposed model to predict the values in the second half of the data is superior in the first two cases, moderately worse in the third case, and comparable in the fourth case. Overall, however, when performing a paired t-test between the actual and the predicted values, the proposed model is superior in three of the four cases (Cases 1, 2, and 4).

Case 3 presents an interesting finding as all models confirm the null hypothesis of no difference between the observed and predicted values. Furthermore, the significance of the results in Case 3 indicate that the Exponential model is superior to the others. Not only does this confirm the findings in the article from which this case originates, but it also highlights the models' capabilities on small projects. Case 3 had only 10 data points – five used for fitting and five used for prediction. The real value of the proposed model is in application to large projects where the learning rate is likely to exhibit changes over time as in Case 1 with 40 floors.

V. Conclusion

Over the last four decades, the learning curve concept has gained popularity in the construction industry. One reason behind this increased popularity lies an industry trend towards cost control as a reaction to the steady increase in labor and construction material costs. As such contractors are developing schedules, checking progress and gathering data about site productivity on a regular basis. These data can be used to develop models describing cost reductions over time due to learning.

Various mathematical models have been developed for modeling learning in different industries. This paper presents a summary of the Learning Models from a mathematical and practical point of view within the construction industry. This review led to the proposal of a new model that accommodates both mechanization and forgetting. The proposed model is similar to the Wright model, but through recursion places more emphasis on the time consumed by the previous unit rather than the time used to construct the first unit. Our model demonstrates less than 1% error in predicting cumulative average unit construction times in three out of the four cases studied.

While these results are encouraging, further research is necessary to resolve multiple outstanding questions in the field. The issue of data representation should be handled with care when applying any learning model. Throughout this paper, the models focused on the use of a cumulative average for the xth unit or after x units of time. Using the cumulative average has the effect of smoothing the data which eases the process of finding a mathematical function to fit the data. However this smoothing process may make the predicted values less useable for the purpose of controlling activities on site. The proposed model, with its recursive structure may be better at supporting unit data as opposed to cumulative data relative to the other models. Further research is needed to verify this assumption.

Quality assurance and quality control are also possible future directions and the following question is to be answered: Does an improvement in quality occur alongside the observed reduction in time associated with learning? Moreover, multiskilling and job assignments could also be considered as managers must decide whether they should allow their workers to learn numerous trades or maintain their dedication to a single task.

References

[1] K. C. Lam, D. Lee, and T. Hu, "Understanding the effect of the learning-forgetting phenomenon to duration of projects construction," *International Journal of Project Management*, vol. 19, pp. 411–414, 2001.

[2] T. P. Wright, "Learning curve," *Journal of the Aeronautical Sciences*, vol. 3, no. 1, pp. 122–128, 1936.

[3] L. Argote and D. Epple, "Learning curves in manufacturing," *Science*, vol. 247, pp. 920–924, 1990.

[4] N. Baloff, "Extension of the learning curve–some empirical results," *Operational Research Quarterly*, pp. 329–340, 1971.

[5] M. B. Lieberman, "The learning curve and pricing in the chemical processing industries," *The RAND Journal of Economics*, vol. 15, pp. 213–228, 1984.

[6] G. F. Nemet, "Beyond the learning curve: factors influencing cost reductions in photovoltaics," *Energy Policy*, vol. 34, pp. 3218–3232, 2006.

[7] J. E. Cook, "A competitive model of the japanese firm," *Journal of Policy Modeling*, vol. 13, pp. 93–114, 1991.

[8] H. Gruber, "The learning curve in the production of semiconductor memory chips," *Applied Economics*, vol. 24, pp. 885–894, 1992.

[9] ——, "The yield factor and the learning curve in semiconductor production," *Applied Economics*, vol. 26, pp. 837–843, 1994.

[10] ——, "Trade policy and learning by doing: the case of semiconductors," *Research policy*, vol. 25, pp. 723–739, 1996.

[11] ——, "Learning by doing and spillovers: further evidence for the semiconductor industry," *Review of Industrial Organization*, vol. 13, pp. 697–711, 1998.

[12] S. Chung, "The learning curve and the yield factor: the case of korea's semiconductor industry," *Applied Economics*, vol. 33, pp. 473–483, 2001.

[13] S. Rosenbaum, M. Toledo, and V. Gonzalez, "Green-lean approach for assessing environmental and production waste in construction," in *Proceedings for the 20th Annual Conference of the IGLC, San Diego, USA*, 2012.

[14] U. Nations, *Effect of Repetition on Building Operations and Processes on Site: Report of an Enquiry*, ser. Document (United Nations). United Nations, Economic Commission for Europe. Committee on Housing, Building and Planning, 1965, no. 438. [Online]. Available: http://books.google.com.lb/books?id=5cV8QgAACAAJ

[15] J. Hinze and S. Olbina, "Empirical analysis of the learning curve principle in prestressed concrete piles," *Journal of Construction Engineering and Management*, vol. 135, no. 5, pp. 425–431, 2009.

[16] A. Jarkas and M. Horner, "Revisiting the applicability of learning curve theory to formwork labour productivity," *Construction Management and Economics*, vol. 29, no. 5, pp. 483–493, 2011.

[17] A. M. Jarkas, "Critical investigation into the applica-

bility of the learning curve theory to rebar fixing labor productivity," *Journal of Construction Engineering and Management*, vol. 136, no. 12, pp. 1279–1288, 2010.

[18] A. Panas and J. P. Pantouvakis, "Simulation-based and statistical analysis of the learning effect in floating caisson construction operations," *Journal of Construction Engineering and Management*, vol. 140, no. 1, 2013.

[19] R. F. Hamade, M. Y. Jaber, and S. Sikström, "Analyzing cad competence with univariate and multivariate learning curve models," *Computers & Industrial Engineering*, vol. 56, no. 4, pp. 1510–1518, 2009.

[20] P. S. Wong, S. On Cheung, and C. Hardcastle, "Embodying learning effect in performance prediction," *Journal of Construction Engineering and Management*, vol. 133, no. 6, pp. 474–482, 2007.

[21] L. Zhang, X. Zou, and Z. Kan, "Improved strategy for resource allocation in repetitive projects considering the learning effect," *Journal of Construction Engineering and Management*, vol. 140, no. 11, 2014.

[22] H. R. Thomas, "Construction learning curves," *Practice Periodical on Structural Design and Construction*, vol. 14, no. 1, pp. 14–20, 2009.

[23] L. Malyusz and A. Pem, "Predicting future performance by learning curves," *Procedia-Social and Behavioral Sciences*, vol. 119, pp. 368–376, 2014.

[24] H. R. Thomas, C. T. Mathews, and J. G. Ward, "Learning curve models of construction productivity," *Journal of Construction Engineering and Management*, vol. 112, pp. 245–258, 1986.

[25] S. C. Gottlieb and K. Haugbølle, "The repetition effect in building and construction works: A literature review," SBI forlag, Tech. Rep., 2010.

[26] M. J. Anzanello and F. S. Fogliatto, "Learning curve models and applications: Literature review and research directions," *International Journal of Industrial Ergonomics*, vol. 41, no. 5, pp. 573–583, 2011.

[27] A. M. Hijazi, S. M. AbouRizk, and D. W. Halpin, "Modeling and simulating learning development in construction," *Journal of Construction Engineering and Management*, vol. 118, no. 4, pp. 685–700, 1992.

[28] S. Globerson and D. Gold, "Statistical attributes of the power learning curve model," *International Journal of Production Research*, vol. 35, no. 3, pp. 699–711, 1997.

[29] J. Vits and L. Gelders, "Performance improvement theory," *International Journal of Production Economics*, vol. 77, no. 3, pp. 285–298, 2002.

[30] A. B. Badiru, "Computational survey of univariate and multivariate learning curve models," *Engineering Management, IEEE Transactions on*, vol. 39, no. 2, pp. 176–188, 1992.

[31] D. A. Nembhard and N. Osothsilp, "Task complexity effects on between-individual learning/forgetting variability," *International Journal of Industrial Ergonomics*, vol. 29, no. 5, pp. 297–306, 2002.

[32] G. Li and S. Rajagopalan, "The impact of quality on learning," *Journal of Operations Management*, vol. 15, no. 3, pp. 181–191, 1997.

[33] G. Knecht, "Costing, technological growth and generalized learning curves," *Operational Research Quarterly*, pp. 487–491, 1974.

[34] S. Kara and B. Kayis, "The effect of the learning process in concurrent engineering projects," *Concurrent Engineering*, vol. 13, no. 3, pp. 209–217, 2005.

[35] J. G. Everett and S. Farghal, "Learning curve predictors for construction field operations," *Journal of Construction Engineering and Management*, vol. 120, no. 3, pp. 603–616, 1994.

[36] J. E. Mazur and R. Hastie, "Learning as accumulation: A reexamination of the learning curve." *Psychological Bulletin*, vol. 85, no. 6, p. 1256, 1978.

[37] M. Uzumeri and D. Nembhard, "A population of learners: A new way to measure organizational learning," *Journal of Operations Management*, vol. 16, no. 5, pp. 515–528, 1998.

[38] P. S. Adler and K. B. Clark, "Behind the learning curve: A sketch of the learning process," *Management Science*, vol. 37, no. 3, pp. 267–281, 1991.

[39] S. H. Farghal and J. G. Everett, "Learning curves: Accuracy in predicting future performance," *Journal of Construction Engineering and Management*, vol. 123, no. 1, pp. 41–45, 1997.

[40] J. G. Everett and S. H. Farghal, "Data representation for predicting performance with learning curves," *Journal of Construction Engineering and Management*, vol. 123, no. 1, pp. 46–52, 1997.
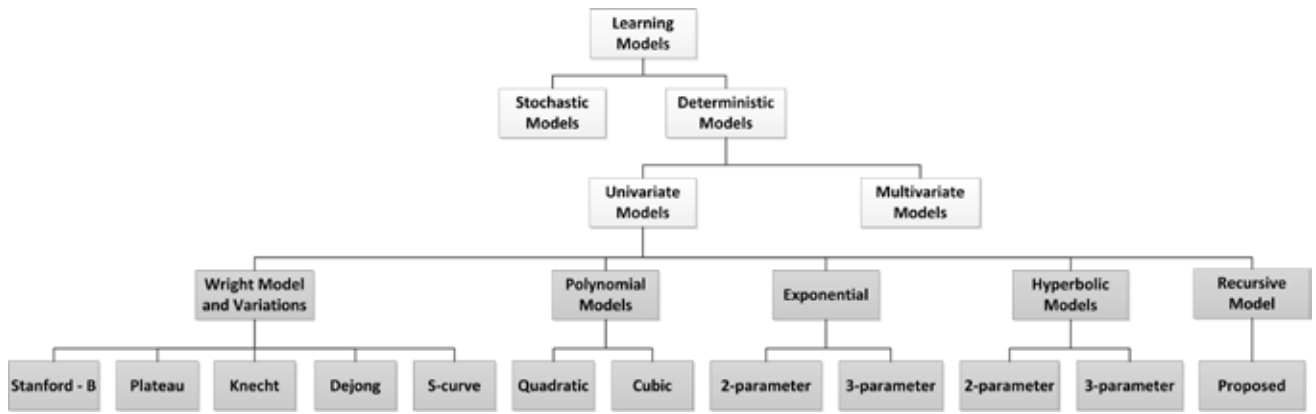
Fig. 1. Learning Models

Model	Case 1, Parameters	Case 2, Parameters	Case 3, Parameters	Case 4, Parameters
Wright, $y = Ax^{-n}$	$A = 10.95$, $n = 0.17$	$A = 26.74$, $n = 0.39$	$A = 100.31$, $n = 0.14$	$A = 221.91$, $n = 0.0074$
Quadratic, $y = A + \beta_0 x + \beta_1 x^2$	$A = 10.71$, $\beta_0 = -0.48$, $\beta_1 = 0.01$	$A = 29.25$, $\beta_0 = -4.35$, $\beta_1 = 0.26$	$A = 113.67$, $\beta_0 = -14.46$, $\beta_1 = 1.62$	$A = 221.90$, $\beta_0 = -0.52$, $\beta_1 = 0.01$
Exponential, $y = Ax^{-n} \exp^{cx}$	$A = 10.99$, $n = 0.17$, $c = 0.001$	$A = 26.63$, $n = 0.43$, $c = 0.013$	$A = 99.10$, $n = 0.18$, $c = 0.016$	$A = 221.91$, $n = 0.01$, $c = 0.00$
2-Parameter Hyperbolic, $y = [k(x/x + r)]^{-1}$	$k = 0.15$, $r = 7.89$	$k = 0.16$, $r = 88.81$	$k = 0.01$, $r = 33.84$	$k = 0.005$, $r = 3.66$
Proposed, $y_n = y_{n-1} 2^{-(r_{n-1} + \varepsilon)} + b_0(y_{n-1}H)^{-a} + M$	$y_0 = 10.97$, $r_0 = 1.55$, $\varepsilon = 0.15$, $b_0 = 0$, $M = 6.40$	$y_0 = 27.00$, $r_0 = 1.05$, $\varepsilon = 0.12$, $b_0 = 0$, $M = 8.74$	$y_0 = 100.63$, $r_0 = 0.83$, $\varepsilon = 0.00$, $b_0 = 0$, $M = 34.25$	$y_0 = 221.43$, $r_0 = 0.12$, $\varepsilon = 0.00$, $b_0 = 0$, $M = 16.94$

Fit, R^2	Wright	Quadratic	Exponential	2-Parameter Hyperbolic	Proposed
Case 1	0.998	0.954	0.998	0.898	0.999
Case 2	0.998	0.957	0.9995	0.999	0.997
Case 3	0.821	0.821	0.831	0.725	0.863
Case 4	0.578	0.651	0.578	0.323	0.652
Predictability, MAPE	**Wright**	**Quadratic**	**Exponential**	**2-Parameter Hyperbolic**	**Proposed**
Case 1	4.56%	58.44%	2.98%	7.53%	0.57%
Case 2	9.59%	146.4%	2.80%	4.47%	0.55%
Case 3	2.23%	36.54%	2.08%	3.48%	2.99%
Case 4	0.54%	0.41%	0.54%	0.94%	0.51%
Paired T-Test, p-value; H0: $\mu_{x_1 - x_2} = 0$	**Wright**	**Quadratic**	**Exponential**	**2-Parameter Hyperbolic**	**Proposed**
Case 1, n = 40	0.00	0.00	0.00	0.00	0.87
Case 2, n = 19	0.00	0.01	0.00	0.00	0.88
Case 3, n = 10	0.14	0.06	0.31	0.13	0.15
Case 4, n = 26	0.12	0.13	0.12	0.00	0.97

Methods to Evaluate and Reduce Data Overload in Electronic Medical Records

Maher Al Ghalayini

Department of Industrial Engineering and Management

American University of Beirut

Bliss Street, Beirut – Lebanon

mbg04@aub.edu.lb

Abstract: **Electronic medical records (EMRs) provide physicians with several important functions that help improve hospital operations. However, the introduction of EMRs has also resulted in unforeseen delays in operations and new types of medical errors. These problems have been found to largely stem from the poor design of the EMR interface, and in particular from the issue of data overload. Nevertheless, it is still unclear how best to design and evaluate an EMR in order to minimize data overload and optimize physician performance. To this end, the goal of this research study is to compare the benefits and limitations of a range of interface evaluation methods for their ability to reflect the performance of physicians using an EMR. In addition, these metrics should be able to 1) pinpoint data overload in EMRs and 2) identify the EMR that best suits the goals and needs of physicians. The application domain for this research will be the AUB Department of Family Medicine. This research will contribute to the literature on interface evaluation and will help usability professionals develop EMRs that maximize safety and efficiency.**

I. Introduction

Data overload, or the presence of high density of data in an interface, is a problem that negatively affects operators in a variety of domains, such as aviation [1], driving [2], medicine [3], graphic design [4], and cartography [5]. The concern about data overload stems from its negative effects on attention and performance [6], namely slowing down visual search [7]. In particular, data overload in the medical domain can have detrimental consequences on patient safety and well-being [3]. The most prominent example of data overload in medicine is in the Electronic Medical Records (EMRs) – digital medical charts – used by physicians [8].

However, to date, researchers have not determined the best way of detecting and measuring data overload in the medical domain. Several methods and metrics have been proposed in the literature, such as subjective assessment [9], [10] and image processing [5], [11], yet it is still unclear which of these measures is best suited to the medical domain. On the other hand, there are alternative metrics whose potential for measuring data overload has not yet been assessed, such as the number of mouse clicks during an operation.

Thus the overall goal of this research is to systematically compare the benefits and limitations of different metrics to measure data overload in EMRs; these metrics will be evaluated based on their ability to reflect physician performance while using EMRs, pinpoint usability problems in EMRs, and identify the EMR that best suits the goals and needs of physicians. The application domain for this research will be the AUB Medical Center (AUBMC) Department of Family Medicine (FM), who have developed an in-house EMR that they use

for patient assessment. The specific aims of this research will then be to 1) outline all the functionalities of the EMR in the FM Department, 2) understand the user needs of physicians in the FM Department with respect to the EMR; and 3) conduct a controlled experiment to gather and evaluate the different suggested metrics in order to identify the most appropriate ones for evaluating data overload.

This work is expected to result in improved guidelines for the evaluation and assessment of data overload in EMRs. In turn, this will allow for increased efficiency and, more importantly, safety in the AUB FM department.

II. Background

A. EMRs

EMRs are medical information technology systems that are used to store medical information, physicians' notes, reports, test results, family medical history, and past medication and treatment [12]. EMRs thus can help make medical operations faster and more efficient, which benefits medical providers and patients alike.

However, current EMRs have failed to fully support this goal. A recent study estimated the number of deaths related to preventable medical errors at 440,000 per year [18], up from the previous estimate of 98,000 [19]. Increasingly, there is a growing realization that errors involving EMRs are not due to a lack of training or negligence on the part of the doctors, but rather stem from poor EMR display design [22]. Design-related issues can thus present a considerable threat to the efficient operation of hospitals and, more importantly, to the health and well-being of patients. The term poor usability (i.e., ease of use of the system) is largely considered the result of poor display design. Several factors can contribute to the poor usability of EMRs, such as poor display feedback [24] and faulty input mechanisms [25], as well as the problem of data overload [12], which will be the main focus of this research.

B. Data Overload

Data overload can be defined as the presence of a high density of poorly-organized data that leads to negative performance decrements [29]. Also referred to as display clutter, data overload is a problem that affects users in many complex domains, such as aviation [1] and medicine [10]. In the medical domain, data overload has figured most prominently in the EMRs used by physicians [3], [8], [22], [30]. EMR data overload could take on many forms, such as a high density of alerts [3] or simply a large amount of irrelevant text and medical information that is poorly organized [31]–[35].

C. Negative Effects of Data Overload

The concern about data overload stems from the reported negative effects it has on performance [36]–[38]. Data overload has been proved to degrade the detection of the change in an interface [39], increase memory load [40], lead to confusion [41] instill confidence in wrong judgments [42], delay visual search [5], [43]–[45], negatively affect the reading and object recognition in the interface [36], and decrease situation awareness [46]. Other negative effects of data overload include committing errors when extracting information [34], higher likelihood of missing data [3] and an inability to obtain the "big picture" of the patient's medical condition [47]. Data overload can also overwhelm one's cognitive capacity [48] and increase mental workload[49].

