IET Smart Grid

Research Article

IET Journals
The Institution of Engineering and Technology

Optimal placement of data concentrators for expansion of the smart grid communications network

eISSN 2515-2947
Received on 28th January 2019
Revised 25th May 2019
Accepted on 1st July 2019
E-First on 12th August 2019
doi: 10.1049/iet-stg.2019.0006
www.ietdl.org

Todd Zhen[1], Tarek Elgindy[2], S.M. Shafiul Alam[2], Bri-Mathias Hodge[2,3] ✉, Carl D. Laird[1,4]

[1]*School of Chemical Engineering, Purdue University, West Lafayette, IN, USA*
[2]*National Renewable Energy Laboratory, Golden, CO, USA*
[3]*Department of Electrical, Computer and Energy Engineering, University of Colorado, Boulder, CO, USA*
[4]*Sandia National Laboratories, Center for Computing Research, Albuquerque, NM, USA*
✉ *E-mail: Bri-Mathias.Hodge@nrel.gov*

Abstract: Evolving power systems with increasing renewables penetration, along with the development of the smart grid, calls for improved communication networks to support these distributed generation sources. Automatic and optimal placement of communication resources within the advanced metering infrastructure is critical to provide a high-performing, reliable, and resilient power system. Three network design formulations based on mixed-integer linear and non-linear programming approaches are proposed to minimise network congestion by optimising residual buffer capacity through the placement of data concentrators and network routeing. Results on a case study show that the proposed models improve network connectivity and robustness, and increase average residual buffer capacity. Maximising average residual capacity alone, however, results in both oversaturated and underutilised nodes, while maximising either minimum residual capacity or total reciprocal residual capacity can yield much-improved network load allocation. Consideration of connection redundancy improves network reliability further by ensuring quality-of-service in the event of an outage. Analysis of multi-period network expansion shows that the models do not deviate significantly from optimal when used progressively (within 5% deviation), and are effective for utility planners to use for smart grid expansion.

1 Introduction

The adoption of new renewable standards and accelerated cost reductions are driving sharp growth in renewable energy technologies. In particular, the number of distributed solar photovoltaic (PV) installations is growing rapidly. However, as distributed solar power becomes an increasing fraction of total energy generation, the electric power grid must ensure continued reliability and cost-effectiveness. The development of the smart grid, as a natural evolution of the electric power grid, seeks to incorporate new technologies to support these distributed generation sources.

New grid devices such as phasor measurement units and smart meters (SMs) will be built to provide the existing supervisory control and data acquisition (SCADA) systems with vital power flow information, allowing utility operators to monitor and control power flow and maintain operational efficiency. The advanced metering infrastructure (AMI) is designed to measure, collect, process, and relay energy consumption data from these SMs to utility operators. The AMI is typically comprised of a multi-tiered network of various communication devices and technologies. The AMI is the most fundamental part of the smart grid and can pose a challenging problem to solve in terms of node placement and routeing design [1].

While technology selection and hardware capabilities are important to the effectiveness of the smart grid, the placement of hardware for AMI applications is equally as vital. Manual reconfiguration of network deployments is often impractical since devices and infrastructure are fixed. Multi-tiered AMI networks require sensors and gateways deployed at specific locations to operate effectively. Routeing assignments between nodes can have significant effects on network quality and performance. Bottlenecks and islanding can occur in the network due to highly congested nodes and poor routeing, leading to difficulties serving customers and ineffective system control. Careful planning of node placement and routeing design can significantly alleviate these quality-of-service (QoS) issues, while at the same time improving performance. Automatic placement and optimisation of network topology are critical to provide a cost-effective, high-performing, reliable, and secure communication network in support of the expanding smart grid.

In this work, we develop a procedure to optimally place communication hardware and route communication within AMI networks in a way that best maximises QoS performance for utility customers, provide redundant network connectivity in the event of a security threat or outage, and provide for effective expansion of the network.

To do so, we first propose three mixed-integer linear programme (MILP) approaches for optimal placement of data concentrators (DCs) to maximise the network residual buffer capacity while ensuring network connectedness between SMs and electrical control centres. We apply the models to a case study with empirical parameters and standards. We analyse the network residual buffer capacities of the results, comparing distributions of connections and placement topologies. We then extend the formulations to improve network resilience by imposing connection redundancy and conclude by testing formulation effectiveness in planning for smart grid communication network (SGCN) expansion.

This paper is organised as follows. In Section 2, we summarise and highlight related work. In Section 3, we describe the SGCN and its supporting technologies, as well as define assumptions about the network that we will use in the optimisation formulation. Section 4 outlines the mathematical formulations for the smart grid expansion problems (SGEPs). Section 5 describes the applied case study in detail. Results of the optimisation are presented in Section 6. Section 7 contains concluding remarks and future research directions.

2 Related work

In this work, we focused on the problem of expanding the distribution grid communications networks as part of a growing,

integrated smart grid. For consistency and conciseness, we will refer to this problem as the SGEP.

The SGEP is two-fold; it consists of (i) a facility location problem and (ii) a routeing assignment problem. Facility location and routeing assignment problems have been studied extensively for applications in supply chain, emergency, pipeline, and telecommunications networks [2, 3]. We refer to topology (or network) design as the process of placing nodes within a network, and to route design as the process of link assignment and bandwidth allocation between nodes.

There have been a number of previous studies on the use of mathematical programming approaches for designing wired/ wireless mesh networks about QoS, cost, and/or traffic allocation. Kojima *et al.* [4] investigated aggregation point placement in a wired SCADA network and proposed an MILP framework involving communication scheduling. Qiu *et al.* [5] presented an MILP formulation based on the capacitated facility location problem to optimally place Internet transit access points in a wireless mesh network, while also handling QoS requirements such as throughput and fault tolerance. Amaldi *et al.* [6] presented an MILP formulation for constructing mesh access points in the form of a multi-commodity network flow problem with consideration to network traffic allocation. Jahromi and Rad [7] designed a communication topology for meshed node placement and assignments of wired connections within the power communication network and formulated a multi-objective, non-linear problem considering cost due to network packet loss, delay, and budget. The previous work on wired networks and mesh networking cannot be directly applied to our context because mesh routeing requires different specifications and characteristics than smart grid AMI for distributed power systems, which are typically multi-layered systems with various interacting technologies.

Also, many related works reported using approximation algorithms and heuristic-based solvers to solve their respective location/routeing problems [5–11], citing high time complexities when solving non-deterministic polynomial-time (NP)-hard problems with exact methods. We instead argue that the problems, while complex, can be handled effectively with exact solution methods (i.e. branch-and-cut as implemented in state-of-the-art commercial solvers), and the benefits outweigh the costs in doing so. Approximation algorithms are typically fast and specifically tailored to the application, but provide minimal optimality guarantees. On the other hand, exact methods, by nature of the methods themselves, automatically provide measures of optimality. Given that the SGEP is a planning and design problem, fast solution times may be less of a concern than the advantages of optimality guarantees provided by exact methods.

There is existing work that investigates the optimal placement problem of DCs within a smart grid context. Huang and Wang [12] evaluate an MILP framework for constructing a three-layer smart grid AMI that focuses on ensuring network demand while minimising total cost and designed an approximation algorithm handle the problem. Klinkert [10] presents an MILP framework for repeater and data collector placement based on fixed-charge network flow within the context of a smart grid communications networks. They also consider a multi-tier AMI involving meters, repeaters, data collectors, and a central data centre, analogous to the AMI presented in this work. Their MILP formulation also focuses on ensuring network feasibility for a minimum cost placement. Tavasoli *et al.* [9] considered the optimal placement of data aggregation points (i.e. DCs) in hybrid wireless and wired communication network by minimising installation costs of the access points, fibre-optic backbone, and worldwide interoperability for microwave access (WiMAX) communication costs. Kong [13] studied the inter-network reliability problem concerning data aggregator placement, specifically the problem of where to place data aggregators to ensure communication and power robustness. They proposed a minimum cost-objective subject to QoS requirements and network robustness assurances. Each of these related works [9, 10, 12, 13] presented models that minimised cost while ensuring network demand is fulfilled. In this work, however, we argue that while optimising for cost alone is an effective short-term solution for cost-limiting scenarios, it is insufficient to guarantee network performance for the long term. A more rigorous objective, namely one that prioritises QoS and network capacity, is required to best ensure network longevity and future performance, especially for planning a long-standing smart grid network.

Gourdin and Klopfenstein [14] have compared different QoS objectives for communication routeing design. They condense the resource management solutions for handling QoS factors (i.e. bandwidth, delay, jitter, and packet loss rate) down to two essential criteria: avoiding congestion and limiting path length. They define congestion as the difference between arc capacity and arc flow and denote this relationship as residual capacity. They also describe three objective functions that can handle such congestion: (i) maximise the minimum arc residual capacity for all arcs in the network, (ii) maximise the total amount of available network bandwidth (also equivalent to maximising the total/average residual capacity across the network), and (iii) minimise the sum of the reciprocal arc residual capacities. They analysed the effectiveness of each of these objectives on a wired communication network. While their methods focus on the optimal routeing of arc flow (spliTable) with consideration to minimal communication path length (i.e. the number of hops required from source to destination), they did not consider node placement (i.e. network design).

Our application focuses on the distribution system communications networks with two-hop communication links and intermediate DC nodes. Owing to the nature of data concentration, we do not consider spliTable routes. Our approach prioritises node placement and introduces binary variables, yielding an MILP design problem. We focus on optimal handling of aperiodic data traffic by optimising the residual buffer capacity of DCs as an analogue to optimise QoS factors such as bandwidth, throughput, packet loss, and delay requirements in the network. We expand on the objectives proposed by Gourdin and Klopfenstein [14] within a topology design framework and extend them with the incorporation of communication redundancy to improve network reliability. Also, our models prove effective for multi-period expansion of the communications network. To the best of our knowledge, no prior work has been done on solving the problem of DC placement while prioritising QoS performance within an AMI network.

