Investigating the Ability of Smart Electricity Meters to Provide Accurate Low Voltage Network Information to the UK Distribution Network Operators

By

Goudarz Poursharif

Thesis submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy

Sheffield University Management School

Faculty of Social Sciences

The University of Sheffield

Engineering and Physical Sciences Research Council

February 2018

"The struggle itself towards the heights is enough to fill a man's heart. One must imagine Sisyphus happy."

Albert Camus (1942)-The Myth of Sisyphus and Other Essays

To My Mother's Memories

To My Father

To My Brother, Sister, and Niece

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor Dr Andrew Brint for his continuous support throughout my PhD and related research. He was always there for me when I needed moral and academic support. His patience, motivation, and deep knowledge of the field helped and inspired me during my research. I believe I could not have had a better supervisor and mentor for my PhD studies.

I would also like to thank my second supervisors Dr John Holliday and Dr Chantal Cantarelli. Dr Holliday provided me with a great learning environment at the Information School and was always available when I needed help and support. His unique taste in music and his great sense of humour made my PhD experience very enjoyable. Dr Cantarelli also provided me with great help and support during the final year of my PhD and always provided me with invaluable guidance and insightful feedback on my work.

I also wish to express my sincere thanks to my industrial supervisors at Northern Powergrid Dr Mary Black and Mr Mark Marshall, without whom this research opportunity would not have been possible. Their knowledge, comments, feedback, guidance drove this work forward and I have done my best so this work reflects the time, energy, and vision that they have dedicated to this research. I would especially like to thank Dr Mary Black who was a constant source of support to me throughout these four years. Besides my supervisors at Northern Powergrid, I would also like to thank Dr Chris Goodhand for providing this research opportunity to me and I also wish to thank Alan Creighton for his great insights and support during this research project. Further thanks must also be given to Mr Andrew Webster and Emma Burton who made it possible for me to access and obtain the necessary data for this project. I would also like to thank all the experts at Northern Powergrid that provided me with great knowledge and guidance during my PhD.

Finally and most importantly, I wish to thank my father Dr Yaghoub Poursharif and dedicate this thesis to him. He has always been there for me in times of difficulty and joy and has never stopped believing in me. He has always inspired me to reach for the stars and has constantly provided me with unconditional support and love. Having him in my life has always reassured me and made me a better and a more confident person.

Thesis Summary

By 2025, most households in the UK are provided with smart electricity meters, which will provide the Distribution Network Operators (DNOs) with new streams of customer data at Low Voltage (LV) network levels. Smart meters will enable the two-way transfer of information between the network operators and end user/generators. In the UK, smart meters will provide the DNOs with half-hourly customer demands which are required to be aggregated at LV level. This in theory can enable a more proactive management of the LV networks, which has not been possible before. However, as the electricity network becomes smarter and accommodates higher amounts of Distributed Generation (DG) and Low Carbon Technologies (LCTs) at lower voltage levels of the network, the questions have been raised whether the provision of smart meter data at half-hourly time resolutions and in aggregated formats can provide the network operators with sufficiently detailed information or more granular smart meter data is required for the smart LV grid applications to become a reality. This research presents a picture of the current status and the future developments of the LV electricity grid and the capabilities of the smart metering programme in the UK as well as investigating the major research trends and priorities in the field of Smart Grid. This work also extensively examines the literature on the crucial LV network performance indicators such as losses, voltage levels, and cable capacity percentages and the ways in which DNOs have been acquiring this knowledge as well the ways in which various LV network applications are carried out and rely on various sources of data.

This work combines 2 new smart meter data sets with 5 established methods to predict a proportion of consumer's data is not available using historical smart meter data from neighbouring smart meters. Our work shows that half-hourly smart meter data can successfully predict the missing general load shapes, but the prediction of peak demands proves to be a more challenging task. This work then investigates the impact of smart meter time resolution intervals and data aggregation levels in balanced and unbalanced three phase LV network models on the accuracy of critical LV network performance indicators and the way in which these inaccuracies affect major smart LV network application of the DNOs in the UK. This is a novel work that has not been carried out before and shows that using low time resolution and aggregated smart meter data in load flow analysis models can negatively affect the accuracy of critical low voltage network estimates.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ADMD	After Diversity Maximum Demand
ADMS	Advanced Distribution Management Systems
AI	Artificial Intelligence
AMIs	Advanced Metering Infrastructures
AMS	Advanced Metering Systems
ANM	Active Network Management
ANN	Artificial Neural Network
APE	Absolute Percentage Error
BAU	Business As Usual
BEIS	(Department for) Business, Energy and Industrial Strategy
BNN	Bayesian Neural Network
CHP	Combined Heat and Power
CLNR	Customer-Led Network Revolution
CREST	Centre for Renewable Energy Systems Technology
CVR	Conservative Voltage Reduction
DA	Distribution Automation
DCC	DataCommsCo.
DECC	Department of Energy and Climate Change
DG	Distributed Generation
DINIS	Distribution Network Information System
DMS	Distribution Management System
DNOs	Distribution Network Operators

DRM	Demand Response Management
DSE	Distribution State Estimation
DSM	Demand-Side Management
DSSE	Distribution System State Estimate
ECRC	Electricity Council Research Centre
ENA	Energy Network Associations
EPRI	Electric Power Research Institute
EPSRC	Engineering and Physical Sciences Research Council
ETP	European Technology Platform
GIS	Geographical Information Systems
GSP	Grid Supply Points
HV	High Voltage
ICO	Information Commissioner's Office
ICT	Information Communication Technology
IEC	International Electrochemical Commission
ISODATA	Iterative Self Organizing Data Analysis Technique
IT	Information Technology
JRC	(European) Joint Research Centre
kV	kiloVolts
kW	kiloWatts
kWh	killoWatt hours
LA	Load Allocation
LCN	Low Carbon Networks

LCNF	Low Carbon Network Fund
LCTs	Low Carbon Technologies
LIM	Loss Incentive Mechanism
LMF	Load Modelling Factors
LV	Low Voltage
MAPE	Mean Absolute Percentage Error
MD	Maximum Demand
MDI	Maximum Demand Indicator
MUA	Monthly Usage Allocation
MV	Medium Voltage
OFGEM	Office of Gas and Electricity Markets
OLTC	On Load Tap Changer
OMS	Outage Management System
PNN	Probabilistic Neural Network
PV	Photovoltaics
RBF	Radial Basis Flow
RMS	Root Mean Square
RTTR	Real Time Thermal Rating
SCADA	Supervisory Control and Data Acquisition
SE	State Estimation
SLC	Standard License Condition
SOM	Self Organising Maps
STLF	Short-Term Load Forecasting

SVR	Step Voltage Regulators
ToU	Time of Use
V	Volts

Chapter 1 Introduction

This thesis analyses how information provided by domestic smart electricity meters can improve the management of the low voltage network in the UK. This is important because of two factors. Firstly, the low voltage network forms a substantial part of the UK's electricity network. It comprises 48% of the total length of the distribution and transmission networks (EurElectric 2013). The amount of energy lost on the low voltage network due to the impedance of the conductors has been estimated as 5% of the energy that the network supplies (Sohn Associates Limited 2009). This compares with estimated losses of 3% on all of the rest of the distribution network.

Secondly, the low voltage network has not been actively managed in the way that the higher voltage networks are. This stems from there being very little knowledge of the sizes of the currents on the low voltage network due to the lack of monitoring points on the network. However, more active management is becoming increasingly desirable as the amount of embedded generation on the low voltage network rises. The United Kingdom government has made a commitment to reduce its CO_2 emission levels by 80% by 2020 compared to the levels in 1990 (DECC 2009). This requires a considerable rise in the amount of electricity produced from renewable sources of energy from 5.5% (in 2009) to 30% (DECC 2009). The latest reports by the Department for Business, Energy, and Industrial Strategy (BEIS) (2016) shows 29% increase in the production of energy from renewable sources of energy since 2014 that includes 87% increase in the amount of energy from solar panels. High levels of this intermittent generation bring the challenges of reversed power flows and rapid changes in the power flows as well as increased levels of network losses and voltage variation (Sohn Associates Limited 2013). This increases the need for higher levels of monitoring at the low voltage side of the network. However, the sheer size of the low voltage network means that the cost of installing more meters to rectify this deficiency is prohibitive.

The roll out of smart electricity meters to the UK's domestic customers is due for completion in 2025 (Smart Energy GB 2017). If the data from these meters can be utilised by the network operators to accurately model the low voltage currents, then this would be a highly cost effective way of solving the information gap, and so would provide the foundations for active management of the low voltage network. However, there are a number of issues that impinge on the accuracy of the currents, voltages and losses that are calculated from the smart meter data. Investigating the accuracy level that can be expected from using new streams of data from smart meters in the UK is the core of this research.

This chapter provides a background into the UK's electricity network and the role and the limitations of smart meters, based upon which the research aims and objectives are derived.

1.1 The UK's Electricity Network

Electrical power systems are usually broken down into 3 element types: generation, transmission, and distribution. In the UK, the larger generating units (power stations) normally deliver their energy at 132 kiloVolts (kV). As conductor losses depend on the square of the current (I^2), the voltage is normally stepped up to 275kV or 400kV before it is transmitted over long distances to reach the demand. This network is the transmission network and is owned by the National Grid. When the energy reaches closer to the demand points, it is stepped down to 132kV and then lower voltages before it reaches the customers.

The network from 132kV downwards is the distribution network and is run (and owned) by the UK's Distribution Network Operators (DNOs). There are now 6 DNOs running the 14 distribution networks in England, Scotland and Wales (see Figure 1-1).

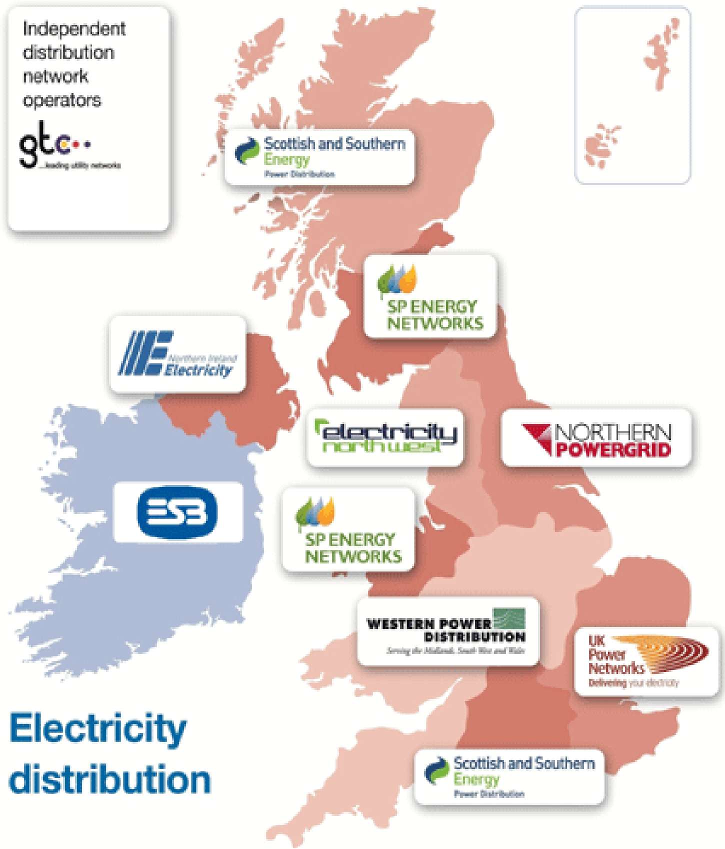


Figure 1-1: The Distribution Network Operators in the UK and Ireland (ENA 2018)

The main voltages present in the distribution network are 132kV, 66kV, 33kV, 11kV and low voltage, although small amounts of other voltages exist (mainly from legacy systems). The higher voltages, i.e. 132kV, 66kV, 33kV and 11kV, are composed of three live conductors – these three phases being labelled red, yellow, and blue. The voltage label, e.g. 132kV, is the phase to phase Root Mean Square (RMS) voltage (Alinjak et al. 2017). However, on the low voltage network, besides the three phases, there is also a neutral (return) phase. The phase to phase voltage is around 400V and the phase to neutral voltage is approximately 230V (Alinjak et al. 2017). Normally, domestic customers are connected between one of the live phases and the neutral phase.

Putting all this together, gives the representation of the network shown in Figure 1-2, although this figure is slightly misleading in that (as mentioned above) the 230V (low voltage) network's length is about half the total network length.

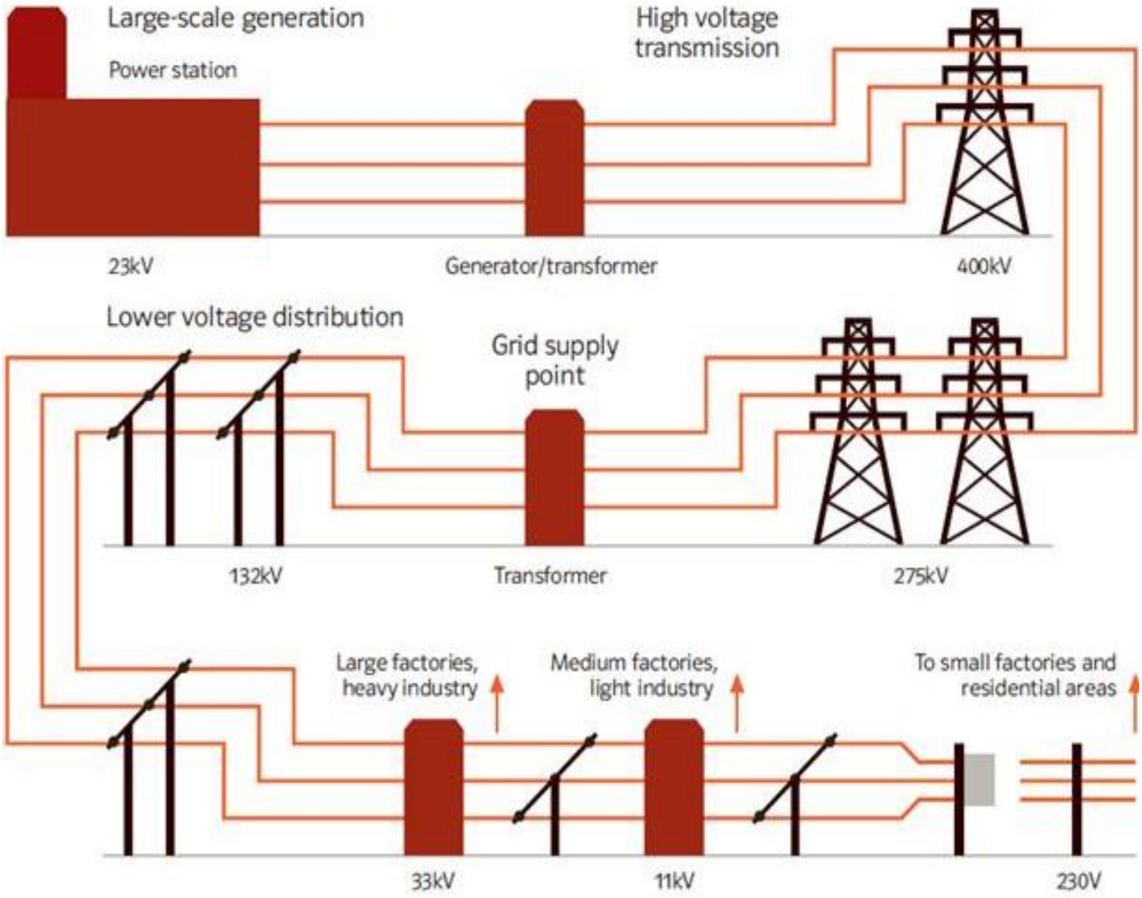


Figure 1-2: GB electricity industry components (EDW Technology 2016)

1.2 Need for Smarter Electricity Distribution Grids

Cities and highly populated urban areas have been and will be the heartbeat of the established and emerging economies. The energy sector and mainly the electricity grid is a major part of these economies and in order to support the move towards smarter cities and also reduce the carbon footprint emitted from the energy sector, the grid has to become more flexible and more intelligent (IEA 2011; Bulkeley et al. 2016). This mainly involves increasing the ability of the electricity network to accommodate more embedded generation, to be more responsive to increasing demands from customers and new demand patterns from widespread low carbon technologies in the future (Bulkeley et al. 2016; Dede et al. 2017).

The existing vertical structure of the electricity network has largely remained unchanged since the early 20th century and this makes meeting new generation (e.g. from solar Photovoltaics) and demand patterns (e.g. from electric vehicles and heat pumps) particularly difficult (Güngör et al. 2011). Traditionally, the electricity grid has not been designed to accommodate bi-directional flow of energy, and active participation by customers, or to provide visible information about the state of the network (Gharavi and Ghafurian 2011). Therefore, these are the main reasons that drive the need for transformation of the traditional electricity networks to a more proactive or smart electricity grid (Gharavi and Ghafurian 2011).

The need for more interconnected or horizontal networks was first pointed out in the 1960s (Amin and Wollenberg 2005) and it has been given further importance since the turn of the century with the gradual introduction of higher proportions of embedded generation and low carbon technologies into the distribution network (see Figure 1-3).

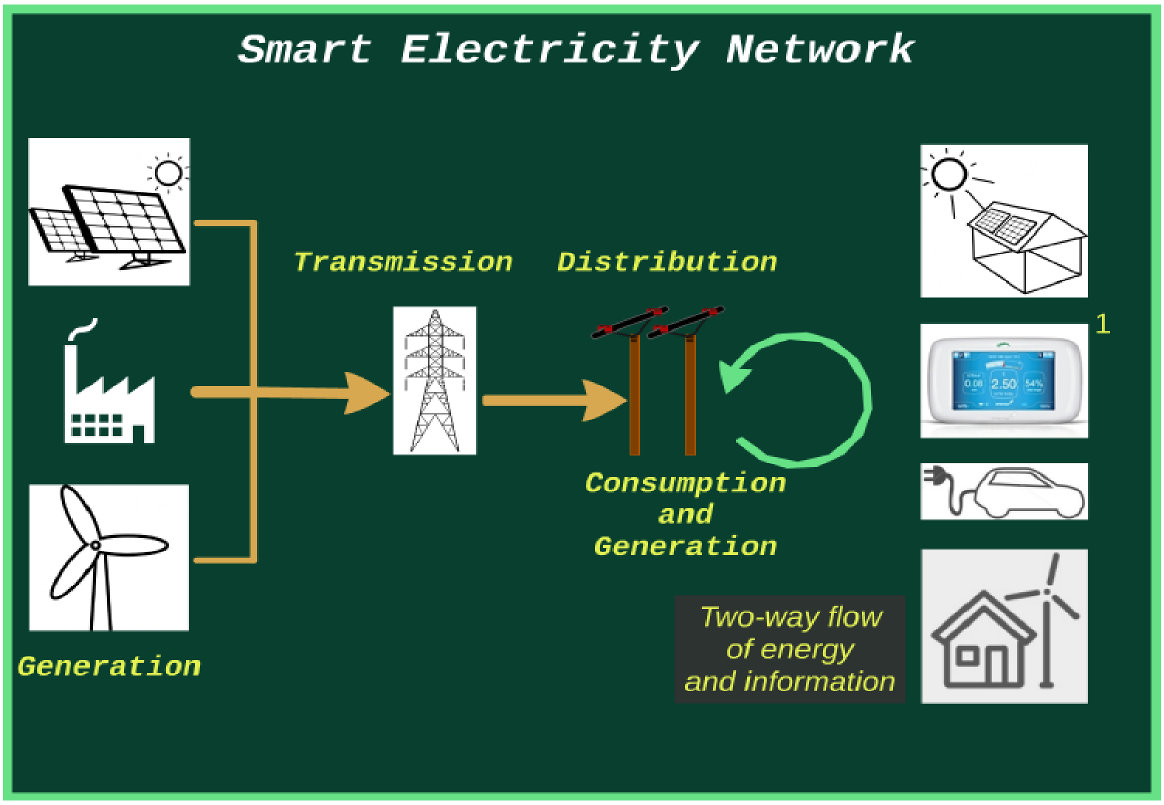


Figure 1-3: A horizontal electricity network (IEEE 2013)

1.2.1 What is a Smart Grid?

Smart grids are at the heart of smart cities. There has been a large number of smart city projects worldwide such as the Australian smart city project (Arup et al. 2014), Japan's smart city project (McGranaghan 2017), India's "Smart Cities Mission", London's Low Carbon Networks (LCN) project, and Malmo's "Smart Climate" project (Bulkeley et al. 2016). Smart grids are one of the key elements of these research projects as the DNOs will have to face the challenges of improving the flexibility and the efficiency levels of the electricity network while at the same time integrating more proactive and complex network applications (Dede et al. 2016).

Momoh (2009) describes the smart grid as a network that provides sufficient behavioural data to the operators and is capable of self-healing. On the other hand, Yu et al. (2011) define a smart grid as a reliable network that engages all the stakeholders such as consumers, generators, and prosumers by means of the state-of-the-art Information Technology (IT) to deliver efficient, cost-effective, and sustainable electrical power to the end users. By the same token, the Smart Grid Forum (2012) and Gharavi and Ghafurian (2011) also emphasise the bi-directional use of IT throughout the various levels of the electricity network in order to supply efficient, cost-effective, sustainable, and secure power (Smart Grid Forum 2014). In other words, the smart grid can be described as the effective integration of IT and power technology into the electricity grid (Momoh 2009; Gharavi and Ghafurian 2011; Yu et al. 2011; Cowan & Daim 2012).

There are many features associated with smart grids, such as reliability, security, flexibility, proactivity, cleanness, and self-healing capability (Farhangi 2010; Zhang et al. 2010; IEEE 2013). The Electric Power Research Institute (EPRI) and the International Electrochemical Commission (IEC) (Li et al. 2011) describe the following characteristics that encompass a smart grid (Yu et al. 2011; Covig et al. 2014) (Table 1-1):

Table 1-1: Smart Grid features based on the American and the European perspectives (IEC 2010; Li et al. 2011)

	American Perspective	European Perspective
Smart Grid Features	Self-Healing Optimization Information Integration Forecasting Interaction Safety Coordination	Advanced Automation Structure Service Support DG Customer Orientation Interaction Reliable Supply Flexibility

As Table 1-1 shows, the European Smart Grid perspective revolves more heavily around accommodating the integration of embedded generation at lower levels of the distribution network, whereas the American perspective represented by the EPRI places more emphasis on features such as self-healing and safety.

The topic of Smart Grid covers a wide range of areas and also varies from country to country in terms of research priorities and implementation level based on differences in electricity networks or market set ups. More importantly, definitions of the "Smart Grid" vary based on the viewpoints of different stakeholders.

For example, from the point of view of DNOs the most important feature of a smart grid is the fact that the traditional unidirectional flow to customers will change into an unfamiliar bidirectional flow of energy and data to and from customers that both consume and produce electricity, otherwise known as "prosumers" (Cowan and Daim 2012). This dramatic change itself can also affect the electricity network companies in a way that their traditional "in front of the meter" role will have to be expanded to "behind the meter" as well or in other words, they will have to take into account what happens to the energy after it has been delivered to consumers or when it is produced by "prosumers" (Cowan & Daim 2012).

Figure 1-4 highlights the complications and new relationships within Smart Grids compared to the traditional electricity grids. Smart substations, smart buildings, smart sensors on the network are also part of the eventual target of the full smart grid implementation to achieve smart cities and these targets all rely on two-way transfer of smart and new kinds of data to enable the DNOs to manage the future network efficiently, effectively, intelligently, and proactively (Energy Network Association 2009; IEEE 2013).

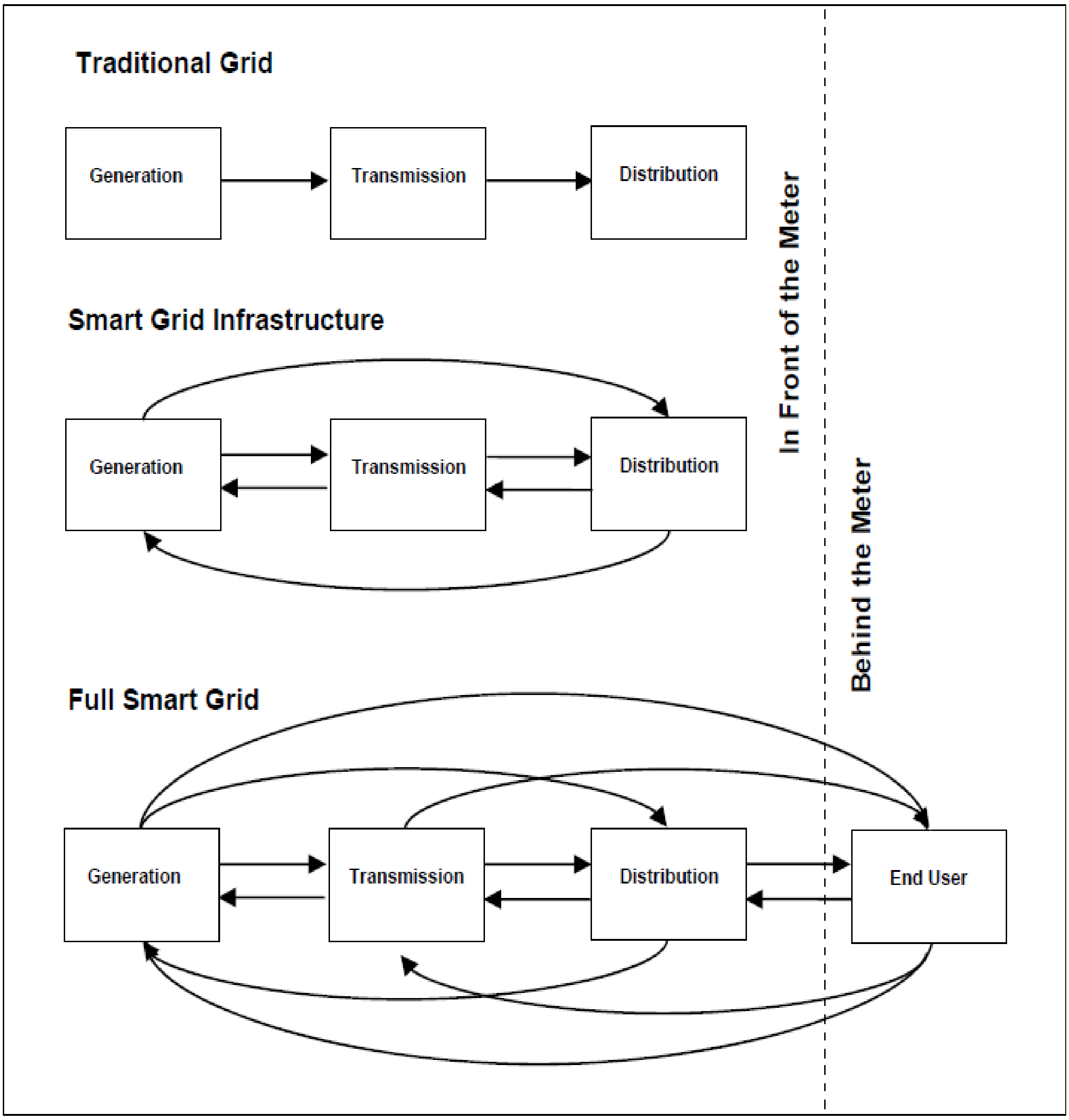


Figure 1-4: From traditional grid to a full Smart Grid (Cowan and Daim 2012)

It is important to note that smart grids do not merely consist of changes in the infrastructure of the electricity network, but that they heavily rely on smart meters, smart appliances, home area networks, and monitoring equipment (e.g. on low voltage feeders). In fact, it can be said that smart meters are the main enablers that drive the transformation of the traditional grid into the Smart Grid (Cowan & Daim 2012; IEEE 2013; Smart Grid Forum 2014), by providing a two-way flow of information between customers and network operators.

Since the plan to provide every customer in the UK with smart meters is already in motion, from the point of view of the DNOs, this could be the most cost effective way of the rectifying the problem of information gaps on the existing low voltage network. The next section explains the potential problems that this information gap can pose to the DNOs.

1.3 Challenges of Monitoring Embedded Generation and New Demand Patterns on Low Voltage Networks

According to the Office of Gas and Electricity Markets (OFGEM), embedded generation or Distributed Generation (DG) is the term used for smaller and newer electricity power generation units (e.g. solar Photovoltaics (PVs), wind farms, hydroelectric power, Combined Heat and Power (CHP) plants) that are connected to the distribution network directly and insert power into the system from various locations on the distribution network (OFGEM 2017a).

The DNOs have a responsibility to ensure that the capacity of the network is well designed to meet the various customer demands at the present and in the future (Sallam & Malik 2011) so as to be able to deliver high quality and reliable power to customers in an efficient manner. Operation of networks, which in the past did not require close monitoring, will have to cope with the introduction of new loads related to various Low Carbon Technologies (LCTs) such as electric vehicles, heat pumps, and micro generators (Smart Grid Forum 2012). This is of a huge significance in relation to balancing the voltage on the network within the statutory requirements and maintaining the voltage quality delivered to customers, or planning for new connections (Smart Grid Forum 2012).

The main reason behind the need for the more proactive management of the distribution networks is that the introduction of embedded generation and new demands from LCTs will increase issues such as voltage variations, thermal stress, reverse power flows, and network losses at the low voltage level of the network (Hollingworth & Miller 2012; Wang et al. 2012; Western Power Distribution 2013).

Figure 1-5 below demonstrates the various low carbon generation and demand sources integrated in smart electricity distribution grids of the future. It shows that the distribution grids of the future will comprise new and interconnected elements such as rooftop solar PVs, electric vehicles, and smart meters, which will introduce new demand types, intermittent generation patterns, and two-way flows of information and power.

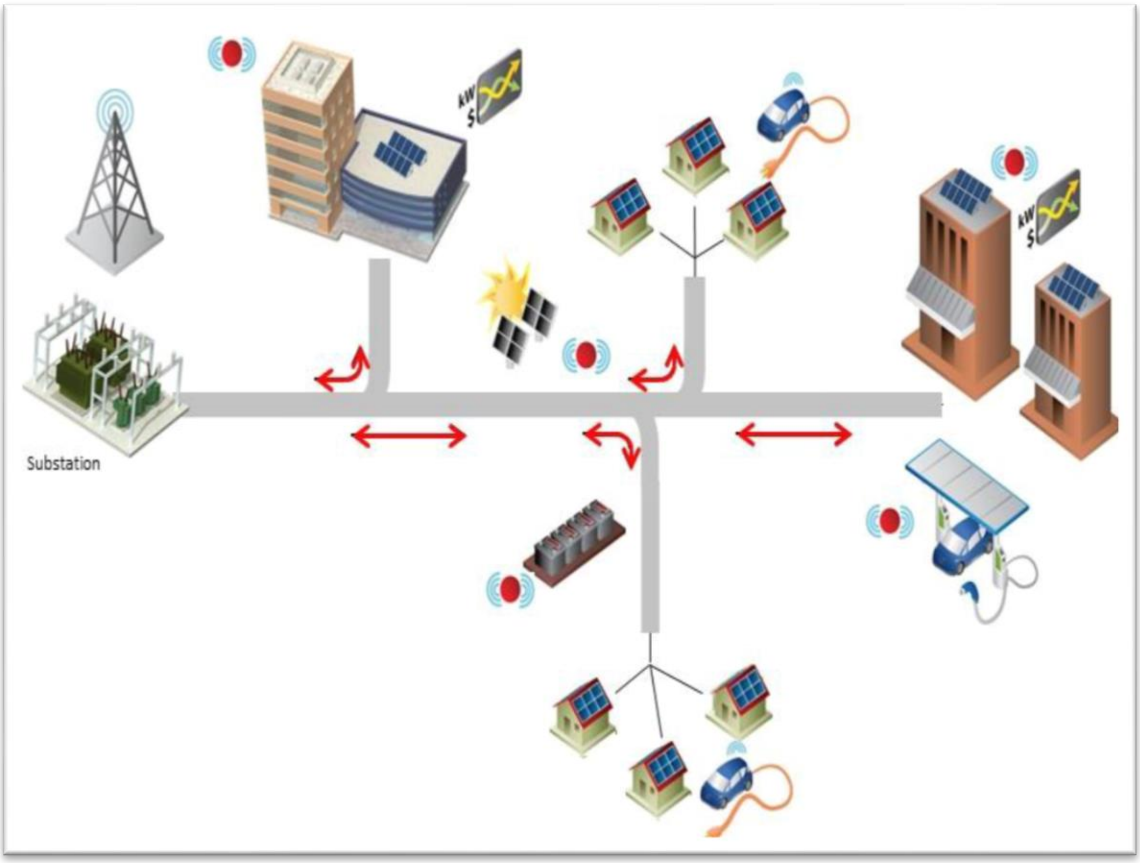


Figure 1-5: A schematic representation of the smart distribution system vision (McGranaghan 2017)

As highlighted earlier, the low voltage side of the distribution network has not been designed based on two-way flow energy or information. In fact as emphasised by Electricity North West Ltd (2014), the main problem is “always” a voltage problem resulting from the future bi-directional flow of energy at low voltage level with thermal stress on the network as a major consequence.

Another major problem flagged by Western Power Distribution (2013) is that different customer demands on different phases of a low voltage network will cause phasing imbalance that in turn will lead to higher technical network losses. Technical network losses are caused by the load being passed through cables with resistance and the heat produced as a result. On average, network losses constitute 7% of consumer bills and approximately 50% of total losses on the electricity network occur on 11kV or lower voltage levels of the distribution network. In theory, smart meter data can offer more accurate ways of determining technical network losses.

At the present time, extremely little monitoring of power flows on the 11kV and low voltage network of the electricity distribution grid is carried out (Lees 2014; Stephen et al. 2014). As Reinders et al. (2017) emphasise, the invisibility of the low voltage network to the DNOs together with the new and unknown loads and generation patterns from higher proportions of low carbon technologies and embedded generation being installed at the lower voltage level, will create a challenging combination for the DNOs.

Although load flow analysis tools such as Distribution Network Information System (DINIS) are currently used at 11kV to set the position of open points so as to, for example, minimise losses, the data fed into these programs are usually only based on rough estimations. On higher voltage levels of the electricity network, Supervisory Control and Data Acquisition (SCADA) is employed to provide information about the High Voltage (HV) and Medium Voltage (MV) parts of the network, but SCADA does not cover the low voltage parts of the network. As Silva et al. (2011) point out, even if SCADA systems were to be available at lower levels of the network (which is very costly and unrealistic due to the size of the low voltage network), the current capabilities of the SCADA systems would not be able to cope with the complexities of the new load/generation patterns on the low voltage networks (Silva et al. 2011).

At present, most low voltage substations are only equipped with Maximum Demand Indicator (MDI) meters which record the Maximum Demand (MD) on each phase (red, yellow, and blue) of the downstream cables (Figure 1-6), but these meters are not very accurate (SP Energy Networks 2015). This means that they cannot provide the DNOs with a granular and detailed knowledge of the sizes of current at various points on the network. Some of the major problems related to the accuracy with these meters according to SP Energy Networks (2015) are that the meters are read manually and only occasionally, MD is recorded at the substation level for 140-200 customers, the time range of peak loads are not recorded, customer peak diversity is neglected, and network faults and/or ambient temperature can influence the MD readings shown on the meters.

Figure 1-6: A typical secondary substation MDI meters on each phase (SP Energy Networks 2015)

These substation meters (shown in Figure 1-6) are the last monitoring point for the DNOs and there still remains a lack of knowledge and visibility downstream of the low voltage substations notwithstanding the presence of these meters.

However, in recent years the twin developments of the introduction of smart meters and the increasing levels of embedded generation, have led to there being significant interest in how the low voltage network can be more actively managed and monitored using data provided at lower voltage levels of the network, namely from customer smart meters.

1.4 Smart Meter Programme and Specifications in the UK

The main driver behind the UK's domestic smart meter programme is the improved energy efficiency that can come from:

- the introduction of time of day tariffs with the hope that some of the peak demand will be shifted to off-peak periods.
- reduction in electricity consumption if users can see the real-time cost of the electricity they are using.

The requirement for better energy efficiency stems from the UK Government's highly ambitious vision for the future of the country's power networks. The deployment of smart meters and Advanced Metering Systems (AMIs) to homes and businesses respectively, is central to this challenging move towards low-carbon power networks, smart grids, and ultimately smart cities (DECC 2010; Smart Grid Forum 2014). The major benefits that the government has identified as stemming from the introduction of smart meters are the ability to read meters remotely and to use the information provided by smart meters to more actively manage the network.

Hence the Department of Energy and Climate Change (DECC) along with OFGEM E-Serve summarised the benefits in the Smart Metering Implementation Prospectus published in DECC (2011) as:

- stimulating the transition to a low carbon economy
- tackling the climate change
- providing consumers with the information they need to reduce their consumption, save money and reduce emissions
- enabling suppliers to read the meters remotely
- enabling Smart grids, Active Network Management (ANM) for renewable energy and electric vehicles

Another essential element that is highlighted in this document is the need for establishing a new GB-wide data entity called DataCommsCo. (DCC), which enables the central communication and data management aspect of smart meters (DECC 2011). Half-hourly data in killoWatt hours (kWh) from smart meters are stored in the meters for up to 13 months and more historic data is kept and managed by the DCC.

The smart metering prospectus (DECC 2011) also touches on the issue of smart grids and how smart metering can pave the way for greater certainty in business decisions and ensure successful future innovations in the grid. A number of benefits to various stakeholders have also been identified in order to help define the functional requirements for smart metering systems.

Table 1-2 below lists a summary of those benefits that apply to the stakeholders. As shown in this table, the benefits to DNOs have not been defined by the DECC. This lack of identification of benefits for the DNOS can also been extended to the DNOs themselves. Therefore, the need to identify the ways in which the full benefits from smart meter data can be achieved has arisen. This knowledge can pave the way to achieve the Smart Grid vision successfully (IEEE 2013; Smart Grid Forum 2014).

Table 1-2: Impact assessment and benefits of smart meter to various stakeholders in the UK (DECC 2011)

Stakeholders	Benefits
Consumers	Energy savings Load Shifting Easy switching Time of use tariffs CO_2 reduction
Suppliers	Avoided meter reading Inbound enquiries Customer service overheads Debt handling Remote disconnection Avoided site visits Reduced losses Reduced theft Micro-generation
Distributors	No information

In 2014, minimum specification requirements of smart meters in UK were decided in consultation with major UK electricity market stakeholders, including the DNOs in the

UK. This consultation took place from 2012 to 2014. According to the Smart Metering Equipment Technical Specifications Version 1.58 published as a result of this consultation in November 2014 by the DECC, the minimum physical requirements for UK smart metering system are as follows (DECC 2010) (Figure 1-7):

- time display
- data storage space
- the physical meter
- Home Area Network (HAN)
- load switch
- a customer interface

Figure 1-7: A typical domestic smart meter customer interface (Ivy Link 2015)

In terms of the type of smart meter data collected, the followings highlight some of the most important aspects of the data communicated by smart meters, which can potentially be used by DNOs to obtain a more in-depth knowledge of the low voltage network and can also be used by the suppliers to provide more accurate customer bills:

- half-hourly consumption and generation
- active energy import and export
- maximum demand import and export data over every half-hour interval

- average RMS voltage
- RMS extreme overvoltage
- RMS extreme undervoltage
- loss and restoration of supply

These new streams of data can significantly upgrade the level of information available to the DNOs on the low voltage side of the network. Therefore, it is very important for the DNOs to understand to what extent these new smart meter data streams are able to provide more accurate network information data and to what extent they fall short due to potential limitations and restrictions either based on current policies in place and/or due to functional limitations of the meters.

The next section describes the various areas of the low voltage network operation that can potentially benefit from the smart meter data from the point of view of the DNOs in the UK.

1.4.1 Potential Smart Meter Impact Areas

In theory, smart meters can provide new levels of low voltage network visibility to the DNOs which have not been available before. Also, smart meters facilitate the bi-directional flow of information and energy in the lower voltages of the electricity network by providing smarter and more granular data about customer demands on the network to the DNOs. Smart meters and AMIs potentially have the capability to continuously measure the voltage, current and phase angle at each and every consumer level. Hence, smart meters have the potential to radically change the knowledge of power flows on 11kV and low voltage networks, and so to have a major impact on how these networks are managed (IEEE 2013).

However, although it is easy to imagine how having perfect meter readings at a particular point in time can allow the DNOs to have a far better grasp of the low voltage network state, in practice there is very little if any knowledge of what is the optimal way to gain benefit from the data that smart meters provide. In fact, turning smart meter data into information and practical knowledge to the DNOs has posed a major challenge to the various stakeholders of the electricity distribution grid, especially to utility companies. Hence, this is a major if not the major current research area in the field of electricity distribution network operational research. Smart meter information is one of the main pieces of the jigsaw in providing the opportunity for more active management of the electricity distribution network. One of the major challenges for the DNOs on the low voltage side is to be able to actively monitor and manage the state of the network which will be experiencing higher shares of embedded generation and low carbon technologies year by year.

A study by UK Power Networks and TE Connectivity (2015) highlights the need for the DNOs to better monitor the power flows on the low voltage networks and use the information to optimise the network capacity as well as identify the areas that are under stress and need reinforcement and replacement. In other words, using new data which provide more visibility on the low voltage network to actively manage the demand and generation on the network and make better informed decision about network planning and design as well as network asset management (UK Power Networks 2015). As Western Power Distribution (2013) emphasises, the role of smart meters and AMIs are becoming increasingly more crucial in providing a clearer picture about the behaviour of customers at the low voltage level and ultimately a more in-depth understanding of demand and generation patterns at the low voltage side of the network.

Potentially, these new sources of information can dramatically enhance conventional low voltage network applications such as network planning and design, asset management, and fault management as well paving the way for the development of modern network applications such as Active Network Management (ANM), Demand-Side Management (DSM), and automatic voltage control (Western Power distribution 2013). Some of the main areas that the DNOs are interested in developing are areas such as ANM, voltage control, fault level management, planning of new connections, dynamic rating of cables, DSM, using network storage, and using smart meter Time of Use (ToU) tariffs to manage customer demands (EA Technology 2016). The overall aim of the DNOs in developing these applications is to reduce network stress and to operate the network more cost effectively.

In theory, high quality smart meter data can help DNOs overcome these challenges as well as helping them with investment and planning decisions (Smart Grid Forum 2014). However, the current policies and specification in the UK may hinder the ability of smart meters to provide optimum levels of accurate network to the DNOs.

1.4.2 Potential Limitations of Smart Meter Data

Over the past few years, there has been a major debate over issues of smart metering data access and handling, mainly due to privacy concerns. The UK Government has been working on the existing legislations, principally the Data Protection Act 1998, to examine whether the current legal frameworks could cover smart metering data or not (DECC 2011). Smart metering data is useful to all of the stakeholders, but it should be noted that the Government's approach aims to protect the privacy of consumers while providing the industry with the appropriate data that they rely on to meet the statutory requirements (DECC 2011). This approach strives to implement the "privacy by design" principle adopted from international best practices and recommendations from the data protection regulator and the Information Commissioner's Office (ICO) (DECC 2011). Essentially, if the data can identify the individuals then the data is considered to be personal data and of a sensitive nature, hence the establishment of the DCC as an independent third-party entity in the autumn of 2013. The DCC is responsible for managing the procurement and contract management of smart metering data and communication services (DECC 2011; OFGEM 2017b).

The programme encourages suppliers and network operators to carry out studies and justify that a higher resolution of data is required for them to meet their operational statutory requirements and this will not invade the privacy framework being designed by the programme (DECC 2011). However, there is a lack of research in this area and it requires more detailed investigations, especially in the case of network operators, where the need for near real-time data will grow as the grid becomes smarter. In the near future aggregated and anonymous data may not suffice (DECC 2011).

The evolving nature of the smart grid applications pose a particular challenge to the DNOs in determining the optimal frequency, time resolution and aggregation of the smart meter data required. For example, access to voltage quality data and customer loads are essential to help network operators manage the local network more efficiently and the EU Smart Grid Task Force considers the data to be of a technical nature, but it must be proved that privacy measures are not affected by granting access to this kind of data at half-hourly resolution or higher (DECC 2011).

The network operators obtain their required data from the DCC as opposed to directly from suppliers and research is needed for the network operators to demonstrate their business dependence on detailed and half-hourly levels of data (DECC 2011). Therefore, careful studies must be conducted about the appropriate frequency and time resolution of data required for the various operational applications within their business.

According to the current regulations, the DNOs in the UK will receive half-hourly averages of customer demands, but the smart meter data is required to be aggregated and anonymised before it is processed within the DNO applications due to privacy issues (OFGEM 2017b). Therefore, the DNOs and the field of electricity network operation research are very interested in finding ways of filling in the gaps where smart meters are not yet available as well as identifying how various aspects of smart electricity distribution grid operation and monitoring benefit or suffer from having aggregated and lower time resolution of smart meter data.

Currently in the UK, the availability of real-time and high resolution smart meter data from all customers to the DNOs is and will be restricted due to:

- the incremental process of smart meter deployment until 2025.
- smart meter communication delays and faults.
- presence of non-smart meters even after 2025.
- averaging of customer demands at half-hourly intervals instead of 1 minute real-time readings.
- anonymization of customer demands from smart meters by aggregating the data from customers.

These limitations can reduce the benefits of smart meter data to the DNOs and consequently affect the quality of smart meter data for distribution network applications such as network planning and design, asset management, fault management, and network capacity management.

This is why Northern Powergrid is supporting this research project in order to determine the impact that various frequency, granularity, and aggregation levels of smart meter data will have on the accuracy of information on the low voltage network. The frequency of smart meter data is the time period of the data that is made available to the DNOs (e.g. daily, weekly, monthly). The time granularity or resolution of data denotes the resolution of smart meter time intervals (e.g. real-time, 1 minute, 5 minute, half-hourly, etc.). The aggregation level of data indicates the number of customers for which the readings have been grouped together as part of anonymization process (e.g. 2 customers, 4 customers, etc.)

The extent to which these issues can affect the accuracy of smart meter data in the context of the DNO application is at the core of this research.

1.5 Research Aims and Objectives

There is uncertainty over the ability of smart meter data relayed to the DNOs to provide accurate low voltage network information. The central focus of this research is to examine the extent that factors such as smart meter time resolutions and frequency, customer data aggregation levels, and the unavailability of smart meter data can affect the estimation accuracy of low voltage network currents, losses, and voltage levels.

It also investigates the effects that these inaccuracies can have on major DNO applications such as network planning and design, asset management, network monitoring, and fault management. Additionally, some recommendations are also made about the ways in which some of the data gaps and inaccuracies can be overcome.

The main aim of this research is to investigate and identify:

- the impact of having partial, half-hourly averaged and/or aggregated smart meter data on the accuracy of the information used by the low voltage network operational applications of the DNOs in the UK.

The main objective of this thesis is to address the following research questions:

- What are the impacts of gaps and inaccuracies in smart meter data, as a result of the averaging and aggregation of customer demands recorded by smart meters, on the accuracy of critical low voltage network information used by the DNOs in the UK?
- To what extent do these inaccuracies affect major DNO applications such as network planning and design, asset management, fault management, and network monitoring in typical low voltage network models?
- What is the optimum frequency, time resolution, and aggregation level which can enhance the management of the low voltage networks?

These research questions are investigated by:

- finding missing values: applying practical low voltage estimation methods to new smart meter data sets to fill in missing half-hourly customer loads using historical smart meter data.
- studying smart meter time resolutions: statistically analysing the effects of various time resolutions of smart meter data from data set to 120 minutes on low voltage network loss, voltage level, and cable load estimates.
- studying smart meter aggregation levels: statistically and graphically analysing the effects of various levels of customer data aggregation from 1 to 10 customers on low voltage network loss, voltage level, and cable load estimates.
- contextualising the impacts: placing the findings in the context of low voltage network applications such as asset management, network planning and design, fault management, ANM, and DSM to propose the optimum levels smart meter data frequency, time resolution, and aggregation level to the DNOs and policy makers.

This study is carried out using actual customer demand data from two smart meter trials in the UK fitted into various test low voltage network models. This work differs from previous studies such as Urquhart & Thomson (2015) and Urquhart et al. (2017) in that the low voltage network models used in this work are representative of typical three phase low voltage networks in practice and comprise of 100 houses. This is a unique approach and significantly improves previous studies in this field such as in Brandauer et al. (2013) and Urquhart and Thompson (2015), which are carried out on single phase low voltage network models with a limited number of houses. Another novelty of this work is that our analysis also includes the impact of smart meter data aggregation, from 1 to 10 customers, on major low voltage network performance indicators which has not been addressed before. To date, the only work in this area is the study presented in EA Technology (2015b) which only estimates the costs that DNOs will incur as a result of having aggregated customer data at 2 or 4 customer aggregation level compared to having disaggregated individual smart meter data.

The models used in this study are also very close to low voltage networks found in the UK both in terms of the network geometry and network component characteristics. However, our work is applicable to most a majority of low voltage networks and our studies can be used by other DNOs or experts in the field. The major novelties of this research is that both the extent to which the accuracy of smart meter data is affected by time resolution and customer data aggregation is investigated and subsequently contextualised with respect to the DNO applications. This research also makes use of descriptive statistical methods, which are the most appropriate and common in the field.

Detailed description of the data sets, methods, and network models employed in this research will be discussed in chapter 3.

1.6 Research Outputs

Findings from this research have been presented and disseminated in academic journals and conferences, industrial reports and workshops, governmental seminars, and public engagement exercises.

The following research outputs have been generated from this thesis:

Journal Papers:

- Chapter 4 Output: The International Journal of Electrical Power and Energy Systems (IJEPES), "Low Voltage Current Estimation Using Smart Meter Data". In press. Reference no. JEPE4507. Impact Factor: 3.289
- Chapter 5 Output: The IET Journal on Generation, Transmission and Distribution, "Using Smart Meters to Estimate Low Voltage Losses". In press. DOI: 10.1049/iet-gtd.2017.1300, Impact factor: 2.213

Conference Publications:

- CIRED 2017, Glasgow, "Analysing the Ability of Smart Meter Data to Provide Accurate Information to the UK DNOs". Published in IET Open Access Journal. DOI: 10.1049/oap-cired.2017.0654
- IEEE PowerTech 2015, Eindhoven, "Geospatial Visualisation of Smart Data for Improved Network Management". Published in IEEE Xplore. DOI: 10.1109/PTC.2015.7232445
- CIRED 2015, Lyon, "Smarter Business Processes Resulting from Smart Data". Published CIRED Conference Proceedings.

Workshops, seminars, and reports:

- Presentation of the findings to the industrial experts in the field by producing reports and organising workshops regularly from 2014 to 2017.
- Presentation of the findings to major policy makers in the UK, the Department for Business, Energy, and Industrial Strategy (BEIS), in December 2016.
- Presentation of our findings about time resolution and loss estimates in chapter 5 in a report published by the Northern Powergrid (Northern Powergrid 2016).

Public Engagement:

- Sheffield Festival of Science and Engineering, Birley Community College, Sheffield, 2014.

1.7 Thesis outline

This work comprises 7 chapters. Chapters 1 to 3 lay the foundation and the theoretical underpinnings of the research as well the methods and the data sets used and chapters 4 to 6 contain the main analysis carried out in this research, followed by chapter 7 which summarises the main findings as well as highlighting conclusions and contributions of this study. The content of each chapter is outlined below.

Chapter 2 describes the literature review process followed by identifying the research trends and priorities in the field of smart grids and smart meters that are found in the literature. In the next step, the research and studies that are relevant to the aims and objectives of this work are reviewed. In the first place, this is carried out by identifying the main approaches in obtaining important low voltage network performance indicators with and without smart meter data. Secondly, the ways in which major low voltage network applications are undertaken in the absence of smart meter data and the ways in which they can be enhanced by having detailed smart meter data are investigated.

Chapter 3 describes the data sets employed in this study and highlights their characteristics. This is followed by presenting the main low voltage network models and the methods used in carrying out the analysis in chapters 4 to 6 and the ways in which they are different from relevant studies.

Chapter 4 presents the results of predicting missing low voltage currents from a portion of customers on the network using historical smart meter data from the neighbouring meters. In this chapter, 5 estimation methods in combination with half-hourly smart meter data from two different data sets are tested on a 20-house low voltage network model and the 2 best preforming methods are then tested on a 50-house network model. The best two methods and the 50-house network model is also ultimately used in order to estimate the missing currents from individual customers.

Chapter 5 demonstrates the correlation between smart meter time resolution and the estimation accuracy of critical low voltage network information such as technical losses, voltage levels, and cable loading percentages. This chapter examines how these estimates are affected by the time resolution of smart meters varying from 1 to 120 minute averages in a balanced and an unbalanced three phase low voltage network using data from 8 sample dates of the two different data sets. Chapter 5 also examines two methods of predicting 1 minute loss estimates based on half-hourly values of losses and the accuracy of the methods is examined.

Chapter 6 investigates the correlation between smart meter aggregation levels of 1, 2, 4, 6, 8, and 10 customers and the accuracy level of critical low voltage network information such as technical losses, and voltage levels. The effects of aggregation scenarios on voltage and loss estimates are examined in two different low voltage network topologies and on 8 different sample dates in a balanced and an unbalanced network model.

Chapter 7 draws conclusions from the findings of this study and the ways in which the findings contribute to the academic and practical knowledge. It also sheds more light on the limitations of the analysis carried out and recommends the future research directions.

Chapter 2 Literature Review

This chapter investigates the literature relating to the various aspects of DNO applications and the role that the smart meter data can play in improving these applications. In the first place, the literature review steps are described. Secondly, the major research trends, priorities, and gaps in the field are identified, followed by an in-depth investigation of low voltage network performance indicators that are important to the DNOs, such as estimation of customer loads, network losses, voltage levels, and cable capacity percentages. These are the types of information that major DNO applications are heavily reliant on. Last but not least, major DNO applications that are relevant to the low voltage side of the network and the ways in which they have been operated prior and after the availability of smart meter data from customers are investigated. These DNO applications consist of asset management, fault location and restoration, network design and planning, network planning, network monitoring, power quality management and integration of embedded generation, and Active Network Management (ANM).

2.1 Literature Review Process

The topic of Smart Grid covers a wide range of areas and also varies from country to country in terms of research priorities and implementation level based on differences in electricity networks or market set ups. Our study focuses on how the DNOs in the UK can benefit from smart meter data based on the current policies in the UK. Considering these issues, it was important to investigate the relevant studies carried out and select the methods and approaches that work in the context of the UK's distribution network. To this end, the literature review process was divided into two stages. An initial narrative review of the wide variety of the works being produced in the field was carried out to identify the general research trends and gaps in the field. The secondary stage of the literature review process comprised a systematic review of the academic literature in order to focus on the latest research and the most relevant studies to the aim and objectives of this thesis.

