ASSESSMENT OF SURFACE WATER POTENTIAL IN DATASCARCE RIVER BASEIN IN ETHIOPIA: CASE OF WABI SHEBELE RIVER BASIN

By:

Ephrem Getachew

May 2019

Addis Ababa, Ethiopia

Addis Ababa University

Addis Ababa Institute of Technology

School of Graduate Studies

School of Civil and Environmental Engineering

ASSESSMENT OF SURFACE WATER POTENTIAL IN DATASCARCE RIVER BASEIN IN ETHIOPIA: CASE OF WABI SHEBELE RIVER BASIN

A thesis submitted and presented to the School of Graduate Studies of Addis Ababa University in Partial fulfillment of the Degree of Masters of Science in Hydraulics Engineering

By: Ephrem Getachew

Advisor: Dr. Belete Berhanu

Addis Ababa University

Addis Ababa, Ethiopia

May 2019

A thesis submitted and presented to the School of Graduate Studies of Addis Ababa University in Partial fulfillment of the Degree of Masters of Science in Hydraulics Engineering

Approved by Members of Examining Board

______________________________	_____________	____________________
Advisor	Date	Signature
______________________________	____________	____________________
Internal Examiner	Date	Signature
_____________________________	____________	____________________
External Examiner	Date	Signature
______________________________	____________	____________________
Chairman	Date	Signature

CERTIFICATION

The undersigned certify that he has read the thesis entitled**: Assessment on Surface water potential in Data Scarce River Basin in Ethiopia: Case of Wabi-Shebele River Basin** and hereby recommend for acceptance by the Addis Ababa University in partial fulfillment of the requirements for the degree of Master of Science in Hydraulics Engineering.

Dr. Belete Berhanu

Advisor

Addis Ababa University

May 2019

DECLARATION AND COPY RIGHT

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Addis Ababa University, I grant to Addis Ababa University the nonexclusive royalty-free right to archive, reproduce, distribute and display the thesis in any and all forms, including electronic format, via any digital library mechanisms maintained by AAU.

I represent and warrant this is my original work, and does not infringe or violate any rights of others.

I acknowledge that I retain ownership rights to the copyright of this work, including but not limited to the right to use all or part of this work in future works, such as articles or books.

Library users are granted permission for individual, research and non-commercial reproduction of this work for educational purposes only.

Any further digital posting of this document requires specific permission from the author. Any copying or publication of this thesis for commercial purposes, or for financial gain, is not allowed without my written permission

Ephrem Getachew

ephgeth@gmail.com

Signature: _______________

Addis Ababa, Ethiopia

ABSTRACT

Quantifying of the potential water resource is the most vital information for appropriate understanding and utilization for diversified economic activities and growth and transformation programs. However, it is hardly challenging, in data scares arid and semi-arid regions. Bias corrected re-analysis climate datasets, like Climate Forecast System Reanalysis (CFSR), became the alternative solutions to model or study the surface as well as the sub surface water resources since the last decade. This study intended to use the bias corrected Climate Forecast System Reanalysis (CFSR) dataset for the assessment of the surface water potential of one of the country data scares basin, Wabi-Shebele River Basin (WSRB). Linear scaling method was applied to avoid the bias on the globally reanalysis CFSR weather dataset to fit for the local conditions. The bias correction was done with limited observed rainfall data collected from National Meteorological Agency of Ethiopia (NMA) holding better records. And the performance of the bias correction evaluated statistically using NSE values ranges from 0.54 – 0.66, for different rainfall stations distributed over the basin. The bias corrected CFSR weather data, currently available local and global soils and land use data sets with better spatial resolution were used to develop the Soil and Water Assessment Tool (SWAT) hydrological model to evaluate the surfaces water potential of Wabi-Shebele River Basin. The statistical model performance evaluation at calibration and validation phase conformed that the model setup to mimic the reality in the ground and be a good tool to quantify the surface water potential of the basin. This model performance described with the statistical measures of NSE and coefficient of determination (R^2) with values 0.40 and 0.44 on calibration and 0.11 and 0.12 on validation phase of the model development. Finally, the surface water potential of the basin estimated as 7.7 BCM, which is annually renewed with the normal hydrological conditions of the basin.

Key Words: CFSR, REANALYSIS, SWAT, SURFACE WATER POTENTIAL, CALIBRATION, VALIDATION

Acknowledgment

Praise be to GOD, who capable me to reach this end. I firstly would like to thank my families and neighboring sisters who were on my side during the past years of study.

I wish to express my utmost gratitude to Dr. Belete Berhanu, for his precious advice from the scratch of the thesis area, encouragement and decisive comment and all rounded support during the research period. His critical comments and valuable advices helped me to take this research in the right direction.

I would like also to thank the Ministry of Water, irrigation& Energy bureau and the National Meteorological Service Agency for their cooperation in providing the required data.

My special thanks go to my friend Dr. Mesfin Sahle for his important contribution on the utilization of GIS tools and his valuable support till the end of this thesis.

Ephrem Getachew

ephgeth@gmail.com

Table of Contents

List of Tables

List of Figures

List of Acronyms

CFSR – Climate Forecast System Reanalysis

DEM – Digital Elevation Model

GIS – Geographical Information System

HRU – Hydrologic Response Unit

IDW – Inverse Distance Weighting

LULC – Land Use Land Cover

MoWIE – Ministry Of Water, Irrigation and Energy

NMA – NATIONAL METEOROLOGICAL AGENCY

NSE – Nash Sutcliffe Efficiency

PBIAS – Percent Bias

RSR – Root Mean Square Error Standard Deviation Ratio

SCS – Soil Conservation Service

SWAT – Soil and Water Assessment Tool

WMO – World Meteorological Organization

WSRB – Wabi Shebele River Basin

1. INTRODUCTION

1.1. Background

Water resource is the most vital resource in the earth's atmosphere which requires detail studies for appropriate understanding and utilization in a diversified economic activity. These may include water supply for domestic or industrial purposes, hydroelectric power production, irrigation or flood control (SUTCLIFFE et al, 1990). As the availability of water is highly uneven in space and time, improper assessment of water resources is potentially devastating (Kindie, 2016). The underestimation of expected flood can lead to overtopping of any hydraulic structures built in the area and consequent its failure. On the other hand, for projects where water potential is overestimated, the system may not come to a position to fill up to the full reservoir level, we can take the example of Jijiga dam in Somali Region, Ethiopia. Growing population demands for the increase of domestic water supply and consumption of big amount of water due to expansion in agriculture and industry. Mismanagement and lack of knowledge about existing water resources and the changing climatic conditions have consequence on the imbalance of supply and demand of water. The problem is prominent in semi-arid and arid areas where the resources are strongly limited.

Wabi-Shebele River Basin (WSRB) is one of the river basins that flow in the eastern part of the country with high population mass. WSRB is the largest river basin in terms of its catchment area but with a least surface water potential (Berhanu B et al, 2013). Due to the lack of enough hydrological data and in depth studies in the surface water potential of the river basin, its surface water potential always reported as a minimum of all river basins. Like in many developing and arid regions of the world, the assessment and management of water resources in WSRB is still suffer with climate and hydrological data scarcity. According to (Elias Nkiaka et al, 2017)the difficulty in collecting data in semi-arid and other remote regions can be attributed to several reasons: (i) lack of reliable equipment; (ii) absence of good archiving system and software to store and process the data, and lack of funds to organize data collection campaigns. Another challenge in these regions is that even when data is collected and archived, the effort and money required to access them can be quite substantial.

Rainfall is the most important data used in hydrological models; hence it is important to obtain it in sufficient temporal and spatial resolution. The spatial and temporal variability of rainfall affects the spatial and temporal distribution of hydrological variables (Berhanu B. et al., 2016). Nevertheless, due to the high spatiotemporal variability of rainfall, it can only be accurately captured by a dense network of rain gauge stations. But, like as for most developing countries, (A. W. Worqlul et al., 2015) dense network is not available in Wabi-Shebele River Basin.

In another hand there are different efforts in the scientific group to overcome the problem of this data scarcity, a multi-year global gridded representations of weather data known as reanalysis datasets became one of the pivot point results in last decade. The large number of variables makes reanalysis datasets ideal for investigating climate variability and to enhance management of water resources (E. Nkiaka et al, 2017).Many researchers all over the world utilized multiyear global gridded representation known as reanalysis data sets for evaluating the surface water potential of a watershed (Tissa et al, 2015; A. W. Worqlu et al, 2015).

Therefore, the purpose of this study is to fill the gap of data scarcity in Wabi-shebele River basin, Ethiopia using Climate Forecast System Reanalysis (CFSR) climate data set and evaluate the surface water potential of the basin.

1.2. Statement of the problem

According to Ayenew et al. (2008) and MoWIE (2014), Ethiopia has three principal drainage systems, which start from the central highlands. The first and largest is the western system, which includes the watersheds of Abay (Blue Nile), Tekeze and Baro-Akobo, all flowing west to the Nile System. The second is the rift valley internal drainage system, which includes the Awash, rift Lakes and the Omo-Ghibe basins.

The Awash River drains to the northeast through the rift floor and remains entirely contained within the boundaries of the country and enters Lake Abbe near Djibouti border. The Rift Lakes basin is a closed system located in central Ethiopia where several rift valley lakes are found. The Omo and Ghibe rivers flow to Lake Turkana in the Ethiopia and Kenya border. The third system is the Wabi-Shebele and Genale-Dawa rivers, which drain to the Indian Ocean through Somalia. (Ethiopian academy of sciences, 2015)

Wabi-Shebele River Basin is one of the largest basins in Ethiopia having a total area of coverage 202,697 km^2, with very low estimated water resource potential for the development of the basin. While having the largest area coverage, its annual runoff reported to be around 3.4 billion m^3 (Abbay Basin Authority, 2016) which is very low as compared to its land mass and the highest rainfall on its upstream area. Hydrological and meteorological data scarcity are major bottleneck that retards advancement of knowledge on water management and climate change in the basin. According to (Adane, 2009) there are about 48 meteorological stations around the basin but most of them are clustered at or near the city of the upper portion of the basin with most of the stations having incomplete and insufficient rainfall records (Fig 1-2). It is also noted that fewer than 19 stream gauging stations are available in the basin but with most of them are not operational at present and few have limited data in the long past year. Kinde (2015) also noted that the river flow data in the watershed is limited to the upstream area and rarely available to downstream part of the basin and the data available are only a few years with most of it missed.

WSRB characterized with fragile arid and semi-arid ecosystem that need well distributed rainfall and evapotranspiration information to understand the ecosystem and its resources potential, particularly water resources. The basin is among the most vulnerable areas due to uncertain water availability, climate change and the frequency of extreme conditions which result in a more frequent hydrologic disasters, such as floods and droughts (MWAR –LAC, 2016). In the other hand the basin has large land area for irrigation development and massive biomass production, which make it one of the growth and development corridors in the Growth and Transformation Plan (GTP) of the country. Therefore, this study tries to enhance the understanding of the spatio-temporal distribution of surface water resources in the basin to complement the implementation of projects in the GTP.

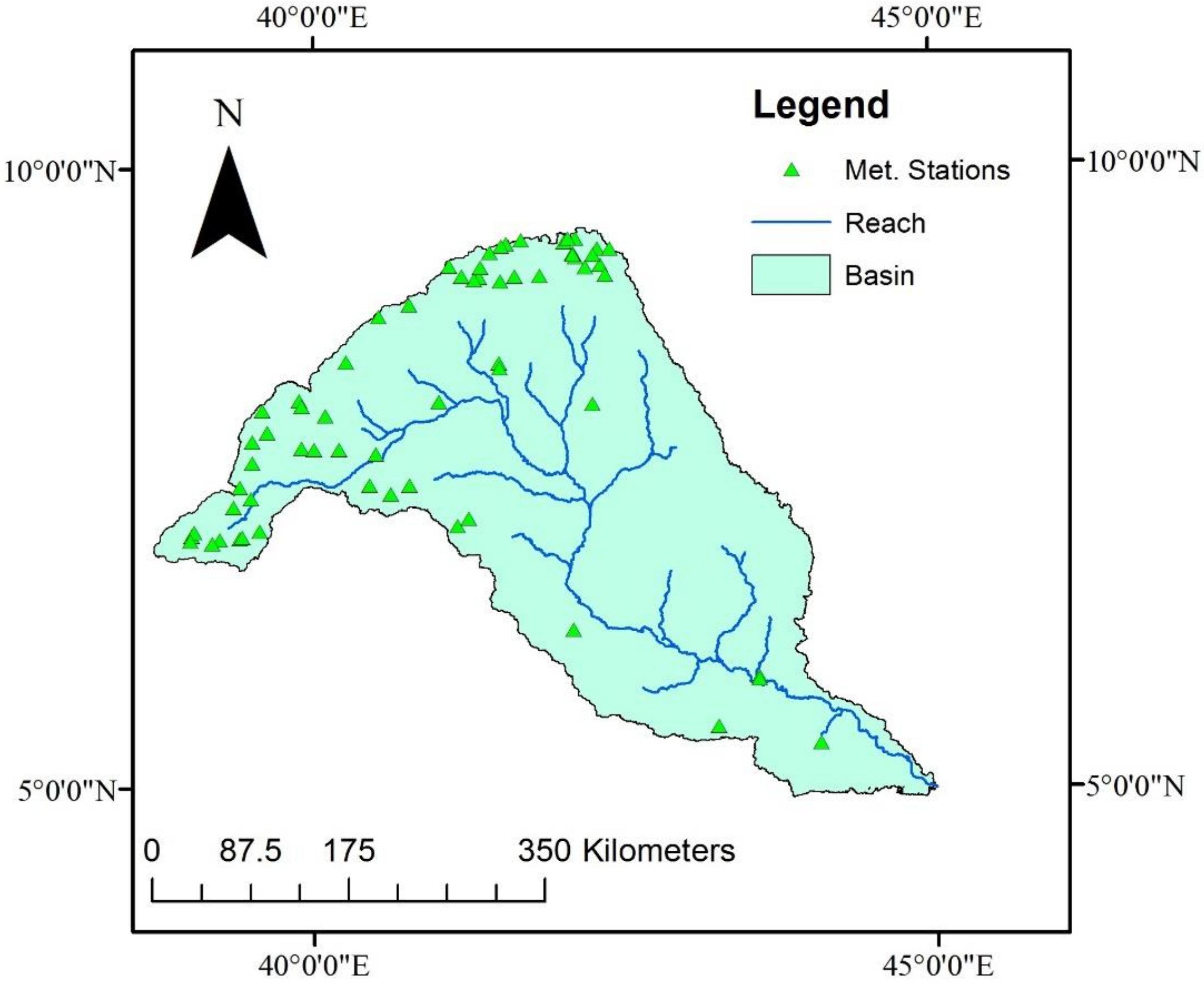


Fig 1-2. Meteorological Stations distribution in WSRB

1.3. Objective of the study

1.3.1. General Objective

The main objective of this study is to evaluate the surface water potential of Wabi-Shebele River Basin using CFSR data sets.

1.3.2. Specific Objective

- Organize the basin CFSR bias corrected and validated rain fall data
- Estimate surface runoff and evaluate the actual water balance of the basin

1.4. Significance of the study

WSRB is the largest river basin out of the twelve river basins in Ethiopia which covers an area of 202,697 km^2. Investigating the history of drought frequency, severity and duration for a region provides a greater understanding of the region's drought characteristics and the probability of drought recurrence at various levels of severity (Tesema, 2015). The settlement of the population can be categorized as highland and lowland residents and large percentage of the population in the highlands depend on agriculture while the lowlanders in general are pastoralists. However, the current estimate of the surface water resources, which account as 3.4 BM^3, doesn’t comparable with the size and has strongly been influenced with data scares hydrological condition of the basin. Therefore, re-evaluating the spatial and temporal distribution and magnitude of surface water has very high significance, to release more accurate estimate of surface water potential of the basin and initiate policy intervention in the development of irrigation schemes, drought mitigation measures and power generation techniques.

1.5. Scope of the study

The scope of this research is limited to assessing the surface water potential of WSRB. There is an increasing demand of irrigation development, hydropower generation and creating climate resilient livelihood. This study used CFSR climate dataset with a resolution of 38 km for data period of 1979 to 2010 and 22 km for data period of 2010 to 2017. Following data arrangement, bias corrections applied to the precipitation data only as this data is the major element of hydrologic cycle. The study does not consider small storage reservoirs in the basin for the general surface runoff estimation. The evaluated surface water resources of the basin is subject to change in the future based on improved data resolution and other research finding on the basin.

1.6. Conceptual framework

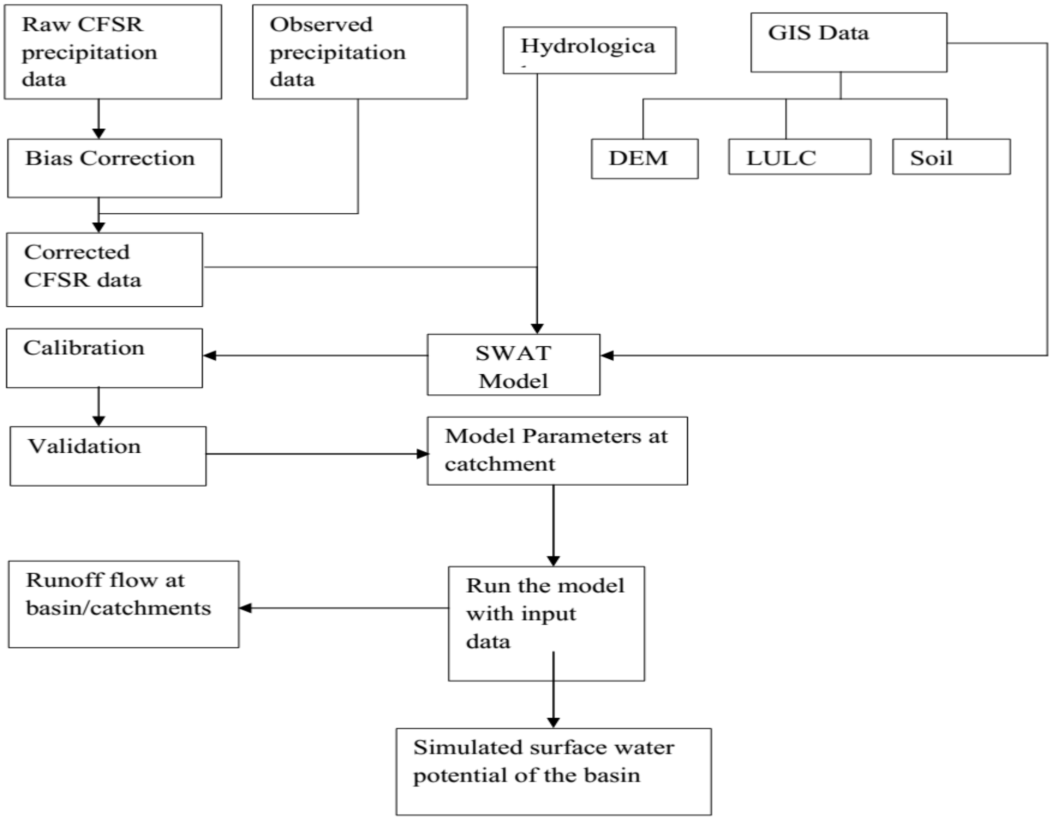


Fig 1-6. General framework of the research

1.7. Thesis organization

The thesis is divided in to six chapters. Chapter one is the introduction part of the study which holds background information about the study, statement of the problem, the study objective, significance and scope of the research. Chapter two is literature review part which collects works of previous studies in related topics. Methodology of the study then followed in chapter three which includes details of the study area, collected data type and procedure of collection, data analysis and SWAT model set up. The fourth chapter is the result and discussion part which presents result of corrected precipitation data, sensitivity analysis of the model, calibration and validation of model output, and the final result of the thesis. Finally, chapter six present conclusions and recommendations provided based on the study finding.

2. LITERATURE REVIEW

2.1. Surface Water Potential Of A Basin

A river basin or watershed is an area over which various hydrologic processes such as precipitation, snowmelt, interception, evapo-transpiration, infiltration, surface runoff, and sub -surface flows are integrated. The water resources of Ethiopia are governed strongly by the amount and distribution of rainfall. The distribution of rainfall over the country is highly variable. Variations in rainfall throughout the country are highly influenced by differences in elevation and seasonal changes in the atmospheric pressure systems that control the prevailing winds. These factors are the drivers of spatial and temporal variability of rainfall distribution and water availability in Ethiopia (Woldeamlak B. et al, 2015). Surface water resources are available although their spatial and temporal distribution and settlement pattern of the population limits their utilization. The mean annual specific runoff varies from zero to 35 l/s per km2. Minimum flows occur in the period from December to March. Apart from the big rivers and their major tributaries, there is hardly any perennial flow in areas below 1500m. In general, perennial streams and springs exist only in the vicinity of mountains with an annual rainfall of more than 1000mm (UN-Water, 2014).