In summary, the placement of DCs to *maximise network residual buffer capacity under the impacts of budget constraints, network connectivity requirements, distance limitations due to path loss, and redundant communication links to enhance robustness for an expanding smart grid communications network* is a unique challenge that we aim to address in this paper.

3 Introducing the SGCN

We first describe the SGCN in detail and then propose an optimisation framework to handle the expansion of the network to ensure QoS guarantees and network connectedness. Specifically, we focus on the AMI, which is the data communication architecture between the SM and the meter data management system. The AMI is used to transfer real-time information from meters including fault, outage, and usage to the utility control centre, where utility operators will use the information to control power flow and maintain operational efficiency. The AMI typically involves a hierarchical network architecture utilising a variety of communications technologies [15]. We will focus on the AMI when referring to SGCN in this paper.

The general framework of the SGCN consists of a multi-tier system: (i) a home area network (HAN) consisting of smart appliances, Internet-of-things (IoT) devices, and distributed energy sources communicating to a nearby SM, (ii) a neighbourhood area network (NAN) consisting of DCs that receive the information from nearby SMs, process and relay the information onto, (iii) the edge router (ER) or a control centre of the energy provider within the wide area network (WAN) [10, 15, 16]. ERs contain secure fibre-optic connections to monitor stations on a SCADA network within the transmission system. A number of communication technologies for the SGCN can be used and can be divided into wired and wireless types [16, 17]. For this paper, we will focus on wireless communications technologies.

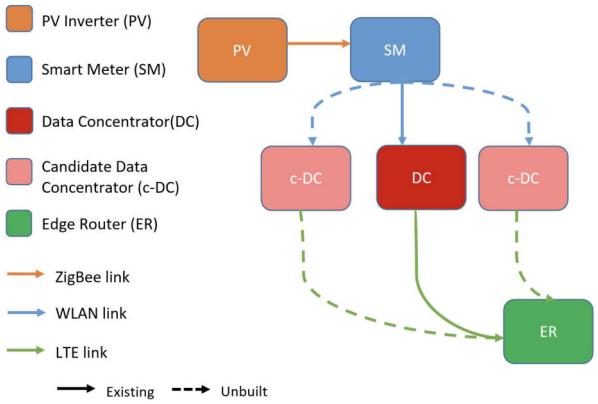

Fig. 1 *Network diagram*

3.1 Home area network

The HAN is localised to the customer domain and encompasses communication within the typical residential home including smart appliances, solar panels, light controls, and various other sensors and actuators. Home devices will send power readings, usage, and outages over the HAN to the SM outside the home for AMI application. A number of wireless technologies can be used within the HAN including Bluetooth, ZigBee, and IPv6 over low-power wireless personal area networks (6LoWPAN) [1].

Bluetooth, Institute of Electrical and Electronics Engineers (IEEE) 802.15.1 standard, is a low power, short-range radio-frequency (RF) standard operating on the 2.4 GHz band [18]. However, Bluetooth communication distances are short (between 1 and 100 m), are highly influenced by IEEE 802.11 networks, and offer weak security compared with similar network standards. 6LoWPAN is an LoWPAN built on the IPv6 routeing protocol and offers interoperability and QoS guarantees within an HAN. Disadvantages include network security and poor service discovery (i.e. automatically locating other nodes and higher-layer services) [1, 15]. ZigBee (built on the IEEE 802.15.4 standard) was developed by the ZigBee Alliance and is recommended as the common choice for HAN communication. ZigBee devices and routeing protocols allow for meshed networking within the HAN and provide QoS guarantees for information relaying across devices. Advantages include low cost, low power consumption, minimal data rates, and widespread usage in existing smart home devices. Additional information on ZigBee communications can be found in [1].

3.2 Neighbourhood area network

The NAN connects SMs to DCs for AMI applications and can involve a series of relays and mesh networking involving technologies such as IEEE 802.11 s RF mesh, IEEE 802.11 2.4 GHz wireless local area network (WLAN), WiMAX, power line communication (PLC), or cellular technologies. Compared to the HAN and WAN, communications technologies can vary significantly within the NAN, and often there is no smart grid-specific standard within this domain. For our purposes, we focus chiefly on wireless communications and will not discuss PLC technologies, though more information can be found in [1].

Mesh networking can present a cost-effective and viable solution for improved network resilience [19]. However, mesh networking can have significant signal interference and fading effects that make the quality of these connections unreliable, presenting crucial security problems. Methods to address these security issues in mesh routeing have been investigated [1]. WiMAX, based on the IEEE 802.16 standard, also provides promise with high-throughput, low-latency, high-security standards, and traffic management tools [20]. However, WiMAX capabilities are typically relegated to specific architectures and would require constructing a utility-proprietary network with high capital costs.

IEEE 802.11-based 2.4 GHz WLAN provides a robust, high-speed, point-to-point communication. Its ubiquity allows for easy installation and low costs. Moreover, WLAN technologies are undergoing continuous improvement of data rates, service, and coverage. High reliability can be achieved by proper routeing design and system design techniques [18]. WLAN is considered an excellent candidate for smart grid systems [15], specifically in remote metering and monitoring applications, and has been recognised by the National Institute of Standards and Technology via the International Electrotechnical Commission 61850 standard for application in smart grid environment [18].

While cellular technologies have been studied for SM transmissions, they are generally preferred for links between DCs and ERs [16].

3.3 Wide area network

The WAN provides connections between the smart grid and the core utility network. Within the WAN, communication between DCs and ERs can include technologies utilised in the NAN as well (e.g. cellular technologies, WiMAX, RF mesh, or broadband fibre-optic connections for wired options [1, 15]).

Cellular and long-term evolution (LTE) technologies [third-generation (3G)/4G on 824–894 MHz/1900 MHz spectrum] are on licenced frequency bands and provide benefits of low interference, high reception rate and information security, extensive data coverage, and no maintenance costs (since operation and network are maintained by carriers) [18]. Since utility control centres require high levels of reliability, LTE connections are suitable for WAN communication [20]. Additionally, the existing infrastructure for cellular communications allows for rapid deployment of smart grid communication hardware.

3.4 Defining the network graph

The network description addressed in this paper is summarised in Fig. 1. The SGCN of interest in this paper considers wireless communications between PV inverters, SMs, DCs, and ERs. In addition, we assume PV inverters transmit to SMs via ZigBee, SMs transmit to DCs via 2.4 GHz WLAN, and DCs relay to ERs via LTE technology. While PV inverters are typically installed close to SMs, DCs and ERs are installed separately from the PV inverter-SM HAN. We do not consider mesh networking or multi-hop communications in this work. Since the network communication is wireless, connections can be established if transmitters and receivers are within communication range of one another, but the assignment of connections will be decided by the optimisation formulation described in Section 4. We use the terms ‘link’, ‘arc’, and ‘connection’ interchangeably to denote the connections in the network. Specific parameters for the network are provided in Section 5. For additional information on smart grid technologies and practises, QoS requirements for smart grid integration to SCADA systems, and background on AMI, we refer the reader to [21].

3.5 QoS factors

There are several critical QoS factors that affect network performance: packet loss, path loss, effective throughput, network criticality, network availability, available bandwidth, latency, and connection outage probability [8]. For our purposes in providing a time-invariant, deterministic routeing assignment, and optimal placement of communication nodes, we will consider only packet loss, path loss, effective throughput, available bandwidth, and latency as QoS factors of interest in this work. We consider the time-varying and stochastic QoS parameters (network availability, criticality, and outage probabilities) as future work.

Packet loss is the fraction of transmitted packets lost in the network and commonly occurs due to queuing congestion and buffer overflows. The path loss depends on factors such as antenna height, frequency, and link distance. Effective throughput, the amount of data successfully passed through a link at a given time, is highly dependent on the processing within each network router to determine the transmission rate to forward the packet with a minimum number of hops [8]. Since packet loss and effective throughput are highly dependent on the effectiveness of underlying

Table 1 Numerical values for SUI model parameters

Model parameter	Value
d_0, m	100
s, dB	8.2–10.6
h_b, m	10–80
f, MHz	2400
h_r, m	6
α	4.0
β, m^{-1}	0.0065
ω, m	17.1
λ, m	0.125

network technologies and routeing protocols [5], we choose to focus on *path loss, available bandwidth*, and *latency* as the principal QoS factors when placing network hardware and assigning network routes.

Latency, or network delay, is defined as the time needed for a bit of data to travel from one node to another. Latency can be divided into four parts: (i) processing delay – the time taken for a node to process the packet header, (ii) queuing delay – the time a packet spends in routeing queues before being processed, (iii) transmission delay – the time taken to push the packet's bits onto the link, and (iv) propagation delay – the time for a signal to reach its destination. The propagation delay is dependent on the distance and geography between communication nodes, while the majority of latency (processing, queuing, and transmission delays) are dependent on the communication hardware and wireless routeing protocols [22].

In the SGEP, propagation delay is the relevant QoS metric of interest within latency, while processing, queuing, and transmission delays are handled by the hardware and wireless technologies (Park *et al.* [23] have recently developed a low-latency congestion control algorithm for wireless LTE networks). Lopez-Aguilera *et al.* [24] have shown that the propagation delay for outdoor IEEE 802.11 wireless networks is dependent on the slot time and translates up to 6 km communication range for IEEE 802.11b without performance degradation. For larger distances, the propagation delay becomes greater than the slot time and performance is degraded.