When it comes to the medical domain, and in particular for EMRs, data overload has been shown to prevent physicians from quickly and accurately extracting EMR information, which can compromise both efficiency and safety in the hospital [3], [29], [50]. For example, research has shown data overload to lead to increased EMR use time and difficulty finding relevant information [44], [45]. In addition, EMR data overload can lead to errors when extracting information [34]. Moreover, data overload in its different forms have led to physicians reporting higher workload ratings [48] and difficulty identifying information among noise [30].

D. Methods for Assessing Data Overload

Despite the negative effects of data overload in EMRs, it is still unclear how best to evaluate data overload in a given display or set of displays. Several different approaches have been used, the most prominent of which rely on subjective input and feedback from physicians. In one study, for example, a survey was used to determine whether practitioners thought there were too many medical alerts in the EMR [3]. In other cases, interviews with physicians were carried out to obtain their insight on data overload [30]. In addition to the more prominent subjective methods, researchers have also occasionally used other techniques. Image processing algorithms (pixel-based algorithms that provide a measure of display usability) have sometimes been used, such as the redundancy measure of Zhang et al. (2011). Only rarely have researchers gathered performance measures such as response time and error rate as a measure of display usability [34], [44], [45].

Despite this attention to EMR usability, there are significant gaps and limitations in the existing literature. The opinions of users, while certainly beneficial and informative, can often be biased and misguided [51]. Rather, from a user-centered or human factors perspective, what matters most when evaluating or rating interfaces is the performance of users. However, relatively few studies have focused on performance measures. At the same time, performance measures alone can only suggest whether a display is potentially useful, but not where the problems within the display are. Other metrics are needed for that; nevertheless, there has not been, to our knowledge, a systematic evaluation of which other measures and approaches are most suitable for evaluating EMRs. There is a need to determine which measures correlate best to performance and, crucially, can provide additional information about the problems within a display. These measures would then form a framework for the systematic and objective evaluation of EMRs.

Eye Tracking

Eye tracking is a non-invasive, infrared-based technique to trace where a user is looking on a screen [52]. Eye tracking has several advantages over other ways of evaluating data overload. In particular, it is a non-invasive and objective measure [57]. Eye tracking also makes it possible to trace changing information access strategies over time at a fine-grained level of analysis [58]. [38] were the first to use eye tracking to systematically evaluate the effects of EMR data overload on physician performance while making diagnoses. They suggested a number of eye tracking metrics that proved to be very beneficial, such as convex hull area and spatial density [55]. However, it is not certain if these metrics generalize to all EMR displays. These metrics will thus be further tested in this research.

Image Processing Techniques

Image processing techniques that measure data overload based on pixel characteristics have also been used to measure data overload [5], [11], [61], [65]. In some approaches, the algorithm produces a scalar value that represents the level of data [11]. Other algorithms create what is called a data overload map by outputting a number of scalar values that represents the data overload in different locations in the display [66], [67]. In addition to that, some researchers have used the color-cluster metric that relies on the definition of data overload as a function of color variation [61]. One of the most cited techniques proposed the feature congestion, sub-band entropy and edge density metrics [5]. This approach will be adopted in this research given its widespread use and validation [68]–[70].

GOMS

Goals, Operators, Methods, and Selection Rules (GOMS) decomposes the task into goals and sub-goals [72] to determine the amount of time needed to complete a task [73]. Goals are the tasks to be completed, operators are the different actions taken to accomplish the goals, and methods are the operators' sequences that should be followed. Finally, selection rules are responsible for choosing the best method to accomplish the goal [73].

III. Methods

This study will be divided into three main phases, which will address the three specific objectives: 1) outline all the functionalities of the EMR in the FM Department, 2) un-

derstand the user needs of physicians in the FM Department with respect to the EMR; and 3) conduct a controlled experiment to gather and evaluate the different suggested metrics in order to identify the most appropriate ones for evaluating data overload. The three phases are explained below.

Phase I

Phase I will tackle the first objective of this research, which is to outline all the functionalities of the EMR. This phase is already underway. As a first step, I have had several meetings with Dr. Joumana Antoun, an Assistant Professor of Clinical Family Medicine, to learn more about the EMR system and how it is used by physicians (see Figure 1 for an example of a page in the EMR). Dr. Antoun explained about the

different tasks that physicians perform as part of their work, and how the EMR supports these tasks. Dr. Antoun also showed and illustrated the use of another version of the EMR that she has developed. This version has not been tested yet. It will be adjusted as part of this study and compared to the original design as a means of assessing the proposed metrics.

Following from my meetings with Dr. Antoun, I conducted a cognitive task analysis (CTA) of the EMR. This is an approach that involves breaking down the use of a system into a series of tasks, and then breaking down each task into its constituent steps. An example of one part of this CTA (for one task) can be seen in Figure 2.

Phase II

In Phase II, the main idea will be to understand the needs of physicians. This include understanding 1) the typical workflow of FM physicians, 2) what specific tasks physicians perform using their EMR as part of this workflow, and 3) exactly what steps are undertaken using the EMR to achieve these tasks. In order to obtain this information, it is necessary to gather input from several FM physicians to build on what was obtained from Dr. Antoun.

Figure 1 EMR Interface for the Progress Notes display

Figure 2 Cognitive Task Analysis for the first task

The participants of this phase will be 20 physicians, residents and experts in the AUB FM Department. Participants for this study will be recruited through email. An email will be sent to all physicians in the FM Department explaining the significance of the research and what the study will entail. Physicians will be asked to respond to the email if they would like to participate in this study and the PI will then get in touch to schedule a session.

I will provide physicians with screen-recording software and ask them to record their screens for one patient session, which typically runs for around four hours. Physicians will be asked to exclude patient names from their screen recordings. Immediately prior to the patient session, I will briefly meet with the physician to provide final instructions, solicit any questions/concerns, and provide the consent form for the physician to sign. Then, immediately following the completion of the patient session, I will meet with the physicians for one hour. The physician will be asked to replay the screen recording and explain the different tasks and objectives at each step of the process. The recorded video will be destroyed at the end of this meeting.

The data that will be collected from this phase will be: the tasks that physicians performed during each patient encounter, what parts of the EMR were used for each task, the sequence in which these tasks were carried out, and the approximate amount of time that physicians took to complete the above discussed five tasks.

At the end of the interview, I will also ask the physicians to tell me what they think should be changed in every display, what areas they think contain data overload, and any other comments that they may have.

The data from these interviews will be used to identify the full list of tasks that can be performed with the EMR, together with all the steps needed or possible to accomplish these tasks. We will also identify the most commonly performed tasks and the pages of the EMR that are most frequently visited. Moreover, the results obtained from this first study will be used to identify five tasks that are used the most or are in significant need of improvement. These tasks will then be tested in Phase II. In addition, the knowledge gained from this phase will be used to improve and adjust the new EMR page developed by Dr. Antoun.

Phase III

In this third phase, a controlled experiment will be conducted with the goal of assessing the proposed data overload evaluation methods. The methods and metrics that will be evaluated are: eye tracking, mouse movements and clicks, image processing algorithms, and GOMS. Of these, only the image processing algorithms and GOMS will be calculated prior to the start of the experiment since they do not require any participant involvement. The GOMS analysis will be done using the CTA that I developed, whereas the algorithms will be done through MATLAB using the freely-available algorithms of [5]. The eye tracking measures and mouse movement clicks, on the other hand, in addition to the performance measures of response time and error rate will all be collected during the experiment.

The location of this experiment will be the Cognitive Psychology and Human Computer Interaction Lab of the Psychology Department (Jesup Hall, Room 107), which contains a Tobii X.2 eye tracker.

Participants. The participants of this phase will also be 20 physicians from the FM Department at AUBMC. They will be recruited via email as in Phase II. In this phase, participants will be awarded a $25 gift card from a restaurant in the Hamra area.

Experiment Design. The independent variables will be the type of task (total of 5 different tasks, as determined from Phase 1) and the amount of data overload (low vs. high). Low data presence refers to the case where only a few data are present on the display and the required information is easily found and extracted from the display. On the other hand, high data levels refer to displays that contain many information present with many distractors and the information extraction process will be harder and will take more time. The new interface developed by Dr. Antoun will play the role of the high data display, whereas the original system will be the one with low data.

In this phase, participants will be asked to complete the set of five tasks listed below.

- Task 1: Read Previous Notes:
- Task 2: Review Patient Disease History
- Task 3: Check if preventive services were needed

- Task 4: Order Lab Reports:
- Task 5: Check Vaccination

For this experiment, participants will have to perform a set of 40 scenarios that represent typical patient cases in the FM Department. For each EMR display, there will be four trials of each task (i.e., each task will be repeated for four different scenarios), for a total of 20 scenarios per EMR display. The same tasks will be repeated in the current and new EMR versions. Each scenario is expected to take a maximum of three minutes. These scenarios and tasks will be developed with the help of Dr. Antoun in order to make sure that they are realistic. The patient names and personal information will be fictitious but the actual patient data will be real data obtained from the Co-PI (the data will be devoid of any names or identifying information). The order of presentation of the tasks will be counterbalanced.

The dependent measures in this phase will be response time and error rate for each task, the number of mouse clicks recorded while performing the different tasks, and eye tracking metrics

Experiment procedure. The researcher will go over the instructions for the experiment with the participants in detail. The researcher will also show the participant the new EMR page and explain how to use it. After this step, the eye tracker will be set up and calibrated, which will involve asking participants to look at a set of points on the screen in order. This first part of the experiment is expected to take around 20 minutes. Next, the eye tracker will be calibrated according, after which the experiment can start when all the tasks on one interface are done, they will be given a 5-minute break and will then continue the experiment with the other EMR interface.

Data Analysis and Results. Repeated-measures ANOVA will be applied to the output of the experiments to determine if there are any significant difference in the results when varying the independent variables, data levels (low, high) and tasks. In addition to that, we will correlate the four metrics with the performance results and use that to develop recommendations for evaluating EMR displays.

IV. Conclusion

This research will add to the human factors, usability, and design literature. A systematic evaluation of these different data overload evaluation metrics has not yet been done, and this research will thus provide an overview of the benefits and limitations of the different approaches.

This research is also expected to help usability professionals select the most appropriate metrics and methods to analyze EMRs. Finally, this research is expected to generalize to other displays in complex domains, such as aviation and process control. The use of the metrics suggested in this research will help prevent the display usability issues, such as data

overload, that have been known to compromise safety and efficiency in complex systems.

ACKNOWLEDGMENT

REFERENCES

[1] A. L. Alexander, E. M. Stelzer, S.-H. Kim, and D. B. Kaber, "Bottom-up and Top-down Contributors to Pilot Perceptions of Display Clutter in Advanced Flight Deck Technologies," Proc. Hum. Factors Ergon. Soc. Annu. Meet., vol. 52, no. 18, pp. 1180–1184, 2008.

[2] K. Yang, S. Member, E. Y. Du, S. Member, E. J. Delp, P. Jiang, and Y. Chen, "A New Approach of Visual Clutter Analysis for Pedestrian Detection," 16th Int. IEEE Conf. Intell. Transp. Syst. (ITSC 2013), vol. 1, no. Itsc, pp. 1173–1178, 2013.

[3] H. Singh, C. Spitzmueller, N. J. Petersen, M. K. Sawhney, and D. F. Sittig, "Information Overload and Missed Test results in Electronic Health record-based Settings," Jama, vol. 167, no. 2, pp. 187–188, 2013.

[4] M. Grahame, J. Laberge, and C. T. Scialfa, "Age Differences in Search of Web Pages: The Effects of Link Size, Link Number, and Clutter," Hum. Factors, vol. 46, no. 3, pp. 385–98, 2004.

[5] R. Rosenholtz, Y. Li, and L. Nakano, "Measuring visual clutter," J. Vis., vol. 7, no. 2, pp. 17.1–22, 2007.

[6] C. D. Wickens and V. Schons, "Visual Separation and Information Access in Aircraft Display Layout," 1993.

[7] M. R. Beck, M. C. Lohrenz, and J. G. Trafton, "Measuring search efficiency in complex visual search tasks: global and local clutter.," J. Exp. Psychol. Appl., vol. 16, no. 3, pp. 238–250, 2010.

[8] H. L. Farley, K. M. Baumlin, A. G. Hamedani, D. S. Cheung, M. R. Edwards, D. C. Fuller, N. Genes, R. T. Griffey, J. J. Kelly, J. C. McClay, J. Nielson, M. P. Phelan, J. S. Shapiro, S. Stone-Griffith, and J. M. Pines, "Quality and safety implications of emergency department information systems," Ann. Emerg. Med., vol. 62, no. 4, pp. 399–407, 2013.

[9] J. M. Schraagen, S. F. Chipman, and V. L. Shalin, "Cognitive Task Analysis," p. 546, 2000.

[10] M. S. Kim, J. S. Shapiro, N. Genes, M. V. Aguilar, D. Mohrer, K. Baumlin, and J. L. Belden, "A Pilot Study on Usability Analysis of Emergency Department Information System by Nurses," Appl. Clin. Inform., vol. 3, no. 1, pp. 135–153, 2012.

[11] R. Van Den Berg, F. W. Cornelissen, and J. B. T. M. Roerdink, "A crowding model of visual clutter," vol. 9, no. 2009, pp. 1–11, 2009.

[12] S. Zakaria and M. K. A. Ghani, "The Impact of EMR User Interface Design on Doctor Satisfaction," e-Proceeding Softw. Eng. Postgraduates Work., p. 94, 2013.

[13] D. A. Ludwick and J. Doucette, "Adopting electronic medical records in primary care: Lessons learned from health information systems implementation experience in seven countries," Int. J. Med. Inform., vol. 78, no. 1, pp. 22–31, 2009.

[14] K. Häyrinen, K. Saranto, and P. Nykänen, "Definition, structure, content, use and impacts of electronic health records: A review of the research literature," Int. J. Med. Inform., vol. 77, no. 5, pp. 291–304, 2008.

[15] N. K. Pera, A. Kaur, and Raveendra rao, "Perception of electronic medical records (EMRs) by nursing staff in a teaching hospital in India Pera NK, Kaur A, Rao R - Int J Adv Med Health Res," vol. 1, no. 2, pp. 75–80, 2014.

[16] J. Heer and M. Agrawala, "Software design patterns for information visualization," IEEE Trans. Vis. Comput. Graph., vol. 12, no. 5, pp. 853–860, 2006.

[17] M. S. Arshad Malik and S. Sulaiman, "Towards the development of an interface model for information visualization in multiple electronic health records," 2013 Int. Conf. Comput. Med. Appl., no. December, pp. 1–5, 2013.

[18] J. T. James, "A new, evidence-based estimate of patient harms associated with hospital care.," J. Patient Saf., vol. 9, no. 3, pp. 122–8, 2013.

[19] L. Kohm, J. Corrigan, and M. Donaldson, "To err is Human: Building a Safer Health Care System," Inst. Med., pp. 26–48, 2000.

[20] Pennsylvania Patient safety authority, "Pennsylvania Patient Safety Authority 2012 annual report," 2012.

[21] "Institute of Medicine," Comm. Patient Saf. Heal. Inf. Technol. Heal. IT patient Saf. Build. safer Syst. better care, pp. 1–8, 2012.

[22] B. T. Karsh, M. B. Weinger, P. A. Abbott, and R. L. Wears, "Health information technology: fallacies and sober realities," J Am Med Inf. Assoc, vol. 17, no. 6, pp. 617–623, 2010.

[23] Maya, "Missing the Point in the Design of Electronic Medical Records," 2009. [Online]. Available: http://maya.com/blog/missing-the-point-in-the-design-of-electronic-medical-records.

[24] U. B. Tasa, O. Ozcan, A. E. Yantac, and A. Unluer, "A case study on better iconographic design in electronic medical records' user interface.," Inform. Health Soc. Care, vol. 33, no. 2, pp. 125–38, 2008.

[25] M. A. Clarke, J. L. Belden, and M. S. Kim, "What Learnability Issues Do Primary Care Physicians Experience When Doing CPOE," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 9171, pp. 373–383, 2015.

[26] J. J. Saleem, E. S. Patterson, L. Militello, S. M. Asch, B. N. Doebbeling, and M. L. Render, "Using human factors methods to design a new interface for an electronic medical record," AMIA ...Annual Symp. Proc. / AMIA Symp. Symp., no. Journal Article, pp. 640–644, 2007.

[27] D. Erol Barkana, A. Açık, D. G. Duru, and A. D. Duru, "Improvement of design of a surgical interface using an eye tracking device.," Theor. Biol. Med. Model., vol. 11 Suppl 1, no. SUPPL.1, p. S4, 2014.

[28] C. M. Johnson, T. R. Johnson, and J. Zhang, "A user-centered framework for redesigning health care interfaces," J. Biomed. Inform., vol. 38, no. 1, pp. 75–87, 2005.

[29] N. M. Moacdieh and N. B. Sarter, "Eye Tracking Metrics: A Toolbox for Assessing the Effects of Clutter on Attention Allocation," Proc. Hum. Factors Ergon. Soc. Annu. Meet., vol. 56, no. 1, pp. 1366–1370, 2012.

[30] T. T. Van Vleck, D. M. Stein, P. D. Stetson, and S. B. Johnson, "Assessing data relevance for automated generation of a clinical summary.," AMIA Annu. Symp. Proc., pp. 761–5, 2007.

[31] F. Bobillo, M. Delgado, and J. G??mez-Romero, "Representation of context-dependant knowledge in ontologies: A model and an application," Expert Syst. Appl., vol. 35, no. 4, pp. 1899–1908, 2008.

[32] K. W. Hammond, E. N. Efthimiadis, and R. J. Laundry, "Efficient de-identification of electronic patient records for user cognitive testing," Proc. Annu. Hawaii Int. Conf. Syst. Sci., pp. 2771–2778, 2011.

[33] C. R. Weir, J. J. R. Nebeker, B. L. Hicken, R. Campo, F. Drews, and B. LeBar, "A Cognitive Task Analysis of Information Management Strategies in a Computerized Provider Order Entry Environment," J. Am. Med. Informatics Assoc., vol. 14, no. 1, pp. 65–75, Jan. 2007.

[34] Q. Zeng, J. J. Cimino, and K. H. Zou, "Providing Concept-oriented Views for Clinical Data Using a Knowledge-based System: An Evaluation," J. Am. Med. Informatics Assoc., vol. 9, no. 3, pp. 294–305, 2002.

[35] R. Zhang, S. Pakhomov, B. T. McInnes, and G. B. Melton, "Evaluating measures of redundancy in clinical texts.," AMIA Annu. Symp. Proc., vol. 2011, pp. 1612–20, 2011.

[36] M. J. Bravo and H. Farid, "Object recognition in dense clutter.," Percept. Psychophys., vol. 68, no. 6, pp. 911–918, 2006.