Most of the Ethiopian rivers originate from highland areas and flow into different directions to lowland areas including to the neighboring countries. This is the reason why Ethiopia has been considered as the water tower of Northeast Africa. The total annual surface runoff from the twelve river basins amounts to about 124 billion cubic meter (Selishe et al, 2007). Although this represents an immense amount, its distribution in terms of time and space is erratic. Out of the 12 river basins eight basins namely Abbay, Baro Akobo, Tekeze, Mereb, WabiShebele, GenaleDawa, Omo Gibe and Awash can be considered as wet basins as they generate considerable quantities of flow. The Lakes Basin (Rift Valley Basin) in which several lakes are fed by numerous rivers and streams and three other basins namely Danakil, Ogaden and Aysha are considered as dry as they receive low amount of rainfall that cannot even satisfy evaporative demands (Woldeamlak

B. et al, 2015). UN-Water (2014) indicated out of the total runoff only 3% of it remain in the country and 97 % of it lost as runoff to the lowlands of the neighboring countries.

One of the most determining important questions in hydrology is how much runoff occurs in a river in response to a given amount of rainfall. To answer this question we need to know how it rains, where it rains, where and how water is stored in the different land surface zones and what pathways water follow to reach the stream channel. These are some of the questions which can be addressed in rainfall-runoff studies. The term runoff is used for overland flow and shallower interflow process including the ground water that takes water to the river approximately a day. The main hydrological mechanisms that generate overland flow are infiltration excess and saturation excess. Infiltration excess is generated when rainfall intensity exceeds the infiltration capacity of the soil causing overland flow to generate. It is based on the concept that runoff is generated when rainfall rates exceed soil infiltration capacity so runoff amount is directly controlled by factors that determine soil infiltration capacity such as land use, soil type and moisture content. Infiltration excess overland flow is commonly referred to as Hortonian Overland Flow (Assayew N., 2016).

Saturation excess is fundamentally different since overland flow is not generated by high rainfall intensities but by land surface saturation as a result of soil saturation by a rising water table. Saturation zones occur close to river reaches at lower laying areas in a basin that are characterized by infiltration zones. Unlike Hortonian flow, where soil type and land use basically play a controlling in runoff generation, landscape position, local topography, and soil depth are some of the major factors control on saturation excess runoff which varies for different catchment. Saturation excess is at the base of the Variable Source Area (VSA) concept that acknowledge the spatial extent of saturation excess varies seasonally, depending on the relative rates of rainfall and evapo-transpiration (Ibid). The runoff generated through the different mechanisms flow over the river and termed as the Surface Water of a basin.

Surface water is a water that is open to the atmosphere and fed by runoff from the surface, such as in a stream, river, lake, or reservoir (Malual, 2015). Among the various resources available in the earth, water is the most important and essential for the life of both human beings and other living things. The expansion of agricultural production is also highly dependent on the availability of water resource. Mbungu et al (2017) indicated that feeding the world population in the recent decades increased the demand for food production and hence causes the expansion of croplands. Approximately 38% of the global land surface is occupied by croplands and grazing lands (Ibid). The study cited the source of World Bank and reported that agriculture employs more than 31% of the world population and this figure even higher for developing countries like Ethiopia. These increasing demands require an appropriate understanding of the water potential of a country in general and the available water potential in specific watersheds in particular. Seleshi (2010) indicated that Ethiopia has an irrigable land potential of around 5.3 million ha and out of which 640, 000 ha of land developed using irrigation systems. Out of the Ethiopia's basin irrigation potential, Wabi-Shebele river basin placed at the 4th largest irrigable land potential which approximate to 237,905 ha of land (Ibid).

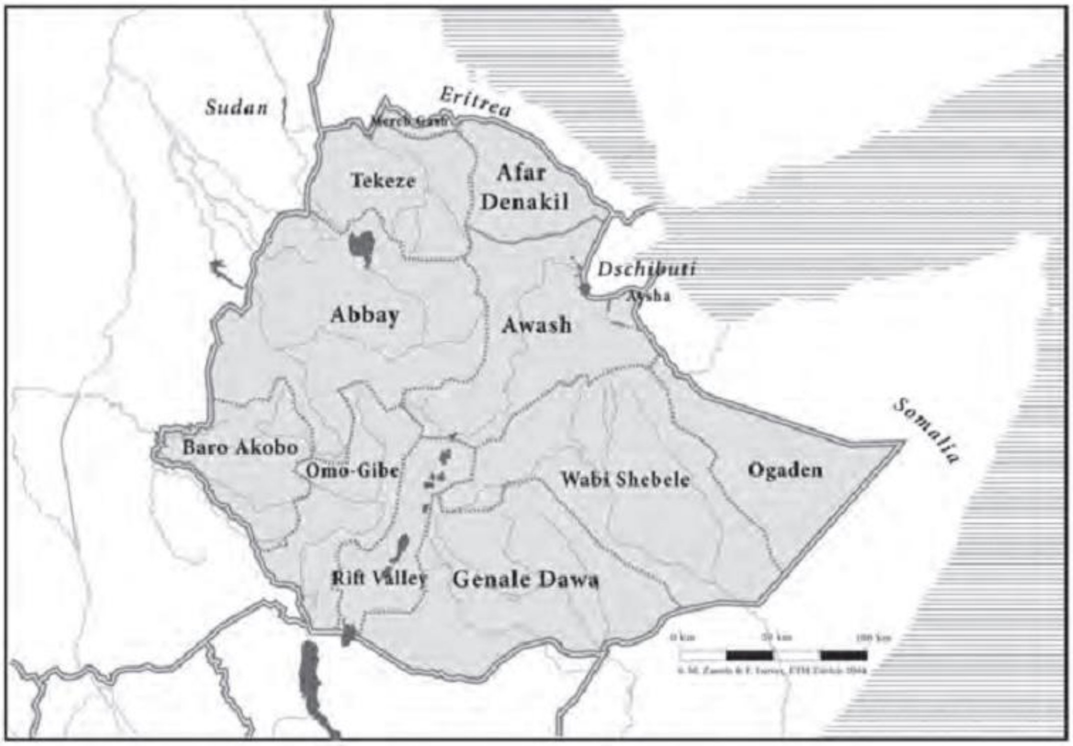


Fig 2-1. River Basins of Ethiopia Separated by broken lines (Source: Woldeamlak B, 2015)

Stream discharge and rainfall data are the essential information required to plan and design any watershed related project. However, in most of the cases these data may not be available prior to the commencement of new project at specific area. For the study area of this research, WabiShebele, the available number of stream gauging stations and rainfall gauging stations are limited and are not well represent the watershed (Tissa, et al, 2015). The problem for efficient utilization of the water resources is the availability of input data for hydrological modeling. According to (A. W. Worqlul et al, 2015) accurate prediction of hydrological models requires accurate spatial and temporal distribution of rainfall observation networks of gauging stations. In developing countries rainfall observation station network are sparse and unevenly distributed. According to the World Meteorological Organization (WMO, 1994) the minimum rainfall station network density for tropical regions is 600 to 900 km^2 per station for flat areas and 100 to 250 km^2 per station for mountainous regions. But, in developing countries like, Ethiopia, such configuration is not available due to economic situations. Recently the availability of satellite rainfall estimation techniques where there is limited or no conventional ground rainfall observation stations has attracted the interest of hydrologist (Worqlul et al, 2015; Tissa et al, 2015; Tewodros Woldemariam et al, 2017). Satellite rainfall estimates have the advantage of high temporal resolution and spatial coverage, even over mountainous regions and sparsely populated areas.

2.2. Global Reanalysis Weather Data

2.2.1. Overview

In climate science, a retrospective analysis – or reanalysis – is a synthesis of both observational and numerical model data, spanning an array of climate variables. This combination of data and modeled fields provides scientists with multi-decadal information on weather patterns, climate variability, and change. Variables include everything from modeled surface fields, such as precipitation, evaporation and radiation fluxes, to components of the stratosphere. The use of atmospheric re-analyses has been adopted amongst a variety of Earth system sciences. (ALLISON R., 2009)

According to (Hu Sheng et al, 2016), Global reanalysis weather data provided by the United States and Europe is currently used for various hydrological applications around the world. Some examples of widely used include: the National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR); the NCEP and US Department of Energy (DOE) NCEP/DOE; the NCEP and the National Center of Atmospheric Research (NCAR) NCEP/NCAR; the European Centre for Medium-Range Weather Forecasts (ECMWF) RA-15/40 (ERA Interim); Modern-Era Retrospective Analysis for Research and Applications (MERRA); and National Aeronautics and Space Administration (NASA) DAO.

2.2.2. Climate Forecast System Reanalysis (CFSR)

The CFSR is a third generation reanalysis product. It is a global, high resolution, coupled atmosphere-ocean-land surface-sea ice system designed to provide the best estimate of the state of these coupled domains over this period. The National Centers for Environmental Prediction (NCEP) Climate Forecast System (CFS) is initialized four times per day (0000, 0600, 1200, and 1800 UTC). NCEP upgraded CFS to version 2 on⊞ March 30, 2011. This is the same model that was used to create the NCEP Climate Forecast System Reanalysis (CFSR). Selected CFS time series products are available at 0.2, 0.5, 1.0, and 2.5 degree horizontal resolutions at hourly intervals by combining either 1) the analysis and one- through five-hour forecasts, or 2) the one- through six-hour forecasts, for each initialization time. This reanalysis will serve many purposes, including providing the basis for most of the NCEP Climate Prediction Center's operational climate products by defining the mean states of the atmosphere, ocean, land surface, and sea ice over the next 30-yr climate normal (1981–2010); providing initial conditions for historical forecasts that are required to calibrate operational NCEP climate forecasts (from week 2 to 9 months); and providing estimates and diagnoses of the Earth's climate state over the satellite data period for community climate research.(NCAR UCAR, 2018)

The NCEP Climate Forecast System Reanalysis (CFSR) was completed over the 39-year period of 1979 to present. The CFSR was designed and executed as a global, high resolution, coupled atmosphere-ocean-land surface-sea ice system to provide the best estimate of the state of these coupled domains over this period. The current CFSR will be extended as an operational, real time product into the future. The CFSR relative to most, if not all, previous reanalysis include (1) coupling of atmosphere and ocean during the generation of the 6hour guess field, (2) an interactive sea-ice model, and (3) assimilation of satellite radiances by the Grid-point Statistical Interpolation scheme over the entire period. (Suranjana, 2010)

2.2.3. CFSR Data Application And Previous Studies

Weather records are seldom complete which required substituting and applying other measurement products. Accordingly, some researchers have utilized radar data to provide precipitation inputs for hydrological modeling studies (Daniel Fuka et al, 2013). Although these data provide important contribution for data scarce regions, they pose their own challenge including discriminating different forms of precipitation such as hail, snow and rainfall and determining the appropriate relationship between radar reflectivity and rain rate (Ibid), to mention that radar data are only available for small fraction of the world’s land surface. Thus, there is a need to consider additional methods to estimate weather conditions for watershed scale modeling.

To remedy the above problem, one possibility is the use of multilayer global gridded representations of weather known as reanalysis data sets, of which there are several types based on their resolution and coverage (Table 2-2-2). Fuka et al (2013) suggested that selecting reanalysis data set for small to medium sized watershed can be done using the following three criteria’s, (1) the dataset should be open and available, including temperature and precipitation; (2) spatial resolution needs to be 30 km; and (3) the length of records should include adequate historical coverage to allow model calibration and validation, and extend to the present. Among all the available data set which meets the above three criteria the NCEP climate forecast System reanalysis (CFSR) data set is the best choice.

The CFSR data set provided by NCEP is available starting from 1979 to present. It is currently used by many researchers all over the world to predict hydrological situations in data scarce watersheds. R. Fuka (2013) indicated that obtaining representative meteorological data for watershed-scale hydrologic models can be difficult and time consuming. Land-based weather stations do not always adequately represent the weather, because they are often far from the watershed of interest, have gaps in their data series, or recent data is not available. The study suggested that utilizing the CFSR precipitation and temperature data provide stream discharge simulations that are as good as or better than simulations using land based weather stations, especially when stations are more than 10 km far from the watershed. The CFSR data could be particularly beneficial for watershed modeling in data-scarce regions and for modeling applications requiring real-time data. The CFSR dataset consists of hourly weather forecasts generated by the National Weather Service's (NWS) NCEP Global Forecast System (GFS). Forecast models are reinitialized every six hours, (analysis-hours = 0000, 0600, 1200, and 1800 UTC) using information from the global weather station network and satellite derived products. At each analysis-hour the CFSR includes both the forecast data predicted from the previous analysis-hour, as well as the data from the analysis utilized to reinitialize the forecast models. This dataset contains historic expected precipitation and temperatures each hour for any land location in the world. Moreover, since the precipitation is updated in near-real-time every 6 hours, these data can provide real-time estimates of precipitation and temperature for hydrologic forecasting.

There are many studies conducted in the watersheds of Ethiopia using CFSR data set for its applicability for hydrological predications (Dile et al, 2014; Worglul, 2015). Dile et al (2014) suggested that the use of CFSR data set without model calibration showed reasonable performance at Gilgel Abay and Gumera River gauging stations at a monthly time step. He used two performance indicators, Nash–Sutcliffe Efficiency (NSE) and Percentage Bias (PBIAS) coefficients. His research found an NSE value of more than 0.75 showed the very good performance of the model in these gauging stations. The PBIAS value for Gilgel Abay also indicated very good performance, while the PBIAS value for the Gumera showed good model performance. Dile et al (2014) in his study he concluded that the quality of data in the case of CFSR data is higher than conventional data which sometimes hard to find in the required quality and proper allocation especially in developing countries. In addition, when the conventional data exist, it may be unreliable because of gaps and random errors. Vinayak S. et al (2009) in their study found that there is high underestimation of rainfall amounts in tipping bucket rain gauges under different rainfall intensities. In addition, the use of CFSR data set in areas where there is data scarcity has an advantage over conventional data set in that it provides complete sets of climatic data. This allows the flexibility to apply different functions pertaining to hydrological models. For example, with the conventional weather, we were limited to using the Hargreaves method to calculate potential evapo-transpiration because this method only requires maximum and minimum temperatures to calculate potential evapo-transpiration. However, availability of wind speed, relative humidity, and solar radiation data in the CFSR weather provides the flexibility to use Penman-Montieth and Priestley-Taylor methods. All in all, while hydrological model simulations should use high-quality observed weather data when available, CFSR weather is a viable option for simulating the hydrology of an area in data-scarce regions (Ibid).

Tewodros W. et al (2017) studied on the evaluation of different precipitation reanalysis datasets over Ethiopia using observed gauged data for a period of 33 years. He pointed out that water resources management and hydrologic modeling studies are limited in Ethiopia mainly because of the unavailability of long term observed data for hydrologic variables. He then, indicated the use of reanalysis data set as a replacement of observed data. He selected 16 representative administrative regions of Ethiopia to collect gauged precipitation data for the analysis. On his evaluation, CFSR data showed a more closed behavior to the observed rainfall distribution in all station in annual rainfall characteristics. However, the study found a poor performance of reanalysis data sets in general in replicating the observed trends in annual rainfalls. The study indicated that observed daily rainfall shown a high correlation value with CFSR rainfall data in most of the stations taken for the investigation (Ibid).

According to the study conducted by Daniel R. (2013) in simplifying watershed modeling using CFSR data, he found a better simulation result than conventional weather data in Gumera watershed in Ethiopia. He indicted a very good performance of a stream simulation with NSE result of 0.71using CFSR data and 0.68 applying the conventional weather data. The paper concluded the reason for the good performance of CFSR data as the weather data are effectively averaged over spatial scales that are more similar to many watershed extents; or, at least more similar than a typical point measurement of a weather station is to a watershed. He further indicated that although most hydrology textbooks note that the magnitude of point rainfall needs to be adjusted when considering the rainfall over a larger surrounding area (Citing Dingman, 2002) few modelers do this explicitly, and often account for these differences during model calibration. Using the spatial CSFR data, such adjustments are not needed. As a result of the difference in spatial scales between CFSR data and weather station data direct comparisons between the two give little correlation. This is not surprising, and indeed has been noted in several other studies. The study cited Vasiloff et al. (2009) and the study point out that wind, hail, missing gauge data, combined with storm paths make comparisons of weather station data against even much higher resolution radar and satellite precipitation products hard. It also cited Mehta et al. (2004), and it demonstrated that weather gauges located closer together than the resolution of the CFSR have a low correlation with each other ($r^2 < 0.3$). However, when the CFSR data are developed there are automatic comparisons between CFSR and the ground based weather data, which ensures some level of agreement (Ibid).

Table 2-2-2. Reanalysis dataset types available

Reanalysis Dataset (CISL ID)	Date Range	Time Step	PPT Field	Resolution	Coverage
NCEP/NCAR (ds090.0)	1948-2010	6hr	PPT Rate	2.5 °	Global
NCEP/DOE R2 (ds091.0)	1979-2012	6hr	PPT Rate	1.875 ° (~209km)	Global
NCEP N. American Regional (ds608.0)	1979-2012	3hr	PPT Rate	~32km	North America
NCEP 51-Year Hydrological (ds607.0)	1948-1998	3hr	Total PPT	0.125°	Continental US

ECMWF 15 Year (ds115.5)	1979-1993	6hr	Strat. + Conv. PPT	1.125^{o}	Global
ECMWF 40 Year (ds117.0)	1957-2002	6hr	Strat. + Conv. PPT	1.125^{o}	Global
ECMWF Interim (ds627.0)	1979-2012	6hr	Strat. + Conv. PPT	0.703^{o}	Global
CFSR (ds094.1)	1979-present	1hr	PPT Rate	0.205^{o} (~22km)	Global
Japanese 25-Year (ds625.0)	1979-2011	6hr	Total PPT	1.125^{o}	Global

NCEP/NCAR is the National Centers for Environmental Prediction
DOE is the Department of Energy.
PPT Rate is the precipitation rate.
Strat. + Conv. refers to strati form plus convective forms of precipitation.
ECMWF is the European Centre for Medium-Range Weather Forecasts

According to the study conducted by Tissa et al (2016) applying CFSR data for discharge simulation in a watershed found in India, he concluded that the number of gauge stations available in the given watershed has impact on accurate estimation of discharge from the watershed. He found that peak discharges estimations are similar in both data sets when the number of rain gauges available is greater than six. Significant decrease in peak discharge estimation observed when the number of gauge stations decrease from six. Hence, he indicated that if sufficient number of rain gauges is available, it is advisable to use the data collected from the stations. However, CFSR data prove better estimate when the number of stations available are 3 or less. For un-gauged watersheds and the watersheds with less number of rain gauges CFSR data gives more reliable results. This is because CFSR data is prepared from the weather data which are effectively averaged over spatial scales (Ibid).

Worqlul et al (2015) conducted a study for comparison of different gridded data sets with ground observed data as an input for hydrological models in data scarce watershed at Upper Blue Nile Basin, Ethiopia. He indicated that CFSR data set captured the gauged rainfall for Gilgel Abay and Main Beles with R^2values of 0.92 and 0.90, respectively. The study also indicated that 75% gauged areal rainfall occurred during the period from June through September and CFSR data hold about 80 % of the data within this period interval. Simulating the stream discharge of the two basins using PED model resulted a good fit between CFSR and observed flow data with very good NSE Values. The CFSR data captured 73% of the flow variation in the daily flow analysis of the observed flow and it is nearly similar with the gauged data which captured 82% of the flow variation. The simulated data for the calibration period using the gauged rainfall and CFSR indicated a fair to good performance with a daily NSE performance indicator equals to0.81 and 0.72 for Gilgel Abay and 0.64 and 0.61 for Main Beles, respectively, and with a reasonable $R^2$and PBIAS values. The simulation for both gauged rainfall and CFSR data captured well the base flow, the rising and recession limb of the hydrograph. The peak flow is better captured by the CFSR data than the gauged rainfall although both simulation by gauged and CFSR rainfall underestimate very high single peaks that are commonly caused by extreme high rainfall events. For the study period, in the Gilgel Abay watershed there are 505 days with observed flow above 200 m^3s^{-1}, the simulation by the CFSR rainfall estimate has captured 340 events and the gauged rainfall has captured 235 events (Ibid).