2.1.1 Initial Literature Review Stage

The first step of the literature review entailed carrying out a narrative literature review of the topics of smart grids and smart meters by investigating the most recent and relevant conference publications from the IEEE Smart Grid and the CIRED conferences (e.g. IEEE Smart Grid conference in Berlin (2013) and CIRED conferences in Frankfurt (2011) and Stockholm (2013)). This was carried out in order to investigate the latest research being produced since the process of publishing in journals usually takes longer than presentation of the work that is underway at the relevant conferences. This is particularly important in our field as research around the topic of smart grids and smart meters is gathering pace every day and the areas that are of interest vary.

Additionally, the journal papers published in the databases relevant to the field of smart grids such as the IEEE Xplore and the IET were also investigated. These were combined with reviewing the most recent policy documents about the Smart Grid vision and the smart meter implementation documents published by the OFGEM in consultation with the UK DNOs and suppliers.

This step was also complemented with visiting various DNO application teams at the Northern Powergrid operation sites, including asset management, network planning and design, fault management, and smart meter implementation. This was carried out in order to obtain the operational needs and expectations from the future smart meter data that are specific to each application context. The findings from reports published by other DNOs in the UK (e.g. Western Power Distribution and Scottish and Southern Electric) were also reviewed to find the common areas that are of interest to the DNOs in the UK.

The next step was to review the major Smart Grid and smart meter projects in Europe and in the UK to further investigate the research priorities and the data sets available to us as well as reducing the chances of conflict in research interests. The most important Smart Grid project in the UK, especially with respect to our research project, is the Customer-Led Network Revolution (CLNR), mainly due to the involvement of Northern Powergrid in the project and the smart meter data sets that were obtained from this project. CLNR was a £31 million project in the UK which aimed at decreasing the level of carbon emissions produced by customers by promoting the use of low carbon technologies and studying Smart Grid solutions and the behaviour of households equipped with smart meters (CLNR 2012). As part of the CLNR project, smart meter roll-out trials were carried out in the UK between 2011 and 2014 with the aim of facilitating the eventual full smart meter deployment in the country.

The project was carried out by British Gas, Northern Powergrid, Durham Energy Institute, the Newcastle University, EA Technology, and Low Carbon Network Fund (LCNF) and through this project 14,000 homes in the North East and Yorkshire were provided with smart meters, out of which 2500 were using solar PVs or heat pumps (Dudeney et al. 2015). Data from these trials were gathered and analysed by the parties involved to examine the new kinds of customer demand which will be new to the electricity distribution network as well as testing the Low Carbon Technology (LCT) solutions and various tariffs (CLNR 2012; Dudeney et al. 2015). CLNR is a major source of data for this research project as well. Smart meter data sets from the trials provide invaluable information to our research purposes.

In summary, the bodies of work that were investigated in this stage of the literature review can be divided into the following three major categories:

- policy, consultation, and advisory documents from major international and European organisations in the field and the UK regulatory bodies.
- academic research.
- learning outcomes and reports from major industrial and academic collaborative trials and studies.

The main reason for investigating these various sources was the fragmented and diverse nature of the research in the field and also the fact that the research priorities and directions are to a large extent driven by the perspectives of the various players in the field.

2.1.2 Secondary Literature Review Process

After the initial review of the literature, a number of meetings were held with the academic and the industrial supervisors and also with members of the CLNR team. During these meeting a number of issues such the research trends and priorities, gaps in knowledge, and the available data sets were discussed in order to narrow down the scope of the research and to identify a set of aims and objectives that offer novel contributions to the realms of academia and the industry. After consolidating the main aims and objectives of this research, a more detailed and systematic review of the literature was carried out by identifying the relevant databases in the first place.

The next step of the literature review process for this thesis was conducted by defining search terms or “search strings” relevant to the research aims and objectives and the research question defined in 1.5. The relevant papers were identified using Boolean searches (Linnenluecke 2017) of the following search strings in Scopus and Google Scholar:

- Smart grid*
- Smart meter* or AMI
- Low Voltage or LV
- Load prediction/Allocation/Forecasting
- Losses
- Voltage or voltage level*
- Asset management
- Network planning
- Network design
- Fault management
- Active network management
- Distributed generation/embedded generation
- Network monitoring

These search strings were used hierarchically starting with searching for papers, conference publications, and book chapters containing the term “smart grid*”. This search returns 28,380 papers. These papers are then narrowed down to only 720 papers by searching the search strings of “smart meter*” and “LV” or “Low Voltage” in the whole text of works. Figure 2-1 below demonstrates the systematic literature review steps taken and the number of papers found at each stage (N).

As Figure 2-1 shows the 720 search returns found in step 2 were then further narrowed down in separate searches using search strings of “load prediction”/” load forecasting, “voltage*”, “losses”, “asset management”, “network planning”, “network design”, “fault management”, “active network management”, “distributed generation”/”embedded generation”.

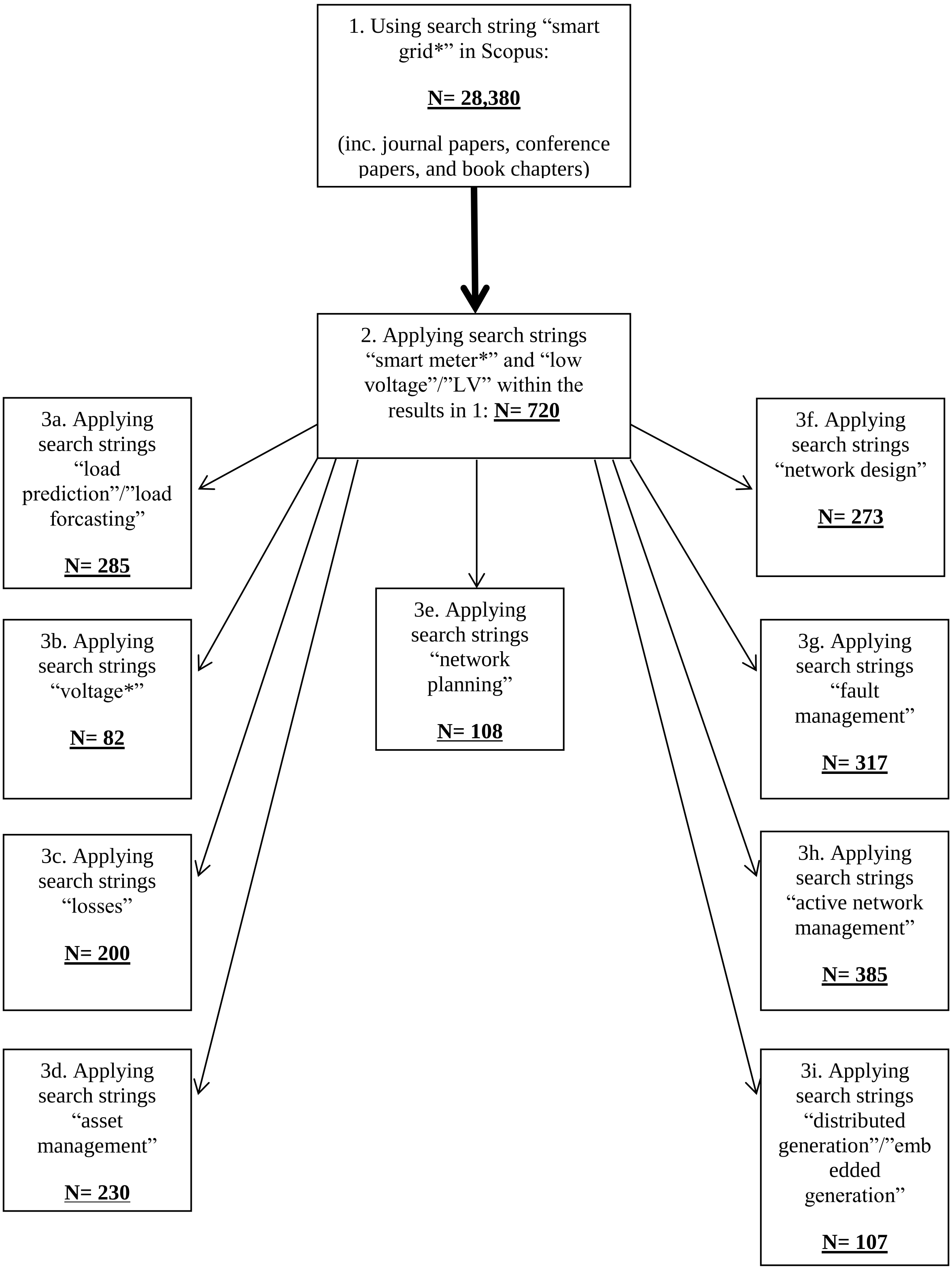


Figure 2-1: The systematic literature review steps

The title and the abstract section of these papers were then read and the studies that were related to this research were then selected to be investigated in full. It was decided to include relevant conference papers from major conferences in the field that ensured a good quality of peer review (e.g. CIRED and IEEE Smart Grid). As Tranfield et al. (2003) highlight, a complete literature review should also include published and unpublished works that have been through credible peer review processes. The selected papers were deemed eligible based on their relevance to the DNO application of interest or whether they contained methods related to studying the relationship between smart meter data frequency, time resolution, or aggregation level and estimation of customer loads, network losses, or voltage levels.

Additionally, various known governmental and organisational websites as well as the websites for major trials and studies were also investigated and tracked to gather the documents and reports related to this research. These websites and databases include CIRED conference proceedings, IEEE Xplore, IET journals, OFGEM, DECC, European Joint Research Centre (JRC), and the Department for Business, Energy and Industrial Strategy (BEIS). This was very helpful to find the relevant studies related to the DNO applications of interest.

Some of the literature on the smart meter projects and regulations around the smart meter time resolutions, gradual implementation, and aggregation of smart meters have been selected to explain the limitations that these issues can cause in the UK. However, the remainder of the literature has not been restricted to the studies in the UK only. This study can be used by the network operators and the experts in the field in other countries that face similar issues such as many European countries. However, it should be noted that the difference between the electricity market set up in the UK and other European countries means that the DNOs in the UK are more limited in their access to smart meter data and customer data information. For example, in the UK the DNOs own and run the distribution networks and the customer billing is carried out by suppliers. However, in most European countries the utility companies carry out both all of these tasks, which provides them with access to more types of data. Therefore, some of the works carried out in other countries were inapplicable to the network set up in the UK.

2.1.3 Selected Academic Papers

In terms of load prediction or load allocation searches (3a), the studies deemed relevant to the scope of this research are McQueen et al. (2004); Kersting and Philips (2008); Arritt et al. (2012); Ferreira et al. (2012); Li et al. (2012); Rossi and Brunelli (2013); Gajowniczeka and Ząbkowskia (2014); Huang et al. (2014); Mirowski et al. (2014); Stephen et al. (2014); Velez et al. (2014); Hayes et al. (2015); Mcloughlin et al. (2015) ; Quilumba et al. (2015); Wong and Chung 2015; Klonari et al. (2016); Valgaev et al. (2016); Vasudevaro et al. (2016).

Considering the studies on voltage levels and losses (3b and 3c), the papers that deemed eligible to be included in our study are Hackman et al. (2013); Rossi and Brunelli (2013); Gajowniczeka and Ząbkowskia (2014); Mahmud et al. (2014); Alzate et al. (2015); Bokhari et al. (2015); Urquhart and Thompson (2015); Garcia et al. (2016); Pompodakis et al. (2016); Wang et al. (2016); Nijhuis et al. (2017); Urquhart et al. (2017); Vasudevaro et al. (2016); Celik et al. (2017); Konstantelos et al. (2017).

The major works related to the DNO applications that were selected are summarised below:

Asset management (3d): Brown and Humphrey (2005); Tor and Shahidehpour (2006); Brint et al. (2008); Black et al. (2009); Brint and Black (2014); Lees (2014); Miller (2015); Sirto et al. (2015); Goyal (2016); Mohsenzadeh et al. (2016).

Network planning (3e) and network design (3f): Brown (2008); Porter and Strbac (2007); Strbac et al. (2010); ENA (2012); Roupioz et al. (2013); You et al. (2014); Nijhuis et al. (2017).

Fault management (3g): Apel et al. (2001); Verho et al. (2004); Gano et al. (2011); Makinen et al. (2013), Zhao et al. (2013), Abusdal et al. (2015); Estebsari et al. (2016); Jiang et al. (2016); Jamali and Bahmanyar (2016); Jamali et al. (2017).

Embedded generation (3h) and active network management (3i): Lees (2014); Mcdonald et al. (2010); Strbac (2010); Repo et al. (2013); Chiandone et al. (2014); Degefa et al. (2014); Zhou et al. (2014); Hattam (2015); Jagtap and Khatod (2015); Neaimeh et al. (2015); Sandulec et al. (2015); Navaro-Espinoza and Ochoa (2016); Paterakis et al. (2016); Phoghosyan et al. 2016; Watson et al. (2016); Barbato et al. (2017).

In summary, the literature review process was dictated by the fragmented, diverse, and transdisciplinary nature of the literature being produced in the field of Smart Grid and smart meter research. The results that were deemed to be relevant to this project are presented in chapters 1, 2, 3, and 7 of this thesis.

2.2 Research Trends in the Field of Smart Grids

The Smart Grid agenda is pursued differently in different countries and regions based on their priorities and network characteristics. The IEA (2012) emphasises that the process of transforming the grid is an ongoing and a gradual process as needs and applications of the DNOs emerge and evolve over time. In a broader context documents and studies such IEEE's "2050 Grid Vision" (2013), European Union's JRC reports (Giordano et al. 2011; Covig et al. 2014), and also European Technology Platform on Smart Grid's report (ETP SmartGrids 2013) draw the main research and development areas of the Smart Grid. In the context of the UK's Smart Grid agenda, the Energy Network Association's report (2009) and the OFGEM's Smart Grid Forum's document (2014) highlight the Smart Grid vision and roadmap for the DNOs in the UK (Energy Network Association 2009; OFGEM et al. 2014). These will be outlined in the following section.

2.2.1 Smart Grid Research Areas and Priorities

According to the latest report from Joint Research Centre (JRC) of the European Union, there has been a growing investment in Smart Grid projects across all European countries (see Figure 2-2 below) reaching £56 billion allocated to 459 projects (Covig et al. 2014). The majority of these projects involve collaborations among various parties from governments, academia, and industry as well as investments from different countries (Covig et al. 2014).

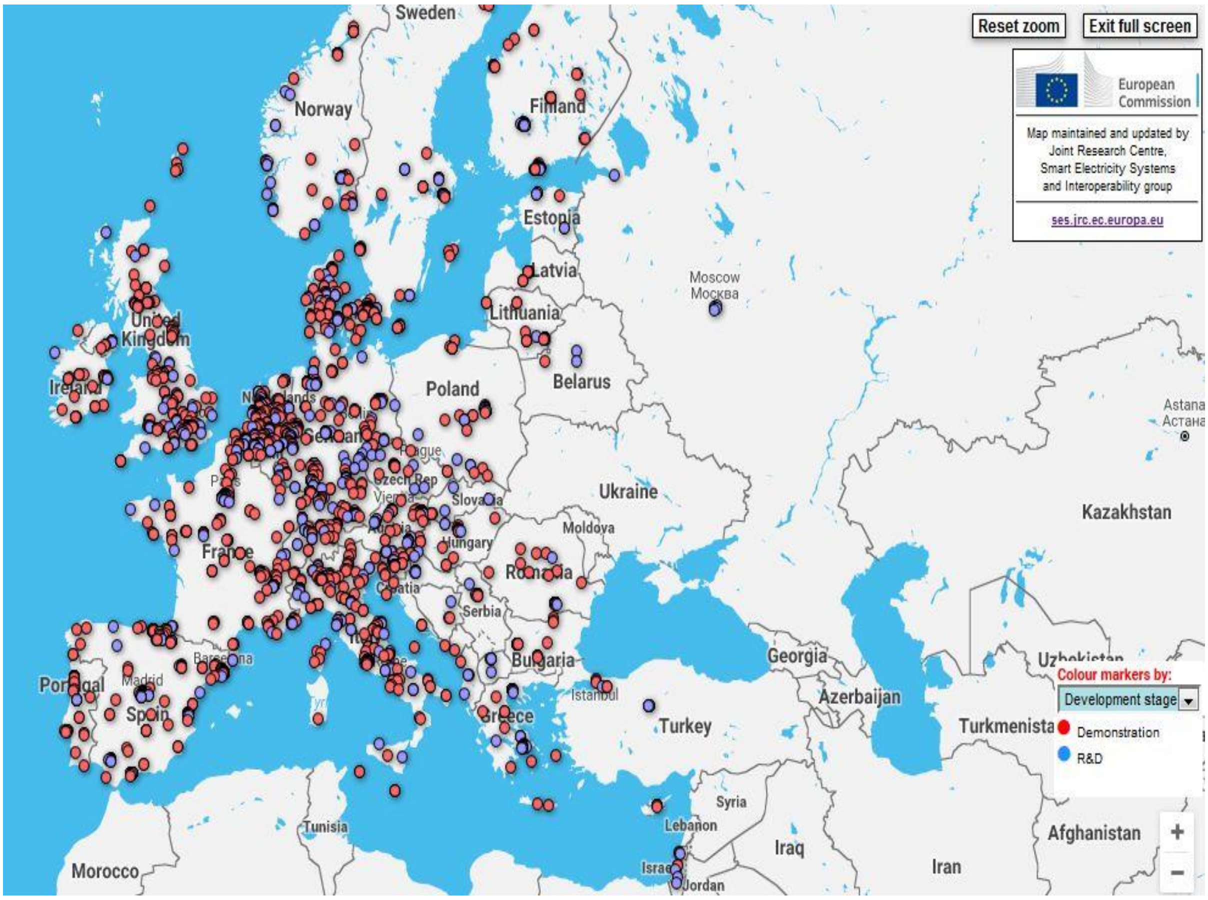


Figure 2-2: Spread of Smart Grid projects in Europe (JRC 2016)

A great majority of investments in the field of smart grids originate from France, United Kingdom, Spain, Italy, and Germany (Covig et al. 2014). Some of the main challenges which these projects encounter are the social, policy and regulatory constraints that vary from country to country and as Covig et al. (2014) also emphasise, these challenges restrict the potential replicability of the results from these projects in different countries. Based on IEEE (2013) and ETP Smart Grids (2015), the top research priorities in the Smart Grid can be summarised as follows:

- **Observability and control:** raising the visibility level of the network operation, especially at the lower levels of the electricity networks to enable a more proactive operation of the grid.
- **Modelling power systems and Information and Communication Technology (ICT):** modelling the various areas of the electricity networks and testing the possibility of utilising advanced information technologies (e.g. smart meters, substation meters, etc.).

- **Consumer maturity:** raising the awareness level to change passive consumers into informed decision makers.
- **Power technology to increase network flexibility:** researching the most effective ways in which embedded generation can be integrated into modern distribution grids
- **Integration of demand-side management:** researching the most practical incentives to reduce the strain on the network at peak consumption times.

A summary of major Smart Grid and smart metering projects are presented in Tables 2-1 and 2-2 below.

As Table 2-1 highlights, most European Smart Grid projects are large-scale and multi-national. Majority of these projects are focused on integration of embedded generation or low voltage technologies in the medium voltage network, while reducing risks and costs.

Table 2-2 shows that smart metering projects are more focused on the lower voltage side of the network with more emphasis on the relationship between customers and smart meter data and the ways in which such data can be used to incentivise higher uptake of solar PVs and shift in consumption at peak times.

Since this thesis focuses on the challenges that smart meter specifications and regulations in the UK can pose to the benefits gained from smart meters (see section 1.4) and also the fact that the electricity market set up in the UK is different from other European countries, Table 2-2 only highlights the smart meter projects in the UK.

Table 2-1: Summary of major Smart Grid projects in Europe

Project Title	Collaborators	Scale	Focus
Meter-On (Marcoci et al. 2013)	Network Operators Academics Technological Institutions	European Level	Minimising and removing technical and market barriers towards widespread implementation of customer smart meters.
Enel Distribuzione (Stein et al. 2013)	Network Operators Academics Suppliers	European Level	Improving the integration of renewables in the medium voltage network.
Grid+ (Losa et al. 2013)	Network Operators Academics Technological Institutions	European Level	Fostering the evolution of the electricity grid toward the European Smart Grid 2020 objectives, especially higher shares of renewables.
InovGrid (Matos et al. 2013)	Network Operators Academics	Country Level: Portugal	Integrating of information systems and the distribution network in order to reduce operating costs and increase efficiency.
Endsea Smart Metering Roll-out (Open meter 2011; Tellechea 2013; Vadacchino et al. 2013)	Network Operators Academics Suppliers	Country Level: Spain	Using innovative technologies in transferring data from smart meters to the utility companies and also in the management of the electricity network.
Smartlife (Soepboer et al. 2013)	Academics Suppliers Technological Institutions	European Level	Developing new ways of asset management and distribution asset management in MV networks.

Table 2-2: Summary of Smart Grid and smart meter projects in the UK

Project Title	Collaborators	Scale	Focus
Lincolnshire Low Carbon Hub (Douglas 2011 and Bale 2015).	Suppliers Academics	Community Level	Active Network Management (ANM) solutions in to facilitate integration of DG into the LV networks.
Hook Norton Low Carbon (Douglas 2011)	Suppliers Academics	Community Level	Monitoring of consumer demands at LV substation and neighbourhood levels within a "smart" community.
Ashton Hayes Project (Kadar 2011)	DNOs Suppliers Academics	Community Level	Transforming the village to a carbon neutral village.
Northern Isles New Energy Solutions (NINES) (Reid 2011)	DNOs Academics	Community Level	Researching proactive network management solutions in order to provide network constraint relief and reduce network peak demands.
Low Voltage Network Solutions (Electricity North West 2014)	DNOs Academics Suppliers	Community Level	Investigate load patterns on networks to accommodate higher shares of DG in the LV network.
Low Voltage Templates	DNOs Academics	County Level	Finding ways of integrating higher shares of DG into the LV network by developing a tool that categorises sections of the LV network into typical LV templates

In the context of the UK electricity networks, the Smart Grid research areas and priorities as highlighted by the ENA (2009) and the OFGEM (2014) can be summarised as follows:

- Rising energy security and integration of embedded generation

- Enhancing energy security and reliability
- Improving network capacity
- Improving network visibility
- Enabling low carbon technology connections (e.g. electric vehicles and combined heat pumps)

Observability and control, especially at the distribution level can potentially be achieved by utilising the smart data received by the DNOs via AMIs and smart meters (ETP SmartGrids 2015; IEEE 2013). This is the key enabler of effective integration of embedded generation and low carbon technologies into the modern grids. However, obtaining a higher level of observability at the low voltage level by installing fine-grained measuring devices at low voltage substations is a very expensive option, mainly due to the tens of thousands of low voltage substations that need to be equipped with these meters and the enormous costs involved (EurElectric 2013).

On the other hand, about 80% of customers in Europe will be provided with smart meters by 2020 (some countries such as the UK until 2025) (EurElectric 2013), therefore it is far cheaper to use these smart meters in a bottom-up approach and draw benefits from the smart data that will be available to the DNOs as a result of widespread smart meter deployment programmes in Europe. It is expected that smart meter data can provide DNOs with a higher level of observability and control at the lower levels of the grid by recording and transmitting the sizes of generation and demand from every customer to the DNOs, and as a result eliminate the need for the DNOs to install high resolution meters at every low voltage substation.

2.3 Smart Meters and Smart Low Voltage Grid Operation

Potentially, having real-time smart meter data, which provide actual customer loads, can improve the accuracy of fundamental network information such as size of the currents on various section of the network and as a result lead to more accurate estimates of losses, voltage levels, and low voltage cable load capacities. This information can in turn increase the visibility of the low voltage network to the DNOs and enhance the distribution network applications such as network planning and design, asset management, network monitoring, fault management, and power quality management. However, in reality it is highly unlikely that the DNOs can obtain high resolution and fine-grained smart meter readings from customers on all of the low voltage networks, due to limitations posed by smart meter specifications, time delays, privacy concerns, and the lack of infrastructure. Traditionally, the data that the smart meters can provide has not been available to the DNOs beyond the low voltage substation level. However, this will potentially change with the widespread deployment of smart meters and the provision of real-time customer demands.

In the traditional electricity network set up, the amount of real-time information that DNOs have about the status of the electricity distribution network decreases as the number of monitoring points drop from high to low voltage networks, to such an extent that it is very rare for the DNOs to have data monitors beyond the primary substations (Karimi et al. 2013). For example, in the UK normally the nearest monitoring point to the low voltage network has usually been at the 33kV to 11kV substation (Lees 2014; Stephen et al. 2014).

Since the early 1990s, experts have been pointing out the limitations of data available to DNOs and have been arguing that merely having voltage and power readings at substations are not sufficient to ensure the quality of supply to customers and more detailed data are required (Baran 1993; Baran and Kelly 1994). In recent years, the DNOs have been studying the benefits of installing advanced meters at low voltage substations (e.g. CLNR project). However, the enormous costs involved in fitting every low voltage substation with advanced meters has encouraged the DNOs to turn their attention to the information data that can be obtained from customer smart meters at no additional costs.

In 2010, Imperial College London and Energy Network Associations (ENA) carried out a study to identify the extent of benefits that will be gained by implementing smart meters to achieve a real-time operational control of the Smart Grids (Strbac et al. 2010). This work was further taken forward in ENA (2012); Balta-Ozkan et al. (2014); and Miller (2015) that argued the importance of smart meters in the future of the electricity distribution grid as the UK Government has been looking to devise ambitious renewable energy strategies in order to decarbonise the grid by 2030 and beyond. Smart meter data is highlighted by Strbac et al. (2010); ENA (2012); Balta-Ozkan et al. (2014); and Miller (2015) to be the main enabler in the transition towards a more proactive network management which provides the capability of utilising schemes such Demand Response

Management (DRM), which could potentially decrease and/or shift the future daily peak demands resulting from widespread use of electric vehicles and combined heat pumps from 50% to just under 30% of the network capacity.

They provide five different electric vehicle and combined heat pump penetration scenarios in Figure 2-3 that shows the various possible penetration levels of low carbon technologies in the UK's electricity grid by 2030, ranging from the most pessimistic (10%) to the most optimistic (100%) scenario. This shows that even in the most pessimistic scenario, there will be a far greater need for network observability and control, most notably via smart meters, in order to manage the new demand patterns on the low voltage network.

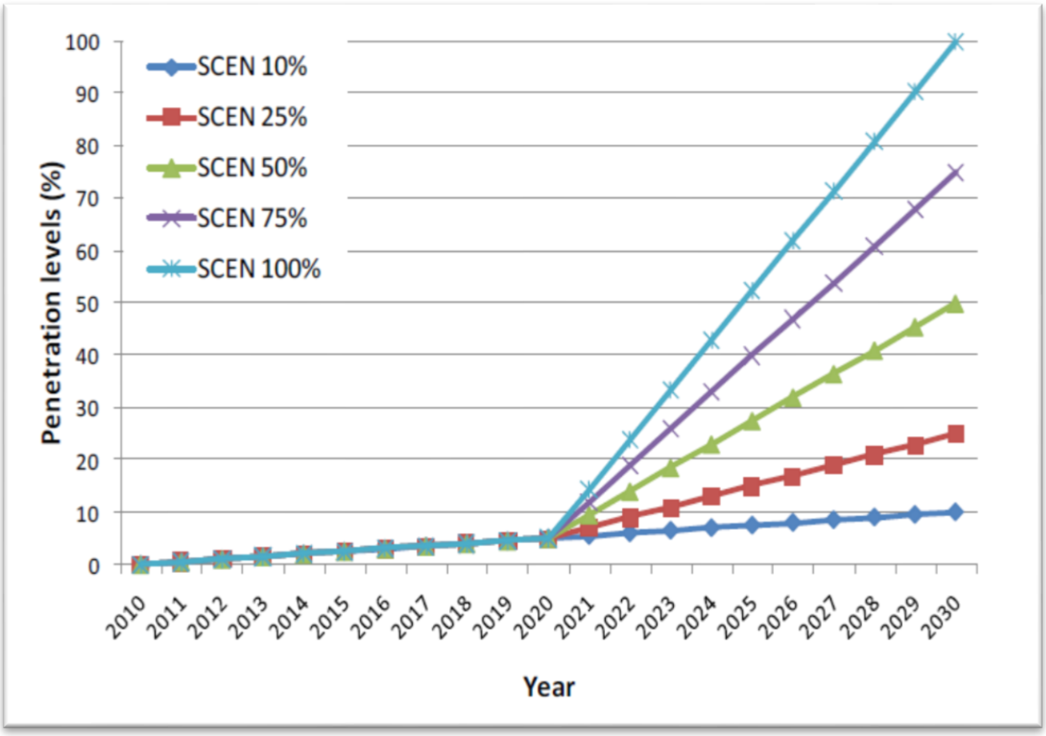


Figure 2-3: Future Low Carbon Technology (LCT) penetration scenarios (Strbac et al. 2010)

EA Technology (2015b) have also made an estimation of the number of units of electric vehicles and heat pumps that will be installed in 4 of the main DNO license areas. These projected numbers are shown in Figure 2-4 below.

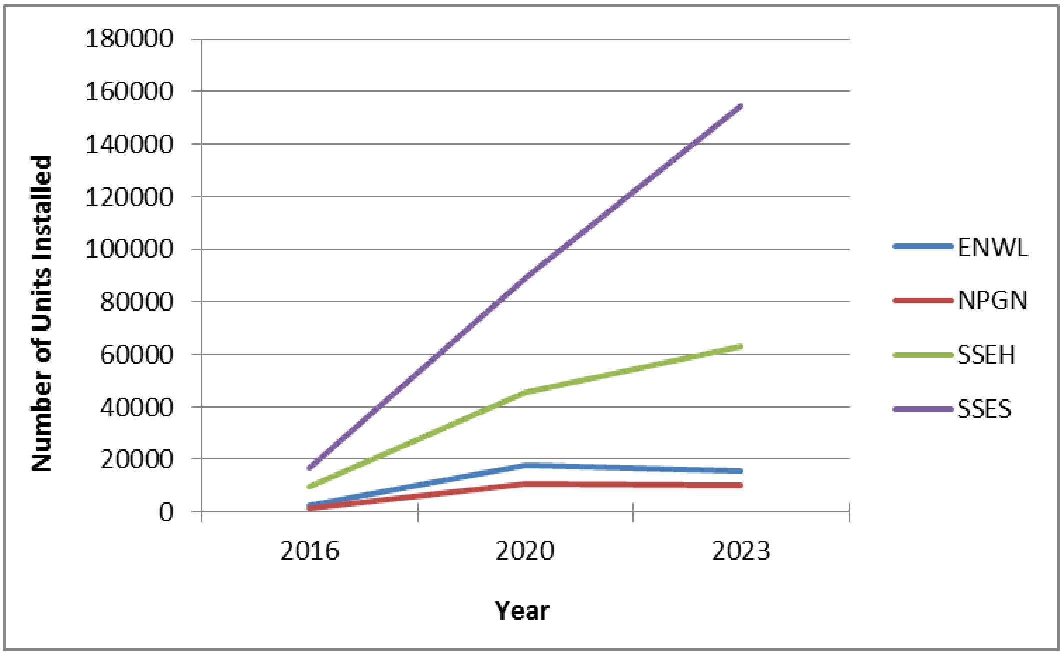


Figure 2-4: Electric vehicle and heat pump installations per license area (EA Technology 2015b)

As Figure 2-4 shows the volume of low carbon technologies installed on the low voltage network will increase significantly and the DNOs need to manage these new patterns more proactively than before.

Strbac et al. (2010) define two network operation modes of Business As Usual (BAU) and "Smart", that represent the passive and active operation of the grid, respectively. They also estimate that the latter (Smart) mode of operation, empowered by smart meter information, will lead to lower costs of network reinforcement in the long run, especially in scenarios with lower penetration levels of electric vehicles and combined heat pumps (Strbac et al. 2010). In their "Smart 2050" scenario, Balta-Ozkan et al. (2014) also define the best case smart grid development scenario in the UK. In this scenario low carbon technologies are fully integrated into the smart grid applications and the DNOs and customers are both fully engaged in active management of the networks. These scenarios are driven by the assumption that smart meters will pave the way for extensive adoption of demand response programmes, which is still open to debate.

In this light, Strbac et al. (2010) and ENA (2012) foresee the following benefits from the deployment of smart meters in the UK:

- Reductions in generation demands by accommodating demand response programmes and utilisation of renewable generation
- Improving the flexibility of the grid
- Improving the decision making in fault restoration and network investment

This will be even more vital as DNOs become more active in the operation of the distribution networks and develop smarter ways of information gathering, control technology and distributed resource integration, which in turn will drive the development of Advanced Distribution Management Systems (ADMS) for the improved operation of electricity networks (Fan and Borlase 2009; Arritt and Dugan 2011; Uribe-Perez et al. 2016). A number of studies have raised the point that the future Smart Grids with embedded generation will be heavily reliant on real-time information from smart meters or AMIs to facilitate a more proactive network operation (Meliopoulos et al. 2011; Karimi et al. 2013; Quilumba et al. 2015; Spencer 2015). It is also highlighted that the current limited Supervisory Control and Data Acquisitions (SCADA) data in relation to MV/LV network information has become less prevalent at lower voltage levels of the distribution network as highlighted earlier (Baran and McDermott 2009; Karimi et al. 2013).

2.3.1 Integration of Smart Meter Data into DNO Applications

The smart meter deployment programme in the UK is a gradual process taking place between 2014 and 2024. Therefore, while the DNOs adapt their applications to integrate new streams of data from smart meters, it is highly likely that smart meters in some areas of the network are not available, so the half-hourly demand/generation data from some parts of the networks will not be available to the DNOs. Also, the smart meter data which are transmitted to the DNOs will be half-hourly averages of the customer loads, which in itself can potentially decrease the accuracy of the data. Additionally, the DNOs are required to anonymise and aggregate these half-hourly averages from the customers (ENA 2015; OFGEM 2017), which can potentially introduce another element of inaccuracy.

These policies and regulations vary from country to country in terms of the time resolution intervals and aggregation requirements of the smart meter data. For example, in the United States the smart meter firmware can be updated by the utilities to obtain higher resolutions of data and there are also no requirements in terms of anonymization of the customer data by the operators (Mirowski 2014). Also some European countries have decided on higher or lower time resolutions of smart meter data and data access arrangements. For example, Belgium, Austria, Italy, and Denmark have chosen higher time resolution of 15 minute intervals, whereas countries such as Estonia and Sweden have opted for lower time resolutions of 1 hour intervals (Ilves 2016). Potentially, these specifications can affect the quality and capability of smart meter data to provide the DNOs with accurate low voltage network information which in turn will limit the ability of the DNOs to play a more proactive role in the operation and control of the low voltage network and the future smart grid applications.

From the viewpoint of the DNOs, the issues which can potentially limit the benefits of the smart meter data in the UK are as follows:

- Missing customer meter data, due to faults or non-existence of smart meters
- Time resolutions of the smart meter data available to the DNOs
- Lack of individual customer data, due to aggregation of the data
- Lack of customer phasing information, due to radio frequency technology used in the UK for the transmission of data from smart meters

Major smart meter and smart grid trials in the UK have attempted to investigate and clarify the changes that will have to take place on the low voltage distribution side of the electricity network in order for the DNOs to be able to overcome the challenges of accommodating new generation and demand patterns as a result of higher take up of embedded generation and low carbon technologies at customer levels. For example, Npower (2011) and Western Power Distribution (2013) put a strong emphasis on the role of the smart metering systems in providing the required sources of information to facilitate the transition to decarbonised distribution grids. A study by the UK Power Networks and TE Connectivity (UK Power Networks 2013) highlights the need for the DNOs to better monitor the loads on the low voltage networks and use the information to optimise the network capacity as well as identify the areas that need reinforcement and replacement. In other words, using new data can provide more visibility on the low voltage network to actively manage the loads and generation on the network and can lead to making better informed decision about network planning and design as well as network asset management (UK Power Networks 2013).

According to a report by EA Technology (2016), the areas that the DNOs are interested in developing in the future or have already developed to some extent are areas such as ANM, voltage control, fault level management, planning of new connections, dynamic rating, Demand-Side Management (DSM), using network storage, and using smart meter Time of Use tariffs to manage customer demands (EA Technology 2016). ANM is particularly important in managing network loading capacity, voltage levels and power flows and many DNOs are integrating the ANM measures to some extent as more renewables are installed in their existing network (EA Technology 2016). For example, better voltage level management can allow a DNO to install higher shares of embedded generation in the low voltage network while maintaining the voltage levels in the statutory limit ranges of 230V +10% -6%. EA Technology (2016) argues that the more in-depth behavioural knowledge of the low voltage networks is becoming ever more pressing to the DNOs as customers' demand and generation are becoming more varied and intermittent.

Western Power distribution (2013) and Barbato et al. (2017) emphasise that the role of smart meters and AMIs are becoming increasingly more crucial in providing a clearer picture about the behaviour of customers at the low voltage level and ultimately a more in-depth understanding of demand and generation patterns at the low voltage side of the network. Potentially, these new sources of information can dramatically enhance the conventional low voltage network. The following sections of this chapter investigate the important and relevant research on the major low voltage network information areas and the DNO applications that will be even more greatly dependent on accurate low voltage network information as the Smart Grids evolve.

The current work addresses the impact of time resolution, missing meter data, and customer data aggregation on the accuracy of major low voltage network information. The information includes network performance indicators such as network losses, voltage levels, low voltage cable loading percentages, and load prediction. Consequently, the findings are contextualised in terms of the impacts on the major low voltage network applications such as fault detection and restoration, asset management, ANM, DSM, and network planning and design.

2.4 Fundamental Low Voltage Network Information Areas

In this section, traditional approaches to estimating missing loads, network losses, voltage variations, and cable capacity percentages and the ways in which smart meter data can improve them are described. These are critical low voltage network indicators that inform and enhance the DNO applications which will be discussed in section 2.5.

The main reasons why parameters such as losses, voltage levels, customer loads and cable loading percentages have been chosen in this study as the main low voltage performance indicators are:

- the direct relationship of these information from these indicators with the objectives of the Smart Grid agenda in the UK.
- the importance of these information areas to the DNOs and the regulatory bodies and the reliance of smart grid applications on them.
- the possibility of quantifying the impact of time resolution, missing loads, and customer data aggregation on the accuracy of the information.

The following sections provide an overview of the research that has been carried out on the ways in which the mentioned low voltage network performance indicators are obtained.

2.4.1 Estimation of Loads on Low Voltage Networks

Since the turn of the century, a great deal of research has been carried out on the short-term forecasting of the loads on the network (24-hour ahead) using Probabilistic Neural Network (PNN), Bayesian Neural Network (BNN) (Laurret et al. 2007), or Artificial Intelligence (AI). However, the majority of the research have been focused on the higher voltage levels of the network (Amjadi 2001; Beccali et al. 2004; Gerbec et al. 2005).

Load forecasting methods, or more specifically Short-Term Load Forecasting (STLF) techniques, have long been devised, developed, and employed by transmission network operators for an effective electricity network planning and operation as well as efficient retail purposes (Gerbec et al. 2005; Lauret et al. 2007; Taylor and McSharry 2007; Hahn et al. 2009). More recently however, more research has been carried out on utilising smart meter/AMI data in the STLF techniques by Baran and McDermott (2009); Ghofrani et al. (2011); Aung et al. (2012); Alzate and Sinn (2013); Mirowski (2014);

Quilumba et al. (2015). An overview of these studies are presented in the following section and a more detailed investigation of the methods used is found in chapter 3.

The methods that have been developed in order to compensate for the lack of distribution network data and network visibility can be divide into three main categories of Distribution State Estimation (DSE), Load Allocation (LA), and load profiling. These methods are discussed in the following sections, respectively.

2.4.1.1 State Estimation (SE):

SE algorithms aim to provide voltage magnitude and phase angles at every bus on the transmission system (Celik and Liu 1998; Yih-Feng et al. 2012). Yih-Feng et al. (2012) describe State Estimation (SE) as a function that can provide more accurate monitoring data to various grid operation applications by obtaining a real-time picture of the network. SE for power systems was first introduced by Scheppe and Wildes in 1970 and it was mainly based on steady state operation conditions (Scheppe and Wildes 1970). Since then, SE has been a major part of network operation of power systems, especially at transmission levels (Celik and Liu 1998). Accurate SE models are useful in optimising power flows, as well as security and reliability analysis. Celik and Liu (1998) point out that at transmission levels, most network analysis tools rely on SE algorithms to obtain a system state picture of the network.

On the other hand, SE techniques on the distribution side have lagged behind compared to the transmission side due to the design and operational differences between the two systems (Hayes and Prodanovic 2014). Transmission networks and distribution networks are different in a number of ways, which make the SE methods for each type of network very different. For example, there are far more measuring points and measurement types available at higher levels of the networks. Also, distribution networks are much more varied, complex, and extensive compared to transmission networks (Hayes and Prodanovic 2014). In the 1990s, the Distribution System State Estimate (DSSE) flourished in order to cope with the growing shares of embedded generation being installed in the distribution network (Hayes and Prodanovic 2014).

On higher voltage levels, state estimator programmes act as providers of information to the SCADA system, which is a major part of the network operating system, especially in relation to Distribution Automation (DA) (Celik and Liu 1998; Baran and McDermott 2009; Yih-Feng et al. 2012). Karimi et al. (2013) explain that a statistical

SE algorithm provides voltage magnitudes, phase angles and power flow information at node points by utilising available data and customers' load data, which are either measured or predicted from historical data, load forecasting methods, or Load Allocation (LA) methods.

A number of researchers have pointed out that most SE approaches are suited for the transmission network, where there are more monitoring points on the network. Ghosh et al. (1997a); Celik and Liu (1998); Yih-Feng et al. (2012); Meliopoulos et al. (2011); Karimi et al. (2013) all raise the issue that the lack of real-time information at the distribution level has been hindering a widespread implementation of SE methods for the purpose of DA. They also highlight the fact that these transmission state estimator approaches fail to take into account the data limitations of traditional distribution networks.

Celik and Liu (1998) also suggests that despite the fact that SCADA systems can contribute to dispatch teams and network engineers by providing some types of real-time data, the distribution network applications are far from integrating these data effectively, especially at lower voltage levels of the network.

These issues that have hindered the implementation of SE methods at lower voltage levels can be overcome with the introduction of smart meters and AMIs to the low voltage networks (Uribe-Perez et al. 2016). More research is now under way to develop distribution SE methods that can be appropriately applied to distribution network scenarios, which are gaining even more significance by the development of the Smart Grid agenda and rising penetration level of embedded generation on the consumer end of the network (Meliopoulos et al. 2011; Karimi et al. 2013). Estimation of customers' loads still remains one of the main aspects of SE. This is known as Load Allocation (LA) and is discussed further in the following section.

2.4.1.2 Load Allocation (LA):

Load Allocation (LA) methods constitute techniques that the network operators have developed over time to assign realistic load estimates to network transformers (Ghosh et al. 1997; Wong and Chung 2015). SE algorithms rely on customers' load data which are not normally available in real-time on electricity distribution networks, hence state estimators require load models which are produced by LA methods (Ghosh et al. 1997b; Poursharif et al. 2015). Research has been carried out on methods that can provide load demands at node points, mainly on MV distribution network scales. Efforts have been focused on moving away from traditional estimation of loads based only on peak loading conditions (Ghosh et al. 1997a). On the other hand, a lack of research in the area of low voltage load modelling is clear, which is an issue considering the need for a more proactive distribution network operation and the fact that the distribution network forms 48% percent of the total length of the UK distribution and transmission networks (EurElectric 2013).

A better knowledge of low voltage network loads is required, and the more realistic the load modelling technique is, the more accurate the results of SE and other advanced applications fed by SE results are, which include estimation of power losses, power and reactive power and can also contribute to voltage optimisation, voltage and reactive power control, feeder reconfiguration, and demand side management (Ghosh et al. 1997a; Ghosh et al. 1997b; Carmona et al. 2010; Baran and McDermott 2009; Sharifian et al. 2012). A number of other processes such as power flows, fault detection, and service restoration benefit from data produced by LA methods (Carmona et al. 2010).

Traditional LA Methods

In the UK, usually the lowest load monitoring point on the distribution network has been of the 11kV feeders at 33kV substations (Karimi et al 2013; Lees 2014). Beyond this point, the approximately yearly consumption (billing) totals from the individual low voltage customers and the maximum currents recorded at distribution substations have been available. The main low voltage current estimation methods have centred on the peak currents for use in the design and extension of low voltage networks. The two approaches that have been widely employed in the UK for estimating these are:

1. After Diversity Maximum Demand (ADMD) (McQueen et al. 2004): this approach was developed in the 1950s. The maximum demand for a customer of a particular type is specified. The maximum demand from a number of these customers is then taken to be this individual maximum demand times a diversity factor that is a function of the number of customers in the group.

2. Customer demand curves (Carson and Cornfield 1973; McQueen et al. 2004; Vélez et al. 2014): allowing customers with peaks at different times to be modelled.

The LA methods have been improving to take into account various factors other than only peak loads based on transformer kVA and customers' peak consumption or predefined demand tables and voltage drops (Ghosh et al. 1997b; Carmona et al. 2010).

The traditional LA approach, that has been used by many network analysis and statistical estimation programmes, utilises the consumer transformer ratings and their nominal kVA to calculate the loads (Kersting and Philips 2008). The accuracy of these values or pseudo measurements can be enhanced by integrating variables such as customers' monthly consumption values, customer types, time of day, and weather conditions (Sharifian et al. 2012; Poursharif et al. 2015).

Improvements on Traditional LA Methods

Lee and Etezadi-Amoli (1993) attempt to improve the traditional LA method for electrical distribution systems with incomplete information by acquiring a ratio factor in relation to customer type coincident factors. They classify customer types into two categories of residential and commercial and assign two different power factor values to them (Lee and Etezadi-Amoli 1993). This power factor is based on average power factor for each customer type (Lee and Etezadi-Amoli 1993). The coincident factor and power ratio factor values are then applied to the substation measurements (Lee and Etezadi-Amoli 1993).

Although this is an improvement on the traditional transformer rate method, this method still relies on peak load measurements at each feeder and substation, which were the available data streams at the time of the study. At the time of that study, information such as power factor of the circuit, load phases, coincident factors, and load distribution patterns were not available to network engineers. Therefore, most network models were not accurate due to the lack of real time information downstream of substations or the costs associated with acquiring more detailed information (Lee and Etezadi-Amoli 1993).

This lack of data would lead to a lack of correct information about problems at lower levels of the network by hiding the individual load pattern variations. This is a major problem when using aggregation of customer loads as opposed to using more fine-grained disaggregated customer data. For example, a linear factor approach commonly known as substation kVA method considers a direct proportional relationship between electrical demand and the total kVA of installed transformers in that area which clearly neglects the variation in demand profiles of different customer types (Lee and Etezadi-Amoli 1993). Lee and Etezadi-Amoli (1993) argue that using different coincident factors for different customer types in conjunction with some meter data instead of kVA, produces more accurate LA results.

Ghosh et al. (1997b) focus further on customer information and the use of customer class curves and consumers' billing information on a 934-bus distribution circuit to present load estimates and a level of uncertainty in the results. They aim to present an LA model that can represent the real time measurements which are very scarce in the reality of distribution network operation and planning. They also try to improve a major drawback of the traditional LA method, which is not providing a level of uncertainty in load estimations (Ghosh et al., 1997b). At the time, there was little literature on this subject apart from the work of Lee and Etezad-Amoli (1993), so Ghosh et al. (1997b) proposed a model to improve the traditional LA method by making use of all power flow measurements and available customer usage information from bills to present two values of expected loads and certainty variance associated with the expected loads. Load curves for each customer types are created by statistical analysis and customer types Load Modelling Factors (LMF) are applied to the values to adjust the expected load results (Ghosh et al. 1997b). While this method makes use of customer class types and time of day information, the various demand profiles within customer types are neglected, which may not be very crucial at MV levels with large number of consumers but it will be at low voltage levels. Early references to the need for the use of AMI data can be found in the works of Brakkan et al. (2006) and Carmona et al. (2010). This is mainly due the investment that utility companies made at the turn of the century to gain more information about their customers (Brakkan et al., 2006).

Brakkan et al. (2006) emphasise the need for the integration of load allocation values with the network models in Geographical Information Systems (GIS), while Carmona et al. (2010) raises the importance of integrating accurate LA results in modern Distribution Management Systems (DMS) of the Smart Grid and the dependency of processes such as power flow analysis, fault management, and Volt/Var control on this integration. Carmona et al. (2010) propose a Load Allocation model based on three customer types of residential, industrial, and services and then divides them up into smaller clusters based on similarities in load types. The model uses the solution of leastsquares and is carried out on MV levels in a context where each MV substation feeder is connected to customers via distribution transformers (Carmona et al. 2010). Carmona et al. (2010) present the best and the worst case scenario ranges, filling in the missing information with forecasted information and daily reference values at certain hours based on historical data.

Kersting and Philips (2008) and Arritt et al. (2012) have presented LA methods which make use of AMI data in some way. They both use 15 minute time resolution meter data but differently and on different scales. Kersting and Philips (2008) work on 3 months of meter data from 314 consumers connected to 23 distribution transformers and make use of a concept termed “diversified demand” which is defined by them as period by period sum of the values of the 23 transformers. They argue that the traditional kVA method of LA can cause overestimation of demand and there is no direct relationship between the transformer rating and the number of consumers served by it. In this study, results from four methods of LA are compared to the values from AMI to determine which method best reflects the true values obtained from AMI readings. The three methods investigated are based on daily kWh, Monthly kWh, transformer kVA which is proved to be less accurate compared to using maximum diversified demands of the transformers obtained from 15 minute AMI data (Kersting and Philips 2008).

On the other hand, Arritt et al. (2012) point out that with the widespread presence of AMI in the future, different load shapes can be drawn for different customers and that information can be fed into network modelling and simulation programmes. They make use of 15 minute resolution AMI data of a network with 1,179 consumers for a period of almost 4 months and also compare the LA results of four different methods including AMI allocation, transformer kVA allocation, monthly usage allocation, and class load shape allocation, but in the context of different network operation and planning processes (Arritt et al. 2012). While Arritt et al. (2012) prove that aggregated AMI load measurement can significantly improve the processes and State Estimation studies of distribution network, this is carried out at MV scale on a network with high numbers of consumers which may not be, at least in theory, susceptible to variation demand profiles of individual consumers.

2.4.1.3 Load profiling and clustering methods

Clustering methods aim at exploring and finding inherent structures in sets of data by merging similar behaviours into clusters (Jain 2010; Klonari et al. 2015). Load profiling methods are based on clustering the data from various numbers of customers and using statistical methods such as mean and meridian to produce representative load curves (consumption patterns) for different customer types based on their location or social status (Klonari et al. 2015; Al-otaibi et al. 2016). As Jain (2010) highlights a number of methods have developed over the years in the field of cluster analysis, yet k-means still remains one of the most important methods which is indicative of the difficulty of devising a high performance clustering algorithm.

Recently, there has been a number of studies on improving the load profiling and clustering of the consumers using high-resolution smart meter data in order to improve the traditional customer type classifications. For example, in the UK, domestic customers have been divided into two broad categories of either consumers on Economy 7 tariff or the non-economy 7 tariff and the industrial customers have been placed into 6 broad categories (Al-otaibi et al. 2016).

Stephen et al. (2014) and Al-otaibi et al. (2016) both highlight the fact that most DNOs have been and are using annual or quarterly consumption data in their network performance indicator estimations. The provision of smart meter data can improve this situation. Mutanen et al. (2011) use Iterative Self Organizing Data Analysis Technique (ISODATA) to group various commercial consumers into representative profile classes at the MV level. There is far less deviation from the average load patterns in large samples of data from a large number of customers. However, this is not the case at the low voltage level and due to smaller number of customers, the individual load patterns become significant (Mutanen et al. 2011).

Stephen et al. (2014) employ various linear Gaussian load profiling methods combined with customer smart meter data to model a particular type of consumer’s load pattern on a given day and they show that the peaks and drops in customer demands can be simulated with a high degree of certainty. Labeeuw and Deconinck (2013) utilise data from 1,300 customers using a top-bottom approach and Markov models within the customer clusters to produce representative load profiles. Ferreira et al. (2012) employ temporal data to examine patterns within the clusters of customers and produce seasonal load profiles. Mcloughlin et al. (2015) use three methods of k-means, k-medoids, and Self Organising Maps (SOM) to examine the best ways of clustering customers together based on their pattern electricity use and then produce typical consumer profile classes. The methods employed in the studies presented in this section will be further investigated in chapter 3 and chapter 4 where the justifications and the explanations for the methodology employed in this thesis are described in detail.

2.4.2 Estimation of Low Voltage Network Losses

Network losses are one of the main network performance indicators for the DNOs and will become increasingly more crucial in the future Smart Grid scenarios. Power losses on electricity networks are strong indicators of the efficiency level of the network system delivering power from generators to customers. The allocation of these losses is not an easy task, because the losses and the customer loads on the network have a non-linear relationship (Slavickas 2000). Network losses make the provision process of electricity to consumers more expensive and the costs are usually passed down to the customers (Slavickas, 2000). For example, in the USA these network losses account for 5 to 10 percent of the power produced to be transmitted and can lead to the losses of millions of dollars worth of energy (Alturki 2011). In Europe, this figure can range between 2.3 to 11.8 percent of the electricity generation (Heckmann et al. 2013). In the UK, network losses constitute approximately 5% of the energy transmitted to customers.

The DNOs, under the new OFGEM Losses Incentive Mechanisms, are having to spend around £100 million per year to reduce the network losses (Sohn Associates 2009; OFGEM 2016). Network losses are far greater on the low voltage side compared to the higher levels of the distribution network, but accurately estimating the amount of losses has been particularly challenging to the DNOs, due to varying characteristics of each low voltage network and the lack of information on them.

Figure 2-5 below shows the proportion of losses experienced on the distribution network in the UK compared to the total losses.