[37] M. Yeh, J. L. Merlo, C. D. Wickens, and D. L. Brandenburg, "Head up versus head down: the costs of imprecision, unreliability, and visual clutter on cue effectiveness for display signaling.," Hum. Factors, vol. 45, no. 3, pp. 390–407, 2003.

[38] N. Moacdieh and N. Sarter, "Clutter in Electronic Medical Records: Examining Its Performance and Attentional Costs Using Eye Tracking," Hum. Factors J. Hum. Factors Ergon. Soc., vol. 57, no. 4, pp. 591–606, 2015.

[39] N. Moacdieh and N. Sarter, "Display Clutter: A Review of Definitions and Measurement Techniques," Hum. Factors J. Hum. Factors Ergon. Soc., p. 0018720814541145–, 2014.

[40] H. Westerbeek and A. Maes, "Referential scope and visual clutter in navigation tasks," Proc. PRE-CogSci 2011, 2004.

[41] G. J. Ewing, C. J. Woodruff, and D. Vickers, "Effects of 'local' clutter on human target detection," Spat. Vis., vol. 19, no. 1, pp. 37–60, 2006.

[42] S. Baldassi, N. Megna, and D. C. Burr, "Visual Clutter Causes High-Magnitude Errors," PLoS Biol., vol. 4, no. 3, p. e56, 2006.

[43] J. M. Henderson and T. J. Smith, "The influence of clutter on real-world scene search: Evidence from search efficiency and eye movements," vol. 9, no. 2009, pp. 1–8, 2009.

[44] G. Duftschmid, C. Rinner, M. Kohler, G. Huebner-Bloder, S. Saboor, and E. Ammenwerth, "The EHR-ARCHE project: Satisfying clinical information needs in a Shared Electronic Health Record System based on IHE XDS and Archetypes," Int. J. Med. Inform., vol. 82, no. 12, pp. 1195–1207, 2013.

[45] D. R. Murphy, B. Reis, D. F. Sittig, and H. Singh, "Notifications Received by Primary Care Practitioners in Electronic Health Records: A Taxonomy and Time Analysis," Am. J. Med., vol. 125, no. 2, pp. 209.e1–209.e7, 2012.

[46] S. Kim and D. B. Kaber, "Assessing the Effects of Conformal Terrain Features in Advanced Head-Up Displays on Pilot Performance," Proc. Hum. Factors Ergon. Soc. 53rd Annu. Meet., pp. 36–40, 2009.

[47] M. P. Tully, Å. Kettis, A. T. Höglund, C. Mörlin, Å. Schwan, and C. Ljungberg, "Transfer of data or re-creation of knowledge – Experiences of a shared electronic patient medical records system," Res. Soc. Adm. Pharm., vol. 9, no. 6, pp. 965–974, 2013.

[48] A. Ahmed, S. Chandra, V. Herasevich, O. Gajic, and B. W. Pickering, "The effect of two different electronic health record user interfaces on intensive care provider task load, errors of cognition, and performance*," Crit. Care Med., vol. 39, no. 7, pp. 1626–1634, 2011.

[49] S. Yang, K. Shukla, and T. K. Ferris, "'Cognitive Efficiency' in Display Media: A First Investigation of Basic Signal Dimensions," Proc. Hum. FACTORS Ergon. Soc. 56th Annu. Meet. - 2012, pp. 1371–1375, 2012.

[50] L. Wu, Z. Zhu, H. Cao, and B. Li, "Influence of information overload on operator's user experience of human–machine interface in LED manufacturing systems," Cogn. Technol. Work, 2015.

[51] A. D. Andre and C. D. Wickens, "When Users Want What's not Best for Them," Ergon. Des. Q. Hum. Factors Appl., vol. 3, no. 4, pp. 10–14, 1995.

[52] A. Poole and L. J. Ball, "Eye Tracking in HCI and Usability Research," in Encyclopedia of Human Computer Interaction, 2006, pp. 211–219.

[53] S. M. Munn, L. Stefano, and J. B. Pelz, "Fixation-identification in dynamic scenes," Proc. 5th Symp. Appl. Percept. Graph. Vis. - APGV '08, vol. 1, no. 212, p. 9, 2008.

[54] J. M. Findaly, "Eye scanning and visual search," in The interface of language, vision and action: Eye movements in the visual world, 2004, pp. 135–159.

[55] J. Goldberg and X. Kotval, "Computer Interface Evaluation Using Eye Movements: Methods and Construct," Int. J. Ind. Ergon., vol. 24, no. 6, pp. 631–645, 1999.

[56] F. Di Nocera, M. Camilli, and M. Terenzi, "A Random Glance at the Flight Deck: Pilots' Scanning Strategies and the Real-Time Assessment of Mental Workload," J. Cogn. Eng. Decis. Mak., vol. 1, no. 3, pp. 271–285, 2007.

[57] K. K. E. Ellis, "Eye tracking metrics for workload estimation in flight deck operations," ProQuest Diss. Theses, vol. 1467693, p. 115, 2009.

[58] G. J. Zelinsky and D. L. Sheinberg, "Eye movements during parallel-serial visual search.," J. Exp. Psychol. Hum. Percept. Perform., vol. 23, no. 1, pp. 244–262, 1997.

[59] J. Sapp, S. Jessee, J. Crutcher, and M. C. King, "Human Factors Engineering Evaluation of the CH-47F Horizontal Situation Display Hover," no. January, 2012.

[60] J. H. Goldberg, M. J. Stimson, M. Lewenstein, N. Scott, A. M. Wichansky, a C. M. S. I. G. on C. G. and I. Techniques, a C. M. S. I. G. on C.-H. Interaction, and A. for C. Machinery, "Eye Tracking in Web Search Tasks: Design Implications," Eye Track. Res. Appl. Symp., no. 650, pp. 51–58, 2002.

[61] M. C. Lohrenz, J. G. Trafton, R. M. Beck, and M. L. Gendron, "A model of clutter for complex, multivariate geospatial displays.," Hum. Factors, vol. 51, no. 1, pp. 90–101, 2009.

[62] M. J. Bravo and H. Farid, "A scale invariant measure of clutter.," J. Vis., vol. 8, no. 2008, pp. 23.1–9, 2008.

[63] M. R. Beck, M. Trenchard, A. Van Lamsweerde, R. R. Goldstein, and M. Lohrenz, "Searching in clutter: Visual attention strategies of expert pilots.," pp. 1411–1415, 2012.

[64] A. J. DeWitt and J. Kuljis, "Aligning usability and security: a usability study of Polaris," SOUPS '06 Proc. Second Symp. Usable Priv. Secur., pp. 1–7, 2006.

[65] H. Chang, J. Zhang, X. Liu, C. Yang, and Q. Li, "Color Image Clutter Metrics For Predicting Human Target Acquisition Performance," pp. 1–4, 2010.

[66] M. Jansen and M. Van Kreveld, "Evaluating the consistency of cartographic generalization," Proc. 8th Int. Symp. Spat. Data Handl., no. 21957, pp. 668–678, 1998.

[67] D. Mušicki, S. Suvorova, M. Morelande, and B. Moran, "Clutter map and target tracking," 2005 7th Int. Conf. Inf. Fusion, FUSION, vol. 1, pp. 69–76, 2005.

[68] E. J. Pereira and M. S. Castelhano, "Peripheral Guidance in Scenes: The Interaction of Scene Context and Object Content.," J. Exp. Psychol. Hum. Percept. Perform., vol. 40, no. 5, p. Advance online publication., 2014.

[69] L. Holy, K. Jezek, J. Snajberk, and P. Brada, "Lowering visual clutter in large component diagrams," Proc. Int. Conf. Inf. Vis., pp. 36–41, 2012.

[70] A. Miniukovich and A. De Angeli, "Quantification of interface visual complexity," Proc. 2014 Int. Work. Conf. Adv. Vis. interfaces, pp. 153–160, 2014.

[71] M. F. Asher, D. J. Tolhurst, T. Troscianko, and I. D. Gilchrist, "Regional effects of clutter on human target detection performance.," J. Vis., vol. 13, no. 5, pp. 1–15, 2013.

[72] A. B. L. Berry, K. A. Butler, C. Harrington, M. O. Braxton, A. J. Walker, N. Pete, T. Johnson, M. W. Oberle, J. Haselkorn, W. P. Nichol, and M. Haselkorn, "Using conceptual work products of health care to design health {IT}," J. Biomed. Inform., vol. 59, pp. 15–30, 2016.

[73] H. Saitwal, X. Feng, M. Walji, V. Patel, and J. Zhang, "Assessing performance of an Electronic Health Record (EHR) using Cognitive Task Analysis," Int. J. Med. Inform., vol. 79, no. 7, pp. 501–506, 2010.

[74] B. E. John and D. E. Kieras, "Using GOMS for user interface design and evaluation: Which technique?," ACM Trans. Comput. Interact., vol. 3, no. 4, pp. 287–319, 1996.

[75] N. Segall, D. B. Kaber, J. M. Taekman, and M. C. Wright, "A Cognitive Modeling Approach to Decision Support Tool Design for Anesthesia Provider Crisis Management," Int. J. Hum. Comput. Interact., vol. 29, no. 2, pp. 55–66, 2013.

Optimization of Annual Maintenance Schedule in Cement Industry

Diana Halawi Ghosn, Dr Ramzi Fayad

Faculty of Engineering, Industrial Engineering and Engineering Management Department

Beirut Arab University

Debbiyeh Branch - Lebanon

Diana.h.ghosn@gmail.com, r.fayad@bau.edu.lb

Abstract

This paper presents the existing situation of annual maintenance scheduling in cement industry. The search shows that there are strong areas such as defined mission and vision, highly experienced maintenance engineers and workers, external labor are always available and spare parts follow up via computerized system. Other areas are found weak such as lack of planning unit, long time in scheduling, non-flexibility of schedule tasks, dependency on experienced personnel to determine tasks requirements, no respect for the schedule allocated and no follow up on schedule results. The result of this study will be significant to cement industry in minimizing scheduling time and external labor and granting flexible scheduling software for their annual maintenances.

1. Introduction

Maintenance is a set of actions necessary for retaining or restoring a piece of equipment, machine or system to a specified operable condition to achieve its maximum useful life. In cement industry the majority of preventive and routine maintenance is carried out during plant operation; however, major maintenance tasks necessitate shutting down the entire production line. Annual maintenance is extremely complex due the amount of repair works to be fitted in into a short interval of time. Which make the whole process of planning and scheduling difficult to manage.

The purpose of this study is to increase the efficiency of the maintenance scheme applied by assessing and improving the deficiencies in the current annual maintenance planning and scheduling. As well as by optimizing the number of maintenance staff to be contracted. To help achieving this outcome, a scheduling program must be coded to optimize the allotting of maintenance tasks taken into consideration time slots, tasks' priorities and resources required to accomplish the job. The main goal is to efficiently maintain the system while minimizing its maintenance cost.

The study aims to map the current situation of the daily workload for maintenance engineers and planners and find a way to optimize the scheduling which leads to lowering annual maintenance costs. The main objectives are

- What the current applied maintenance scheme that cement plants follow?
- To what extent are cement plants working with annual maintenance tasks prioritization?
- What system or software are the cement plants using at the mean time?
- What are the scheduling software characteristics and specifications?
- How bottlenecks are specified and dealt with?
- How to optimize the resources allocation to tasks?

2. Cement Industry

2.1. Cement Manufacturing Process

[1] Cement manufacturing industry is identified by North American Industry Classification System (NAICS) code 32731 (formerly identified as SIC code 3241).

The procedure of cement making follows:

1- The cement manufacturing process starts by extracting the main raw material from quarry and transported to the plant. This basic raw material in known as limestone.

2- The raw material limestone is combined with clay, pounded in a crusher and stored in silos. Sand, iron and bottom ash are then added to the limestone and clay in a precise calculated quantities and all together grinded into a fine powder in a roller mill.

3- Next, the fine powder is heated as it passes through the Pre-Heater Tower into a large kiln. The kiln is at the heart of the manufacturing process. Once inside the kiln, the raw meal is heated to around 1,500 degrees C - a similar temperature to that of molten lava. Upon exiting the kiln, the clinker is cooled and stored, ready for grinding, to produce cement.

4- The clinker is combined with small amounts of gypsum and limestone and finely ground in a finishing mill. The mill is a large revolving cylinder containing tons of steel balls that is driven by a motor. The finished cement is ground so fine that it can pass through a sieve that will hold water.

5- Finally, the cement is stored in silos before being shipped in bulk or in bags to the sites where it will be used.

2.2. Maintenance in Cement Industry

Khan and Darrab [2] reported that the purpose of maintenance is not only to upkeep the plant machinery and equipment preventing from failure and breakdowns, increasing reliability, maintainability, and availability of the operating

system for maximizing production, but also to improve quality and boost higher productivity through improving capacity, faster and more dependable throughput, reducing inventory, and lowering operating cost.

In cement industries, maintenance cost consumes approximately 20-25 % of the total production system, which comes in second rank after the energy cost [3]. Due to that fact, cement industries are focusing more on the maintenance area in the last few years.

In the past the equipment in cement industry was not so multifaceted while on the other hand more operators where required to keep the line operating. At that time maintenance had no large influence on production. Quantity and quality of the cement was proportional to the skills of the operators and how fast they can react. As technology was improving and increase in automation, more focus was shifted on maintenance and non-planned stoppages became so unfavorable to happen as they could highly affect the productivity. Achieving that goal was not possible without proper planning and scheduling for all resources. "Fig.1" summarizes all types of maintenance [4].
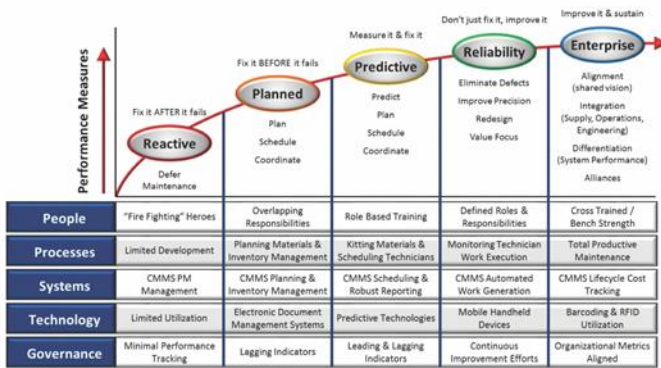

Fig. 1: Maintenance Maturity Continuum

2.3. Maintenance Planning and Scheduling

Planning is a detailed program on how to do the job.

Scheduling is writing the process and specifying when to do the job. It specifies the tasks to be done, all the resources required, durations on which the task can be executed, and the equipment and tools.

Matt Midas [4] in his article Best Practices of Maintenance Planning and Scheduling, 2015 mentions that without planning and scheduling, the wrench-on time for a company is on average only 35%. That means that for every technician working an 8-hour day, only 2.8 hours of actual work on assets is done.

As per Bruce Hawkins and Timothy C. Kister [5], Maintenance planning consists of the following points:

- Maintenance Planning is arranging specific tasks, ahead of time, that require to be done in an efficient and effective manner when the job is required to be implemented in the future.

- Maintenance Planning is a process of detailed analysis to first deter- mine and then to describe the work to be performed, by task sequence and methodology.
- Maintenance Planning provides for the identification of all required resources, including skills, crew size, labor-hours, spare parts and materials, special tools and equipment.

Maintenance scheduling process is allocating selected resources and resource skills to complete a specific job. The purpose of maintenance scheduling is to make sure that specific resources such as material and personnel are available at the same time when the equipment is ready to be worked on.

Scheduling objectives could be summarized as below:

- Minimize overall maintenance cost
- Minimize equipment downtime delays
- Insure that all maintenance needs are prepared (spare parts, tools, personnel…)
- Tasks prioritization with respect to the whole operation
- Perform the work in safely manner
  In order to achieve the objectives of scheduling below prerequisites need to be taken care of:
- Lead time: the time that is allocated for the maintenance shall be well known and set ahead of time so that the tasks can be effectively planned and scheduled

3. Present Situation

The offered information is selected depending on the areas covered in this study but with some complementing areas to help capture the whole picture. The former are data collected from interviews and mails with the employees of cement industries; in addition to internal documents collected.

3.1. Annual Maintenance Procedure

In cement plants, each production line shall undergo at least one total shut down maintenance period throughout the year, this period is referred to the "Annual Maintenance". Annual maintenance is the most expensive of all other maintenance jobs in the cement industries, not only because of the production loss, but also due to the major maintenance being performed. Industry surveys report that between 25 and 52 % of maintenance budgets are expended in individual area or whole plant shutdowns [6]. The complexity of maintenance planning is though higher because of some characteristic that distinguish from other types of scheduling [7].

Annual maintenance is the most complex maintenance period due to:

- Short duration
- Critical equipment maintenance tasks are usually more complex than small repeatable equipment maintenance tasks
- Pressure from trying to plan and schedule as much tasks as possible during the stoppage period
- Huge number of tasks to be executed

- Numerous contractors on site at same time

3.2. Annual Maintenance Planning and Scheduling

In order to understand the present situation regarding how the annual maintenance planning and scheduling take place at the mean time; visits were conducted to several Lebanese cement factories, and e-mails conversations were organized with contacts at several Portuguese cement factories.

3.2.1. Annual Maintenance Planning Process

In all the cement plants that where contacted during this study, it showed that 100% follow the same planning procedure while preparing for their annual maintenance.

Each cement plant has personnel known as "Inspectors" whose daily role is to travel around the plant check over equipment's situation. Inspectors always write notes about the cases they witness and report to their managers. Some cases are critical, need to be acted upon directly while other cases are shifted to the annual maintenance. When the annual maintenance date and time is agreed upon by top management, the maintenance team meets to decide on the tasks they need to implement. In all the plants it was clear that the tasks are being chosen based on the experience of the personnel. Where either the experienced supervisors or / and inspectors are deciding on the following:

- Choosing the tasks (no priorities indicated, while they mention all possible tasks)
- Specifying the required skills for each task
- Specifying the number of personnel required for each task
- Specifying the duration of the tasks

After all tasks are ready, the plants that have a planning department, the tasks are sent to the named department for scheduling, while in the plants that lack this department, usually the head of units arrange their own schedule via Microsoft Projects software.

3.2.2. Annual Maintenance Scheduling Process

At this stage the management had set the duration for the annual maintenance and the maintenance shift timing. The three possible shift timing observed are:

- 2 shifts – 12 hours each (most commonly implemented)
- 2 shifts – 8 hours each
- 3 shifts – 8 hours each

It's time for the scheduler to distribute the tasks over the stoppage period in the most efficient way taking into consideration the available internal personnel and decreasing as much possible the external personnel required. All the contacted plants use the MS Project as main scheduling software for their annual maintenance. All of them expressed the restrictions and the time consuming that the system obliges as mentioned below:

- Several tasks exceeds 12 hours and could be implemented only during day shifts, in this case the task shall be split over many day shifts. This can be only done manually on MS which is a real time consuming and tough job for the planner.