2.2.4. Bias Correction And Performance Evaluation

Throughout Africa the most important meteorological parameter is rainfall. This is because the economy of the continent is largely dependent on rain fed agriculture and in many areas rainfall is marginal with respect to crop production. Anomalously low or high rainfall can give respectively to drought or floods, both with disastrous economic and humanitarian consequences (Washington et al., 2006 cited in G. T. Diro et al, 2008). Belete B. et al (2016) mentioned that Rainfall is the core element of the hydrological cycle that drives energy circulation in the atmosphere. The spatial pattern of rainfall fields are also the fundamental inputs of hydrological models and stream flow analysis (Ibid). He further mentioned studies which highlight the importance of the spatial and temporal variation of precipitation which in turn proved to affect the accuracy of runoff prediction in gauged and un-gauged catchments. Moreover, spatial and temporal variability of rainfall affects the spatial and temporal distribution of hydrological variables in the system (Ibid).

Precipitation data are generally recognized as the most important driving data for hydrologic models. However, constrained by the sparse distribution of observation stations, the applicability of such models is limited. With development of modern observation and massive computing technologies, the estimation of precipitation based on combination of multi-source data (historical observed, radar and satellite) has become a feasible means for extending model applications(Yan Yang et al, 2014). Berhanu B. et al (2016) mentioned (Lemma, 2013)and indicated that although more than 1200 rainfall gauging stations are available in Ethiopia they do not serve well the demands of hydrological models due to their sparse location, gaps in temporal data coverage and quality of rainfall data. Moreover, accessibility of daily rainfall data for large river basins and regional analysis is limited by the rules and regulations of the National Meteorological Authority (Ibid). Many studies suggested the use of satellite sourced rainfall data for the benefit of predicting runoff and hydrological variables in the case of data scarce watersheds (Berhanu B. et al, 2016; G. T. Diro, 2008; Yan Yang et al, 2014). However, Berahnu et al (2016) mentioned (Vergaraet al., 2014) and indicated that most of the remotely sensed rainfall data lack accuracy due to uncertainties caused by the use of statistical methods and sampling. Yan Yang et al (2014) also indicated that errors in gridded precipitation data have a high probability for inducing errors and uncertainties in hydrologic simulations.

(Assayew N., 2016) tries to summarize the causes of magnitude differences between satellites sourced data and meteorological measured rain gauge data. As his summary the causes are 1) satellite rainfall estimates are indirect estimates of rainfall from cloud properties observed space in case of Geo-stationary satellites and derived from microwave emissions from raindrops and scattering from ice in orbiting Satellites 2) assumptions like the surface emissivity, neglecting evaporation below clouds, and empirical relationships are the driving factors of error 3) satellite rainfall estimates errors are caused by various factors like sampling frequency, field of view of the sensors, and uncertainties in the rainfall retrieval algorithms. Habib et al (2014) cited in Assayew N. (2016) identify errors can be random or systematic. It is the systematic error that is commonly referred to as bias and reflects errors which are systematically distributed over time and space. Bias in satellite rainfall products can cause large uncertainties in hydrological modeling.

(Margaret Wambui, 2018) suggested to reduce the errors these satellite products hold to make them more representative with the local rainfall variability as they often exhibit large discrepancies with ground measurements. Although rain gauge data have low spatial distributions, their direct way of measuring rainfall are still vital as a reference to the local rainfall variability. For better representations of local rainfall processes, the inclusion of all available quality controlled rain gauge data merged with satellite products can enhance the products' future applications (Ibid).

Berhanu B. et al (2016) indicated that many studies tried to evaluate CFSR by comparing its accuracy to ground observed precipitation with different spatial and temporal scale. Different methods have been proposed to reduce satellite rainfall estimates errors. A study by (Margaret Wambui, 2018) applied bias correction using empirical cumulative distribution (CDF) maps on a seasonal basis for hydrological applications in the upper Blue Nile in Ethiopia. To reduce temporal rainfall variability, a seasonal timescale was utilized. However, the study indicated that in high elevated areas, areas near inland water bodies and those with maritime influences high rainfall variability's are experienced. Hence the study identified the choice of temporal scale which may differ from place to place. It is worth noting that the effectiveness of bias correction on rainfall products may also differ from location to location and consideration of spatial scale is of great importance.

Berhanu B. et al (2016) showed that CFSR data set is a good alternative dataset for large scale hydrological modeling. Hence, his study indicated that a bias correction method which focuses mainly on the magnitude of the data is only important rejecting the pattern and trend of the dataset based on the research finding of Worqlul et al (2014) which reported CFSR precipitation can reproduce the observed rainfall pattern in the Blue Nile, Ethiopia, but it overestimated or underestimated the observed values. From among the various arrays of rainfall data adjusting techniques, most of which are statistical, the linear scaling (LS) bias correction method was selected for this study as it aims to match the monthly mean of corrected values perfectly with that of the observed ones (Lenderink et al., 2007 cited in Berhanu et al, 2016). It operates with monthly correction values based on the difference between observed and raw data. The change factor for precipitation is a multiplier that is computed from the ratio of the monthly mean of the observed to the raw dataset:

$$P_{d,cor} = P_{d,raw} * \left(\frac{\mu\ (P_{m,obs})}{\mu\ (P_{m,raw})}\right) \text{............... Equation 2.2.3.1}$$

Where $P_{d,cor}$ is the corrected daily precipitation and $P_{d,raw}$ is the daily raw precipitation data from CFSR. In this case, ($P_{m,\ obs}$) is the long-term mean monthly rainfall of observed data, and ($P_{m,raw}$) is the long-term mean value of the monthly/yearly raw rainfall data (Ibid).

(G. H. Fang et al, 2015) mentioned Schmidli et al (2006) and presented another bias correction method known as Local intensity scaling (LOCI) of precipitation. The LOCI method corrects the wet-day frequencies and intensities and can effectively improve the raw data which have too many drizzle days (days with little precipitation). It normally involves two steps: firstly, a wet-day threshold for the mth month P_{thres},m is determined from the raw precipitation series to ensure that the threshold exceedance matches the wet-day frequency of the observation; secondly, a scaling factor

$$S_m = \left(\frac{\mu\ (\frac{P_{obs,m,d}}{P_{obs,m,d}} > 0)}{\mu\ (\frac{P_{raw,m,d}}{P_{raw,m,d}} > P_{thres,m})}\right) \text{............. Equation 2.2.3.2}$$

is calculated and used to ensure that the mean of the corrected precipitation is equal to that of the observed precipitation:

$$P_{cor,m,d} = \begin{cases} 0, & if P_{raw,m,d} < P_{thres,m} \\ P_{raw,m,d}\ x\ \boldsymbol{S}_m & otherwise \end{cases} \text{ ... . Equation 2.2.3.3}$$

The other method of bias correction according to (G. H. Fang et al, 2015) is Power transformation (PT) of precipitation. While the LS and LOCI account for the bias in the mean precipitation, it does not correct biases in the variance. The PT method uses an exponential form to further adjust the standard deviation of precipitation series. Since PT has the limitation in correcting the wet-day probability (Teutschbein and Seibert, 2012 cited in G. H. Fang et al, 2015), which was also confirmed in our study (not shown), the LOCI method is applied to correct precipitation prior to the correction by PT method.

Therefore, to implement this PT method, firstly, we estimate b_m, which minimizes

$$f(\, b_m) = \frac{\sigma\,(\, P_{obs,m})}{\mu\,(\, P_{obs,m})} - \frac{\sigma\,(P_{LOCI,m}{}^{b_m})}{\mu\,(P_{LOCI,m}{}^{b_m})} \text{ ……….. Equation 2.2.3.4}$$

Where b_m is the exponent for the mth month, σ (.) represents the standard deviation operator, and $P_{LOCI,m}$ is the LOCI corrected precipitation in the mth month. If b_m is larger than1, it indicates that the LOCI-corrected precipitation underestimates its coefficient of variance in month m. After finding the optimal b_m, the parameter

$$S_m = \left(\frac{\mu(P_{obs,m})}{\mu\,(P_{LOCI,m}{}^{b_m})} \right) \text{ .... Equation 2.2.3.5}$$

is then determined such that the mean of the corrected values corresponds to the observed mean. The corrected precipitation series are obtained based on the LOCI corrected precipitation $P_{cor,m,d}$:

$$P_{cor,m,d} = \; S_m x P_{LOCI,m,d}^{b_m} \text{ ……… Equation 2.2.3.6}$$

In General the different bias correction for precipitation and temperature can be summarized as:

Table 2-2-3-1. Bias correction methods for precipitation and temperature

Bias correction for precipitation	Bias correction for temperature
Linear scaling (LS)	Linear scaling (LS)
Local intensity scaling (LOCI)	Variance scaling (VARI)
Power transformation (PT)	Distribution mapping for temperature using Gaussian distribution (DM)
Distribution mapping for precipitation using gamma distribution (DM)	
Quantile mapping (QM)	

The performance evaluation of these corrected precipitation datasets was tested with frequency-based indices and time series performances against observed precipitation data. The frequency-based indices include mean, median, standard deviation (SD), 90thpercentile, probability of wet days and intensity of wet days. The time series-based metrics include the Nash–Sutcliffe measure of efficiency (NSE) (Nash and Sutcliffe, 1970), the root-mean-square error (RMSE) and the percent bias (PBIAS) (Berhanu B. et al, 2016).

The Nash-Sutcliffe efficiency (NSE) is a normalized statistic that was calculated to determine the relative magnitude of the residual variance compared to the measured data variance. NSE indicates how well the plot of observed versus simulated data fits the 1:1 line. NSE is computed by:

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(P_{obs}^{i} - P_{cor}^{i}\right)^2}{\sum_{i}^{n}(P_{obs}^{i} - P_{obs}^{mean})^2} \text{ .... Equation 2.2.3.7}$$

Where $P_{i,obs}$ is the i^{th} observation for the precipitation being evaluated, $P_{i,cor}$ is the i^{th} corrected value for the precipitation being evaluated, P mean is the mean of observed data for the precipitation being evaluated, and n is the total number of observations. NSE indicates how well the simulation matches the observation, and it ranges between –∞ and 1.0, with NSE=1 indicating a perfect fit. The higher this value, the more reliable is the model (Ibid).

The PBIAS is the other performance evaluation index and given by (Belete B. et al, 2016):

$$PBIAS = \frac{\sum_{i=1}^{n}\left(P_{obs}^{i} - P_{cor}^{i}\right)*100}{\sum_{i}^{n}(P_{obs}^{i})} \text{ .... Equation 2.2.3.8}$$

The Root Mean Square Error (RMSE) is given by (Belete B. et al, 2016):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_{obs}^{i} - P_{cor}^{i})^2}{N}} \text{ .... Equation 2.2.3.9}$$

PBIAS measures the average tendency of the simulated data to their observed counterparts. Negative values indicate an overestimation (i.e., the simulated, dataset is higher than the observed dataset), while positive values indicate an underestimation (the simulated dataset lower than the observed dataset). The optimal value of PBIAS is 0.0, with low-magnitude values in both directions, possibly indicating accurate model simulations (Belete B. et al, 2016).

G. H. Fang et al (2015) mentioned mean absolute error (MAE) for corrected precipitation performance indicator. It is given as

$$MAE = \frac{\sum_{i=1}^{n} |P_{obs}^{i} - P_{cor}^{i}|}{n} \text{ .... Equation 2.2.3.10}$$

E. Tarawneh et al (2016) indicated the root mean square error standard deviation ration (RSR) as performance evaluator and it is given as:

$$RSR = \sqrt{\frac{\sum_{i=1}^{n}(P_{obs}^{i} - P_{cor}^{i})^2}{\sum_{i=1}^{n}(P_{obs}^{i} - P_{obs}^{mean})^2}} \text{ .... Equation 2.2.3.11}$$

Daniel de et al (2014) cited Moriasi et al (2007) and indicated recommended performance ratings (Table 2-2-3-2) below.

The guidelines used for evaluation is based on the performance ratings (Table 2-2-3-2). If one or more of the ranges of RSR, NSE and PBIAS indicated an "unsatisfactory" result, then the performance is determined unsatisfactory (Ibid).

Table 2-2-3-2. Performance Rating values for estimating methods

Performance Rating	RSR	NSE	PBIAS %	Grading for each
Very good	$0.00 \le RSR \le 0.50$	$0.75 < NSE \le 1.00$	$PBIAS < \pm10$	3
Good	$0.50 < RSR \le 0.60$	$0.65 < NSE \le 0.75$	$\pm10 \le PBIAS < \pm15$	2
Satisfactory	$0.60 < RSR \le 0.70$	$0.50 < NSE \le 0.65$	$\pm15 \le PBIAS < \pm25$	1
Unsatisfactory	$RSR > 0.70$	$NSE \le 0.50$	$PBIAS \ge \pm25$	Unsatisfactory

2.3. Hydrological Models

According to Gayathri K Devi et al. (2015), a model is a simplified representation of real world system. The best model is the one which give results close to reality with the use of least parameters and model complexity. Models are mainly used for predicting system behavior and understanding various hydrological processes. A model consists of various parameters that define the characteristics of the model. A runoff model can be defined as a set of equations that helps in the estimation of runoff as a function of various parameters used for describing watershed characteristics. The two important inputs required for all models are rainfall data and drainage area. Along with these, water shed characteristics like soil properties, vegetation cover, watershed topography, soil moisture content, characteristics of ground water aquifer are also considered. Hydrological models are now a day considered as an important and necessary tool for water and environment resource management.

Hydrology has a long history dating back to several millennia (Biswas 1970). However, the birth of hydrologic modeling can be traced to the 1850s when Mulvany (1850) developed a method for computing the time of concentration and hence the rational method for computing peak discharge which is still used for urban drainage design, Darcy (1856) who conducted experiments on flow-through sands and developed what is now referred to as Darcy's law which laid the foundation of quantitative groundwater hydrology, and Fick's first law which states that under steady-state conditions the diffusive flux is proportional to the concentration gradient(spatial) which laid the foundation of water quality hydrology. Hydrological modeling involves formulating a mathematical model which represents hydrologic processes and the interaction between them. So in hydrologic modeling the inter-relationship of soil, water, climate, and land use are considered and represented through mathematical abstraction (Gosain et al., 2009).This can be challenging because it involves highly nonlinear processes, complex interactions and high spatial variability at basin scale.

Table 2-3-1. Selected physically based hydrologic model, spatial description and discretization type

Model Acronym	Model Definition	Semi/Fully Distributed	Discretization Type *
TOPMODEL	Topography based hydrological MODEL	Semi	HRU
WATBAL		Semi	OG
SHE	European Hydrologic System	Fully	OG
ISBA	Interaction Soil Biosphere Atmosphere	Fully	OG
IHDM	Institute of Hydrology Distributed Model	Fully	HRU
THALES		Fully	IE
SLURP	Semi-distributed Land Use-based Runoff Processes	Semi	GRU
MIKE SHE		Fully	OG
SWAT	Soil and Water Assessment Tool	Semi	HRU
WATFLOOD/ SPL9	Waterloo Flood Forecasting Model	Fully	OG
HRCDHM	Hydrologic Research Center Distributed Hydrologic Model	Semi	HRU
DPHM-RS	Semi-Distributed Physically based Hydrologic Model using Remote Sensing and GIS	Semi	HRU
R.WATER.FEA		Fully	IE
tRIBS	TIN-based Real-time Integrated Basin Simulator	Fully	TIN
TOPNET		Semi	HRU
MISBA	Modified Interaction Soil Biosphere Atmosphere	Fully	OG
LISTFLOOD		Fully	OG
HydroGeoSphere		Fully	OG
PAWS	Process-based Adaptive Watershed Simulator	Fully	OG
CREST	The Coupled Routing and Excess Storage	Fully	OG

Source: Zahidul (2011)

*Abbreviation used for Discretization type: OG= Orthogonal Grid, HRU= Hydrologic Response Unit, GRU= Grouped Response Unit, IE=Irregular Elements, TIN=Triangulated Irregular Network.

Hydrologic models can be classified according to the physical processes involved in modeling as conceptual and physically based (Refsgaard, 1996). In conceptual models each of the hydrologic processes, that we read into our observations of the catchment, are represented by simplified mathematical relationships, where as in physically based model the detail physical processes can be represented in a deterministic way by representations of mass, momentum and energy conservation (Refsgaard, 1996). According to the spatial description of the watershed process, hydrologic models can be classified as lump and distributed models. In a lumped model the spatial variability of watershed characteristics are ignored, while in a distributed model the spatial variability of vegetation, soil, topography, etc are taken into account. The conceptual models are usually lumped while the physically based model in practice has to be distributed in manner (Refsgaard, 1996). In physically based hydrologic modeling the hydrologic process of water movement are modeled either by the finite difference approximation of the partial differential equation representing the mass, momentum and energy balance or by empirical equations (Abbott et al., 1986b). Typically the primary components of hydrologic cycle related to the land phase are taken into consideration. These are: interception, snowmelt, evapo-transpiration, sub -surface runoff, groundwater flow, surface runoff and channel routing.

2.3.1. SWAT Model

2.3.1.1. **Model Description**

The Soil and Water Assessment Tool (SWAT) has been applied in many studies around the world, especially in research related to include hydrology, erosion, climate, soil, temperature, plant growth, nutrients, pesticides and land management. Stream processes considered by the model include water balance, routing, sediment, nutrient and pesticide dynamics (Gokhan Cuceloglu et al, 2017).The model was selected because of its robust approach of soil water balance at the watershed scale. The SWAT model has been used to study the impacts of environmental change in several parts of the world. SWAT is a process-based model that operates at a daily time step and uses a modified Soil Conservation Service-Curve Number (SCS-CN) from the United States Department of Agriculture Soil Conservation Service (USDA-SCS) to estimate surface runoff, and peak runoff rates using a modified rational method (W. B. Mbungu et al, 2016).

The model was designed to assess long term impact of land management on water balance, sediment transport and non-point source pollution in river basins. In the SWAT model, a watershed is divided into homogeneous hydrological response units (HRUs) which are a combination of land use, management practices, topographical and soil characteristics. The HRUs are represented as a percentage of the sub watershed area and may not be contiguous or spatially identified within a SWAT simulation. Alternatively, a watershed can be subdivided into only sub watersheds that are characterized by dominant land use, soil type, and management. Water balance is the driving force behind all the processes in SWAT because it impacts plant growth and the movement of sediments, nutrients, pesticides, and pathogens. Simulation of watershed hydrology is separated into the land phase, which controls the amount of water, sediment, nutrient, and pesticide loadings to the main channel in each sub basin, and the in-stream or routing phase, through the channel network of the watershed to the outlet. Plant growth is estimated under optimal conditions, and then computes the actual growth under stresses inferred by water and nutrient deficiency. Subdividing the watershed allows users to analyze hydrologic processes in different sub-watersheds within a larger watershed and under localized land use management impacts (W. B. Mbungu et al, 2016).

SWAT is a comprehensive hydrologic model that allows for numerous physical processes to be simulated in a watershed. These processes may be separated into two coarsen divisions of the hydrologic cycle: the land phase and the routing phase. These divisions include important processes such as precipitation, surface runoff, evapo-transpiration, groundwater flow, snowmelt, and flood routing. The model is physically based; meaning each hydrologic process is directly modeled by SWAT, as opposed to empirical, regression-based equations. Thus, specific information about weather, soil properties, topography, vegetation, and land management practices occurring in the watershed is required (most of which can be directly measured in the field). This deterministic approach allows the user to study the relative impact of alternative input data on particular variables of interest. SWAT is driven by a water balance equation which relates individual components of the hydrologic cycle. Additional details including specific equations associated with the water balance and the individual hydrologic processes may be found in the SWAT Theoretical Documentation, Version 2009 (Aaron Havel, 2015).

The SWAT model uses Manning's equation to define the rate and velocity of stream flow. Both routing options within SWAT, variable storage and Muskingum, are distributed flow routing models, meaning variables may be determined as functions of space and time. Both of these methods are based on variations of the kinematic wave model. SWAT assumes a trapezoidal channel shape with 2:1 side slopes for stream flow routing calculations. The user may enter the width and depth of the channel when filled to the top of the banks. The user is also required to enter the channel length, which in many cases is quite long, depending on the stream initiation threshold and sub-basin outlet locations (Aaron Havel, 2015).

The SWAT model can be applied to support various watershed and water quality modeling studies. Examples of such studies include the following:

- National and regional scale water resource assessment considering both current and projected management conditions.

- Bosque River TMDL in Erath County, Texas. The project determined sediment, nitrogen and phosphorus loadings to Lake Waco from various sources including dairy waste application areas, waste treatment plants, urban areas, conventional row cropsand rangeland. Numerous land management practices were simulated and analyzed (Saleh et al., 2000 cited in M. WINCHEL et al, 2013)

- Poteau River TMDL in Oklahoma/Arkansas. This project assessed sediment, nitrogen and phosphorus loadings to Wister Lake and dissolved oxygen, temperature, algae, and CBOD in the river. Management scenarios regarding poultry waste were analyzed (Srinivasan et al., 2000 cited in M. WINCHEL et al, 2013).