For our proposed approaches, we ensure path loss and propagation delay in the network are within QoS requirements by defining the effective communication range of wireless technologies to within the limits that ensure minimal performance degradation (see Section 3.6). Once the communication ranges are established, the optimisation can then consider optimal placements for wireless hardware and routeing decisions by maximising available network bandwidth.

Following the convention of [14], we refer to available network bandwidth as residual capacity, specifically on the capacities of buffers at each DC. Each DC has a buffer size that can handle incoming messages. Exceeding this buffer limit negatively impacts the packet loss rate, and increases communication delay [14]. Therefore, routeing assignments are made by maximising residual capacities across the network.

3.6 Path loss propagation model for communication radius

The maximum allowable distance between nodes (communication radius) can be predicted using a minimum threshold of received power required for signal reception and a model for signal propagation. The Stanford University Interim (SUI) propagation model [25, 26] is chosen to represent the 2.4 GHz WLAN link between SMs and DCs in our case study.

The SUI models are specified for three types of terrains (A, B, and C), where Type A is associated with hilly terrain, and moderate to heavy foliage with maximum path loss, Type B is associated with mostly either flat terrain with moderate to heavy tree densities or hilly terrains with light tree densities, and Type C is associated with flat terrain and light tree densities with minimum path loss. More information and associated parameters on the SUI models can be found in [25].

Given that the case study we will be considered in Section 5 is representative of a typical suburban area, we have elected to use the terrain Type B specification, though the proposed approach is general and can be used with Type A and Type C terrains. The path loss equations with correction factors and other parameters are defined below:

$$P = X_g + 10\gamma\log_{10}\left(\frac{d}{d_0}\right) + X_f + X_h + s \tag{1a}$$

$$X_g = 20\log_{10}\left(\frac{4\pi d_0}{\lambda}\right) \tag{1b}$$

$$\gamma = \alpha - \beta h_b + \frac{\omega}{h_b} \tag{1c}$$

$$X_f = 6.0\log_{10}\left(\frac{f}{2000}\right) \tag{1d}$$

$$X_h = -10.8\log_{10}\left(\frac{h_r}{2000}\right) \tag{1e}$$

where P is the path loss (defined as the ratio of the transmitted to received power); X_g is the free space or line-of-sight (LOS) gain; d is the distance between the transmitter (SM) and the receiver (DC) antennas; d_0 is the reference distance ($d > d_0$),;s is a shadow fading factor; h_b is the base station height above ground; γ is the path loss exponent; X_f and X_h are correction factors for the operating frequency and the antenna height, respectively; f is the operating frequency of the transmission signal; λ is the wavelength from the communication frequency; and h_r is the receiver antenna height above ground. The terrain constants α, β, and ω used in our study are given in Table 1, along with ranges and values for the other parameters in the path loss equation.

Bounding (1a) for maximum path loss and expanding P_{max}, we have

$$P_{max} = P^{trans} - P^{rec}_{min} \tag{2}$$

where P_{max} is the maximum path loss incurred when the received power is at its minimum; P^{trans} is the transmitted power (in dBm); and P^{rec}_{min} is the minimum received power (in dBm). From (1a)–(1e), we obtain a relationship for the communication distance in terms of P_{max}

$$d \le d_0 10^{P_T/10\gamma} \tag{3}$$

where $P_T = P_{max} - X_g - X_f - X_h - s$. SM transceivers can operate at peak transmitted power up to 1 W (30 dBm), but typically averages low transmitted power since the transmission intervals are very short and infrequent [27]. We, therefore, assume average transmitted powers of $P^{trans} = 0.2$ W (23.0 dBm) for our test case [28]. Minimal received signal powers for wireless networks (IEEE 802.11 variants) are typical $P^{rec}_{min} = -100$ dBm. Combining these power specifications leads to $P_{max} = 123$ dB. Therefore, the SM communication radius, by the SUI propagation model parameters in Table 1, is $d \le 0.93$ km with $s = 9.0$ and $h_b = 10$, to remain within minimal received signal power range for all receivers. Besides, since this communication range for 2.4 GHz WLAN is less than the reported distance to ensure minimal propagation delay (i.e. 6 km according to Lopez Aguilera *et al.* [24]), setting this as the distance limit between SMs and DCs will ensure acceptable QoS in the optimal design.

4 Smart grid expansion problem

A larger number of SMs will eventually be constructed as the smart grid grows, distributed generation becomes more widespread, and

an increasing number of IoT devices become commonplace. The SGEP focuses on placement of additional DCs (by selecting from a set of candidates denoted in Fig. 1 as c-DCs) and assignment of communication links between SMs and DCs to handle the load on the network while minimising congestion and improving network connectivity and robustness. We consider the residual buffer capacity of the network as the key metric to define congestion levels in the network.

In this section, we show three formulations that focus on improving network residual buffer capacity within an MILP to form the SGEP.

The SGEP–average (SGEP–A) focuses on maximising average residual buffer capacity and is defined by the following MILP:

$$\underset{x,\,y}{\text{maximise}} \quad \sum_{l \in L} r_l \tag{4a}$$

$$\text{subject to} \quad \sum_{l \in L} y_l \leq K \tag{4b}$$

$$x_{a,\,l} \leq y_l, \qquad \forall l \in L,\, a \in \mathscr{A}_l \tag{4c}$$

$$\sum_{l \in \mathscr{L}_a} x_{a,\,l} = 1, \quad \forall a \in A \tag{4d}$$

$$r_l = b_l y_l - \sum_{a \in A_l} x_{a,\,l} f_a, \quad \forall l \in L \tag{4e}$$

$$y_l \leq \sum_{a \in \mathscr{A}_l} x_{a,\,l}, \quad \forall l \in L \tag{4f}$$

$$y_l,\, x_{a,\,l} \in \{0,\, 1\}, \quad \forall l \in L,\, a \in \mathscr{A}_l \tag{4g}$$

$$r_l \geq 0, \quad \forall l \in L \tag{4h}$$

where A is the set of all SMs that must be serviced; L is the set of DCs (existing and candidate); and $\mathscr{L}_a$ is the set of candidate DCs within communication range of SM a, as determined by the path loss model from (3). Here, $\mathscr{A}_l$ is the set of all SMs that are within the communication range of candidate DC l [i.e. the distance between SM a and DC l is less than or equal to d defined in (3)]. The binary variable y_l indicates if a DC at a location l is selected for installation, binary variable $x_{a,\,l}$ indicates if SM a is assigned a connection to DC l, and r_l is a variable defined as the residual buffer capacity for DC l. Parameters f_a denote the flow of data or amount of data to be processed daily (depending on units of buffer capacity) from SM a; K defines the limit on the total allowable number of DCs in the network (includes existing DCs); and b_l is the buffer capacity on DC l.

Problem formulations (4a)–(4h) seek to select the set of DCs that should be built (y_l) along with the connection routeing that maximises the residual capacity of the constructed network. Constraint (4b) limits the total number of newly installed DCs to be within budget K. Constraint (4c) enforces that a connection between SM a and DC l can occur only if DC l is installed and is within communication range of a. Constraint (4e) describes the residual capacity, defined as the difference between the allowable bandwidth and the total throughput across each DC. This definition corresponds to residual bandwidth in [14]. Constraint (4f) enforces that a DC will not be installed unless at least one SM is assigned. Note that if $y_l = 0$ (the DC l is not selected), then no connections are possible for that concentrator (i.e. $x_{a,\,l} = 0\ \forall a \in A$ and $r_l = 0$ from constraint (4e). Also, y_l is fixed to 1 for all existing DCs. Constraint (4g) defines the binary variables in the optimisation, while constraint (4h) binds the residual capacity to be non-negative.

While the average residual buffer capacity seems such as a reasonable metric, it does nothing to ensure bandwidth is balanced. Consider a simple case of two DCs, each with bandwidth b. A design that results in each DC at 50% residual bandwidth has the same average as a design, where one DC is at 10% and the other at 90%. Therefore, this formulation would conceivably yield saturated nodes due to the nature of the averaging metric. Therefore, we also reformulate SGEP–A to (i) maximise the minimum residual buffer capacity and (ii) to minimise the sum of the reciprocal residual buffer capacities. These alternative objectives present options for penalising oversaturated nodes in the network, while also handling node placement.

The reformulation to maximise the minimum residual buffer capacity, called the SGEP–maximin (SGEP–MM), is as follows:

$$\underset{x,\,y,\,z}{\text{maximise}} \quad z \tag{5a}$$

$$\text{subject to} \quad r_l \geq z, \quad \forall l \in L \tag{5b}$$

$$\sum_{l \in L} y_l \leq K \tag{5c}$$

$$x_{a,\,l} \leq y_l, \quad \forall l \in L,\, a \in \mathscr{A}_l \tag{5d}$$

$$\sum_{l \in \mathscr{L}_a} x_{a,\,l} = 1, \quad \forall a \in A \tag{5e}$$

$$r_l = b_l - \sum_{a \in A_l} x_{a,\,l} f_a, \quad \forall l \in L \tag{5f}$$

$$y_l \leq \sum_{a \in \mathscr{A}_l} x_{a,\,l}, \quad \forall l \in L \tag{5g}$$

$$y_l,\, x_{a,\,l} \in \{0,\, 1\}, \quad \forall l \in L,\, a \in \mathscr{A}_l \tag{5h}$$

$$r_l \geq 0, \quad \forall l \in L \tag{5i}$$

Objective (5a) now maximises an auxiliary variable z, defined as the minimum residual capacity over all connections in the network, enforced through constraint (5b). The objective function in SGEP–A was focused on maximising the sum of the residual buffer capacity. For SGEP–A, if a DC was not selected, r_l must be forced to zero for that specific DC. In the SGEP–MM formulation, we are maximising the worst-case buffer capacity. This worst-case should be identified for only the selected DCs. Therefore, (5f) is written to ensure that unselected DCs do not impact this worst-case objective. That is, numerically, the residual capacity r_l for unselected DCs are set equal to the total bandwidth to ensure that they are not impacting the objective function as written. All other constraints remain the same as those of SGEP–A.