- MS is non-flexible software in cement plants' annual maintenance scheduling. Whenever there is a predecessor link for a task, it is so hard to change the task place among the schedule since this will cause change in all the remaining tasks that follows. These cases are faced a lot due to several reasons.
- Since all the tasks are suggested by inspectors depending on some written notes that have been taken or depending on their memory, some tasks are missed. When these tasks are asked to be entered in the schedule later on, this causes a chaos in the schedule. That's why whenever a task is being missed, most of the time the schedule is not revised and the task is not included in the schedule.

3.3. Strengths and Improvements

The strengths in planning and scheduling found in the cement industries will be mentioned below; in addition to weak points witnessed that has room for improvement.

Strengths
- Defined mission and vision for the plant
- Highly experienced maintenance engineers and maintenance workers which have great knowledge about the processes and equipment
- External contractors are always available to be contracted for the additional tasks
- Software for following up on spare parts
- Trainings for all employees at all levels each in his domain

Weaknesses
- Lack of planning unit
- Dependency on experienced individuals to indicate tasks duration
- Dependency on experienced individuals to indicate skills and number of personnel required for each maintenance task
- Long time for scheduling
- Non flexibility in schedule tasks
- No respect for the schedule allocated
- No task prioritization
- No schedules follow up
- Planners rank in plant hierarchy

4. Improvement Proposals

The proposals have considered the fact that the spare parts for the maintenance are available at premises and there are no restrictions on the number of external maintenance staff and the duration of hiring.

4.1. Establish Maintenance Planning and Scheduling Unit

Without proper planning and scheduling, maintenance, at best is hazard; at worst, it is costly and ineffective. The completeness and complexity of maintenance planning and scheduling improves and increases the level of production as it reduces the halting time. Well-Planned and scheduled tasks have many advantages over non-planned tasks:

- Efficient performance
- Less cost
- Decrease operation stoppages
- Better work quality
- Higher workers' satisfaction

4.2. Facilitate Communication with all Other Units

The maintenance planning and scheduling goals should be circulated among all other units, directly or indirectly related to maintenance department.

During the annual maintenance preparation period, plans and schedules could be improved by having maintenance and production representatives attending single planning meeting where the plan and schedule adjusted in the presence of all parties, which lead into massive time saving in revising the schedule over and over again.

When the schedule is finalized, checked and approved it needs to be communicated into all other maintenance and production units. Team work is highly required among the planning and scheduling unit and other units, therefore to communicate all the time is a critical point for success.

4.3. Tasks' Information Standardization

Usage of tasks' history for launching standard tasks durations, parts, and consumables for repetitive tasks is a must, that's what is referred to by tasks' information standardization. Each plant shall get the most out of its CMMS / ERP system they are using. Moreover, a data base may be built out that includes all the possible tasks that may occur during an annual maintenance including the critical information required by the planners. In order to obtain world-class maintenance, a system for data collection is required, that can handle data related to breakdown frequency, duration of maintenance and so on [8]. It is advantageous to have a system that can record data in real-time which continually updates the information in the system [9].

4.4. Tasks Criticality and Prioritization

Each task to be scheduled depends on the equipment criticality and its priority among other tasks. Annual maintenance by definition it is mainly for the maintenance of critical equipment, while it also includes several other equipment benefiting from the outage. Task priority is a factor for sequencing work planning, where the highest priority tasks are the first to be planed and scheduled. Usually the requested of the task shall indicate the task priority for the planners and schedulers.

Equipment criticality is to rank each equipment in relation to how much it affect the production process. Usually equipment's' criticality is set by the production department as they are responsible for the process in cement plants. Bruce Hawkins and Timothy C. Kister [5] had defined equipment criticality and job priority assignments in their book Miantenance Planning and Scheduling: Streamline your organization for a Lean Environment published on February 2006 as below:

Table 1: Equipment Criticality

Criticality	Description
10	Shuts down entire plant
9	Shuts down more than one line
8	Not spared production equipment / shuts down one line
7	Mobile equipment
6	Spared production equipment / not spared support equipment – product can be made on one or more lines
5	Support equipment spared
4	Infrequently used production equipment
3	Miscellaneous equipment
2	Roads and grounds
1	Buildings and offices

Table 2: Job Priority Assignments

Index	Title	Job Planning
E	Emergency	Un-Plannable
1	Urgent	Un-Plannable
2	Critical	Usually Un-Plannable
3	Rush	At least partially plannable
4	Essential-but deferrable	Plannable
5	Desired	Plannable
6	Shutdown	Plannable
7	Routine	Plannable

4.5. New Scheduling Software

Studying and assessing the currently applied maintenance scheduling software in all 5 factories considered in this study, MS project software is not giving the required results in scheduling the annual maintenance. A scheduling program must be coded to optimize the allotting of maintenance tasks taken into consideration the time slots, tasks priorities and resources required to accomplish the job. The software details mentioned further in section 5.

5. Scheduling Software

First of all it is necessary to acquire complete and accurate reading concerning equipment operation, accidents and their consequences, maintenance operations and their costs, in order to know real operational conditions. So, there is a need for reliability data collection in relation to all types of components from the field records of installations and operations. An interface for entering maintenance tasks has been created as shown in "Fig. 2".
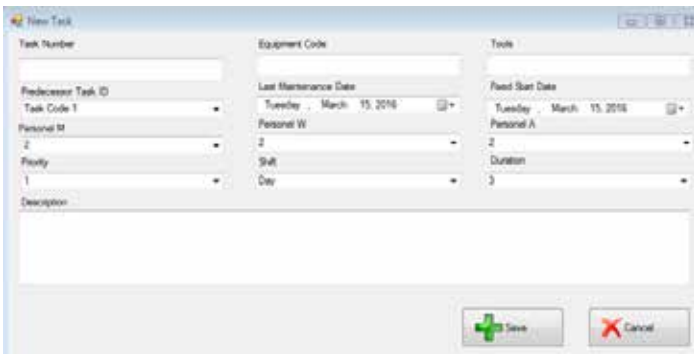
Fig.2: Maintenance Task

The database is built including the following information:
- Equipment number
- Maintenance tasks for each equipment
- Tasks' durations
- Personnel required to implement the task
- Tools required for the task

Tasks duration and personnel required has been calculated from the average duration history collected from the past 5 years data collected at Sibline cement factory in Lebanon.

When the planner receives the list of tasks required for the maintenance, they choose them from the database and forward them to the scheduling interface.

The schedule can have input from 2 sources: either from the database or directly from an excel sheet.

The scheduling interface as shown in "Fig. 3" obliges the planer to enter the following requirements:
- Annual Maintenance period start and end dates
- The shifts required and the hours per shift
- Choose which days are non-working days
- Task priority
- Task shift (day-evening-night or multi-shift)
- Specify if any task has a predecessor
- Specify task date "if the task has a specific date to start with"
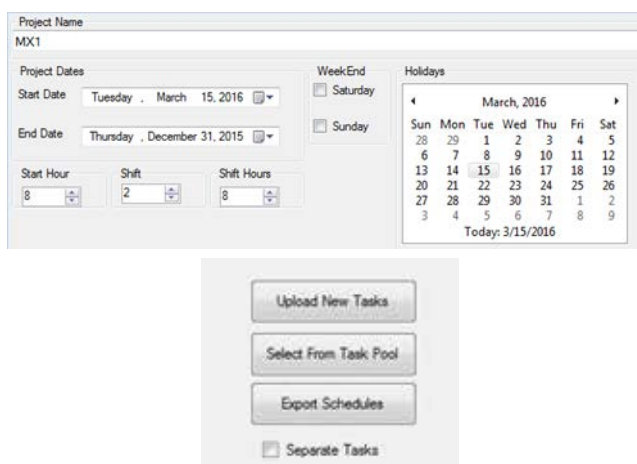
Fig. 3: Schedule Interface

5.1. Scheduling Software Algorithm

The schedule algorithm is transferred from the real manual situation that the planners work in. the steps are shown in below "Fig. 4":
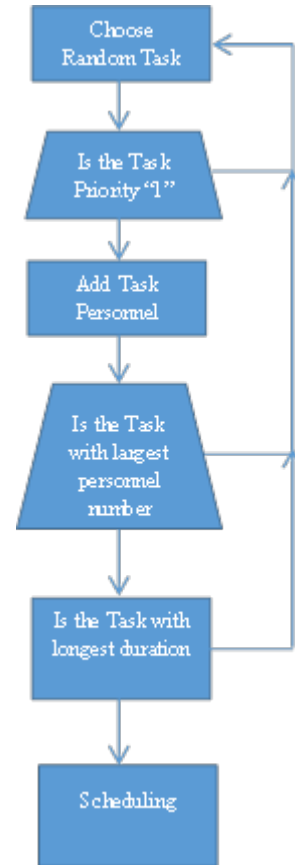

Fig. 4: Scheduling Software Algorithm

Scheduling: Start with the task having maximum number of staff and having the longest duration. This task will be followed by task with same number of staff, same or shorter duration; in condition they both could be performed during same shaft. If it is not possible; the second task will be deferred to the next day or night shift, and the system will search for another task from the database to fit with the remaining time of the same shift.

Another scenario, supposing that we already assigned task at the start of a certain shift, this task doesn't require all member of staff to be involved to perform the task, the requiring staff would be allocated to do task requiring less number of people in parallel time slot. In one condition, the personnel should be able to complete the task. This is all in function of the number of shifts per day, whether 2 shifts of 12 hours each or 3 shifts per day of 8 hours each.

5.2 Data Collection

As mentioned earlier, cement industries are the target of this study. The study included one known cement industry in Lebanon, Sibline Cement, from which the maintenance tasks

data was collected. The data collected are: maintenance areas and their codification, all equipment in the plant and their codification, all maintenance tasks applied at Sibline since 2013, timing of the tasks, personnel type and number for each task, standard tools used per task. In addition to several schedules that where already implemented at Sibline in previous annual maintenances. All attained tasks where inserted in the data base created and transferred into a test project. In addition for testing a complete schedule in order to compare the real life one that was implemented in Sibline with the software results.

5.3 Results and Discussion

Real life annual maintenance project that has been implemented at Sibline cement plant, 10 days project, was used as a test project to analyze the software results. Software schedule output shown in "Fig. 5":

"Fig. 5": Software Schedule Output

Table 3: shows a comparison between the time taken with the planner to prepare the schedule on MS projects manually and the new software scheduling.

Table 3: Time Comparison

	MS Projects	New Software
Time to insert tasks (mints)	30	5
Time to finalize the scheduling (mints)	300	5
Revision 1	180	5
Revision 2	120	5
Total (mints)	630	20
Saved Time (mints)		610

Result 1: Approximately ten hours were gained for the planner during preparing the annual maintenance schedule. Indirect cost saving 10 hrs* 14 $/hr = 140 $.

Result 2: During the following annual maintenance project, Sibline had contracted 35 external staff per day to help over in the job (based on duration and personnel number estimated by the experienced inspectors, in addition to the manual schedule prepared via MS projects). After running the software, the total required external to help was decreased into 27 per day. Calculation below will show the cost reduction that the software had resulted.

35 external * 12 hours = 420 hrs*10.5 $/hr= 4,410 $/day

27 external * 12 hours = 324 hrs*10.5 $/hr= 3,402 $/day

In such a small project, in addition to the simplicity and easy work brought into a planner the software has saved approximately one and a half working day for the planner, despite the frustration they have with each revision that leads to major changes. Moreover, cost wise, the software was able to attain 10,000 $ direct cost saving for the total annual maintenance in just 10 days.

6. Conclusion

Throughout this work, it has been shown that the problem is in using the MS Project as main scheduling software in addition for the dependability on the experienced staff estimation of duration and staff required for a task. The software proposed was evaluated through case studies and proven to give better results from both the planners using the software and annual maintenance cost.

REFERENCES

[1] Hani Shafeek, "Maintenance Practices in Cement Industry," Asian Transactions on Engineering (ATE ISSN: 2221-4267) Volume 01 Issue 06

[2] Khan, M.R.R. and Darrab, I.A., "Development of analytical relation between maintenance, quality and productivity," Journal of Quality in Maintenance Engineering, 16: 341–353, 2010

[3] Sumit Sharma

[4] MATT MIDAS, "Best Practices of Maintenance Planning and Scheduling," Inspectioneering Journal, April 2015

[5] Bruce Hawkins, Timothy C. Kister, "Maintenance Planning and Scheduling Handbook Streamline Your Organization for a Lean Environment" ISBN 0750678321, February 2006

[6] Richard D. Palmer, "Maintenance Planning and Scheduling Handbook," 2006.

[7] Noemi, P. M., & William, L., "Maintenance Scheduling: Issues, Results and Research Needs. International Journal of Operations & Production Management," 1994

[8] Latino, K., "Understanding event data collection: Part 2," Plant Engineering, 2004

[9] Labib, A. W., "World-class maintenance using a computerized maintenance management system," Journal of Quality in Maintenance Engineering, 1998

PERT-IA: A New Heuristic for Schedule Risk Analysis
in Project Management

Mays Assad Hmadh
Industrial Engineering & Managemant
AMERICAN UNIVERSITY OF BEIRUT
Mah111@mail.aub.edu

Abstract- The delay in project schedule is a serious problem, and almost all projects suffer from such delays whether they are small, medium, or large ones. Delays might cause financial losses especially in big projects. The question is why this problem is still there, with all the advancements in project scheduling techniques? One important reason seems to be related to the methods that are utilized in analyzing project schedules such as CPM and PERT, which fail to capture the risk accurately. This paper attempts to estimate the probability of a project being late by accounting for the topology of the whole project network instead of just looking at one path, as in CPM or PERT. The method we develop, PERT-IA, is a hybrid of PERT, which considers the path with the longest duration, and an independence analysis (IA) that assumes that all paths of the network have independent durations. The method estimates the lateness probability as a weighted average of that from PERT and IA, with the weight on PERT being high for project networks having paths exhibiting high commonality (i.e., have many common activities). Results in this paper show that using PERT-IA improves schedule risk analysis, over PERT, by yielding better estimates of the probability of being late. This is validated using Monte Carlo simulation.

I.   INTRODUCTION

Completing project on time is the target goal that all of project managers focuses on, no matter how large or small this project is. The aim to complete all the project tasks remains the same by using project management, but the question that should be asked is why almost all projects usually exceed their schedule. Many reasons causes projects exceed their schedule like in sufficient schedulers, relying just on software, not having a proper description for the project requirements, and one important reason is understanding the influence of risk attitude on decision making [4]. Schedulers don't care much about the input data and its quality; all they care about is having an output that satisfies the customer or the manager. Ignoring the quality of the input data will lead to a big problem, which is project schedule is not realistic or achievable.

As in [3] it is critical for the success of the company to apply project management, and the most important methods used are PERT and CPM. Even though these two methods made a revolution in project management the concept of these two methods was not complete, many enhancement should be applied on them. The proposal suggests increasing the efficiency of scheduling, and to do so we tested the probability of being late for some networks and we compared them to simulation results. To see for what types of networks the difference between results of probability of being late increases, to analyze and figure a solution to minimize this difference, and to conclude that how several characteristics of the network could affect the probability of being late calculations.

PERT-IA is a new approach form PERT; the unique concept behind this approach is totally new. Research on PERT falls in so many domains, and in order to come up with PERT-IA we worked on the mathematical assumptions behind PERT. In order to give more accurate results, many researchers focused on the task time distribution for PERT. This paper shows by numerical comparison, how can we get more accurate results when using the network topology in PERT calculations. In this paper we show how Using network features can give different results than regular PERT.

With these motivations, this paper develops the hybrid method PERT-IA, which estimates the lateness probability as a weighted average of, PERT and IA. What makes PERT-IA really different is how the network topology features are represented in the calculations.

This paper is organized as follows: states the first approach which is commonality percentage; section 3 states the second approach which is potential critical path and weight of commonality approach; section 4states the third approach Pert-IA and the three methods that go under this approach, Pert-IA regular, Pert-IA weighted average, and Pert-IA maximum; section 5 states the methodology to be followed; section 6 states the comparison between the results of Pert-IA approach methods.

II.   LITERATURE REVIEW

Our work contributes to four streams of literature. One is on PERT. The second is about PERT limitations, why it is needed to make changes on PERT. Attempts to improve PERT. The last section is on Monte Carlo simulation.

A.   PERT

For over four decades the Project Evaluation and Review Technique (PERT) has been the most effective tool in project management. As in [2], PERT is an approach that enhances Critical Path Method (CPM), because it takes into consideration the uncertainty in activity durations. PERT and CPM have been used since the 1950s, both methods were

developed after the Second World War to estimate the time needed to complete a project. PERT is used in calculating the probability to finish a project by a given time that is why it's the probabilistic approach from CPM, which considered a deterministic one [3]. CPM was developed in the 1957 in order to help in scheduling maintenance shutdowns of chemical processing plants by J.E. Kelly from Remington-Rand and M.B. Walker from DuPont. In a short time after CPM was developed, the U.S. Navy in 1958 invented PERT in order to manage the progress Polaris missile [1]. Although PERT considered an advanced version of CPM, both methods are network-based techniques. Applying PERT on a certain project includes the following steps:

The first step is Building the network model by determining or estimating duration times for each activity, and precedence relationships. In this phase nodes are used to represent the activities and arrows used to connect the nodes. Dummy activities are used in the network construction, because it is not allowed for two different activities to have the same starting and ending nodes. What differentiate PERT from CPM is dealing with the uncertainty in activity duration; to deal with this PERT uses the three-point estimation technique as in [5] PERT model contains estimation for three times for each activity in the network:

Pessimistic time (b): the longest time that an activity might be completed in.

Optimistic time (a): the shortest time in, which the activity can be, completed in.

Most likely time (m): the time needed to complete the activity that has the highest probability.

The expected time or the mean completion time for each activity is then estimated using the following weighted average: mean=[(a + 4m + b) / 6]; and the standard deviation as (b-a)/6. Then the network is analyzed using the estimated mean of each activity time.

The second step is to determine the critical path. In order to do so we have to calculate the slack time for each activity, and if it is zero then this activity belongs to the critical path. It is required to find out four values for each activity, to calculate the slack times. The values are Earliest Start time (ES), Earliest Finish time (EF), Latest Start time (LS), Latest Finish time (LF). By working forward through the network, and considering the predecessor activities, the ES and EF of each activity are determined. The LS and LF are calculated by working backward through the network. The activity slack time is then the difference between LS and ES, if its value is zero then this activity belongs to the critical path. The longest path in the project is the critical path, which determined by adding the times for the activities in each path. The sum of means for activates on the critical path are normally distributed, which is an assumption in PERT method. This previous assumption allows the probability calculations of the project completion by a given time, as stated in [4].

B. PERT Limitations

Schedules must be realistic and in order to do so PERT limitations should be enhanced. PERT has many limitations; one of the most important limitations is data quality [21]. Using estimated data for activities duration in PERT, is considered to be the main cause of the unrealistic output [14]. So it is normal to have a misleading, wrong result, and projects exceeding the scheduled completion date [15]. The reason behind that is most people care only about the output, and they pay no attention to the input as in [21]. Determining the proper distribution to use for estimating activity durations is one more problem that faces PERT [21]. People behind project scheduling usually don't spend the proper time to determine which distribution among Triangular, Beta, Normal, or Uniform is the best for each project activities [16]. Project scheduler does not spend much time in choosing the proper distribution because of many reasons, being afraid from the project manager, owner, customer, or competency [17]. One more limitation is the critical path is not the shortest path, because this fact may change during the project execution [18]. Non-critical path may become critical and determine project date of completion, this is because the uncertainty assigned to each path in the project [19].