- DDT in the Yakima River basin, Washington. SWAT was used to simulate past and future sediment contamination by DDT in the Yakima River basin (M. WINCHEL et al, 2013).

- The EPA office of pesticide registration is evaluating SWAT for use in landscape/watershed scale evaluation for pesticide registration.

- SWAT is being used extensively in the U.S. and Europe to assess the impact of global climate on water supply and quality (Rosenberg et al, 1999 cited in M. WINCHEL et al, 2013).

2.3.1.2. Advantages and Disadvantages

First of all, SWAT is an open source tool and detailed online documentation, user groups, video tutorials, international conferences and a unique literature database (more than 2700 papers) are available. This all makes the tool user-friendly, which can explain, at least partly, the fact that it is one of the best known and most widely used tools to develop water quality models at the watershed scale (Cambien,2016-2017).

Secondly, the tool is continuously improved, supported by the core developmental team and as a response to shortcomings demonstrated by the many users (Cambien,2016-2017). This results in the development of new tools, e.g. GIS interface tools, pre- and post-processing tools and statistical evaluation tools (Gassmann et al., 2010). In addition, a trend to interface SWAT with other environmental or economic models enlarges its application range (Gassman et al., 2007). Moreover, a complete revision of SWAT will be available soon, solving certain limitations of the tool and enhancing its capabilities. Innovations of this revision are the new approach for watershed discretization and a modification of the input files (Cambien, 2016-2017).

A third advantage is its comprehensive code. SWAT is proven to be an effective and flexible tool for a wide range of applications, watershed scales and environmental conditions (Gassman et al, 2014; Krysanova and White, 2015; Tuppad et al., 2011). Moreover, the semi-distributed structure makes the model computationally efficient and enables to generate spatially explicit outputs. Finally, the tool is suitable for large, complex watersheds (Gassman et al., 2014).

However, every tool has its shortcomings and these are often linked with its advantages. The constant improvements, for example, have led to a difficult code and a high number of parameters, requiring expertise to run the model and complicating the calibration process (Arnold et al., 2012a; Vigerstol and Aukema, 2011). Additionally, the tool is highly data intensive. Although SWAT is said to run on readily available input data this is not always the case, especially in developing countries. Certainly the data accuracy and precision might be an issue, as expressed by the rule “garbage in is garbage out” (Estrada et al., 2009; Gassman et al., 2007; Querner and Zanen, 2013).

Another limitation is the use of non-interacting HRUs. Because the HRUs are not linked to each other, routing processes of flows and pollutants and pollutant attenuation within the sub watershed are lacking (Arnold et al., 2010; Krysanova and Arnold, 2008). Moreover, the semi-distributed structure impedes to generate spatially explicit outputs at the HRU level (Cambien, 2016-2017).

Furthermore, SWAT is said to be a physically based model but also empirical equations like the curve number method are implemented (Bauwe et al., 2016). The curve number method has been applied successfully and enables the adaptation to the study-specific conditions during calibration (Arnold et al., 2012a). On the other hand, this controversial method is not developed for the application at individual HRUs, does not account for all runoff generating processes and causes poor results in some cases (Cambien, 2016-2017; Gassman et al., 2007).

Lastly, there are some limitations with respect to the simulation of pesticide dynamics. First of all, pesticide input into the rivers via point sources, drift and groundwater up flow is not taken into account. Especially with respect to point sources this is an important disadvantage, as they can contribute largely to the occurrence of pesticides in surface water (section 2.2). To tackle this issue, Holvoet (2006) and Gevaert et al. (2008) extended the SWAT code to implement point sources and droplet drift. This modification, however, has not been incorporated into the standard SWAT versions (Cambien, 2016-2017). In addition, Cambien (2016-2017) reports flow partitioning problems and the need to decouple pesticide transport from tile drainage and lateral flow. The last limitation is the fact that SWAT can only route one pesticide at a time (Neitsch et al., 2011).

2.3.1.3. Runoff Generation

Surface runoff will occur when the precipitation rate exceeds both infiltration capacity and surface detention capacity. It is partly governed by factors such as precipitation intensity and duration that are external to the slope system.

However, the most important determining factors are those of the slope system itself such as soil type, the amount and type of vegetation, slope form and angle, and relative position on the slope. When the rainfall intensity is greater than infiltration capacity overland flow will occur and will move down slope increasing in volume as it moves. There are empirically (SCS,1972) and physically based infiltration models (Kinde, 2015).

Physically based infiltration models can be transferred to other regions but they need boundary conditions and high computation time. Empirical Hydrological methods are transferable through calibration on another scale and climate region (Dilnessaw, 2006 cited in Kinde, 2015)

It was the product of more than 20 years of studies involving rainfall-runoff relationships from small rural watersheds across the U.S. The equation was developed to provide a consistent basis for estimating the amounts of runoff under varying land use and soil types (Evans, 1988 cited in Kinde, 2015). According to (Catherine Kuhn, 2014) the model uses a master water balance approach to compute runoff volumes and peak flows. This is given by the following equitation:

$$SW_t = SW_0 + \sum_{i=1}^{t} (R_{day} - Q_{surf} - E_a - W_{seep} - Q_{gw}) \text{ ……. Equation 2-4-1}$$

Where SW_0 is initial soil water content and SW_t is the final soil water contents on day i. All other measurements are taken in millimeters and time (t) is in days. The equation subtracts all forms of water loss on day i from precipitation on day i (Rday) including surface runoff (Qsurf), evapo-transpiration (Ea), loss to vadose zone (Wseep) and return flow (Qgw) (Neitsch, Arnold et al. 2009). By manipulating this equation the model can predict changes in variables of interest like runoff and return flow (Catherine Kuhn, 2014).

According to (Dina K. et al, 2004) The Soil Conservation Service (SCS) curve number (CN) equation is used to estimate surface runoff. This method was developed from many years of stream flow records from agricultural watersheds in many parts of the United States. CN is a function of soil group, land cover complex, and antecedent moisture conditions. The curve number method was adopted in the SWAT model because it (1) is used widely throughout the United States, (2) has been tested on watersheds of varying sizes, and (3) requires only easily available input data. The SCS curve number method uses two equations for runoff. The first relates runoff to rainfall and retention parameter as:

$$Q = \frac{(R-0.2S)^2}{(R+0.8S)} \text{ , R> 0.2S ……. Equation 2-4-2}$$

Where

Q = daily surface runoff (in mm)

R = daily runoff (in mm)

S = retention parameter, the maximum potential difference between rainfall and runoff (in mm) starting at the time the storm begins.

The second equation relates retention parameter to curve number as:

$$S = 25.4\ (\frac{1000}{CN} - 10) \text{ …… Equation 2-4-3}$$

Where, CN= curve number ranging from $0 \leq CN \leq 100$

2.3.1.4. Model Input

The model used in this study was built using the SWAT (2012) version using ArcSWAT. Building a SWAT model requires availability of spatially distributed information on Digital Elevation Model (DEM), land cover and land use and soils. Data on climate and river discharge were also important for prediction of stream flow and calibration purposes. (Dina K. et al, 2004) summarizes the three main watershed level input files as follows: a) Watershed configuration file (fig.fig): The watershed configuration file contains information used by SWAT to simulate processes occurring within the HRU/sub-watershed and to route the stream flow and constituent loads through the channel network of the watershed, b) The master watershed file (file.cio): The master watershed file contains information related to modeling options, climate inputs, databases, and output specifications. Information in this file includes number of calendar years simulated, beginning year of simulation, the beginning and ending Julian day of simulation, and weather and rain station information. This file also contains links to files needed for the watershed definition and delineation, as well as links to the files holding the precipitation and temperature information needed for the simulation and c) The basin input file (basins.bsn): General watershed attributes are defined in the basin input file. These attributes control a diversity of physical processes at the watershed level. The attributes initially are automatically set to “default” values. Examples of attributes in the basins.bsn file are specification of the method used for estimating ET, initial soil water storage, and surface runoff lag time, and other parameters used in the SWAT simulation on the watershed scale. Users can use the default values or change them to better reflect conditions in a specific watershed.

2.3.2. Model Sensitivity Analysis, Calibration and Validation

SWAT input parameters are process based and must be held within a realistic uncertainty range. The first step in the calibration and validation process in SWAT is the determination of the most sensitive parameters for a given watershed or sub watershed. The user determines which variables to adjust based on expert judgment or on sensitivity analysis. Sensitivity analysis is the process of determining the rate of change in model output with respect to changes in model inputs (parameters). It is necessary to identify key parameters and the parameter precision required for calibration (Ma et al., 2000). In a practical sense, this first step helps determine the predominant processes for the component of interest. Two types of sensitivity analysis are generally performed: local, by changing values one at a time, and global, by allowing all parameter values to change. The two analyses, however, may yield different results. Sensitivity of one parameter often depends on the value of other related parameters; hence, the problem with one-at-a-time analysis is that the correct values of other parameters that are fixed are never known. The disadvantage of the global sensitivity analysis is that it needs a large number of simulations. Both procedures, however, provide insight into the sensitivity of the parameters and are necessary steps in model calibration (Arnold et al, 2012).

The second step is the calibration process. Calibration is an effort to better parameterize a model to a given set of local conditions, thereby reducing the prediction uncertainty. Model calibration is performed by carefully selecting values for model input parameters (within their respective un-certainty ranges) by comparing model predictions (output) for a given set of assumed conditions with observed data for the same conditions. The final step is validation for the component of interest (stream flow, sediment yields, etc.). Model validation is the process of demonstrating that a given site-specific model is capable of making sufficiently accurate simulations, although “sufficiently accurate” can vary based on project goals (Refsgaard, 1997). Validation involves running a model using parameters that were determined during the calibration process, and comparing the predictions to observed data not used in the calibration. In general, a good model calibration and validation should involve: (1) observed data that include wet, average, and dry years (Gan et al., 1997); (2) multiple evaluation techniques (ASCE, 1993; Legates and McCabe, 1999; Boyle et al., 2000); (3) calibrating all constituents to be evaluated; and (4) verification that other important model outputs are reasonable. In general, graphical and statistical methods with some form of objective statistical criteria are used to determine when the model has been calibrated and validated. Calibration can be accomplished manually or using auto calibration tools in SWAT (van Griensven and Bauwens, 2003; Van Liew et al. (2005) or SWAT-CUP (Abbaspour et al., 2007).

3. METHODLOLGY

3.1. Description of the Study Area

The area of study in this research is Wabi Shebelle River Basin (WSRB) in Ethiopia. The basin spread over three regional states namely; Oromia region in the North West part of the basin, Somali Region in the south East, and Harari region fully in the middle of the basin. Oromiya and Somali regional states cover about 38% and 60% of the basin area respectively. (Adane Abebe, 2009). The higher elevation area is found in Oromiya regional state which is the major contributor of the surface water of the basin.

Geographically, the area is bounded between latitudes 4° 55' 31'' N and 9° 34' 59'' N; and longitudes 38° 41' 38'' E and 45° 28' 11'' E with north to south average length of 530 km and east to west average width of 450 km having areal extent of about 202,697 sq. km out of 1,127,261 sq. km (the country landmass). It covers about 17% of the country landmass. (Tesema Kebede, 2015). It is bounded by Genale basin in the South West, rift valley in the West and North West, Awash basin in the North, Aysha Dewele in the North East, Ogaden in the East and Somalia in the South.

About 12% of the population is urban. The major accesses to the basin are a gravel and asphalt road along the highlands of the Bale Mountains and Addis Ababa respectively. The different governments in the last three decades have recognized the need for irrigation development. Accordingly about 75 modern schemes have been constructed. These cover 5,630 ha benefiting 21,246 households as per the data collected from Zonal offices in Phase I study excluding sites which do not have such details. The average size developed per scheme is about 100 ha (Ibid).

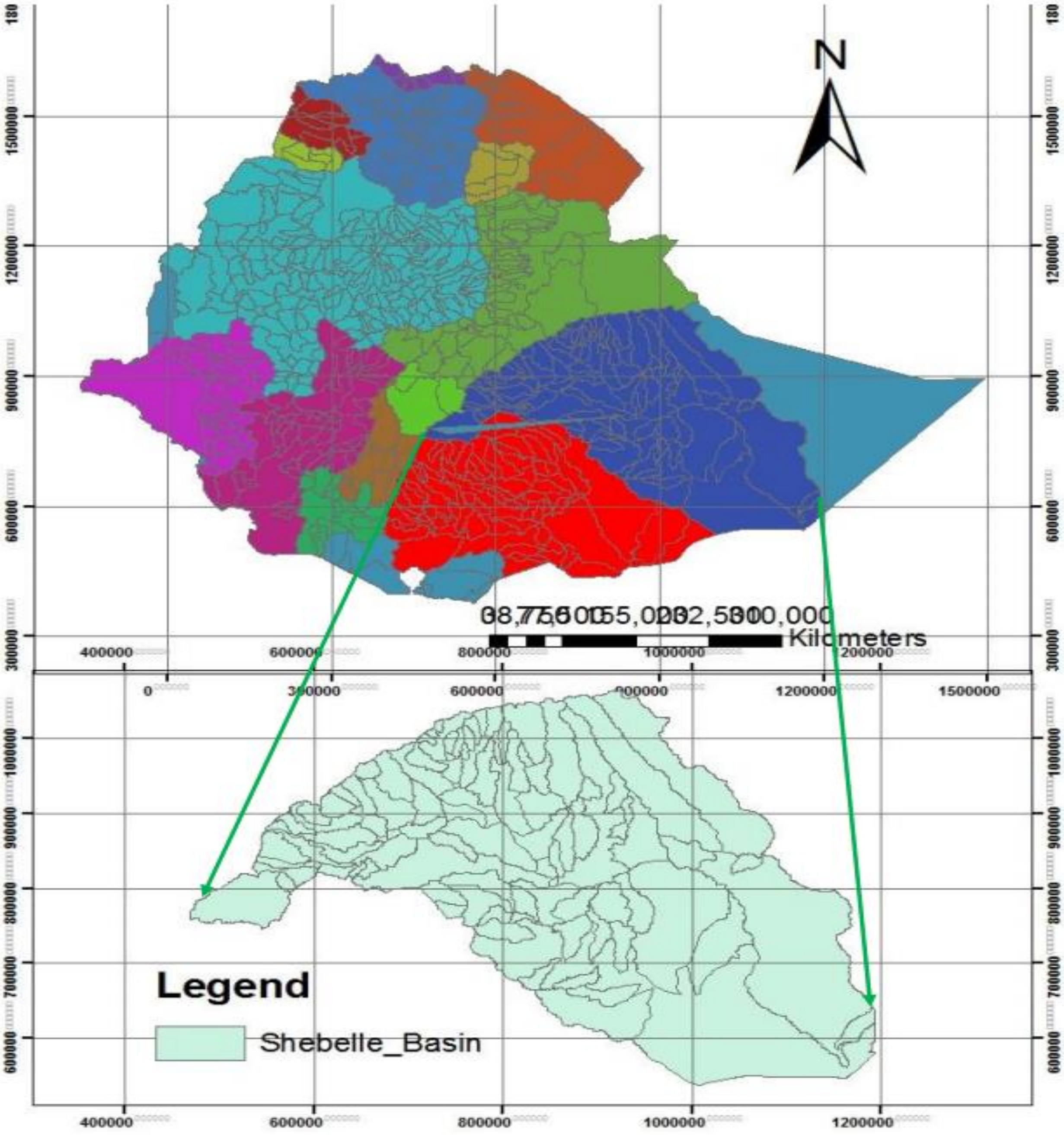


Fig 3-1-1. Location of Wabi-Shebele River Basin in Ethiopia

3.1.1. Topographic Feature

Ethiopia has four major physiographic regions namely the north western plateau, the south eastern plateau, the main Ethiopian Rift and the Afar depression. The study area, Wabi-Shebele River Basin, found in the south eastern plateau. The Arsi and Bale highlands, western and eastern Hararghe plateaus are the margin that forms the north and northwestern watershed divide of the basin (Tesema Kebede, 2015). Wabishebelle River emerges from the mountainous areas of the North Western borders of the river basin near a place called Hebeno. The altitude varies from 73 m above mean sea level (msl) to about 4216 mm on the highlands of Bale. The basin has great geographical diversity with high and rugged mountains, flat topped plateaus including river valleys and rolling plains. These major physical features are the result of uplift of the land mass and the emission of huge quantity of lava along extension and spectacular faulting. The basin consists of lowlands and highlands. The physical condition and variation in altitude have resulted in great diversity of climate, soil and vegetation (Tesema Kebede, 2015).

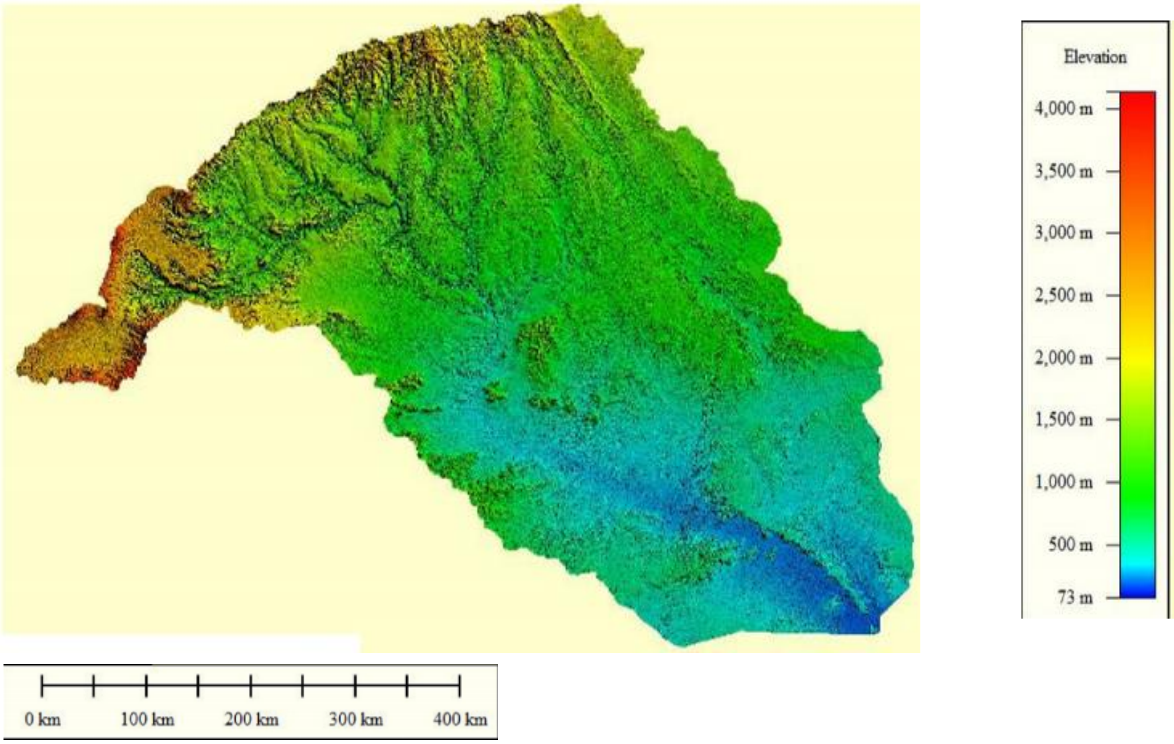


Fig 3-1-1-1. Topography of WSRB

3.1.2. Geologic feature

The area is dominated by Mesozoic sedimentary formations, to some extent there are also volcanic rocks at the North West of the basin and isolated ridges and hills within the sedimentary basin. Metamorphic rocks outcrops in a small extent at the northern part of the study area. Alluvial deposits are also distributed linearly along the Wabishebelle Erer and Fafen rivers and fan deposits of seasonal floods and stream beds .The volcanic rocks of Arsi-Bale basalt bordering the rift valley are highly fractured . Numerous springs outcrops along faults and fractures in this area and form substantial part s of the base flow of Wabishebelle River. The Southern part of the basin is over laid by thick gypsum and limestone. The water level monitoring for one hydro logic cycle on two wells at Gode showed that the water level is deep always lower than the river bed and the phreatic water level is practically the same during the hydrologic cycle and no interaction with Wabi Shebelle River water indicating the permeability is very low. On the other hand at Kelafo there is interaction of the alluvial ground water with Wabi Shebelle River. At Muslahil there is infiltration of the river flood waters in to the alluvial ground water. (Kinde Engdaw, 2016)

3.1.3. Land Use/Land Cover

WSRB has a small coverage of cultivated land on the North West part of the basin. Large part of the basin is covered with bush, wood or grass land. The northeastern portion of the basin is covered with shrubs and bush mainly utilized for pasture, however, crop cultivation is not common on this part of the basin. (Tesema, 2015)

The North West part has also a small coverage of shrub land. Patches of exposed rock or sand surface are found in parts of Bale and Hararghe lowlands in the southeast. Parts of central Arsi and northern Bale have afro-alpine and sub-afroalpine vegetation. These consist mostly of short shrub and heath vegetation used partly for sedentary grazing and browsing. Riparian woodland and bush land occur along the river banks and on floodplain sand are important in the semi-arid and arid parts of the basin where they are used for grazing and browsing and scattered seasonal crop cultivation on some of the flood plains.