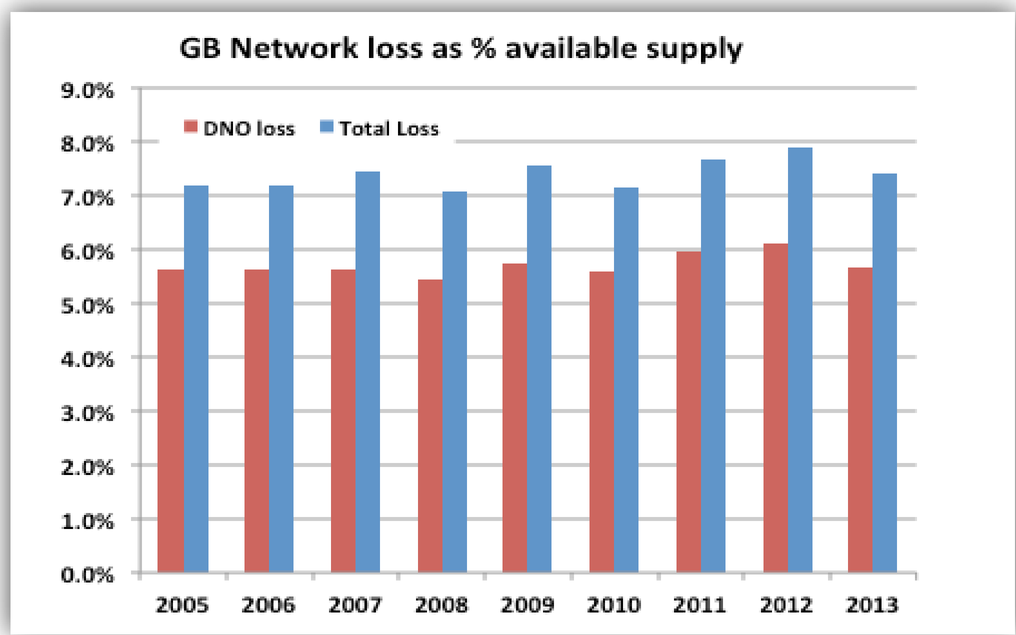


Figure 2-5: Proportion of losses on the electricity distribution network in the UK (Sohn Associates 2013)

As Figure 2-5 above demonstrates that a large proportion of losses on the electricity network occurs on the distribution network side managed by the DNOs.

2.4.2.1 Loss Estimation Approaches

The simplest approach is to measure the difference between the input and output energy at each end of the low voltage network as suggested by Urquhart et al. (2017), but this requires measuring devices on both ends of the low voltage network. This method also fails to distinguish between technical and non-technical losses or localise where on the network the losses are occurring. Therefore, various studies, especially in the field of energy management research have been carried out over time to (Fourie and Calmeyer 2004):

- estimate technical/non-technical losses of the power networks
- decrease the technical losses
- investigate the effects of embedded generation on the power losses

For example, square root of mean current (I^2), Artificial Neural Network (ANN), regression analysis, clustering models, and Radial Basis Flow (RBF) network, and forward backward sweep are amongst the methods used in the absence of smart meter data (Hui-Ian et al. 2005; Zhang and Bo 2010).

Non-technical losses are mainly related to illegal withdrawal of power from the network as well as unmetered loads such as street lights. However, technical losses are mainly caused by the resistance (impedance) in the cables, losses from heat, or electromagnetic fields (Fourie and Calmeyer 2004; Diop and Cercle 2005). Before the availability of smart meter data, the energy usage from each customer had been measured over long periods, e.g. 6 months or a year, and these periods were extremely unlikely to be the same for all customers on a low voltage circuit, e.g. the starting day of the period will vary for different customers. Even if the metering periods for all customers did coincide, the breakdown of the losses to individual conductors and time periods would not be known (Dashtaki and Haghifam 2013).

A new element that can influence the network losses is the introduction of embedded generation and low carbon technologies into the distribution grid. Salman (1996) argues that the introduction of embedded generation into the low voltage distribution system can either increase or decrease the power losses depending on factors such as the location of the generators, topology of the network, and the ratio of generation to the load demand on the network. While Bell et al. (2009) discuss that the introduction of generation into the distribution system can decrease the network energy losses, Fourie and Calmeyer (2004) point out that there is only a certain degree to which technical network losses can be reduced, due to the physical restrictions of the electricity network.

This is particularly challenging to assess for the DNOs, because the sizes of currents on the low voltage network with or without embedded generation or low carbon technologies is relatively unknown to them (Costa and Matos 2004; Carpaneto et al. 2006; Frame et al. 2012; Jagtap and Khatod 2015a; Jagtap and Khatod 2015b). Studies such as Beddoes et al. (2007) and Karimi et al. (2013) demonstrate that the distribution networks with embedded generation normally experience a higher percentage of peak losses, which signifies the importance of the size, location and the operating mode of the embedded generation units connected to this level of the network (Zhang and Bo 2010). In fact, Zhang and Bo (2010) demonstrate that network losses can be minimised by almost 76% by choosing the appropriate size, location, and operating mode of the units in the distribution network. In Heckmann et al. (2013), the impact of embedded generation on losses is examined at various levels of the distribution network. They argue that although initially the introduction of distributed generation reduces the network losses, in the long run and with more penetration levels, this decreasing effect slows down or is cancel out.

Loss Estimation in the Absence of Smart Meter Data

In the absence of detailed customer load data from smart meters, Diop and Cercle (2005) use a regression analysis model to link the annual load measurement at MV and low voltage substations to model the total network losses including both technical and non-technical losses. The main drawback of this approach can be its restricted applicability to the networks with similar topology and consumption patterns (Diop and Cercle 2005). Beddoes et al. (2007) take an interesting approach to estimate the technical losses of a distribution network and the ways in which the penetration of embedded generation can affect the technical losses in all voltage levels ranging from 132 kV to 0.4 kV. In this work, the element of time is introduced and the fact that technical loss calculations become more intricate as the network and the level of distributed generation integration increase due to consumption and generation variations in time (Beddoes et al. 2007). Three main elements that were used in this study are Grid Supply Points (GSP) on the network, distributed generation penetration level, and real customer demand data (Beddoes et al. 2007). Three types of urban, rural, and mixed low voltage networks are synthesized based on annual demand data from EA Technology and various embedded generation penetration scenarios to examine the impact of distributed generation on annual or daily technical losses of different low voltage networks (Beddoes et al. 2007).

Au et al. (2008) use the feeder characteristics and general customer type load profiles to calculate the technical losses on the low voltage network. Characteristics such as feeder length, ratio of demand to maximum capacity and the load distribution profile on the low voltage cables are used by Au et al. (2008). This approach calculates the maximum percentage of power losses along the 11 kV and low voltage cables based on the mentioned characteristics (Au et al. 2008).

Another approach has been to estimate low voltage losses using loss factors (Oliveira and Padilha-Feltrin 2009; Quiroz et al. 2012). This factor is used to multiply the peak load losses to the given average losses. The attraction of this approach is that maximum demand is often measured (or estimated) for low voltage circuits and so, it provides a straightforward way to estimate the losses. However, this method only provides a rough estimate of the losses as the relationships between the peak demand and the peak losses, and between the peak losses and the average (or total) losses, are very dependent on the circuit's characteristics and the shape of the load curves at different points on the circuit. Quiroz et al. (2012) argue that using the average demand rather than the maximum demand is better as it reflects a period of time rather than one time instant.

An alternative approach estimates a low voltage circuit's losses by matching the circuit with a set of benchmark circuits. Various features can be used for the matching, for example, Dashtaki and Haghifam (2013) use the main feeder length, the length of branches, the number of branches, customer information and conductor sizes. The approach relies on the benchmark circuits having been modelled in detail, and so their calculated loss values are regarded as being accurate. However, as previously mentioned by Dortolina and Nadira (2005) a flaw of this type of method is that low voltage networks vary greatly in size and topology.

Improvement of Loss Estimation Methods Using Smart Meter Data

A general weakness of the mentioned approaches is that they only provide a single figure for the losses rather than providing a geographical and temporal breakdown of the losses. Brandauer et al. (2013) note that breaking down the losses is becoming more important due to decentralised generation and the move towards smart grids. They look at the consequences of simplifications such as using mean or peak loads by combining existing standard load profiles with smart meter data. Brandauer et al. (2013) found that existing loss estimation approaches had particular problems in low density rural areas. These branches were also sensitive to the time resolution of the data used to calculate losses with losses calculated using one second mean values being up to 20% higher than those calculated using 15 minute values.

Having customer smart meter readings available for a low voltage circuit will allow the "copper" or technical losses to be estimated using a load flow analysis (Carpaneto et al. 2006; Quiroz et al. 2012). Not only will this avoid the coarse approximations involved in using loss factors and allow a temporal breakdown of the losses, but the consequences of phase imbalance (Frame et al. 2012) and embedded generation can be accounted for. However, although their measurement time periods are much shorter than the months or years of the meters that they are replacing, the typical measurement time periods of 15, 30 and 60 minutes (McKenna et al. 2011) mean that the losses calculated using smart meter data underestimates the true losses (Brandauer et al. 2013; Urquhart and Thompson 2015).

Urquhart and Thompson (2015) investigated the effect of the time period length on the calculated losses and considered the losses from a single appliance switching on and off at random. For short time periods, i.e. in terms of seconds rather than minutes, the underestimation was modelled (and validated) as being a linear function of the time period. As the time period increases, the assumption of there being at most a single switching event (either off to on or on to off) in any time interval breaks down and the relationship stops being linear. Comparing the summed demands from between 1 and 22 dwellings indicated that the relationship between losses calculated using 1 minute resolution and larger time resolutions became closer to a linear one as the number of dwellings increased. Figure 2-6 from Urquhart and Thompson (2015) shows the ratio of estimated losses at different averaging periods for different sizes of customer groups.

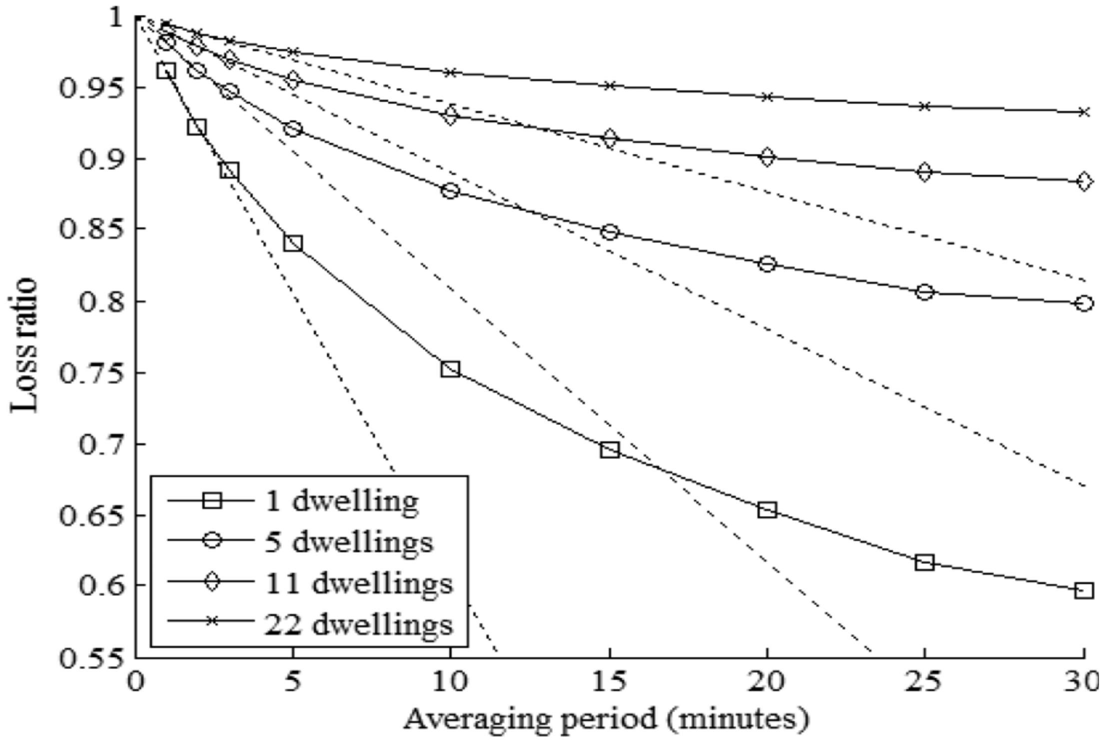


Figure 2-6: Loss ratio for grouping of dwellings (Urquhart and Thompson 2015)

The weakness of this work is that the network model is limited to 22 houses and the houses are arranged in a single phase model. This results in the losses being calculated for each house without considering the neutral phase currents and the effects of the loads from other customers on the various phases on the calculation of overall low voltage network losses.

Clearly, the estimation of network losses is a pressing issue for the DNOs and having information about the outgoing power and the delivered power, which can be obtained or allocated from smart meter data can help measure the minimum and maximum technical losses experienced on the network. Cao et al. (2009) highlight the fact that the ageing UK electricity network structure can greatly benefit from having accurate information about the power deficiencies of the existing networks. This was also highlighted earlier by Davidson and Ljumba (2002). They also argued that network operators can pursue more advanced asset management, Research and Development (R&D), and personnel allocation if they have accurate network loss information in hand (Davidson and Ljumba 2002).

It is also extremely likely that the network operators may soon face statutory requirements to maintain their network losses within certain limits and this has already begun in countries such as Germany and Sweden (Heckmann 2013).

2.4.3 Estimation of Voltage Drops on Low Voltage Networks

The issue of voltage variations on the low voltage network is becoming more significant as new generation and demands such as embedded generation and low carbon technologies (e.g. electric vehicles and combined heat pumps) are introduced to the low voltage side of the network. Kaspirek (2013) and Konstantelos et al. (2017) argue that with rising levels of embedded generation installed in the distribution grid, it will become even more problematic for the DNOs to maintain the voltage levels within the statutory limits of 230V +10%/-6% in the UK (Miller 2015).

Often, integration of embedded generation in the distribution system can cause problems for the DNOs in terms of voltage regulations and meeting the statutory voltage limits of between 216V and 253V (Salman 1996; Conti 2001; Caire 2002; Scott et al. 2002; Caldon et al. 2005; Tonkoski et al. 2008; Wang et al. 2016; Konstantelos et al. 2017). Voltage variations can also affect the network losses which were discussed previously (Desmet et al. 2007). It will be increasingly crucial for the DNOs to be able to anticipate the voltage variation ranges experienced on the low voltage network, because voltage drop levels of a network system is a significant indicator of the quality of power that is delivered by the system to the end consumers (Vujošević et al. 2002; Konstantelos et al. 2017). Klonari et al. (2016) argue that when the demand on the low voltage is low and solar panels are producing high amounts of power (e.g. summer time), the overvoltage occurs in the system that not only affects the quality of power delivered to the end users, but can also disrupt the injection of power from the prosumers into the low voltage network.

To a large degree, voltage levels on the low voltage networks are affected by the load on each section of the network, i.e. the load variations can affect the voltage behaviour on the cables (Vujošević et al. 2002; Konstantelos et al. 2017). Therefore, a good knowledge of customer loads on the network can lead to accurate voltage drop estimations. Lin et al. (2012) argue that the most significant factor in voltage variation is load power factor, in particular at the end of the supply networks where the substation reactive power and load power drop. Emelin et al. (2013) argue that the voltage variations are directly affected by load density on that section of the distribution network. Konstantelos et al. (2017) argue that the voltage variations on the low voltage networks can directly impact the ability of the DNOs to accommodate embedded generation on the networks. Smart meter data collected or allocated, which can be determined using methods such as in chapter 4, can introduce new possibilities of greater voltage variation monitoring and control to the DNOs.

In the past, a number of studies have been carried out to investigate the ways in which voltage levels in distribution systems with distributed generation can be regulated. For example, Vujošević et al. (2002) estimate voltage variations using global parameter method based on length and section of the distribution cable applicable to all levels of the distribution network.

Caldon et al. (2005) assume that active power factors of the embedded generation on the low voltage network cannot be managed with great certainty using the known methods such as HV/MV substation On Load Tap Changer (OLTC) transformers, or shunt compensators, however they can be used in conjunction with HV/MV substation OLTC transformers, Step Voltage Regulators (SVR), and static Var compensator to control the voltage levels at strategic points of the network (Caldon et al. 2005; Doumbia and

Agbossou 2007; Senjyu et al. 2008). In Bokhari et al. (2015) the role of Conservative Voltage Reduction (CVR) and embedded generation in urban low voltage systems is investigated.

In Mahmud et al. (2014), an analytical approach to study how distributed generation units can affect voltage variations in the distribution systems is presented, defining the worst case scenario where the penetration level is at maximum allowed. Penetration level of distributed generation is defined as the ratio of power generated by embedded generation to the total power generated in a particular time loading (Bokhari et al. 2015). Mahmud et al. (2014) and Pompodakis et al. (2016) also argue that the reactive power from distributed generation sources can reduce network losses and voltage variations. On the other hand, insertion of generation from such sources at low voltage levels can cause high voltage levels at that part of the network, higher than other points where power is only withdrawn by customers (Pompodakis et al. 2016).

Tonkoski et al. (2008), Mahmud et al. (2014), and Pompodakis et al. (2016) argue that the capacity of a distribution network to integrate embedded generation effectively depends on a number of variables such as voltage variations on that system, particularly at the receiving end, the size and the distance of the distributed generation, and the loads on the system. This signifies the importance of smart meter data and geographical representation of the networks even further.

Sexauer and Mohagheghi (2013) highlight the fact that the uncertainty that is associated with the rising shares of embedded generation and low carbon technologies in the system effectively requires an operational change from deterministic approaches to probabilistic approaches, taking into account what ranges of voltage variations different sections of the distribution network are likely to experience under different operating conditions. Navarro-Espinoza and Ochoa (2016) argue that the DNOs can use new data to set voltage variation thresholds at various points of their network to flag and monitor potentially problematic areas.

ENA (2012) recommend that DNOs be provided with half-hourly consumption and generation information from smart meters as well as half-hourly voltage information at important points along the low voltage network, which can be obtained from half-hourly smart meters. This data can be used to ensure that the loadings on various sections of the low voltage network do not exceed the thermal ratings of the low voltage cables as well as monitoring the voltage variations that the low voltage network experiences.

Therefore, it is important to examine to what extent does real-time or predicted load readings from customers via smart meters can provide the DNOs with information about the voltage levels that is being experienced by the customers and to what extent variations in time resolutions, network arrangements, and aggregation patterns can influence the accuracy of such information. The reliance of smart grid applications on these estimates will be discussed further in the following chapters.

2.4.4 Low Voltage Cable Loading Percentage Estimation

Knowledge of capacity of low voltage underground cables to accommodate additional loads is becoming increasingly more important as Smart Grid management solutions such as load shifting and demand response management at community levels, are becoming more widespread (Miller 2015). The question is whether the real-time demand on each phase of the low voltage cables can be monitored using smart meter data and to what extent the information are improved or distorted as the granularity of smart meter data is decreased from data set to 120 minute averages.

This knowledge will then determine the headroom that is available on each phase which can potentially facilitate more proactive load management of the low voltage network by the DNOs and also optimal use of the distributed generation in the system. One of the main characteristics that is of particular interest to the DNOs is the peak demands at feeder level, which is the accumulation of loads withdrawn by customers on that feeder. This is mainly due to the fact that the higher the load peaks are, the higher the network losses and voltage drops experienced on the piece of low voltage network (Ijumba et al. 1999).

On the other hand, the DNOs need to design and manage the distribution network in a way that can cope with the Maximum Demand (MD) and demand diversity that will occur on the network (Strbac 2008; Barteczko-Hibbert 2015). Also, the peaks can be managed using embedded generation installation near the areas of the network that experience higher peak demands. As Ijumba et al. (1999) demonstrate, this will lead to loss reduction on the network. Additionally, with the increasing shares of distributed generation installed in the distribution network a need for more accurate knowledge of the network capacity is required (Strbac 2008).

The knowledge of the headroom available at peak times or off-peak times on a feeder can enable a more effective embedded generation and low carbon technology integration as well as a more efficient network asset management and investment strategies. Also, having knowledge of the network capacity on each phase of the various feeders leaving a low voltage substation can provide the network designers with useful information in terms of allocating new connections to various phases without the network being overloaded.

One of the main Smart Grid applications that can benefit from the knowledge of cable loading percentages is DRM. From the point of view of the DNOs, DRM involves measures that enable the shifting of peak loads to minimise network losses, voltage drops, and network strain, and maximise embedded generation integration (Infield et al. 2007; Strbac 2008). This will be discussed in more detail in section 2.5, which is about the distribution grid applications, but it is important to emphasise the role of customer loads obtained from smart meter data to estimate the peak loads that the various low voltage cable phases are likely to experience. This knowledge can provide the DNOs with invaluable information which has not been available to them before.

The following sections explain the ways in which the knowledge of customer loads, losses, voltage levels, and cable capacity usage at the low voltage level can greatly improve smart grid applications.

2.5 Major DNO Applications Impacted by Smart Meter Data

This section of the thesis introduces the major smart grid applications that are vital to the operation the distribution networks. Traditional and recent methods in conducting the DNO applications are introduced here as well as how these applications have coped with a lack of detailed low voltage data and the ways in which they can be improved with the availability of smart meter data.

2.5.1 Asset Management

In one sense, asset management is the core business of a DNO (Brint et al. 2008) with over a billion pounds being spent on it annually in the UK (OFGEM 2004). As Banyard and Bostock (1998) point out, distribution utilities such as those in the electricity, water and gas industries can be differentiated from many other industries by the high value of their asset base compared with their turnover. Hence ensuring that this asset base stays in a good operating condition is fundamental to a distribution utility's long-term future. The amount of money that the network operators in the UK spend every year on replacing their aging and underperforming assets constitute about 50% of the network status maintenance related investments, which can rise above tens of millions of pound per year (Black et al. 2009). Therefore, asset management has become an integral part of the utility sector. As Brint et al. (2008) point out, the rise in importance of asset management has been driven by the following three factors:

- the ageing of the distribution networks
- the advent of new technology
- the changes in the structure of the industry

Since the mid-20th century major changes in the electricity networks have taken place, assets have naturally aged and coupled with the customer load growth gave rise to a greater need for monitoring, maintaining and improving the network (Brint et al. 2008). Also, the changes in record keeping practices and the move from paper based network schematics to more computerised databases in the 1990s and even more recently GIS based records, have enabled more intricate asset condition monitoring practices (Brint et al. 2008). Additionally, the utility market changes as a result of privatisation in the UK and the fact that investment decisions and priorities are investigated by market regulators, also reinforced the major role of asset management in the DNOs business structure (Brint et al. 2008).

The term asset management started off in the financial sector, but began to be widely used by distribution utilities in the late 1980s and early 1990s (Brown and Humphrey 2005). However, asset management in the context of the utilities is much more complex due to the complex nature of the industry, performance related issues such as refurbishment and renewal of assets, and the intricate relationships in the utility networks (Brown and Humphrey 2005). A number of definitions of Asset Management when applied to distribution networks have been put forward with perhaps the clearest one being (Howard 2001):

> "Simply the way we look after the assets around us, both day to day in maintenance and operations, and medium to long-term in strategic asset planning." (Howard 2001)

In the context of Northern Powergrid, asset management has been defined as:

> "Ensuring that the distribution asset performs its required function safely, within the law and at a minimum lifetime cost." (Hammond and Jones 2000)

Or as it is defined by Sallam and Malik (2011) asset management:

> "aims to manage all distribution plant assets through their lifecycle to meet customer reliability, safety, and service needs." (Sallam and Malik 2011, p.9)

Monitoring the equipment installed on the network, replacing old or obsolete pieces, and investing in new technologies are the subtasks of the asset management team (Carer 2006). The tasks can be divided into short-term tasks such as maintaining the healthy status of the network, which is reliant on real-time and near real-time data, or more medium and long-term tasks such as maintenance or replacement tasks, which are more dependent on monthly, seasonal, and annual data (Tor and Shahidehpour 2006).

Historically, many assets have been managed on the basis of their "asset life". However, Clutterbuck et al. (2005) point out that a number of "asset lives" have been used, such as the manufacturer's recommended life, the financial life, the commercial life and the technical asset life. As the networks have aged and the pressure for efficiency improvements has increased, there has been a move away from using asset lives to using condition information or condition monitoring. Using only age to predict asset condition ignores the fact that condition is affected by manufacturing and installation quality, along with the asset's operating history and environment (Morton 1999).
As Black et al. (2009) emphasise the relationship between age and status and failure of utility network assets is very complex. OFWAT has commented that the prediction models are too often hampered by the poor quality of data that is input to them (Parsons 2006). Therefore, attention has concentrated on estimating conditions by periodically sampling the assets coupled with simple deterioration functions.

Hughes (2003) uses an exponential ageing curve, while Black et al. (2005) use a semi-Markov model. Brint and Black (2014) investigate the best way to handle the data when it is regarded as a sampling on two occasions problem. As Brown and Humphrey (2005) point out, a reliable information system should be at the heart of asset management procedures of utility network operators and these information sources should supply engineering and management decisions.

More accurate information about the way in which distribution network assets are used in terms of their capacity and loading percentages and also vital performance indicators such as network losses and voltage variations, can help DNOs take more proactive asset management approach as opposed to the current reactive approach (Brown and Humphrey 2005).

2.5.1.1 Use of Smart Data in Asset Management

Smart data offers the opportunity to model the asset's operating history and environment. The data from smart meters will allow much better estimates of current flows to be made and these can be combined with the knowledge of ambient temperatures (and where appropriate, ground conditions). This will allow much more accurate estimation of the stress that an asset has been experiencing, and so improve the estimates of the conditions of the assets that have not been sampled recently. As Sirto et al. (2015) point out, data from the customer end of the network can contribute greatly to better predict and subsequently manage the loads that will be experienced by various assets of the low voltage network, especially the transforms. Low voltage transformers are known to be the most expensive components of the distribution network and a more detailed knowledge of the operating capacities and conditions of these transformers can lead to a better preservation of them (Mohsenzadeh et al. 2016).

Two studies have been carried out in recent years to incorporate various types of customer meter data with the aim of improving the existing asset management approaches. Goyal (2016) use characteristic data of 10 sample 400kV transformers, along with minimum and maximum daily temperature data and annual 15 minute smart meter data from the customers connected to the transformers to calculate the electrical age of the transformers. The typical electrical age of a transformer of this size is usually about 180,000 hours or 21 years, but this can be affected by factors such as the number of hours during which the maximum capacity of the transformer is used and the temperature conditions. In Goyal (2016) these operational conditions are examined by having detailed smart meter information and temperature data from the nearest weather station. Mohsenzadeh et al. (2016) use similar types of data to Goyal (2016), but from 1000 kV transformers. They also incorporate demand response measures into their models to examine to the extent to which decreasing the load at peak times effects the longevity of the transformers.

As can be seen asset management to some extent is also very closely linked to network design and also network planning. Miller (2015) argues that smart data can help the DNOs reduce the amount of network component replacement and strengthening operations if they have enough smart information about demands on the network. For example, new research by CLNR shows that the domestic peak demand from households is around 0.9 kW, which is far lower than 1.5 to 2 kW figures that are traditionally used by the DNOs. However, with the introduction of LCTs into the low voltage network and the rising uptaking of such technologies in the future the demand diversity and the behavioural changes need to be understood in order for the asset management team to reinforce the capacity of the network where it is required to avoid both power quality and asset life deterioration. Additionally, Miller (2015) argues that with rising levels of embedded generation units in the low voltage network, knowing detailed voltage levels up to the customer end of the network is necessary, so that the potentially problematic areas can be identified and resolved by the asset management and network design teams.

2.5.2 Network Design

A key parameter in low voltage network design is accommodation of peak demand on the network (Brown 2008). It is uneconomic to design electricity networks so as to be able to supply every customer on the basis that they would all simultaneously have a Maximum Demand (MD) equal to the rating of their premises. For example, a domestic customer may have a rating of 75 amps, and there may be 100 customers on a feeder, but the feeder's thermal rating will be nowhere near 7,500 amps as it is assumed that people will not all require 75 amps at the same time. In fact, the cable rating is likely to be just a few hundred amps (Note that the situation is not quite as precarious as it may appear as most assets can run considerably over their thermal rating for short periods of time as it takes time for them to overheat and each phase is fused at the substation).

Let n be the number of customers on a feeder and D_n be the design demand from each customer, then the feeder is designed to meet a demand of

$$n \times D_n$$

Diversity between customers means that D_n is a strictly decreasing function of n. The After Diversity Maximum Demand (ADMD) approach that was developed by the UK electricity supply industry in the 1950s and 1960s formalised how D_n should be calculated.

However, the ADMD method is based around having a single MD figure for each customer. With the introduction of more complex tariffs such as Economy 7 in the 1960s, the Electricity Council's Load Research Unit produced tables of demand and the variability in the demand for the most important annual 60 half hours for each customer type. This in turn led to the development by the Electricity Council Research Centre (ECRC) of a design method based on using these half hours which was implemented in the program DEBUT (Carson and Cornfield 1973).

In essence, the approach assumes that the demand from each customer type for each half an hour can be modelled as a normal distribution. The independence of the demand from different customers is assumed. Hence the demand from a group of customers can be calculated by adding the means and the variances. A windows version, WinDebut, was created in 1997.

2.5.2.1 Smart Data and Network Design

Smart Data is likely to have a significant impact on low voltage and 11kV network design. Amongst other things:

- Feedback – The better knowledge of low voltage and 11kV power flows will allow the quality of the network (and so its design) to be assessed. Currently, only those networks that end up with a problem, e.g. voltage drops, provide any feedback.
- Knowing the existing power flows will allow better design of network extensions.
- General knowledge of typical power flows in low voltage and 11kV networks will also lead to an order of magnitude improvement in the assessment of alternative network designs such as how many voltage levels to have, the amount of interconnection that is optimal, and comparisons with the US pad mounted transformer designs. The comparisons that have been carried out (e.g. Brint et al. (1998); Porter and Strbac (2007)) have been severely hampered by a lack of detailed knowledge of the power flows.
- An understanding of how loads change as the network matures.

- Insights into how to modify designs to make them appropriate for embedded generation networks.

ENA (2012) points out that future low carbon technologies such as electric vehicles and combined heat pumps will introduce higher peak demands as well as unknown demand and generation diversity, consumption patterns, and more fluctuant voltage variations and/or reverse power flow due to the presence of embedded generation in the low voltage network, all of which will require greater knowledge of loading percentages on the low voltage network. For example, CLNR studies by Miller (2015) shows that demand of customers with low carbon technologies can be assumed to be twice as much as regular domestic customers.

ENA (2012) recommends that the DNOs should be provided with half-hourly consumption and generation information from smart meters as well as half-hourly voltage information at important points along the low voltage network, which can be obtained from half-hourly smart meters. This data can be used to ensure that the loadings on various sections of the low voltage network do not exceed the thermal ratings of the low voltage cables as well as monitoring the voltage variations that the low voltage network will experience to ensure that the statutory voltage limits of 230V +10%/-6% (216V and 253V) are met by the DNOs (Strbac et al. 2010; ENA 2012; DNV KEMA 2013; and Miller 2015). This is particularly important in the context of low voltage network design to ensure that the network is set up in a way that meets the demands of new connections as well as providing high quality service to the existing consumers.

Lees (2014) notes that the type of data that is required for network design is different from the type of data that is required for planning or monitoring. For planning purposes, longer intervals of data at lower resolution will suffice, but for design purposes it is ideal for the DNOs to have the most recent demand and voltage data at high resolution at critical points of specific types of low voltage networks (Lees 2014).

You et al. (2014) argue that smart meter data can provide the DNOs with the essential network planning and design parameters such as system losses, customer loads, and voltage levels. They point out that the availability of fine-grained smart meter data can shift the network planning methods toward “Stochastic” methods that incorporate uncertainties and variations in load demands, which are neglected in traditional methods carried out based on peak demands (You et al. 2014). Roupioz et al. (2013) have made some initial recommendations about using smart meter data in long term network planning of the networks and the studies presented in section 2.4.1 can also be used to further improve the loading profiles on the network using smart meter data. However, the use of such data in short-term network planning, expansion, and design has been neglected to a great extent (Nijhuis et al. 2017). Nijhuis et al. (2017) present a method for low voltage network planning and expansion that takes into account various sources of information, including smart meter data. They use a bottom-up approach of assigning smart meter load profiles to customers randomly in order to estimate typical low voltage feeder profiles and then compare planning methods based on this type of data compared to only using peak loads (Nijhuis et al. 2017).

There have also been a number of other studies using smart meter data in predicting demand and generation pattern on low voltage networks with embedded generation. These studies will be presented in section 2.5.4.1.

In theory, smart meter data can either provide the necessary information directly or yield the process of retrieving them. For example, the DNOs that have accurate smart meter information can run the required load flow analysis and models to determine what parts of the network are under stress and where is likely to experience over/under voltages. This information can also be used by the DNOs to proactively predict the parts of the network that are likely to be experiencing faults as a result of overloading and/or flag up the parts of the network that are under performing due to faults.

2.5.3 Fault Location and Restoration

Fault management is a very important part of a DNO's business remit as some of the key regulatory measures relate directly to faults, e.g. the number of customer minutes lost. Significant resources and investment are required to keep up with changes in the industry as well as dealing with the everyday task of maintaining a reliable supply to customers (Apel et al. 2001). The poor visibility that the DNO have of their low voltage network means that they are not usually aware of faults on the low voltage network until they receive notification from the customer end. The fault management team's task then consists of localising the fault, separating the affected area from the rest of the network, and restoring the power to the customers and all of these tasks need to be carried out as quickly and cost effectively as possible to benefit all stakeholders (Apel et al. 2001).

In addition to these tasks, incidents and interventions are required to be reported to the regulator (Verho et al. 2004). Problems on the network can occur due to changes in weather conditions and natural disasters or technical failures, in which case the affected part of the network is localised, isolated by tripping the system protection at circuit breakers or fuses (at low voltage level) (Verho et al. 2004; Estebsari et al. 2016).

Makinen et al. (2013) point out that the traditional ways of fault management in the low voltage networks need improving, especially in terms of collecting more detailed data on the incident location and the nature of the faults. The process has traditionally been based on trial and error until the protective relay tripped the feeder, but with the advent of numerical relays, which enabled reading the current at substations, finding the location of faults has gradually become less manual (Lehtonen et al. 2000). Relays work in milliseconds with two-minute response time in some cases such as first automatic switching and restoration. Historic information can be used in cases of voltage quality complaints as well as enabling the dispatchers to make decision about fault locations. Reliability of supply and ensuring an acceptable voltage quality standard are also the other main targets of the fault management team (Gono et al. 2007).

2.5.3.1 Fault Location at Low Voltage Level in the UK

The current system still relies on phone calls from customers to report the problems (Thomas et al. 2012). Call takers find the name and postcodes of the customers. The system stores call history, incidents, causes, ready messages to affected customers, and it also has the ability to update the messages as maintenance works progress. When customers report a fault, the incident is dealt with based on the three priority categories below:

P1: Urgent (e.g. smoke)

P2: Loss of supply

P3: Quality of supply complaints (e.g. flickering lights)

Once the phone calls are logged in, dispatchers send the appropriate team to deal with the issue by usually sending a “rapid” (a general technician) to carry out an initial assessment and then more specialised teams such as diggers, jointers, etc. Dispatchers use GIS, schematics and diagrams (see Figure 2-7) and use their own knowledge and experience in sending teams to attend faults and they use a system called “phone tracker” to make contacts with the teams.

The information about the nature and causes of the problem and works being carried out by the team are passed to the dispatchers on the phone and records are updated. Regulation requires faults to be addressed within the first three hours from the time that they are reported and in practice, the typical restoration time is between 45 minutes to 1 hour. In the near future, restoration teams will be equipped with smart tablets that will allow them to log in information, access job details, and update the status of the job easily. These interventions and steps are recorded in the Outage Management System (OMS) as job reports and then the Quality of Supply team will produce incident reports and pass them to the regulator, OFGEM. At low voltage level, 100% of incidents must be reported to OFGEM.

Currently, dispatchers deal with 20 calls per day on average. 20% of these calls are related to the quality of supply, which in theory can be predicted, diagnosed, and rectified with greater ease if high quality smart meter data are available to the DNOs. In addition to that, loss supply alerts, last gasp, and the restoration of the supply can be reported to the DNOs by smart meters. Presently, the network companies would not know if customers are back on supply unless the maintenance team informed the dispatchers, but smart meters are equipped to notify the DNOs when the customers are back on supply.

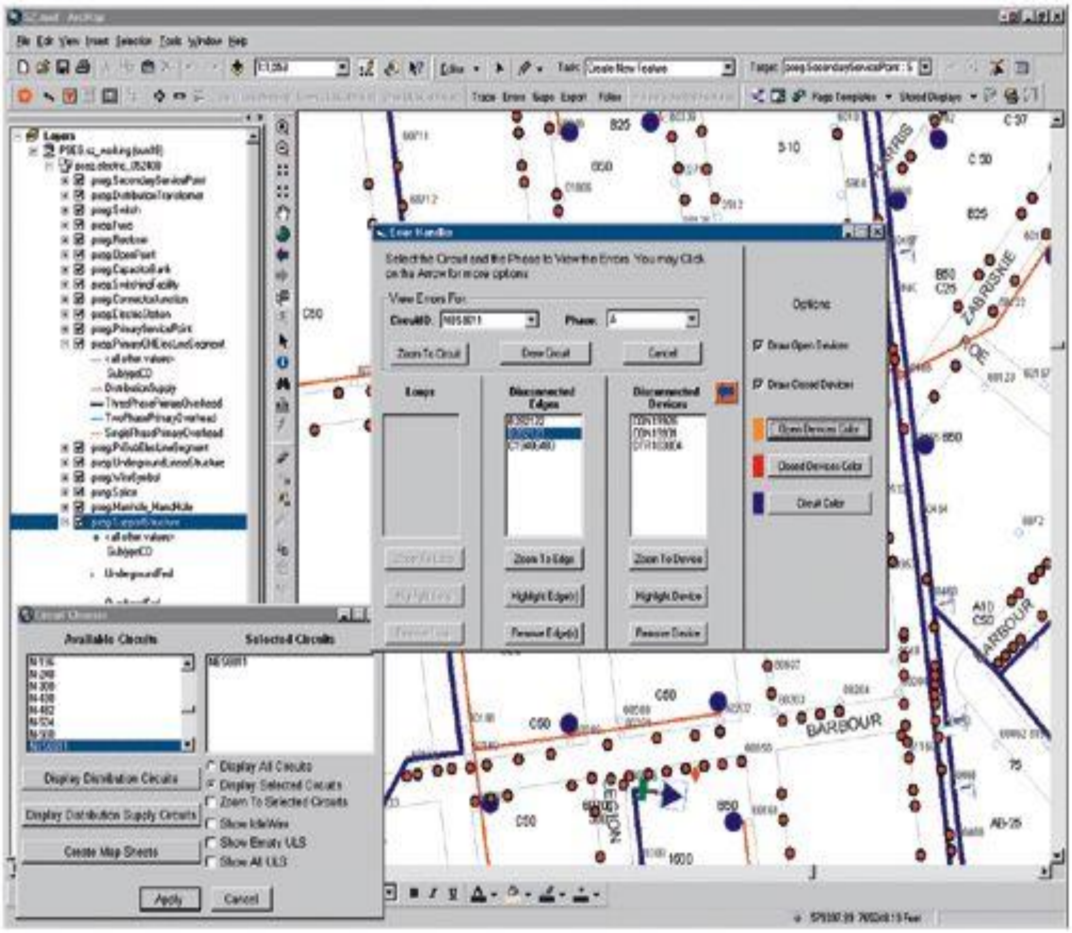


Figure 2-7: an example of GIS enabled OMS (Esri 2011)

2.5.3.2 Smart Data and Fault Management

The problem with the smart meter readings transmitted to the DNOs is that they are not always in real-time and one of the major aspects of fault management is fault location and this is about fixing a real-time problem. One specific capability of smart meters is the ability to send a (last gasp) signal to the DNOs when it loses power. Last gasp alerts are created when the supply of voltage in meters are lost (Estesbari et al. 2016). The DNOs will receive last gasp alerts from meters that lose their power, but there could be issues with this, for example if two separate faults within the same low voltage area occur, the DNOs may not recognise that these are two separate incidents.

An alternative would be for the DNOs to monitor the low voltage feeders leaving their 11kV substations. This could possibly alert them to a fault having occurred and then the smart meter data could be used to detect its likely location, given the impact on the power flow in the feeder that is affected. However, it is likely to be very expensive for the DNOs to equip their substations with separate network meters.

Theoretically if the DNOs have enough information about the various parts of the low voltage network that are overloaded for long periods of time or are not performing well in terms of showing unusual amounts of losses and voltage variations, they may be able to identify the potential problematic areas of the network where faults are more likely to occur. This aspect of the application ties into the information from the asset management team as well.

Historic customer-based and system-based data such as failure and outages of particular pieces of equipment, total and average interruptions experienced by customers, and average duration of interruptions can also help the fault management team in identifying the affected areas of the network (Gono et al. 2011). Liu et al. (2002) propose a model to integrate various data sources such as SCADA, customer calls, and network meter polling to predict faults on the network, while Zhao et al. (2013) present a fault prediction model based on simulated AMI load data and power flow data from distribution network sensors. Mäkinen et al. (2013) and Jiang et al. (2016) present models that combines smart meter trial data in Distribution Management System (DMS). This is to send the DNOs automatic fault alarms, which has now been resolved with the introduction of last gasp functionality on all smart meters. In Abusdal et al. (2015) smart meter voltage information are used to detect faulty conductors. In Jamali and Bahmanyar (2016) load demand are forecasted using partial smart meter data on various branches of a model distribution network in conjunction with network and feeder information to detect faults.

In recent years, DNOs have invested heavily in integrating IT and GIS into the fault management processes. These can help them locate faults on the network (using customer postcodes and meter registration numbers), monitor lines and equipment, dispatch personnel more quickly and effectively through an integrated system called OMS, hence shorten the outage time (Li et al., 2012). For example, Estesbari et al. (2016) introduces a fault detection method which uses smart meter data and network topology data from the GIS system to run power flow simulations at different nodes of particular low voltage networks and identify the nearest node to the affected area based on the changes between the measured and the calculated voltages on each phase of the network. The success of this method relies on the availability of good quality and complete smart meter data as well as great knowledge of the low voltage network topology and customer phases. Jamali et al. (2017) proposes a fault detection algorithm which incorporates current and voltage data from substations and partial field measurements from the low voltage network downstream of the substation. This method is less reliant on real-time household smart meter data in comparison to the method highlighted in Estesbari et al. (2016). On the other hand, it assumes the availability of high resolution data from low voltage substations meters which is very expensive for the DNOs to install.

2.5.4 Network Monitoring

The ability to better monitor the low voltage networks directly contributes to one of the most important targets of the Smart Grid agenda, improved observability and control. It also enhances a majority of DNO applications such as asset management, network design, power quality management and embedded generation integration, and active network management.

ENA (2012) argues that achieving the higher visibility of power flows and voltage behaviour on the low voltage side of the network via smart meters is the cheapest and the most efficient way. On the other hand, DNV KEMA (2013); Lees (2014); Miller (2015) argue for more monitoring points to be devised on the network in order to collect a higher granularity of data (e.g. 5 to 10 minute granularity or smart meters on low voltage feeders). They also raise this issue that various low voltage network applications require various granularities, latency and accuracy of data. For example, the data requirements for network control, network planning, and network design are different as mentioned earlier (DNV KEMA 2013; Lees 2014; Miller 2015).

Sanduleac et al. (2015) point out that the smart meter data are required to be highly accurate to be able to enable an effective monitoring of the low voltage networks. Barbato et al. (2017) argue that the evolution of the low voltage side of the distribution network to be able to facilitate and manage the integration of higher shares of embedded generation and low carbon technologies is highly reliant on the ability of the DNOs to monitor power losses, voltage levels, phase imbalances, and network losses.

2.5.4.1 Effective Integration of embedded generation

Salman (1996); Mcdonald et al. (2010); Chiandone et al. (2014); Jagtap and Khatod (2015b) reiterate the issue that the transition from a passive and rigid traditional electricity distribution grid with one one-way flow of energy and information to an active/smart distribution grid with two-way flow of energy and information is mainly driven by the need to accommodate higher proportions of embedded generation in the system. According to Salman (1996) and Hattam and Vukadinovic (2017) some of the main areas of the network operation and management such as network planning and design, asset management, and fault management can be affected by higher shares of distributed generation in the network due to the following factors:

- loss of supply due to faults
- harmonic distortions
- volatile reliability of the network
- voltage variations
- changes in network losses
- phase imbalance

Inevitably, the rising penetration level of low carbon technologies and embedded generation in the low voltage network will lead to voltage sag and swell (dip and rise) and harmonic distortion on the low voltage network and ENA (2012) argues that although smart meters cannot in theory provide power quality data, DNOs can utilise the smart meter information to pinpoint the problematic areas of the network.

Miller (2015) argues that in scenarios with a lower penetration level of electric vehicles and heat pumps, serious problems seem unlikely, but this can change as the penetration level of low carbon technologies rise in the future (Degefa et al. 2014; Zhou et al. 2014). To this end, a number of studies in recent years have focused on investigating the impact of electric vehicles and small scale distributed generators such as rooftop PVs on the low voltage network. In Neaimeh et al. (2015), the authors combine 1-minute resolution electric vehicle charging data and driving information with half-hourly smart meter data from the CLNR project to analyse the impact of electric vehicle charging on the low voltage network loads. Watson et al. (2016) use GIS and SCADA data to create a low voltage model of the network which can be populated with customer loads and PV data to run power flows of a realistic three phase low voltage network. They mainly focus on the impact of PV generation at various levels on the capacity of the low voltage network in New Zealand (Watson et al. 2016). Navaro-Espinoza and Ochoa (2016) also use probabilistic methods to look at the impact of PVs, electric vehicles, and heat pumps on the voltage levels of a realistic low voltage network. The smart meter profiles used in this study were simulated at 1-minute time resolutions, but were then adopted to create 5-minute demand profiles (Navaro-Espinoza and Ochoa 2016). Phoghosyan et. al (2016) use agent based bottom up simulations of electrical vehicle loads in accordance with the UK Government's prediction of various levels of electric vehicle uptakes in the next 10 years in order to study their impact on the peak loads in the low voltage networks. They use half-hourly smart meter data and daily electrical vehicle charging patterns (Phoghosyan et. al 2016). A similar approach has also been used in Hattam and Greethard (2015).

2.5.4.2 Active network Management (ANM)

Active Network Management (ANM) comprises of a set of measures to monitor and manage the capacity of the network in accommodating more low carbon technologies and embedded generation in the system without the low voltage cables being overloaded or strained or the quality of supply being affected (ENA 2012 and Miller 2015). Mcdonald (2008) points out the importance of the term "active" in the application's name, contrasting it with the current passive design of the traditional medium and low voltage electricity networks. ANM aims to optimise the network capacity to meet the new demands instead of using more expensive and less environmentally friendly methods of adding new circuits to expand the networks (Mcdonald 2008; Zhou et al. 2014).

The main purposes of ANM are to integrate embedded generation in the distribution grid more effectively and provide flexible measures for the DNOs to accommodate new patters of demand from consumers with electric vehicles and other forms of low carbon technologies (Mcdonald 2008; ENA 2012; Miller (2015).

According to Strbac (2010); DNV KEMA (2013); Lees (2014); and Miller (2015) ANM could include:

- active voltage monitoring of the low voltage network using half-hourly voltage information from strategic points on the network.
- control of the power factor.
- flexible changing of the low voltage open points.
- warning systems in the cases of peak demands and overloading of the low voltage cables.

- better investment decisions resulting from power flow and voltage information, estimated about 5% or £3.5m reduction in annual network investments (DNV KEMA 2013 and Miller 2015).
- using demand response schemes to manage the loading capacity of low voltage cables
- active voltage control measure.

Miller (2015) argues that these measures can be achieved without the DNOs having to invest more money in the communication infrastructure of the low voltage network as the smart metering implementation programme is already on the way to deliver new voltage and power flow data on the low voltage network. ANM is particularly important in managing network loading capacity, voltage levels and power flows and many DNOs are integrating the ANM measures to some extent as more renewables are installed in their existing network (EA Technology 2016). For example, better voltage level management can allow a DNO to install more distributed generation on the low voltage network while maintaining the voltage levels in the statutory limit ranges of 230V +10% -6%.

Zhou et al. (2014) carried out a study integrating various electric vehicle charging data into the distribution management system in order to enhance the network visibility and observation of the DNOs. Degefa et al. (2014) focus on the relationship between thermal stress and stochastic generation patterns from distributed generators on a test low voltage network and examine the role Real Time Thermal Rating (RTTR) in managing the thermal stress and increasing the loading capacity of conductors in the network. Also, studies such as Repo et al (2013) and Paterakis et al. (2016) have focused on market incentives in actively managing a group of customers with low carbon technologies and embedded generation in smart neighbourhood scenarios.

EA Technology (2016) argues that the requirement for more in-depth behavioural knowledge of the low voltage networks is becoming ever more pressing to the DNOs as customers' demand and generation are becoming more varied and intermittent. However, as it was highlighted earlier critical network information that are required for ANM such as voltage levels will be somewhat distant from the reality of the situation as smart meter data are aggregated and averaged over half-hourly intervals.

2.6 Summary

Smart grid applications are heavily reliant on key low voltage network information parameters such as power flows, network losses, voltages, and cable loading percentages. It is likely that their reliance level will increase further as smart grids evolve and more embedded generation and low carbon technologies are connected to the low voltage levels of the network.

Traditionally, it has not been easy for the DNOs to obtain accurate estimations of these key network performance indicators as the information on the low voltage side of the network has been limited. In the absence of data on the low voltage networks, different methods were devised to fill this information gap, including these described in this chapter. However, smart meters are regarded as the game changer which can either improve some of these methods that are still in practice, or completely change the way in which the key low voltage network information are obtained by the DNOs. In theory, smart meter data can provide the DNOs with high resolution data which in turn will lead to more accurate estimations of power flows, losses, voltage levels, phasing imbalance, and customer loads. This can lead to the enhancement of DNO applications such as asset management, network planning and design, fault location and restoration, and network monitoring. These applications were traditionally carried out in the absence of detailed information on the low voltage side of the electricity distribution network. However, this will change with the availability of smart meter data. Accurate information on the low voltage network can help the DNOs manage embedded generation and customer loads consumed and/or generated through programmes such as ANM and DSM, as well as managing areas of the network that will encounter reverse power flows.

Operation and maintenance of low voltage networks, which in the past did not require close monitoring, will also benefit from more accurate information resulting from smart meter data in coping with the introduction of new loads related to electric vehicles, heat pumps, and micro generators. This is of a huge significance when it comes to balancing the voltage on the network within the statutory requirements and maintaining the voltage quality delivered to customers, or planning for new connections (Smart Grid Forum 2014).

Based on the literature presented above, the extent to which DNO applications are and will be reliant on high resolution smart meter data is presented in Table 2-3 below.

Table 2-3: Reliance level of major low voltage network applications on high resolution smart meter data

Applications	**Accuracy of Loss Estimates**	**Accuracy of Voltage Estimates**	**Accuracy of Cable Loading Estimates**
Asset Management	High	High	High
Power Quality Management and Integration of DG	High	High	High
Active Network Management	High	High	High
Network Design and Planning	Medium	High	High
Network Monitoring	Medium	High	High
Fault Management	Low	Medium	Low

The question still remains how much in reality smart meter data can contribute to the accuracy of key low voltage network information required by DNOs to run their smart grid applications, especially if the data are of low resolution, low frequency, or are aggregated. This is the main question this thesis is aiming to answer. This a question that is particularly important in the UK as the potential benefits that the DNOs can gain from the smart meter data can be affected by factors such as time resolution intervals and aggregation of smart meter data and the gradual implementation of smart meters. However, the answer to this question can apply to other countries that have set similar smart meter data specifications.

The next chapter presents the methods and models used in conducting this study and the ways in which they compare to the relevant research that has been carried out in recent years in the field.

Chapter 3 Methods

This chapter describes the methods and the models that are used in the light of the aims and objectives identified in chapter 1 in more detail. This is carried out by explaining the main differences between the methods employed in this work and similar studies. In the first place, the main data sets used in the various aspects of this study are described and their strengths and weaknesses are summarised. Secondly, the methods and the ways in which the data sets are used in each study model are described.

3.1 Data Sources

The data sets used in this research were:

1. Loughborough University trial data

> Obtained from a study carried out by Loughborough University in 2008 and 2009. This trial was sponsored by E.ON UK and the Engineering and Physical Sciences Research Council (EPSRC) (Richardson and Thomson 2010). The data sets are free to download and publicly available from Richardson and Thompson (2010).

2. CLNR customer data:

> Collected by British Gas during the Customer-Led Network Revolution (CLNR) project from 2011-2014 (Customer-Led Network Revolution 2017). The collection of data from customers involved in these two trials were ethically approved and the data were obtained, anonymised, stored, and released to the interested researcher in accordance with the Data Protection legislation. The data sets are free to download and publicly available from CLNR Project Library (2017).

The use of these data sets for the purpose of this research has also been approved by the data providers and the research project itself has been ethically approved by the University of Sheffield. The specific ways in which the Loughborough and CLNR data sets were used will be discussed in more depth in the following sections of this chapter.

3.1.1 Loughborough 2008-2009 Data Sets

This data set contains 1-minute customer consumption readings in kiloWatts (kW). These consumption readings were collected from 22 domestic smart meters from 2008 to 2009. Along with the customer electricity consumption loads in kW, a set of questionnaires were also filled out by the participants that indicates the type of the houses that the smart meters are installed in (e.g. detached, semi-detached, etc.) as well as the characteristics of the houses, for example the appliances in use (Richardson and Thomson 2010).

A summary of properties of the data in this data set can be found in Table 3-1 below.

Table 3-1: Summary of the characteristics of the Loughborough data set

File Number	**1**	**2**
Date of collection	2008	2009
Number of houses	22	20
Consumption data	Yes-kW	Yes-kW
Time resolution	1 Minute	1 Minute
House type information	Yes	Yes
Number of occupants	Yes	Yes
Geographical information	Approximate	Approximate

The data from the 22 meters are stored in two separate Excel files for each house in 2008 and 2009, accompanied by a separate excel spreadsheet containing information about the house types and the appliances available in each house.

In the case of a majority of the 22 meters, the consumption data in both 2008 and 2009 worksheets are clean and easily accessible, containing consumer loads in kW from January 2008 to December 2009. A perfect 1 minute time resolution data set should contain 525,600 recorded data points for a whole year for each meter. However, the data points recorded for the 22 meter in 2008 range from just above 260,000 to just under 525,600 (1 data point for each minute). This shows that in the case of some meters, some data points have not been recorded and are missing.

Meters 2 and 16 do not have any customer load records for 2009, which makes them redundant for some of the analysis that was the subject of this research, such as current estimation studies in section 3.3. On the other hand, the number of data points recorded for the remaining 20 meters in 2009 shows a greater range varying from minimum data points at just under 70,000 to maximum data points at just above 525,000, which is indicative of the inconsistency of data points stored in the data recorded from the houses in 2009. Regardless of these irregularities, these data sets provide one of the most complete smart meter data sets available for research purposes at the moment.