As in [3] it is critical for the success of the company to apply project management, the most important methods to do so are PERT and CPM. Even though these two methods made a revolution in project management the concept of these two methods was not complete. As was also stated by [5], it is crucial for all projects to use PERT and CPM and many enhancements should be applied on them. The success of any organization needs clarity in goals and objectives, it's also very important to identify a schedule for each project or it will be an endless open process. PERT makes heave use of network, and it could be applied on any project no matter how huge or small project is. Also [6] describes how we become in the 21-century, and we still lack a project scheduling system that is valid for all types of networks similarly. In [7] PERT has been used for four decades, and even though it has many shortages PERT networks ignores the stochastically nature of activity times, in addition to that PERT ignores the effect of reducing some activity times on reducing the mean project completion time. As for [4] the delays in project schedule can cause losing one million per day for some projects and less amounts for other but for sure it will cause a non-efficient project execution.

C. Attempts to Improve PERT

Most of the research on PERT is related to task time distribution, as stated in [8]. Activity durations in PERT are beta distributed, and the mean and variance are estimated based on three values, which are also estimated. Because of the previous many researchers tried to improve PERT accuracy by modifying the original three-point estimator. In [20] determine which distribution among the four possible distributions should be used in representing activity durations. The most commonly distributions used are the triangular and beta distributions

Hillison (2004). Triangular and beta distributions are the best for activity durations because they have the asymmetrical property, that matches the risk related to activity durations as [21]. Uniform and normal distributions could be used also, but the problem is that they must be symmetrical which limits the representation of project risk [21].

Several papers discuses the attempts to improve PERT, which proves that PERT needs an improvement to increase its efficiency. As in [9] in order to improve the accuracy of PERT, new mean and variance approximation has been proposed. One of the approximation advantages that was stated by [9], that no more estimation IS required for "pessimistic", "optimistic" and "most likely" time estimates.

While for another approach like [10], they tried to present an approach that try to make sure the activity duration most likely value falls in to the mode of the beta distribution. [11], presented the same method but their method was not consistant.

D. Monte Carlo Simulation

The most important way to overcome PERT, and CPM limitations is Computer simulation. Many difficulties that face PERT could be overcome by using simulation. [12] was one of the first and many researchers that pointed out to how much it is important to use Monte Carlo simulation in studying PERT. Simulation can enhance PERT in many ways, like flexibility in choosing any distribution for activity times, better accurate estimates of the project length, and a better evaluation for risk in a network schedule.

Another literature shows the attempts to apply a graphical modification on PERT called GERT [13]. The previous approach was made to give the ability for the activity time to follow other distributions than beta. [14] Presented a modification on PERT that allowed each activity completion time to follow a distribution. Monte Carlo simulation is widely used in computing the project completion time distributions.

Monte Carlo simulation helps to overcome the uncertainty in PERT activity durations. It calculates the uncertainty associated with results for a system that has many uncertain components in it, by operating and iterating under specific rules to overcome the uncertainty. Monte Carlo simulation can help effectively in determining and deriving the three activity durations estimates (optimistic, pessimistic, and most likely) in PERT. Monte Carlo simulation can determine the Pessimistic, optimistic, and most likely durations in order to the given embedded risk in them [20]. It was founded in 1940s but it has not been applied on project management that since. Project management uncertainties like the completion time and coast estimates used to be solve using Monte Carlo simulation many years ago. As in [21] Monte Carlo simulation computes the uncertainty of project completion time and coast. Mainly Monte Carlo simulation performs the computations of PERT schedule many times, because of the uncertainty associated with each activity durations. All possible combination of the activity durations are used in Monte Carlo simulation to compute PERT schedule times.

A. COMMONALITY Percentage APPROACH

At the first approach in this paper the commonality percentage factor was the factor that has been used to determine which method gives better results in calculating the probability of being late.

- **Pert and IA Calculations**

In this document we are going to use PERT to calculate the probability of being late for a project, in this method the normal distribution is used to calculate the probability, to do so three values are needed the first value is the constant value that we are calculating the probability for being late for, the second value is the mean for this path which is the sum of means for all the activities in this path, the third value is the variance which is also the sum of variances for all the activities in this path. The excel function to calculate PERT probability is

1-NORMSDIST ((the value that we want to calculate the probability for-sum of means for this path)/sum of variances for this path ^0.5)

Before that we calculated the mean and the variance for each activity, like the following (it was justified that it's a Beta distribution with mean and variance from Beta density function).

Mean=μ=(Optimistic+4*likely +Pessimistic)/6=(a+4m+b)/6

The verification for this is that we have (a+b)/2 which it is the mid point, and the mean the weight given for mean=(a+4m+b)/6 is twice the weight for the mid point=(a+b)/2 and the hall is divided by 3 so we get [(a+b)/2+2*(a+4m+b)/6]/3.

Variance=σ^2=((Pessimistic-Optimistic)/6)^2=[(b-a)/6]^2

For the variance it comes from that for any uniform distribution if we took the minimum a and the maximum b, there is a 6σ distance between them as shown in figure1, so 6σ=b-a => σ=(b-a)/6 => σ^2=[(b-a)/6]^2 so its based on a theorem for a uniform model distribution, and 99.98 percent of the measurements says that a, and b are 3σ units from the mean, and in this case it is 100 percent because it is bounded.

For IA: There is another theory to calculate probabilities, which comes from an approximation that the paths are independent; we assumed the paths in the network are independent to be able to calculate the probabilities of being late, and to compare the results with the results from PERT. The method is to multiply the results of Probability of being early which equals 1-Probability of being late from the previous PERT method so we get the Probability of being early, and to get the probability of being late in Max or Independent approximation (IA) method we subtract the previous result from one, the function used in Excel is

Prob of being early in IA=PRODUCT (Prob of being early from Pert)

Prob of being late in IA=1- PRODUCT (Prob of being early from Pert)

We compared the results from the both previous methods to make sure that the results of the work are right, then the last we did is simulating project durations and look at the maximum

project time; we have to get the same previous probabilities that we got using formulas in PERT and IA. We are going to generate 20000 paths and see how many ones of them exceed the value we want; the number of paths that exceed this value is the probability that we are going to compare to our previous results. So we found the probability of being late using simulation.

At last we compared the probabilities that we got from PERT, IA, and Simulation to see how much are the results close or far from each other.

- *THE Approach DESCRIPTION*

A set of illustrated examples for different types of networks was used. Five networks each one has different characteristics used to proof how the commonality percentage can determine which method in between Pert and IA is better for a network. The empirical results showed that when the commonality percentage is high which means there are many activities that are common between paths an the network, Pert gives better results because Pert results in calculating the probability of being late are closer to the simulation results which considered the most accurate method in calculating the probability of being late. While for the networks that have parallel paths, which means low commonality percentage and low dependency IA method gives results that are closer to simulation results which means for network that has low commonality percentage IA is better than Pert.

A final result for this approach is, IA is better than Pert when the paths are parallel, when there is no common activities, and when the dependency is low, no matter whether the length of paths is close or far from each other.

Another result from this approach results is big difference between PERT result and IA means that there is a high dependency between the paths in the project; so PERT is better because of the high dependency.

The last result from this approach is that, when paths are parallel, and highly independent, PERT gives results that are closer to simulation results.

B. *POTENTIAL CRITICAL PATH AND WEIGHT OF COMMONALITY APPROACH*

This approach is an advanced version of the previous approach; at the first approach using the commonality percentage to determine which method in between Pert and IA gives closer results to simulation is not accurate enough.

To get more accurate classification two factors are being used in this approach, these two factors reflects how dependency and paths length affect the determination of when IA is better than Pert and visa versa.

The two factors are Potential Critical Path (PCP), and Weight of Commonality (WC). Four points summarize the final result for this approach.

If PCP is high and WC is low, IA gives better results than Pert for this network.

If PCP is high and WC is high, Pert gives better results than IA for this network.

If PCP is low and WC is low, Pert and IA gives similar results for this network.

If PCP is low and WC is high, Pert gives better results than IA for this network.

In this approach two network characteristics, are used to determine which method is better for which network.

In this approach the main goal is to determine which method is better in calculating the probability of being late in between Pert and IA for different types of projects, two factors are going to do so. Several steps should be applied to reach the final goal, which is determining if Pert is better for the determined project or IA. The first step is eliminating the paths that has possibility to be over dominated by the critical path in this step the aim is to find the Potential Critical Paths (PCP) and a methodology is being used to do the previous. The second step is calculating the weight of commonality, which means calculating the percentage of common activities in a project between all paths except the baths that were eliminated using the elimination methodology. The final step is to determine depending on the previous results of step one and step two which is the best method to use to calculate probability of being late for a project, will it be Pert or IA.

Note: we also can determine weather to use IA, or Pert to calculate the probability of being late, P(Late)=WC*P(Late in Pert)+(1-WC)*P(Late in IA), so after calculating WC for each path, then the average for WC for all paths are calculated, but this method is not going to be used.

A. *Potential Critical Path*

The method used to find Potential_Crtical_Paths (PCP) is as follows: at first mean and standard deviation should be calculated for each path, the path with the biggest mean is the critical path, for this path mean-3*standard_deviation, $(\mu-3*\sigma)$ is calculated while for all other paths mean+3*standard_deviation, $(\mu+3*\sigma)$ is calculated, then if $(\mu-3*\sigma)$ for Critical_Path (CP) is smaller than $(\mu+3*\sigma)$ this Non_Critical_Path (NCP) should be eliminated, so if $(\mu-3*\sigma)$ CP<$(\mu+3*\sigma)$ NCP => this paths is a PCP if not then it is Non_Critical_Path (NCP) that should be eliminated.

B. *WEIGHT OF Commonality*

To find Weight_of_Commonality (WC), all common activities for all activities for each PCP and CP should be counted so if an activity belongs to more than one path it takes a number that refers to how many paths it belongs to, for example if this activity belongs to 4 paths, the weight for it will be 4, and this weight is called Activity_Commonality (AC), then all weights for activities which have AC>1 are counted to get the Number of Activities with AC>1, then this number is divided over Number of Activities-Number of activities with AC=0 (Activities does not belong to any PCP), so for each path which is critical we calculate the ones then we count the ones for each activity, so for activities which belongs to more than one path, and the path should be critical, we count them then we divide number activities with AC>1, and we divide it over number of total activities-activities with AC=0

Note: ones for activities that belong to non-critical paths should not be count.

WC=Number of Activities with AC>1/ (Number of Activities-Number of activities with AC=0).

In this step the best method for calculating probability of being late is determined (Pert or IA) for the project. If WC and PCP percentages are both low then Pert and IA could be used (quarter1), if WC and PCP percentages are both high then Pert should be used (quarter4), if WC percentage is high and PCP percentage is low then Pert should be used (quarter2), and if WC percentage is low and PCP percentage is high then IA should be used to calculate the probability for being late for the project (quarter3). At the last step the error was calculated using two methods the first method is SSE (Sum of Squares Error), in this method the squared error is calculated for PERT and IA, the following function is used in Excel =SQRT(P{Late in PERT}-P{Late in Simulation}), while for IA =SQRT(P{Late in IA}-P{Late in Simulation}).Then the sum of all square errors for PERT, and then the same for IA.The second method is absolute error, for PERT the function is =ABS(P{Late in PERT}-P{Late in Simulation})/P{Late in Simulation} then the average is calculated.The same should be done for IA =ABS(P{Late in IA}-P{Late in Simulation})/P{Late in Simulation} then the average is calculated.
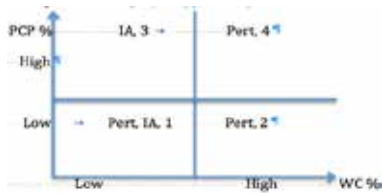

Figure 1: Potential Critical Path and Commonality approach Diagram.

IV. PER-IA

A. *PERT-IA Regular*

In this approach a new method to calculate probability of being late was invented this method is called Pert-IA which is a weighted average for Pert and IA. In the two previous approaches the methods were used are the common methods that are used in calculating the probability of being late, while in this approach the method Pert-IA most of the times give results in between Pert and IA, and for some networks it gives similar results.

Pert-IA method uses the weight of commonality (WC), Pert probability, and IA probability in order to give the probability of being late. The new concept behind this method is using the network WC inside the calculations; while for Pert and IA the network characteristics do not affect their calculations.

Three points summarize this method outcome:

- Pert-IA gives better results than IA and Pert for short deadlines.
- Pert-IA gives similar results or in between Pert and IA when WC is low, or high.

- Pert-IA does not give worse results than Pert and IA in any case.

Methodology

In this approach the paths of the network are determined at first, then categorized in to three categories PCP, Not_PCP, and CP so the critical path is the longest one which has the biggest mean, PCP is the path that its Mean+3σ >Mean-3σ for critical path its and Excel Formula for this is

=IF (Mean-3σ for critical path< Mean+3σ for the other path,"PCP","Not_PCP")

Then PCP percentage is calculated PCP%, which is equal to number of PCP plus one for the critical path over the total number of paths. The next step in this approach is to calculate Activity Communality (AC), and for that all PCP's and CP activities should be set then if the activity belong to one path then it the weight one is assigned to it, and if it belongs to more than one activity the number of paths that this activity belongs to should be assigned to it. Then Weight of Commonality is calculated by dividing the number of activities that have AC>1 over the total number of activities. After calculating WC the Weight of Commonality this value then is multiplied by the value of probability of being late from Pert then it is added to the multiplication of 1-WC by probability of being late from IA to get the value of probability of being late of Pert-IA.

For PCP =IF(Mean-3σ for critical path< Mean+3σ for the other path,"PCP","Not_PCP"). For WC Weight of Commonality =Number of activates with AC>1 over total number of activities (2/6). For PERT-IA WC*Pert+(1-WC)*IA.

B. *Pert-IA Weighted Average*

In this approach the paths of the network are determined at first, then categorized in to three categories PCP, Not_PCP, and CP so the critical path is the longest one which has the biggest mean, PCP is the path that its Mean+3σ >Mean-3σ for critical path its and Excel Formula for this is

=IF(Mean-3σ for critical path< Mean+3σ for the other path,"PCP","Not_PCP")

Then PCP percentage is calculated PCP%, which is equal to number of PCP plus one for the critical path over the total number of paths. The next step in this approach is to calculate Activity Communality (AC), and for that all PCP's and CP activities should be set then if the activity belong to one path then it the weight one is assigned to it, and if it belongs to more than one activity the number of paths that this activity belongs to should be assigned to it. Then all of AC's are added up together to get the Sum of AC's, after that Toatal Activity Commonality (TAC) is calculated which is the sum of AC's for all activities in each path. The next step in this method is to calculate Path Weight (PW) for each path that is equal to TAC/Sum of AC's . After that Average Path Commonality (APC) is calculated for each pat by dividing the mean over TAC for this path, at last the sum of APC for all paths is divided over the multiplication of sum of the means for all paths and the sum of AC's. This final number is used in Pert-IA probability calculation because to get Pert-IA probability of being late this number is multiplied by probability of being late from Pert plus 1-this number is multiplied by the probability of being late from IA. Finally SSE is calculated for Pert, IA, and

Pert-IA to determine for this network which method in calculating the probability of being late gives the best results for this type of network. Then the sum of SSE for each method (Pert, IA, and Pert-IA) to determine which method is doing better for each type of networks. Notations were used in this approach PW: Path Weight, TAC: Total Activity Commonality, and APC: Average Path Commonality

C. Pert-IA Maximum

This method is similar to Pert-IA weighted average in all steps except that the maximum Path Weight (PW) value in Pert-IA calculations Pert-IA Weighted Average.

REFERENCES

[1] Wikipedia contributors, "Program evaluation and review technique," Wikipedia, The Free Encyclopedia.
[2] Haga, Wayne A., and Tim O'keefe. "Crashing pert networks: A simulation approach." In 4th International conference of the Academy of Business and Administrative Sciences Conference, no. 1998, pp. 1-15. 2001.
[3] Islam, Nafish Sarwar. "Complex Project Crashing Algorithm."
[4] Herroelen, W. S. "Practical schedule risk analysis, by David T. Hulett."International Journal of Production Research 48, no. 18 (2010): 5543-5544.
[5] Premachandra, I. M. "An approximation of the activity duration distribution in PERT." Computers & Operations Research 28, no. 5 (2001): 443-452.
[6] Stelth, P., and Guy Le Roy. "Projects' analysis through CPM (critical path method)." School of Doctoral Studies (European Union) Journal 1 (2009): 42.
[7] Haga, Wayne A., and Tim O'keefe. "Crashing pert networks: A simulation approach." In 4th International conference of the Academy of Business and Administrative Sciences Conference, no. 1998, pp. 1-15. 2001.
[8] Avots, Ivars. "THE MANAGEMENT SIDE OF" PERT"." California Management Review 4, no. 2 (1962): 16-27.
[9] Herrerı, José Manuel, Rafael Herrerı, and Johan Rene Van Dorp. "Revisiting the PERT mean and variance." European Journal of Operational Research210, no. 2 (2011): 448-451.
[10] Schonberger, Richard J. "Why projects are "always" late: a rationale based on manual simulation of a PERT/CPM network." Interfaces 11, no. 5 (1981): 66-70.
[11] RIGGS, LELAND S. "Risk management in CPM networks." Computer-Aided Civil and Infrastructure Engineering 4, no. 3 (1989): 229-235.
[12] Van Slyke, Richard M. "Letter to the Editor-Monte Carlo Methods and the PERT Problem." Operations Research 11, no. 5 (1963): 839-860.
[13] Pritsker, A. B., and WHITEHOU. GE. "GERT-GRAPHICAL EVALUATION AND REVIEW TECHNIQUE. 2. PROBABILISTIC AND INDUSTRIAL ENGINEERING APPLICATIONS." Journal of Industrial Engineering 17, no. 6 (1966): 293.
[14] Kennedy, K. W., and Robert M. Thrall. "PLANET: A simulation approach to PERT." Computers & Operations Research 3, no. 4 (1976): 313-325.
[15] Hall, D., Hulett, D. and Graves, R., 2002. Universal risk project final report.Available from the PMI Risk SIG website www. risksig. com, or www. techriskmgt. com/home2link. html.
[16] Hillson, D., 2000. Project Risks: Identifying Causes, Risks, and Effects Accuracy is essential; and good techniques exist that simplify the process and eliminate confusion. PM NETWORK, 14(9), pp.48-51.
[17] Hillson, D., 2003. Effective opportunity management for projects: Exploiting positive risk. CRC Press.
[18] Taleb, Nassim Nicholas. The black swan: The impact of the highly improbable. Random House, 2007.
[19] Williams, T. M. "Criticality in stochastic networks." Journal of the Operational Research Society (1992): 353-357.
[20] Williams, T. M. "What is critical?." International Journal of Project Management 1
[21] Hulett, David. Practical schedule risk analysis. Gower Publishing, Ltd., 2009.