The land use consists of large part of silvipastoral type. Areas of intensively cultivated land are found on the highlands of Arsi and parts of highland Harerge, and northern Bale. The major seasonal crops in the basin include maize, barley, wheat and sorghum while the perennial crops include coffee, chat and fruit trees (Kinde, 2015).

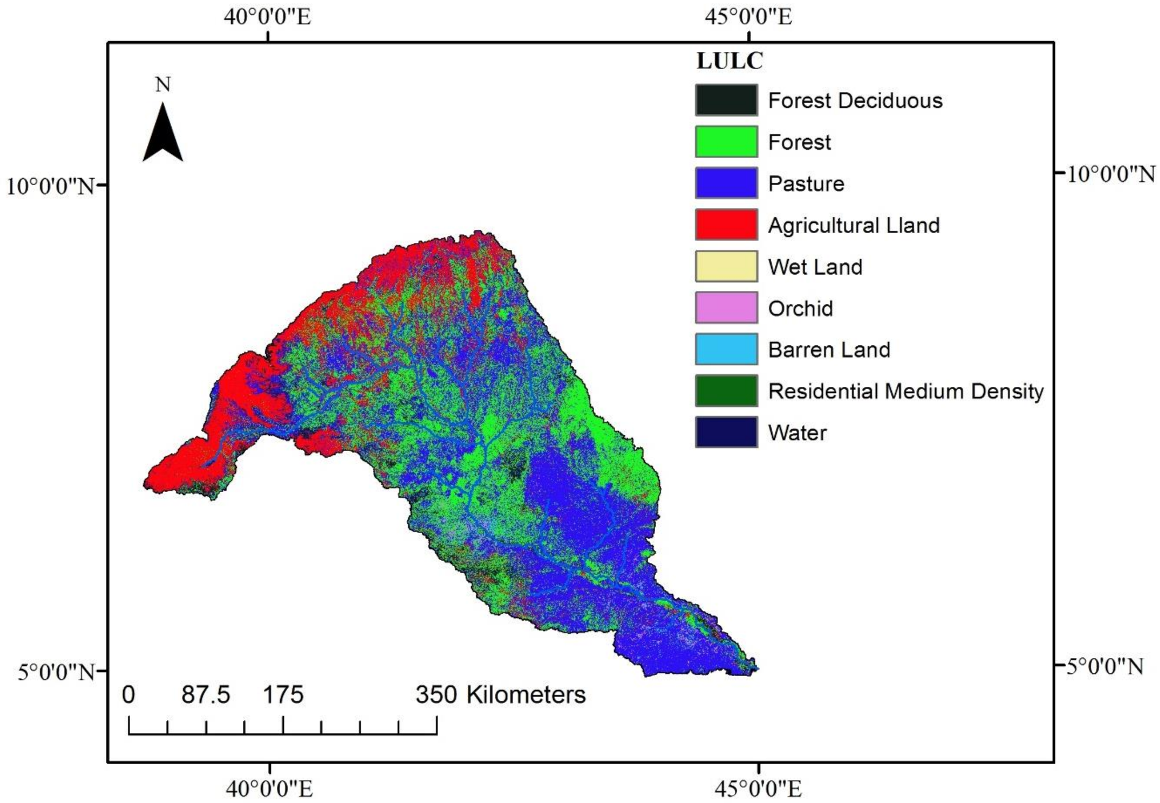


Fig 3-1-3-1. Land Use Land Cover of WSRB

3.1.4. Cultivation

According to (Tesema, 2015) the north western part of the basin has high potential of cereal production while the north east part of the basin has low potential of cereal production. The south and south eastern portion of the area have livestock population, identified as pastoral and agro-pastoral lowland area with low and variable rainfall amount.

People living on the highlands of the north western part of the basin engaged in agricultural practices like cultivating different kinds of crops such as wheat, teff, maize, sorghum etc. In addition to crop production, most of the farmers raise domestic animals. Whereas, those living in the lowlands are living in a scattered manner and they are semi pastoralists (Tesema, 2015).

According to MoWIE report; in south central part there is scanty in cropping while the border to Wabi-shebelle River is covered with irrigated crop. But larger portion of the central part inhabited with pastoral farming life style. The higher elevation part is covered with crop like barely, wheat, sorghum, maize; with lower part is covered by other cultivation. National Meteorology Agency Jijiga Branch Office reported, in the southern and south eastern parts, because of insufficient water and grazing resources people in the area are forced to move from one place to place. This, of course, means that the people are not sedentary, but neither is they nomads in the strict sense of the term. They move in such a way that they match their needs to the periodic supplies of nature. These areas are pastoral areas having an altitude from 1000 m to 166 m from sea level. Annual precipitation less than 700mm and the area are dominated by Somali peoples (Tesema, 2015).

3.1.5. Soil Types

The wide ranges of topographic and climatic factors, parent material and land use have resulted in extreme variability of soils. In different parts of the country, different soil forming factors have taken precedence. WSRB has different composition in terms of soil types throughout the basin. 50 % of the soils are soils of calcarious or gypseous differentiation types. Vertisols comprise 12% of the basin covering significant areas at the middle belt. The soils with calcarious dierentiation are characterized by the redistribution of lime in the prole while the soils with gypseous dierentiation are characterized by the redistribution of gypsum in the prole. The latter predominantly occur in the southern arid part of the basin. The water-holding capacity of the soils is subdivided into nine classes, from wetlands which are given a water-holding capacity of 1000 mm to soils with a water-holding capacity of < 20 mm. (Abebe, 2009)

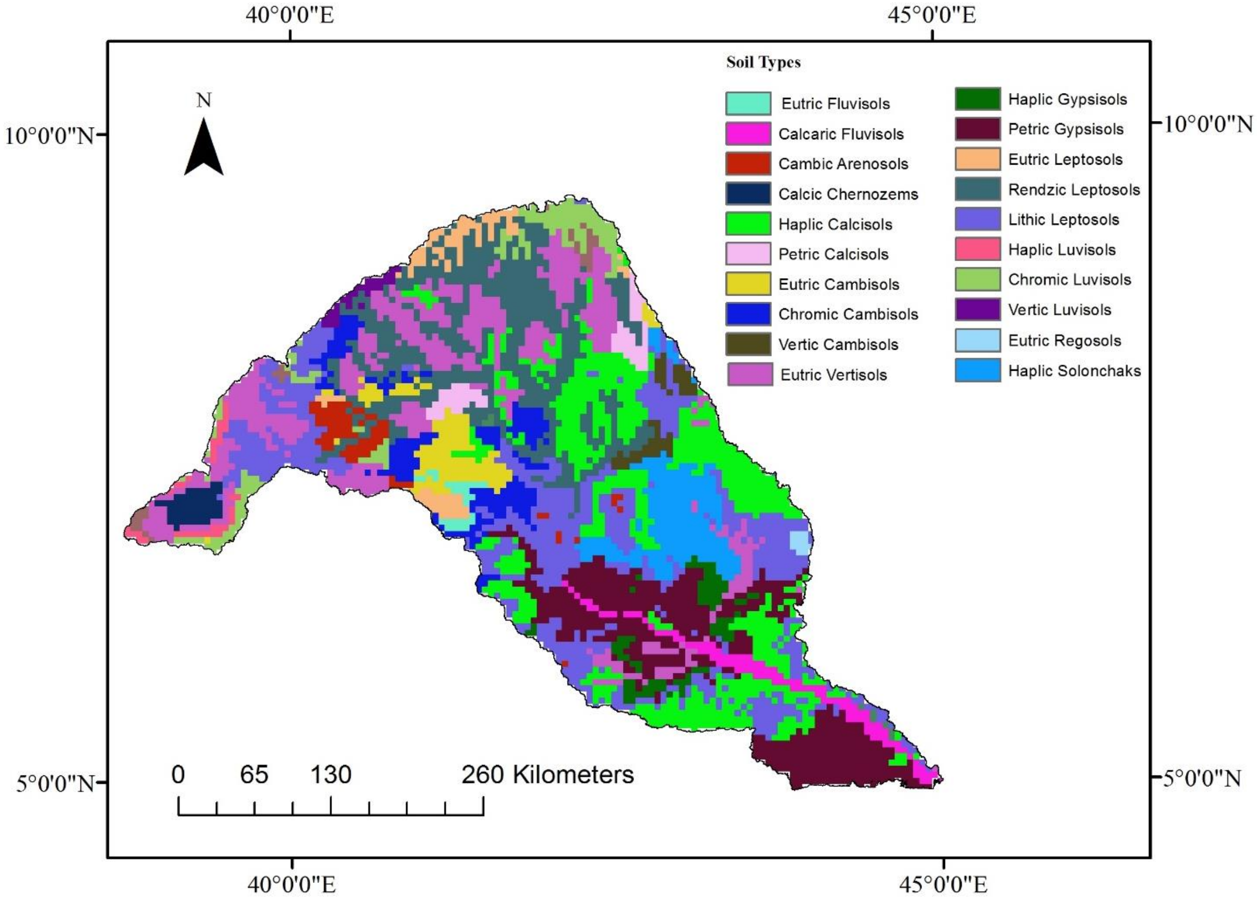


Fig 3-1-5-1. Soil types of WSRB based on Harmonized World Soil Database soil data

3.1.6. Hydro Climatic Conditions

3.1.6.1. Climate

Water is indispensable for life, but its availability at a sustainable quality and quantity is threatened by many factors, of which climate plays a leading role. Climatic elements such as precipitation, temperature, humidity, sunshine and wind are affected by geographic location and altitude. Ethiopia in general, and the basin in particular being near the equator and with an extensive altitude range, has a wide range of climatic features suitable for different agricultural production systems. Climatic heterogeneity is a general characteristic of the country as well the basin (Tesema, 2015).

The climate of Wabi-Shebele River Basin is dependent on the basin altitude. The highland areas are cool and suitable for people settlement while the lowland areas are arid and not suitable for settlement (Adane, 2009). There are some meteorological stations placed around the basin very few of them are located within the basin and most are located around the cities of the basin. The rainfall amount within the Wabishebelle river basin ranges from 200 mm on the arid part of the basin to 1250 mm towards the upper part of the basin (Ibid). The major reason for the variability of the rainfall amount is the difference in altitude which ranges from 73 m at the lower part to 4216 m to the highlands of bale.

3.1.6.2. Hydrological situation

Most of the rivers in the area arise along the northern and north western margins of the basin (Arsi Bale and Hararghe Plateau) and descend with a steep gradient into the eastern and south eastern direction and joining to the Wabi-Shebelle river which flow first to eastern direction and then south eastern and finally to Somalia territory. Currently some areas in the upper catchments have gauging station. According to MoWIE, the lower catchments are remote areas and not accessible to have gauging station. In the upper catchment there are more than 15 stations, among these Lelisso, Hararo, Assassa, Weyib at Agarifa, Ukuma, Maribo near Adaba, Jewis near Bedesa, Wabi at bridge and Wabi at Melkawakena, Robe station at Robe town, Jijiga River at Jijiga and Lake Adelle at Adelle are operational stations. But in most of the stations the data in these stations are not complete to define the hydrograph separation of the area (Tesema, 2015).

According to (Kinde, 2015) the available hydrological data in the basin are very sparse. There are some stream gauging stations in the basin and fewer are operational at present. A significant portion of which have operated intermittently, the distribution of the meteorological and hydrological gauging stations is not well integrated reducing the plausibility of such data. Adane (2009) selected the basin for research area due to the fact that the river basin is one of the water scarce basins in the country. While having the largest area coverage, its annual runoff and water availability are one of the lowest among the major river basins. Stream flow data at Gode (lower reach) is important for stream flow estimation, however, the available data is not complete, not enough record and hence cannot be used for hydrological prediction.

Inspection of most of the stations has shown that the stations are not reliable to capture the flow characteristics of the basin as most of them are clustered in the upstream and only surrounding one location. Some of them are installed on raised structures anchored to bridges, while others are settled in pockets of the main course of the stream. Low flow series often contain years with zero values. In some arid areas, zero flows are recorded more often than non-zero flows. Stream flows recorded as zero imply either that the stream was completely dry or that the actual stream flow was below a recording limit. Zero-values should not simply be ignored, nor do they necessarily reflect accurate measurements of minimum flow in a channel. Based on the hydraulic configuration of a gauge, and knowledge of the rating curve and recording policies, one can generally determine the lowest discharge which can reliably be estimated and would not be recorded as zero.

3.1.6.3. **Meteorology**

Rainfall is the most important part of the atmospheric precipitation in the hydrologic cycle that falls on the earth surface in the form of water droplets and its amount is one of the most fundamental factors to determine the density and distribution of vegetation. The rainfall in WSRB varies from less than 200 mm in arid zones (the south east part) to 1250mm in upper catchment. This is due to altitude variations over the basin from about 73m above sea level in the south east border up to 4137m above mean sea level in the upper side or at Bale Mountain Massif (Tesema, 2015).

3.1.6.4. **Meteorological Setting**

The distribution of the meteorological stations is not well integrated reducing the credibility of the available data. There are about 50 meteorological stations within and around the basin (Seleshi et al, 2007). Most of them are clustered in or near urban centers of the upper portion of the basin. Meteorological data are taken from the station based on their availability within the recording year. The stations are not evenly distributed and sometime lack continuous meteorological records. However, the available few with well record stations for recent data are enough for the purpose of this study.

As mentioned earlier, the fact that the study area is an area of various hydro meteorological characteristics it needs as many station as that of topographic variation. However achieving a well network of stations are impossible in WSRB due to difficulty of the basin nature. According to (Tesema, 2015), the National Meteorological Agency (NMA) of Ethiopia indicates that there are four class of station in the country. These are First class (synoptic station), second class (principal station), third class (ordinary station), and fourth class (rainfall recording stations). The distribution of these stations depends on the topographical point and accessibility. On the study area the meteorological station distribution are parallel with the population distribution and are condensed at the upper catchments.

3.2. Materials Used In the Study

3.2.1. Data Collection

3.2.1.1. Meteorological Data

The meteorological data required for the estimation of surface runoff includes precipitation, temperature, relative humidity, solar radiation and wind speed. The data collected from National Meteorological Agency (NMA) of Ethiopia for 17 stations located around the watershed and within the watershed for the available period of records. However, out of the 17 stations 6 of them found to have relatively better recorded data during the recent years of record. The stations are Chelenko, Karamaile Harar, Golocha, Ime and Kebri dehar. These stations have relatively good quality data for calculating the bias corrected values of the main SWAT driving input data, CFSR weather data. Most of the records does not have totally values for wind speed and solar radiation.

3.2.1.2. Climate Forecast System Reanalysis (CFSR) Data

All the CFSR climate data set taken from globally accessed web https://rda.ucar.edu/datasets/ds094.1/ address for the Wabi-Shebele location area within latitude between 4 and 9 N and longitudes 38 and 46. The CFSR global atmosphere resolution is approximately 38 km for the years 1979 to 2010 and 22 km resolution released for the year 2010 onwards with 64 levels extending from the surface to 0.26 hPa. The global ocean's latitudinal spacing is 0.25 deg at the equator, extending to a global

0.5 deg beyond the tropics, with 40 levels to a depth of 4737m. The data is downloaded for the period 1979 to 2017 at hourly basis. The hourly data then converted to daily data with the support of Visual Basic and Excel applications.

The data set clipped for only the watershed boundary of Wabi-Sheble river basin with the support of GIS application. The whole basin covered by 281 CFSR stations (Fig 3-2-1-2-1) placed in the boundary of the watershed and then the precipitation data corrected and tested for its performance based on statistical parameters. The corrected daily data for each of the 281 stations arranged sequentially in a format (.txt) which is acceptable in the SWAT hydrological model. The swat model requires daily data for Maximum and minimum temperature, precipitation, relative humidity, solar radiation and wind speed.

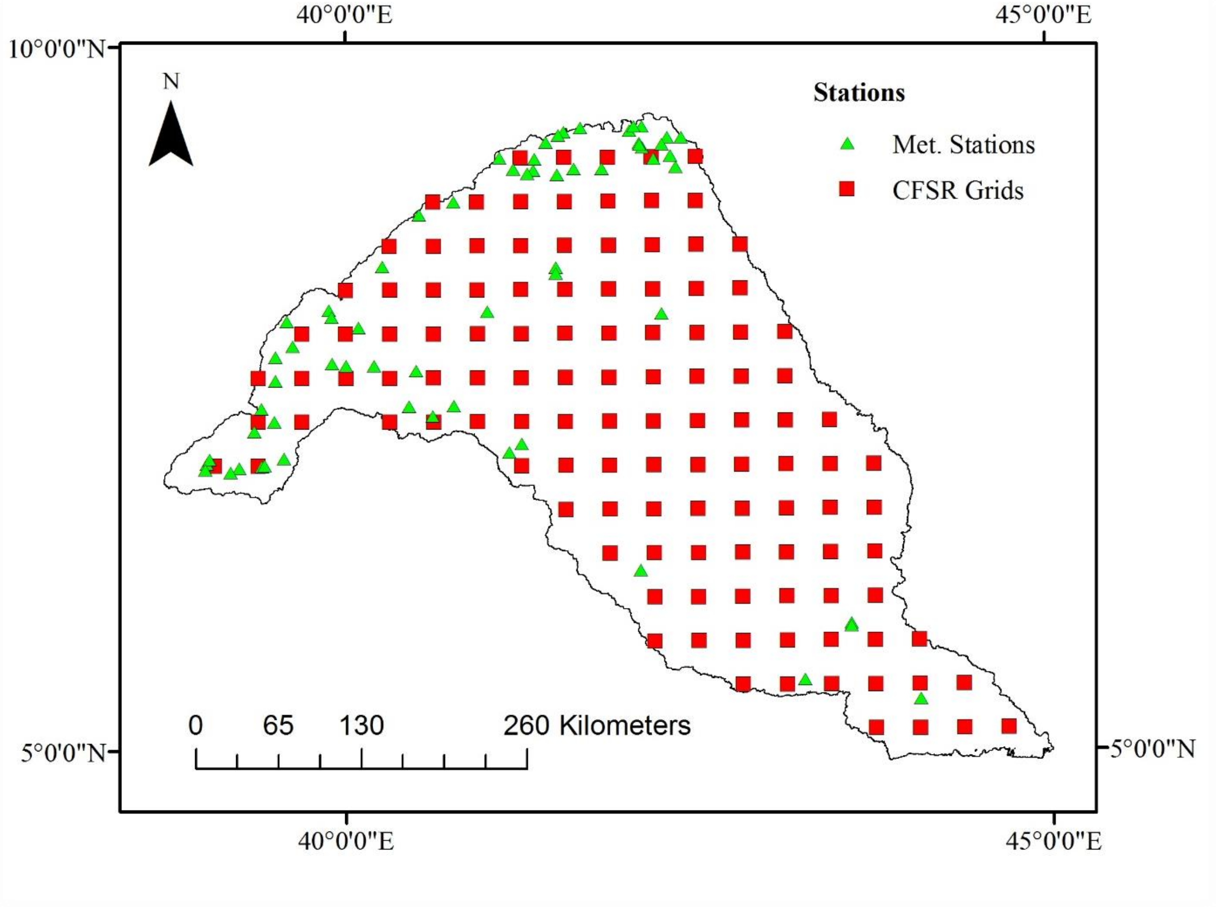


Fig 3-2-1-2-1. The distribution of CFSR climate data in the WSRB (Source: NCEP)

3.2.1.3. **Hydrological Data**

Flow data records are required for the purpose of model calibration and validation. WSRB has few gauging station which are also located in cluster in the left upper side of the basin. Although these stations are available most of the station record are missed and not enough for the calibration and validation purposes. A better record station obtained only at Wabi at bridge (wabi @ bridge) station which can be used as representative for the left upper section of the watershed. However, the right upper, middle and lower portions of the watershed cannot be represented in the modelling process as these places lack well recorded and enough flow data. Hence, the behavior of the watershed only well represented in the upper left part of the basin but the remaining portions of the basin are assumed to have the same behavior as in the upper left part. The data is collected from the Ministry of Water, Irrigation and Energy (MWIE) for the period of 1990 to 2008.