In relation to the research objectives of this project, the limited number of households (22 smart meters) and the inconsistency between the 2008 data and 2009 data were found to be particularly challenging for studies of current estimation in section 3.3, due to the lack of sufficient numbers of meters with at least 13 months of data. Ideally, at least 50 meters would be required to populate a two branch low voltage network with 25 houses on each branch. However, the good quality of the data set resolution smart meter data available in these data sets provided an appropriate platform for carrying out the analysis regarding the effects of smart meter time resolution and aggregation variations in estimation of electricity network loss estimation, voltage drops, and cable loading percentages. This data set is referred to as “Loughborough data set” throughout this thesis. The data sets are free to download and publicly available from Richardson and Thompson (2010).

3.1.2 CLNR Data Sets

Customer-Led Network Revolution was launched in 2011 and completed in 2014. This project brought together academic and industrial players in the field of electricity network operation and supply, such as OFGEM, LCNF, Northern Powergrid, British Gas, EA Technology, the University of Durham, the Newcastle University, and 13,000 customers to investigate the major challenges that the UK electricity network will encounter as the transition to a smarter and greener grid operation takes place (Customer-Led Network Revolution, 2016).

The customers that were part of the CLNR project were equipped with smart metering systems and some groups of these customers were also provided with emerging low carbon generation and/or consumption technologies (e.g. solar PVs, electric vehicles, and heat pumps) (Customer-Led Network Revolution 2016). Therefore, the data sets produced by this project fall into the 9 different categories of data below, based on the categories of the customers (Bird, 2015):

1. Basic profiling of domestic smart meter customers

2. Basic profiling of small and medium sized enterprise (SME) customers

3. Domestic smart meter customers on time of use tariffs

4. Domestic solar PV customers using In-Home Displays (IHDs) for manual in-premises balancing

5. Domestic solar PV customers with automatic in-premises balancing for hot water charging

6. Enhanced profiling of domestic customers with air source heat pumps

7. Enhanced profiling of domestic customers with Electric Vehicles (EVs)

8. Enhanced profiling of domestic customers with solar photovoltaics (PV)

9. Enhanced profiling of domestic smart meter customers

These data sets were released in December 2015 for the use of the parties that were not involved in the CLNR project, but the data sets for each category were stored in a format that was not easy to access using conventional Microsoft Office software programmes such as Excel or Access, therefore they were divided into smaller chunks using Python 2.7.2 (see Appendix A).

There are a number of reasons for selecting data sets no.1 and no.8 from the nine categories of smart meter data available from the CLNR trials. Data set no.1 contains the most number of domestic smart meter consumption data at half-hourly time resolution, which will be the resolution at which such data will be available to the DNOs. Although this data set does not contain the data points for all the dates between 2011 and 2013, the fact that the data recorded for matching meter IDs are present improves the chances of developing a more accurate low voltage network load prediction studies.

Data set no.8 provides data set smart meter data from 150 meters. This is the largest data set with time resolution higher than 30 minutes within the CLNR data sets. This data set is ideal for studying the effects of smart meter time resolution variations on the accuracy of loss estimates, voltage levels, and cable loading percentage estimation.

Although this data set also contains some data gaps and the data points for some dates are missing, there are sufficient dates between 2012 to 2014 on which smart meter data from at least 100 meters is recorded.

The other CLNR data sets listed earlier were not deemed to be suitable for the smart meter time resolution studies or aggregation studies, because either the smart meter data stored were recorder at half-hourly intervals in the first place or there were less than 100 meter IDs with data set data recorded.

The selected CLNR data sets are referred to as CLNR data set no.1 and no.8 throughout this thesis. The properties of these two CLNR data sets that are used in this research are presented in Table 3-2 below.

Table 3-2: Summary of the characteristics of the CLNR data sets

File name and number	**1. Basic Profiling of domestic smart meters (data set no.1)**	**8. Enhanced profiling of domestic customers with PVs (data set no.8)**
Date of collection	2011-2013	2012 to 2014
Number of houses	5000	150
Consumption data	Yes-kW	Yes-kW
Time resolution	30 Minute	1 Minute
House type information	No	No
Number of occupants	No	No
Geographical information	Approximate	Approximate

It should be noted that the CLNR data sets do not contain any demographical information and the approximate geographical information about the trial locations made it possible to obtain the maximum daily temperature data from the nearest weather station. The Loughborough data set also included approximate geographical information that made it possible to acquire the maximum daily temperature data from the nearest weather station. The Loughborough data set also includes information about the house types and the appliances in each house. However, this type of data is not usually transmitted to the DNOs due to privacy reasons and are therefore not accounted for in the methods used by the DNOs.

In summary, the current estimation studies in chapter 4 are carried out using both the Loughborough data sets and the CLNR data set no.1 and the time resolution and aggregation studies use both the Loughborough data sets and the CLNR data set no.8. Figure 3-1 shows the data sets used in each study.

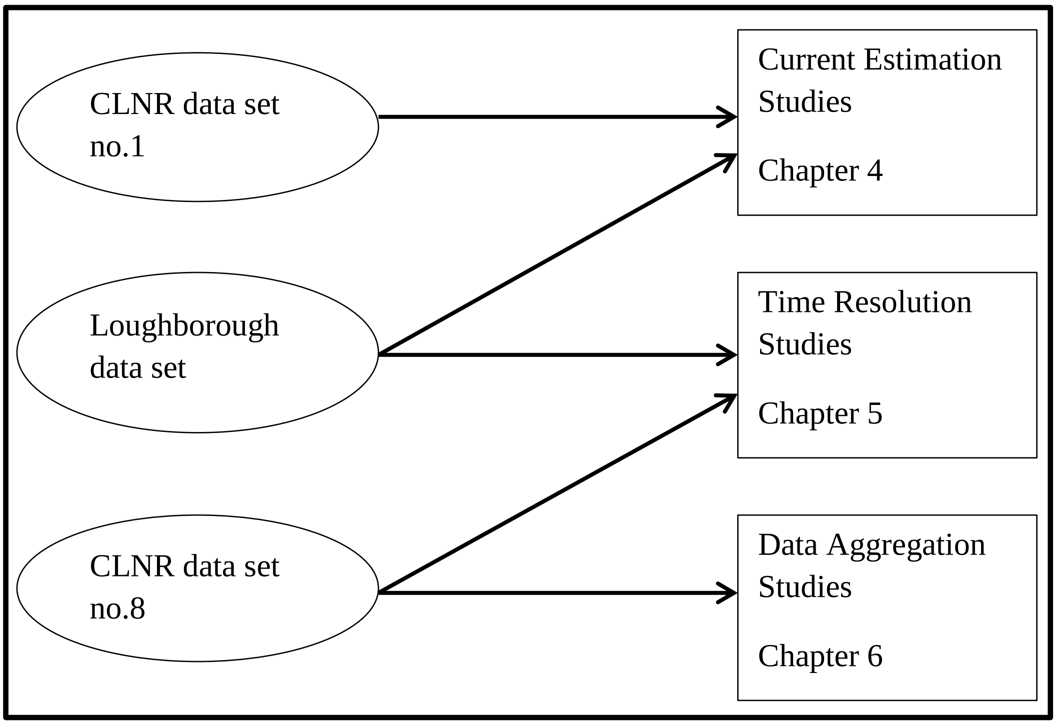


Figure 3-1: The data sources used in the three main studies in this thesis

The ways in which the network models are populated using these data sets are discussed in more detail in the following sections of this chapter.

3.2 Low Voltage Current Estimation Using Historical Smart Meter Data

Most load prediction and short term load forecasting methods have been developed at higher levels of the electricity networks (Hayes et al. 2015; Hong and Fan 2016; Valgaev et al. 2016) (see section 2.4.1). There has been a growing need for adopting these methods at lower voltage levels of the network due to the increasing integration of embedded generation. However, the adaptation of the methods to the low voltage side has proven to be challenging (Hayes et al. 2015; Hong and Fan 2016; Valgaev et al. 2016). This is mainly due to the fact that as the currents are disaggregated from higher levels to the lower levels of the network, the load profiles become noisier and more diverse (Hayes et al. 2015; Hong and Fan 2016; Valgaev et al. 2016). For example, load shapes of aggregated loads from customers at 11 kV substations show more homogenous trends compared to those obtained at low voltage substations, and subsequently at individual customer levels.

From the point of view of the DNOs, the importance of estimating the loads of individual customers is to be able to carry out realistic load flows of the low voltage networks and use such analysis in the network applications. Since the smart meter implementation in the UK is a gradual process and the communication infrastructure can be prone to lags and faults, it is important to be able to estimate the missing smart meter data from the available data on the network. Also, it is highly unlikely that the DNOs will have any information about individuals' lifestyle patterns, due to privacy reasons. Therefore, it is important to devise load estimation methods that are realistic in the context of the UK distribution network operation.

The following sections describe the methods used in this thesis to calculate missing real time customers loads on the low voltage network using historical smart meter data from the neighbouring customers and the historical smart meter data.

3.2.1 Data Used

In order to carry out the analysis, a 20-house model (Model A) and a 50-house model (Model B) were created and populated with half-hourly smart meter consumption data. Model A-1 was populated with data from the Loughborough data sets and model A-2 was populated with the CLNR data set no.1 data. Since the Loughborough data sets only contained 13 months data for 20 meters, model B was only populated with the CLNR data set no.1 data.

Consequently, two sample dates in each data set were selected, for which there were sufficient number of smart meters with historical smart meter data (13 months). Figure 3-2 below shows a representation of Model A, comprising households with 20 smart meters.

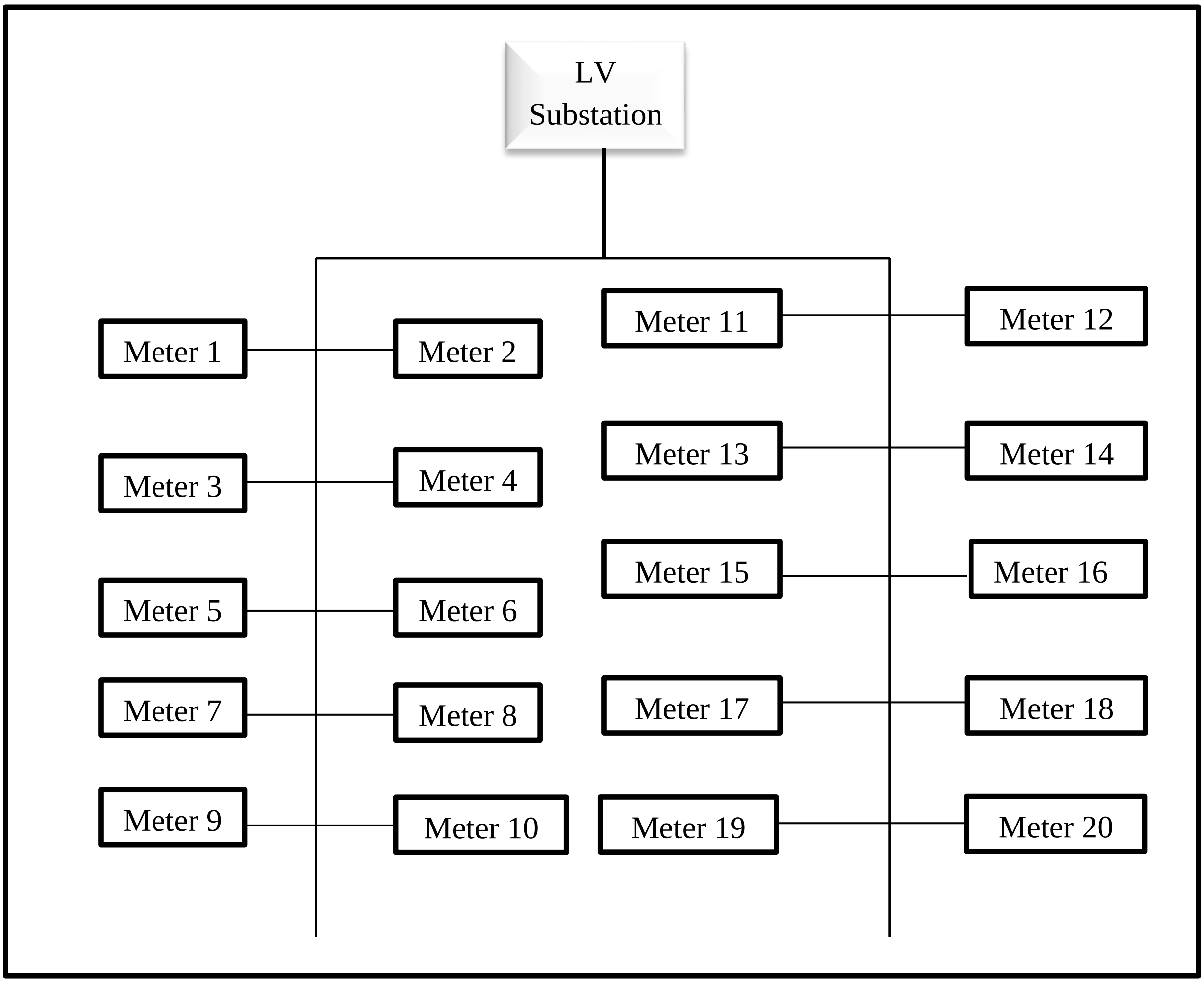


Figure 3-2: A schematic representation of Model A

Additionally, since the region from which the CLNR data were obtained was approximately known, maximum daily temperature recorded at one the nearest weather stations to North Yorkshire was also obtained (Weatheronline 2016) to be used alongside the CNLR data to examine whether such data can improve the estimation accuracy of the methods. More granular temperature data such as half-hourly temperature data was not available.

3.2.2 Load Estimation Methods

Load estimation approaches can be divided into four major categories of:

- Traditional methods such as After Diversity Maximum Demand (McQueen 2004) and customer load curves (Kersting and Philips 2008) or Standardised Load Profiles (SLPs) (Valgaev et al. 2016).

- Methods using substation readings such in Kersting and Philips (2008) and Arritt et al. (2012) and a combination of substation data and smart meter data as in Mirowski et al. (2014). These approaches aim at predicting the peak loads, so methods used in this thesis have been devised to predict the load shapes of the customers for 24 hours using approaches similar to using billing data (methods 1 to 4 below)
- Methods to predict loads of a particular building type by using machine learning techniques such as Artificial Neural Network (ANN) such as in Wong et al. (2010) and Chitsaz et al. (2015). These methods are not generally used by the DNOs are at early stages of development at low voltage levels.
- Methods using time series analysis and k-nearest neighbours such as in Iwafune et al. (2014) and Valgaev et al. (2016). Valgaev et al. (2016) use a combination of k-nearest neighbours and functional time series and a relatively recent work by Chaouch (2014) use functional time series to predict the load shape of individual customers. Our approach however, combines substation data and smart meter data and the predictions are point based in that each half-hour from historical data is used to estimate the missing half-hour load on the sample date, therefore k-nearest weighted average of the nearest neighbours was selected as the statistical approach of choice.

Hayes et al. (2015) argue, simpler statistical methods are more effective at the lower voltage level of the network as the customer loads are disaggregated. Considering that the DNOs in the UK are not provided with customers' lifestyle information due to privacy issues and will not have widespread low voltage substation metering infrastructure in place due to the costs involved, the following load estimation methods were tested on models A-1 and A-2 and the most accurate methods were then applied to model B with a larger sample size:

1. Prediction of missing customer loads on the date based on real-time smart meter data from the neighbouring meters on the sample date and the historical data from a week earlier on a similar day.

2. Prediction of missing customer loads on the date based on real-time smart meter data from the neighbouring meters on the sample date and the historical data

from a month earlier on a similar day.

3. Prediction of missing customer loads on the sample date based on real-time smart meter data from the neighbouring meters on the sample date and of average of historical data from four weeks before on a similar day.

4. Prediction of missing customer loads on the sample date based on real-time smart meter data from the neighbouring meters on the sample date and the historical data from a similar type day from a year before on a similar day.

5. Prediction of missing customer loads at peak times on the sample date based on real time smart meter data from the neighbouring meters on the sample date and the k-nearest weighted averages of the closest historical data values from the neighbouring smart meters. A recent work by Valgaev et al. (2016) also use a similar approach relying only on k-nearest smart meter values without the use of substation data. Also, our approach also incorporates the maximum daily temperature in calculating the distance between the nearby points.

Central to methods 1 to 4 is the concept that although an individual consumer's energy usage pattern is volatile, a customer's energy consumption pattern on a specific day is likely to be similar to be to their load pattern on the same day type a week prior to the sample date or weeks further back. Figures 3-3 and 3-4 show the load shape for 6 customers from the Loughborough data set (model A-1) and CLNR data set no.1 (model A-2) on the two sample dates, respectively. These figures demonstrate the difference between the consumption patterns of individual customers.

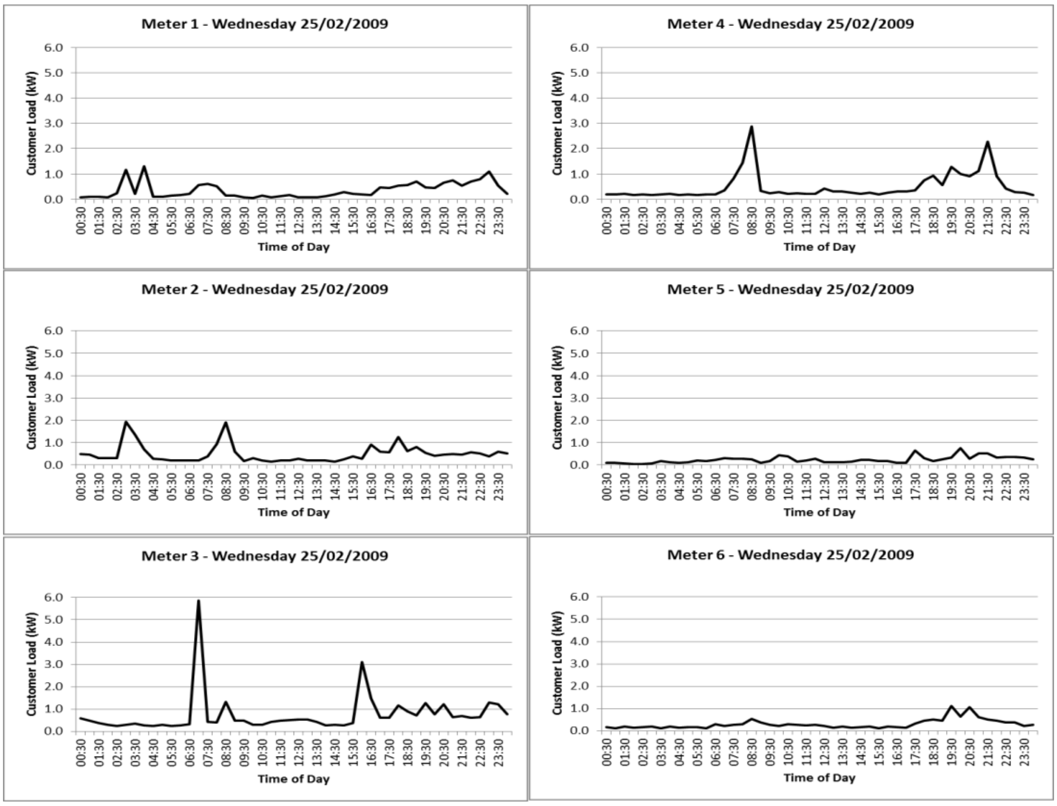


Figure 3-3: Load shapes of 6 different customers on the sample date (Loughborough data set)

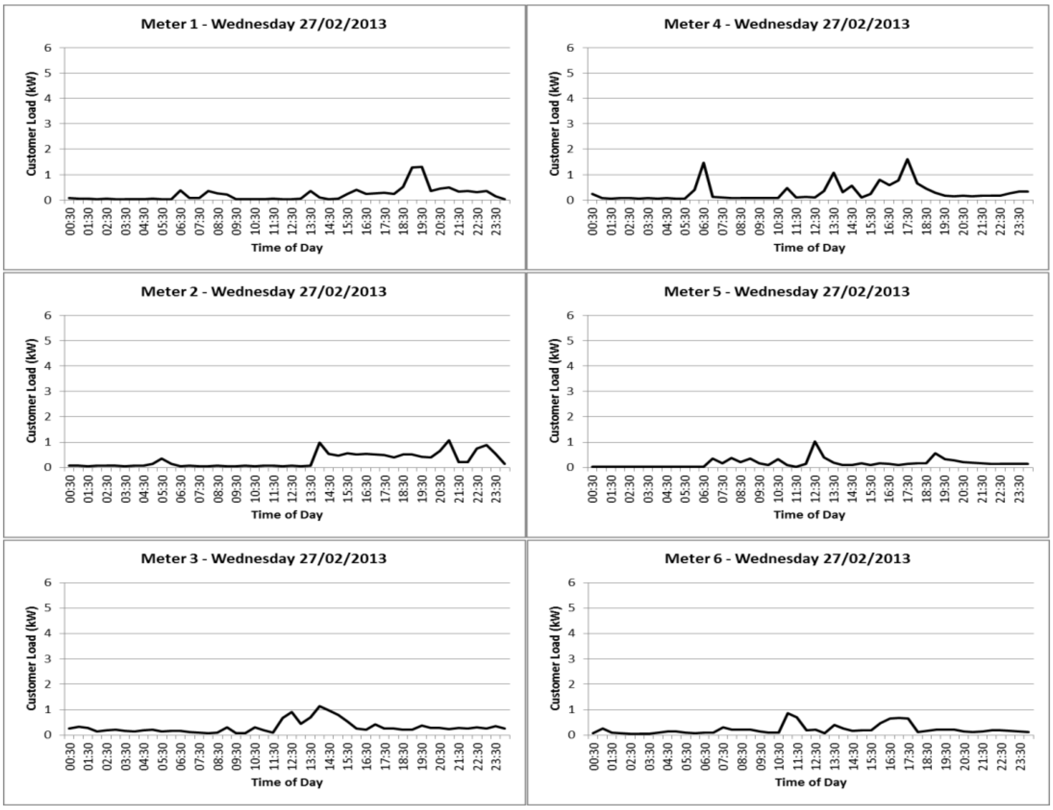


Figure 3-4: Load shapes of 6 different customers on the sample date (CLNR data set no.1)

Figures 3-5 and 3-6 show the load shape for meter ID 1 from the Loughborough data set and CLNR data set no.1, respectively.

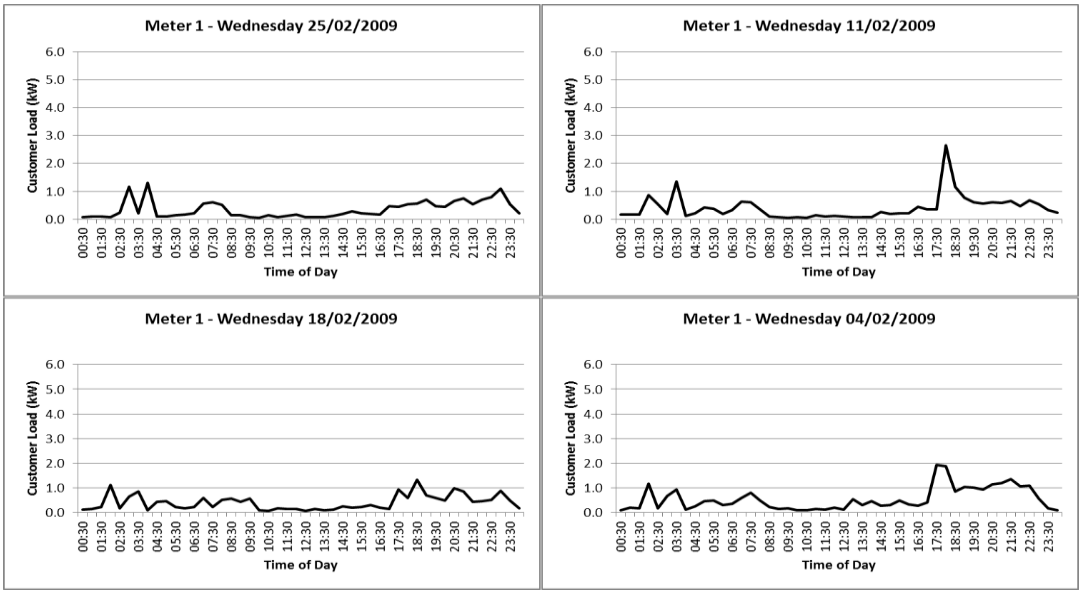


Figure 3-5: Load shapes of meter 1 on the sample date and the historical similar days (Loughborough data set)

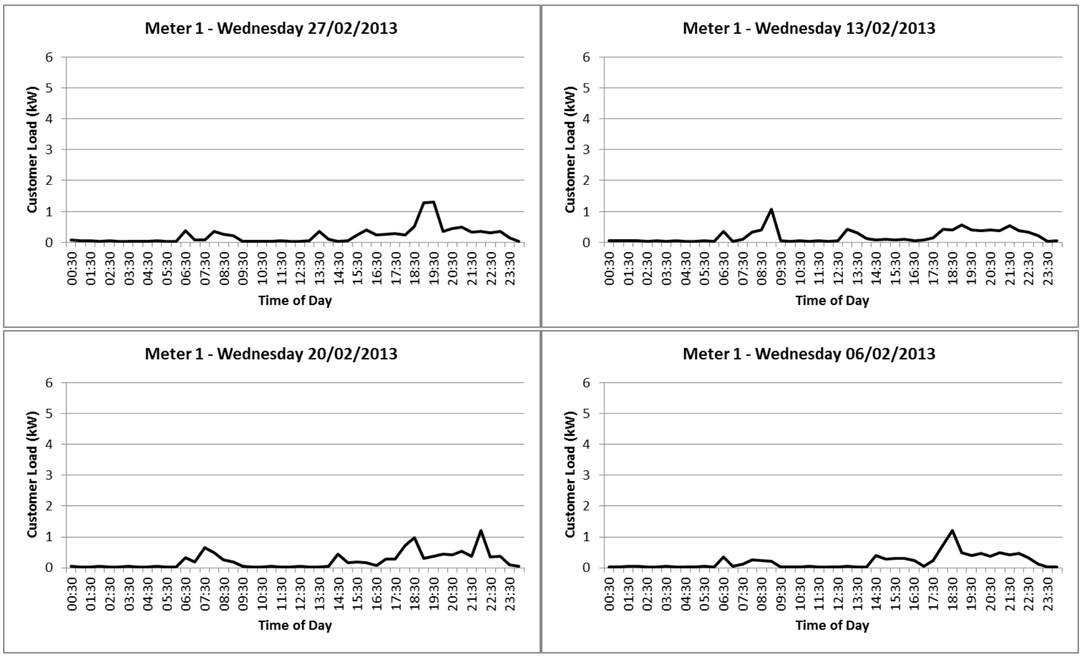


Figure 3-6: Load shapes of meter 1 on the sample date and the historical similar days (CLNR data set no.1)

Figures 3-5 and 3-6 show that there is a higher degree of similarity between a customer's consumption patterns from week to week compared to month to month. On the other hand, Figures 3-3 and 3-4 show that different customers have very different consumption patterns, especially during peak times.

To examine this further, methods 1 to 4 are used in order to determine the best input of historical smart meters to be combined with substation meter readings. The difference between the estimated half-hourly values and the measured half-hourly values on the sample dates (which are initially deleted) is calculated using Absolute Percentage Error (APE) for each half-hourly estimation. Subsequently Mean Absolute Percentage Error (MAPE) is calculated to determine the average estimation error using each method. APE for each half hour can be calculates using the equation below:

$$APE = 100 \times \left| Recorded\ load - Predicted\ load \,/\, Recorded\ load \right|$$

The best two performing methods are then applied to model B that contains data for 50 customers. The results of these analysis are presented in chapter 4 and the findings will be discussed in detail.

3.3 Impact of Smart Meter Time Resolutions on Estimation Accuracy of Losses, Voltage Levels, and Low Voltage Cable Currents

This section describes the methods that are employed in this work to model a typical low voltage network model in order to study the ways in which the accuracy of vital low voltage performance indicators such as technical losses, voltage levels, and cable loading percentages are affected as the time resolution of smart meter data is decreased from data set intervals to 120 minute averages.

3.3.1 Time Resolution and Customer Data

To date, only a limited number of studies have focused on the effects time resolution of data on the information that can be acquired from customers. Wright and Firth (2006) argue that while having half-hourly interval billing data from customers are sufficient for suppliers and energy traders in the UK, this time resolution of data hides critical fluctuations in customer demands. The examine the impacts of having 1 minute data resolution from 2 houses on the load shapes and compare them to having half-hourly data from the 2 dwellings (Wright and Firth 2006). They show that when the data are averaged over a half an hour, the demand from customers are underestimated and sudden spikes in customer demands (mainly caused by switching on high power devices) are flattened or are ignored as shown in Figure 3-7 (Wright and Firth 2006).

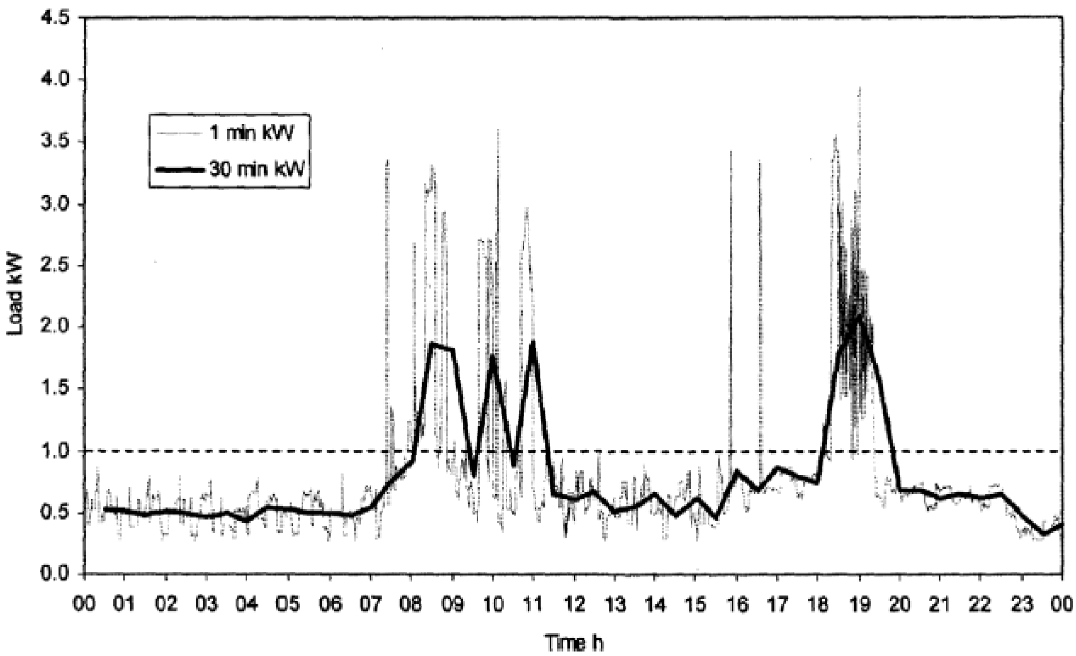


Figure 3-7: Customer loads at 1 minute and half-hourly time resolution (Wright and Firth 2006)

This issue has also been investigated in Richardson et al. (2010); Brandauer et al. (2013); McKenna and Thompson (2015). Richardson et al. (2010) examines 1 minute data from 22 houses to investigate the effects of customer behaviour and their use of appliances on the demand curves from customers. Brandauer et al. (2013) investigate the impact of having data at 15 minute time resolution intervals compared to 1 second time resolutions on low voltage networks. McKenna and Thompson (2015) use high resolution smart meter data to investigate the effects of thermal-electrical usage from households on load demands.

In terms of the relationship between the estimation accuracy of losses and the time resolution of smart meter data, Urquhart and Thompson (2015) focus on the relationship between smart meter time resolution from1 to 30 minute intervals and loss estimates. They investigate the effect of smart meter time resolution on the estimation of losses in a 22 house single phase network model. The weakness of this work lies in that the losses are calculated for each house without considering the neutral phase currents and the effects of the loads from other customers on the various phases on the calculation of overall low voltage network losses. However, our study focuses on the effects of time resolution on the estimation accuracy of three critical low voltage network information areas of losses, minimum voltage levels, and cable capacity percentages in a balanced and an unbalanced three phase low voltage network model with 100 customers. The importance of using the three phase model is that most low voltage networks in the urban areas in the modern economies, especially in the UK use three phase feeders with earth return. Also, the analysis in the unbalanced network model replicates the situation that the DNOs are likely to find in some of their low voltage networks. Additionally, our model uses different sample dates from 2 different data sets and uses them in the low voltage network model. The losses estimated on the various dates are also used in predicting the 1 minute losses based on estimated losses at half-hourly averages. The data used in our analysis is described in the next section and the way in which the critical low voltage network indicators were estimated at various time resolutions are explained in section 3.3.3.

3.3.2 Data Used

Both the Loughborough and the CLNR data sets were utilised for these studies which will be discussed in detail in chapters 5 and 6. As one of the most important parts of this research is studying the relationship between varying time granularity of the smart meter data and the accuracy of major low voltage estimations, the data available in the Loughborough data set and the CLNR data set no.8, which are 1 minute data, were selected as the most suitable options. The consumption readings from these two data sets were then used to populate a 100-house low voltage network.

Since the Loughborough data sets only contained consumption information of 22 households, the remaining 78 household consumption data profiles were required to be sampled using the closest dates to the chosen sample dates. The reason behind choosing the neighbouring dates was to maintain the daily consumption pattern of the households as much as possible.

The CLNR data set no.8 contained reading from 150 meter IDs, but there were gaps in recorded data on some dates. Therefore, a query was run in Excel using Visual Basic to determine the dates on which data from at least 100 meters were available. The 52 dates on which data from at least 100 customers were recorded were chosen as the pool of sample dates. Initially, 4 sample dates were selected from each data set were selected. In the next step and to carry out 1 minute loss predictions based on half-hourly estimates in section 5.3, 48 more sample dates from the CLNR data set no.8 were added to the 4 initial sample dates. The data from the 8 sample dates from the Loughborough data set and the CLNR no.8 data set were then added to a simulated 100-house three-phase low voltage network model. Table 3-3 lists the 8 sample dates chosen.

Table 3-3: The 8 selected sample dates from the two data sets

Sample Day	Loughborough Data Set	CLNR Data Set no.8
Day 1	Wednesday 16/01/2008	Saturday 12/01/2013
Day 2	Wednesday 02/07/2008	Wednesday 12/02/2013
Day 3	Wednesday 09/04/2008	Wednesday 10/04/2013
Day 4	Saturday 06/09/2008	Wednesday 20/02/2013

For the regression models used in section 5.3, 48 additional sample dates from the CLNR data set no.8 were selected from the 100 dates available to provide a representative mixture of:

- Working days
- Non-working days
- Various months

The 100-house three-phase low voltage model created comprises of a low voltage substation, three pieces of main low voltage cables (A, B, and C), a hundred service cables, and a hundred household smart meters. Thirty households were assigned to cable A and thirty and forty households to cables B and C, respectively. The low voltage cables were divided into 5-meter sections in order to facilitate the calculations at each section of the network. The service cables were allocated a constant length of six meters. The cable characteristics were obtained from the Northern Powergrid technical data documents, which can be found in the Appendix B.

Table 3-4: Cable types and their characteristics

Cable Name	Cable Type (mm)	Nominal Rating (amps)	Phase Resistance per Km (Ω)	Neutral Resistance per Km (Ω)
Mains-A	185	304	0.19	0.81
Mains-B	95	208	0.38	1.57
Mains-C	95	208	0.38	1.57
Service Cables	35	140	1.11	3.47

Figure 3-8 on the next page shows the schematic representation of the low voltage model created. This model is colour-coded based on the phases allocated to each customer. The phase allocation was carried out on the basis of the common practice popular in the industry otherwise known as “balanced network”, which assigns each customer to one of the three phases starting from the red phase, followed by yellow and blue.

This phasing pattern is used for the majority of the analysis. However, in section 5.2 this balanced phasing pattern is changed to model an “unbalanced network”, in order to replicate the unbalanced low voltage network that are found in practice. Throughout the next chapters the former phasing pattern will be referred to as “balanced network” and the latter as “unbalanced network”.

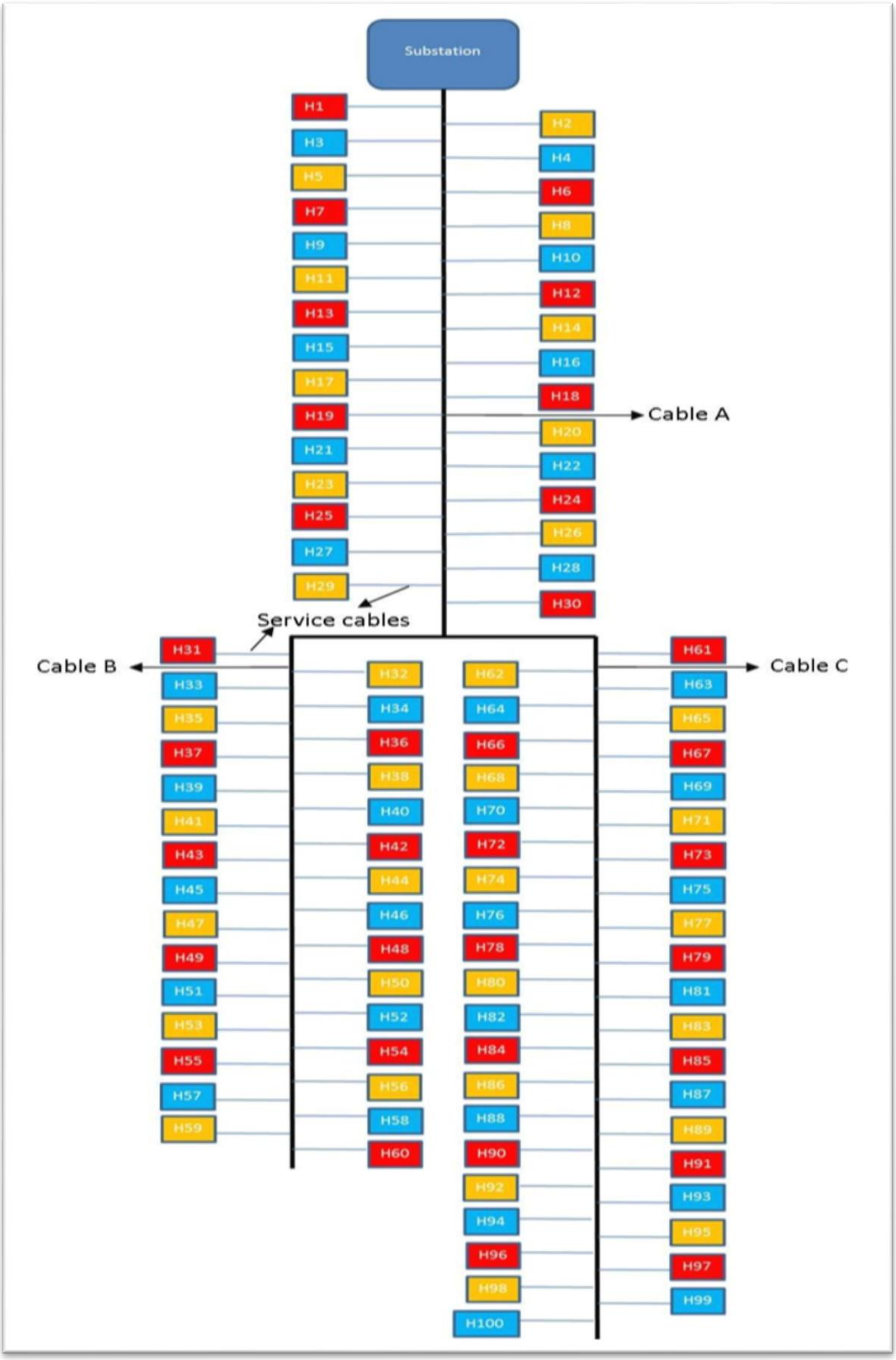


Figure 3-8: The low voltage model colour-coded based on balanced phasing

3.3.3 Estimation Methods

This section of the thesis describes the way in which the technical losses, voltage levels, and cable loading percentages were estimated on the low voltage network model in Figure 3-8.

3.3.3.1 Estimation of losses

The following steps were carried out to optimise the raw smart meter electricity consumption data, which were available at data set intervals, using Microsoft Excel and the built-in Visual Basic programme in Excel.

1. 1-minute currents (I) in Amps are estimated from customers' demand readings (P) in Watts based on constant voltage of 240V (equation 1). The current measurements are then calculated based on 5 minute, 10 minute, 15 minute, 30 minute, 60 minute, and 120 minute averages of customer demands:

$$(1).\quad I_{(A)} = P_{(W)} / V_{(V)}$$

2. Current measurements of the households are added on a three phase low voltage network model (see Figure 3-8) to calculate the total demand on each phase at the substation.
3. Customer demands on each phase at each household downstream of the substation are also calculated.
4. The currents on the red (R), yellow (Y), and blue (B) phases are then used to calculate the Current on the neutral phase (N) at each household (equation 2):

$$(2).\quad N = \sqrt{(R^2+Y^2+B^2-RY-RB-BY)}$$

5. Main phase and neutral phase resistance values are also calculated for each section of the network by multiplying the constant cable length (5 m) of the main cables and neutral phase resistance figures based on each cable type (see Table 3-4).
6. Losses are then calculated at each section of the low voltage network using the load on each phase at that particular section and the resistance values (Salman 1996 and Slavickas 2000), as calculated in equation 3:

(3). *Network loss at each section:*

$$\text{Main phase resistance} \times (R^2+Y^2+B^2) + \text{Neutral phase resistance} \times (N^2)$$

7. Network losses at each section are then added to estimate the losses of the low voltage network.
8. The same process is then repeated using decreased time granularity of currents from 1 minute to 5 minute, 10 minute, 15 minute, 30 minute, 60 minute, and 120 minute averages of customer demands.
9. The results of network losses at various time granularities for each day are then plotted for data analysis.

3.3.3.2 Estimation of Voltage Levels

1 minute currents in Amps are estimated from customers' demand readings in Watts based constant voltage of 240V (equation 1). The current measurements are then calculated based on 5 minute, 10 minute, 15 minute, 30 minute, 60 minute, and 120 minute averages of customer demands:

(1). $I_{(A)} = P_{(W)} / V_{(V)}$

1. Current measurements of households are added on the three phase low voltage network to calculate the total customer demands on each phase at the substation.
2. Loads on each phase at every household downstream of the substation are also calculated.
3. Main phase resistance is also calculated for each section of the network by multiplying the constant cable length (5m) and the main phase resistance based on each cable type (see Appendix B).
4. Voltage levels on each phase at the end of each section of the network is calculated by multiplying the household loads on each phase and the main phase resistance and the results are added together to obtain minimum and maximum voltage drop on each phase based on load pattern change throughout a 24 hour period for the representative dates.
5. A similar process is then repeated using decreased granularity of current data from 1 minute to 5 minute, 10 minute, 15 minute, 30 minute, 60 minute, and 120 minute time resolution of data.

6. The results of voltage level estimates at various time granularity for each day are then plotted.

As anticipated the highest voltage drops are experienced at the end of each branch of the network on branches B and C.

3.3.3.3 Estimation of Cable Loading Percentages

In order to calculate the percentage of loads on each phase of the low voltage network the following steps are carried out:

1. Current in Amps are estimated from customers' demand reads in Watts based constant voltage of 240V (equation 1). The current measurements are then calculated based on 5 minute, 10 minute, 15 minute, 30 minute, 60 minute, and 120 minute averages of customer demands:

$$(1).\quad I_{(A)} = P_{(W)} / V_{(V)}$$

2. Current measurements of households are added based on a balanced three phase low voltage network model and unbalanced phasing arrangements to calculate the load on each phase at the substation.
3. Loads on each phase at each household downstream of the substation are also calculated.
4. The sections of each branch A, B, and C were examined to and the sections with maximum loads on each phase are chosen.
5. The loads are divided by the maximum rating of each cable type to calculate the loading percentage of each phase of each cable.
6. These are calculated and plotted based on time granularities decreasing from 1 minute to 120 minute averages.
7. The results are plotted to compare the shapes based on granularity level.

3.4 Impact of Smart Meter Data Aggregation on Estimation Accuracy of Losses and Voltage Levels

This section describes the methods that are employed in this work in order to study the ways in which the accuracy of vital low voltage performance indicators such as technical losses and voltage levels are affected as the smart meter data from customers on the model low voltage network (see Figure 3-8) are aggregated together at various levels ranging from no aggregation to 10 customers on the same network.

3.4.1 Preservation of Consumer Data

Preserving the privacy and the individual lifestyle pattern of consumers is of utmost concern in relation to deployment of smart meters (Saputro and Akkaya 2013). In Saputro and Akkaya (2013), four main threats to customers' privacy, which can result from having detail smart meter data, are identified as follows:

- Obtaining individual behavioural data
- Obtaining information about the appliances in use
- Performing Surveillance
- Attacking homes when unoccupied

Under the Standard License Condition (SLC) 10 introduced by the OFGEM, the DNOs are not provided with load profiles of individual customers, due to privacy concerns (EA Technology 2015b; OFGEM 2017). Therefore, the DNOs are in the UK are asked to anonymise the data as soon as they receive them from the DCC in order to preserve customers' privacy.

3.4.1.1 Anonymization Methods

Previously, data aggregation methods were proposed in Armknecht et al. (2008); Garcia and Jacobs (2010); Marmol et al. (2012a); Marmol et al. (2012b); Onen and Molva (2012); Biselli et al. (2013) to investigate the ways in which the smart meter data can be aggregated in order to preserve the lifestyle patterns of the individual customers. Garcia and Jacobs (2010); Marmol et al. (2012a); Marmol et al. (2012b); and Biselli et al. (2013) propose various methods of encrypting the smart data using computational techniques and then transferring the data key to an aggregator meter, while Armknecht et al. (2008) and Onen and Molva (2012) focus on aggregating the meter data at various node points based on the topology of the network. These studies focus on the ways of preserving the privacy of customers, but they do not evaluate the impact of these methods of aggregation on the operation side of the network. Also as Saputro and Akkaya (2013) and EA Technology (2015a) point out, due to the fact that the DNOs in the UK do not own the raw customer data in the first place and only obtain them from the DCC, they usually opt for methods of anonymization that result in dissociation of customer IDs from the individual consumption data. Customer data aggregation based on phasing, which results in grouping the consumption data from a number of customers on a similar phase and within close proximity is deemed to be the most effective way of customer data anonymization (Saputro and Akkaya 2013; EA Technology 2015a).

Most recently, EA Technology investigated the effects of customer data aggregation in low voltage networks. EA Technology (2015a) focuses on the reduction of customer identification risk percentage as the customer data are aggregated. EA Technology (2015a) uses the CREST model to simulate typical 24 hour half-hourly household loads for customers on 10 cable with 9 to 124 connections. The training customer loads are then aggregated at feeder level and verified using the CLNR network monitoring data from low voltage network substations (EA Technology 2015a). The similarity between the aggregated load profiles and the individual load profiles are then examined using graphical representation of loads, correlation investigation of a random load to the aggregated load, and the k-means clustering approach (EA Technology 2015a). The results from this study indicate that as the aggregation level increases from the 1 to the 5 house level, the risk of customer identification drops from 100% to 15% (EA Technology 2015a), with the sharpest drop occurring between 1 and 2 house level of aggregation from 100% to 22% (EA Technology 2015a). EA Technology (2015b) also attempts to quantify the cumulative monetary losses to the DNOs as a result of customer meter data aggregation. The work carried out by the EA Technology (2015a and 2015b) investigate three network aggregation scenarios at aggregation points of looped services, and rural and urban feeder sections and examines the decrease in financial benefits in distribution grid investment strategies to the DNOs as customer data are aggregated. In EA Technology (2015b), the reduction in benefits as a result of customer data aggregation is expressed in monetary terms. The estimated financial impact involves the cost of disaggregating the aggregated customer profiles.

Our work investigates the impact of customer load aggregation on loss and voltage estimates in an urban low voltage network model with 100 houses at 5 different aggregation levels of 2, 4, 6, 8, and 10 houses. Also analysis presented in this thesis presents the effects of data aggregation in the context of changes in the accuracy levels of important low voltage network performance indicators such as technical losses and voltage levels.

3.4.2 Data Used

The smart meter data used in this part of the study is the half-hourly smart meter data obtained from averaging 1 minute data from the Loughborough data sets and the data set no.8 of the CLNR data sets. Our research uses actual customer demands from the CLNR and Loughborough trials for 8 sample dates from the two data sets between 2012-2013 and 2008 and 2009.

3.4.3 Aggregation Methods

Since the smart data will be transmitted to the DNOs in the UK are in half-hourly format, it was decided to investigate various house aggregation scenarios of the half-hourly averages of the demand data. In order to achieve this, the smart meter data used in a balanced 100-house three phase low voltage model (see Figure 3-8) are aggregated based on 5 different scenarios (see Figure 3-9). The smart meter data that are used in the studies presented earlier are aggregated at 2-house, 4-house, 6-house, 8-house, and 10-house aggregation points on the network. Figure 14 below shows a representation of some of the aggregation points for the customers on the red phase of cable A (as shown previously in Figure 3-8). The results of network loss and voltage estimates at each aggregation level are then compared with zero aggregation (1-house) results.

The technical network losses on the main three phase low voltage cables are calculated by using the load of a single customer on each phase and 5 meter section length of the low voltage cable where that load is connected to before the next customer is connected, which means the losses at each section are then added together to calculate the total loss of the network. However, when aggregation points are placed on the network model (as in Figure 3-9), the customer demands from two neighbouring houses on a similar phase are added together at the end of a 15 meter cable which is the middle point between the two houses.

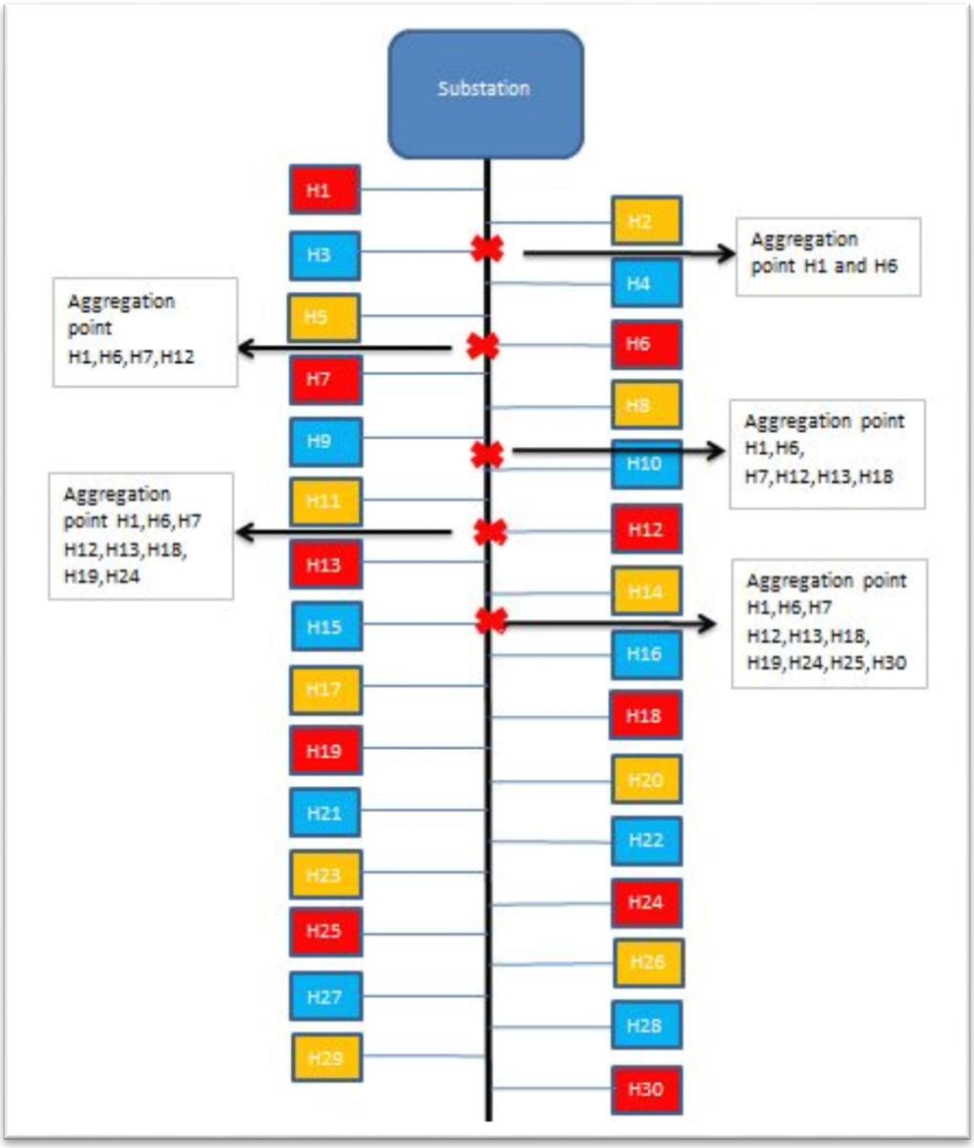


Figure 3-9: A sample of the aggregation points for customers on red phase (cable A)

In order to further investigate the impact of the placement of aggregation points on the estimate values, the network model represented in Figure 3-8, which is a 100-house balanced three phase low network with three cables and two branches, is changed to a 100-house balanced three phase low network with only one main cable and no branches. This new model is presented in Figure 3-10.

Throughout this thesis, the model in Figure 3-10 is referred to as “alternative” low voltage network model.