Queueing-Based Design and Analysis of Traffic Systems to Balance Cost and Emissions

Johnny Gemayel
Industrial Engineering and Management
American University of Beirut
Beirut, Lebanon
jeg03@mail.aub.edu

Abstract- **A road cost model is analyzed in this paper based on established queueing theory models of traffic systems. Various model parameters are estimated based on studies that utilize real data. Key Performance Indicators (KPIs), such as the throughput, the mean number of vehicles and average waiting time are calculated and the optimal number of lanes which minimizes costs from building new roads and from delays due to traffic is obtained. The model also incorporates the costs resulting from the environmental impact, in terms of carbon emissions, of road construction and traffic congestion. The latter, being a recent topic of interest, was mostly unaccounted for in previous road design. A sensitivity analysis is also presented, showing what environmental variables are the most influential on the cost function.**

Keywords- **Traffic modelling, Emission factors, Queueing theory**

I. INTRODUCTION

Queueing occurs ubiquitously, from transportation to communication and computer networks, to manufacturing, logistical and service systems. Queueing theory targets the design and control of queueing systems, for example, by determining the right level of resources (servers) and the right service discipline in a way that meets a certain service requirement, or minimizing an expected cost (typically composed of the costs of waiting and the cost of providing service). (See [6] and [11] for background.)

The analysis of queueing systems however seems to have ignored the effect of congestion on emissions, despite the natural link between the two factors. Simply, more congestion leads to more emissions, which is obvious for example, in a highway system, where the longer the time cars spent in a traffic jam, the higher the emissions. The same is true in other transportation systems such as airplanes experiencing taxi delay on a runway, or ships queued in a maritime passage (such as phosphorus or Suez canals). Emissions have also been, somewhat surprisingly linked to congestion in computer networks. For example, [28] reports that Google emits 1.5 million tons of carbon annually; this is slightly higher than the country of Laos. Much of the emissions are linked to the heavy usage of Google search engine and the related congestion.

Given the strong motivation demonstrating the correlation of emission and congestion, we propose to develop queueing models that account for this important dependency. What is interesting about queueing systems is the strong sensitivity of their performance to operational adjustments. For example, adding one more resource or server could reduce delay by ten

folds, in contrast, to the two folds, "linear" intuition. This is promising as simple, highly inexpensive, operational changes could perhaps reduce emissions significantly.

The focus of this paper will be solely on the transportation field, where various road costs are realistically approximated and a cost model is devised. It is to be noted that there are two kinds of traffic that are most commonly studied using queueing systems: Interrupted and un-interrupted traffic flow. Queue length is mostly used with interrupted traffic flow facilities, and is more difficult to define and use in uninterrupted flow facilities as per [9], which are the main focus of this paper.

Queueing theory approaches focused on traffic modelling and analysis using queueing theory, which used deterministic models in the past in [17]. Consequent work in uninterrupted traffic flows used basic queueing models in [12], which were extended to general queueing models [29]. These models were validated in [30] with actual traffic data.

[15] introduced state dependent queueing models to analyze traffic. These models were previously used to analyze pedestrian movements within critical facilities such as schools and hospitals in [32], [4], [24] and [5].

However, these models fail to provide operational procedures focusing on reducing emissions. [23] used actual data to model and reduce taxi times or airplanes which in turn would reduce emissions, without using mathematical models which could be extended to other fields.

Operational work on emissions using mathematical models mainly focused on single and multi-item inventory management previously. In single-item inventory management, [3] use the Economic Order Quantity (EOQ) model to show how operational policies reduce emissions with minimal cost increase. [14] and [27] extend the EOQ model to account for carbon emissions policies. Other useful extensions were obtained in [2] and [25]. Multi-item inventory management was initially investigated in [26] and [22]. More recent works were presented in [18], [7], Zhang and Xu (2013) and [20].

This section presents the cost model and the main theories and studies behind it. Section 2.A formulates a mathematical model showing the relationship between vehicle speed and carbon emissions. Section 2.B explains the theory behind the queueing traffic model used in the cost model originally developed by [15]. Section 2.C presents the cost model without accounting for carbon emissions. Finally in Section 2.D, the cost model is extended to include costs due to carbon emissions.

A. Carbon Emissions and Vehicle Speed

Literature studies of [1] and [19] show a parabolic curve for the car speed vs. carbon monoxide emissions, where the emissions per kilometer tend to infinity at a speed of zero, and then decrease until around 50 mph (80km/h) after which the emissions increase again, as shown in Figure 1. However, the results of Figure 1 were obtained using a software simulator (Mobile5b), which takes as inputs "average speeds, vehicle fleet characteristics, ambient conditions, and trip duration distribution to estimate emission factors" [19]. The wide range of factors that affect car emissions as well as the different types of cars have prevented previous literatures from obtain a clear mathematical equation that models emissions in a car.

Car emissions, however, are mostly carbon monoxide gas (CO), which solely has low greenhouse effects. However, its reactions that produce CH4 and O3 increase its greenhouse effect, where the Global Warming Potential (GWP) of CO could reach up to 3 in the long-run, i.e. 1 kilogram of CO emitted is equivalent to the emission if 3 kilograms of CO2 as per [10].

B. M/G/C/C State-Dependent Queueing System [15]

For modelling and analyzing traffic flows, the established state-dependent M/G/C/C queueing model was used in this paper. The model assumes a Markovian (exponential) arrival rate and a general distribution state-dependent service rate, with 'C' servers and a limited capacity equal to the number of servers. The "state dependent" term accounts for the deterioration in the service rate of the road as the number of cars in the system increase.

Such models were previously used to analyze movements of pedestrians in congested locations, but can also be applied to un-interrupted traffic flow on a highway, as shown in [15]

The limiting probabilities of the M/G/C/C state-dependent models are given by (1), where

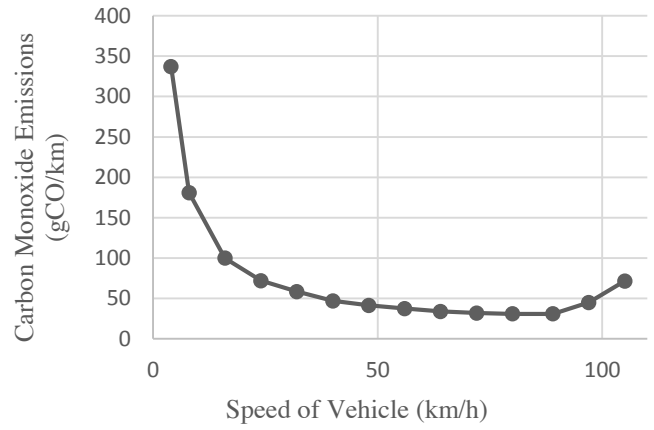- $P0$: Probability of having an empty system, given by (2)

Figure 1. Carbon Monoxide Emissions for a car on the road as a function of speed [19].

- λ : Arrival rate of cars to the road segment in study, practically ranging from 1000 to 6000 vehicles/hour (veh/h) in previous literatures, and assumed to be 4000 veh/h as a base case
- L : length of the road segment in study, which is assumed to be 1 km for generality
- A : Average travel speed of a lone occupant on the road, practically ranging from 60 to 120 km/h, and assumed to be 88 km/h as a base case following the value of [15]
- n : Number of cars in the system, thus Pn is the probability of having 'n' cars in the system
- C : Capacity of the road in veh/km (and the number of servers & capacity in the M/G/C/C model), is given by (3), where
  - o k : Jam density, or the maximum number of cars that 1 lane can accommodate. This value ranges from 115 to 165 veh/lane-km according to [15], depending on varying factors and road conditions. 'k' is assumed to be 138 veh/lane-km as a base case.
  - o N : Number of lanes of the road segment in study, which will be the main decision variable in our cost model

The road segment under study along with the parameters explained earlier are shown in Figure 2, where μ(n) or f(n) as defined in [15] is the state-dependent service rate of the road segment, which deteriorates and the number of cars in the system (n) increases.

$$P_n = \left[\frac{[\lambda L/A]^n}{\prod_{j=1}^{n} j[(C+1-j)/C]}\right] P_0 \qquad (1)$$

$$P_0^{-1} = 1 + \sum_{i=1}^{c} \left[\frac{[\lambda L/A]^i}{\prod_{j=1}^{i} j[(C+1-j)/C]}\right] \qquad (2)$$
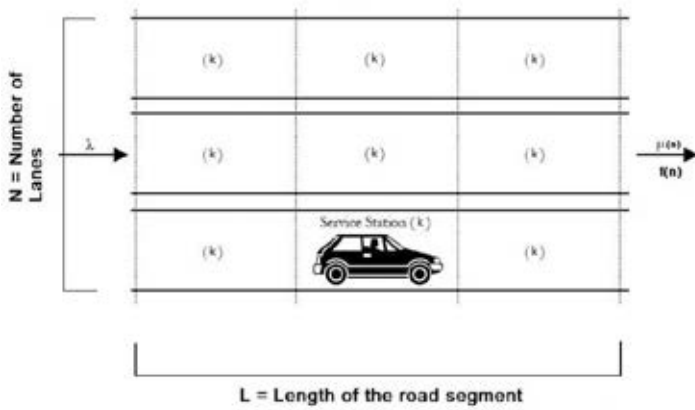
$$C = k \times N \times L \qquad (3)$$

Figure 2. Road segment used for traffic analysis using the M/G/C/C model

C. Our Cost Model without Carbon Emissions

To build our model, we used the cost function stated in (4), without considering carbon emissions. The cost model is expressed in ($/km.h) and consists of two main costs:

1. The cost of providing service (C_S. D(C)) is the cost of building new lanes to accommodate for more vehicles per hour.
2. The cost of waiting (C_w.L) which is the cost of delays on the road due to traffic congestion.

The model parameters are explained as follows.

- C_S : Hourly cost of providing service, which is calculated as the present cost of building a new lane per km, divided by the expected lifetime of a lane ($/lane.km.h). According to [16], expansion of urban highways in the US typically costs around $300,000-700,000 annually per lane-mile including land acquisition price and interests. This is equivalent to $187,500-437,500 per lane-km annually. Assuming the costs lie in the middle of the range specified ($312,500) and dividing by 365 days and 24 hours per day, we obtain an hourly cost of around $36.6 per lane-km. Accounting for the time-value of money over 20 years (240 months), with a 0.5% monthly-compounded interest rate, 6.25M$ will be equivalent to paying equal monthly instalments given in (5). Dividing the above by 30 days and then 24 hours, we get an hourly cost of 62.19$ per lane-km, which is our service cost (CS).
- D(C) : Discrete function of the capacity C representing the number of lanes for a set of values of C as follows, where k is defined above and assumed to be 138 veh/lane-km
    - If $1 \leq C \leq k$, D(C) = 1
    - If $k \leq C \leq 2k$, D(C) = 2
    - Etc…
- Cw : Cost of waiting of 1 vehicle per unit time ($/veh.h). According to [31], the average monthly wage for Lebanese citizens in a study done in 2007 (excluding construction industry) was 690,000 Lebanese Pounds, or 457.71$ as per the currency equivalence at the time of publication of this paper. Assuming an average of 40 working hours per week, or around 160 working hours a month, the Lebanese get paid an average of 2.86$ an hour. Such a wage however is too low in practice, thus we shall choose the cost to be 5$/veh.h as a base case.
- L_S : Mean number of vehicles in system per kilometre (veh/km). After calculating the limiting probabilities given by (1) and (2), the mean number will be calculated as shown in (6), where n is the number of cars in the system and Pk is the probability of having 'k' cars in the system.

The cost model in (4) ignored road costs related to carbon emissions. However, the emissions, as shown later, have a big impact on the cost model and the decision variable, especially when varying some parameters.

To incorporate the emissions cost to the model of (4), two cost factors are added:

1. The emissions cost of providing service, which estimates the emissions resulting from adding new lanes, and evaluates their cost based the carbon tax, which the tax that some developed countries are currently imposing on activities that lead to carbon dioxide emissions.

2. The emissions cost of waiting, which evaluates the costs of the emissions due to traffic congestion.

The updated model shown in (7) adds the costs described earlier to the cost model of (4).

D. Our Cost Model with Carbon Emissions

The model parameters are explained as follows.
- CS : Hourly cost of providing service ($/lane.km.h)
- D(C) : Discrete function of the capacity C representing the number of lanes (lane)
- Cw : Cost of waiting of 1 vehicle per unit time ($/veh.h)
- L : Mean number of vehicles in system per kilometre (veh)
- Ct : Carbon tax or cost of emitting 1 kg of CO2, in ($/kgCO2). The carbon tax varies between countries. On average, each litre of gasoline emits 2.35kg of CO2

$$EC(C) = C_S. D(C) \ (cost \ of \ service)$$
$$+ \ C_W. L_s (cost \ of \ waiting) \qquad (4)$$

$$A = 6.25M \times \frac{1.005^{240} \times 0.005}{1.005^{240} - 1} \qquad (5)$$

$$= 44.78k\$ \ (A/P, 0.5\%, 240 \ months)$$

$$L_s = \sum_{k=0}^{n} k.P_k \qquad (6)$$

As per [13], and the tax is 1.2 cents for each kg of CO2 emitted as per [8]. That make Ct around 0.028\$/L of gasoline on average, which will be the base case in the model.

- Lc : Carbon emissions due to the addition of 1 lane-km or road. According to [21], it is approximated that the construction of a 4-lane, 1km road would emit 2438.5 tons of CO2 during its 20-year life span. This leads to an approximate hourly rate of 3.47 kgCO2 per lane-km.
- I(v) : Hourly rate of CO emissions per vehicle as a function of speed (kgCO$_2$/veh.h), and is obtained using (8), where:
  - o H(v) : Modelled function of carbon emissions vs effective speed (kgCO$_2$/veh.km)
  - o v : Average speed of the vehicles in the system (km/h)

- CO2$_e$: GWP of CO, i.e. the kgCO$_2$ equivalency of emitting 1 kg of CO gases, which has a value of 3 as shown earlier.

H(v) was modelled using the data obtained from [19]. Figure 3 shows the actual data obtained from that paper and the modelled H(v), which is obtained using (9). The RMS error resulting from modelling H(v) was 497, which is considered acceptable and thus the approximation is valid.

III. RESULTS & INTERPRETATIONS

This section evaluates the cost model and generates insights that may affect design studies in the future. Section 3.A evaluates the cost model without accounting for the costs from carbon emissions, which are later added to the model in Section 3.B. Section 3.C shows a sensitivity analysis on the various elements of the cost model and presents significant insights.

$$\boldsymbol{ECT(C) = C_S.D(C)}(Cost \ of \ Service)$$
$$+ \boldsymbol{C_W L_s}(Cost \ of \ Waiting)$$
$$+ \boldsymbol{C_t D(C) L_c}(Emissions \ Cost \ of \ Providing \ Service) \qquad (7)$$
$$+ \boldsymbol{C_t I(v) L}(Emissions \ Cost \ of \ Waiting)$$

$$I(v) = H(v).v.CO2_e \qquad (8)$$

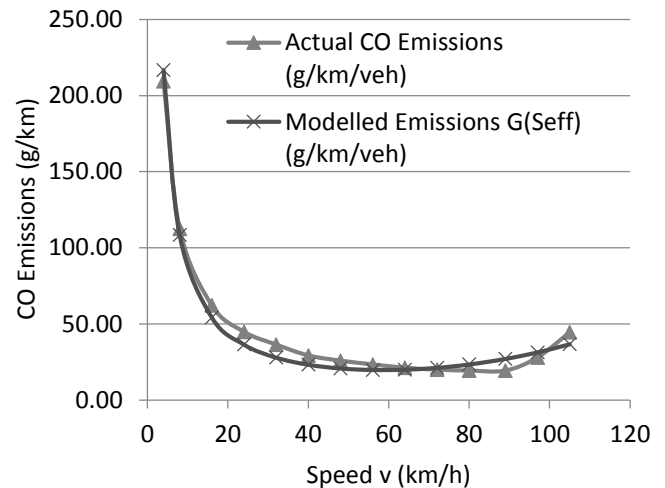$$H(v) = \frac{866.8}{v} + 2.45 \times 10^{-5} \times v^3 \qquad (9)$$

Figure 3. Actual vs Modelled Emissions Function

A. Results of Cost Model without Emissions Cost

Position figures and tables at the tops and bottoms of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be centered below the figures; table captions should be centered above. Avoid placing figures and tables before their first mention in the text. Use the abbreviation "Fig. 1," even at the beginning of a sentence.

The limiting probabilities related to the cost model stated in (4) were obtained using MATLAB, and the cost excluding emissions was calculated and plotted in Figure 4 against the number of lanes (varying from 1 to 10 lanes), which is the decision variable in this case. All the parameters were evaluated at their base case, where

- λ = 4000 veh/h
- L = 1 km
- A = 88 km/h
- k = 138 veh/lane.km
- C$_S$ = 62.19 \$/lane.km.h
- C$_w$ = 5 \$/veh.h

The optimal number of lanes which minimizes the cost in Figure 4 is at N = 2, where the minimum cost is 437.58\$/km.h.

B. Results of Cost Model with Emissions Cost

The cost including emissions was calculated and plotted in Figure 5 against the number of lanes (varying from 1 to 10 lanes), which is the decision variable in this case. All the parameters were evaluated at their base case, where

- λ = 4000 veh/h
- L = 1 km
- A = 88 km/h
- k = 138 veh/lane.km
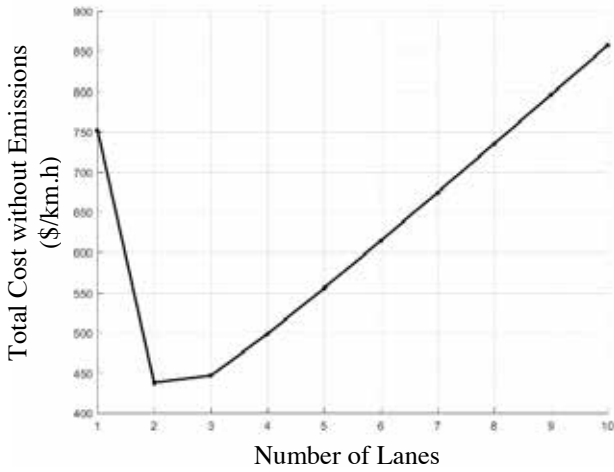- C$_S$ = 62.19 \$/lane.km.h
- C$_w$ = 5 \$/veh.h

Figure 4. Total Cost vs. Number of Lanes without Emissions

Figure 5. Total Cost vs. Number of Lanes with Emissions

- $C_t = 0.028$ \$/kgCO$_2$
- $L_C = 3.47$ kgCO$_2$/lane.km.h

The optimal number of lanes which minimizes the cost in Figure 5 is at N = 2, where the minimum cost is 445.16\$/km.h, with an increase of 1.73% from the minimum cost excluding emissions evaluated earlier.

C. Sensitivity Analysis

The optimal number of lanes which minimizes the cost in Figure 5 is at N = 2, where the minimum cost is 445.16\$/km.h, with an increase of 1.73% from the minimum cost excluding emissions evaluated earlier.