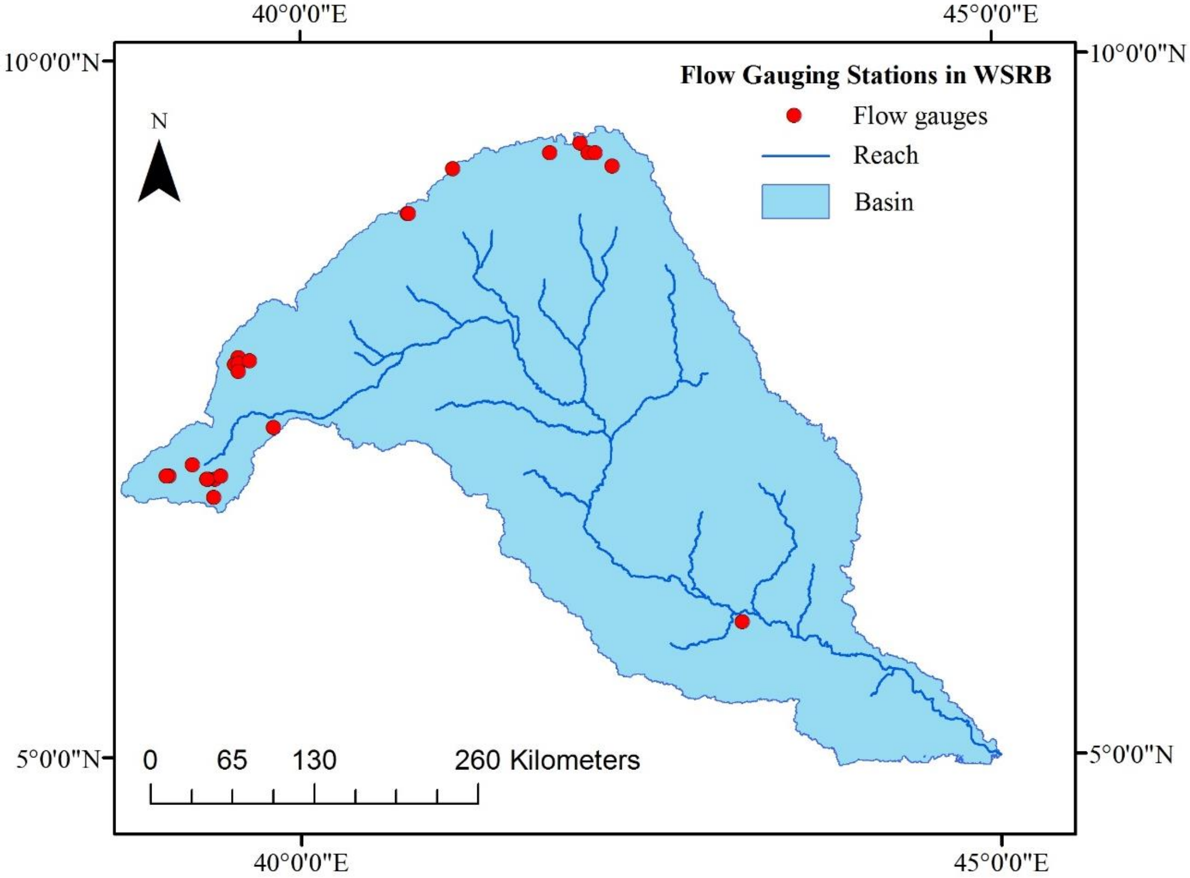


Fig 3-2-1-3. Flow gauges distribution at the WSRB

3.2.1.4. Digital Elevation Model (Data)

Digital Elevation Model (DEM) is the digital representation of the land surface elevation with respect to any reference datum. DEM is frequently used to refer to any digital representation of a topographic surface. DEM is the simplest form of digital representation of topography. DEMs are used to determine terrain attributes such as elevation at any point, slope and aspect. Terrain features like drainage basins and channel networks can also be identified from the DEMs. DEMs are widely used in hydrologic and geologic analyses, hazard monitoring, natural resources exploration, agricultural management etc. Hydrologic applications of the DEM include surface runoff modeling, groundwater modeling, estimation of the volume of proposed reservoirs, determining landslide probability, flood prone area mapping etc. (A. Balasubramanian, 2017). The digital elevation model (DEM) is specifically made available in the form of raster or regular grid of spot heights and it is the basic input data for SWAT hydrological model. The Wabi-Shebele River Watershed was delineated and River networks were generated from DEM that has resolution of 30m x 30 m.

3.2.1.5. GIS Application

Geographic information systems have emerged in the last decade as an essential tool for urban and resource planning and management. Their capacity to store, retrieve, analyze, model and map large areas with huge volumes of spatial data has led to an extraordinary proliferation of applications. A GIS is a system of hardware, software and procedures to facilitate the management, manipulation, analysis, modeling, representation and display of georeferenced data to solve complex problems regarding planning and management of resources (Escobar et al). A GIS 10.1 version software used in this study.

3.2.1.6. Land Use/Land Cover Data

A raster land use land cover data (LULC) prepared in 20 m resolution used for this study. This is downloaded from the European Space Agency web address available at (http://2016africalandcover20m.esrin.esa.int/) and the data for WSRB was clipped from it with the support of GIS application. The legend of the data includes 10 generic classes that appropriately describe the land surface at 20m: "trees cover areas", "shrubs cover areas", "grassland", "cropland", "vegetation aquatic or regularly flooded", "lichen and mosses / sparse vegetation", "bare areas", "built up areas", "snow and/or ice" and "open water". The legend defined using the look up table which is the appropriate (.txt) format for SWAT application. A look-up table that identifies the 4-letter SWAT code for the different categories of land cover/land use was prepared so as to relate the grid values to SWAT land cover/land use database.

3.2.1.7. Soil Map Data

A raster soil map data was used for this study which is prepared in 100 m resolution. The raster soil data downloaded from the Harmonized World Soil Database (available at http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/). As for the land use, the soil layer in the map was identified in a separate lookup table which enables easier allocation of soil group information for the watershed.

3.3. CFSR Data Analysis

3.3.1. Bias Correction and Validation

CFSR climate data set has to be adjusted to the local scale through appropriate bias correction method. Linear scaling method (Equation 2-2-3-1) of correcting raw CFSR data was used as it is one of precipitation correcting method and previously used and tested in Ethiopia. The raw monthly and observed monthly precipitation data first checked to prove the applicability of the method for Wabishebele river basin. Once the correction method checked and verified, the long term mean monthly observed precipitation data of the stations was used for surface interpolation of rainfall data over the entire watershed with the support of GIS tool. The mean monthly observed values for corresponding to every CFSR data stations extracted and the ratio of these values with that of the mean of the long term mean values of raw CFSR data taken to correct the daily raw data. The monthly values of the corrected precipitation data was used to test its performance with the monthly values of observed values for the selected gauging stations in the watershed.

The performance of the corrected precipitation values were checked using time series based indices which includes Nash–Sutcliffe measure of efficiency (NSE) indicated in (Equ 2-2-3-7) , percent bias (PBIAS) indicated in (Equation 2-2-3-8) and the ratio of mean square error to standard deviation (RSR) indicated in (Equation 2-2-3-11) . GIS based techniques used to select which CFSR station is near to the observed station up on which the bias correction and performance evaluation applied.

3.3.2. Interpolation

Interpolation is a method or mathematical function that estimates the values at locations where no measured values are available. Spatial interpolation assumes the attribute data are continuous over space. This allows the estimation of the attribute at any location within the data boundary. Another assumption is the attribute is spatially dependent, indicating the values closer together are more likely to be similar than the values farther apart. These assumptions allow for the spatial interpolation methods to be formulated. The goal of spatial interpolation is to create a surface that is intended to best represent empirical reality thus the method selected must be assessed for accuracy (Azpurua et al, 2010). There is different interpolation technique applicable to estimate unknown values. This study used IDW (Inverse Distance Weighting) which work based on the assumption that the nearby values contribute more to the interpolated values than distant observations. IDW is a good interpolator for phenomena whose distribution is strongly correlated with distance and it does not create a surface which exceeds the known values of the sample points.

The monthly average precipitation data of the observed data set of each station interpolated over the surface of the basin to create average precipitation data for the rest of the area within the boundary of the river basin. Based on CFSR data station distribution, values of observed monthly precipitation extracted from the interpolated surface using the GIS extraction technique for the corresponding CFSR stations.

3.4. SWAT Model Set up

3.4.1. Watershed Delineation

The watershed area of WSRB delineated using SWAT 2012 application interface on GIS program. This is done using 30 m resolution DEM. The watershed delineation interface in Arc-View is separated into five sections including model Set Up, Stream Definition, Outlet and Inlet Definition, Watershed Outlet(s) Selection and Definition and Calculation of Sub basin parameters. The watershed delineation dialog box allows for the calculation of flow direction and accumulation. Then the steam and sub basin outlet definition has done by fixing the area of sub-basin in the suggested threshold area. Then outlet points for the possible flow gauging station locations provided using the add point tab of this window and the whole watershed outlet point selected to delineate the watershed area. At the end of this process the number of sub basins, the watershed area, minimum and maximum elevation and the number of outlets calculated. The watershed is divided into 40 sub basins.

3.4.2. Hydrologic Response Unit (HRU) Definition

Hydrologic response units are portions of sub basin that possess unique land use/management/soil attributes (Neitsch et al, 2004). The HRUs are normally defined by lumping similar land use, soil type, and optionally slope characteristics within a given sub basin based on user-defined thresholds for each category. In this standard method, the user can control the number of HRUs by applying a threshold on land area permitted for a given land use or soil type within a sub basin. Fewer HRUs may be desirable for achieving computational efficiency. At the small watershed to field scale, however, individual field management may become an important consideration and field-based outputs and potentially inputs may be necessary depending upon simulation objectives. In particular, if SWAT model results are to be communicated to stakeholders such as farmers, landowners, or land managers, outputs should match socially meaningful area units such as parcels, fields, or even counties (Margaret M. et al, 2015).

The overlay of land use, soil and slope maps resulted in the definition of 553. HRUs enable the model to reflect differences in evapotranspiration and other hydrologic conditions for different land covers and soils. The runoff was estimated separately for each HRU and routed to obtain the total runoff for the watershed. This increases the accuracy of inflow prediction and provides a much better physical description of the water balance. The land use and soil data in a projected raster file format was loaded into the SWAT interface to determine the area and hydrologic parameters of each land-soil category simulated in each sub-basin according to the percentage of each HRU definition. Accordingly, the model set up is done based on many studies suggestion as 10% land use, 20 % soil and 10% slope definitions which give more emphasis to soil unit in the simulation process.

3.4.3. Weather Data Definition

The weather data of each of the 281 CFSR stations arranged according to the appropriate format used by SWAT model which is the (.txt) file format. The 38 km resolution dataset changed to 22km grid dataset applying square polygon technique in GIS application. The daily corrected and validated precipitation data for the period starting from 1979 to 2017 prepared on (.txt) format with the corresponding station location and elevation. The same format prepared for the daily values of minimum temperature, maximum temperature, solar radiation and relative humidity with the respective header files name identity.

Before loading the (.txt) files of each of the variables of the weather data, the weather generator table of the swat database located at WEGN user was filled for all stations calculating the long term monthly averages of the required fields of climate data listed on the database table and calculating statistical parameters namely Pstat and dew point. Pstat values are calculated based on precipitation data whereas dew point calculated based on temperature data and humidity data.

3.4.4. Sensitivity, Calibration And Validation

After all the data filled in the SWAT model interface and the setup completed, the model subject to run to generate the simulated result. The acceptability of the model output is based on the fitness of the simulated output with the observed data. The fitness of the simulation result depends on the model parameters and these must held within a realistic uncertainty range. Before calibration and validation of the SWAT model, sensitivity analysis of parameters conducted to identify the most sensitive parameter representing the river basin. This helps to make a focus on the most sensitive parameters during the calibration and validation process.

Generally there are two types of sensitivity analysis methods namely global and one a time (local) sensitivity analysis. The global sensitivity analysis implemented by allowing all the parameters change at a time and one a time sensitivity done by allowing one parameter to change at a time. Since sensitivity of one parameter depends on the value of other related parameters, the one at a time method of analysis is not advised to follow. The SWAT cup SUFI-2 is used for the sensitivity analysis of the most sensitive parameter and for further calibration and validation processes.

3.4.5. Model Simulation and Surface Runoff Estimation

After calibrating the model, the parameter of best fit is taken to the SWAT model to amend the parameters of the model which make it a representative model of the river basin. Based on this arrangement the model set to run again to generate the outputs of the watershed. The result of the calibrated model is then validated using observed flow data. The result of the output from the reach file taken and the total surface runoff amount of the total river basin calculated.

4. Results and Discussion

4.1. Precipitation Bias Correction And Validation Result

The daily raw CFSR data collected for the time period 1979 to 2017 is adjusted for its systematic error and validated to ground observed values. The adjustment is to hold the variability between CFSR raw data and observed rainfall data with small number and unevenly distributed Rainfall gauging stations (Table 4-1-1). Observed gauge stations data for the correction and validation of CFSR data is taken from the available records based on better recorded data both in terms of similar period years and relatively less missing values. Accordingly gauging stations located at Chelenko, Harrar, Karamaile , Gololcha, Ime and kebridehar are selected from the available stations used for correction.

Linear scaling is applied for the bias correction of all of the 281 CFSR stations covering the whole area of the watershed. The CFSR data has time coverage from 1979 to 2017 with no missing data and achieve even distribution throughout the watershed. The corrected daily precipitation data is then checked for their validation using time series statistical tests as Nash-Sutcliffe Efficiency (NSE), Percent bias (PBIAS) and the root mean square error standard deviation ratio (RSR). The performance evaluation of the corrected values of precipitation data showed good result for the stations located at Chelenko and Harar (Table 4-1-2.). However, the statistical test result a satisfactory result for the station located at Golocha, Ime, Kebridehar and Karamaile (Table 4-1-2). As chelenko and Harar stations are located nearby and their statistical test showed similar result. The fact that the CFSR dataset showed a consistent result on the two stations revealed that a bias corrected CFSR weather data could be applicable to the assessment of the water resource situation of a river basin when and if the watershed is under data scarce condition.

Table 4-1-1. Precipitation information collected from NMA

Station Name	latitude	Longitude	Elevation	Percentage of missing data	Data Period
Arabi	9.93225	42.7202	1097	10.96	2012 -2014
Bedeno	9.13333	41.6333	2050	8.54	1997-2013
Chelenko	9.3833	41.55	2178	5.5	2003-2014
Degahabur	8.21667	43.55	1070	10.96	1997-2015
Dire Dawa	9.9667	42.5333	1180	1.3	1997-2014
Ejersagoro	9.4833	43.2167	2161	9.1	1999-2013
Erer	9.55	41.3667	1088	33.1	1999-2014
Fedis	9.1333	42.0833	1690	4.22	1997-2014
Gode	5.9	43.5833	295	29.8	2004-2012
Hakimgara	9.28333	42.1167	2159	14.55	1997-2015
Harar Indicative	9.3	42.0833	1977	15.94	1998-2015
Jijiga	9.33333	42.7833	1775	11.4	1997-2016
Karamaile	9.36667	41.35	2264	53.2	1998-2015
Kebri Dehar	6.73333	44.3	505	47.13	1997-2015
Kobo	9.2	41.35	1768	9.78	1997-2015
Ime	6.28	42.09	584	37	1988-1991
Gololcha	8.01	40.09	1773	6.7	1990-1991 & 2007-2009

Table 4-1-2. Performance evaluation result of corrected CFSR precipitation data at the corresponding observed stations.

Evaluation Result

Station Name	**NSE**	**PBIAS**	**RSR**
Chelenko	0.66	-1.69	0.58
Harar	0.65	4.41	0.58
Karamaile	0.54	-13.5	0.67
Gololcha	0.58	18	0.64
Ime	0.61	-14	0.61
Kebri Dehar	0.58	-20.7	0.63

The mean annual rainfall distribution over the basin is demonstrated in the figure below (Fig 4-1-1). The estimated corrected annual CFSR data shows a 23 % more in the minimum and a 15% more in the maximum annual values comparing with observed records.

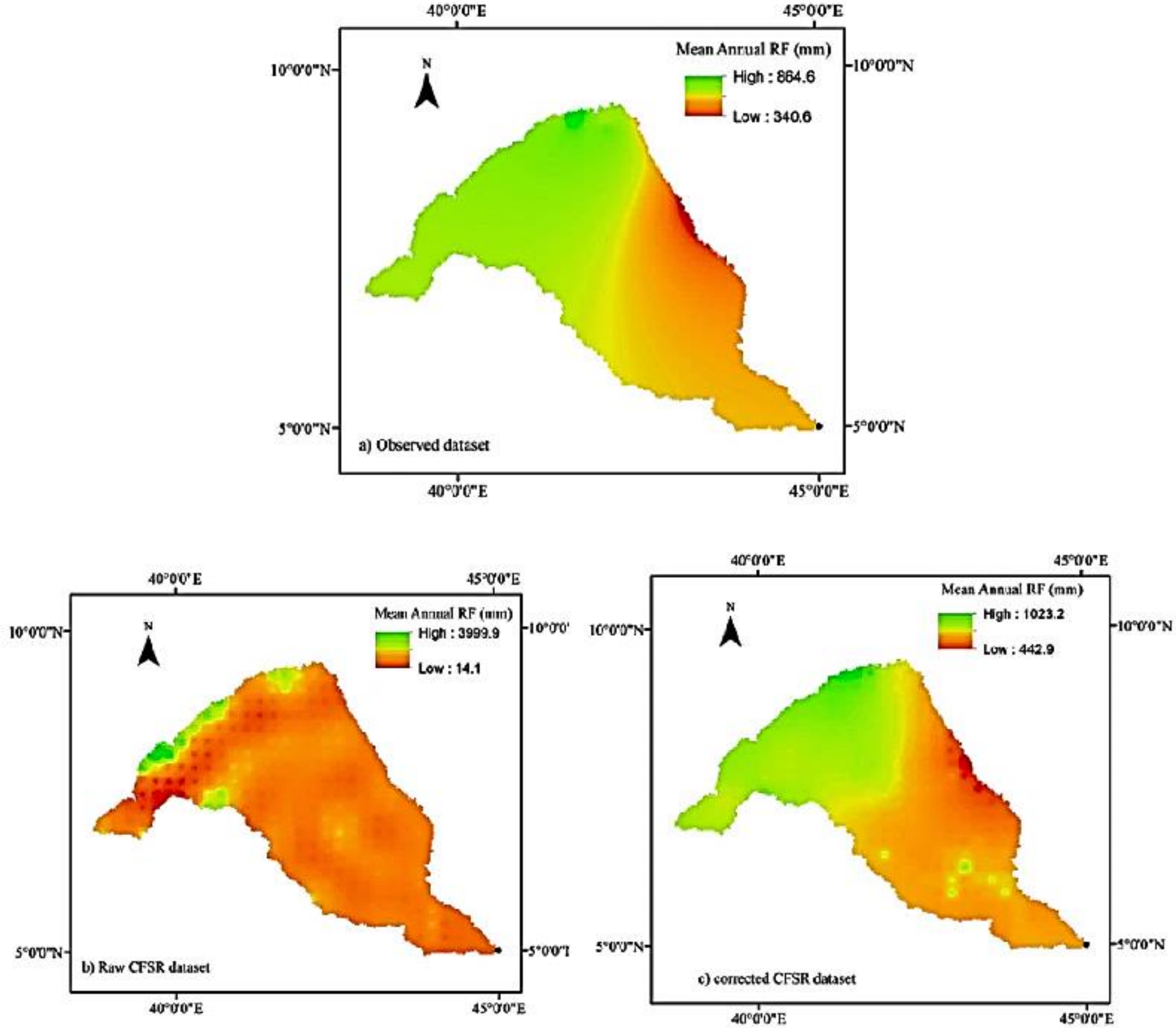


Fig 4-1-1. Comparison of spatial distribution of mean annual rainfall in the WSRB with all types of data set

As it can be seen from the figures below (Fig 4-1-2), the raw CFSR dataset has different distribution as compared with that of the observed rainfall records. The average mean monthly value of raw CFSR data shows relatively highest distribution in almost all months of the year than that of the observed values at all gauging stations except that of Gololcha. At Gololcha the raw CFSR data distribution shows lower distribution than the observed rainfall distribution throughout the year.

The raw CFSR data shows a different mean monthly distribution at Harar gauging station throughout the year. It has higher distribution for months from January to June and it holds almost similar distribution for the months from July to December with that of the observed records. The figure also depict the average monthly raw CFSR data has almost similar distribution throughout the year with that of the observed records at Kebri dehar and Ime rainfall gauging stations. In all of the stations, the estimated mean monthly corrected precipitation shows a good distribution of throughout the year.