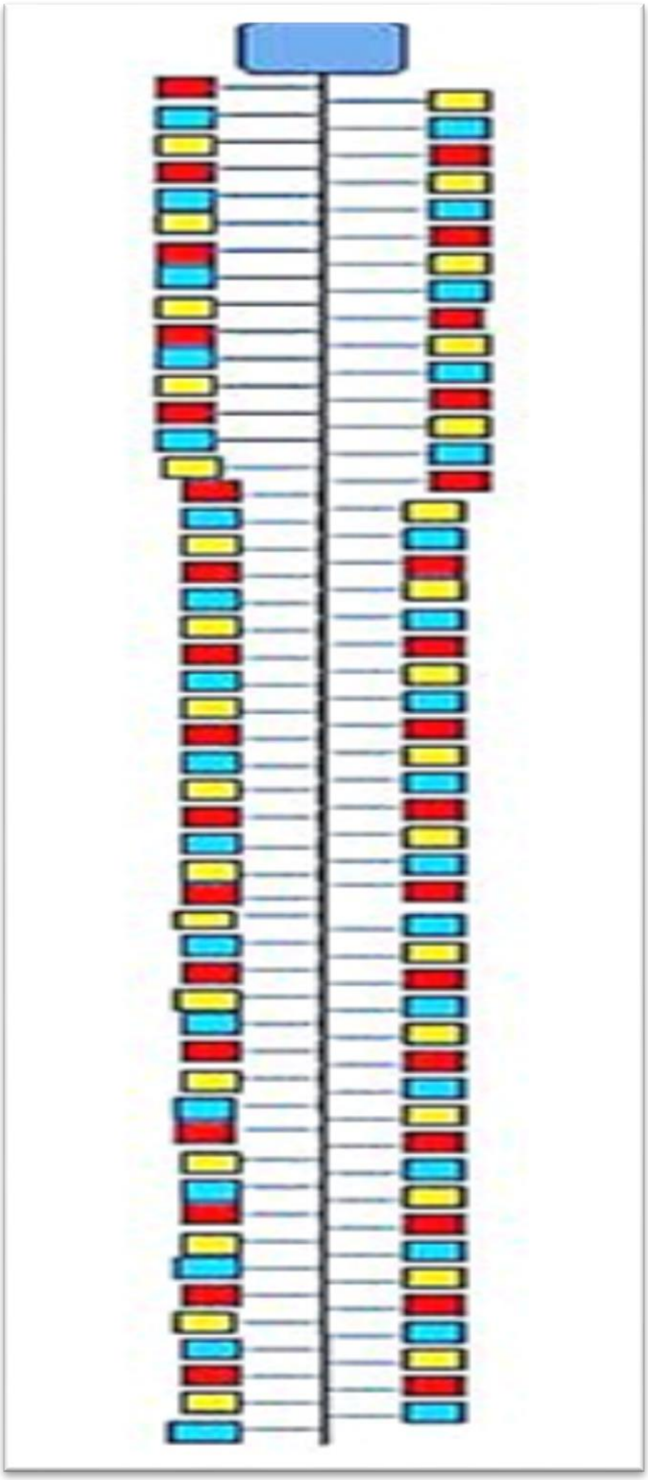


Figure 3-10: The alternative low voltage network model for customer aggregation studies (the colour codes refer to the customers' designated phases)

The model is then populated with smart meter data used in the first model and the aggregation points for the 5 levels of aggregation are placed on the new network as previously shown in Figure 3-9.

Chapter 4 Estimation of Missing Smart Meter Readings Using Smart Meter Data

In the UK, about 53 million smart gas and electricity meters will be deployed at households by 2025 (OFGEM 2016), with small businesses required to install Advanced Metering Infrastructure (AMI), and also more low voltage substations are likely to be equipped with feeder phase current measuring devices (Lees 2014). Not all smart meter consumption/generation data on every specific low voltage network will be available to the DNOs in the UK in real-time, due to the need for customer data aggregation for privacy reasons, time delays in processing and transmission, or lack of smart meter installations (Lees 2014). Therefore, as more smart low voltage grid applications are developed to utilise smart meter data, a need for determining the missing smart meter customer loads in low voltage networks will arise. This is becoming ever more important, especially with the growing reliance of smart grid applications such as network planning and design and Active Network Management (ANM) on more detailed and real-time customer demand information. Traditionally, DNOs have been using Maximum Demand (MD), load profiles, and annual billing data to predict consumers' demand patterns.

These approaches have been practical until recent years, due to the fact that they can provide accurate results in balancing the high voltage supply with the aggregated demands from end users (Valgaev et al. 2016). However, with the increasing share of embedded generation in the low voltage network, the demands need to be balanced at local levels and for disaggregated customers loads (Valgaev et al. 2016). Prediction of individual customer demands at local level is a difficult task, because as the demands are disaggregated from substation levels to end users, the load shapes become more volatile and diverse (Hayes et al. 2015). This is particularly challenging in the case of residential loads, where diversity between various lifestyle patterns and appliances in use can often create differences between two individual customers that may live in the same building but in two different flats. On the other hand, loads from commercial buildings usually follow a certain pattern during working hours and non-working hours (Valgaev et al. 2016).

This chapter evaluates the applicability of smart meter data to a number of approaches in estimating the missing half-hourly customer currents on the low voltage network using historical smart meter data and substation currents. 5 estimation methods have been selected as baseline approaches as the most that are deemed to be most practical within the DNOs applications and in relation to the data types that will be available to them. A number of other load forecasting methods have been tested previously. Short-Term Load Forecasting (STFL) methods have been used by the energy market suppliers in predicting aggregated loads at higher levels of the network and they have not been used at lower levels, because they do not perform as well in predicting disaggregated loads at low voltage levels (Valgaev et al. 2016). Machine learning methods and Artificial Neural Network (ANN) methods have also been used at high voltage levels. These methods are more suited to the transmission level of the network due to the fact that there are higher levels of information about the network and loads available to the transmission network operators. In addition to the lack of monitoring points on the low voltage network, the computational complexities of such methods have led to a reluctance by the DNOs to develop such methods at low voltage levels.

In this chapter, approaches 1 to 4 have been selected with regards to the types of data that will be available to the DNOs in the UK and the methods that are established in their network applications. These methods combine historical smart meter data with the transformer kVa allocation (Kerstin and Philips 2008; Arritt et al. 2012) and Monthly Usage Allocation (MUA) methods (Arritt et al. 2012), which are suited to the DNO applications at low voltage levels. Method 5 is a statistical approach based on the k-nearest weighted average of the closest points to the missing peak demand points. This approach is different from the k-nearest method used in Valgaev et al. (2016) in that it incorporates parameters such as half-hours, day types, weekly separation, substation load, and maximum temperature in calculating the Euclidean distance from the missing demand points, and it uses the weighted average of 5 nearest historical data points.

Methods 1 to 5 are also preferable to methods such as clustering and using class load shapes in (Velez et al. 2014), because they are easily applicable to any type of low voltage network. On the other hand, the clustering and load shape methods are defined for a limited number of specific low voltage networks and the great diversity in the low voltage network types restricts the use of such methods.

In the next sections of this chapter, the selected estimation approaches (described in section 3.2.2) are tested on two model networks of A-1 and A-2 that are populated with readings from two different data sets. A schematic representation models A-1 and A-2 is shown in Figure 3-2. These low voltage network models are created using smart meters from 20 households. Model A-1 is populated using data from the Loughborough Data set and Model A-2 is populated using data from the CLNR data set no.1. After testing the 5 estimation methods on the test networks of A-1 and A-2, the top 2 best performing estimation methods are applied on a network model with a larger number of customers, model B, and the results are reported.

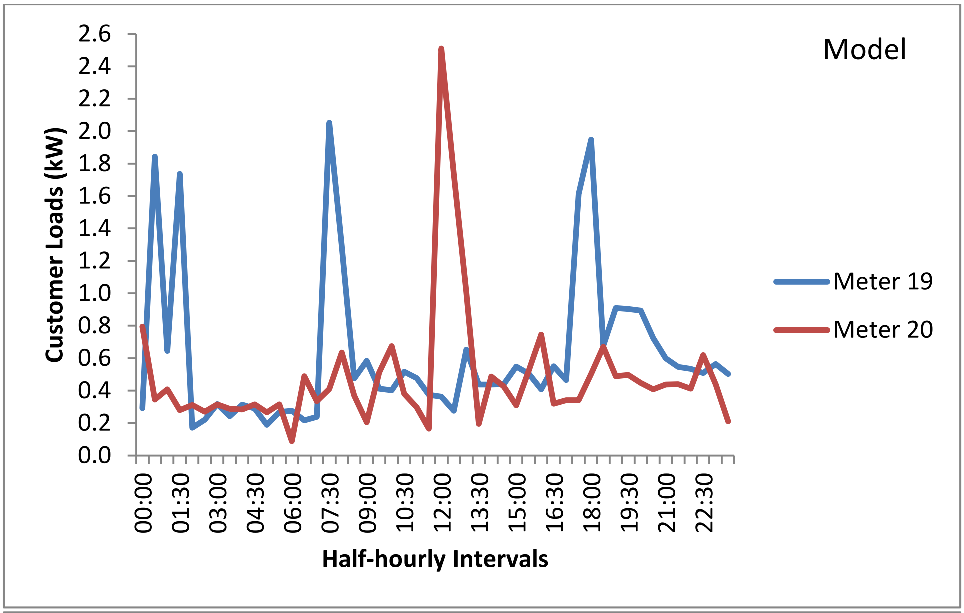
4.1 Load shapes at Substation Levels and End User Levels

When customer data are measured in aggregated form at substations, the diversity in individual customer demand patterns is eliminated and the load shapes become smoother, with clear peaks and drops, compared to individual customer demands patterns.

Figure 4-1 shows the individual load patterns for meters 19 and 20 from the Loughborough data set and CLNR data set no.1 used in models A-1 and A-2, respectively.

The load shapes are for the sample dates of 25/02/2009 and 27/02/2013. These sample dates were selected based on the following factors:

- Having the same day type (e.g. Wednesday).
- Being from the same month (preferably in Winter when the demand is the at the highest).
- The dates for which there is at least 13 month historical smart meter data are recorded.



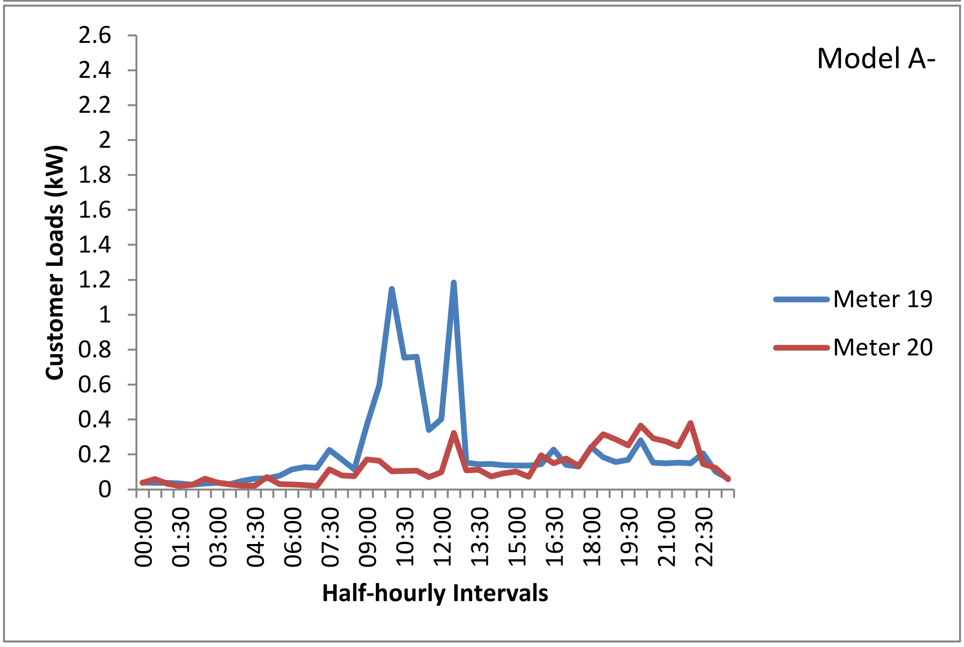


Figure 4-1: Individual demand patterns for sample meters from model A-1 and A-2 on the sample dates

As Figure 4-1 shows, investigating individual consumption patterns of 4 different consumers from 2 different data sets demonstrate a great degree of diversity. The load shapes from the two customers in model A-2 show a more homogenous pattern at non-peak times compared to the customers in model A-1. However, the way in which the four users consume energy at peak times seem to vary.

Figure 4-2 below shows the aggregated loads from all 20 meters in models A-1 and A-2 on the sample dates of 25/02/2009 and 27/02/2013.

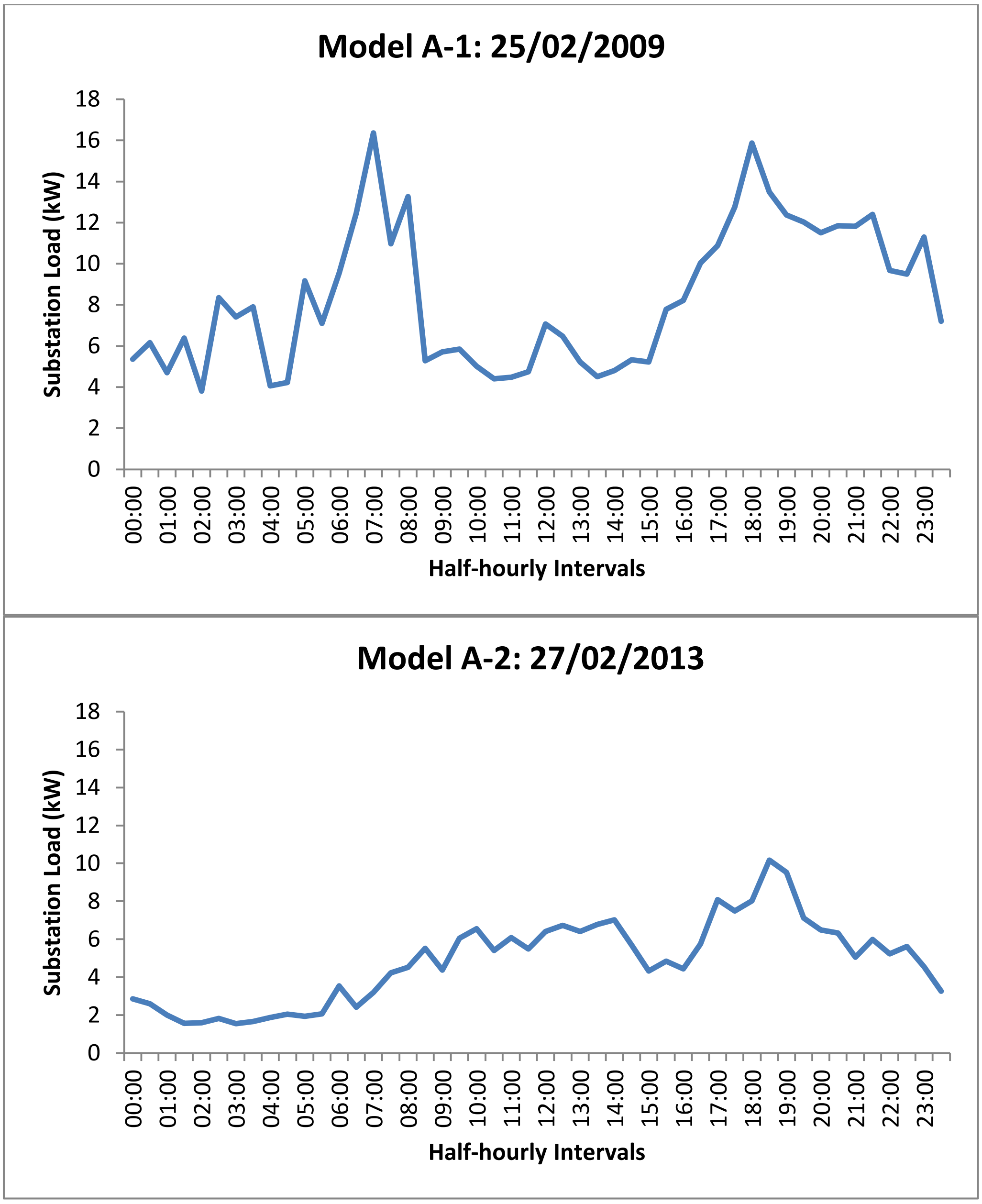


Figure 4-2: Substation demand patterns for models A-1 and A-2 on the sample dates

In Figure 4-2, a clear trend can be observed with peaks and drops in the demands. Also, Figure 4-1 clearly demonstrates that various customers have diverse lifestyle and usage patterns. In Figure 4-2, clear consumption trends can be observed with conventional peak times in the early morning and in the afternoon in model A-1 and a continuous increase in consumption from 06:00 to 14:00 and a peak later in the afternoon for customers in model A-2. A comparison between Figures 4-1 and 4-2 clearly demonstrate that the individual customer demand patterns are noisier and more volatile than aggregated demands at substations.

The estimation approaches presented in this thesis are based on the assumption that the proportion of loads from a group of customers on the low voltage substation is likely to be close to the proportion of loads from the same group of customer for "similar times and situations" in the past. For example, the split of currents from the meters 19 and 20 compared to the total currents at the substation from a week before the sample date is a good indicator of the split of the current from the same group of customers (meters 19 and 20) on the sample date. This is more likely to be accurate if similar day types (e.g. week days or weekends) are used. Figure 4-3 demonstrates the proportion of loads of meters 19 and 20 as percentages of the total loads at the substations on the sample dates and a week prior to the sample dates for models A-1 and A-2.

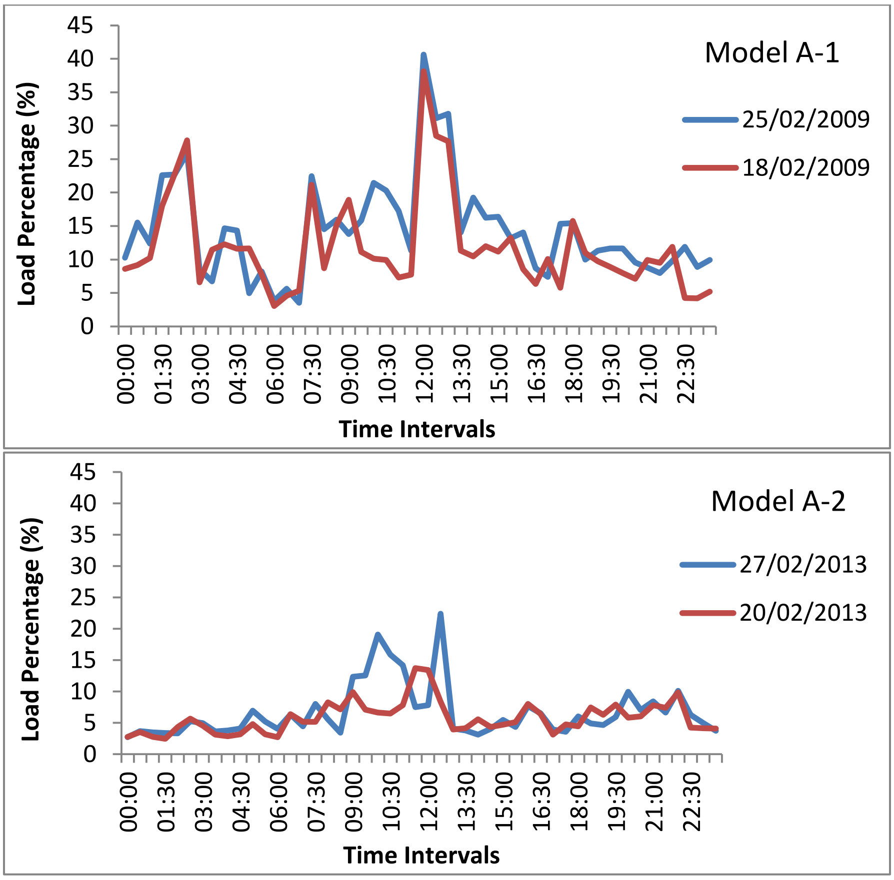


Figure 4-3: Loads on meters 19 and 20 as percentage of substation loads on the sample dates and from a similar day a week prior to the sample dates

As Figure 4-3 shows, for a majority of half-hourly intervals in a 24 hours cycle the load percentages of meters 19 and 20 as proportions of substation currents are similar to a similar day and a similar time interval from a week before. This is central to the estimation approaches presented in the next section.

In theory, the total currents at the substation should be a helpful measure in estimating the missing smart meter data on the sample dates in that through scaling, the volatility in individual smart meter demand patterns can be accounted for when the substation data and historical smart meter readings are combined. Also, the effect of seasonality is factored in the total substation currents and using the half-hourly substation current in the scaling calculations reflect the times of high or low demand if they are caused by seasonal effects.

4.2 Analysis and Results

The following load estimation methods have been tested on two models of low voltage network containing 20 meters (models A-1 and A-2). The estimation approaches assume that the real-time smart meter data for 90% of the customers are available on the sample date and the remaining 10% of the network do not communicate any smart meter data in real-time. However, historical data up to 13 months earlier are available for 100% of the network. This assumption is based on the prediction by the OFGEM that between 2% to 10% of the customers will not be provided with smart meters in the UK as the smart meters are gradually implemented by 2025 (OFGEM 2016). The approaches also assume that the total currents at the low voltage substations upstream of the network are available.

4.2.1 Methods 1 to 4: Estimation of Low Voltage Currents Using Data from Similar Historical Dates

These methods, previously described in section 3.2.2, assume that the total of the customer currents from the substation in models A-1 and A-2 are available to the DNOs. This data is combined with the substation readings from the other 18 meters on the network and smart meter readings of meters 19 and 20 from a similar day a week prior, a month prior, and a year prior to the sample dates of 25/02/2009 and 27/02/2013, respectively. Method 4 is employed by using the average of values recorded on similar dates to the sample dates up to 4 weeks earlier. The aim is to estimate the missing half-hourly values of the loads for meters 19 and 20 on the sample dates.

Predictions are made for loads from meters 19 and 20 at each of the 48 half-hourly intervals on the sample dates using the total loads recorded at the substation on the sample dates, the historical loads from meters 19 and 20, and the historical total substation loads from the other 18 meters on the network. The predicted half-hourly loads for meters 19 and 20 are then compared to half-hourly measured load values for the two meters on the sample dates and the Absolute Percentage Error (APE) for each half-hour prediction is calculated followed by calculation of the Mean Absolute Percentage Error (MAPE) for each method.

Estimation Equation:

Let S_{p1} be the total loads at the substation at the first half-hour of the day (00:00-00:30) on the sample date, S_{h1} be the total historical loads (e.g. a week, a month, or a year before) at the substation at the same half-hour, and Lh_1 be the historical loads recorded for meters 19 and 20 at a similar half-hour, then the missing half-hourly load of meters 19 and 20 on the sample date, Lp_1 can be estimated by the following formula:

$$Lp1 \; = \; Sp1 \times \frac{Lh1}{Sh1}$$

This can be used to estimate the missing values of currents for every half an hour on the sample dates (i.e. 1-48).

For example, using **method 1** the equation can be updated for model A-1 to:

$$Lp1\ (25/02/2009) \; = \; Sp1\ (25/02/2009) \times \frac{Lh1\ (18/02/2009)}{Sh1\ (18/02/2009)}$$

Using **method 2** the data used in the equation are obtained from a similar day a month prior to the sample date, so the equation will is expressed as follows:

$$Lp1\ (25/02/2009) \; = \; Sp1\ (25/02/2009) \times \frac{Lh1\ (28/01/2009)}{Sh1\ (28/01/2009)}$$

Using **method 3** the data used in the equation are obtained from a similar day a year prior to the sample date, so the equation is updated to:

$$Lp1\ (25/02/2009) \; = \; Sp1\ (25/02/2009) \times \frac{Lh1\ (27/01/2008)}{Sh1\ (27/01/2008)}$$

In **method 4**, the average values for each half-hour from 4 weeks prior to the sample dates are used in equations.

Results:

Figure 4-4 shows the sum of demands recorded and predicted for meters 19 and 20 on the sample date of 25/02/2009 in model A-1.

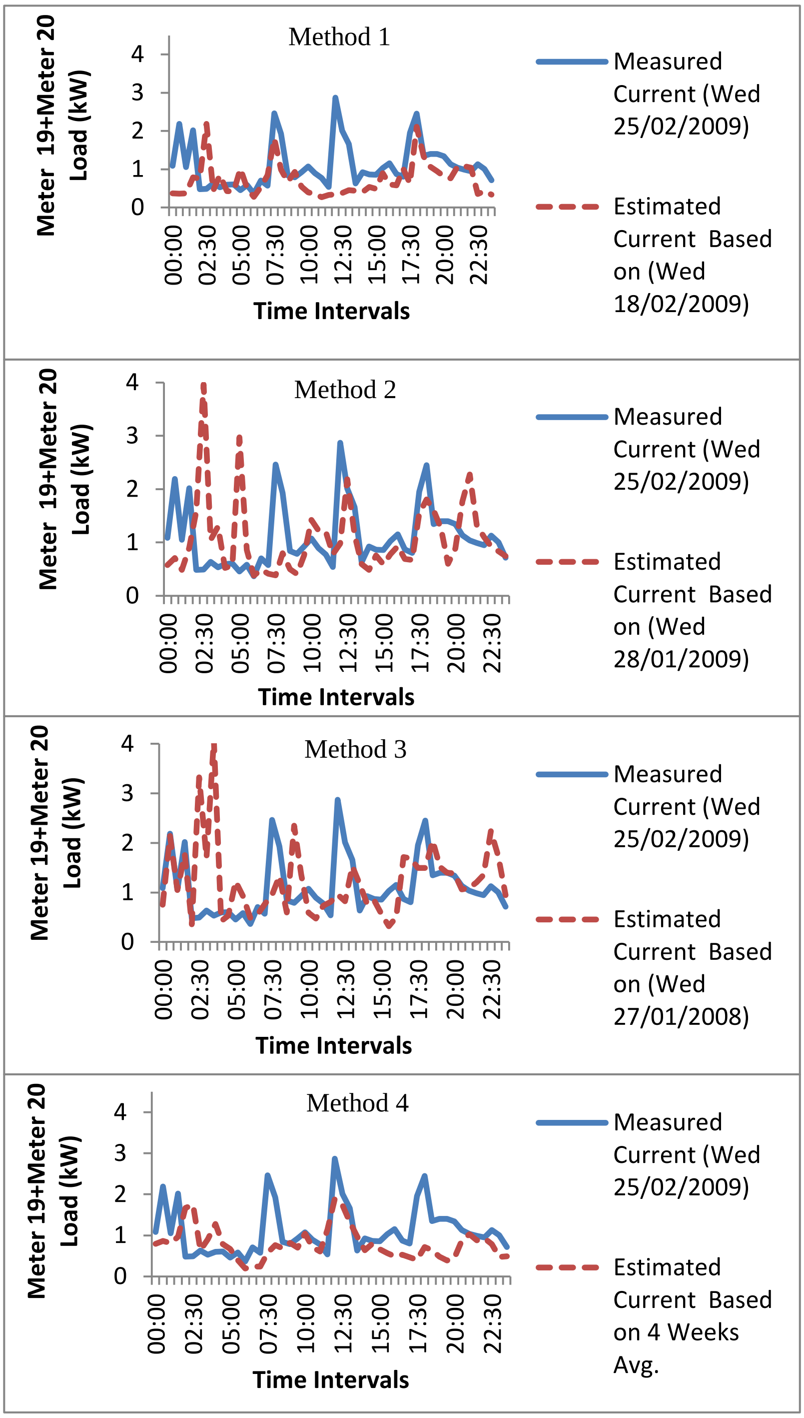


Figure 4-4: Estimated loads v predicted sum of loads from meters 19 and 20 using methods 1-4 (Model A-1)

Figure 4-5 shows the results of a similar analysis on model A-2 and on the sample date of 27/02/2013.

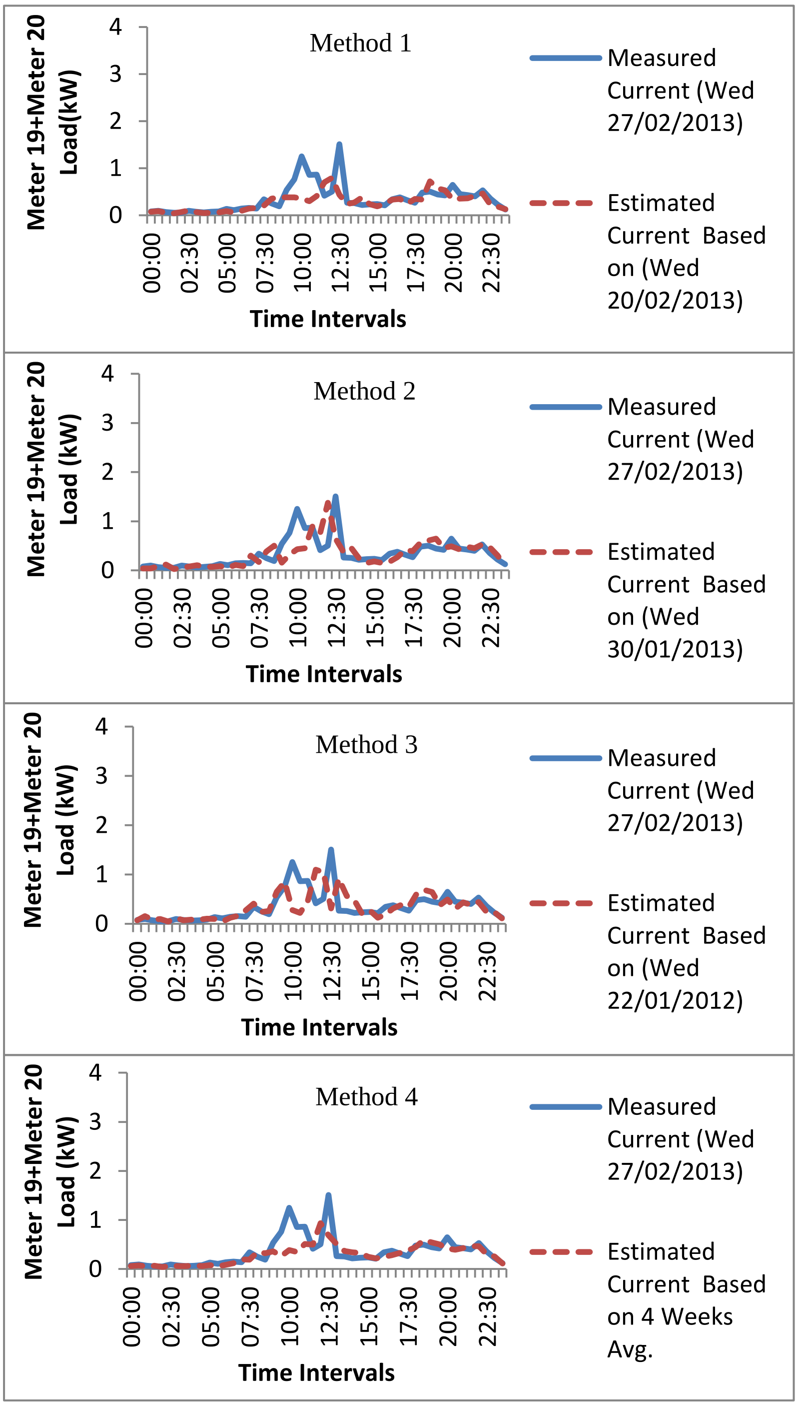


Figure 4-5: Estimated loads v predicted sum of loads from meters 19 and 20 using methods 1-4 (Model A-2)

Figure 4-4 indicates that the measured load pattern for meters 19 and 20, that represent 10% of the meters on the low voltage test network of A-1, show some unusual peaks at early hours of morning of the sample date. This can be due to anomaly in the data or an individual customer's behavioural pattern. A close look at the data set indicates that this is not an anomaly and similar peaks in consumption at meter 19 can are observed in the early hours of the morning of a week before the sample date and also weeks earlier. This becomes marginal compared to the usual peaks observed in aggregated residential currents (e.g. in Figure 4-2).

In relation to the three main peaks periods observed in the recorded loads in Figure 4-4 at time intervals of 07:00-09:00. 11:30-13:30, and 17:00-19:00, none of methods 1-4 successfully predict all three peaks. Method 1 using data from a week before, correctly estimates the peak in the morning and in the afternoon, but fails in predicting the peak demand at midday. Method 2 using data from 1 month earlier of the sample date, underestimates the first peak demand between 07:00-09:00 but provides a good estimation of the second and the third peaks. Method 3 using data from a year before, estimate the first and the second peak demands, but fails to accurately estimate the time periods at which these maximum demands take place. Method 4 using the average of the readings from 4 weeks prior to the sample date, fails to estimate 2 out of the 3 peak demand periods.

Figure 4-5 shows less complex load shape for the sum of currents from meters 19 and 20, that represent 10% of the meters on the low voltage test network of A-2. For these customers, the consumption peaks in the morning at around 09:00 and peaks again at around midday. This trend is also observed in the substation currents (see Figure 4-2). However, the substation currents also show a high consumption value in the afternoon from 17:30-19:30. Methods 1 to 4 all provide fairly accurate estimations of the general load shape on the sample date. However, the peak demands are either underestimated or predicted at a different time interval. The exception is method 2 which provides a fairly accurate estimation of the peak consumption readings.

Looking at the errors between estimated and the measured loads at each 48 half-hour intervals reveals that Methods 1 and 4 proved the most accurate estimation of the missing loads. The APE values are calculated between the actual and the estimated current at every half-hour and then the average of the 48 values are calculate to find the

MAPE for each method. The spread of the APEs for each method is shown in Figures 4-6 and 4-7 below. Figure 4-6 shows the APEs for model A-1 and Figure 4-7 shows the APEs for model A-2. The top whisker shows the maximum APE value, the bottom whisker shows the minimum APE value, and the box shows the interquartile range between Q1 and Q3. The line in the middle of the box shows the median APE value.

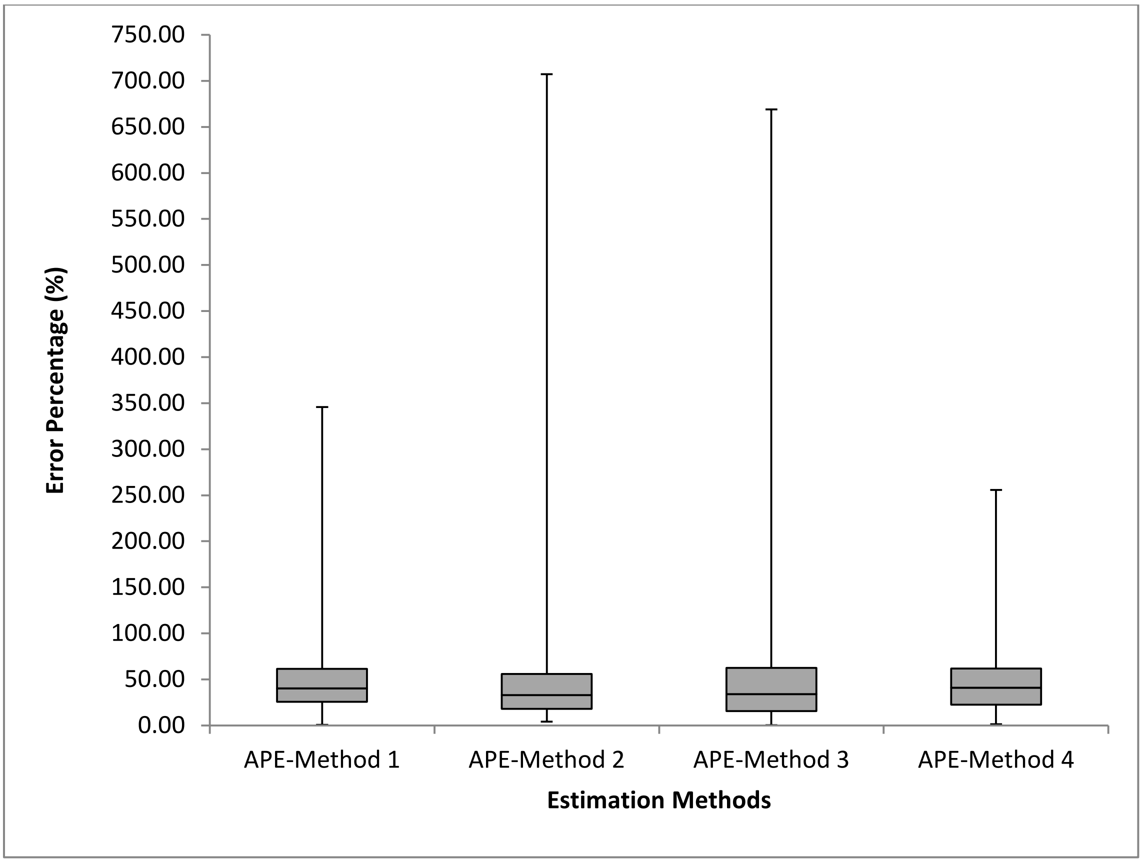


Figure 4-6: The spread of APEs for the 4 estimation methods in model A-1

The unusually large maximum errors shown in this figure should be treated as an outlier as a close look at the data shows that they are caused by estimation error at only one half-hour. Interestingly, the half-hour time period is the early morning peak for meter 19. This indicates that since this unusual consumption pattern is not represented by other customers on the network, these methods all fail to predict it. Notwithstanding this, most prediction errors range between just above 30-50%. On average in the case of model A-1, methods 1 to 4 provide MAPE values of 49.61%, 66.41%, 70.20%, 49.99%, respectively.

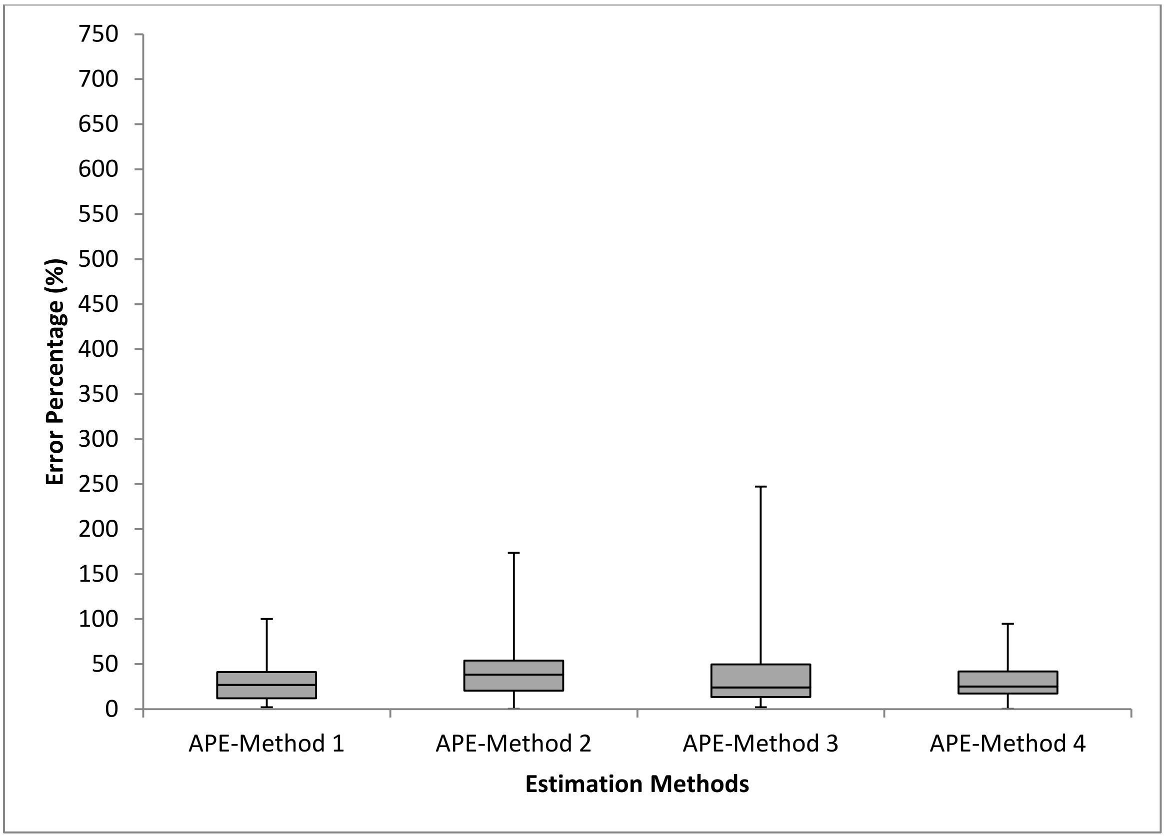


Figure 4-7: The spread of APEs for the 4 estimation methods in model A-2

In the case of model A-2, looking at the APEs in Figure 4-7 shows that most prediction errors are under 50%. In this case, methods 1 to 4, provide MAPE values of 29.83%, 44.51%, 42.74%, and 31.39%, respectively.

Clearly, method 1 that uses data from a week earlier provides the most accurate estimation of loads from the missing meters on the sample date. More importantly, by looking at the graphical representation of the load shapes in Figures 4-4 and 4-5, it can be observed that the methods described above do not provide the most accurate estimation of peak consumption values and/or time periods. Hence, method 5 below is tested to examine whether the peak demands, which are the most important elements in designing network capacities and managing the embedded generation on the network at local levels, can be estimated using half-hourly smart meter and substation data.

4.2.2 Methods 5: Estimation of missing currents using k-nearest weighted averages

In spatial statistics, K-nearest methods are employed in order to find the values with the least geographical distance from the target values, because these are likely to have a close value to each other as well (Oliver & Webster 1990, Hofstra et al. 2008). The k-nearest weighted average approaches used in this section are carried out in three steps:

1. Defining the separation parameters: instead of selecting the nearest points based on Euclidean distance (as done in geographical settings), various influential elements such as time and temperature are converted into weights (Wu et al. 2014).
2. Selecting the nearby points: k number of nearby points are chosen based on the lowest distance score (Ledolter 2013). The value of k is usually decided based on the best results produced on the test data (Fotheringham et al. 2002).
3. Estimating the values: the separation parameters are converted to weights (Fotheringham et al. 2002) and the weighted average of the k-nearest historical values are used in order to estimate the peak demand values on the sample dates.

The parameters that were used in step 1 were:

- Three peak times of 8:00, 12:30, and 18:30 are chosen based on the peak times observed from the substation load shapes.
- Half-hour in the day (H): each half hour in the day is given values between 1 to 48 and the separation of each half-hour from the three peak times are calculated.
- Day type (D): the separation between day types is calculated by assigning the value of 1 if the days are both on the same weekdays, 2 if they are different week days (e.g. Monday and Wednesday), and 4 if they are weekend days.
- Week separation (W): the number of weeks that two situations are distant from each other is calculated by assigning from 1 to 52 to the situations.
- Substation load difference (S): this is calculated by taking the difference the substation load on the sample date and the response date
- Temperature separation (T): this is also the difference in the maximum temperature recorded on the dates.

The Separation between each half hour and the peak times on the sample dates can be defined as:

$$Separation = H^h \times D^d \times W^w \times S^s \times T^t$$

In step 2, various k numbers were used and the k number that resulted in most accurate predictions were taken forward to be used in model B.

In step 3, the reciprocal of the separations is used to calculate the weighted averages.

Results:

Initially, power factors in calculating the separation were set as 1 (e.g. h =1, d=1, etc.). The results show various accuracies in models A-1 and A-2 when calculating the MAPE for the three estimated peak demand points on the sample dates. Table 4-1 below shows the MAPE results found for various k numbers of points for the two models.

Table 4-1: MAPE results using different k-nearest average numbers for models A-1 and A-2

	MAPE (k=5)	MAPE (k=10)	MAPE (k=15)
Model A-1	73.85%	74.08%	74.65%
Model A-2	44.84%	45.01%	46.76%

Based on these results k number was decided to be set as 5. So the average of values from 10 nearby points to the peak values on the sample dates were observed to provide the most accurate values. In the next step, the power values in the separation calculation were changed to examine whether the accuracy levels in predicting the peak consumption values can be improved. Since the weighted averages are carried out using the reciprocal of the separation and since it is more likely that the values from half-hours are close to the half-hour next to them, the separation measure were recalculated to emphasise this issue. To this end, the power factor for H was changed from 1 to 4 and the most accurate results were produced when h was set as 4 (h=4). Changing the temperature power factor (t) in the separation measure was found not to significantly improve the estimates. This could stem from the fact that the half-hourly temperature data was not available for these data sets and the maximum daily temperature values were the same for each half-hour. However, the fact that the temperature changes are already factored in the substation loads and since the substation load differences are also used in the separation, it can be argued that the impact of temperature changes has been accounted for. As Chaouch (2014) argues, aggregated loads at substation or node points along the network are more reflective of influencing factors such as seasonality and weather conditions, whereas disaggregated customer loads are more influenced by individual lifestyle patterns of the consumers.

Table 4-2 below shows the MAPE result of the k-nearest weighted averages when h is set as 4.

Table 4-2: MAPE results using different k-nearest average numbers for models A-1 and A-2 (h=4)

	MAPE (k=5)	MAPE (k=10)	MAPE (k=15)
Model A-1	48.80%	55.00%	55.26%
Model A-2	36.51%	37.53%	38.17%

Figure 4-8 below shows the load shapes of the missing meters 19 and 20 on the sample dates of 25/02/2009 and 27/02/2013 and the peak estimates using the k-nearest weighted approach with k set as 5.

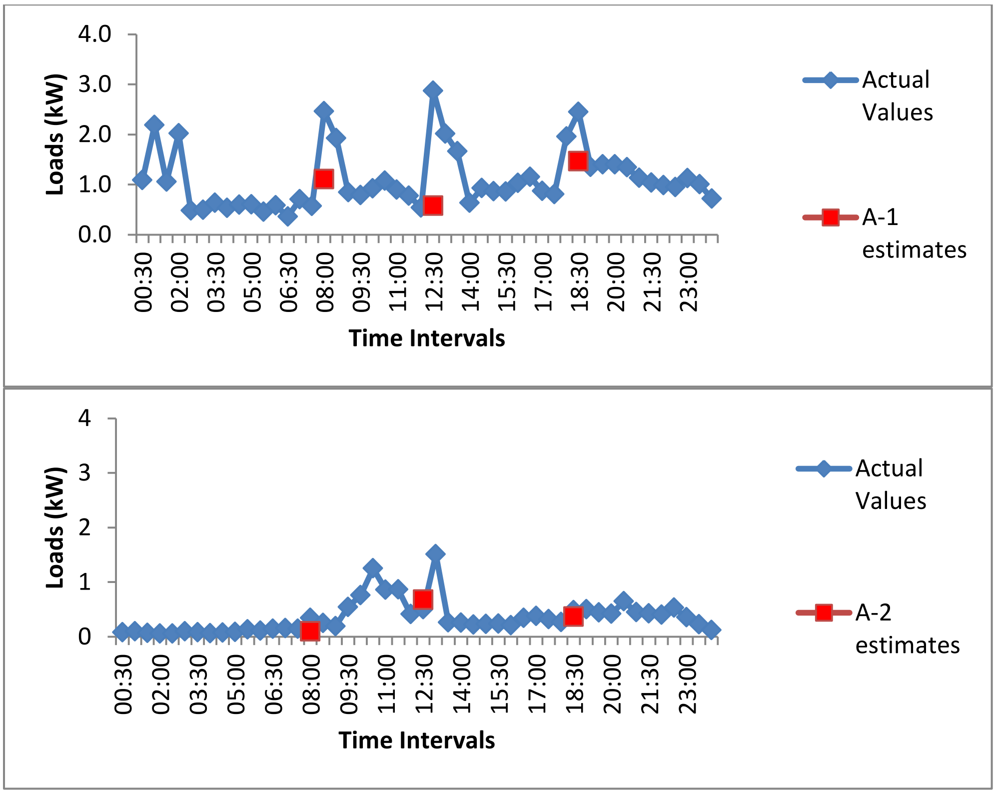


Figure 4-8: Measured loads compared to estimated peak loads using k-nearest weighted average (k=5)

As observed in Figure 4-8, the k-nearest approach does not provide perfect estimates of the peak consumptions. In theory, the models could have been improved by having half-hourly temperature data instead of maximum daily data, but this data was not available. The errors produced using the k-nearest weighted approach is only slightly lower than that of methods 1 and 4 discussed before. Table 4-3 below compares the MAPE results for the 5 methods.

Table 4-3: MAPE results for various methods used in estimating the missing loads on samples dates in models A-1 and A-2

Method	Method 1	Method 2	Method 3	Method 4	Method 5
MAPE A-1	49.61%	66.41%	70.20%	49.99%	48.80%
MAPE A-2	29.83%,	44.51%	42.74%,	31.39%	36.51%

As Table 4-3 shows, in the case of model A-1, method 5 offers just above 1% more accuracy compared to methods 1 and 4, but in the case of model A-2, methods 1 and 4 provide a better estimates of the loads by 5-7%. This can be due to the fact that the outlier observed in data used in model A-1 (see Figure 4.6) is taken into account in methods 1 and 4 when MAPE of 48 half-hourly APEs is calculated, but method 5 only estimates the peak values at 3 peak half-hours. Table 4-4 below compares the MAPE results for the load estimates at the three daily peak times.

Table 4-4: MAPE results for various methods used in estimating the missing loads on samples dates at peak times in models A-1 and A-2

Method	Method 1	Method 2	Method 3	Method 4	Method 5
MAPE A-1	43.61%	28.65%	47.26%	42.31%	48.80%
MAPE A-2	30.31%,	43.58%	40.81%,	35.91%	36.51%

The results presented in Table 4-4 shows that methods 1 and 4 perform best in estimating the demand at peak times as well.

In the next step, methods 1 and 4 are used in a larger network model, model B, with 50 meters to examine to what extent does having a larger customer group size can increase or decrease the accuracy levels of these estimation methods.

Estimation of missing loads in Model B

Model B is comprised of 50 customers. The smart meter data are obtained from CLNR data set no.1 and it is assumed that on the sample date of 27/02/2013, the substation currents are available but the smart meter data from only 90% of the network are communicated to the DNOs. Therefore, on the sample date, the smart data from meters 1 to 45 are assumed to be available in addition to the total currents from the substation, but the data from meters 46 to 50 are missing and are required to be estimated. The most accurate methods in the previous step, which were methods 1 and 4 are applied to examine whether the accuracy level is improved if larger sets of data are available to the network operators. Figure 4-9 below shows the predicted loads based the historical data from a week before the sample date and the averages of the meter readings at each half-hour from the same day type (Wednesdays) of the previous 4 weeks.

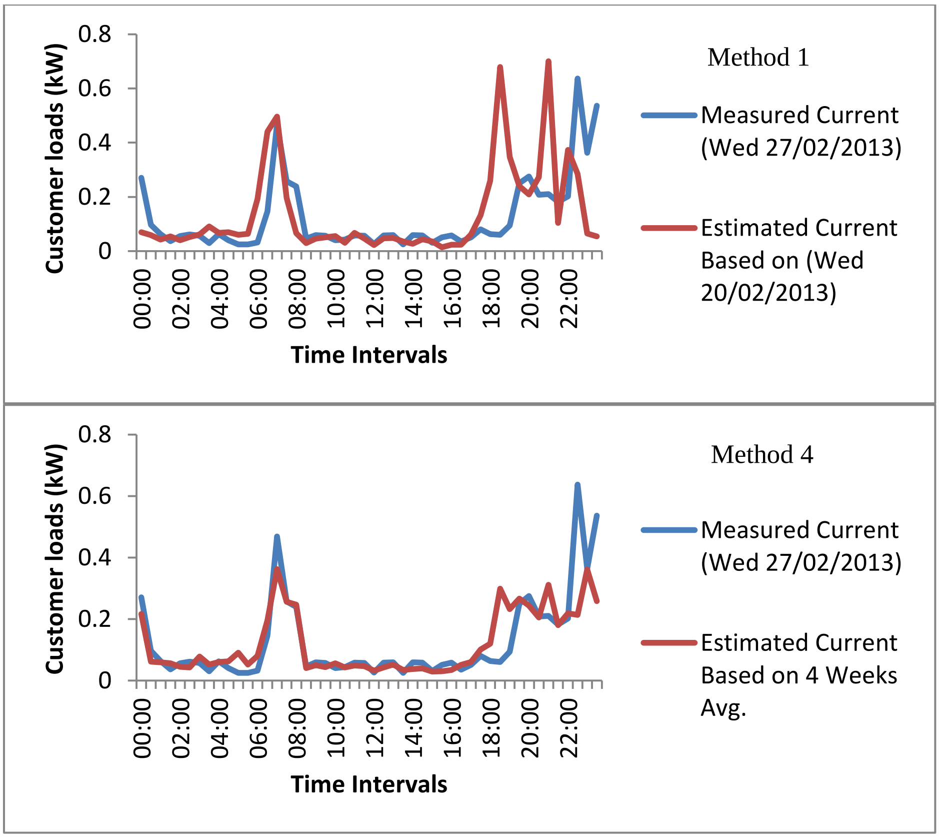


Figure 4-9:Estimated loads and measured loads for meters 46-50 in model B

Figure 4-9 shows that using the historical smart meter data from a week earlier to the sample date in combination with the substation data can accurately estimate the early morning peak demand between 07:00 to 09:00. Also, the first afternoon peak demand is shifted in the estimated currents from 19:30 to 18:30. The peak in the evening from 22:00 is also shifted to 20:00.

Figure 4-9 shows that using method 4 slightly underestimates the peak demand in the early morning hours from 07:00 to 09:00, but accurately estimates the demand at midday from 12:00 to 13:00. This method also underestimates the peak demand at 22:00 and shifts the afternoon peak demand at 19:30 forward by an hour to 18:30.

MAPE results applied to model B for methods 1 and 4 are 29.32% and 28.15%, respectively. These show a slight improvement when compared to MAPE results of methods 1 and 4 applied to model A-2 that are 29.83% and 36.51%, respectively. This shows that a larger sample size provides marginal improvements in the estimation approaches, as long as the proportions of missing current that are estimated are similar. In this case, both models A-2 and B assume that on the sample date, 10% of customer meters do not communicate real-time smart meter data to the network operators. In the next section the accuracy of approaches 1 and 4 in predicting the load demand of a single meter is examined.

Using methods 1 and 4 to estimate individual meter loads:

In this section, methods 1 and 4 which were the best performing methods in the previous models are tested on model B. However, this section assumes that smart meter data from one individual customer (e.g. meter 50) is missing on the sample date. This equates to 2% of the low voltage network in model B. The estimation of loads is repeated for each of the 50 meters in model B. Each time one of the meter’s data on the sample date is deleted and the loads for the meter is estimated using methods 1 and 4 and the errors between the estimated and the measures loads for the meter are calculated using MAPE. Figure 4-10 below shows the MAPE results obtained for the 50 estimation iterations carried out.