As opposed to the SGEP–A formulation, which optimises for maximum average residual capacity over the entire network, the proposed formulation SGEP–MM improves reliability by assigning connections to DCs in an attempt to maximise the worst-case or *minimum* residual capacity. By ensuring that the minimum residual buffer capacity is maximised, utility planners can be confident about minimum QoS guarantees the system will be able to provide.

The reformulation to minimise the sum of the reciprocal residual buffer capacities, called the SGEP–reciprocal (SGEP–R), is as follows:

$$\underset{x,\,y}{\text{minimise}} \quad \sum_{l \in L} \frac{1}{r_l} \tag{6a}$$

$$\text{subject to} \quad \sum_{l \in L} y_l \leq K \tag{6b}$$

$$x_{a,\,l} \leq y_l, \quad \forall l \in L,\, a \in \mathscr{A}_l \tag{6c}$$

$$\sum_{l \in \mathscr{L}_a} x_{a,\,l} = 1, \quad \forall a \in A \tag{6d}$$

$$r_l = b_l - \sum_{a \in A_l} x_{a,\,l} f_a, \quad \forall l \in L \tag{6e}$$

$$y_l \leq \sum_{a \in \mathscr{A}_l} x_{a,l}, \quad \forall l \in L \tag{6f}$$

$$y_l, x_{a,l} \in \{0, 1\}, \quad \forall l \in L, a \in \mathscr{A}_l \tag{6g}$$

$$r_l \geq 0, \quad \forall l \in L \tag{6h}$$

Objective (6a) now minimises the sum of the reciprocal residual buffer capacities in the network. Constraint (6e) has been modified in the same manner as SGEP–MM. SGEP–R follows the same assumption that the residual capacities of unselected DCs are equal to its bandwidth rather than zero for the reciprocal objective to perform as expected. All other constraints remain the same as those of SGEP–A.

While SGEP–A maximises the total residual capacity of the network, it fails to consider the distribution of links to each node. In other words, the SGEP–A objective is indiscriminant of a network with oversaturated nodes and concentrated links versus a network with evenly distributed links. The SGEP–R objective, however, does seek an even distribution of links in the network while also maximising the total residual capacity. The reciprocal objective function used in SGEP–R has been referred to as a tractable penalty model, falling between the extreme approaches used in SGEP–A and SGEP–MM in terms of the number of connections routed to DCs [14].

Owing to the reciprocal objective, SGEP–R becomes a mixed-integer non-linear programming problem and can be challenging to solve. Fortunately, the non-linearity is convex, and this problem can be approximated to arbitrary accuracy by applying linear underestimators to approximate the non-linear function. First, linear underestimators can be applied to the non-linear function using a Taylor series expansion as follows:

$$\frac{1}{r_l} \simeq \frac{1}{r^*_{l,m}} - \frac{1}{r^{*2}_{l,m}}(r_l - r^*_{l,m}), \quad \forall l \in L, m \in M_l \tag{7}$$

where $r^*_{l,m}$ denotes the points along the domain of r_l (denoted by $m \in M_l$) to apply the underestimators. These points can be generated uniformly or algorithmically to improve convergence times. In our cases, we have chosen to generate them uniformly. Besides, because the non-linear function is convex for the minimisation objective, the constraint can be converted to an inequality. The result is the following relaxed SGEP–R formulation:

$$\underset{x,\,y,\,v}{\text{minimise}} \quad \sum_{l \in L} v_l \tag{8a}$$

$$\text{subject to} \quad v_l \geq \frac{1}{r^*_{l,m}} - \frac{1}{r^{*2}_{l,m}}(r_l - r^*_{l,m}), \quad \forall l \in L, m \in M_l \tag{8b}$$

$$\sum_{l \in L} y_l \leq K \tag{8c}$$

$$x_{a,l} \leq y_l, \quad \forall l \in L, a \in \mathscr{A}_l \tag{8d}$$

$$\sum_{l \in \mathscr{L}_a} x_{a,l} = 1, \quad \forall a \in A \tag{8e}$$

$$r_l = b_l - \sum_{a \in A_l} x_{a,l} f_a, \quad \forall l \in L \tag{8f}$$

$$y_l \leq \sum_{a \in \mathscr{A}_l} x_{a,l}, \quad \forall l \in L \tag{8g}$$

$$y_l, x_{a,l} \in \{0, 1\}, \quad \forall l \in L, a \in \mathscr{A}_l \tag{8h}$$

$$r_l \geq 0, \quad \forall l \in L \tag{8i}$$

where v_l is an auxiliary variable for the reciprocal residual buffer capacity, which is approximated by constraint (8b). This relaxation allows the problem to be solved as an MILP. For the rest of this paper, SGEP–R will refer to the formulation comprising (8a)–(8i).

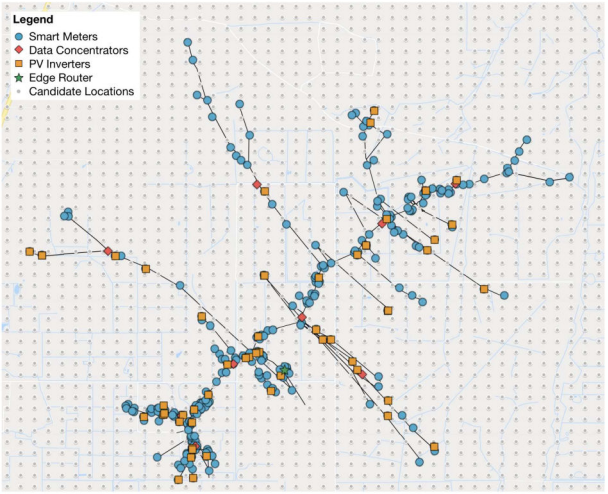

Fig. 2 *RTCA network diagram*

Table 2 RTCA topology

Node type	Total count
PV inverters	57
SMs	275
DCs	10
ER	1
candidate locations for DCs	1936

4.1 Redundancy

Also, due to the flexibility of the MILP framework, constraints in all three formulations can easily be modified to consider connection redundancy. To do so, we can modify constraints (4d), (5e), and (8e) and apply a redundancy threshold to the right-hand side as follows:

$$\sum_{l \in \mathscr{L}_a} x_{a,l} = C, \quad \forall a \in A \tag{9}$$

Constraint (9) provides redundancy if a DC is unavailable for a connected SM, either because of component failure or unexpected congestion, ensuring that each SM can connect to at least C DCs. For example, if $C = 2$, then this formulation ensures that each SM can connect to one of two DCs (e.g. primary and secondary) while ensuring that the DC can feasibly handle the combined bandwidth from all potential connections identified.

5 Case study

To show the application of the SGEP formulations for an SGCN, and to evaluate the capabilities of the modelling method in answering smart grid communication system problems, a case study designing the topology of a semi-urban SGCN was conducted.

The reference test case A (RTCA), shown in Fig. 2, was adapted from Pacific Northwest National Laboratory taxonomy feeder R2-25.00-1 [29] and represented a moderately populated suburban area. It was placed arbitrarily in an area with good solar resource data to provide representative scale for the network, but has no relation to actual locations depicted. The characteristics of the network are listed in Table 2.

The RTCA network spans 10.0 km from East to West and 8.0 km from North to South.

A grid layout of candidate locations for DCs was chosen, where new nodes are allowed to be placed at any vertex on the grid. The density of the grid can be assigned as a parameter to the optimisation. In this case, a 44×44 matrix of candidate locations

was used, totalling 1936 candidate locations with ∼200 m between each candidate node.

The communication range of the SMs via 2.4 GHz WLAN is dependent on the transmit power of the transmitters (SMs), the received power at receivers (DCs), terrain, elevation, and LOS. Sources report widely ranging values for physical ranges of 802.11 signals, from 25 m for 802.11a/b/g [17] to 80–450 m for 802.11ah (for smart cities in urban settings) [30] to 1.0 km for 802.11ac [17]. Instead, we will utilise a path loss propagation model to predict the communication range of WLAN, described in Section 3.6, particularly for a suburban topology that reflects the RTCA test case. Nevertheless, the approach taken in this research can be easily extended to urban environments that utilise closer range protocols (as in the IEEE 802.11ah standard for smart cities). The communication range is vital to construct the sets $\mathscr{L}_a$ and $\mathscr{A}_l$, which determine what SM–DC pairs are communicable.

Given the size of the system under consideration, the LTE communication range allows DCs to communicate with the ER with minimal loss regardless of their proximity to each other. Besides, each PV inverter is tied to an SM in the network, but additional SMs exist for which no PV inverters are linked. These additional SMs represent residents who have not yet integrated solar generation into their household, but will shortly.

Therefore, the SGEP is formulated to place additional DCs to ensure network connectivity for these emerging sources of distributed generation. A multi-period expansion will be performed using the proposed formulations to evaluate the effectiveness of the methods for network expansion over time. The optimisation is focused on the placement of the DCs and the assigned connections between the SMs and DCs.

Table 3 summarises the parameter values used in the optimisation. To facilitate the case study, the buffer capacities and data transfer sizes are kept constant across all DCs and SMs in the network, but the proposed formulations are amenable to varying values for these parameters.