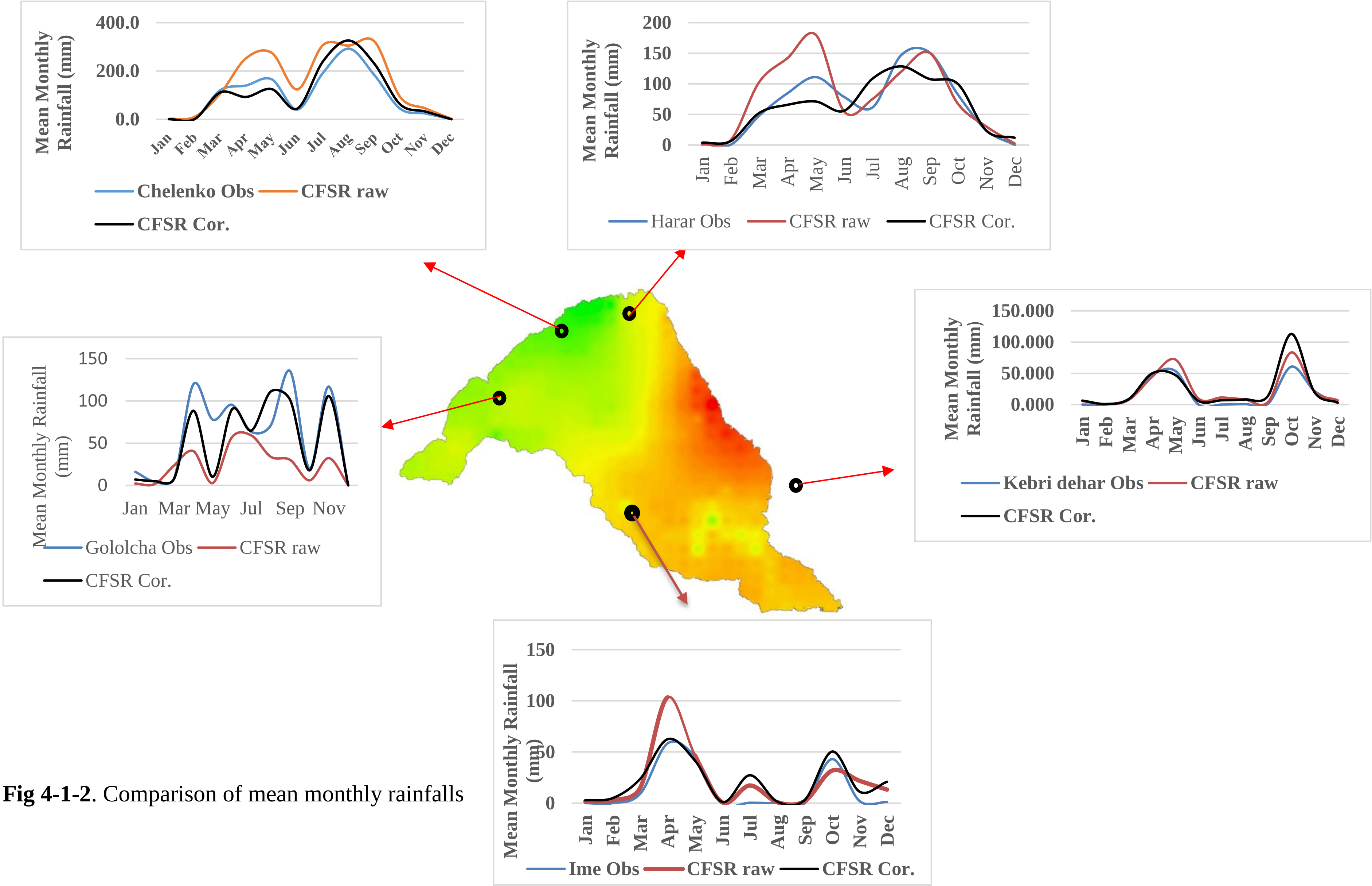


Fig 4-1-2. Comparison of mean monthly rainfalls

4.2. Stream Flow Analysis

4.2.1. Sensitivity , Calibration and Validation

Any given parameter can be sensitive within a particular range and during specific seasons or time periods. The sensitivity of a given parameter may vary depending on one or multiple other parameter values. Parameter sensitivity can be defined as the change in outcome in response to a specified change in a particular value. Sensitivity analysis of SWAT input parameters were conducted first to identify the most sensitive parameter involved in the process of calibration and validation. SWAT CUP SUFI-2 sensitivity, calibration and validation tool provides lists of parameters with their respective uncertainty range to be maintained to achieve an acceptable result of simulation. In SUFI-2, uncertainty in parameter, establish range of uniform distributions of parameters which captures the uncertainty stem from parameters, measured values, input variable and on the conceptual model. Propagation of the uncertainties in the parameters leads to uncertainty in the model output variables which is termed as the 95% prediction uncertainty (95PPU).

The parameters involved in the calibration and validation process tested for their sensitiveness for WSRB. In SUFI -2 sensitivity algorithm there are two types of working sensitivity analysis, these are Global and Local analysis, the former works by changing values one at a time and the later works by allowing all parameter values to change at a time. Since the sensitivity of a parameter often depends on the values of other related parameters the local sensitivity analysis commonly not advised. Here, we use Global sensitivity analysis method to keep the interdependence of parameters. In SUFI-2 sensitivity analysis there are two statistical measures used to identify the most sensitive parameter, namely t-stat and p-values. The most sensitive parameter exhibit a higher t-stat value in absolute value and lower p-values.

With this general overview of the sensitivity analysis 16 selected parameters was evaluated in the Wabishbele River Basin SWAT model (Table 4-2). Although, these 16 parameters have different t-stat values ranging from -2.2 to 0.15, all have p-values closer to zero, which shows the significance of all parameters. Therefore, none of them rejected out. All are kept on the process of calibration and validation.

Table 4-2-1. Sensitivity measures of parameters

No	Parameter Name	t-Stat	P-Value	Range	Calibrated Values
1	V__GW_DELAY.gw	-2.2	0.03	[0, 500]	230.7
2	V__RCHRG_DP.gw	-1.8	0.07	[0, 1]	0.05
3	R__SOL_AWC(..).sol	1.48	0.14	[-0.3 , 0.3]	0.13
4	V__DEP_IMP.hru	-1.46	0.14	[0 , 6000]	5583
5	V__REVAPMN.gw	-1.2	0.24	[0 , 1000]	993.5
6	R__SOL_K(..).sol	-1.1	0.28	[-0.3 , 0.3]	0.24
7	V__ESCO.hru	-1.06	0.29	[0, 1]	0.16
8	V__ALPHA_BF.gw	-0.98	0.33	[0, 1]	0.01
9	R__SOL_Z(..).sol	0.85	0.4	[-0.3 , 0.3	-0.22
10	R__CN2.mgt	-0.63	0.5	[-0.3, 0.3]	-0.29
11	V__CH_N2.rte	0.55	0.58	[-0.01 , 0.03]	0.2
12	V__CH_K2.rte	0.48	0.6	[0 , 500]	216
13	V__GWQMN.gw	0.31	0.76	[0, 5000]	4901
14	V__GW_REVAP.gw	-0.23	0.8	[0.02 , 0.2]	0.18
15	R__SOL_BD(..).sol	-0.2	0.84	[-0.3, 0.3]	0.03
16	V__SURLAG.bsn	0.15	0.88	[0.5 , 24]	22.5

Based on the above sensitive parameters, calibration of the model set to run. Calibration is an effort to better parameterize a model to a given set of local conditions, thereby reduce the prediction uncertainty. Once the sensitive parameters for this study selected, then an iteration of the SWAT CUP conducted setting the parameters within a specified range of variation. The mean monthly observed flow data for 5 years (1999-2003) at Wabi-bridge gauging station used for calibration. This is the only station we have with complete records for the period of calibration and validation. Calibrated parameters based on this gauging station applied for the entire watershed to estimate the river basin flow which leads to underestimation of actual flow. This is due to lack of data in the watershed and it is the reason that makes previous attempts of estimation unreliable. SUFI-2 calibration and validation best fit model judged through statistical measuring criteria's such as coefficient of determination R^2 and NSE. Accordingly, calibration resulted R^2 of 0.44 and NSE value of 0.40 on the calibration period (Fig 4-2-1).

After calibration, the model final best fit parameters used for validation of the model. A 5 year (2004-2008) observed flow data used for the validation step and the model consistency evaluate based on the statistical measurements. The model showed a performance of R^2 0.12 and NS value of 0.11 during validation period (Fig 4-2-2).

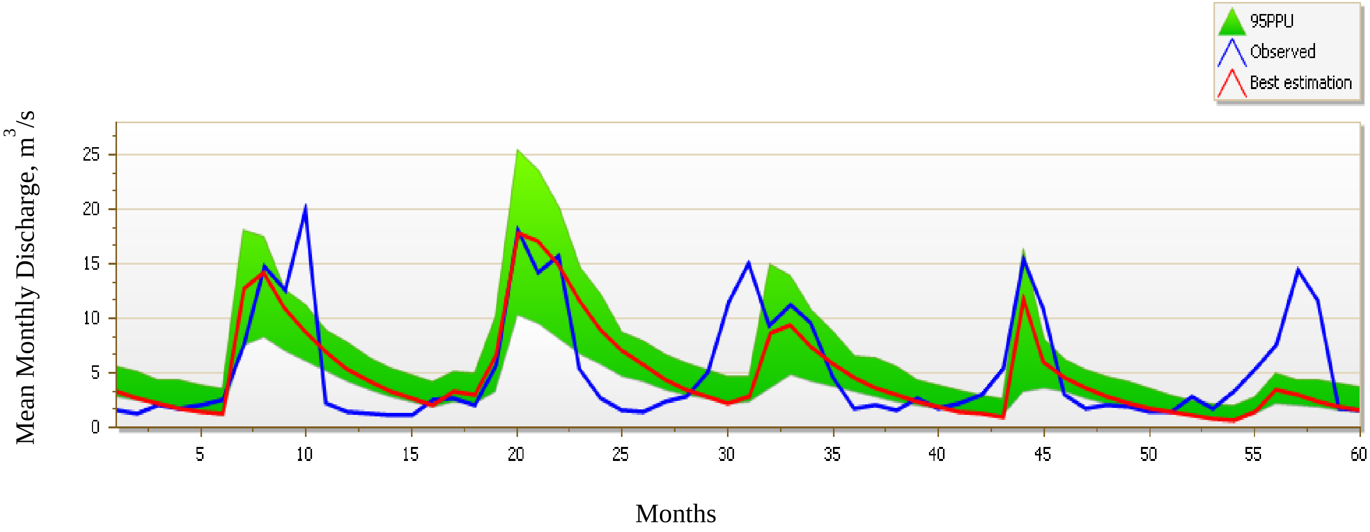


Fig 4-2-1. Calibration period observed and simulated result

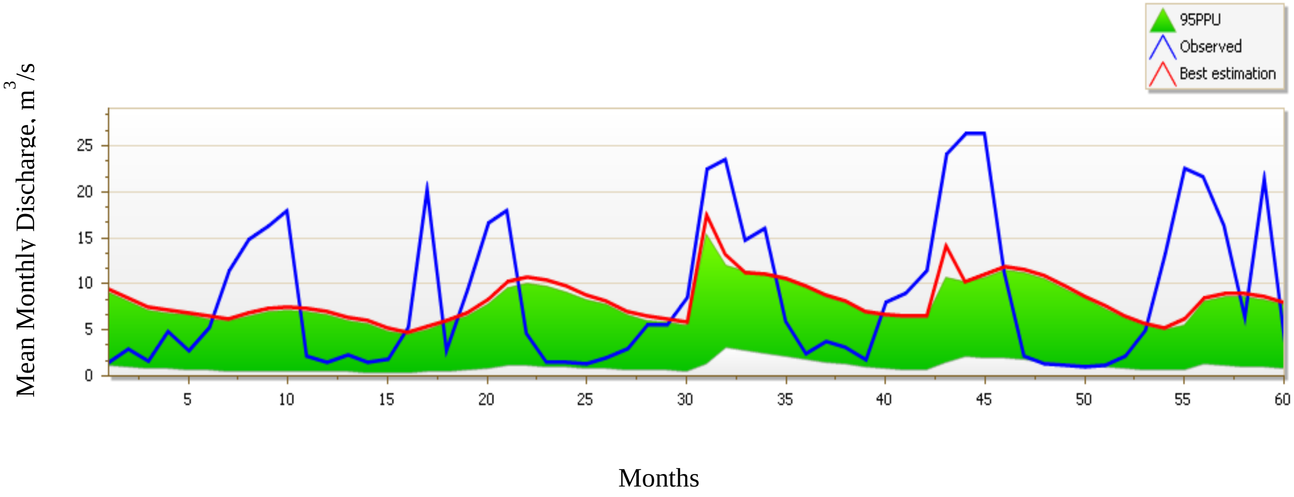
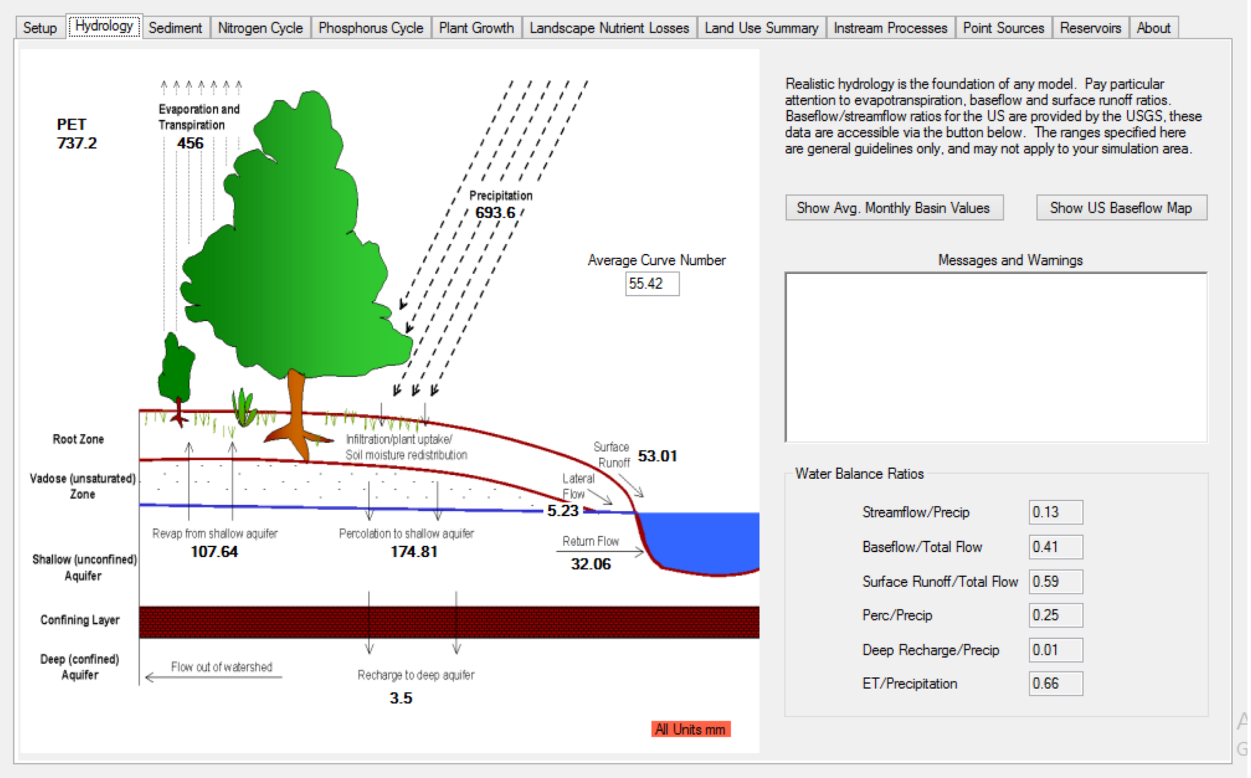


Fig 4-2-2. Validation period observed and simulated result

4.3. Basin Water Balance

SWAT is a semi-distributed continuous widely used hydrological model to examine stream flow in the process of hydrologic cycle based on the water balance equation. The water balance of the Wabi-Shebele River Basin evaluated using the different SWAT evaluation outputs. First the water balance of Wabi-Shebele basin evaluated based on the SWAT checker proportional values of hydrological elements (Figure 4-3). In this water balance evaluation the proportion share of Evapotranspiration, stream flow, infiltration and deep percolation per precipitation stated as 0.66, 0.13, 0.25, 0.01 respectively, which show the relative distribution of the water share over hydrological elements satisfied the scientific ranges of share of hydrological cycle elements. This mimic of the natural hydrological system confirms the accuracy of the SWAT model setup of the basin.



Realistic hydrology is the foundation of any model. Pay particular attention to evapotranspiration, baseflow and surface runoff ratios. Baseflow/streamflow ratios for the US are provided by the USGS, these data are accessible via the button below. The ranges specified here are general guidelines only, and may not apply to your simulation area.

Average Curve Number: 55.42

Water Balance Ratios

Streamflow/Precip	0.13
Baseflow/Total Flow	0.41
Surface Runoff/Total Flow	0.59
Perc/Precip	0.25
Deep Recharge/Precip	0.01
ET/Precipitation	0.66

Fig 4-3. Hydrological water balance as expressed in SWAT checker

Secondly the water balance of the basin evaluated with mean monthly areal values of hydrological elements (Table 4.3). In this evaluation the water yield accounts the sum of the stream flow (QSF) and the ground water recharge. Therefore, the mean annual water volume of surface runoff, lateral flow, stream flow and the total water yield of the basin is accounted as 6.8 BCM, 0.7 BCM, 7.7BCM and 13 BCM respectively.

Table 4-3. Areal mean monthly values of Hydrological water balance elements in the basin

Month	Rain (mm)	Q surf (mm)	Q Lat (mm)	Q SF (mm)	Q Loss (mm)	Water Yield (mm)	ET (mm)	PET (mm)
Jan	8.4	0.1	0.4	2.5	1.7	4.2	22.5	71.3
Feb	21.8	3.4	0.3	4.0	2.7	6.7	18.9	64.8
Mar	54.3	3.6	0.4	3.9	2.7	6.7	35.0	72.4
Apr	124.9	6.5	0.4	5.5	3.8	9.3	48.1	60.5
May	72.5	2.0	0.4	2.9	2.0	4.8	56.9	62.7
Jun	37.6	1.0	0.4	2.2	1.5	3.7	40.3	50.4
Jul	89.3	7.4	0.4	6.0	4.1	10.0	38.3	51.6
Aug	88.9	15.0	0.5	10.6	7.3	17.9	41.4	59.7
Sep	82.6	7.1	0.6	6.2	4.3	10.5	40.2	59.2
Oct	76.7	4.2	0.6	5.0	3.4	8.4	47.1	58.9
Nov	27.3	2.5	0.5	4.1	2.8	6.9	38.0	58.8
Dec	9.2	0.2	0.5	2.9	2.0	4.9	29.5	66.1
Annual	693.5	53.0	5.2	55.9	38.2	94.0	456.0	736.4
(in BCM)	98.4	6.8	0.7	7.7	5.3	13	62.9	102

Thirdly the basin water balance expressed based on the land use types of the basin. The basin classified into 4 land use/cover types. This land use/cover has varied water holding characteristics, which accounted with different magnitudes for different hydrological elements (Table 4-3-1). As presented in the table, Forest Deciduous land cover take the largest share of Evapotranspiration and the agricultural land use/covers accounted the large share for surface and ground water flows.

Table 4-3-1. Basin water balance information with respect to land use/covers

LULC	Area (%)	CN	Rainfall (mm)	Evapotranspiration (mm)	Surface Runoff (mm)	Ground Water (mm)	AWC (mm)
AGRL	17.78%	59.7	915.51	495.71	103.5	100.75	184.45
FRST	36.03%	52.3	688.8	479.87	42.95	26.14	166.17
PAST	45.88%	56	609.85	420.98	41.67	21.14	158.99
FRSD	0.31%	54	926.53	601.20	8.03	86.51	242.72
Mean		55.5	785.17	499.44	49.04	58.64	188.08

4.4. Stream Flow Characteristics

The spatial and temporal stream flow characteristics are basic information that guides the surface water utilization in a given river basin. Therefore, the spatial and temporal stream flow characteristics of Wabi-Shebele River Basin were analyzed using semi-distributed hydrological model; SWAT. As given in the above sections the spatial distribution of the stream flow conditions of the basin modelled with 40 sub basins. However, to see sectional characteristics of the flow 3 sub basins are selected based on the main tributary and gauging sites in the basin. The mean monthly stream flow of these sub basins are analyzed to see the spatial variability of flows in the basin.