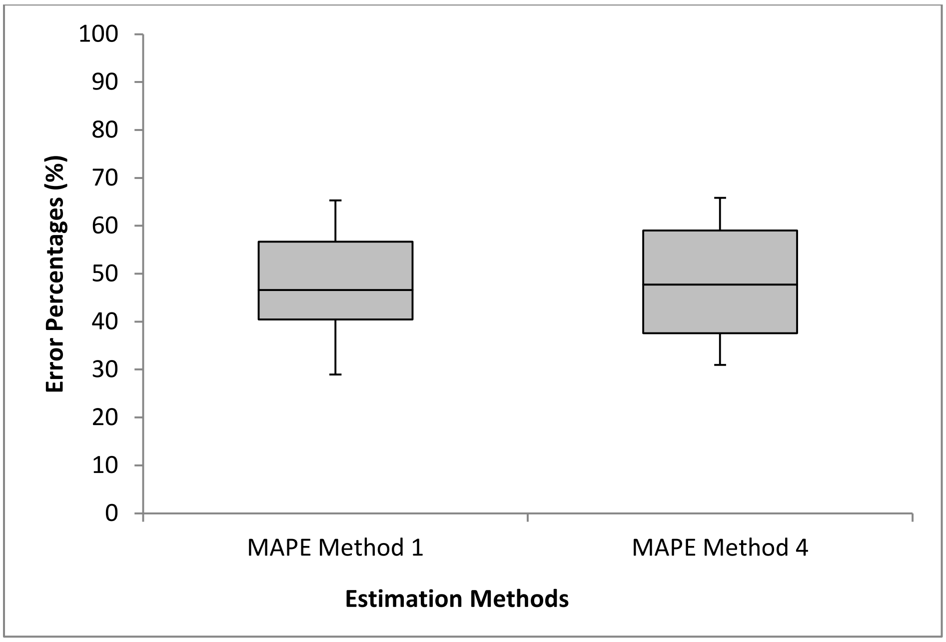


Figure 4-10: MAPE results for each 50 meters using methods 1 and 4

As Figure 4-10 indicates, the maximum errors observed for methods 1 and 4 are 65.30% and 65.80%, respectively and the minimum errors are 28.97% and 30.99%. These results show the difficulty of predicting an individual meter's load demand curve, due to the volatile and unpredictable nature of individual customers. Figure 4-10 also shows that there is less variation in results of method 4 as most estimation errors are in the interquartile range of 21.48%.

4.3 Summary

In this chapter, 5 different methods of estimating missing smart meter loads were initially tested on two different low voltage network models of A-1 and A-2 with 20 customers. On two different sample dates, the smart meter data from 90% of the customers of the low voltage network models were combined with the sum of all loads at the low voltage substation on the sample dates and historical smart meter data up to a year earlier than the sample dates. Based on the theory that the ratio of currents from the same group of customers on the sample date is likely to be close to the ratio of their loads in previous similar times and situations, the following methods of estimating the missing loads were tested:

- Combination of smart meter data from the metered customers and substation data on the sample dates with smart meter readings and substation data from a week, a month, or a year earlier, respectively.
- Combination of smart meter data from the metered customers and substation data on the sample dates with the average of smart meter readings and substation data from similar days up to 4 weeks earlier than the sample dates.
- K-nearest weighted average of 5 closest readings to the loads on the sample dates. Four elements of half-hourly distance, day type distance, substation load distance, and temperature difference were incorporated into the calculation of separation scores. This method was used in estimating the three main peak times observed at the substation level.

In section 4.2, the accuracy of these 5 methods were then calculated using the APE for each half hour and the MAPE for each method. It was found that methods 1 and 4 provide the most accurate results by producing MAPE results of just under 50% in both cases for model A-1, while for model A-2, method 1 produced MAPE results of just under 30% and method 4 produced MAPE of just above 31%. All methods performed well in replicating the overall load shape, especially at non-peak times. However, the estimated peak demands were sometimes under estimated, overestimated, or peak times were shifted by half an hour to hour. Using method 5 did not show any significant improvements over the other method in predicting the correct peak demand values.

In the next step, the top two performing methods, methods 1 and 4 were used in larger model (model B) with 50 customers to examine the effects of the having a larger sample size on the accuracy level of the estimation approaches 1 and 4. Since the missing portion of the sample size still constituted 10% of the low voltage network model, it as found that a using a larger sample size slightly improves the accuracy level of the estimations. The accuracy level of method 1 is increased by less than 1% and the accuracy level of method 4 is improved by 8% when errors are expressed as MAPE. However, method 1 and 4 in the context of model B produce far accurate estimation of the peak demand values and peak times, especially method 4 (see Figure 4-9).

In the next stage, methods 1 and 4 were tested in estimating individual customer loads on the sample date in model B. The process involved predicting the missing loads from 1 meter out of 50 meters in model B using both methods 1 and 4. This process was iterated 50 times, where each time one meter was considered as the non-metered customer with missing smart meter data. The results of MAPE showed that for method 1 the MAPE results range between 28.97% and 65.30% and for method 4 they range between 30.99% and 65.80%. However, in the case of method 4, more estimation errors lie in the interquartile range of 21.48% between just under 40% to just under 60% MAPE.

4.4. Conclusions

From the analysis and results presented above, it is clear that smart meter data can be utilised to increase the visibility of the customer load patterns on the low voltage networks beyond the low voltage substation levels. The 5 methods used in this chapter were devised to anticipate realistic situations in which portions of a meter data or a meter data as a whole is missing or is not available to the DNOs due to time delays, privacy concerns, faults, lack of coverage, or network errors.

Methods 1 and 4 build upon the established methods of kVa transformer allocation in Kersting and Philips (2008) and Arritt et al. (2012), Monthly Usage Allocation (MUA) in Arritt et al. (2012), and the annual billing approaches used by the DNOs (Stephen et al. 2014), by combining the half-hourly smart meter data with substation data to predict missing loads on segments of the low voltage network. These approaches are also more easily applicable to various types of low voltage network compared to load profiling and clustering methods in Stephen et al. (2014); Velez et al. (2014); Klonari et al. (2015); and Al-otaibi et al. (2016). Method 5 that is a statistical method based on k-nearest neighbours weighted averages alaso contributes to the k-nearest method used in Valgaev et al. (2016) by incorporating factors such as day types, week separation, substation load, and maximum daily temperatur to calcualte the closest historical half-hourly demand values to the missing half-hourly peak deamand.

Our analysis on two different sets of data shows that historical smart meter data from 90% of the network model in combination with substation data can provide accurate estimation of the missing currents from the 10% of the customers that are communicating smart meter data in real time. The accuracy level is higher at non-peak times, but the prediction of the peak demands is found to be a more challenging task.

The best two methods of estimating the missing currents based on the Absolute Percentage Error (APE) scores were found to be Methods 1 and 4 which use historical smart meter data from a week prior to the sample dates and the average of historical readings from 4 weeks prior to the sample dates, respectively. It was found that using k-nearest weighted averages of historical current does not perform better than methods 1 or 4.

These are simple approaches that provide fairly accurate results when employed in estimating missing currents from a number of customers. Both methods produce Mean Absolute Percentage Error (MAPE) results of about 30%. However, the accuracy of the estimation methods drop when they are employed to estimate individual customer loads. In this scenario, the MAPE results while using methods 1 and 4 in estimating 50 different individual loads range from just under 30% to 65%. However, in this case, method 4 produces more accurate estimations more consistently. This is mainly due to the variations in individual customer's load patterns. Our results show that using the average currents from 4 similar times and dates from 4 weeks earlier to the sample date better replicates the behavioural patterns of individual customers, compared to using the data from 1 week earlier.

The findings in chapter show that, half-hourly smart meter data can be used in simple statistical methods by the DNOs to predict the missing data from the smart meters that are not communicating their values to the DNOs in real-time. In relation to DNO applications, the estimated half-hourly values can be used in load flow analysis models to calculate network losses, voltage levels, and network capacity percentages. This information can subsequently be used in network planning and design and asset management. However, in relation to applications concerned with the monitoring of the network, a more accurate knowledge of peak demands can be highly beneficial to the DNOs. For example, Demand Response Management (DRM) or Active Network Management (ANM) rely heavily on the accurate knowledge of the customer peak consumptions at low voltage levels. This can be achieved with a more detailed knowledge of individual customer consumption patterns or more metering points at node points of the network.

Chapter 5 Studying the Relationship Between Smart Meter Time Resolutions and Important Low Voltage Network Performance Indicators

Traditionally, low voltage networks have been invisible to the DNOs, due to the way in which they were designed. Low voltage networks were designed to supply end users in the most cost effective fashion with minimum wasted energy. However, the introduction of embedded generation and low carbon technologies by customers have encouraged the network operators to design, operate, and monitor the low voltage networks in a more proactive way. This requires having access to important detailed information at various location on the low voltage network that were not deemed to be vital in the past.

The three performance indicators of technical losses, voltage levels, and network capacities can provide the DNOs with detailed picture of the low voltage networks. This knowledge can in turn improve various DNO applications such as asset management, network planning design, fault management, network monitoring, and active network management. For example, detailed knowledge of low voltage losses which in the UK constitutes for 5% of the energy delivered to the customers, can provide the DNOs with information about the underperforming parts of the network and the areas that need to be reconfigured and/or reinforced (Sohn Associates 2009; Dashtaki and Haghifam 2013; Poursharif et al. 2017).

Also the DNOs have financial incentives to reduce the losses on their operational networks. The incentives known as Loss Incentive Mechanism (LIM) were introduced by the regulator OFGEM in 2015 (OFGEM 2017) and a close look at the data published by Sohn Associates (2009) reveals that the amount of losses experienced on the low voltage network (5%) is far higher than the amount of losses experienced on the other parts of the electricity network (3%). A more detailed knowledge of voltage levels at various locations on the low voltage networks have also become increasingly pertinent due to higher proportions of embedded generation and low carbon technologies being installed on the customer end of the grids. The areas of the network that experience voltage rise and drop need to be identified by the DNOs in order to maintain the statutory voltage limits of 230V +10% -6%, which becomes more challenging as embedded generation can introduce reverse voltage into the system and low carbon technologies can introduce high voltage drops on different location on the networks.

Low voltage network feeder capacities are also indicators for the DNOs that are becoming increasingly important due to the rising proportions of low carbon technology demand and embedded generation at lower levels of the grid. Balancing the demand and generation on various phases of the network and identifying the headroom and ration of demand to the network capacity can provide significant information to the DNO applications.

In theory, smart meter data can provide the DNOs with estimates of the mentioned performance indicators by complementing the low voltage load flow analysis models used in the DNO applications. However, one of the limitations that can potentially decrease the benefits gained from the smart meter data is the low time resolution of such data in the UK.

In this chapter, the effects of decreasing the smart meter time resolution data from 1 minute intervals to 5, 10, 15, 30, and 120 minute intervals are investigated. This is carried out by identifying the impacts on the estimation accuracy of network losses, voltage levels, and low voltage feeder capacities in the context of a balanced and an unbalanced 100-meter three-phase low voltage network model. The data are acquired from two different sources of smart meter data sets. CLNR data set no.8 and the Loughborough data set. 56 sample dates across the two data sets were selected and the consumption data from the customers were used to populate the low voltage network model. The sample dates were selected from the dates that have at least data from 100 customers and can reflect different day types (working and non-working days) and different months. The information such as day and month type were used in order to estimate the 1 minute network losses based on loss estimates obtained at lower time resolutions of smart meter data in section 5.3 (e.g. a minute losses estimated using loss estimates at half-hourly averages).

5.1 Impact of Varying Smart Meter Time Resolutions on Loss Estimations in Balanced Low Voltage Networks

The main objective of this section is to investigate the impact of various smart meter data time resolutions on the accuracy of loss estimates in a balanced network model. A balanced low voltage network is the ideal situation for the DNOs. In a balanced low voltage network, customers are allocated equally to the three phases of red (R), yellow

(Y), and blue (B) to create a balanced network, in which customer loads are equally spread across the three phases. This ensures a lower amount of technical losses and a more efficient and cost effective operation of the low voltage networks.

In the model presented in this section, 100 customers are allocated to the three phases on the 3 main cables of A, B, and C (see Figure 3-8). 30, 30, and 40 customers are connected to cables A. B, and C, respectively via service cables. On each cable, the customers are allocated to the three phases of red, yellow and blue. The model is populated by data set smart meter data on 4 sample dates from the Loughborough data sets and 4 sample dates from the CLNR data set no. 8. In the next step, the smart meter readings are averaged to calculate the customer loads at 5, 10, 15, 30, 60, and 120 minute time resolutions. Finally, the technical losses for the entire network on each date is estimated by adding all the individual losses at each of the 100 houses.

The main question that this section is aiming to answer is to what extent does changing the time resolution of smart meter data from 1 minute time resolution to 5, 10, 15, 30, 60, and 120 time resolutions change the accuracy of the estimated losses and whether having smart meter data at time resolutions higher than 30 minutes (as will be the case for the DNOs in the UK) will improve the accuracy of loss estimations.

Figures 5-1 and 5-2 show the losses estimated at different time resolutions on the 8 sample dates from the Loughborough data set and the CNLR data set no.8 as a fraction of the losses estimated using 1 minute time resolution of smart meter data on the sample dates. The sample dates are presented in Table 5-1.

Table 5-1: The 8 sample dates selected from the two data sets used in time resolution studies

Sample Day	Loughborough Data Set	CLNR Data Set no.8
Day 1	Wednesday 16/01/2008	Saturday 12/01/2013
Day 2	Wednesday 02/07/2008	Wednesday 12/02/2013
Day 3	Wednesday 09/04/2008	Wednesday 10/04/2013
Day 4	Saturday 06/09/2008	Wednesday 20/02/2013

The loss ratio at every time resolution interval is calculated by dividing the estimated loss value at a particular time resolution by the estimated loss value at using 1 minute time resolution intervals.

Let L_1 be the estimated losses at 1 minute interval on one of the sample dates and L_5 be the estimated losses at 5 minute time resolution on the same specific date. The loss ration of L_r can be calculated as follows:

$$L_r = L_5 / L_1$$

L_5 in the above formula is then replaced by L_{15}, L_{30}, L_{60}, and L_{120} in order to calculated the loss ratios at the lower time resolution of smart meter data as shown in Figures 5-1 and 5-2.

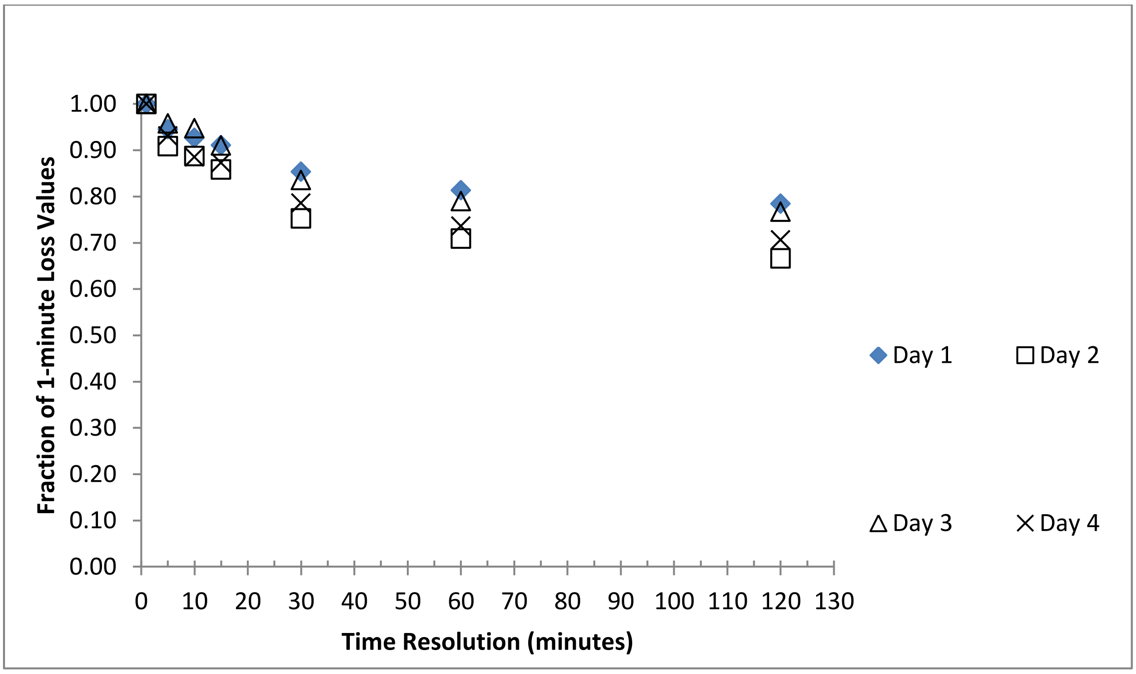


Figure 5-1: Fraction of estimated losses for different time resolutions (Loughborough data set)

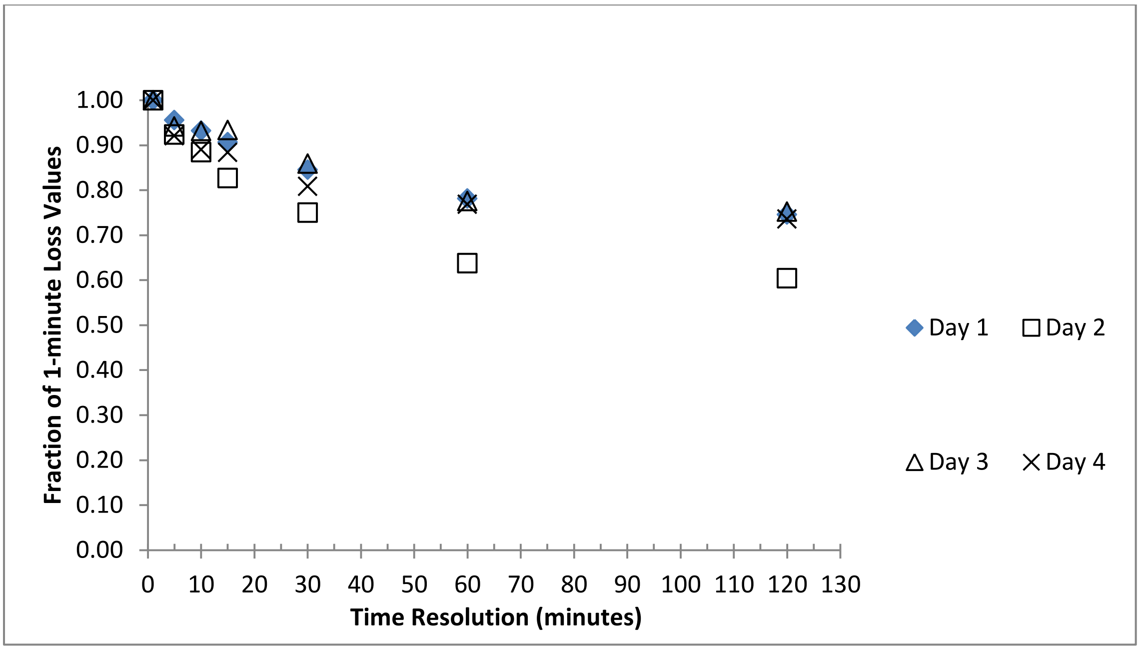


Figure 5-2: Fraction of losses for different time resolutions (CLNR data set no.8)

As Figures 5-1 and 5-2 demonstrate, as the time resolution of smart meter data is decreased from 1 minute to 120 minute intervals, the estimated values for losses also decrease. Significantly, the sharpest fall in the loss estimate values (mean = 11.2%) occur between the data set and 30 minute intervals, which is highly important to the DNOs, due to the fact that the smart meters in the UK will collect, store, and transmit the consumer data on half-hourly averages. This underestimation trend is consistent across both sets of data and all of the 8 sample dates. The underestimation percentage at each time resolution interval compared to the estimated losses at 1 minute time interval is calculated as follows:

$$\left| L_t / L_1 - L_1 / L_1 \right| \times 100$$

Where L_t represents loss estimates at different time resolution intervals (e.g. t = 5, 10, 15, 30, 60, or 120) and L_1 represents loss estimates at 1 minute time interval.

Table 5-2 shows the average underestimation percentages for the 4 sample dates from each data set as the time resolutions are decreased from data set to 120 minute intervals.

Table 5-2: Mean underestimation percentages of loss estimates on the 8 sample dates

Time Resolution (minutes)	Average Underestimation % Loughborough Data Set	Average Underestimation % CLNR Data Set No.8
5	6.4	6.4
10	8.8	9.0
15	11.2	11.2
30	19.3	18.4
60	23.7	25.9
120	26.9	29.0

As Figures 5-1 and 5-2, and Table 5-2 show, on average, changing the time resolution of smart meter data from 1 minute intervals to 30 minute intervals leads to underestimation of losses by 19.33% and 18.44% on the sample dates from the Loughborough data set and CLNR data set no.8, respectively.

In the case of the sample dates from the Loughborough data set, the underestimation percentages at 30 minute intervals range between 14% to 24% and in the case of the sample dates from CLNR data set no.8 these numbers range between 15% to 25%. Significantly, the highest percentage of underestimation occurs when the smart meter data time resolution is decreased from 1 to 15 minutes. This is significant, because many European countries such as Denmark and Italy have opted for smart meter collection at 15 minute intervals.

Looking at the underestimation percentages in Table 5-2, the average underestimation for the sample dates from both data sets is just above 11%. In the case of the sample dates from Loughborough data set, the underestimation percentages at 15 minute intervals range between 8% to 16% and in the case of sample dates from CLNR data set no.8 these numbers range between 6% to 17%. This indicates that having smart meter data at 15 minute intervals instead of 30 minute intervals does not significantly improve the accuracy of losses and time resolutions of higher than 15 minutes are required to minimise the loss of estimation accuracy. On the other hand, the results above indicate that having smart meter data at time resolutions of 60 minutes does not severely decrease the accuracy of loss estimates compared to having smart meter data at 30 minute time resolutions.

The results discussed above were also confirmed when 52 additional sample dates were tested using the data in CLNR data set no.8. The box plots in Figure 5-3 show the underestimation percentages occurred at each time resolution interval for all of the 56 sample dates. The top whisker shows the maximum value, the bottom whisker shows the minimum value, and the box shows the interquartile range between Q1 and Q3. The line in the middle of the box shows the median value amongst the 56 values for each time resolution interval.

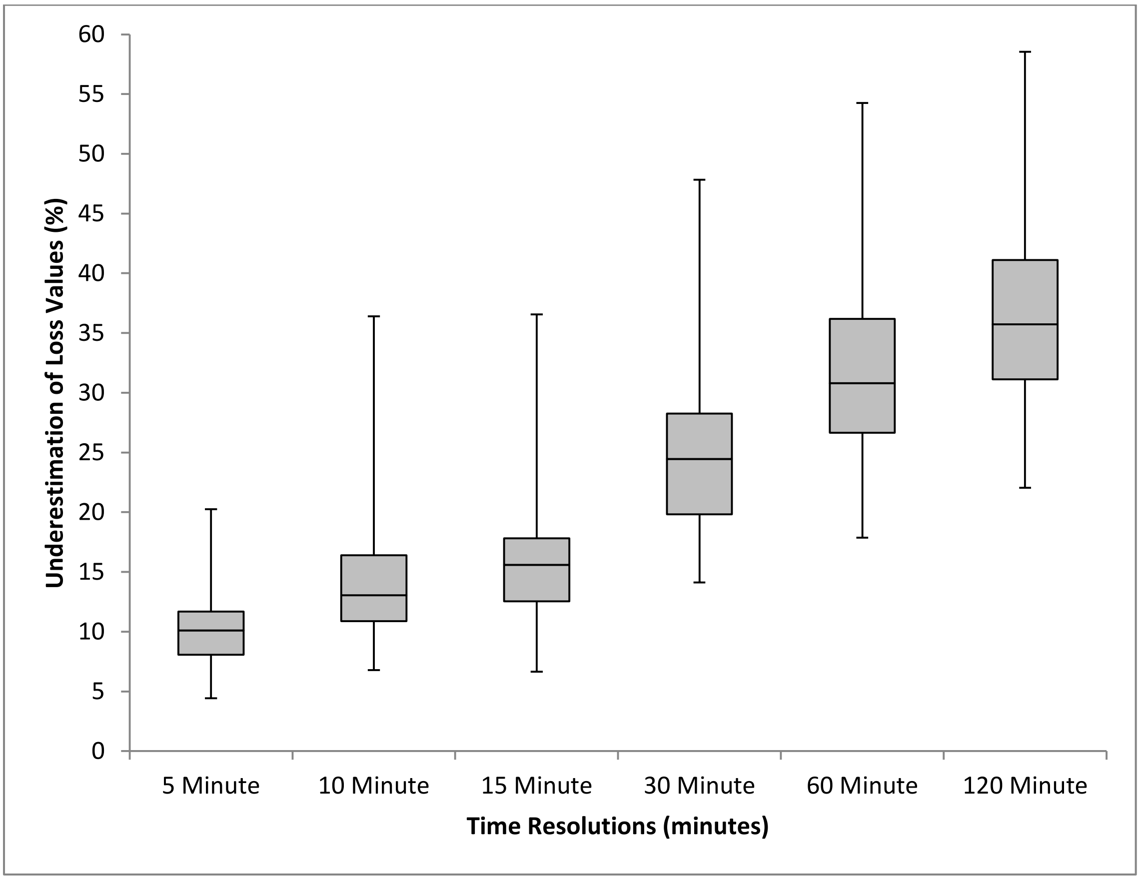


Figure 5-3: Variations in underestimation percentages of losses at different time resolutions compared to data set time resolution losses

As Figure 5-3 above highlights, median underestimation percentages at 15 minute and 30 minute time resolutions are just above 15% and just under 25%, respectively. At 15 minute time resolution, the minimum and maximum underestimation percentages are 6.6% and 36.5%, respectively and the numbers at 30 minute intervals are 14.1% and 47.8%.

The underestimation of losses at lower times resolution intervals is caused when the high demand spikes in the customer loads are smoothened when currents are averaged over time (Quiroz et al. 2012; Urquhart and Thompson 2015; Poursharif et al. 2017). This occurs in loss estimation models based on power flow analysis. Power flow models take into account the impedance, individual customer demands, and network topologies and these advantages allow the network operators to calculate losses at various points on the network, whereas the method using power difference (input-output) do not provide the breakdown of losses on the network and require meters at both ends of the network (Poursharif et al. 2017; Urquhart et al. 2017).

Power flow analysis models such as the model used in this thesis use the square of the current in I^2R to estimate the technical losses. Therefore, as the spikes in customer demands are decreased when the currents are averaged, the losses estimated at each section of the network are underestimated.

Based on common occurrences of spikes in the customer in customer loads, three cases can be defined:

- Single high demand spike for a short period of time:
  - Let a current value be I for a time period of D and let us consider the current outside of D to be zero, then for the time interval of W ($W \geq D$) the technical loss is calculated by using the average of current is calculated using $I*D/W$. Hence using the average of current in the loss calculation equation will lead to the underestimation of losses at the rate of $1/W$.

- Very low variations in demand:
  - Let us consider the current to have a constant value of I over the entire time interval. Then for a specified period of time, the calculated losses are independent of the number and size of the time intervals.

- Smooth with linear trend:
  - Let us consider the situation in which there is constant rise in the current over the period of time. If the time period is divided into equal interval of

n, as the width of the intervals (W) increases, the losses calculated using the average of current in that time period decreases in line with W^2.

In practice, a mixture of these demand patterns occur in low voltage networks connected to multiple customers and therefore a combination of these relationships between the time intervals and the amount of losses estimated is likely to be present in the low voltage networks.

5.2 Impact of Varying Smart Meter Time Resolutions on Loss Estimations in Unbalanced Low Voltage Networks

Although the DNOs strive to design a balanced low voltage network, in reality some low voltage networks are unbalanced. Fleckenstein (2013) defines an unbalanced three-phase low voltage network as a situation in which at least one of the phases experience higher currents. This could be due to a higher number of customers being connected to one of the three phases (usually the red phase) when network extensions are carried out in the field or due to higher proportions of low voltage technologies being installed on one of the phases compared to the other (Pezeshki and Wolfs 2012; Ahmadi et al. 2016). Phase imbalance can lead to higher technical losses, lower quality of supply, and lower longevity of network assets (Pezeshki and Wolfs 2012).

In order to create an unbalanced low voltage network model, the customers that were previously allocated to the three phases of red, yellow, and blue in 33:33:34 ratios were reallocated to these phases in 40:30:30 ratios with 40 out of the 100 customers on the red phase. The technical losses for this unbalanced low voltage network were calculated using the data from both Loughborough and CLNR no.8 data sets on the 8 sample dates which were previously used in section 5.1.

Figure 5-4 below compares the averages of estimated losses for these 8 sample dates in the balanced and unbalanced network arrangements.

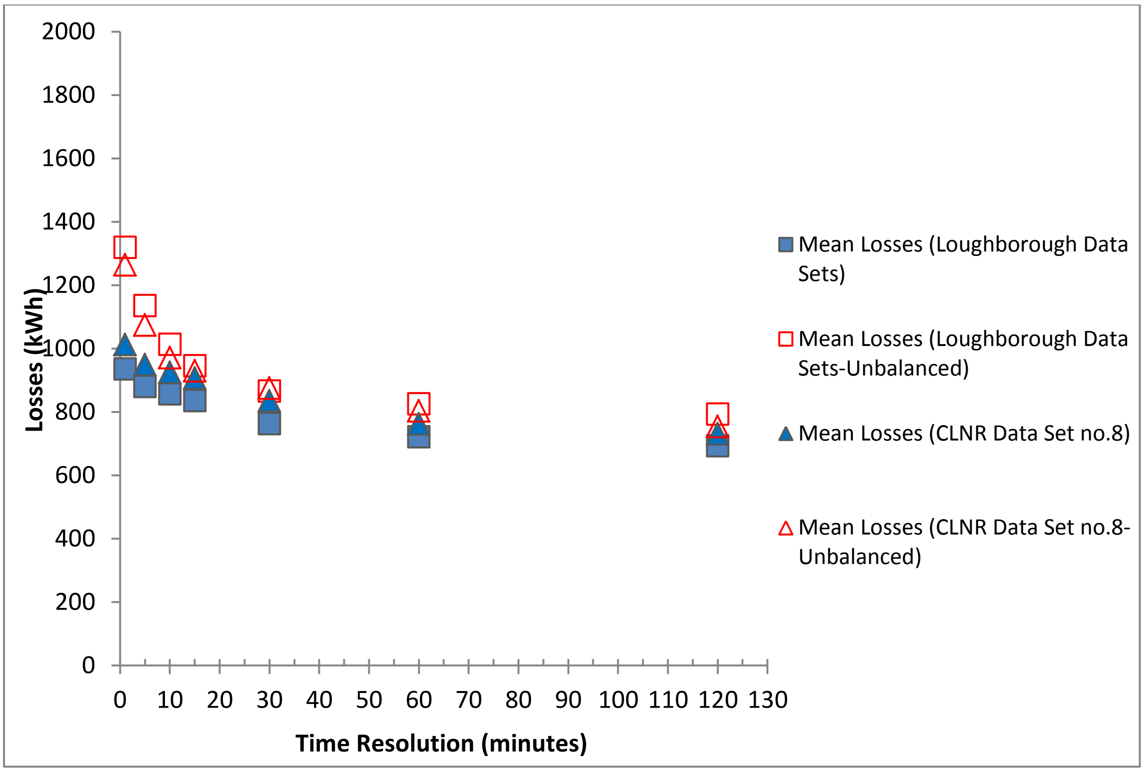


Figure 5-4: Average of technical losses on the 8 sample dates in the balanced and unbalanced models

As Figure 5-4 shows, the unbalanced network arrangement leads to higher voltage losses as previously discussed in Pezeshki and Wolfs (2012). However, the question is whether the relationship between smart meter time resolution intervals and estimated losses follow the same pattern as in the balanced network model discussed in section 5.1.

Figures 5-5 and 5-6 below show the estimated losses in the unbalanced network model at various smart meter time resolution intervals on the 8 sample dates as a fraction estimated losses at data set time resolutions on each date.

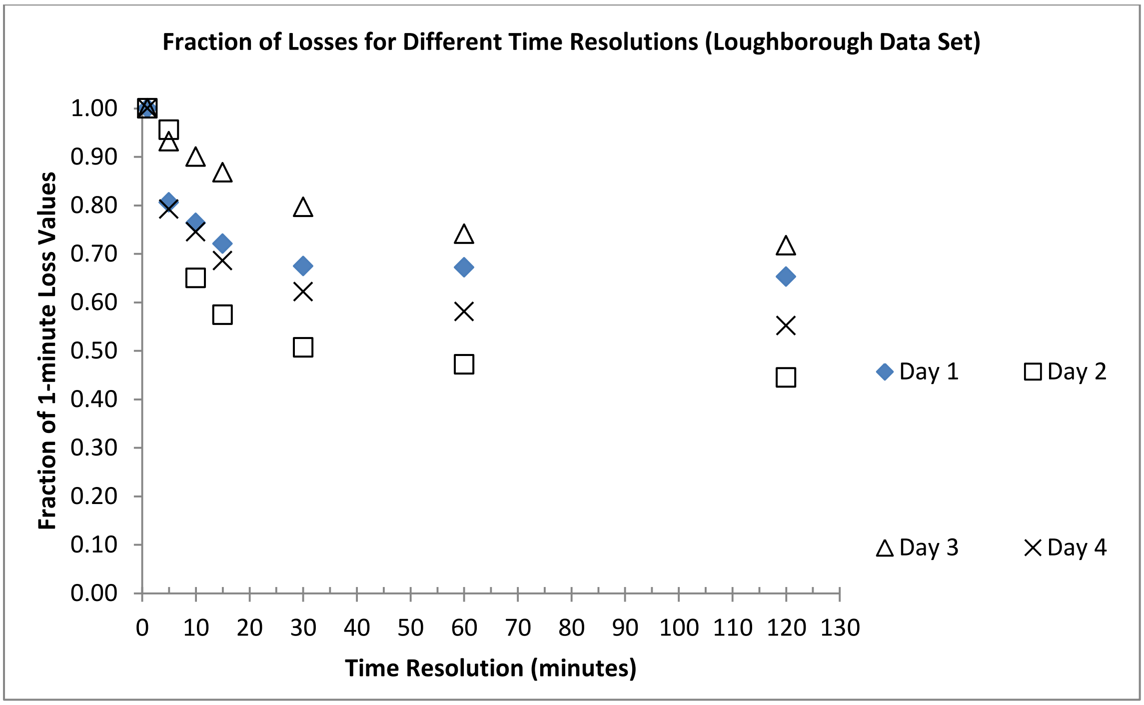


Figure 5-5: Fraction of losses for different time resolutions in an unbalanced network (Loughborough data set)

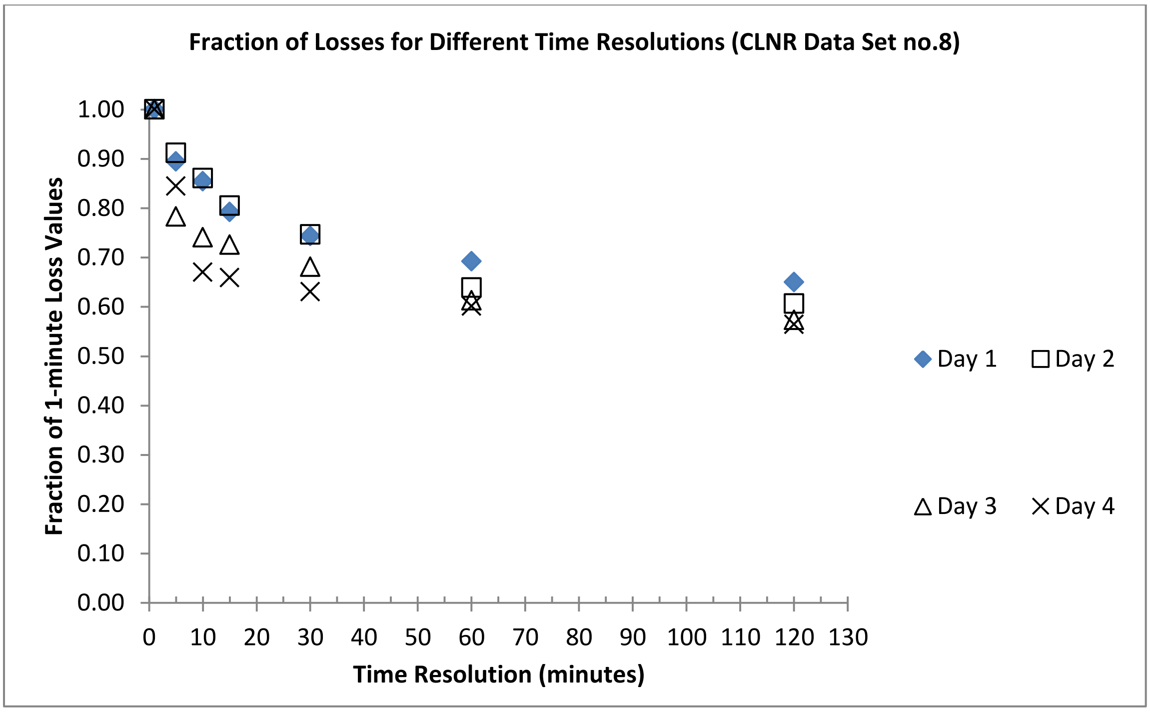


Figure 5-6: Fraction of losses for different time resolutions in an unbalanced network (CLNR data set no.8)

Figures 5-5 and 5-6 both confirm that a similar trend to the balanced network model can also be observed in the unbalanced network model in that as the time resolution of the smart meter data decreases from data set to 120 intervals, the technical losses are underestimated for all the 8 sample dates. Similar to the previous model in section 5.1, the larger amount of underestimation occurs when the smart meter data are averaged over half an hour intervals and the underestimation trend slows down from 30 minute to 120 minute time intervals.

As Table 5-3 below indicates, similar to the balanced network model the largest amount of loss underestimation occurs when the time resolution of the smart meter data is decreased from 1 to 15 minutes. On average, this figure is 25.46% and 28.75% for the sample dates from the Loughborough and CLNR no.8 data sets, respectively.

Table 5-3: Comparison of average underestimation percentages losses in both balanced and unbalanced models

Time Resolution (minutes)	Average Underestimation % Loughborough Data Set (Balanced)	Average Underestimation % Loughborough Data Set (Unbalanced)
5	6.4	14.1
10	8.8	21.8
15	11.2	25.4
30	19.3	29.9
60	23.7	36.3
120	26.9	40.1
Time Resolution (minutes)	**Average Underestimation % CLNR Data Set No.8 (Balanced)**	**Average Underestimation % CLNR Data Set No.8 (Unbalanced)**
5	6.4	12.8
10	9.0	23.5
15	11.2	28.7
30	18.4	34.9
60	25.9	38.3
120	29.0	40.8

Interestingly, as Table 5-3 shows, the underestimation percentages at each time intervals are more severe in the unbalanced network compared to the corresponding time interval in the balanced low voltage network model. For example, in the case of the balanced network model, for the sample dates from the Loughborough and CLNR no.8 data sets, the average underestimation percentage of the calculated losses from 1 to 15 minutes intervals is just above 11%. However, this figure changes to 25.4% and 28.7% underestimation for the Loughborough and CLNR no.8 data sets, respectively. A similar trend can also be observed at the remaining time intervals.

In the next section, methods are presented to estimate loss estimates at data set time resolution using the estimates at time resolutions of 30 minutes and/or higher.

5.3 Prediction of 1 minute Losses Based on Half-Hourly Loss Values

Investigations in sections 5.1 and 5.2 show that the values of technical network losses based on half-hourly averages of smart meter data, which is the way in which the DNOs in the UK receive them, are underestimated compared to the network loss values calculated using smart meter data at 1 minute time resolutions. This section presents a method to estimate the data set loss values based on loss values at lower time granularity of data which will be available to the DNOs such as loss values for 30 minute, 60 minute, and 120 minute time periods. This is carried out in two ways:

- by fitting the best curve to the loss values of each date that are available at higher granularity of 30, 60, and 120 and using the constants to estimate the data set value for each date
- by obtaining an average constant values from various day types in the same month and apply the constant to half-hourly loss values in order to obtain the data set losses.

For this study, the loss values for 52 sample dates from the CLNR data set no. 8 at various granularities of 1, 5, 10, 15, 30, 60, and 120 were selected and used in the model below to fit to the time interval width of **t** and losses data of ***L*** in equation 1 below:

$$(1)\ L = \beta \times t^{\alpha} \quad where\ \alpha \in (0, 1)$$

The value of α determines the shape of the loss curve, while β is the loss estimated at time resolution of data set. L is the predicted losses for a specific time resolution, t is the time resolution (e.g. 1, 5, 10, 15, 30, 60, and 120).

Given the losses L_1 and L_2 at times t_1 and t_2, then the value of α is calculated by:

$$(2)\ \alpha = \log (L_1/L_2) + \log (t_2/t_1)$$

Values of α and β are constant and are obtained by fitting the best curves to the loss values of various granularity for each date in Microsoft Excel Solver. Excel Solver determines the optimum value of α for each date based on the non-linear regression model described above using loss values at a time resolution of 30 minutes and lower. This method in Excel solver minimises the sum of squared error for the equation 1 and fits the bust curves to the data (Horton and Leonard 2005). The data set loss values are then calculated using either the average of α values for each day type (e.g. the average α values on four Fridays in April) from the 30, 60, and 120 minute loss values or using the α value for each specific date and using the 30 minute, 60 minute, and 120 minute loss values. Let the former be called $\overline{\alpha}$ and the latter be called $\widetilde{\alpha}$ from now on in this section.

The results of predicted data set losses based on both values of α are then compared to the actual measured curves in figures below. Figure 5-7 below shows the results of data set loss predictions based on $\widetilde{\alpha}$ based on half-hourly data using constant value of α obtained from the measured loss values at 30, 60, 120 time granularities on other similar day types in the sample dates.

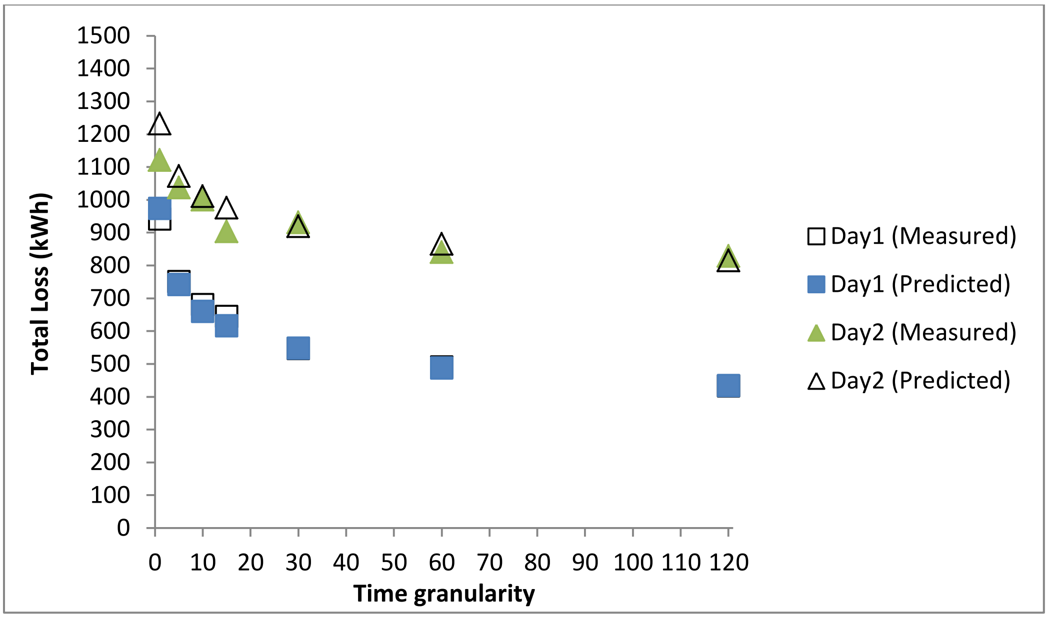


Figure 5-7: Predicting 1 minute losses based on $\tilde{\alpha}$ compared to the measured losses

Figure 5-7 compares the measured total network losses and the fitted curves of predicted total losses based on loss values at 30, 60, and 120 minute granularity of data for 05/04/2013 and 12/04/2013. The data set loss prediction value of 974 kWh for the sample date of 12/04/2013 is very close to the measured 1 minute loss of 941 kWh. In this case, the prediction method which uses the constant $\tilde{\alpha}$ value performs better than the method that uses $\overline{\alpha}$ that returns the value of 813 kWh. The prediction on 05/04/2013 is less close to the measured data set loss value of 1,121 kWh than the previous example, but still using the $\tilde{\alpha}$ value for this specific date returns a closer value (1,232 kWh) than using the average of all α ($\overline{\alpha}$) values for Fridays in April 2013 (1,383 kWh).

However, investigating the predicted losses for all of the sample dates indicate that using an average of α from each day type in each specific month ($\overline{\alpha}$) produced more accurate data set loss estimations compared to using α values for each date based on the correlation between lower time resolution intervals of 30, 60, and 120 minutes and the losses recorded at the time resolution intervals ($\tilde{\alpha}$). Figure 5-8 below confirms this observation by showing the spread of individual APE scores recorded for all the sample dates.

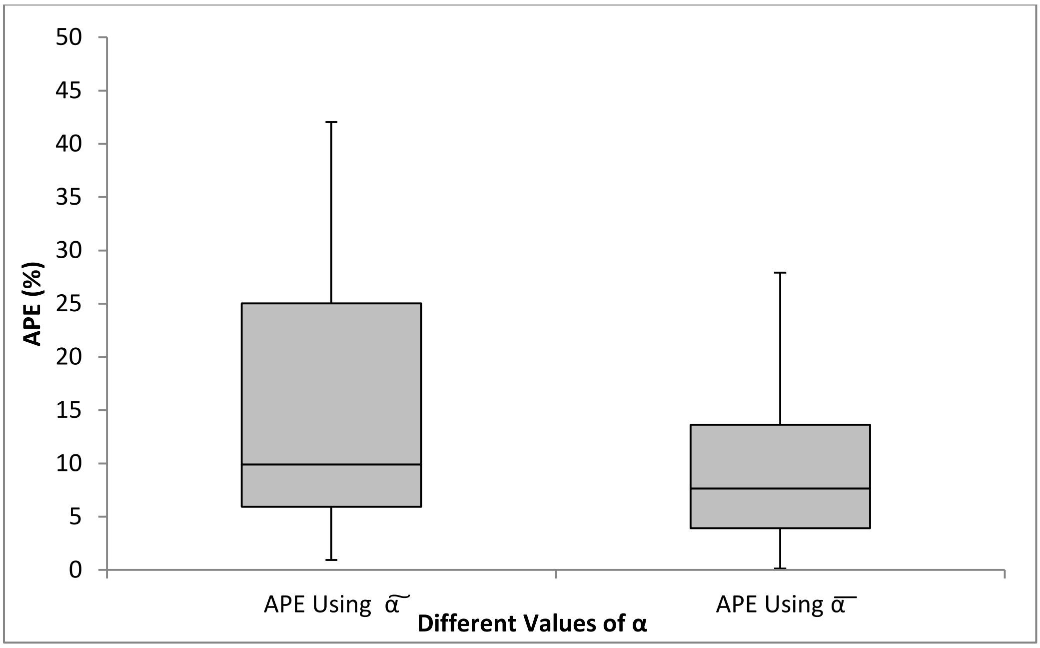


Figure 5-8: APE scores using different values of α

As Figure 5-8 shows, using the mean of α values from 4 sample dates similar to the target date ($\bar{\alpha}$) produces median APE value of 9.3%, across the 56 sample dates, with the minimum APE of 0.1% and the maximum APE of just under 28%. Eliminating the outlier values shows that a majority of prediction errors lie between the 1st quartile of just under 4% and the 3rd quartile of just under 14% APE values.

On the other hand, using an α values based the losses recorded on each sample date at lower time resolution intervals of 30, 60, and 120 minutes ($\tilde{\alpha}$) produces less accurate data set loss estimates on average compared to the previous method ($\bar{\alpha}$). In this case, the minimum APE error value is just under 1% and the maximum APE error value is just above 42%, with the 1st quartile value of just under 6% and 3rd quartile value of just above 25%.

5.4 Impact of Varying Smart Meter Time Resolutions on Estimation of Voltage Levels in a Balanced Low Voltage Network

Voltage level variations in a low voltage network can directly impact the power quality that the customers connected to that network experience. Hence, the DNOs in the UK have the statutory obligation to maintain the voltage levels within the thresholds of 230V +10%/-6% or they will be penalised by the regulator OFGEM. Maintaining the voltage levels within the statutory limits is becoming more challenging for the DNOs with the installation of higher proportions of embedded generation and low carbon technologies on the lower voltage levels of the distribution networks. This is mainly due to intermittent reverse power flows and new demand patterns being introduced to the low voltage network which is not generally observable to the DNOs.

In practice, as the distance from the low voltage substation increases, the voltage level on the network decreases. This is known as voltage drop and it can lead to lower quality of power (e.g. flickering light, etc.) supplied to the customers distant from the substation. This section examines the relationship between the various time resolution intervals of smart meter data and voltage levels estimated at the end of the balanced low voltage network model shown in Figure 3-8.

The voltage levels at the end of cables B and C are calculated for each of the three phases by subtracting the sum of all the voltage drops at each section of the low voltage network from the voltage level at the substation level. The maximum voltage drop subtracted from the voltage level at the substation shows the lowest voltage level recorded on the low voltage network on each sample date.

Figures 5-9 and 5-10 below, show the minimum voltage estimated, which occurs at the end of cable C, on the red phase for the sample dates from the Loughborough and CLNR data set no.8, respectively. The estimated voltages for each time resolution interval is displayed and the minimum voltage levels estimated for the customers on the yellow and blue phase can be found in Appendix C.

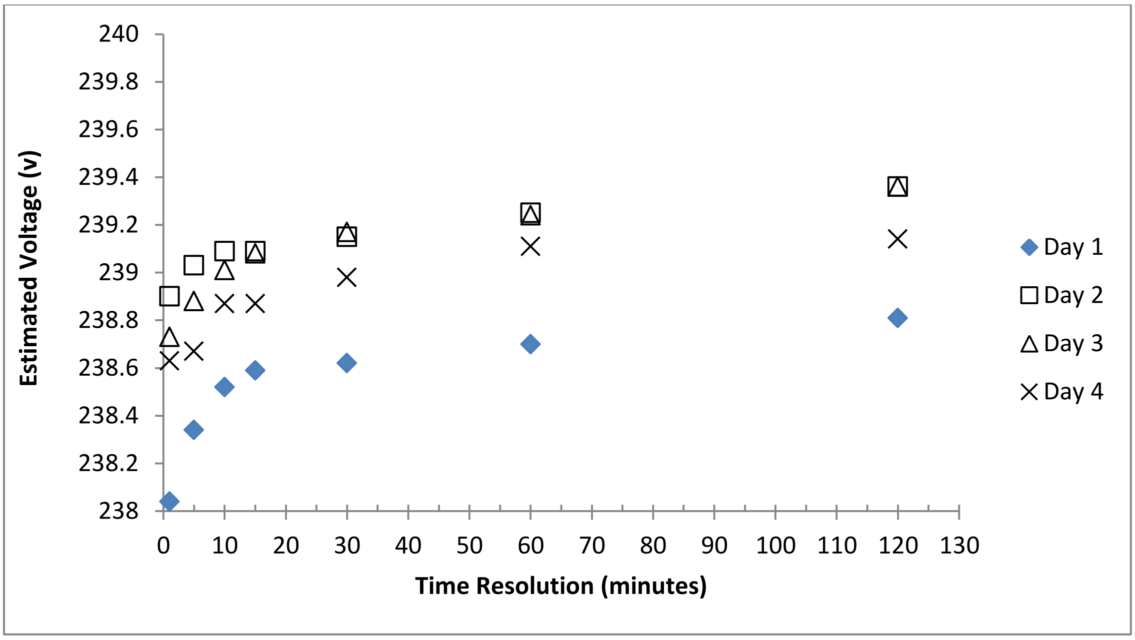


Figure 5-9: Minimum voltage levels on the red phase estimated using various smart meter time resolution intervals (Loughborough data set)

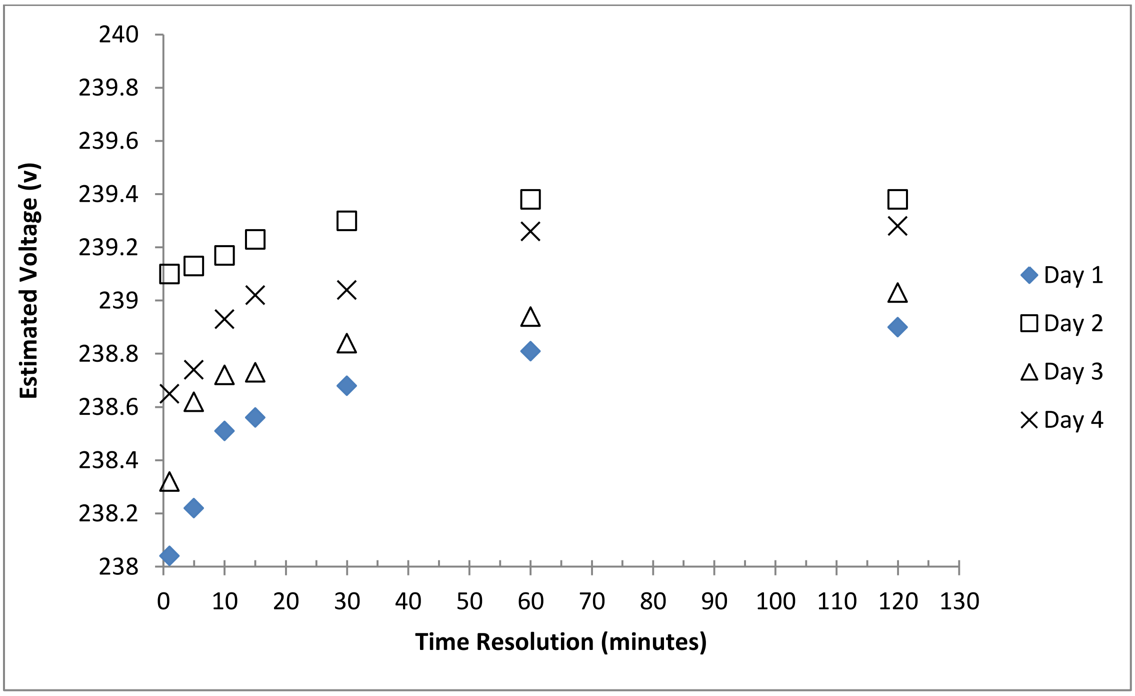


Figure 5-10: Minimum voltage levels on the red phase estimated using various smart meter time resolution intervals (CLNR data set no.8)

As Figures 5-9 and 5-10 demonstrate, the estimated minimum voltage levels on the red phase increase as the customer smart meter time resolution intervals decrease from data set to 120 minutes. For all of the 8 sample dates from both of the data sets, the overestimation mainly occurs when the customer currents at data set intervals are averaged to 30 minute intervals. As both Figures 5-9 and 5-10 show, the highest overestimation percentages take place in the first 15 minute averages. Table 5-4 shows the average overestimation percentage at each time resolution interval for the 8 sample dates from the two data set.