Andreadou *et al.* [16] studied SM traffic, and DC processing for AMI systems, particularly the effects on the network for SM message sizes from 4 to 12,000 B and transmission intervals from 4 to 24 times/day. They showed that for an urban network of 101 m to 1 DC, 12,000 B transmitted 24 times/day from each meter required about 11 Mbits of data to be processed daily at each DC. For our case, the amount of data or size of packets to be processed daily from each SM was set to be at this maximum (11 Mbits) to simulate a scenario where a large amount of daily traffic occurs at each DC in the network. This also ensures that the resulting network will be able to handle increases in data rates from future demand. Since the approach relegates the handling of packet transmission and reception rates to the hardware technologies, the placement problem needs to only consider best locations to place the hardware and the specific routes to ensure network capacity can handle high-traffic events.

An LTE range of 50.0 km was chosen to represent a moderately populated suburban area for RTCA from [17]. For more rural environments, evolved universal mobile telecommunications system (UMTS) terrestrial radio access network can be used, which is an air interface for LTE cellular network that supports coverage up to 100 km with the acceptable performance [31]. For denser urban environments, LTE coverage ranges are typically only reported up to 20 km, according to the COST 231-Hata model [32]. The proposed approach is general and can be extended to urban or rural environments with an appropriate choice of LTE coverage.

Table 3 Parameter values for SGEP–A/–MM/–R

Model parameters	Value
b_l	640 Mbits [16]
f_a	11 Mbits [16]
WLAN range	0.93 km (see Section 3.6)
LTE range	50.0 km [17, 31]

Table 4 Numerical results for SGEP–A, SGEP–MM, and SGEP–R solutions

Budget K	Model	Objective value	Residual buffer capacity	
			Max, %	Min, %
8	SGEP–A	73.74	98.28	0.31
8	SGEP–MM	55.31	94.84	55.31
8	SGEP–R	1329.30	93.12	53.59
10	SGEP–A	76.37	98.28	0.31
10	SGEP–MM	67.34	93.12	67.34
10	SGEP–R	1328.56	93.12	63.91
25	SGEP–A	86.50	98.28	0.31
25	SGEP–MM	84.53	96.56	84.53
25	SGEP–R	1327.42	93.12	84.53
50	SGEP–A	92.12	98.28	0.31
50	SGEP–MM	89.69	98.28	89.69
50	SGEP–R	1327.07	94.84	84.53

6 Numerical results

In the first set of numerical results, we consider all three formulations (SGEP–A, SGEP–MM, and SGEP–R) with a single period, no redundant connections. The optimisation problems were constructed using Pyomo [33, 34] and solved with Gurobi Optimiser. Gurobi Optimiser is a commercial optimisation solver primarily used for linear, quadratic, and mixed-integer linear or quadratic programming problems [35]. Gurobi Optimiser is well-established in the operations research literature as a state-of-the-art solver for mixed-integer problems of this class. The results were obtained on a computer running a 2.4 GHz Intel Core i5, dual-core processor with 8 GB RAM. Table 4 lists the computational results for SGEP–A, SGEP–MM, and SGEP–R with different budgets of DCs (budget K) based on the RTCA network. A total of ten uniformly distributed linear underestimators were used for the reciprocal function for each DC in constraint (8b).

Given a budget limit, fixed assignments from constraint (4c), and the assumption of constant parameter values in this case study, the average (or total) residual buffer capacity of the network can be pre-computed assuming the budget allows a fully connected network. As such, SGEP–A [formulation (4)] reduces to a pure feasibility formulation. The SGEP–A formulation leads to both oversaturated connections and underutilised DCs, as evidenced by the unchanged maximum and minimum residual buffer capacities of 98.29 and 0.31%, respectively. The SGEP–MM formulation, on the other hand, seeks to maximise worst-case residual buffer capacity, yielding values of 55.31, 67.34, 84.53, and 89.69% for budget limits of 8, 10, 25, and 50 DCs, respectively. The SGEP–R formulation seeks to minimise the sum of the reciprocal of the residual buffer capacity and performs similarly to SGEP–MM, yielding minimum values of 53.59, 63.91, 84.53, and 84.53% for budget limits of 8, 10, 25, and 50 DCs, respectively.

The histogram comparisons in Figs. 3–5 show clearly the distribution of connections per DC in the solutions provided by SGEP–A, SGEP–MM, and SGEP–R for an increasing budget size. Since SGEP–A only considers feasibility, the distributions shown in Fig. 3 are not surprising – a large number of connections to a few DCs along with many underutilised DCs with only one or two connections. On the other hand, formulations SGEP–MM and SGEP–R result in more balanced distributions of data flow with the same average residual buffer capacity over the network.

The reciprocal objective in SGEP–R penalises designs with DCs that have low residual buffer capacity. As Fig. 4 shows, SGEP–R attempts to equalise the number of connections per DC across the network to reduce bottlenecks. This behaviour effectively reduces the maximum number of connections per DC by consolidating the distribution. As the budget limit for DCs is increased, this consolidation becomes more pronounced, yielding a progressively tighter distribution. By distributing connections effectively, the network robustness and QoS are improved in the case of node outages.

SGEP–MM behaves similarly to SGEP–R. SGEP–MM maximises the minimum residual buffer capacity of the network, which, in test cases, where the amount of data to be processed daily

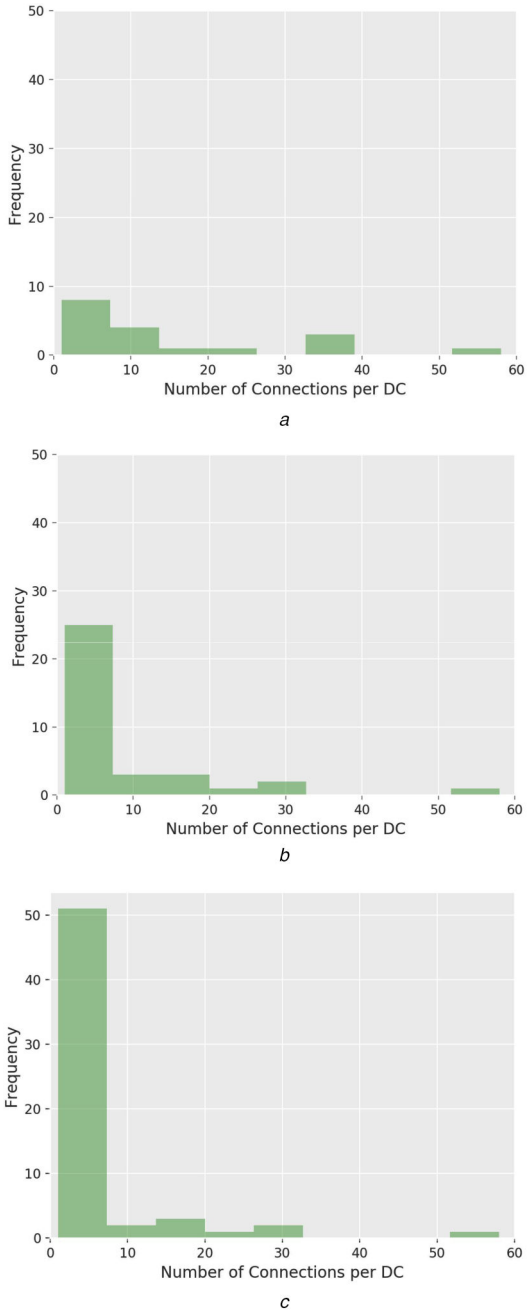

Fig. 3 *Distribution of connections per DC for SGEP–A solutions with budget limit (a) $K = 8$, (b) $K = 25$, and (c) $K = 50$*
(a) SGEP–A with $K = 8$, ***(b)*** SGEP–A with $K = 25$, ***(c)*** SGEP–A with $K = 50$

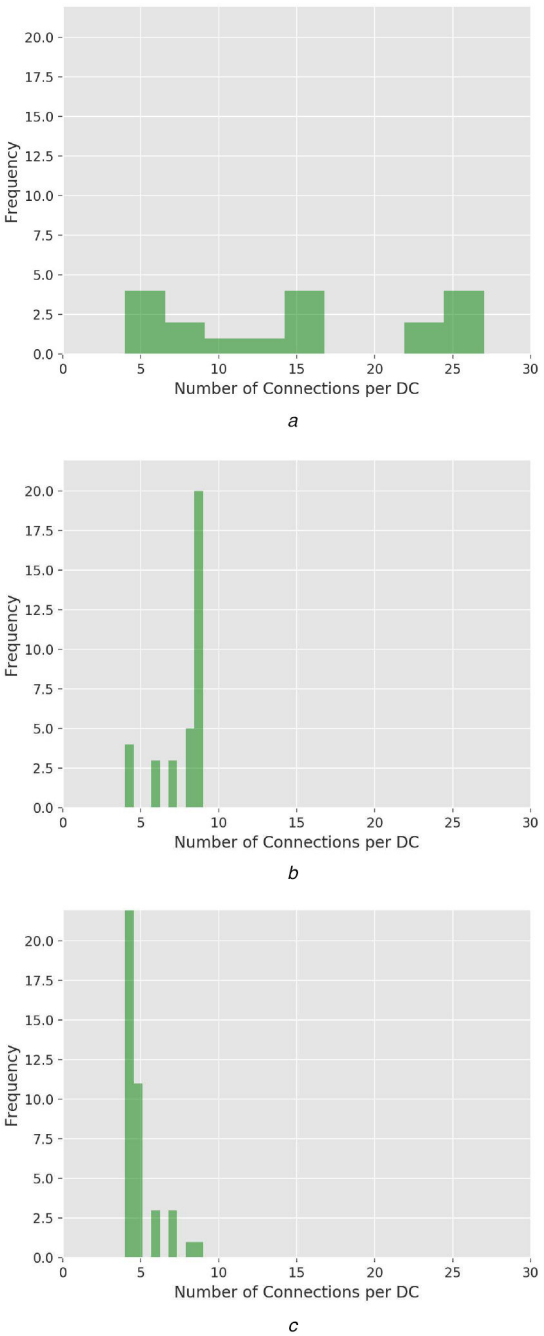

Fig. 4 *Distribution of connections per DC for SGEP–R solutions with budget limit (a) $K = 8$, (b) $K = 25$, and (c) $K = 50$*
(a) SGEP–R with $K = 8$, ***(b)*** SGEP–R with $K = 25$, ***(c)*** SGEP–R with $K = 50$

f_a for each SM and the buffer capacities b_l of each DC are equal, is equivalent to minimising the maximum number of connections per DC. Rather than reducing the maximum number of connections as a result of consolidating in the case of SGEP–R, SGEP–MM directly minimises the maximum value. By doing so, SGEP–MM not only distributes connections similar to SGEP–R to reduce congestion bottlenecks, but also provides the guarantee that the minimum residual buffer capacity is maximised. One disadvantage of SGEP–MM is that, by prioritising the minimisation of the high end of the distribution, SGEP–MM neglects the low end, as shown in Fig. 5*c*. This can lead to a few underutilised nodes. Compared to Fig. 4*c*, SGEP–R provides better management of the low end of the distribution, but does not provide the guarantees of SGEP–MM.