Although the base case plots of Figure 4 and Figure 5 show an insignificant effect of adding the costs of emissions to the cost function of (4), the base case values assumed for both plots may differ in many circumstances, and this difference will cause a greater impact of emissions on the cost function, and even to the point of affecting the decision variable. That could impact road designs in the future if emission costs were accounted for.

Upon varying the parameters λ, A, C_S, C_w, C_t and L_C, the variable L_C had the least impact on the cost function, while variables λ and A affected the decision variable and forced increasing the number of lanes as the variables increase, and that is to maintain an acceptable level of service on the road, which is indicated by the throughput of the system shown in (10) as per [15].

Pbalk is the probability of cars arriving at the road segment under study, but failing to enter the system since the capacity 'C' was already reached.

An acceptable level of service is assumed when λeff is at least 90% of λ. Although variables λ and A altered the decision variable, this impact was obvious and the variables failed to produce any interesting results. The variables which yielded interesting results were C_S, C_w and C_t, which happen to be costs that will vary according to countries, road types, obstacles, etc… Thus, varying these parameters around the base case is realistic,
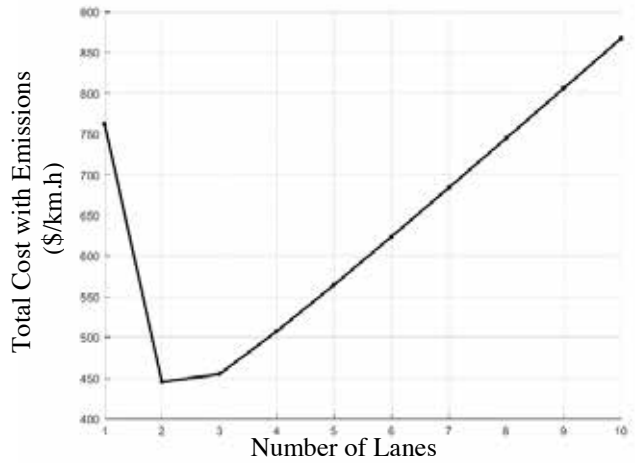
and yields practical results rather than results that may only occur in theory.

Variation of the cost of providing service yielded interesting results, since changing C_S from 62.19 \$/lane.km.h to 52.5 \$/lane.km.h changes the decision variable N from 2 to 3 lanes as shown in Figure 6. The minimum cost corresponding to N=3 lanes was 425.43\$/km.h.

Also, setting C_S to 53 \$/lane.km.h, the minimum cost excluding emissions cost occurs at 3 lanes, and was 419.05\$/km.h. Including emissions cost, the minimum cost occurs at 2 lanes instead of 3 lanes, and was 426.78\$/km.h. This shows how carbon emissions cost can change significant decision variables such as building a 2-lane road instead of a 3-lane road to reduce harmful emissions.

Variation of C_w tends to significantly increase the cost function, especially when the number of lanes is less than 3. This is logical, since for a high waiting cost, each additional vehicle in the system will incur a high waiting cost which dominates the service cost (which is directly proportional to the number of lanes) at low number of lanes. At higher number of lanes however, the service cost dominates the waiting cost.

Also, increasing C_w from 5 to 6\$/veh.h changes the decision variable from the previously optimal value of 2 to 3 lanes. Similarly to the case with C_S, the decision variable depends on the emissions cost for $C_w \approx 7$\$/veh.h.

As for the carbon tax, all our previous analysis are based on a tax of 0.028 \$/kgCO2 emitted. Since this cost is much lower than the service and waiting costs, it usually has a small impact on the cost function, excluding special cases similar to the ones stated earlier. However, this carbon tax is bound to increase in the future, as more efforts are put into decreasing emissions to prevent from global warming effects. This increase in carbon tax will produce more interesting results in the cost model developed, and will eventually be a main component to account for during road design.

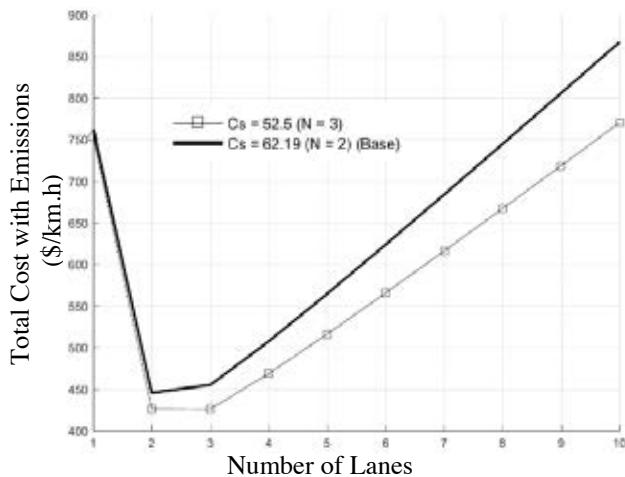$$\Theta = \lambda_{eff} = \lambda(1 - P_{balk}) \qquad (10)$$

Figure 6. Changing Cs from 62.19 to 52.5 vs. Number of Lanes

IV. CONCLUSION

In this paper, a road cost model is developed based on established traffic queueing models and by using realistic parameters. The cost model mainly contributes to studies related aiming to decrease greenhouse gases through cost-effective procedures. It was shown that the cost of emissions could be significant in some cases, and would affect road design decisions if incorporated in the projected road costs. The cost model is also very flexible, allowing it to be applied for any road simply by varying the related parameters.

V. REFERENCES

[1]   André, M., & Rapone, M. (2009). Analysis and modelling of the pollutant emissions from European cars regarding the driving characteristics and test cycles. Atmospheric Environment, 43(5), 986-995.

[2]   Benjaafar, S., Li, Y., & Daskin, a. M. (2013). Carbon Footprint and the Management of Supply Chains: Insights From Simple Models. IEEE transactions on automation science and engineering, vol. 10, no. 1, 99-116.

[3]   Chen, X., Benjaafar, S., & Elomri, A. (2013). The carbon-constrained EOQ. Operations Research Letters 41, 172–179.

[4]   Cheah, J. (1990). State dependent queueing models. Master's thesis, Department of Industrial Engineering and Operations Research, University of Massachusetts, Amherst, MA, 1003.

[5]   Cheah, J. Y., & Smith, J. M. (1994). Generalized M/G/c/c state dependent queueing models and pedestrian traffic flows. Queueing Systems, 15(1-4), 365-386.

[6]   Cooper, R.B (1981). Introduction to Queueing Theory. North-Holland (Elsevier).

[7]   Dekker, R., Bloemhof, J., & Mallidis, I. (2012). Operations Research for green logistics – An overview of aspects, issues, contributions and challenges. European Journal of Operational Research 219, 671–679.

[8]   Dhar, P. (2011). Carbon taxes: An inquiry into their impacts on competitiveness and distribution of income. International Journal of Marketing and Management Research, 2(9), 238-256.

[9]   Elefteriadou, L. (2014). An introduction to traffic flow theory. New Work: Springer.

[10] Fuglestvedt, J. S., Isaksen, I. S. A., & Wang, W. C. (1996). Estimates of indirect global warming potentials for CH4, CO and NOx. Climatic Change, 34(3-4), 405-437.

[11] Gross, D., Shortle, J. F., Thompson, J. M., Harris, C. M. (2008). Fundamentals of Queueing Theory. Wiley.

[12] Heidemann, D. (1996). A queueing theory approach to speed-flow-density relationships. In International symposium on transportation and traffic theory (pp. 103-118).

[13] Hensher, D. A. (2008). Climate change, enhanced greenhouse gas emissions and passenger transport–What can we do to make a difference? Transportation Research Part D: Transport and Environment, 13(2), 95-111.

[14] Hua, G., Cheng, T., & Wang, S. (2011). Managing carbon footprints in inventory management. International Journal of Production Economics 132, 178–185.

[15] Jain, R., & Smith, J. M. (1997). Modelling vehicular traffic flow using M/G/C/C state dependent queueing models. Transportation Science, 31(4), 324-336.

[16] Litman, T. (2009). Transportation cost and benefit analysis. Victoria Transport Policy Institute, 31, 5.6-5 – 5.6-6. Available online at http://www.vtpi.org/tca/tca0506.pdf

[17] May, A. D., & Keller, H. E. (1967). A deterministic queueing model. Transportation research, 1(2), 117-128.

[18] Moon, I., & Cha, B. (2006). The joint replenishment problem with resource restriction. European Journal of Operational Research 173, 190–198.

[19] Sbayti, H., M. El-Fadel, and I. Kaysi. "Effect of roadway network aggregation levels on modelling of traffic-induced emission inventories in Beirut. "Transportation Research Part D: Transport and Environment 7.3 (2002): 163-173.

[20] Schaefer, B., & Konur, D. (2014). Joint Replenishment Problem with Carbon Emissions. Industrial and Systems Engineering Research Conference (pp. 1-9). Montreal: Institute of Industrial Engineers.

[21] Seo, H., Hwang, Y., Seo, S., & Park, K. (2003). Quantitative assessment of environmental impacts on life cycle of highways. Journal of Construction Engineering and Management, 129(1), 25-31. doi:10.1061/(ASCE)0733-9364(2003)129:1(25)

[22] Shu, F. (1971). Economic ordering frequency for two items jointly replenished. Management Science, 17(6), B-406.

[23] Simaiakis, I., & Balakrishnan, H. (2009, August). Queuing models of airport departure processes for emissions reduction. In AIAA Guidance, Navigation and Control Conference and Exhibit (Vol. 104).

[24] Smith, J. M. (1991). State-dependent queueing models in emergency evacuation networks. Transportation Research Part B: Methodological, 25(6), 373-389.

[25] Song, J., & Leng, M. (2012). Analysis of the Single-Period Problem under Carbon Emissions Policies. In Handbook of Newsvendor Problems (pp. 297-315). New York: Springer.

[26] Starr, M. K., & Miller, D. W. (1962). Inventory control: theory and practice. Englewood Cliffs, N.J.: Prentice-Hall.

[27] Toptal, A., Özlü, H., & Konur, D. (2014). Joint decisions on inventory replenishment and emission reduction investment under different emission regulations. International Journal of Production Research, 52, 243–269.

[28] The Guardian (2008). Google discloses carbon footprint for the first time. Available online at http://www.theguardian.com/environment/2011/sep/08/google-carbon-footprint.

[29] Vandaele, N., Van Woensel, T., & Verbruggen, A. (2000). A queueing based traffic flow model. Transportation Research Part D: Transport and Environment, 5(2), 121-135.

[30] Van Woensel, T., & Vandaele, N. (2006). Empirical validation of a queueing approach to uninterrupted traffic flows. 4OR, 4(1), 59-72.

[31] Yaacoub, N., & Badre, L. (2011). The Labour Market in Lebanon. Central Administration of Statistics, part of the EU Twining project to support the Central Administration of Statistics in Lebanon.

[32] Yuhaski Jr, S. J., & Smith, J. M. (1989). Modeling circulation systems in buildings using state dependent queueing models. Queueing Systems, 4(4), 319-338.

Validating the Integration between Project Risk Management and Project Quality Management in Practice
A case Study of Lebanese Projects

Engineer **Amina Ashi**
Department of Industrial and Engineering Management
Beirut Arab University

Abstract

This paper assesses the correlation between project quality management and project risk management in construction projects. The relationship has been tested in terms of a set of indicators chosen based on previous review. Those previous reviews tackled the relationship between quality and risk management in different service sectors. However, none of these reviews addressed the construction sector. Results showed that project risk management and project quality management are cross-linked in real practice. Therefore, one can conclude that they have partially been integrated together.

Keywords: project quality management- project risk management- construction projects

Introduction

The discipline of Project Management in modern times has been undergoing extensive research in its both academic and practical dimensions looking for scientific establishment and best practices that are ever possible. The Project Management Institute (PMI) was one of the few professional institutes that have been taking the major responsibility in proliferating such research findings and best practices to professionals all over the globe. PMI [1] indentified nine different knowledge areas related to Project Management that are divided in four Core Functions; namely: Scope Management, Time Management, Cost Management, and Quality Management, and four Facilitating Functions; namely: Human Resource Management, Communication Management, Risk Management, and Procurement Management. The ninth area is the Project Integration Management area that is aimed at fusing all eight functions identified above into a comprehensive set of working tools and techniques that help Project Managers achieve the goals of their projects successfully. One important question that can be asked is to how much the integration among the eight

functions has been actually attained in real practice. More interestingly is to which degree it is valid to assume that Project Risk Management measures are effective to Project Quality Management results, and vice versa.

The purpose of this paper is to detect the potential link between the processes of project risk management and those for project quality management. The importance of such investigation is due to the possible fact that if quality requirements were not fulfilled, they may lead to compromising the project success, and thus ultimately leading to the higher risk of project failure. Moreover, improper management of potential risks may contribute to poor quality in the project final product.

Literature review

According to Mick McKeown et al [2], Quality and risk were discussed to be directly linked together in the services sector, mainly health care services, where many risks may threaten human lives. Therefore, quality assurance of the service provided in health care is necessary. Another example of expected association between risk and quality management practices is the service of e-business, where the risks inherent in the internet often discourage consumers from using it as a shopping channel. Therefore, the quality of delivering this service is critical as argued by L.W. Turley [3]. Other examples include consumer risk and the quality of food [4], banking [5], and army [6]. All of those examples are clearly born in the service sector domains, the main interest of this paper; however, is the quality and risk management in projects, particularly construction projects.

On the other hand, a great value of focus has been added to the project risk management process in the academic research as well as in practice. For example, a Risks Breakdown Structure (RBS) is introduced as a tool to assist in understanding the distribution of risks

across a project. A Risk Identification Checklist, which is another proposed tool in Risk Management that is based on the RBS, provides a common framework for risk reporting in all similar projects. Standardized RBS forms are suggested for different sectors (construction, hospitals...) based on the most frequently occurring risks in that sector [7]. Regarding construction projects, the most frequently occurring risks has been identified and assessed according to their impact. Having conducted a comprehensive factor analysis on a large number of potential risks in construction, San Santoso et al [8] listed the top risks involved in construction management and ranked them with respect to their impact and their frequency of occurrence in construction projects, among which are:

- subcontractors fail to meet deadlines
- problems in coordination with subcontractors
- errors in drawings
- late in material delivery
- failure of equipment
- slow payment by client
- Lowering Specifications

A great value has been added too to the Project Quality Management processes through research which has helped in identifying the degree of contractors' compliance in some countries to the ISO 9000 quality standards in construction. Bubshait et al [9] showed that the most ISO quality clauses complied with are Inspection and testing and Control of non conforming product, while the least ISO quality clauses complied with are Training and Document and data control. In their work, Barad and Raz [10] reveal that the most critical quality components are Employee Training, Employee Improvement, and Employee Empowerment, Management Commitment, Customer Focus, and Suppliers Management. Moreover, the impact of the existence of a TQM program on a project performance was studied by Bryde and Robinson [11]. Their results show that respondents working in companies with a recognized TQM program are more likely to be customer focused in their practices than respondents in companies with no recognized TQM program.Finally, the question of how can Project Quality actually benefits from Project Risk Management was addressed by Williams el al [12], and that is by prioritizing quality risks. They argued that the most important risks in project management that affect quality used to lie in Poor Processes and Poor Designs, but currently, the important causes of poor quality happen to lie in Weak

Internal and External Relationships with key personnel and partners.

Research Aim

The main interest in this paper is to use a case study to investigate the interdependence effects between the outcomes of Risk Management policies and those of Quality Management procedures in construction projects. The assumption in this study is, as illustrated in Figure 1, that Risk and Quality are interlinked in all projects, especially construction projects so that when quality degrades then there is higher risk that the project will end up with failure, and when more risks are involved in projects then there is a tendency in lowering the quality standards of the final product of projects. This interlink constitutes a positively reinforcing loops in the language of systems engineering, which means that the dynamics of this interdependency tend to contribute even more to strengthen each other without reaching a common balance.
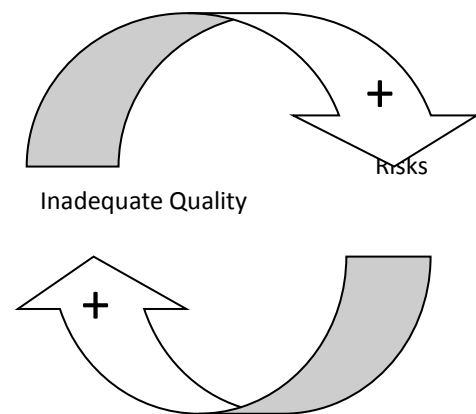
Figure 1: risk and quality are cross-linked

In this work, Lebanese projects are considered for the purpose of validating the above assumption. Lebanon, like other countries in the Middle East and the World has witness in the recent decade a huge boom in construction projects, mainly for residential endeavors.

Therefore, the hypothesis of this study is set as follows:

"A risk compromises the final quality of the project".

The variables that were elected to be used in are selects from the work by Williams et al [10] as deemed to be most particularly applicable in the case of Lebanon after thorough discussion with expert and professional in the construction industry in this country. Not all the variables suggested in the

referenced work were used in order to make the statistical questionnaire simple and not burdensome to the targeted project managers in the sample. The main quality components to be studied are:

- Management Commitment
- Benchmarking
- Employee Training and Empowerment
- Supplier Quality Management
- Customer Focus
- Testing and Inspection
- Reporting and Documentation

The rational for selecting these components is that it was proposed that when Management Commitment is lacking, there will be no focus on the needs of the project clients, so there will be a risk of losing these clients. Moreover, lacking Management Commitment may affect the proper management of supplier. Benchmarking is important too, in order to keep the standard of performance as high as possible. If the organization conducting the project does not improve continuously, it may face the risk of running broke.

Additionally, when employees are, not highly involved, empowered and trained, they will not be able to respond quickly and effectively when a deviation from the plan occurs. However, if the employees are highly trained to respond to risks, but aren't empowered and thus, don't have the authority to take decisions, or react instantly to the risk and take corrective actions once needed, the consequences of the risk may get worse.Insufficient evaluation of suppliers and of the supplied material quality, leads to the risk of delivering a poor final product of the project. Inadequate designs and product quality that are below stakeholders' expectations, may lead to the dissatisfaction of those stakeholders, consequently, their future investments will no longer be capable to compete. On the other hand, risks, if not properly mitigated, can harm the quality of the whole project.

Personal injury is a risk, for example, if it happens during the project, it can affect the morale of the rest of the personnel, thus affecting their productivity and adversely affecting the quality of their work. Errors in drawings, as a risk mentioned earlier, can directly affect the quality of the project. Rework would be required. This means more waste of time and material. Late material delivery, modifications, and coordination problems between different sub-contractors would cause delays in finishing the work. All these delays are a waste of time and money. To compensate for the lost time, work might be rushed, thus inadequate quality

may arise. In addition, testing the supplied material should be regularly done, as well as inspecting used equipment so that the final quality of the product matches the needed specifications.