Table 5-4: Mean voltage overestimation percentage across the sample dates at each smart meter time resolution interval

Time Resolution (minutes)	Mean Overestimation Percentage (%) Loughborough Sample Dates	Mean Overestimation Percentage (%) CLNR Data Set no.8 Sample Dates
5	0.065	0.063
10	0.125	0.128
15	0.139	0.150
30	0.170	0.184
60	0.210	0.239
120	0.248	0.260

The major rise in overestimation occurs when the granularity of data decreases from data sets to 30 minutes, with the main increase between 1 and 15 minute time period, which is consistent in the case of all representative dates, and all phases of the network. The trend slows down when the granularity of data decreases from 30 minutes to 120 minutes. In the case of the sample dated from the Loughborough data set, the average overestimation from 1 to 15 minute time resolution intervals is 0.139%, but this overestimation percentage only rises by just over 0.040% when the smart meter time resolutions decrease from 15 to 30 minutes. A similar trend can be observed for the sample dates from the CLNR data set no.8. This indicates that although having smart meter data at 15 minute resolutions improves the voltage level estimations compared to having 120, 60, or 30 minute resolutions, the most sever inaccuracy occurs when the data is averaged from data set intervals to 15 minute intervals.

The overestimation of minimum voltage levels is caused by averaging the customer currents over time resolution intervals higher than data set time intervals. As shown in section 5.1, the average of currents over higher time resolution intervals flattens the spikes in the customer demands and in the power flow models leads to the underestimation of voltage drops at each section of the network. The underestimation of voltage drop in turn will contribute to the overestimation of minimum voltage levels.

Given the fact that this model network is not particularly overloaded and the network capacity is not pushed to the limits, the underestimation of voltage drops estimation can pose some major challenges to the DNOs, who have the statutory duty to keep the voltage variations on their network within acceptable limits. Also the overestimation of these minimum voltage levels can lead to the misestimating of embedded generation and low carbon technology capacity of each network branch since the DNOs will be likely to allow for less overvoltage expected from embedded generation and more under voltage expected from low carbon technologies installed in that particular section of the network.

5.5 Impact of Varying Smart Meter Time Resolutions on Estimation of Voltage Levels in an Unbalanced Low Voltage Network

This section presents the results from estimation of minimum voltage levels at various smart meter time resolution intervals in the unbalanced low voltage network setting described in section 5.2. In this model, the ratio of the 100 customers on the three phases of red, yellow, and blue was changed from 33:33:34 to 40:30:30 to create an unbalanced low voltage network model with more loads being allocated to the red phase.

The minimum voltage levels at the end of the network were calculated for all of the three phases. Figure 5-11 and 5-12 show the minimum voltage levels recorded on the red phase for the 8 sample dates from the Loughborough data set and the CLNR data set no.8. The results from the yellow and the blue phase which experience a lower voltage drop percentage can be found in Appendix D.

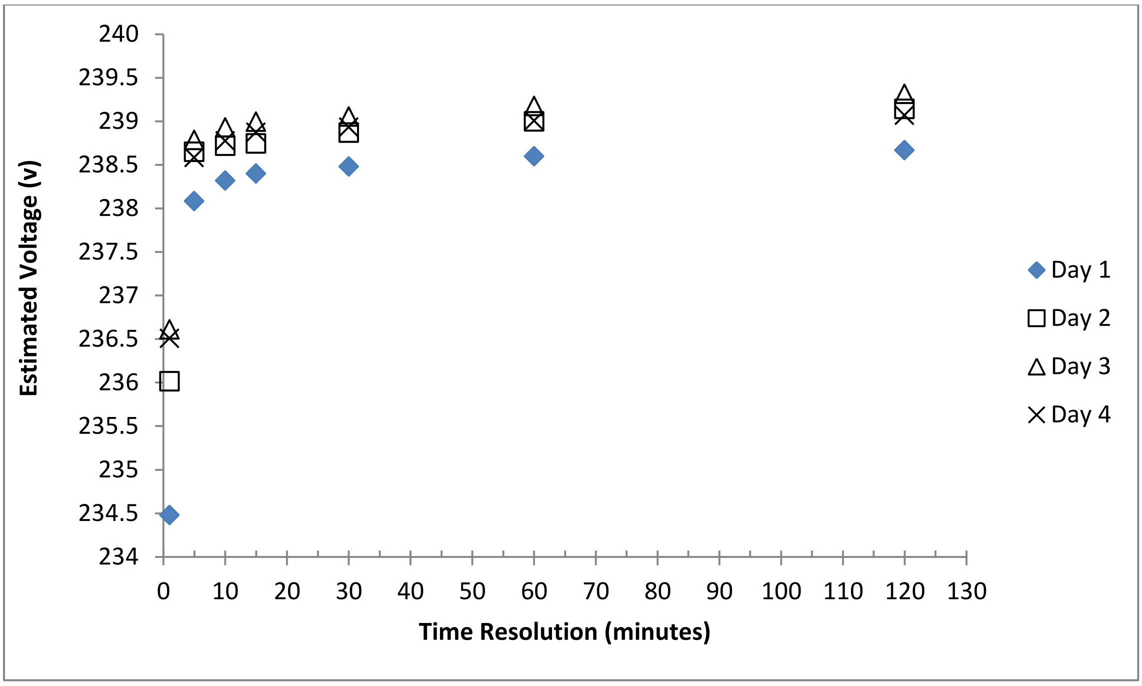


Figure 5-11: Minimum voltage levels on the red phase estimated using various smart meter time resolution intervals in an unbalanced network (Loughborough data set)

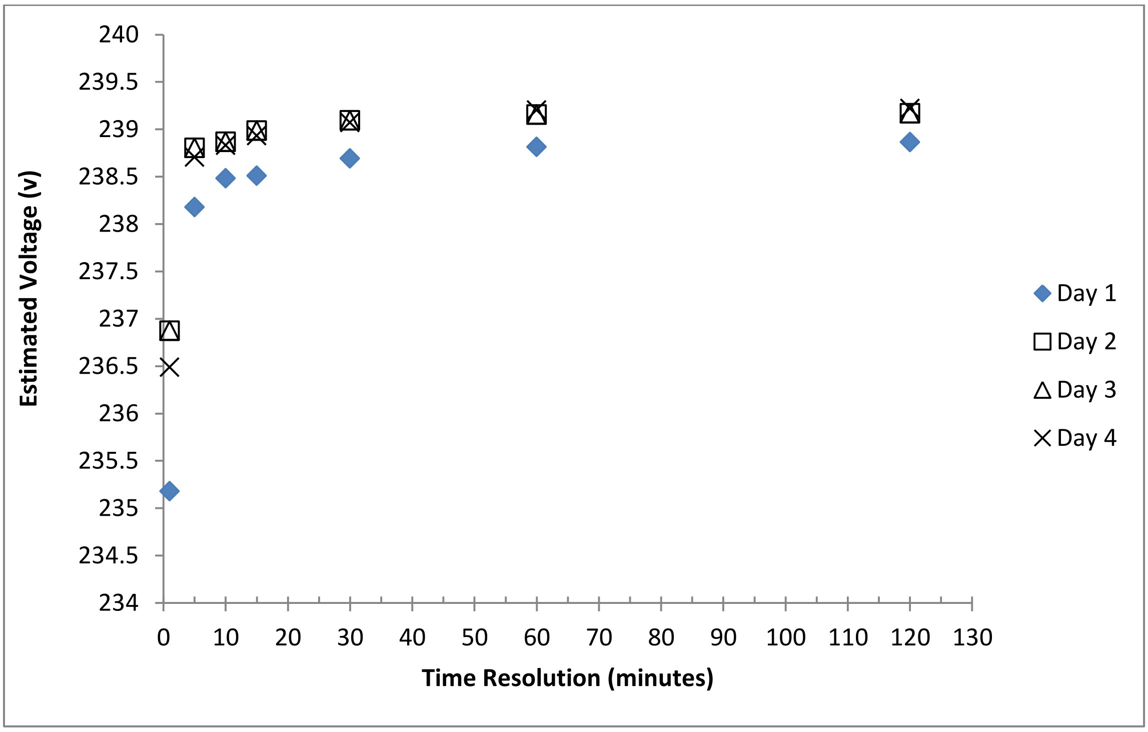


Figure 5-12: Minimum voltage levels on the red phase estimated using various smart meter time resolution intervals in an unbalanced network (CLNR data set no.8)

As Figures 5-11 and 5-12 show, the relationship between the estimated voltage levels and the time resolution intervals of smart meter data in unbalanced networks follow a similar trend to balanced networks. Clearly, as the time resolution of smart meter data decreases, the estimated minimum voltage figures on the red phase increase for all the 8 sample dates from both data sets. Table 5-5 below shows how the overestimation in this model compares to those of the balanced model presented in section 5.4.

Table 5-5: Comparison of mean voltage overestimation percentages between the balanced and the unbalanced models

Time Resolution (minutes)	**Mean Overestimation Percentage (%) Loughborough Sample Dates-Balanced Model**	**Mean Overestimation Percentage (%) Loughborough Sample Dates-Unbalanced Model**
5	0.065	1.112
10	0.125	1.180
15	0.139	1.209
30	0.170	1.243
60	0.210	1.290
120	0.248	1.334
Time Resolution (minutes)	**Mean Overestimation Percentage (%) CLNR Sample Dates no.8 Balanced Model**	**Mean Overestimation Percentage (%) CLNR Sample Dates no.8 Unbalanced Model**
5	0.063	0.961
10	0.128	1.020
15	0.150	1.059
30	0.184	1.116
60	0.239	1.155
120	0.260	1.165

As Table 5-5 shows, the overestimation of minimum voltage levels is the highest when the data set times resolution is averaged over 15 minute time intervals and the trend slows down from 15 minute time resolution intervals to 120 minute time intervals. Significantly, in the unbalanced model with higher loads on the red phase the overestimation percentages rise compared to the balanced model. For example, in the case of the Loughborough data set, the mean overestimation percentage going from 1 to 15 minute averages rises from 0.065% to 1.112%. In the case of the CLNR data set no.8, the mean overestimation percentage increases from 0.063% to 0.961% as the time resolution decreases from 1 to 15 minutes.

5.6 Impact of Varying Smart Meter Time Resolutions on Estimation of Low Voltage Cable Current Capacities

Knowledge of the capacity of low voltage underground cables to accommodate additional loads is becoming increasingly important as Smart Grid management solutions such as load shifting and Demand-Side Management (DSM) at community levels are becoming more widespread. The question is whether the real-time demand on each phase of the low voltage cables can be monitored using smart meter data and to what extent the information are improved or distorted as the granularity of smart meter data is decreased from data set to 120 minute averages. This knowledge will then determine the headroom that is available on each phase which can potentially facilitate more proactive load management of the low voltage network by the DNOs.

In this section, initially the maximum loads capacities of the low voltage cables A, B and C in the balanced model (Figure 3-8) are calculated using the cable specifications in Appendix B. Secondly, the model is populated with smart meter data recorded on the sample dates and the ratio of currents on each phase of the cable to the maximum capacity of the cables, is expressed in percentages over 24 hours on the sample dates. A similar procedure is then repeated using lower time resolutions of smart meter data up to 120 minute. This is only carried out for the balanced network model, because the cable capacities are calculated at the starting point of the cable and the imbalance in phases does not affect the relationship between the time resolution intervals and the cable capacity percentage estimates.

Figures 5-13 to 5-15 show the load curve for the 24 hour period on the sample date of 16/01/2008 from the Loughborough data set. The graphs show the load percentages on the red, yellow and blue phases of cable A, which experiences the highest load proportion compared to cables B and C, as the percentage of the red phase load capacity at time resolution ranging from data set to 120 minute intervals. The capacity of loading percentages of cables B and C are shown in Appendix E.

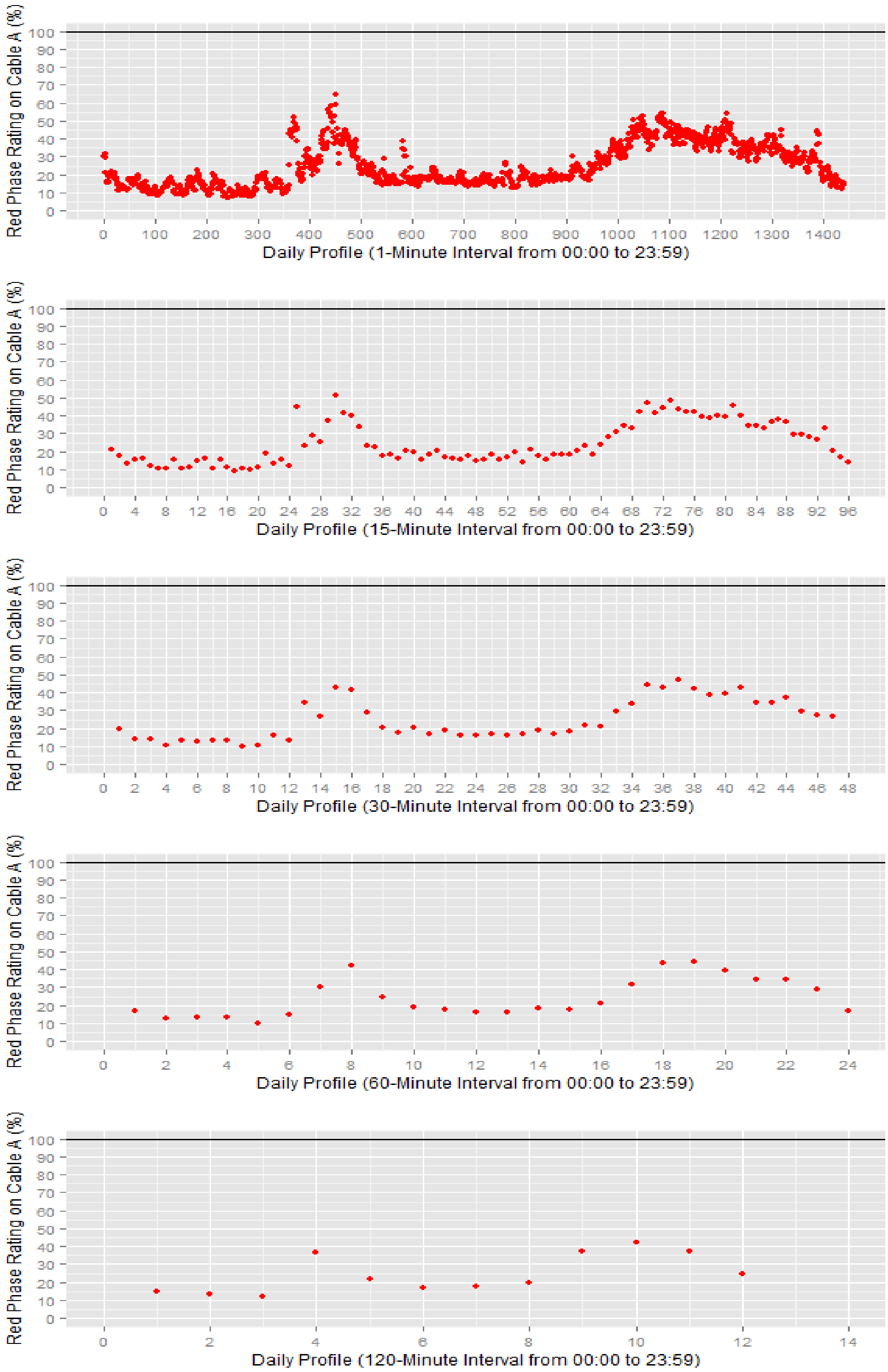


Figure 5-13: Red phase cable loading percentages on cable A (16/01/2008)

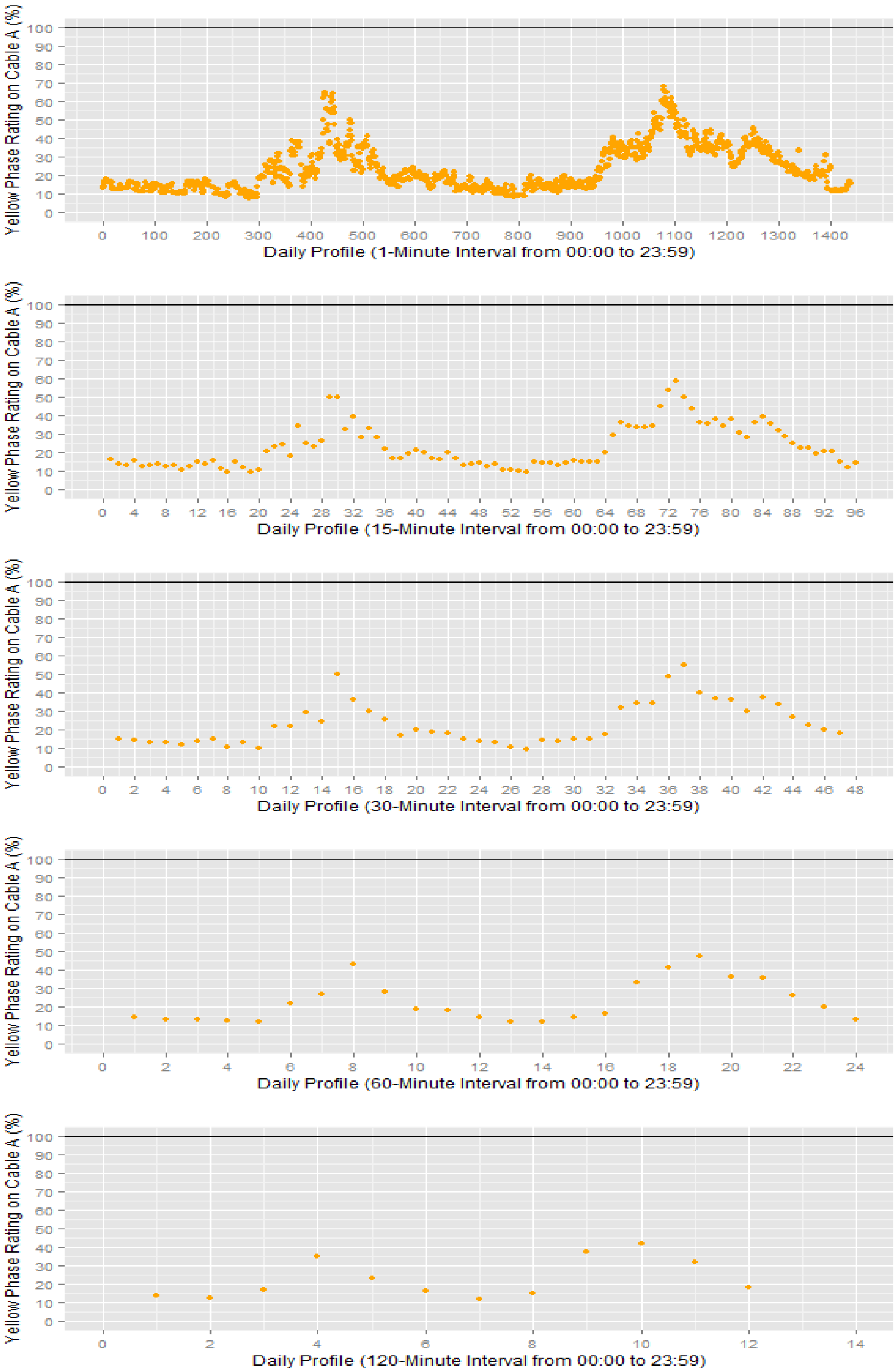


Figure 5-14: Yellow phase cable loading percentages on cable A (16/01/2008)

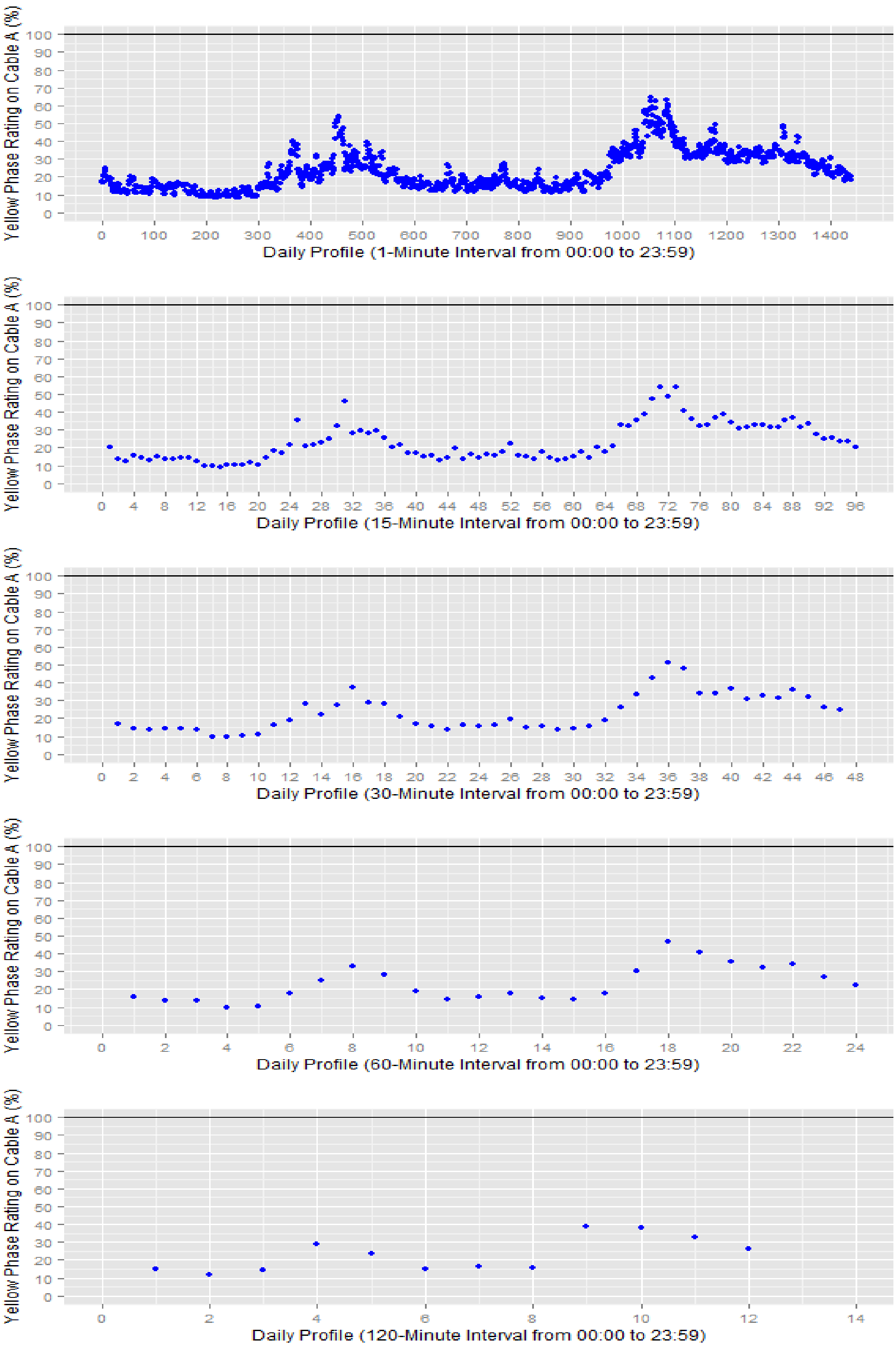


Figure 5-15: Blue phase cable loading percentages on cable A (16/01/2008)

Studying the above graphs of cable loading percentages on the underground low voltage cables show that although decreasing the granularity of data from data set to 120 eliminates much of the noise and clutter in the data, crucial information such as maximum load percentages at peak times are underestimated. This reconfirms the results found in sections 5.1 to 5.5 in that it shows to the extent to which averaging the time resolution intervals underestimates the peak customer loads.

The figures above show that while moving away from 30-minute data to 60 minute and then to 120 minute does not result in a sharp underestimation of peak loads, it still leads to some peak points being neglected. A very similar trend is also observed in the following graphs in Figure 5-16 to 5-18. These figures show the density plots of occurrence of each percentage point in the whole sample of data points for a particular time resolution interval.

The reason why these density plots were chosen was due to the fact that the data models of different granularities of smart meter data contain different numbers of data points. Therefore, a histogram would not be suitable for these data sets. This sample reflects all the data points recorded on all the representative dates from both data sets.

The Figure 5-16 to 5-18 show the density of various loading percentages at each time resolution interval. The (x) axis of the density plots show the low voltage cable loading percentages (0-100%) and the (y) axis show the fraction of data points in the overall data sample that represent a specific loading percentage. The 7 density plots in each figure represent the 7 time resolutions of the smart meter data from 1 to 120 minute averages of customer demands. The time resolutions are shown on top axis of each of the 7 plots.

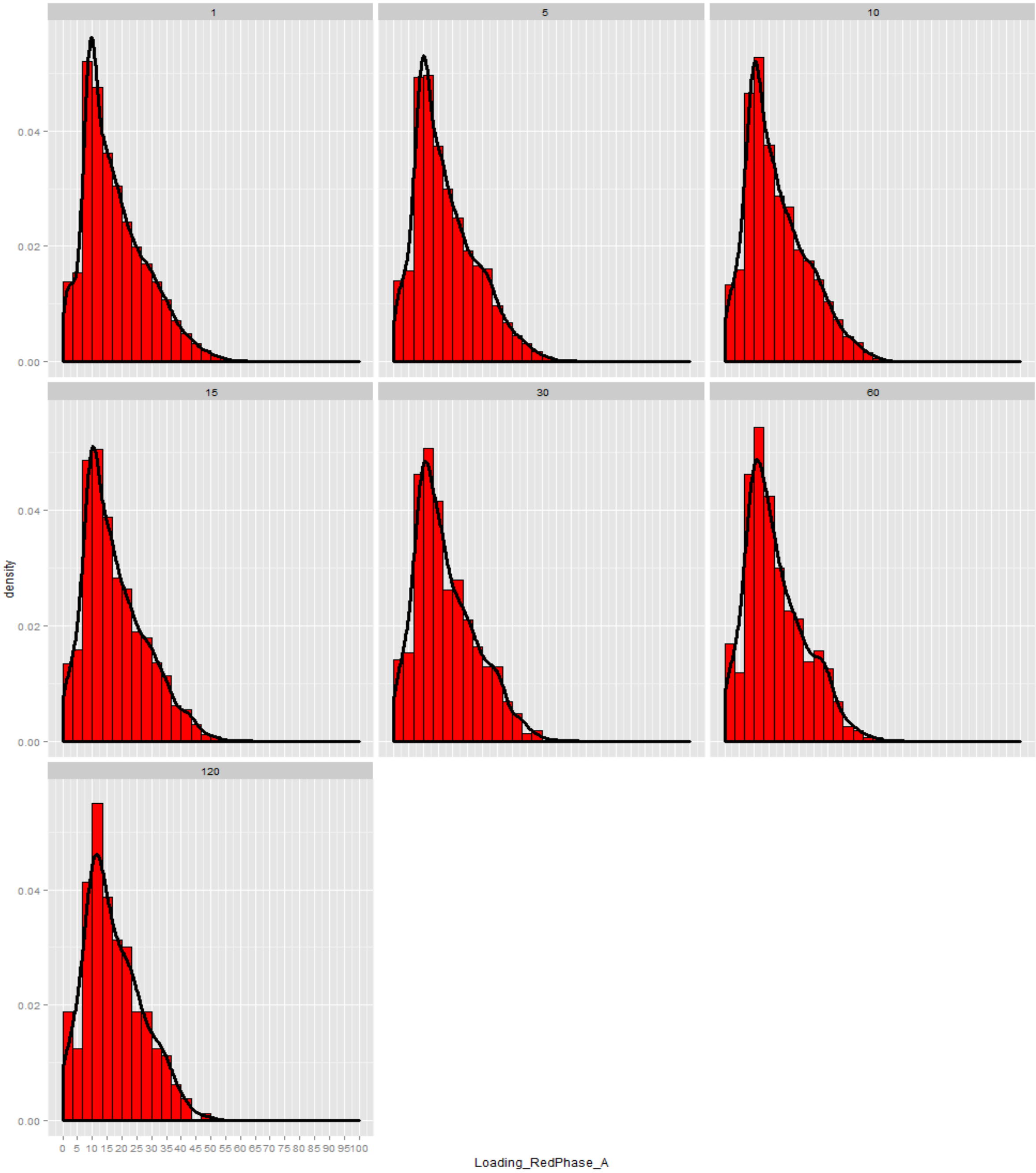


Figure 5-16: Density plots of loading percentages frequency at each time resolution interval on cable A (Red phase)

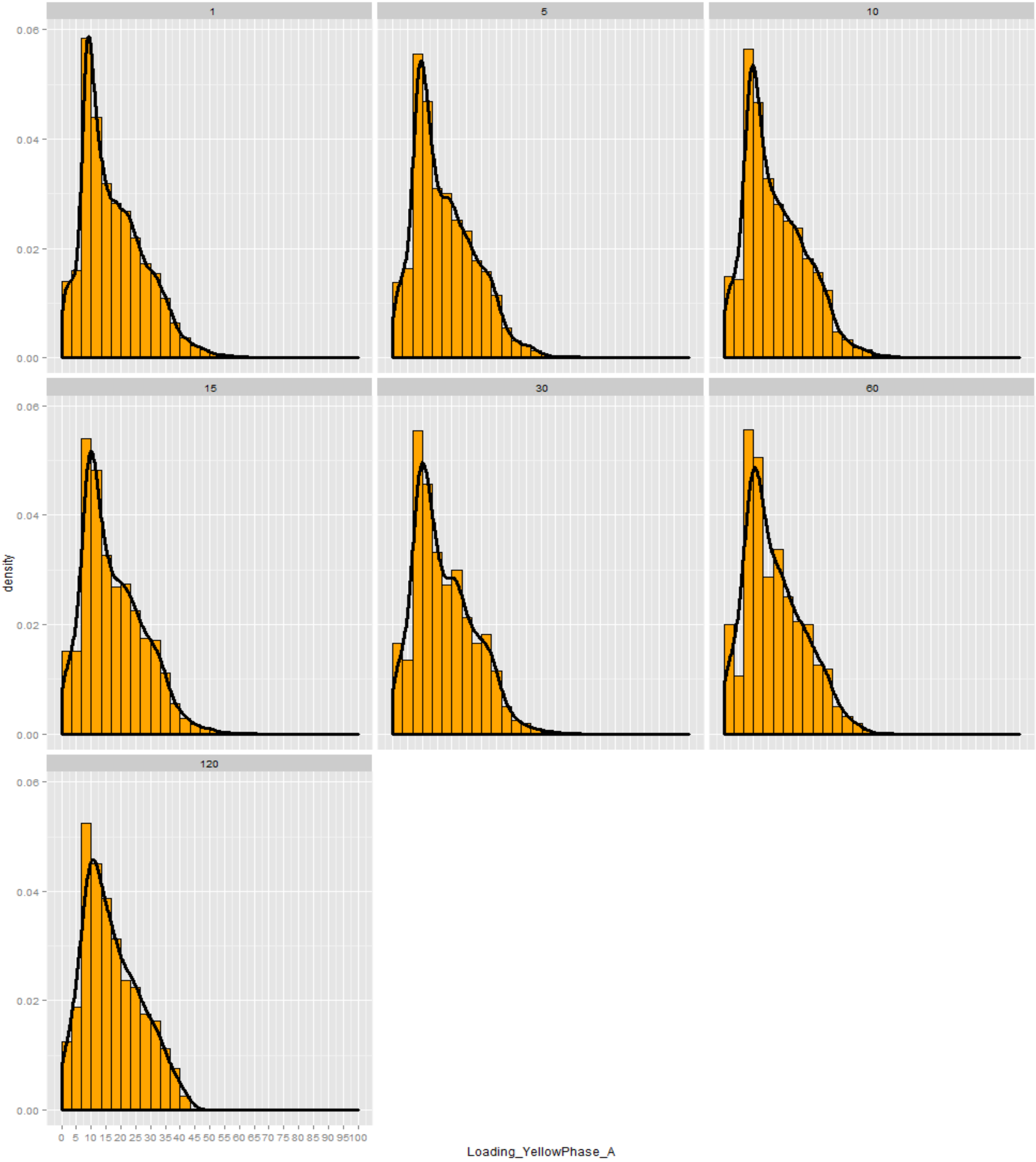


Figure 5-17: Density plots of loading percentages frequency at each time resolution on cable A (Yellow phase)

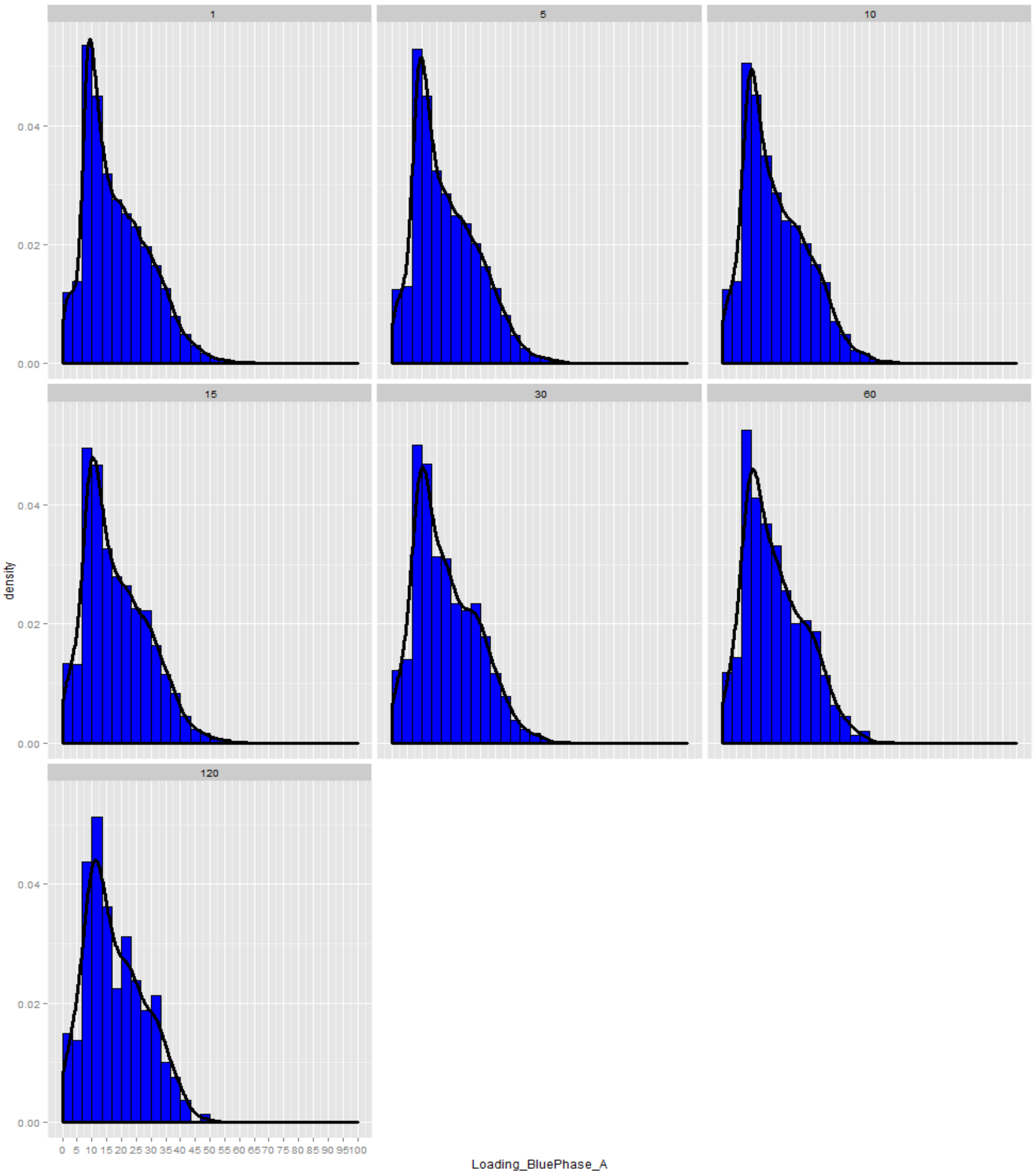


Figure 5-18: Density plots of loading percentages frequency at each time resolution on cable A (Blue phase)

Figure 5-16 to 5-18 show that as the time resolution of data decreases from 1 to 120 minute averages, the density of peak loading percentages decreases. This can be observed where the columns of higher percentages of cable loadings are eliminated in density plots for lower time resolution of data and the columns of lower percentages rise, which confirms the flattening of loads as the data are averaged to lower granularity.

As the granularity of data decreases from data set to 120 minutes, the frequency of middle data points such as 10-20% loading increases and the peak loading percentages such as 50-60% are neglected. This trend is also observed in Appendix F, but the effects are less dramatic on cables B and C, because they are less heavily loaded compared to cable A, so they experience less variety in peak demands.

5.7 Summary

In this chapter, the relationships between having smart meter data at various time resolution intervals and the estimation accuracy of critical low voltage network performance indicators such as losses, voltage levels, and cable capacity percentages are investigated. The main driver behind carrying out this analysis is that the smart meter data are transmitted to the DNOs at half-hourly averages, so it is very important to examine how accurate analysis driven from these half-hourly averages are compared to having high resolution data set smart meter data in providing the DNOs with appropriate levels of visibility and information.

In the first place, the effects of varying the smart meter time resolution intervals from 1 to 120 minutes on technical loss estimates are examined in a balanced and an unbalanced low voltage network model using readings from 100 customers on 8 different sample dates and from two different sets of data. Secondly, a model is devised to predict the data set losses when having loss estimated at time resolution intervals of 30 minutes and lower. This is carried out for 52 different sample dates from the CLNR data set no.8.

Thirdly, the impact of varying the smart meter time resolution intervals from 1 to 120 minutes on minimum voltage estimates on each phase are examined in the balanced and the unbalanced low voltage network models using the data from the 8 samples dates obtained from the two different data sets.

Finally, the effects of varying the smart meter time resolution intervals from 1 to 120 minutes on the cable capacity percentages of each phase of the balanced low voltage network model are examined.

5.8 Conclusions

These series of analyses in sections 5.1 and 5.2 demonstrated that as the time resolution intervals of smart meter data is decreased from 1 minute to half-hourly and 120 minute averages, the network loss estimations become less accurate with the sharpest decrease in accuracy occurring from data set to 15 minute intervals (Poursharif et al. 2017). In the balanced model the mean underestimation percentage when moving away from data set to 15 minute time resolutions is about 11% in the balanced model and this changes to 26% and 29% in the unbalanced model for the samples dates from the Loughborough data set and the CLNR data set no.8, respectively. The underestimation trend can also be seen when moving away from 15 minute time resolution intervals to 30 minutes. However, this trend slightly decreases compared to the first 15 minutes. The mean underestimation percentage in the balanced model is just under 20%, but this changes to approximately 30% and 35% underestimation, in the unbalanced model, for the sample dates from the Loughborough data set and the CLNR data set no.8, respectively.

This was an improvement on methods used in Oliveira and Padilha-Feltrin (2009) and Quiroz et al. (2012) that use loss load factors to estimate losses on model low voltage networks. It also improved on the work carried out by Brandauer et al. (2013) that highlight the effects of short-term high demands on loss estimates. Our studies were also improvements on top-bottom loss predictive methods in the absence of detailed data used in Dashtaki and Haghifam (2013) at medium voltage levels. More importantly, our studies improved the model used by Urquhart and Thompson (2015) by highlighting the impact of smart meter time resolution intervals from 1 minute to 120 minutes on the estimation accuracy of losses in a balanced and an unbalanced 100 house three phase network models.

The analysis in section 5.3 showed that a regression model to fit the best curve to the estimated losses at 30, 60, and 120 time resolution intervals can be used in predicting the actual data set estimated losses. Fitting this model to the losses calculated on the sample dates, produced MAPE of just under 10% which indicates a 50% improvement instead of having loss estimates at 30 minute time resolution intervals that on average underestimate the data set losses by just under 20%. This was a novel approach that has encouraged the DNOs to invest in load forecasting models to overcome the information gap resulting from having lower time resolution of smart meter data (Northern Powergrid 2016).

The series of analyses in sections 5.4 and 5.5 were never done before and demonstrated that as the time resolution interval of smart meter data decreases from 1 to 120 minutes, the minimum voltage levels experienced on each phase of the balanced and the unbalanced low voltage network is overestimated. Similar to the estimation of losses, the most severe overestimation occurs in the first 15 minutes. Our analysis show that moving from 1 to 15 minute averages overestimates the minimum voltage levels on the red phase in the balanced network by about 0.15%. While this underestimation percentage might not see significant, the unbalanced network shows that this figure changes to just over 1% when the load on the red phase become heavier in the unbalanced low voltage network model.

The results from sections 5.1 to 5.5 show that as the time resolution of smart meter data is decreased from data set to 120 minutes, the demand peaks are flattened, hence the network losses are underestimated and voltage levels are overestimated. This is clearly observed in section 5.6. In the case of cable loading percentages, it was demonstrated that with the decrease of the smart meter data granularity major demand peaks are flattened, which both hides great benefits of the smart meter data and also contributes to the underestimation of losses and overestimation of voltage levels.

The results in this chapter show that half-hourly smart meter data do not provide the DNOs with accurate estimates of critical low voltage network information such as losses, voltage levels, and cable loading percentages. Our analysis also shows that even having higher resolution of smart meter at 15 minute intervals fail to provide accurate low voltage network information to the DNOs. Regression models can be used to predict 1 minute losses and improve the half-hourly estimates by approximately 50%, but previous studies by Urquhart and Thompson (2015) and Urquhart et al. (2017) show that even 1 minute smart meter data underestimate the estimated losses compared to having smart meter data at 1 second time resolution intervals or when the losses are calculated using the difference between output and input power on transmitted through the low voltage network.

The lack of inaccuracy in the estimation of critical low voltage information resulted from having lower time resolution of data can affect a number of different operation applications. The underestimation of losses can hide the areas of the low voltage network that require reconfiguration or reinforcement. It also fails to pinpoint the areas of the network that are inefficient. The overestimation of minimum voltage levels can misestimate the capacity of the network to host embedded generation or low carbon technologies and can also fail to identify the areas of the network in which the power quality delivered to the customers is affected. It can also fail to identify the areas of the low voltage network which are likely to experience faults. The underestimation of peak demands in the low voltage cables can also adversely affect the ability of the network operators to carry out applications such as Demand-Side Management and Active Network Management.

Chapter 6 Studying the Relationship Between Smart Meter Data Aggregation and Important Low Voltage Network Performance Indicators

This chapter investigates the effects of aggregating customer loads at the low voltage network level on the accuracy of important low voltage network estimates such as losses and voltage levels. The approaches employed in this thesis are based on the requirement for the DNOs to anonymise the individual customer data as soon as they receive them from the DCC. It is believed that aggregation of customer demands is the most efficient and cost effective method of preserving individual consumer lifestyle information at low voltage levels, compared to methods such as using encryption methods and data aggregators at higher levels of the network (EA Technology 2015a).

Aggregation studies in this chapter is carried out by grouping the half-hourly smart meter data from customers at 2, 4, 6, 8, and 10 houses and comparing the loss and minimum voltage estimates with the loss and minimum voltage estimates at 30 minute intervals when no aggregation takes place. In the UK, the DNOs receive the smart meter data of individual customers at half-hourly intervals. The customers are grouped together based on the phasing and proximity, i.e. half-hourly loads from 2 customers that are closest and are on the same phase are added together to form new demand points on the low voltage network model. The aggregation points are presented in Figure 3-9. The aggregation studies are carried out in a balanced network model and an unbalanced network model using smart meter data on 4 sample dates from the Loughborough data sets and 4 sample dates from the CLNR data set no.8. The sample dates are presented in Table 6-1.

Table 6-1: The 8 sample dates from the two data sets used in the aggregation studies.

Sample Day	Loughborough Data Set	CLNR Data Set no.8
Day 1	Wednesday 16/01/2008	Saturday 12/01/2013
Day 2	Wednesday 02/07/2008	Wednesday 12/02/2013
Day 3	Wednesday 09/04/2008	Wednesday 10/04/2013
Day 4	Saturday 06/09/2008	Wednesday 20/02/2013

The various levels of aggregation are also tested in two different network topologies. Initially, the network topology presented in Figure 3.8, which has three cables with two branches, is employed for comparison of loss and voltage estimates. Secondly, the network topology is changed to the network model arrangement in Figure 3.10 where all the 100 customers are connected to a single main cable with specifications of cable A.

6.1 Effects of Varying Customer Aggregation Levels on Loss Estimation in Balanced Low Voltage Networks

In this section, the results of network loss estimates using 30 minute time resolution intervals at each aggregation level are compared with zero aggregation (1-house) results using scatter plots in the following Figures 6-1 and 6-2, which represent the values of losses for each aggregation level on the sample dates. Figure 6-1 shows the loss estimates at 1, 2, 4, 6, 8, and 10 house aggregation levels on 4 sample dates from the Loughborough data set. Figure 6-2 shows the loss estimates at similar aggregation levels on the 4 sample dates from the CLNR data set no.8. These figures show the estimated losses at each aggregation level as a ratio of losses estimated at half-hourly intervals with no aggregation (1-house).

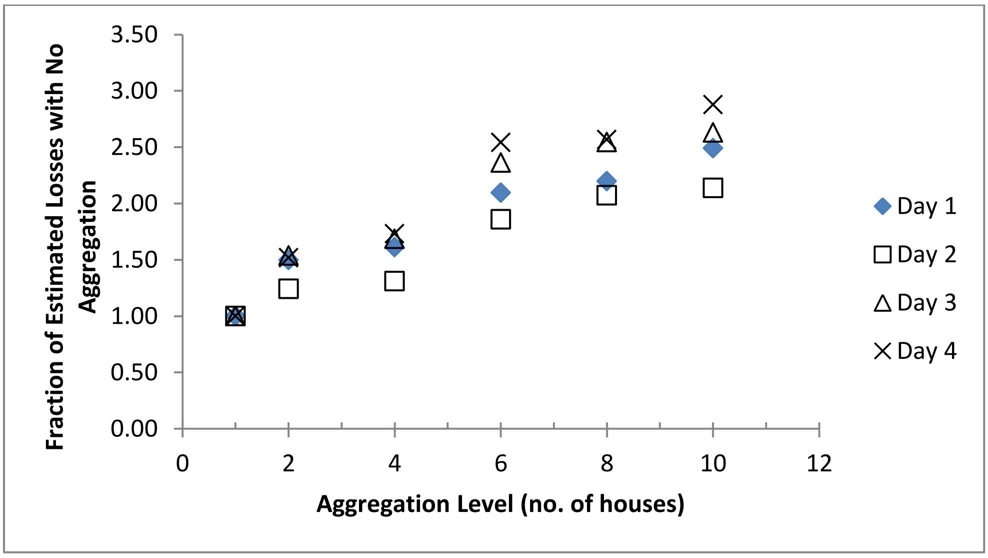


Figure 6-1: Fraction of estimated losses for different aggregation levels (Loughborough data set)

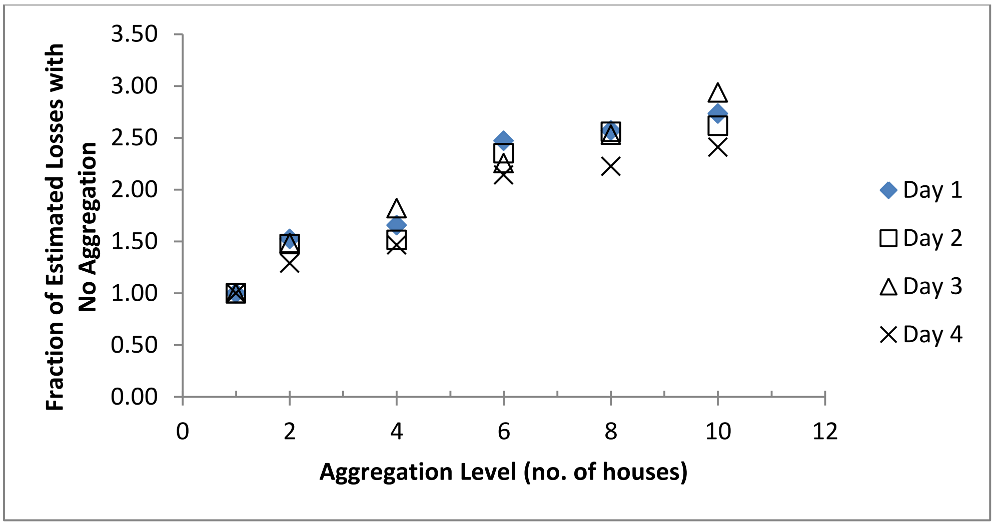


Figure 6-2: Fraction of estimated losses for different aggregation levels (CLNR data set no.8)

Figures 6-1 and 6-2 show that that as the customer loads are aggregated, the network loss estimate values also increase. For both data sets, there is a significant rise in loss estimates when the loads from 2 customers are added together. In the case of the sample dates from Loughborough data set, the increase in loss estimates at 2-house aggregation level ranges from just under 30% to approximately 50% and in the case of the sample dates from the CLNR data set no.8 it ranges from just under 40% to approximately 50%. The rate overestimation in losses then slows down from 2-house aggregation to 4-house aggregation level on all 8 sample dates. Moving from 4-house to 6-house aggregation causes another large increase in loss estimates of about under 100% and this trend is then slows down when the data from more houses are added together in 8 and 10 house aggregation scenarios.

The average of overestimation percentage for each aggregation level across the 8 sample dates are presented in Table 6-2.

Table 6-2: Average overestimation in loss estimates at different aggregation levels

Aggregation Level	Average Overestimation % Loughborough Data Set	Average Overestimation % CLNR Data Set no.8
2-house	44.9	44.2
4-house	58.5	61.3
6-house	121.3	130.3
8-house	134.5	146.9
10-house	153.4	167.2

Table 6-2 shows that on average there is just under 45% loss overestimation at the first stage of aggregation from no aggregation to 2-house level. This trend then slows down to a steadier level as more customers are aggregated with another large increase of just under 70% overestimation between 4-house to 6-house levels of aggregation. The overestimation trend slows down again from this point to the 10-house aggregation level. It can be observed from the figures and tables above that as more customer demand data are added together the network loss estimation increase. The sharpest rise in levels occur between no aggregation and 2-house aggregation and 4-house and 6-house aggregation levels.

This is due to the fact that in the network models with no customer data aggregation, the network losses on the main three phase low voltage cables are calculated by using the load of a single customer on each phase and 5 meter section lengths of the low voltage cable where that load is connected to leading to the next customer connection, which means that the losses at each section are then added together to calculate the total loss of the network (see section 3.3). However, when aggregation points are placed on the network model (as in Figure 3-9), the customer demands from two neighbouring houses on a similar phase are added together at the middle point between the two houses. The changes in cable lengths in the aggregation models and aggregated loads contribute to higher loss estimate values. For example, Figure 6-3 shows a representation of how the losses on the main cable are calculated for customers on the red phase in Figure 3-8 (no aggregation) and 3-8 (with aggregation points).

Using the formula for calculation of losses in a load flow model that uses I^2 (square of the current) and the resistance on the main cables, the losses on the red phase is calculated by:

$$L = I^2 \times r$$

Resistance (r) is calculated using the length of the cable that a demand point is connected to and the impedance of the cable. This equates to 0.0002 Ω per meter of main cable length.

Figure 6-3 shows the cable length and the demand points in a network model with no aggregation and a network model with 2-house aggregation demand points for customers on the red phase.

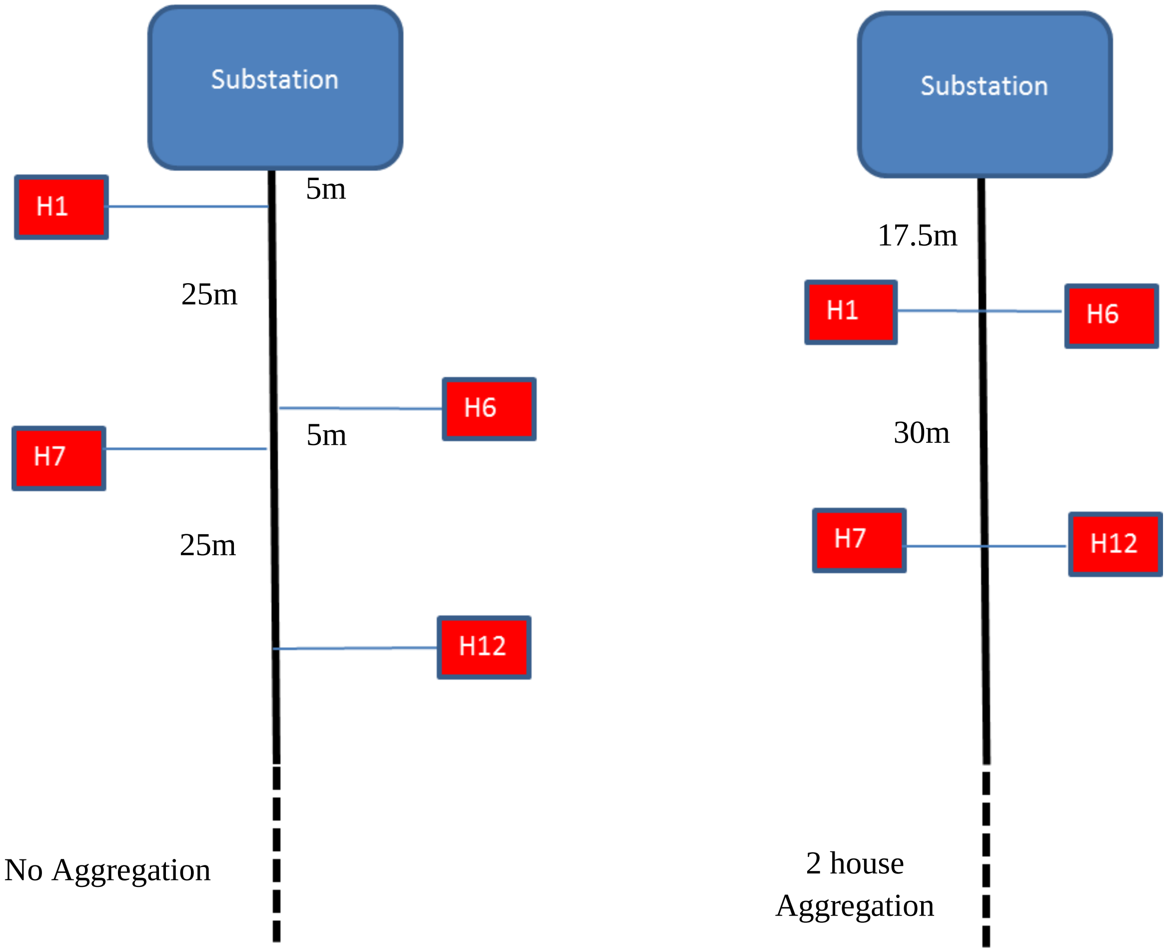


Figure 6-3: A representation of the placement of demand points and the changes in cable lengths in 2-house aggregation scenarios

Let us assume that all the customers in Figure 6-3 have the current of 2amps in one of the 48 half-hours on one of the sample dates. Using the customer currents and the cable resistance and cable lengths, the losses calculated for the 4 customers on the red phase in the two scenarios can be calculated as follows, where L_0 and L_2 denote estimated half-hourly losses with zero and 2-house aggregation, respectively:

$$L_0 = (4\times5\times0.0002) + (4\times25\times0.0002) + (4\times5\times0.0002) + (4\times25\times0.0002) = 0.047$$

$$L_2 = (16\times17.5\times0.0002) + (16\times30\times0.0002) = 1.511$$

As can be seen in the example above, despite the decrease in the cable lengths and therefore the resistance in the aggregation model, the fact that the square of the aggregated currents are used in the load flow loss calculation models contributes to the overestimation of losses. This overestimation trend is then continued as more customer demands are added together as observed in Figures 6-1, 6-2, and Table 6-1. One of the main reasons for the higher increase in network estimates at 6-house aggregation and beyond is that in the case of 6-house, 8-house, and 10-house some of the aggregation points are placed on cables B and C which are low voltage cables with smaller diameters compared to cable A and therefore in these aggregation scenarios there is higher demands on cables with higher resistance which contribute to higher rate of loss estimate values.