Fig. 6 shows the resulting placement of DCs following the SGEP–A (Fig. 6*a*), SGEP–MM (Fig. 6*b*), and SGEP–R (Fig. 6*c*) formulations for budget $K = 8$ DCs. This limit is also the fewest number required for the problem to remain feasible (i.e. fully connected). As observed, the placements can be distant from the feeder network. Since we have assumed a radius of communicability for each SM rather than distance-dependent performance, this result is not surprising.

Both SGEP–MM and SGEP–R assign connections to more effectively utilise each DC. As such, not only do they handle network demand more efficiently by spreading the load across DCs, but the resulting placement locations are also closer to the feeder network as a consequence of increasing the number of connections to DCs. By staying within vicinity of the feeder network, the solutions of SGEP–MM and SGEP–R become more practical for utility operators and contractors, since the infrastructure on which the DCs can be installed may already be in place. SGEP–R shows a slightly better result than SGEP–MM in terms of adhering to the feeder network, though the difference is minimal.

The SGEP–R formulation was also tested on a feeder-centric candidate topology, where the DCs were restricted to be placed only along existing utility poles. The findings from this test showed that the placements from the gridded topology were very similar to those from the feeder topology. The restricted feeder topology

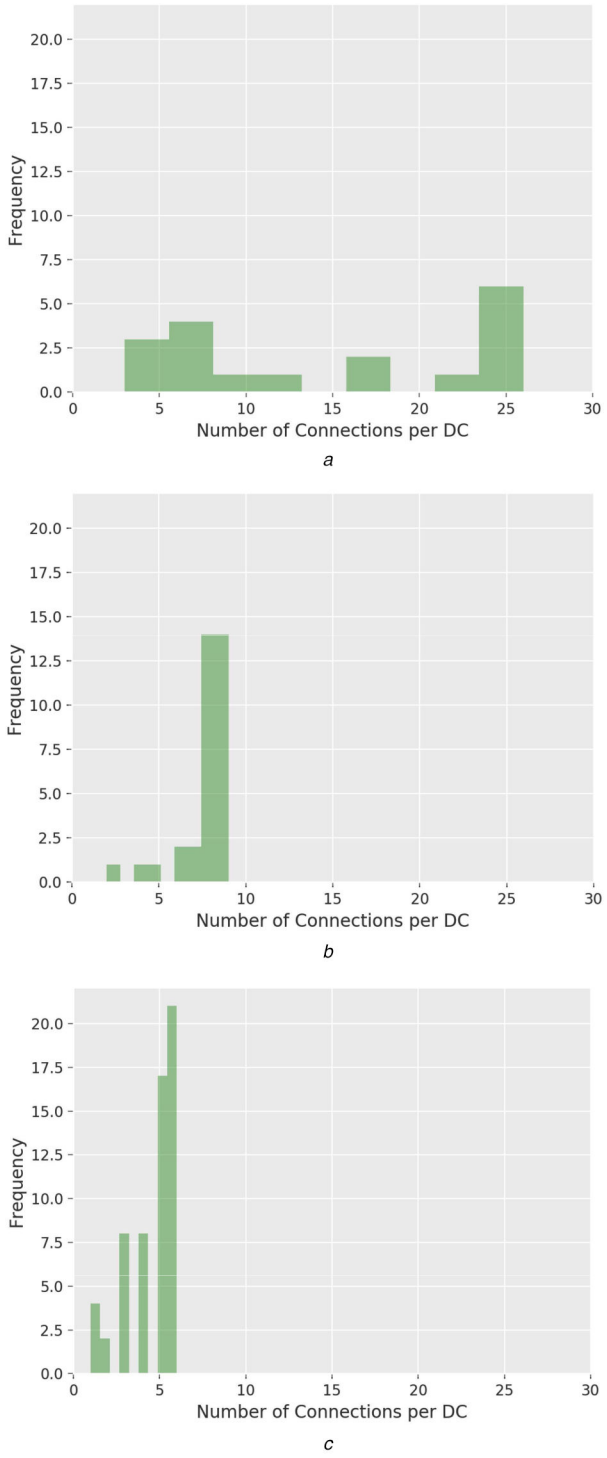

Fig. 5 *Distribution of connections per DC for SGEP–MM solutions with budget limit (a) $K = 8$, (b) $K = 25$, and (c) $K = 50$*
(a) SGEP–MM with $K = 8$, ***(b)*** SGEP–MM with $K = 25$, ***(c)*** SGEP–MM with $K = 50$

results in a similar number of connections per DC. Minor discrepancies in location of placement between the two were because the optimisation model considered any connecting node within communication radius to be of equal value. In other words, as long as an SM and corresponding DC were within communication range, there was no added value in locating the DC closer to the SMs. This is evident in the results of the gridded case, where placements of DCs can be seen in locations that extend outwards in directions away from the utility lines, in intermediate distances between its connected SMs.

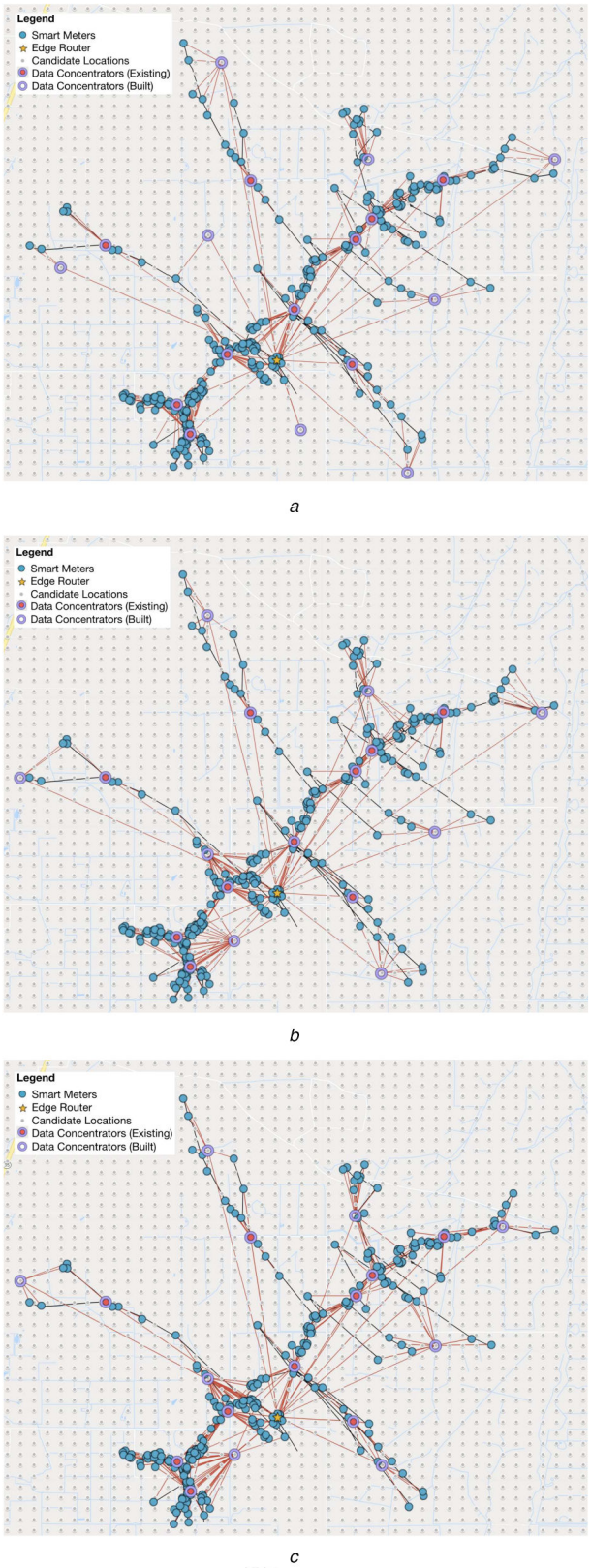

Fig. 6 *SGCN solution for eight newly installed DCs with (a) SGEP–A, (b) SGEP–MM, and (c) SGEP–R formulations. Black links denote the underlying power line feeder. Red links denote the connections from SMs to DCs and from DCs to the ER*
(a) SGEP–A, ***(b)*** SGEP–MM, ***(c)*** SGEP–R

6.1 System resilience

To address system resilience, we use the modified formulations that enforce redundancy. This formulation enforces multiple possible connections for each SM. With backup connections, we improve the resiliency and robustness of the network in the face of

Table 5 Numerical results for SGEP–A, SGEP–MM, and SGEP–R with redundancy $C = 2$

Budget K	Model	Objective value	Residual buffer capacity	
			Max, %	Min, %
20	SGEP–A	68.49	98.28	0.31
20	SGEP–MM	43.28	93.12	43.28
20	SGEP–R	1339.31	93.12	41.56
25	SGEP–A	72.99	98.28	0.31
25	SGEP–MM	62.19	93.12	62.19
25	SGEP–R	1335.84	91.41	62.19
30	SGEP–A	76.37	98.28	0.31
30	SGEP–MM	70.78	94.84	70.78
30	SGEP–R	1334.75	93.12	63.91
50	SGEP–A	84.24	98.28	0.31
50	SGEP–MM	82.81	96.56	82.81
50	SGEP–R	1333.10	94.84	81.09

possible security threats, outages, or node failures. In the case of a node outage, the formulation ensures that a nearby DC will be able to continue processing the information originally handled by the damaged unit since routeing assignments and placement strategies have accounted for the additional buffer capacity necessary.