Table 4-4. Mean monthly and annual stream flows for selected sub basins

Months	Sub Basin Name			
	Gobele	Daketa	Wabi	Gode
Jan	7.32	2.27	34.93	88
Feb	6.94	2.04	31.97	82
Mar	8.22	7.21	31.51	88
Apr	8.71	3.68	32.43	298
May	8.53	3.87	32	452
Jun	8.27	2.44	33.62	220
Jul	10.09	5.65	43.55	160
Aug	11.3	5.72	59.56	235
Sep	10.9	6.65	55.25	264
Oct	10.61	3.61	50.61	217
Nov	10.98	3.29	44.82	215
Dec	9.79	2.9	40.97	132
Annual (cms)	9.3	4.11	40.94	204
Annual (BCM)	0.3	0.13	1.29	6.4

Gode hydrological gauging station is located at the downstream of the WSRB which is the best to capture the flows generated upstream of it. However, it is one of the stations in the river basin which lacks complete and sufficient daily stream flow data for calibration and validation of the model. Hence, here the available monthly average observed flow data of Gode hydrological station used as a control station to compare the model predication and assure the result of model output is within the range of the reality on the ground.

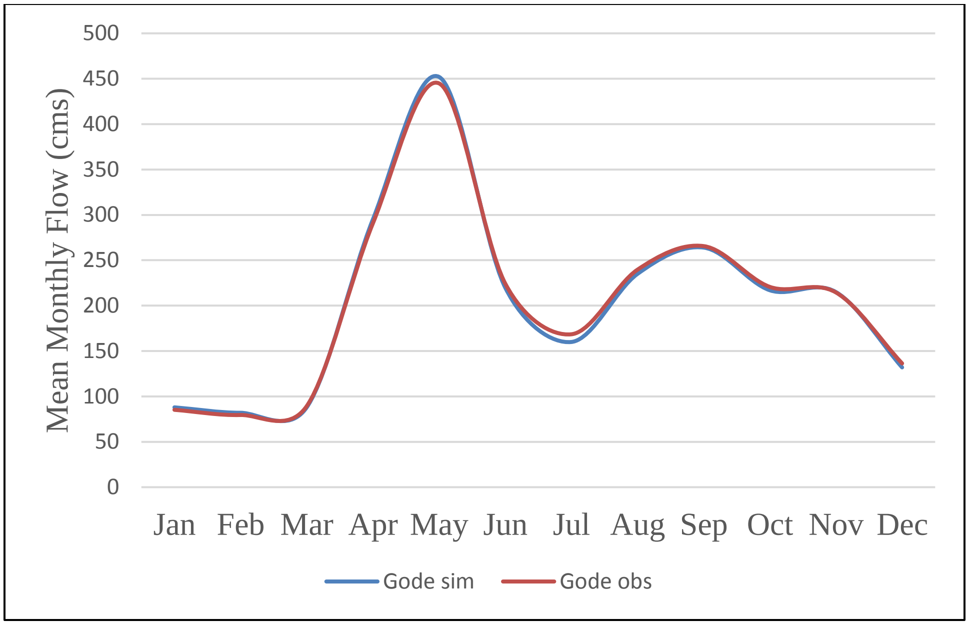


Fig 4-4. Flow Comparison at Gode Hydrological station

5. Conclusion and Recommendation

The study used bias corrected CFSR weather data to drive SWAT hydrological model. The bias corrected CFSR data showed good correlation with ground observed rainfall data with an NSE value 0.54-0.66 on the selected rainfall gauging stations. The rainfall gauging stations on this basin have limitations on proper records of rainfall both in terms of years of record and complete records. The fact that statistical performances resulted on the selected stations are different, highlights lack of well recorded and sufficient data in the river basin. Hence, previous studies which attempts to assess the water potential of the river basin using observed rainfall data is uncertain.

The stream flow data used for the calibration and validation were only from one end of the upstream side of the watershed. The simulated stream flow resulted a performance of R^2 value of 0.44 and NS of 0.41 during calibration years of 1999 to 2003. Whereas, during the validation years of 2004 to 2008 the simulated result an R^2 to 0.12 and NS value of 0.11. The values are less as the station is located at the upstream area which has low and undependable flow records.

The study also confirmed the application of SWAT hydrological model for water assessment of a basin. In line with this the surface water potential of Wabi-Shebele River Basin more reliably estimated as 7.7 BCM annually. However, this output should be taken with great care for other applications as it is much more than the master plan output which is 3.4 BCM. Generally, we concluded that the application of bias corrected CFSR data is suitable to support any hydrological analysis in data scarce areas.

From the perspective of this research we highly recommend to evaluate the surface water potential of the WSRB with better resolution data sets. Climate Forecast System Reanalysis (CFSR) and any other data set again recommended for the study of other hydrological analysis like flood forecasting and drought analysis of a basin.

REFERENCES

Aaron Havel, 2015. Hydrologic and Hydraulic Response to Wildfires in the Upper Cache La Poudre Watershed Using SWAT and HEC-RAS Model Cascade. A thesis submitted to the Department of Civil and Environmental Engineering, Colorado State University, Fort Collins, Colorado, Fall2015.

Abbaspour K. C., M. Vejdani, and Haghighat S, 2007. SWATCUP calibration and uncertainty programs for SWAT. In Proc. Intl. Congress on Modelling and Simulation (MODSIM'07), 1603-1609. L. Oxley and D. Kulasiri, eds. Melbourne, Australia: Modelling and Simulation Society of Australia and New Zealand.

Abbaspour Karim C., 2015. SWAT CUP SWAT calibration and Uncertainty Program. A user manual.

Abbaspour Karim C., Saeid Ashraf Vaghefi and Raghvan Srinivasan, 2017. A Guideline for Successful Calibration and Uncertainty Analysis for Soil and Water Assessment: A Review of Papers from the 2016 International SWAT Conference, Water 2018, 10, 6; doi:10.3390/w10010006

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E., Rasmussen, J., 1986b. An introduction to the European Hydrological System—Systeme Hydrologique Europeen 'SHE'. 2: Structure of a physically based, distributed modelling system. J. Hydrol. 87, 61–77.

Adane Abebe, 2009. Hydrological Drought Analysis-occurrence, severity, risks: the case of Wabi Shebele River Basin, Ethiopia. A dissertation vom Fachbereich Bauingenieurwesen der University at Siegen genehmigte.

ALLISON R. HURLEY, 2014. Comparison and Validation of Arctic Precipitation Fields from Three Atmospheric Reanalysis: CFSR, MERRA, ERA-Interim. A thesis submitted to the Faculty of the Graduate School of the University of Colorado in partial fulfillment of the requirement for the degree of Master of Arts, Department of Geography, 2014.

Andreas Bauwe, Petra Kahle, Bernd Lennartz, 2016. Hydrologic evaluation of the curve number and Green and Ampt infiltration methods by applying Hooghoudt and Kirkham tile drain equations using SWAT. Journal of Hydrology, 537(2016) 311-321.

Arnold J. G. et al, 2010. Assessment Of different representations of spatial variability On Swat Model performance.

ASCE. 1993. Criteria for evaluation of watershed models. J. Irrig. Drainage Eng. 119(3): 429-442.

Arnold J. G., Moriasi D. N., Gassman P. W., Abbaspour K. C., White M. J., Srinivasan R., Santhi C., Harmel R. D., van Griensven A., Van Liew M. W., Kannan N., Jha M. K., 2012.

SWAT: MODEL USE, CALIBRATION, AND VALIDATION, Vol. 55(4): 1491-1508, American Society of Agricultural and Biological Engineers ISSN 2151-0032

Assayew Nebere, 2016. Using satellite based rainfall estimates for runoff modeling with the REW approach: the case of Upper Gilgel Abay catchment. A thesis submitted to the faculty of geo-information science and earth observation of the University of Twenty in partial fulfillment of the requirement for the degree of masters of Science in geo-information science and earth observation.

Atinkut M., 2015. College Of Dryland Agriculture and Natural Resources Department of Land Resources Management and Environmental Protection (LaRMEP). Module on Integrated Water Resouces Management (WRIM2062), Mekele University.

Balasubramanian, 2017. Digital Elevation Model (DEM) in GIS, Centre for Advanced Studies in Earth Science, University of Mysore, Mysore.

Berhanu B., Yilma Seleshi, Solomon S. Demisse and Assefa M. Melesse, 2016. Bias correction and characterization of climate forecast system re-analysis daily precipitation in Ethiopia using fuzzy overlay. Meteorological Applications, Meteorol. Appl. (2016).

Biswas AK, 1970. History of Hydrology, North Holland Publishing Company, Amsterdam, p 336

Boyle, D. P., Gupta H. V., and Sorooshian S., 2000. Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. Water Resour. Res.36 (12): 3663-3674

Catherine Kuhn, 2014. Modeling rainfall-runoff using SWAT in a small urban wetland. Tests of scale for the Soil and Water Assessment Tool hydrology model, Yale University School of Forestry and Environmental Studies.

Claudia Teutschbein, 2012. Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. Journal of Hydrology journal homepage: www.elsevier.com/locate/jhydrol

Daniel de Almeida Bressiani, 2014. Effects of spatial and temporal weather data resolutions on stream flow modeling of a semi-arid basin, Northeast Brazil.

Daniel R. Fuka, 2013. Using the climate forecast reanalysis as weather input data for watershed model. Hydrological processes, hydrol. Process (2013).

Daniel Richard Fuka, 2015. Simplifying Watershed Modeling. A Dissertation Presented to the Faculty of the Graduate School of Cornell University in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy.

Darcy H, 1856. Les fontaines publiques de la ville de Dijon. Victor Dalmont, Paris

Dina K. Saleh, Charles R. Kratzer, Colleen H. Green, and David G. Evans, 2004. Using the Soil and Water Assessment Tool (SWAT) to Simulate Runoff in Mustang Creek Basin, California. Scientific Investigations Report 2009–5031 U.S. Department of the Interior U.S. Geological Survey.

Dina K. Saleh, Charles R. Kratzer, Colleen H. Green, and David G. Evans, 2009. Using the Soil and Water Assessment Tool (SWAT) to Simulate Runoff in Mustang Creek Basin, California. USGS, Scientific Investigations Report 2009–5031

Dirol G. T., 2008. Evaluation of reanalysis rainfall estimates over Ethiopia. International Journal of Climatology, Int. J. Climatol. 29: 67–78 (2009).

Elias Nkiaka, Nawaz N. R. and Jon C. Lovett, 2017. Evaluating Global Reanalysis Datasets as Input for Hydrological Modelling in the Sudano-Sahel Region. Hydrology 2017, 4(1), 13; doi:10.3390/hydrology4010013

Escobar F., Assoc Prof G. Hunter, Assoc Prof I. Bishop, Dr A. Zerger. Introduction to GIS, Department of Geomatics, the University of Melbourne

Fang G. H., 2015. Comparing bias correction methods in downscaling meteorological variables for a hydrologic impact study in an arid area in China. Hydrol. Earth Syst. Sci., 19, 2547–2559, 2015www.hydrol-earth-syst-sci.net/19/2547/2015/doi: 10.5194/hess-19-2547-2015.

Gan T. Y., Dlamini E. M., and Biftu G.F., 1997. Effects of model complexity and structure, data quality, and objective functions on hydrologic modeling. J. Hydrol.192 (1): 81-103.

Gassman P. W. et al, 2007. The Soil and Water Assessment Tool: Historical Development, Applications, and Future Research Directions. Invited Review Series

Gokhan Cuceloglu, Abbaspour Karim C. and Izzet Ozturk, 2017. Assessing the Water-Resources Potential of Istanbul by Using a Soil and Water Assessment Tool (SWAT) Hydrological Model

Gosain A.K., Mani A. and Dwivedi C., 2009. Hydrological modeling literature review: Report No.1. Indo-Norwegian Institutional Cooperation Program 2009-2011.

Holvoet K., (2006). Monitoring and Modeling the Dynamic Fate and Behavior of Pesticides In River Systems At Catchment Scale, Thesis submitted in fulfillment of the requirements for the degree of Doctor (PhD) in Applied Biological Sciences.

HU Sheng et al, 2017. Evaluation of the applicability of climate forecast system reanalysis weather data for hydrologic simulation: J. Geogr. Sci. 2017, 27(5): 546-564.

Kindie, 2016. Assessment on Surface Water Potential and Demands of Wabishebele Basin in Ethiopia. A thesis submitted and presented to the School of Graduate Studies of Addis Ababa University in Partial fulfillment of the Degree of Masters of Science in Civil Engineering (Hydraulic Engineering).

KRYSANOVA et al, 2008. Advances in eco-hydrological modeling with SWAT—a review. Hydrological Sciences Journal

Legates, D. R., and McCabe G. J., 1999. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. Water Resour. Res.35 (1): 233-241.

Ma L., Ascough J. C. II, Ahuja L. R., Shaffer M. J., Hanson J. D., and Rojas K. W., 2000. Root Zone Water Quality Model sensitivity analysis using Monte Carlo simulation. Trans. ASAE43 (4): 883-895.

Malual Deng, 2015. Assessment of Surface Water Resources and its Allocation: Case Study of Bahr el-Jebel River Sub-Basin, South Sudan. A thesis submitted in Partial Fulfillment of the Requirements for the Master of Science Degree in Integrated River Basin Management (IRBM).

Margaret Wambui, 2018. Bayesian Bias Correction of Satellite Rainfall Estimates for Climate Studies. Faculty of Geo-Information Science and Earth Observation, University of Twente, 217 7500 AE Enschede, The Netherlands; j.c.b.hoedjes@utwente.nl (J.C.B.H.); z.su@utwente.nl (Z.S.).

Mulvany TJ, 1850. On the use of self-registering rain and food gauges. In: Proceedings of the Institute Civil Engineers 4(2): 1–8, Dublin, Ireland.

MWAR –LAC, 2016. Managing water resource in Arid and Semiarid regions of Latin America and Caribbean. United Nations Educational, Scientific and Cultural Organization. International Hydrological Programe

Naomi Cambien, 2016-2017. Valuation of the Soil and Water Assessment Tool (SWAT) To Simulate Pesticide Dynamics in the Guayas River Basin (Ecuador). Master’s Dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Master of Science in Bioscience Engineering: Environmental Technology.

NCAR UCAR, 2018. Research data archive computational and information systems lab. https://rda.ucar.edu

Neitsch S.L, 2009. Soil and Water assessment tool theoretical documentation, version 2009. Agri life research and extension.

Refsgaard, J.C., 1996. Terminology, modelling protocol and classification of hydrologic model codes. Distributed Hydrologic Modelling, edited by Michael B. Abbott and Jens C. Refsgaard, 41-54.

Refsgaard, J. C. 1997. Parameterization, calibration, and validation of distributed hydrological models. J. Hydrol. 198(1): 69-97

Roth V. and Lemann T., 2016. Comparing CFSR and conventional weather data for discharge and soil loss modelling with SWAT in small catchments in the Ethiopian Highlands.

Ruben Dario Estrada, Marcela Quintero, Alonso Moreno, and Helle Munk Ravnborg, 2009. Payment for Environmental Services as a Mechanism for Promoting Rural Development in the Upper Watersheds of the Tropics, CPWF Project Report.

Seleshi Bekele Awulachew, 2010. Irrigation potential in Ethiopia Constraints and opportunities for enhancing the system. International Water Management Institute.

Srinivasan R., 2000. ArcGIS Interface for Soil and Water Assessment Tool (SWAT), Blackland Research and Extension Center and Spatial Sciences Laboratory Texas Agricultural Experiment Station Texas A&M University, http://www.brc.tamus.edu/swat
Stefan Liersch, 2003. The programs dew.exe and dew02.exe, user's manual, Berlin.

Suranjana Saha, 2010. The NCEP Climate Forecast System Reanalysis, Submitted to the Bulletin of the American Meteorological Society.

TENA B.A et al, 2015. Assessment of Spatio-Temporal Occurrence of Water Resources in Didessa Sub-Basin, West Ethiopia. International Journal of Civil, Structural, Environmental and Infrastructure Engineering Research and Development (IJCSEIERD) ISSN (P): 2249-6866; ISSN (E): 2249-7978 Vol. 5, Issue 1, 105-120.

Tesema Kebede, 2015. Characterizing the Groundwater Potetial of Wabi Shebelle River Basin. A thesis submitted to school graduate studies, Addis Ababa institute of Technology, in partial fulfillment for the degree of masters of Science in Civil Engineering under Hydraulic Engineering.

Tewodros Taffese, 2012. Physically Based Rainfall- Runoff Modelling In the Northern Ethiopian Highlands:-The Case of Mizewa Watershed. M.Sc Thesis, Water Resource Engineering Institute of Technology Bahir Dar University.

Tewodros W. et al, 2017. Evaluation of ERA-Interim, MERRA, NCEP-DOE R2 and CFSR Reanalysis precipitation Data using Gauge Observation over Ethiopia for a period of 33 years. AIMS Environmental Science, 4 (4): 596-620. DOI: 10.3934/environsci.2017.4.596

The Federal Democratic of Ethiopia Abbay Bassin Authority, 2016.

Tissa Tomya and Sumam K. S., 2015. Determining the Adequacy of CFSR Data for Rainfall-Runoff Modeling Using SWAT. International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST - 2015).

Tuppad et al, 2011. Soil and Water assessment tool (SWAT) hydrological /water quality model: extended capability and wider adoption.

UN-WATER, 2014. United Nations Educational, Scientific, and Cultural Organization World Water Assessment Program. National Water Development Report for Ethiopia (Final) Addis Ababa December 2004.

Van Griensven, A. and W. Bauwens, 2003. Multi objective auto calibration for semi distributed water quality models. Water Resour. Res. 39(12): 1348-1356.

Van Liew, M. W., J. G. Arnold, and D. D. Bosch. 2005. Problems and potential of autocalibrating a hydrologic model. Trans. ASAE48 (3): 1025-1040.

Vigerstol et al, 2011. A comparison of tools for modeling freshwater ecosystem services. Journal of Environmental Management.

Vijay P. Singh, 2018. Hydrologic modeling: progress and future directions. Singh Geosci. Lett. (2018) 5:15 https://doi.org/10.1186/s40562-018-0113-z

Vinayak S. et al, 2009. Measurement Errors in Tipping Bucket Rain Gauges under Different Rainfall Intensities and their implication to Hydrologic Models. An ASABE Meeting Presentation Paper Number: 097368.

Vincent Roth et al, 2016. Comparing CFSR and conventional weather data for discharge and soil loss modelling with SWAT in small catchments in the Ethiopian Highlands. Hydrol. Earth Syst. Sci., 20, 921–934, 2016www.hydrol-earth-syst-sci.net/20/921/2016/ doi: 10.5194/hess-20-921-2016.

Winfred B. Mbungu, 2017. Assessing the Hydrology of a Data-Scarce Tropical Watershed Using the Soil and Water Assessment Tool: Case of the Little Ruaha River Watershed in Iringa, Tanzania. Open Journal of Modern Hydrology, 2017, 7, 65-89 http://www.scirp.org/journal/ojmh

Worqlul A. W., Collick A. S., Tilahun S. A., Langan S., Rientjes T. H. M., and Steenhuis T. S., 2015. Comparing TRMM 3B42, CFSR and ground-based rainfall estimates as input for hydrological models, in data scarce regions: the Upper Blue Nile Basin, Ethiopia. Journal of Hydrology and Earth System Science. HESSD 12, 2081–2112, 2015.

WINCHEL M. et al, 2013. Arcswat Interface for Swat2012, User's Guide. Blackland Research And Extension Center Texas Agrilife Research 720 East Blackland Road - Temple, Texas 76502 Grassland, Soil And Water Research Laboratory Usda Agricultural Research Service808 East Blackland Road -Temple,Texas 76502.

Woldeamlak Bewket et al, 2015. Climate Change Impact, Vulnerability, Adaptation and Mitigation, IV, Water and Energy. Ethpiopian Acedamy of Science.

Yan Yang et al, 2014. Evaluation of Gridded Precipitation Data for Driving SWAT Model in Area Upstream of Three Gorges Reservoir. PLoS ONE 9(11): e112725. doi:10.1371/journal.pone.0112725

Yihun Taddele Dile, 2014. Evaluation of CFSR Climate Data for Hydrologic Prediction in Data scarce Watersheds: An Application in the Blue Nile River Basin. Journal of the American Water Resources Association.

Zahidul Islam, 2011. Literature Review on Physically Based Hydrologic Modeling. Department of Civil and Environmental Engineering, University of Alberta.