Research Methodology

The basic approach in this work is to use an analytical survey in order to test the proposed hypotheses, and that is by exploring the association between Risk Outcomes, and Quality Components. A questionnaire was designed and contained 18 statements about both Quality Components (11 statements) and Risk Outcomes (7 statements). The Quality Components that will be considered as the dependent variables in the regression analysis are:

1. The project team members are frequently provided with appropriate training to keep executing their tasks in the best way
2. The Management is continuously evaluating its suppliers regarding how well they meet their material delivery schedule
3. Quality of supplied material is regularly tested to see how well they match the required material quality specification
4. The Management is continuously comparing its performance against competitors
5. The Management is continuously comparing its performance with client requests
6. The working drawings are rarely modified after beginning with the execution phase of the project
7. Weekly reports of achievements are issued
8. Rectifications to control conformance of project execution to required technical specification are documented
9. Inspection and testing of used equipment is regularly done to ensure their efficiency
10. Management is continuously trying to minimize rework, and consequently waste
11. The number of corrective actions decreases when efficient supervision increases

While the Risk Outcomes, which will be the independent variables, are

1. Working drawings usually contain errors (Errors in Drawings)
2. Sub contractors always fail to finish their work on time (Subcontractors Being Late)
3. Coordination problems between different subcontractors always occur during project execution (Coordination Problems)

4. The Management is frequently allocating additional resources to meet critical deadlines (Allocating Additional Resources to Meet Deadlines)
5. The failure of equipment being used frequently occurs (Failure of Equipment)
6. Slow payment by client affects the progress of work (Slow payments by Client)
7. A drop in market price forces developers to lower the finishing specifications (Lowering Specifications)

The questionnaire targeted the population of contractors and developers of construction projects in Beirut city (the Capital of Lebanon) where the best managed projects are expected to take place. The aim was to identify such developers that are expected to have knowledge about Project Management various areas as determined by PMI, and may have been certified accordingly. A pilot study was made to test how well the questions asked in the designed questionnaire, were understood and proper adjustments were done to fine-tune the questionnaire. The survey is based on 5 points Lickert scale answers that varies from Strongly Disagree of "1" to Strongly Agree of "5", allowing a neutral opinion of "3".

The data was collected and run on SPSS™ statistical analysis package that is mostly used in Statistical Analyses. Multiple Regression Analyses were conducted for each of the dependent variables over all independent variables using STEPWISE technique in order to investigate the dependency among these variables and identify the statistically significant independent variables that affect each dependent variable. First, however, Factor analysis was conducted in order to describe the interdependency among each group of variables, i.e Risk outcomes and Quality components as listed above.

Findings and results

Factor analysis

When analyzing the data collected about the group of Quality Components, factor analysis showed a value of the Kaiser-Meyer- Olkin (KMO) measure equal to 0.620. Bartlett's test of sphericity also showed a significance of zero. This means that the data collected about these components are commonly interdependent and thus is useful for analysis.

The extraction method of principle component analysis revealed that there are four factors that can be extracted, which are:

- Factor 1 grouped the following variables together: Employee training, Supplier Management, Testing Materials, and Customer Focus.

- Factor 2 grouped the following variables together: Benchmarking, Reporting, Testing Equipment, and Management Minimizing Rework and Waste

- Factor 3 grouped the following variable: Documenting Non-Conformance

These groupings of variables are logical, since factor 1 includes variables that can be explained as the Management Commitment, factor 2 includes variables that are mainly related to the Practice of Work. In Factor 3, the variable of Documenting Non-Conformance stands alone. When analyzing the data concerning the group of variables measuring the Risk Outcomes in projects, factor analysis technique showed a value of the KMO equal to 0.580. Bartlett's test of sphericity also showed a significance of zero. This means that the data collected about these variables is useful for analysis.

The extraction method of principle component analysis revealed that among the independent variables three factors that were extracted, which are:

➢ Factor 1 grouped the following variables together: Errors in Drawings, and Lowering Specifications.

➢ Factor 2 grouped the following variables together: Subcontractors Being Late, Coordination Problems, Failure of Equipment, and Slow Payment by Client.

➢ Factor 3 grouped the following variable: Allocating additional Resources to Meet Deadlines

These groupings are also logical since the variables that are included in factor 1 are related to "Problems in Execution." While the variables grouped in factor 2 are related to the "Progress of Work and Practice" problems that is the main indication for possible delays. Failure to Meet Deadlines stood alone in Factor 3.

Regression analysis

In order to investigate the validity of integrating Project Risk Management tools and Project Quality Management policies in practice, and specifically the

hypotheses in this study Stepwise Regression is applied in order to discover the main independent variables that are actually affecting each dependent variable of the analysis in the case of each hypothesis.

Testing the hypothesis that "Project risks compromise the quality of a project"

The results showed a few of the quality requirements actually happen to be related to a few risk factors. The significance of these findings is the discovery of risks that affect the quality requirement that the analysis revealed. This should give insights on what risks project managers are induced by in their work in delivering the quality measures as defined in this analysis. Other quality requirements that happen not to be affected by any of the risks considered prove that further efforts in defining the proper risks that may have an effect on these requirements so that the integration function of the PMI model [1] is better enhanced between risk management and quality management procedures in the general project management practices.

In the following the statistically significant regression results are shown and later the quality requirements that showed no significance dependency with the risks under consideration are discussed.

Significant dependencies

The regression analysis that was conducted in order to investigate the effects of risks on each of the quality requirement revealed that the most significant set of Risks (at a statistical significance of 0.05) that are associated with each Quality Requirement are summarized in the following table:

Quality Requirement	Risk(s) Mostly Affecting the Quality Requirement
Customer Focus	Additional Resources to Meet Deadlines
Weekly Reports	Failure of Equipment, Slow Payment by Client
Inspection of Equipment	Additional Resources to Meet Deadlines
Minimizing Rework	Lowering Specifications

Table 1: Risks Affecting Quality Requirements

The ANOVA test showed that for the quality requirement of Customer Focus as the dependent variable that a regression of significance=0.025 is attained with the risk factor of allocating Additional

Resources to Meet Deadlines (Table 1). This means that management efforts are actually made in order to avoid a delay (which is a risk) by allocating additional resources to meet deadlines. This shows that management's best approach is to focus on keeping customers satisfied, by delivering its work to the customer on time .Other quality components, according to respondents' replies, failed to be significantly linked to Customer Focus requirement, and this is quite expected as other quality components deal mostly with internal operations and supply management.

Another significant result in Regression Analysis shows

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	4.483	1	4.483	5.552	.025[b]
Residual	25.032	31	.807		
Total	29.515	32			

a. Dependent Variable: customer focus
b. Predictors: (Constant), meeting deadlines

Table 2: ANOVA results for Customer Focus

that Stepwise Regression attained a significance=0.005 for the quality component of Weekly Reports of Achievement when it is related to the risk factor of Failure of Equipment. It seems that in the view of Project Managers, that reporting is mainly due when to report problems that may lie beyond their personal control.

This is shown below (Table 3), where it happened that stepwise first identified a link between Weekly Reporting and the Risk of Slow Payment by client but

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	5.962	1	5.962	5.449	.026[b]
Residual	33.917	31	1.094		
Total	39.879	32			

a. Dependent Variable: inspection of equipment
b. Predictors: (Constant), meeting deadlines

Table 3: ANOVA results for Inspection of Equipment

was removed in the final most significant model

This may suggest that the former risk outcome affects the latter quality requirement and for the reason that it was correlated with other independent risk variable that is Weekly Reporting was not highly dependent on.

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	5.717	1	5.717	6.869	.013[b]
Residual	25.798	31	.832	25.798	31
Total	31.515	32			
Regression	9.306	2	4.653	6.285	.005[c]
Residual	22.209	30	.740		
Total	31.515	32			

a. Dependent Variable: reporting

b. Predictors: (Constant), meeting deadlines

c. Predictors: (Constant), failure of equipment

Table 4: ANOVA results for Reports of Achievement

Upon applying stepwise regression analysis between the quality component of Inspection of Equipment and the risk factors, only the risk of allocating Additional

Resources to Meet Deadlines happens to be the most significant at a of significance= 0.026 (Table 3). This can be explained in that since when management is always trying to avoid delays, it will make sure to test and inspect the used equipment regularly as a proactive precaution in order to prevent their failure and avoid potential delays.

The last statistically significant result for regression analyses of the quality requirements and risks shows that there is a regression dependency at a significance= 0.041 (Table 4) between Management effort to Minimize Rework and Waste as a quality requirement and Lowering Specifications as a risk outcome. So, it is likely that more rework is needed when high quality is required.

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	3.858	1	3.858	4.564	.041[b]
Residual	26.203	31	.845		
Total	30.061	32			

a. Dependent Variable: minimizing rework and waste

b. Predictors: (Constant), lowering specifications

Table 4: ANOVA results for Minimizing Rework and Waste

Lack of dependencies

In the Regression analysis that was summarized above, not all the quality components (dependent variables) showed significance dependency on the risk variables. More specifically, the quality requirements that lacked dependencies with the risk factors under consideration are:

- Employee Training and Empowerment (Q1)
- Supplier Quality Management (Q2 and Q3))
- Benchmarking (Q4)
- Minimizing Rework (Q6 and Q11)
- Documentation (Q8)

It become very clear that these requirements happen to be not related to the risks introduced in this study, and thus one can conclude that in depth investigation should be conducted to identify what true risks are candidates to affect the quality requirements under consideration. This is a potential for future studies.

On the other hand, some of the risks that were considered happen not to affect any of the quality requirements that were interested in, namely:

- Errors in Drawings
- Subcontractors Being Late
- Coordination Problems
- Lowering Specifications

The above fact leads to the conclusion that the advanced tools and technology happen to effectively mitigate these risks, especially information and communication technologies that provided accurate drawing tools and better communication with suppliers and within the project teams as well. This by no mean suggests that no further risks may affect quality, rather that new directions of investigations should take place in order to identify potential new risks in the project management practices.

Conclusion and Recommendations

This study showed in part that the processes of project risk management and project quality management are cross-linked in real practice. Therefore, one can conclude that they have partially been integrated together. Further investigations are requires to identify potential risks that affect existing as well as new quality requirements in project management practices.

Future research should as well be conducted towards validating the required integration among all core and support functions of project management for the purpose of consolidating the knowledge base of the practice. Rules and standards need also to be developed for this integration to be more successful in the profession of project management.

References

1- A Guide to the Project Management Body of Knowledge- Fourth Edition

2- Mick McKeown, Martin Hinks, Mark Stowell-Smith, Dave Mercer, Joe Forster, (1999) "Q methodology, risk training and quality management", International Journal of Health Care Quality Assurance, Vol. 12 Iss: 6, pp.254 - 266

3- L.W. Turley, (1990) "Strategies for Reducing Perceptions of Quality Risk in Services", Journal of Services Marketing, Vol. 4 Iss: 3, pp.5 - 12

4- Ruth M.W. Yeung, Wallace M.S. Yee, (2003) "Risk reduction: an insight from the UK poultry industry", Nutrition & Food Science, Vol. 33 Iss: 5, pp.219 - 229

5- Tser-yieth Chen, Pao-Long Chang, Hong-Sheng Chang, (2005) "Price, brand cues, and banking customer value", International Journal of Bank Marketing, Vol. 23 Iss: 3, pp.273 - 291

6- Jacob Kashiwagi, Kenneth Sullivan, Dean T. Kashiwagi, (2009) "Risk management system implemented at the US Army Medical Command", Journal of Facilities Management, Vol. 7 Iss: 3, pp.224 - 245

7- David Hillson , "Using a Risk Breakdown Structure in Project Management", Henry Stewart publications, Journal of Facilities Management Vol. 1 Iss: 1, pp. 85-97

8- Djoen San Santoso, Stephen O Ogunlana and Takayuki Minato, "Assessment of Risks in High Rise Building Construction in Jakarta", Journal of Construction and Architectural Management, Vol. 10, Iss:1, pp. 43-55

9- Abdulaziz a. Bubshait and Tawfiq h. al Atiq, "ISO 9000 Quality standards in Construction", Journal of Management in Engineering, 1999

10- Miryam Barad and Tzvi Raz, " Contribution of Quality Management Tools and Practices to Project Management Performance", International Journal of Quality and Reliability, vol. 17, pp. 571-583

11- David James Bryde and lynne Robinson, "The Relationship Between Total Quality Management and the Focus of Project Management Practices", The TQM Magazine, Vol. 19, pp. 50-61

12- Roger Williams, Boudewijin Bertsch, Barrie Dale, Ton Van Der Wiele , Jos Van Iwaarden, Mark Smith and Rolf Visser, "Quality and Risk Management: What Are The Key Issues?", The TQM Magazine, Vol. 18, pp. 67-86

INDEX

570	EL JURDI, JULIE	436	HAMMOUD, HUSSEIN
221	EL KAISSI, DANI	322	HARMOUCHE, MAHER
502	EL KAWAM, IBRAHIM	185	HARRAN, RAY
346	EL KHOURY, ALAIN	574	HASHISHO, EZZEDINE
243	EL RIFAI, HOUSSAM	140	HATOUM, ALI
197	EL SABBAGH, GHINA	666	HMADH, MAYS
542	EL SAFADI, WALAA	243	HOBALLAH, NADER
322	EL SAMAD, GHALI	285	HOUSAIN , HASNAA
362	EL SAYED AHMAD, EL MOATAZ MOHAMAD	429	HUSSEINI, MOHAMMAD
519	EL-HAJJ, ZACHARIAH	072	HUTET, ISMAIL
449	ELIAS SLEIMAN, FADI	313	IBRAHIM, ELIO
496	ELKIK, JAD	171	ISMAEL, HAMZY
608	ELSHAR, IBRAHIM	496	ISSA, RAWANE
301	EZZEDDINE, REINA	423	ISSA, FARAH
449	EZZIDINE, HASAN	555	ISSA, ISRAA
419	FAKHER, GHASSAN	145	ITANI, MOHAMMAD
140	FAKHR, MARWAN	293	ITANI, DANA
307	FAKHREDDINE, MOHAMAD	269	JABER, ALI
542	FANOUS , YAZAN	285	JABER, REEM
293	FAYED, LYNN	523	JABER, TAREK
577	FAYSSAL, IYAD	277	JALWAN, LARA
171	FAYYAD, SARY	191	JAMAL, ROBERT
564	FLEIFEL, MOHAMAD	564	JANNOUN, BASSEL
269	FOUANI, HASSAN	519	JOUDI, IBRAHIM
564	GALI, AMR	419	KAAFARANI, KARIM
527	GEBAI, SARAH	307	KACHOUH, HUSSEIN
185	GEBARA, LARA	436	KANAWATI, OUSAMA
227	GEMAYEL, KAREN	496	KARAM, RAMZI
672	GEMAYEL, JOHNNY	261	KASSAR, TANIA
654	GHALAYINI, MAHER	587	KASSAR, SARI
489	GHAMLOUCH, MOHAMAD	354	KAYALI EL-ALEM, YARA
285	GHANDOUR, HOUSSAM	598	KHALIFE, SALAM
628	GHIZZAWI, FARAH	301	KHALIFEH, FAWZI
660	GHOSN, DIANA	346	KHALIL, JAD
102	GHOSSOUB, CHRISTINE	519	KHAYYAT, MICHAEL
185	HABIB, MAYA	423	KHODR, HALA
374	HACHEM, HANADI	362	KHORIATY, ANTHONY
414	HACHEM, LAYALE	616	KHOURY-HANNA, ANTHONY
523	HACHEM, RIBAL	648	KIOMJIAN, DAOUD
436	HAIDAR, MORTADA	409	KOUZ, ELIE
449	HAIDAR, NOURHANE	301	LADAN, TAMER
152	HAIDAR AHMAD, SERINE	145	LAHOUD, TAREK
549	HAJJ ALI, WAEL	307	LAHOUD, MAKRAM
261	HAKIM RAHME, TAREK	221	LAKKIS, BACHIR
330	HALABI, BASSEL	191	LEBBOS, MICHEL
140	HALAWI, JAD	191	MAAMARI, PAUL
301	HALLAK, BASMA	261	MANSOUR, KARIM
118	HAMMOUD, RABAB	313	MANSOUR, ELIE

253 MAROUN, MICHEL	301 SHOKOR, NABEEHA
382 MASRI, OMAR	374 SHTAYYEH, ZAID
253 MASSAAD, JIMMY	587 SIBLINI, SARAH
285 MATTAR, ISAMAR	354 SINJABA, LEEN
277 MDEIHLI, NOUR	293 SKAYAN, NAYIRIE
338 MEHDI, REEM	180 SLIKA, WAEL
277 MENASSA, NOUR	338 SMOUNEH, SALLY
354 MIKATI, ABDULRAHMAN	471 SONJI, IMAN
269 MOHAMMADIEH, MIRA	221 SOUKI, MARIANA
423 MOHSEN, MOU'MEN	462 SROUR, MALEK
454 MOODAD, JOUMANA	253 STEPHAN, DANIELLE
185 MOUAWAD, RITA	368 SWAID, SOMAR
330 MOUGHRABY, OUSAMA	235 TABAJA, KARIM
213 MOUSTAFA, RAED	197 TABBARA, REINA
346 MUSAED, REHAM	110 TAKKOUSH, SARA
269 NAAMAN, RAZAN	527 TAKKOUSH, SOBHI
382 NACHABE, NOUR	429 TALL, IBRAHIM
382 NAHAS, OMAR	502 TARABAY, JAD
293 NARCISS, TONI	213 TAYYAN, OMAR
368 NASREDDINE, NADER	285 TEDDY, YARA
213 NASSAR, NOURHAN	392 TURKMANI, ABDEL WAHAB
322 NASSER EL DINE, ALI	502 VARTABEDIAN, FREDDY
409 NASSREDDINE, FARAH	145 WEHBE, MALAK
235 NEHME, SAEED	587 WEHBI, BILAL
205, 227 OBEID, HASSAN	166 YAHFOUFI, REEM
235 OUEIDAT , HELMI	180, 330 YAHIA, CESAR
429 OUIEDAT, GHASSAN	072 YASSIN, MARIAM MARIA
243 RAGHEB, JULIEN	368 YASSINE, MOHAMAD
152 RAJEH, CAROLINE	514 YASSINE, REEM
549 RAMLAWI, NABIL	346 YAZBECK, THERESEA
462 SAAB, RASHAD	159 YOUNES, GEORGES
549 SAAB, ALI	392 YOUSSEF, IBRAHIM
159 SAAD, SANDRA	392 ZALLOUM, JEFFRY
362 SAAD, JOY	374 ZALZAL, JAD
636 SAAD, EMILIO	185 ZARIFEH, ADAM
622 SABA, ALIF	098 ZAWAYDEH, SAMIH
330 SADEK, BASSEL	132 ZEIN, MAZEN
374 SAKAYA, KHALED	403 ZEINEDDINE, JAD
191 SAKR, NASSIB	
338 SALAMEH, MAHER	
152 SALAMOUN, LYNN	
514 SALMAN, FATEN	
253 SASSINE, JOE	
159 SEBAALY, CHARBEL	
140 SEHNAOUI, NOUR	
338 SENAN, MOHAMMAD HASAN	
471 SHAABAN, MOHAMAD	
471 SHARAF ALDINE, AHMAD	