Table 5 lists numerical results for SGEP–A, SGEP–MM, and SGEP–R with redundancy threshold $C = 2$, where $K = 20$ is the minimum number of additional DCs required for a feasible, fully connected network. Fig. 7 shows the resulting placement of DCs following the SGEP–A (Fig. 7*a*), SGEP–MM (Fig. 7*b*), and SGEP–R (Fig. 7*c*) formulations for budget $K = 20$ DCs and redundancy threshold of $C = 2$.

Since SGEP–A focuses on feasibility only, DCs are overutilised with very low residual buffer capacity (0.31%), along with underutilised DCs yielding high residual buffer capacities of 98.28%. On the other hand, SGEP–MM produces solutions with higher-minimum residual buffer capacities (43.28, 62.19, 70.78, and 82.81% for $K = 20$, 25, 30, and 50, respectively) while maintaining high-maximum residual buffer capacities as well (93.12, 93.12, 94.84, and 96.56% for $K = 20$, 25, 30, and 50, respectively). SGEP–R performs close to SGEP–MM, producing minimum residual buffer capacities of 41.56, 62.19, 63.91, and 81.09% for $K = 20$, 25, 30, and 50, respectively, and maximum residual buffer capacities of 93.12, 91.41, 93.12, and 94.84% for $K = 20$, 25, 30, and 50, respectively.

In comparison with the numerical results from $C = 1$ solutions, the redundancy solutions for $C = 2$ yield lower residual buffer capacities overall. This is not surprising, since consideration of redundancy requires the potential of additional throughput to DCs, leading to lower overall residual buffer capacities for the same number of nodes (minimum residual capacities of 84.53% for SGEP–MM with $C = 1$ versus 62.19% for SGEP–MM with $C = 2$ for $K = 25$, for example). However, as the number of DCs increases, this difference decreases (minimum residual capacities of 89.69% for SGEP–MM with $C = 1$ versus 82.81% for SGEP–MM with $C = 2$ for $K = 50$, for example). This effect is because the redundancy threshold, in our case, does not scale with the number of DCs, but with the number of SMs. Since the number of SMs in the network is fixed, increasing the number of allowable DCs will always serve to alleviate congestion by increasing total buffer capacity.

While this phenomenon holds for SGEP–MM and SGEP–R, the same cannot be said of SGEP–A. SGEP–A could still assign the maximum number of connections to a few DCs, regardless of the budget limit, while sparsely assigning the minimum number of connections to the rest of the DCs.

Analysing Fig. 7 shows all models placing DCs in locations that can be very close to others. This result is since, if a DC fails, the SMs in the area must be able to reconnect to an alternative DC nearby. An important trade-off occurs between the desire to prevent against node failure by applying redundancy versus effective utilisation of DCs. Nevertheless, SGEP–MM and SGEP–R handle

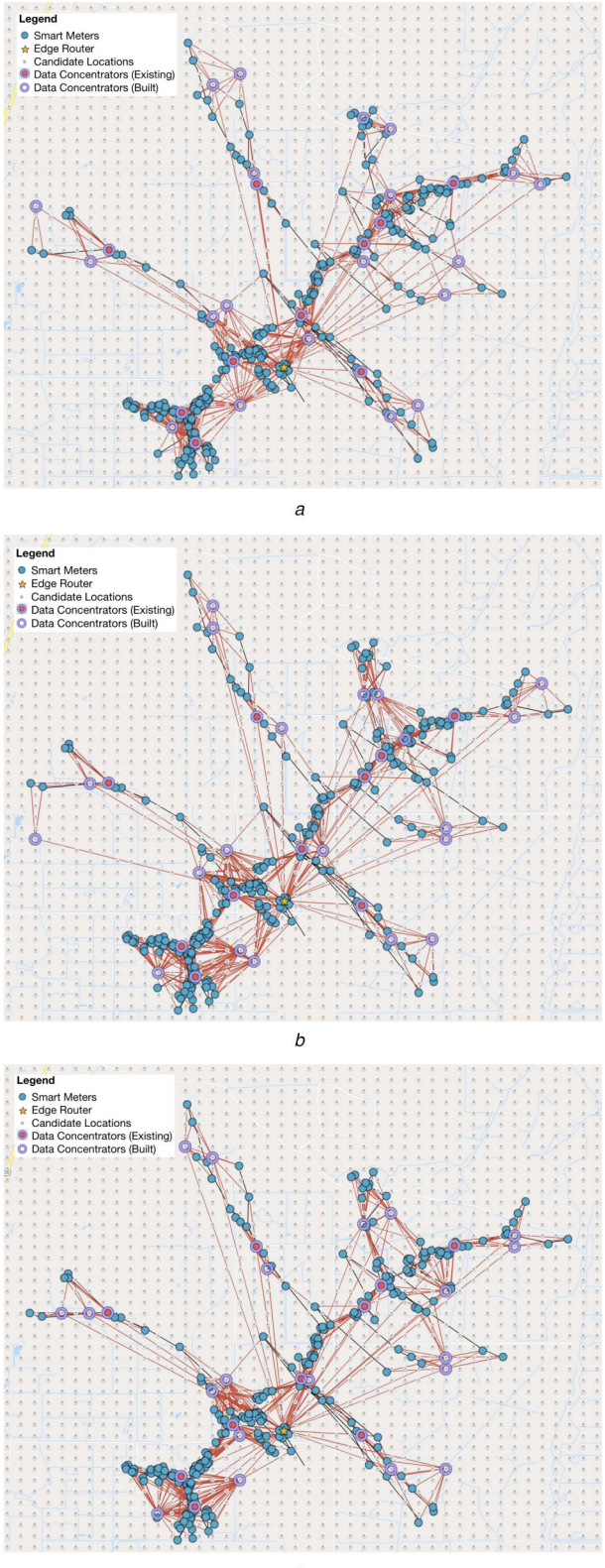

Fig. 7 *SGCN solution with (a) SGEP–A, (b) SGEP–MM, and (c) SGEP–R formulations for 20 newly installed DCs (redundancy $C = 2$)*
(a) SGEP–A with $C = 2$, ***(b)*** SGEP–MM with $C = 2$, ***(c)*** SGEP–R with $C = 2$

this trade-off well by more evenly distributing route assignments among DCs despite placing ones that can be near each other. Future formulations should consider addressing the assumption of independence of failure and spatial distribution of DCs.

6.2 Multi-period expansion

As this paper focuses on the expansion problem, it may be reasonable to assume that utility planners will need to resolve the

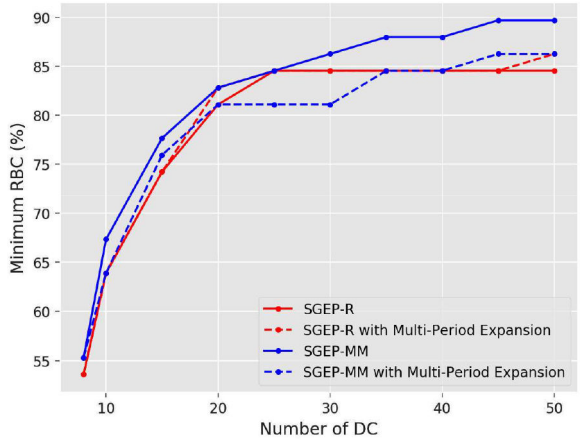

Fig. 8 *Comparison of SGEP–MM/–R versus SGEP–MM/–R with multi-period expansion*

expansion problem multiple times over decades as the smart grid grows and distributed generation spreads. Rather than solving a large-scale multi-period problem, we verify the effectiveness of a sequence of single period solves in a rolling horizon. In terms of the optimisation, we model this rolling horizon approach by progressively solving the expansion problem with an increasing budget while requiring that nodes from each prior placement is considered built.

Fig. 8 shows the system performance for a series of single-period optimisations (with five DCs added in each period). This is compared with the best possible solution obtained if the placements were re-optimised for the new number of DCs. From this comparison, we see that while placing a full set of DCs at once (with the basic SGEP–MM formulation) yields a higher-minimum residual capacity, the difference compared with the rolling horizon solution is minimal, with an average gap of 4%. This gap is dependent on the number of nodes utilised for each progressive addition. The minimum residual buffer capacities of the rolling horizon solutions and re-optimised solutions for SGEP–R are nearly identical and lie close to those of the rolling horizon solutions of SGEP–MM.

Investigation of SGEP–A was omitted here since the minimum residual buffer capacities of SGEP–A solutions remain unchanged, falling automatically to their extremes (minimum of 0.31%).

7 Conclusions and future work

A central issue in the expansion of the smart grid is the placement of additional DCs to service a growing network of SMs while maintaining a high level of reliability and QoS. In this research, three network design formulations based on MILP approaches were proposed to handle smart grid expansion with QoS guarantees: (i) SGEP–A – maximising the average residual buffer capacity, (ii) SGEP–MM – maximising the minimum residual buffer capacity, and (iii) SGEP–R – minimising the total reciprocal residual buffer capacity. We applied the design formulations to a case study with empirical parameters and solved them to optimality using Gurobi.

Our key conclusions in this work are as follows. First, our proposed formulations worked to design optimal placement topologies of DCs in smart grid communications networks. Second, the SGEP–A model can be useful for targeting (i.e. the minimum number of DCs for feasibility), but is poor for routeing assignment, since the total (or average) residual buffer capacity can be computed directly. Third, the SGEP–MM and SGEP–R models provide better solutions in terms of routeing assignment and leads to better placement within the vicinity of the feeder network, a practical consideration useful for ensuring that chosen placements can be built. Also, SGEP–MM provides minimum residual buffer capacity guarantees that are useful for establishing and ensuring QoS standards. Fourth, these formulations can address resilience concerns and are effective for rolling horizon network expansion.

Future avenues of study include evaluation of the proposed approaches on a variety of urban environments and network sizes, as well as showing the scalability of the approaches across a larger span of problems. Besides, investigation of the optimal spatial distribution when placing nodes is important, since the clustered placement of DCs can reduce their effective utilisation despite provisioning for node outages. Also, multi-period placement approaches with stochastic growth models for the expansion of the smart grid network over time can address concerns such as worst-case scenarios and placement under uncertainty. Finally, the investigation of statistical metrics such as conditional value-at-risk can be useful when considering alternative objectives for optimal placement prioritising risk mitigation.

8 Acknowledgments

This work was authored in part by the Alliance for Sustainable Energy, LLC, the manager and operator of the National Renewable Energy Laboratory for the U.S. Department of Energy (DOE) under Contract no. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Solar Energy Technologies Office. The views expressed in this paper do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting this paper for publication, acknowledges that the U.S. Government retains a non-exclusive, paid-up, irrevocable, worldwide licence to publish or reproduce the published form of this work or allow others to do so, for U.S. Government purposes. The Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under Contract no. DE-NA-0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in this paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

9 References

[1] Saputro, N., Akkaya, K., Uludag, S.: 'A survey of routing protocols for smart grid communications', *Comput. Netw.*, 2012, **56**, (11), pp. 2741–2771

[2] Daskin, M., Snyder, L., Berger, R.T.: 'Facility location in supply chain design', in: Langevin, A., Riopel, D. (Eds.): *Logistics Systems: Design and Optimization* (Springer, Boston, MA, 2005), pp. 39–65

[3] Fernández, E., Landete, M.: 'Fixed-charge facility location problems', in: Laporte, G., Nickel, S., Saldanha da Gama, F. (Eds.): '*Location science*' (Springer International Publishing, Cham, 2015, 1st edn.), pp. 47–77

[4] Kojima, H., Tsuchiya, T., Fujisaki, Y.: 'The aggregation point placement problem for power distribution systems', *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, 2018, **E101A**, (7), pp. 1074–1082

[5] Qiu, L., Chandra, R., Jain, K., *et al.*: 'Optimizing the placement of integration points in multi-hop wireless networks'. Proc. ICNP, Berlin Germany, 2004, vol. 4, pp. 1–12

[6] Amaldi, E., Capone, A., Cesana, M., *et al.*: 'Optimization models and methods for planning wireless mesh networks', *Comput. Netw.*, 2008, **52**, (11), pp. 2159–2171

[7] Jahromi, A.E., Rad, Z.B.: 'Optimal topological design of power communication networks using genetic algorithm', *Sci. Iranica*, 2013, **20**, (3), pp. 945–956

[8] Rastgoo, R., Sattari Naeini, V.: 'Tuning parameters of the QoS-aware routing protocol for smart grids using genetic algorithm', *Appl. Artif. Intell.*, 2016, **30**, (1), pp. 52–76

[9] Tavasoli, M., Yaghmaee, M.H., Mohajerzadeh, A.H.: 'Optimal placement of data aggregators in smart grid on hybrid wireless and wired communication'. 2016 Fourth IEEE Int. Conf. Smart Energy Grid Engineering, Sege, 2016, pp. 332–336

[10] Klinkert, A.: 'Designing smart-grid telecommunications systems via interval flow network optimization', Southern Methodist University, 2018

[11] Hassan, A., Zhao, Y., Pu, L., *et al.*: 'Evaluation of clustering algorithms for DAP placement in wireless smart meter network'. 2017 Ninth Int. Conf. Modelling, Identification and Control (ICMIC), Kunming, China, 2017, pp. 1085–1090

[12] Huang, X., Wang, S.: 'Aggregation points planning in smart grid communication system', *IEEE Commun. Lett.*, 2015, **19**, (8), pp. 1315–1318

[13] Kong, P.Y.: 'Cost efficient data aggregation point placement with interdependent communication and power networks in smart grid', *IEEE Trans. Smart Grid*, 2019, **10**, (1), pp. 74–83

[14] Gourdin, E., Klopfenstein, O.: 'Comparison of different QoS-oriented objectives for multicommodity flow routing optimization'. Proc. Int. Conf. Telecommunications, Funchal, Portugal, 2006, pp. 1–4

[15] Niyato, D., Wang, P.: 'Cooperative transmission for meter data collection in smart grid', *IEEE Commun. Mag.*, 2012, **50**, (4), pp. 90–97
[16] Andreadou, N., Kotsakis, E., Masera, M.: 'Smart meter traffic in a real LV distribution network', *Energies*, 2018, **11**, (5), pp. 1–27
[17] Kuzlu, M., Pipattanasomporn, M., Rahman, S.: 'Communication network requirements for major smart grid applications in HAN, NAN and WAN', *Comput. Netw.*, 2014, **67**, pp. 74–88
[18] Parikh, P.P., Kanabar, M.G., Sidhu, T.S.: 'Opportunities and challenges of wireless communication technologies for smart grid applications'. IEEE PES General Meeting, Minneapolis, USA, July 2010, pp. 1–7
[19] Xu, S., Qian, Y., Qingyang Hu, R.: 'Reliable and resilient access network design for advanced metering infrastructures in smart grid', *IET Smart Grid*, 2018, **1**, (1), pp. 24–30
[20] Patel, A., Aparicio, J., Tas, N., *et al.*: 'Assessing communications technology options for smart grid applications'. 2011 IEEE Int. Conf. Smart Grid Communications, Brussels, Belgium, October 2011, pp. 126–131
[21] Budka, K.C., Deshpande, J.G., Thottan, M.: '*Communication networks for smart grids making smart grid real*' (Springer, London, 2014, 1st edn.)
[22] Nagai, Y., Zhang, L., Okamawari, T., *et al.*: 'Delay performance analysis of LTE in various traffic patterns and radio propagation environments'. 2013 IEEE 77th Vehicular Technology Conf. (VTC Spring), Dresden, Germany, 2013, pp. 1–5
[23] Park, S., Lee, J., Kim, J., *et al.*: 'ExLL: an extremely low-latency congestion control for mobile cellular networks'. Proc. 14th Int. Conf. Emerging Networking Experiments and Technologies, Heraklion, Crete, 2019, pp. 307–319
[24] Lopez Aguilera, E., Casademont, J., Cotrina, J.: 'Propagation delay influence in IEEE 802.11 outdoor networks', *Wirel. Netw.*, 2010, **16**, (4), pp. 1123–1142
[25] Abhayawardhana, V.S., Wassell, I.J., Crosby, D., *et al.*: 'Comparison of empirical propagation path loss models for fixed wireless access systems'. 2005 IEEE 61st Vehicular Technology Conf., Stockholm, Sweden, 2005, vol. **1**, pp. 73–77
[26] Hari, K.V.S., Smith, M.S.: 'Channel models for fixed wireless applications', IEEE 802.16 broadband wireless access working group, 2003
[27] 'Smart meters: compliance with radio frequency exposure standards'. Available at https://www.gsma.com/publicpolicy/wp-content/uploads/2015/06/gsma_smart-meters_2015.pdf, accessed 25 October 2018
[28] Rupasinghe, N., Guvenc, I.: 'Licensed-assisted access for WiFi-LTE coexistence in the unlicensed spectrum'. 2014 IEEE Globecom Workshops (GC Wkshps), Austin, USA, 2014, pp. 894–899
[29] Schneider, K.P., Chen, Y., Chassin, D.P., *et al.*: 'Modern grid initiative: distribution taxonomy final report', Pacific Northwest National Laboratory, 2008
[30] Bellekens, B., Tian, L., Boer, P., *et al.*: 'Outdoor IEEE 802.11ah range characterization using validated propagation models'. GLOBECOM 2017 IEEE Global Communications Conf., Singapore, 2018, pp. 1–6
[31] Remy, J.G., Letamendia, C.: 'LTE standards and architecture', LTE standards, 2014, pp. 1–112
[32] Hamid, M.: 'Path loss models for LTE and LTE-A relay stations', *Univ. J. Commun. Netw.*, 2013, **1**, (4), pp. 119–126
[33] Hart, W.E., Watson, J.P., Woodruff, D.L.: 'Pyomo: modeling and solving mathematical programs in python', *Math. Program. Comput.*, 2011, **3**, (3), pp. 219–260
[34] Hart, W.E., Laird, C.D., Watson, J.P., *et al.*: '*Pyomo–optimization modeling in python*', vol. **67**, (Springer, Cham, 2017, 2nd edn.)
[35] 'Gurobi optimizer reference manual'. Available at http://www.gurobi.com, accessed 18 November 2018