Alternative Network Topology

In order to further investigate the impact of the placement of aggregation points on the loss estimate values, the network model represented in Figure 3-8, which is a 100-house balanced three phase low voltage network with three cables and two branches, is changed to a 100-house balanced three phase low voltage network with only one main cable and no branches. This new model is presented in Figure 3-10. In theory, this should minimise the influence of aggregation point placements on the accuracy of loss estimates.

Figures 6-4 and 6-5 below show the loss estimates at various aggregation levels for the sample dates from the Loughborough and CLNR data set no.8 data sets. The loss estimates are expressed as the fraction of loss estimates when there is no smart meter data aggregation of the customers.

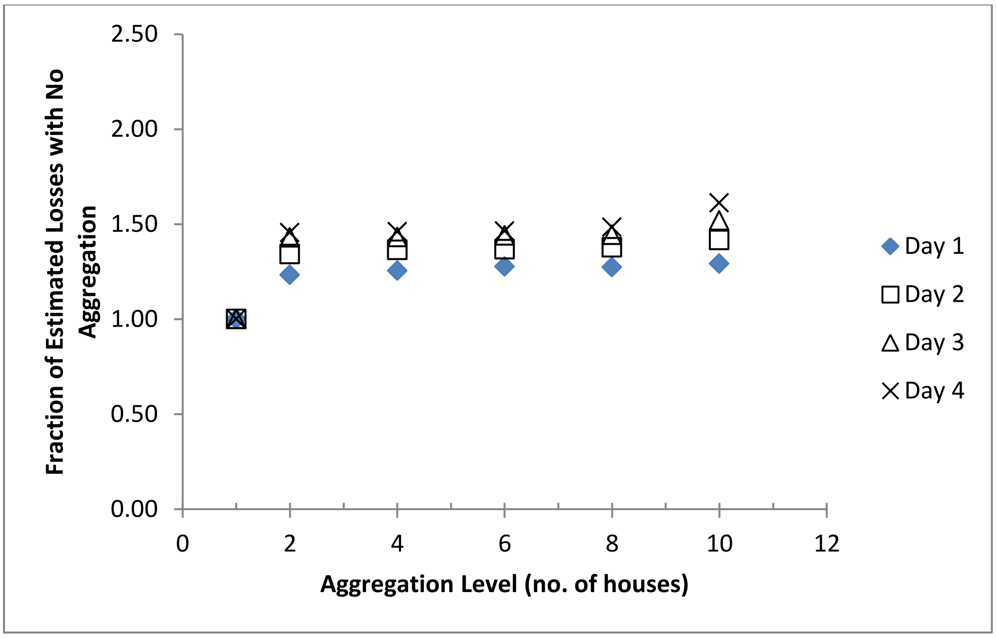


Figure 6-4: Fraction of estimated losses for different aggregation levels-alternative topology (Loughborough data set)

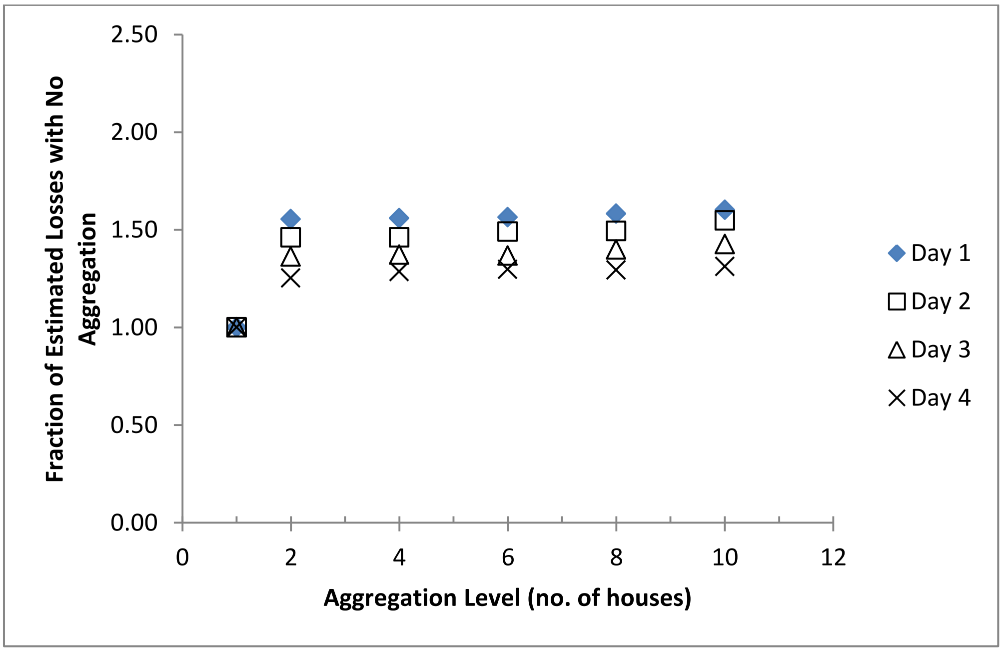


Figure 6-5: Fraction of estimated losses for different aggregation levels-alternative topology (CLNR data set no.8)

Figures 6-4 and 6-5 show that in this alternative network arrangement the largest increase in the loss estimates take place when the data from two customers are grouped together and further aggregation only slightly increases the estimates. The rise in estimated losses at 2-house aggregation level ranges between 25-50% across the 8 sample dates. Table 6-2 below shows the average overestimation percentages at each level of aggregation for the sample dates from the two data sets.

Table 6-3: Average overestimation in loss estimates at different aggregation levels-alternative topology

Aggregation Level	Average Overestimation % Loughborough Data Set	Average Overestimation % CLNR Data Set no.8
2-house	36.55	40.80
4-house	37.75	42.00
6-house	38.72	42.97
8-house	39.46	44.21
10-house	46.00	47.25

The overestimation percentages in losses now follow a more logical pattern compared to the earlier model. The highest overestimation of the percentage of losses occur when customer data from two customers are added together with the average of 36.55% and 40.80% increase in loss estimates for the sample dates from the Loughborough data set and the CLNR data set no.8, respectively.

After this point, when customer data from 4, 6, 8, and 10 customers are added together, the overestimation percentage only rises marginally by about 2% for each level of aggregation on average, up to the 8-house aggregation level and between 3% to 7% rise in the estimate values on average from 8-house to 10-house aggregation levels. This shows that if all the households are served by a low voltage cable network or with similar cables with similar impedance characteristics, the aggregation of customers produces a more predictable trend in the results.

The difference in results is mainly due to the placement of the aggregation points and the topologies of the low voltage network models. Once the customer demands are added together on the various aggregation levels, the sum of demands is the same, so the only contributing factor to the changes in loss estimates of the low voltage network are the square of load values and different the cable section lengths used in network loss estimation based on the new aggregation points.

6.2 Effects of Varying Customer Aggregation Levels on Loss Estimation in Unbalanced Low Voltage Networks

In order to study the effects of customer data aggregation on the accuracy of loss estimates, the balanced low voltage network presented in Figure 3-8 was changed to an unbalanced low voltage network with a higher number of customers connected to the red phase. On the unbalanced low voltage model, the ratio of customers on the three phases has been changed from 33:33:34 to 40:30:30 on the three phases of red, yellow, and blue respectively. The smart meter data from all the 8 sample dates that were used in section 6.1 are also used in this unbalanced model.

Figures 6-6 and 6-7 below show the estimated losses at various levels of aggregation in the unbalanced low voltage network model, using smart meter data recorded on the 8 sample dates from the Loughborough data set and the CLNR data set no.8. The losses are expressed as a fraction of estimated losses with no aggregation (1-house).

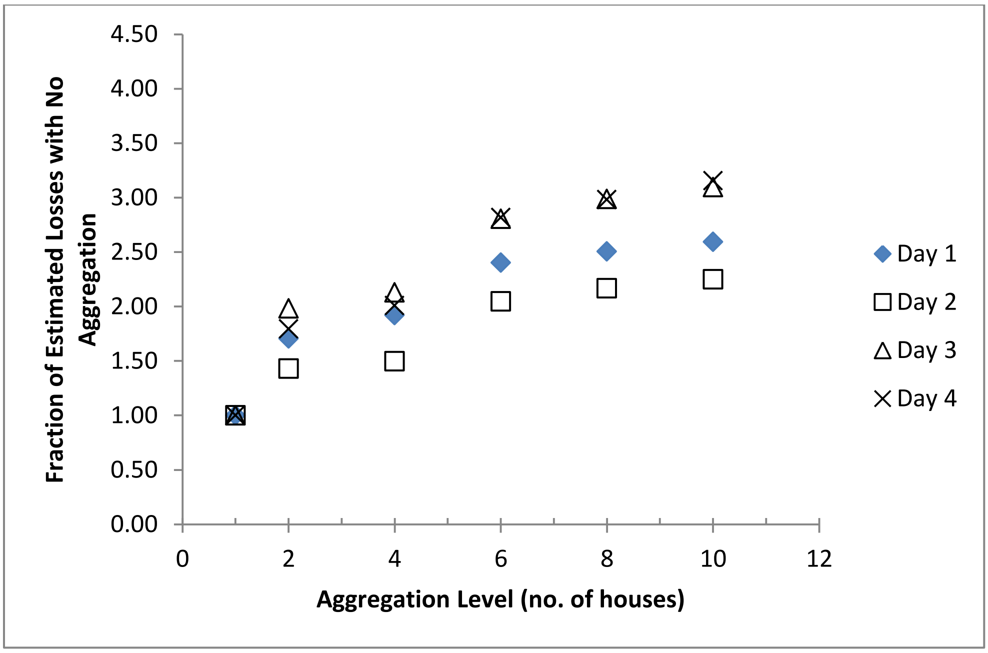


Figure 6-6: Fraction of estimated losses for different aggregation levels-unbalanced model (Loughborough data set)

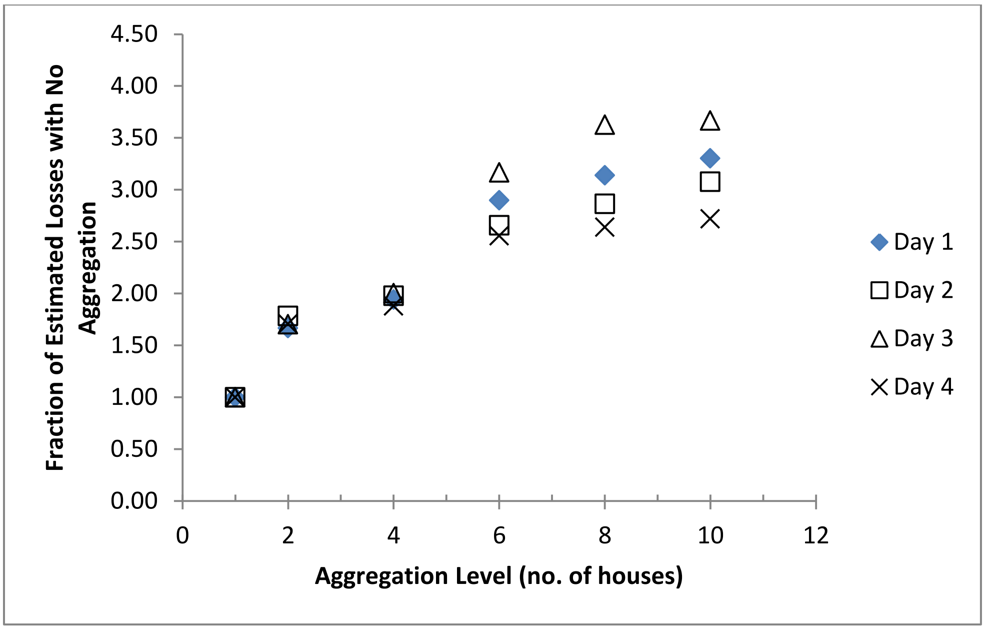


Figure 6-7: Fraction of estimated losses for different aggregation levels-unbalanced model (CLNR data set no.8)

Figures 6-6 and 6-7 show that a similar trend to the balanced network model occurs as the smart meter data are aggregated in the unbalanced low voltage network with the highest rates of increase in the loss estimates taking places at 2-house and 6-house aggregation levels. However, as Table 6-4 indicates the overestimation percentages are further increased in the unbalanced model compared to the balanced low voltage network model. Table 6-5 compares the average overestimation percentages between the balanced and the unbalanced model at each aggregation level.

Table 6-4: Comparison of overestimation percentages at various aggregation levels

Aggregation Level	Average Overestimation % Loughborough Data Set-balanced	Average Overestimation % Loughborough Data Set-unbalanced
2-house	44.93	72.69
4-house	58.53	88.86
6-house	121.39	151.72
8-house	134.52	165.97
10-house	153.40	177.30
Aggregation Level	**Average Overestimation % CLNR Data Set no.8-balanced**	**Average Overestimation % CLNR Data Set no.8-unbalanced**
2-house	44.20	71.32
4-house	61.38	95.00
6-house	130.34	181.92
8-house	146.92	206.63
10-house	167.24	219.08

As Table 6-4 demonstrates, in the unbalanced low voltage network model the average overestimation percentage at 2-house level increases by just under 30% compared to the balanced model for the sample dates from both data sets. This figure changes to just over 30% for the sample dates from the Loughborough data set and about 60% for the sample dates from the CLNR data set no.8 when data from 6 customers are added together. The overestimation trend is consistent with the trend observed in section 6.1, where the largest overestimation percentages of loss estimates take place at 2-house and 6-house aggregation levels. However, the overestimation percentages are more severe in the unbalanced network model. This is caused by the extra load inserted by the customers on the red phase in the unbalanced network model compared to the balanced network model, while the length of the cable remains unchanged.

Alternative Network Topology

A similar trend is also seen when the network topology is changed and all the 100 customers are connected to one main cable in the alternative network model. Figures 6-8 and 6-9 below show the estimated losses at various aggregation levels, for the 8 sample dates, as a fraction of estimated losses with no aggregation.

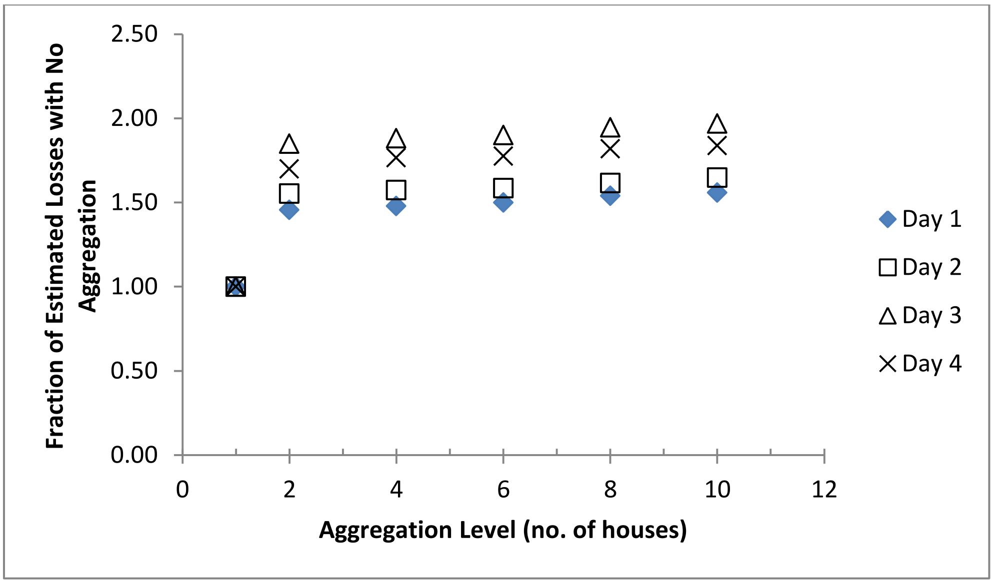


Figure 6-8: Fraction of estimated losses for different aggregation levels-unbalanced model–alternative topology (Loughborough data set)

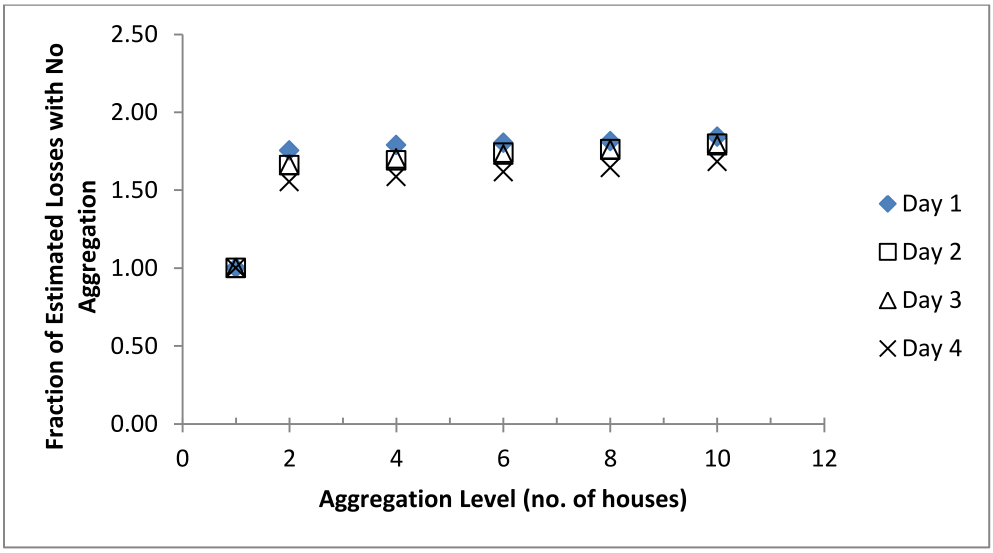


Figure 6-9: : Fraction of estimated losses for different aggregation levels-unbalanced model–alternative topology (CLNR data set no.8)

Figures 6-8 and 6-9 demonstrate that a similar overestimation trend to Figures 6-4 and 6-5 is observed in the unbalanced low voltage network model. In both balanced and unbalanced network models, the largest overestimation occurs when 2 customers are grouped together. However, as these figures and Table 6-5 below show, on average the overestimation percentages increase in the unbalanced network setting as compared to the balanced network setting.

Table 6-5: Comparison of overestimation percentages at various aggregation levels-alternative topology

Aggregation Level	Average Overestimation % Loughborough Data Set-balanced	Average Overestimation % Loughborough Data Set-unbalanced
2-house	36.55	63.92
4-house	37.75	67.44
6-house	38.72	68.99
8-house	39.46	72.98
10-house	46.00	75.34
Aggregation Level	**Average Overestimation % CLNR Data Set no.8-balanced model**	**Average Overestimation % CLNR Data Set no.8-unbalanced model**
2-house	40.80	65.80
4-house	42.00	69.25
6-house	42.97	72.22
8-house	44.21	74.46
10-house	47.25	77.75

Table 6-5 shows that on average the loss estimates in the unbalanced network with an alternative topology also increase as the aggregation level increases. Although the trend in the balanced network is similar to the unbalanced network with the largest rise in the loss estimates occurring at 2-house aggregation level, the results in Table 6-5 show that this increase is more severe when the low voltage network is unbalanced. For example, compared to the balanced network, in the unbalanced network the loss estimates for the 2-house aggregation level rise by just under 30% and just above 25% for Loughborough and CLNR data sets, respectively.

6.3 Effects of Varying Customer Aggregation Levels on Estimation of Voltage Levels in Balanced Low Voltage Networks

In this section of the thesis, maximum voltage drops along the low voltage network models are calculated on each phase in different aggregation levels. The maximum voltage drops, which are calculated at each section of the network models, are subtracted from the starting voltage at the substation. This provides the minimum voltage levels that is experienced by the customers at the end of the network on each phase. The minimum voltage levels are important to the DNOs in maintaining the statutory limits of 230V +6% -10% and also in determining the capacity of various phases of the low voltage network to host embedded generation and low carbon technologies units.

Previous studies in sections 5.4 and 5.5 showed that as the time resolution of smart meter data is decreased from 1 minute to 120 minute intervals, the minimum voltage estimates are overestimated (Poursharif et al. 2017). In this section, the effect of customer data aggregation on the accuracy of minimum voltage level estimates is investigated by comparing the minimum voltage level estimates on each phase of the low voltage network at 2, 4, 6, and 8 house aggregation level to 1 house level (no aggregation).

Figures 6-10 and 6-11 below show the minimum voltage levels experienced at the end of cable C on the red phase at various aggregation levels. The results for the yellow and blue phases can be found in Appendix G.

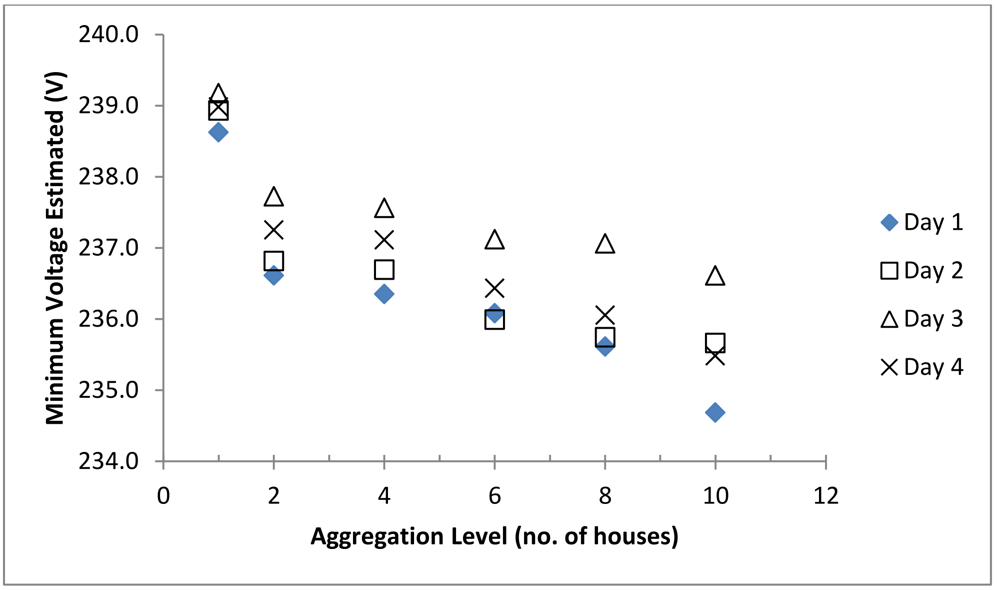


Figure 6-10: Minimum voltage levels at various aggregation levels (Loughborough data set)

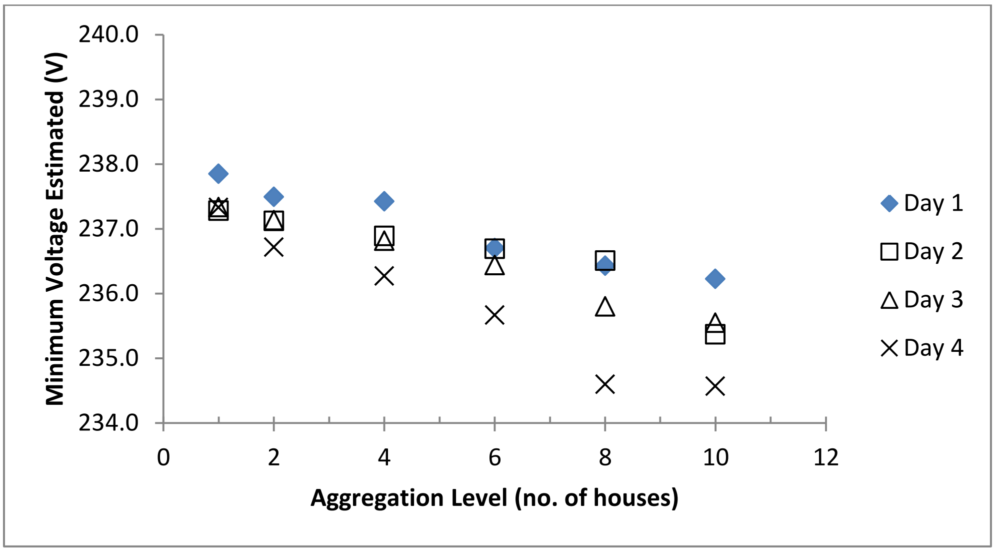


Figure 6-11: Minimum voltage levels at various aggregation levels (CLNR data set no.8)

Figures 6-10 and 6-11 show that as the customer loads are aggregated, the minimum voltage estimations are underestimated. The 2-house and the 4-house aggregation levels provide the closest estimations to the actual voltage drops along the low voltage network. Table 6-6 below shows the average underestimation percentages at each level of aggregation across the 8 samples dates from both data sets.

Table 6-6: Average underestimation percentages at various aggregation levels

Aggregation Level	Average Underestimation % Loughborough Data Set	Average Underestimation % CLNR Data Set no.8
2-house	0.76	0.14
4-house	0.84	0.25
6-house	1.05	0.45
8-house	1.18	0.68
10-house	1.39	0.85

It can be observed from Table 6-6 that the highest underestimation percentages at 2-house and 6-house aggregation points with the averages of 0.76% and 0.14% at 2-house and 19% and 20% at 6-house aggregation levels for the sample dates from the Loughborough data set and the CLNR data set no.8, respectively. The trend then slows down to a steadier underestimation trend as customer aggregation levels increase to 10.

While the underestimation of voltage levels at 2-house aggregation can be justified by the fact that larger loads are experienced at the sections of the low voltage network as a result of placing the aggregation points mid-way between the two neighbouring customers on the same phase, the large voltage drop experienced at 6-house aggregation level can be explained by the fact that some of loads that were previously on cable A with lower resistance are now shifted to cable B or C as a result of aggregation and this causes another large decrease in the voltages.

In order to validate this, an alternative network model (see Figure 3-10) was created, where all the 100 customers are connected to a main cable with the specifications of cable A.

Alternative low voltage network model:

This section presents the results of minimum voltage levels estimated at various aggregation levels on the red phase of the low voltage network shown in Figure 3-10. Figures 6-12 and 6-13 below shows the results using the smart meter data on the sample dates from the Loughborough and CLNR data sets, respectively.

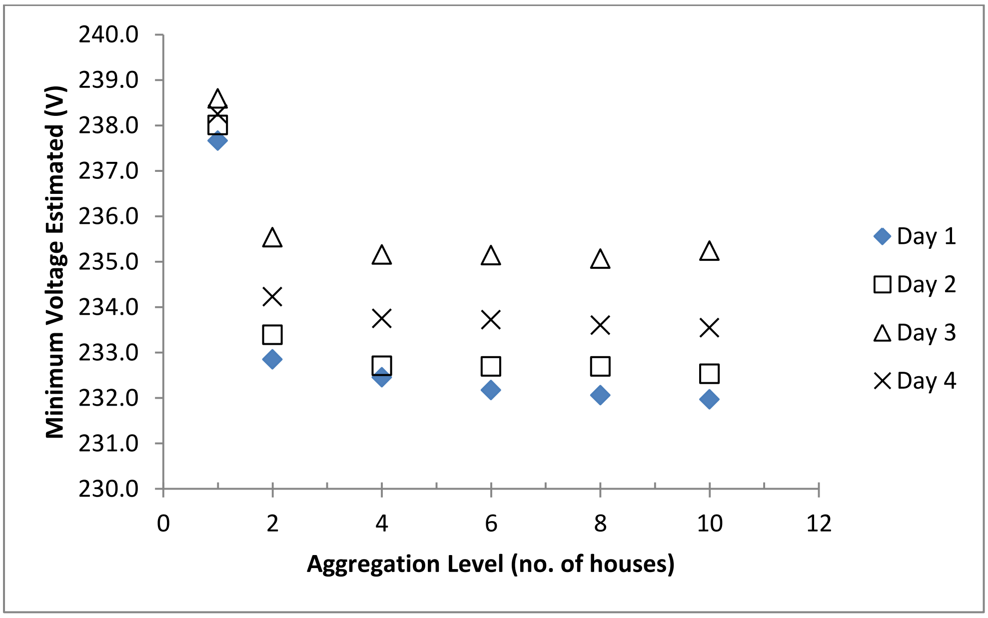


Figure 6-12: Minimum voltage levels estimated for different aggregation levels-alternative topology (Loughborough data set)

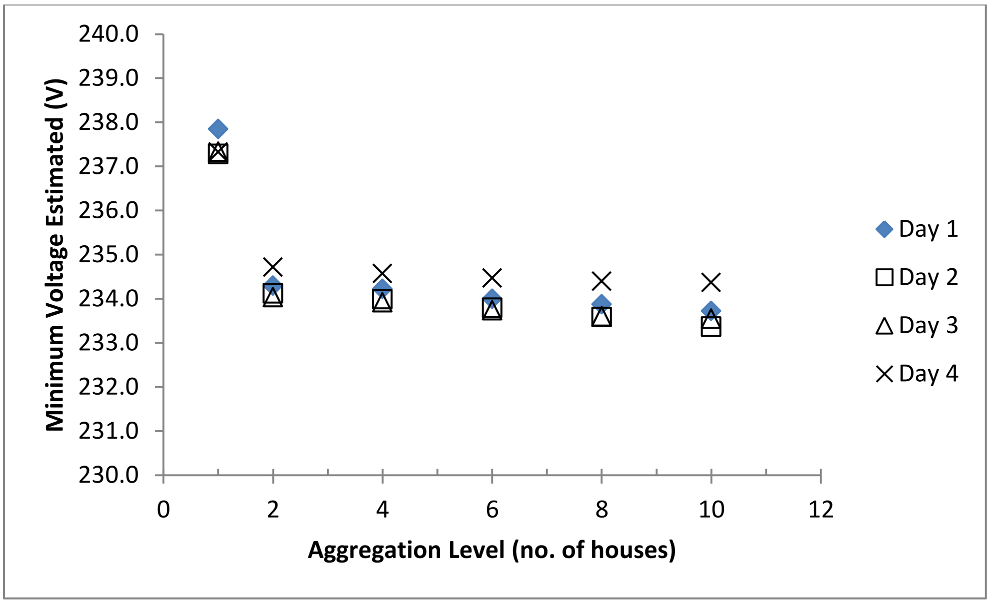


Figure 6-13: Minimum voltage levels estimated for different aggregation levels-alternative topology (CLNR data set no.8)

As Figures 6-12 and 6-13 show as data from more numbers of customers are aggregated together, the minimum voltage levels estimates are decreased. The largest underestimation occurs when data from 2 customers are aggregated. Table 6-7 below shows the average underestimation percentage at different levels of aggregation across the 8 sample dates from both data sets.

Table 6-7: Average underestimation percentages at various aggregation levels-alternative topology

Aggregation Level	Average Underestimation % Loughborough Data Set	Average Underestimation % CLNR Data Set no.8
2-house	1.73	1.33
4-house	1.94	1.38
6-house	1.97	1.45
8-house	2.01	1.51
10-house	2.02	1.56

As Table 6-7 above shows, now that all customers are connected to one main cable the largest share of underestimation in voltage estimates only take place at 2-house aggregation level with 1.73% and 1.33% for the Loughborough and CLNR data sets, respectively. Further aggregation from 2-house level to 10-house level only introduces marginal underestimation percentages.

6.4 Effects of Varying Customer Aggregation Levels on Estimation of Voltage Levels in Unbalanced Low Voltage Networks

In this section, the effect of customer data aggregation on the accuracy of voltage estimated is investigated in an unbalanced network arrangement. In this network model, there are more customers on the red phase and the allocation ratio of customers to phases has been changed from 33:33:34 to 40:30:30 for customers on the red, yellow, and blue phases, respectively.

Figures 6-14 and 6-15 below show the changes in minimum voltage levels estimated for the customers on the red phase as smart meter data from customers are aggregated from no aggregation to 10-house aggregation level.

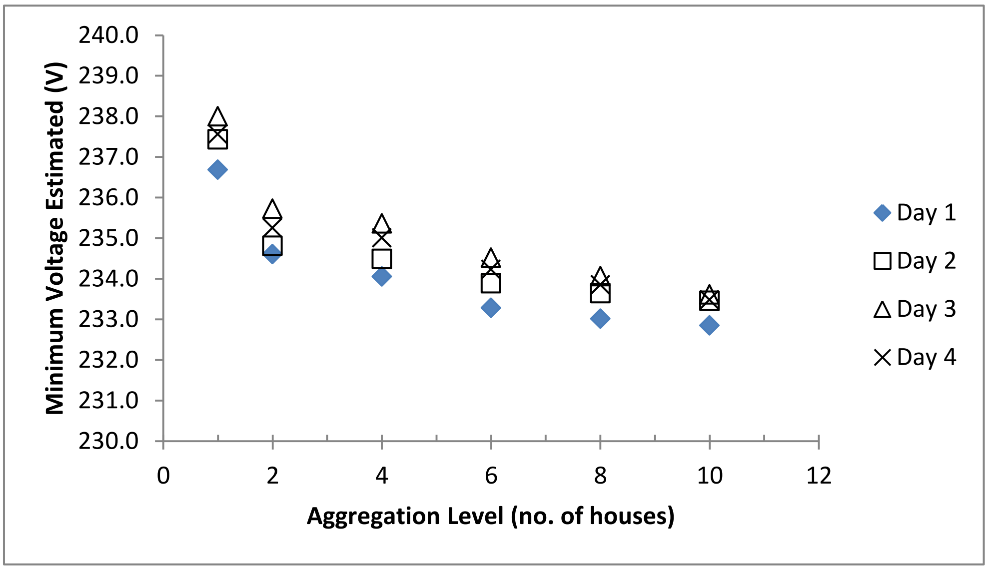


Figure 6-14: Estimated minimum voltage levels on the red phase at various aggregation levels-unbalanced network (Loughborough data set)

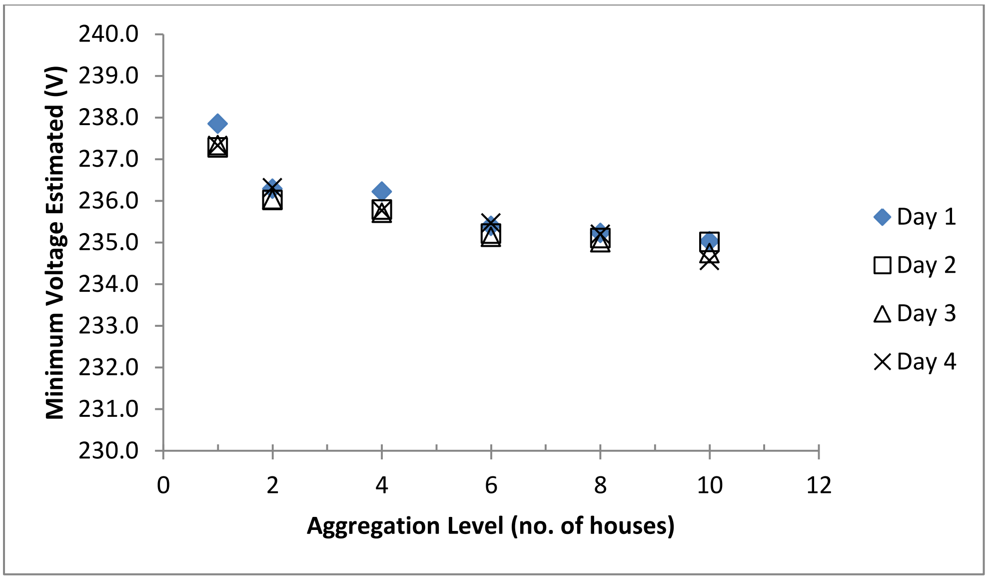


Figure 6-15: Estimated minimum voltage levels on the red phase at various aggregation levels-unbalanced network (CLNR data set no.8)

As Figures 6-14 and 6-15 above show that similar to the balanced network model, in the unbalanced network model the largest decrease in minimum voltage levels occur at 2 and 6 house aggregation levels.

However, as Table 6-8 below shows, in a situation where the low voltage network model is unbalanced the estimated minimum voltage levels are underestimated at a higher rate compared to the balanced model. For example, at the 2-house aggregation level, minimum voltage estimates on the red phase in the balanced network setting experience 0.76% and 0.14% underestimation on average, but this changes to 0.98% and 0.54% when the network model is unbalanced .

Table 6-8: Comparison of average underestimation percentages of minimum voltage levels in the balanced and unbalanced models

Aggregation Level	Average Underestimation % Loughborough Data Set-balanced	Average Underestimation % Loughborough Data Set-unbalanced
2-house	0.76	0.98
4-house	0.84	1.13
6-house	1.05	1.45
8-house	1.18	1.59
10-house	1.39	1.71
Aggregation Level	**Average Underestimation % CLNR Data Set no.8-balanced**	**Average Underestimation % CLNR Data Set no.8-unbalanced**
2-house	0.14	0.54
4-house	0.25	0.66
6-house	0.45	0.90
8-house	0.68	0.97
10-house	0.85	1.10

This trend is also observed at other aggregation levels, especially at 6-house aggregation level that experiences about 0.45% higher underestimation in the unbalanced network model compared to the balanced network model.

Alternative Low Voltage Network Model

Similar results are also found when the network arrangement is changed to the alternative low voltage network model with all 100 customers connected to one main cable. Figures 6-16 and 6-17 below show that similar to the alternative balanced low voltage network model, in the unbalanced network model the highest underestimation of minimum voltage levels take place at 2-house level aggregation and further aggregation only causes relatively minor underestimation.

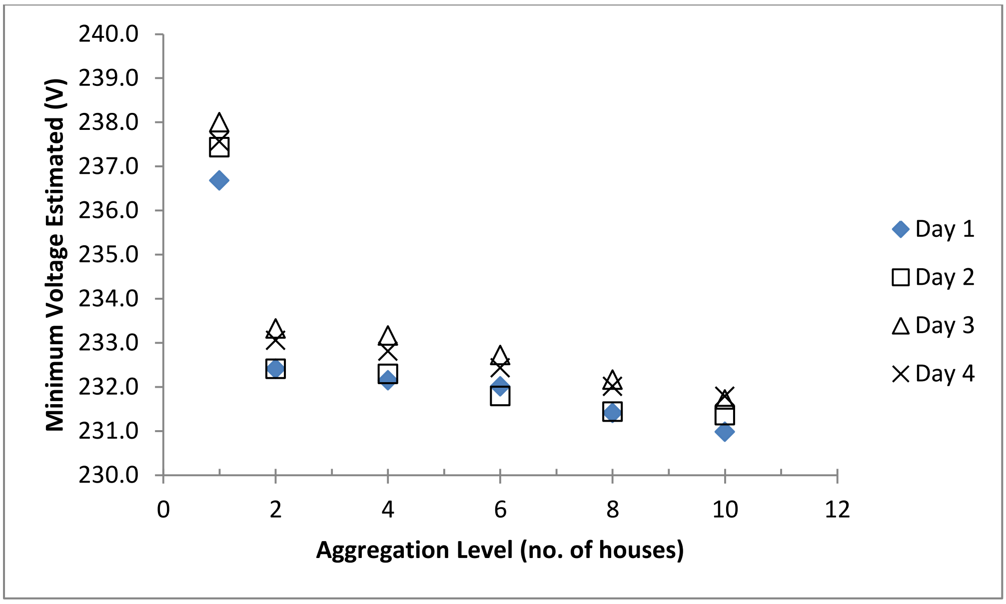


Figure 6-16: Minimum voltage levels estimated for different aggregation levels-alternative topology-unbalanced model (Loughborough data set)

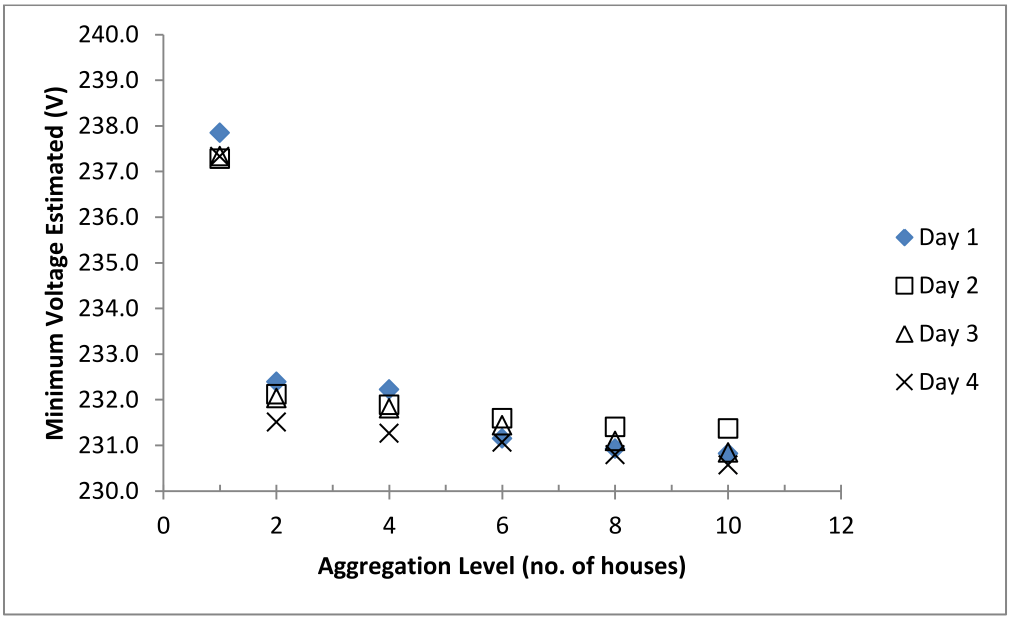


Figure 6-17: Minimum voltage levels estimated for different aggregation levels-alternative topology-unbalanced model (CLNR data set no.8)

A comparison of average underestimation percentages between the balanced and the unbalanced alternative model at various aggregation levels in Table 6-9 below shows that the rate of underestimation increases in the unbalanced model.

Table 6-9: Comparison of average underestimation percentages of minimum voltage levels in the balanced and unbalanced models-alternative model

Aggregation Level	Average Underestimation % Loughborough Data Set-balanced	Average Underestimation % Loughborough Data Set-unbalanced
2-house	1.73	1.93
4-house	1.94	2.03
6-house	1.97	2.18
8-house	2.01	2.38
10-house	2.02	2.51
Aggregation Level	**Average Underestimation % CLNR Data Set no.8-balanced**	**Average Underestimation % CLNR Data Set no.8-unbalanced**
2-house	1.33	2.29
4-house	1.38	2.38
6-house	1.45	2.58
8-house	1.51	2.69
10-house	1.56	2.76

The results in Table 6-9 confirm the results previously seen in section 6-3 where the highest underestimation rate is observed at 2-house aggregation level. However, this rate of underestimation is higher in the unbalanced model. For example, in the case of the Loughborough data set, the average underestimation percentage rises from 1.73% to 1.93% at 2-house aggregation level in the unbalanced network and in the case of the CLNR data set no.8 it rises from 1.33% to 2.29%. The minimum voltage levels estimated on the yellow and the blue phase can be found in the Appendix H.

6.5 Summary

In this chapter, the relationship between low voltage aggregation of customer loads on a similar phase and the accuracy of half-hourly loss and voltage estimates were investigated. Aggregation levels are defined as no aggregation, 2, 4, 6, 8, and 10 house aggregation levels. In each aggregation scenario, neighbouring customers on the same phase are grouped together and an aggregation point between the neighbouring meters replaces the single household demand points. Examples of the aggregation point placements can be seen in Figure 3-0.

Additionally, an alternative low voltage network model is created as shown in Figure 3-10, where 100 customers are connected to a single long main cable with the specification of cable A in the previous model. This model is referred to as the alternative low voltage model. Load flow analysis is carried out on both models in a balanced and unbalanced setting and the losses and minimum voltage levels are estimated for 8 sample dates from the two data sets.

In the first place, the impact of aggregation levels on the accuracy of loss estimates is examined in a balanced low voltage network model on the sample dates. This is also tested in the alternative low voltage network model to investigate the effects of aggregation point placement and network topology on the accuracy of loss estimates. These steps are also repeated on an unbalanced low voltage network model that has more loads connected to the red phase.

Secondly, the minimum voltage levels on the various phases of the network models are estimated in each aggregation scenario. This analysis is carried out in the balanced and the unbalanced network setting. Also, in both cases the impact of aggregation point placement and network topology is tested in the alternative network model. The results from of the minimum voltage level estimates for the customers on the red phase are presented in this chapter and the results of the analysis carries out on the yellow and blue phases can be found in Appendices G and H.

6.6 Conclusions

Our analysis in section 6.1 showed that in the balanced network models as the customer loads are aggregated, the loss estimates are overestimated. In the network with two branches the highest overestimation percentages occur at 2-house and 6-house aggregation levels with the average overestimation percentages of 44.6 % and 125.9% across the 8 sample dates, respectively. The results in the alternative model that contains 100 customers connected to one main cable and no branches show that the largest overestimation percentage takes place only at 2-house level of aggregation with the average overestimation percentage of about 38%. This demonstrates that the large overestimation at 6-house level aggregation is due to the placement of aggregation points as a result of the network topology. A similar result is also seen in the unbalanced network model in section 6.2. The findings in section 6-3 shows that a similar trend to the balanced network can be seen in unbalanced networks with higher percentages of overestimation as customer data are aggregated compared to the balanced network model.

Analysis in sections 6.3 and 6.4 showed that as the number of customers grouped together increases the estimates of voltage drops on the low voltage networks rise. This leads to the underestimation of minimum voltage levels on each of the phases. Similar to the loss estimates, the highest underestimation percentages in the low voltage network with two branches take place at 2-house and 6-house levels. The average underestimation percentages at 2-house and 6-house aggregation levels across the 8 sample dates are 0.45% and 0.75%, respectively. The alternative low voltage model shows that in a different topology the effect of 6-house aggregation level is minimised. The results in the unbalanced network setting also indicate that the rate of underestimation increases from no aggregation to 2-house aggregation as the loads on the phases are unbalanced.

Our studies are novel in that for the first time the effects of aggregation of customer data on the low voltage network estimates are investigated. Previous studies in EA Technology (2015a) and EA Technology (2015b) show that aggregation of data from two customers can preserve the privacy of the consumers to highest degree while being the most cost effective aggregation level in terms of the costs involved to disaggregate the aggregated data. Our research contextualises the impacts on the accuracy of important low voltage network performance indicators such as loss and voltage estimates. Our findings also show factors such as low voltage network topology and placement of the aggregation points are also very important factors when aggregating customer smart meter data. This is particularly important due to the fact that each DNO owns and operates thousands of different low voltage networks with different topologies and characteristics.

Chapter 7 Conclusions

The nature of electricity distribution network operation in the UK and throughout the rest of the World is changing. This is mainly due to the fact that the distribution grid operators need to ensure swift accommodation of the rising amounts of embedded generation in the electricity network, especially at lower voltage levels.

A great many of the new sources of energy generation are being installed in the electricity distribution grid either at consumer premises or at small generation sites. This requires the low voltage side of the network, which has traditionally been designed only to transmit energy to consumers, to be transformed to host embedded generation units. As a result, the conventional uni-directional flows of information and energy in the traditional distribution grid set ups is changing to bi-directional flows of energy and information, which in turn will transform the existing passive distribution grid to an intelligent grid. The proactive distribution network operation requires great understanding of voltage levels and customer demands on the low voltage side of the network. To this end, information systems and various sources of monitoring such as smart meters and Advanced Metering Infrastructure (AMI) are being widely implemented in the UK and various different countries to enable these changes as they can provide the required information from the low voltage side of the network to DNOs.

By 2025, half-hourly smart meter data from most domestic households and businesses in the UK will be available and can potentially provide the DNOs with more detailed information about the low voltage network. However, as our extensive study of the documents and literature in chapters 1 and 2 showed there is a significant gap in knowledge in both fields of academic and industrial research about the extent to which the smart meter data in the UK can provide accurate information about the low voltage network to the DNOs.

Based on the research aims and objectives of this research, the following aspects of smart meter data were investigated in this thesis:

- the effectiveness of historical half-hourly smart meter data from a group of smart meters in predicting the customer demands from unavailable smart meters on the same low voltage network (chapter 4).

- the impact of smart meter time resolutions on the estimation accuracy of critical low voltage network information such as losses, voltage levels, and cable loading capacity (chapter 5).
- the impact of customer smart meter data aggregation on the estimation accuracy of losses and voltage levels (chapter 6).

The following sections highlight the main findings and conclusions learned from the analysis carried out in chapters 4, 5, and 6 and the ways in which the DNO applications are affected by these results.

7.1 The Use of Historical Smart Meter Data to Predict Missing Low Voltage Loads

Knowledge of the customer currents on the low voltage network is a key foundation for smart grid applications. However, knowledge of low voltage currents is generally poor. Accurate knowledge of the currents is a key element of Advanced Distribution Management Systems. Traditionally, the DNOs were interested in peak demands from clusters of customers, but with the transition to smarter grid applications, details of customer demand and generation patterns on specific low voltage networks are necessary in order for the DNOs to be able to monitor power flows and voltage levels at critical points on the network (Lees 2014). However, research into various methods of load prediction has been mainly focused on the higher levels of the electricity network.

In chapter 2, we looked at some of the major methods of load allocation which are used in state estimation methods at higher levels of distribution networks, followed by investigating demand prediction methods that are used by the DNOs on the low voltage levels such as:

- After Diversity Maximum Demand (ADMD) (McQueen et al. 2004)
- Customer demand curves (Carson and Cornfield 1973; McQueen et al. 2004; Vélez et al. 2014)
- Transformer kVA allocation (Kersting and Philips; Arritt et al. 2012)
- Monthly usage allocation (Kersting and Philips; Arritt et al. 2012).

We then selected and examined 5 methods of customer load prediction in chapter 4 that were based on historical smart meter data and k-nearest weighted averages. Methods 1 and 4 were based on a combination of transformer kVA and monthly usage allocation approaches used in conjunction with smart meter data sets. Method 5 (k-nearest weighted averages) was based on a statistical approach used in a new context to find the closest historical half-hourly values to the missing values. The situation investigated was where not all the smart meters are reporting their values in real-time.

The 5 devised methods of customer load prediction were as follows:

1. Prediction of missing customer loads on the date based on real-time smart meter data from the neighbouring meters on the sample date and the historical data from a week earlier on a similar day.

2. Prediction of missing customer loads on the date based on real-time smart meter data from the neighbouring meters on the sample date and the historical data from a month earlier on a similar day.

3. Prediction of missing customer loads on the sample date based on real-time smart meter data from the neighbouring meters on the sample date and of average of historical data from four weeks before on a similar day.

4. Prediction of missing customer loads on the sample date based on real-time smart meter data from the neighbouring meters on the sample date and the historical data from a similar type day from a year before on a similar day.

5. Prediction of missing customer loads at peak times on the sample date based on real time smart meter data from the neighbouring meters on the sample date and the k-nearest weighted averages of the closest historical data values from the neighbouring smart meters.

These methods were initially tested on a 20-house network model (models A-1 and A-2) and the best two methods were then applied to a larger network model with 50 meters (model B).

7.1.1 Key Findings

Our analysis showed that methods 1 and 4 produce the best results, based on the Mean Absolute Percentage Error (MAPE) of the predictions for 48 half-hours on the sample dates. The results indicated that the best performing predictions take place when the prediction are either based on measured data from one week earlier or the average of customer loads from four weeks earlier. It was also found that in the absence of half-hourly temperature data, method 5 performs worse than methods 1 and 4, but only slightly.

When comparing the MAPE results from the three daily peak times on the sample dates in model A-1 (model with 20 meters from Loughborough data set), the MAPE values for methods 1, 4, and 5 were 43.61%, 42.31%, and 48.80%, respectively. These figures for model A-2 (model with 20 meter from CLNR data set no.1) were, 30.31%, 35.91%, 36.51%. These results showed that methods 1 and 4 that use historical smart meter data from 1 week earlier and the average of 4 weeks earlier perform better in predicting the peak loads, compared to method 5 that uses k-nearest weighted averages.

Applying methods 1 and 4 to a larger sample of meters in model B showed that increasing the size of the smart meters with real time data does not significantly improve the load estimates of the missing meters. In both models (A and B), 10% of the meters were assumed not to be communicating the customer currents to the DNOs in real time, which is the worst case scenario predicted by the OFGEM. In this model, the average MAPE does not significantly change for method 1, but in the case of method 4, MAPE results improved from just over 36% to just above 28%. However, applying methods 1 and 4 to predict loads from 50 different individual households in model B showed a greater range of errors as a result of volatility in individual lifestyle of consumers.

7.1.2 Conclusions

The focus of this chapter of the thesis was to investigate how the estimation of the customer demands at any time of the day and year, can be improved by using data from distribution substation monitoring along with either annual consumptions or historical smart meter data. Our approaches considered how smart meter data can be used to improve on these methods and how effective are simple approaches, which assume demands from “similar times and situation” are effective indictors of missing demands, based on smart meter data?

It was found that on average, using smart meter data from a week prior and the average of half-hourly values from 4 weeks earlier on a similar day type, performed the best with the k-nearest weighted average method performing slightly worse than the two methods mentioned (methods 1 and 4). These methods can successfully predict the load shapes of the missing customer currents, but in some instances peak demands are misplaced by an hour or are underestimated. Also, the consistency in the results from the two data sets show that apart from the natural differences in every data set, the methods can be used in other data sets as well.

The approaches presented in chapter 4 improve the current practices in the field and are easy to implement by the DNOs. The fact that methods 1 and 4 are simpler and computationally less demanding than complex prediction methods mean that this could well be the choice of network operators in practice, especially in applications such as asset management network planning and design by providing a more detailed load flow analysis models to the DNOs. Also, installing network meters by DNOs on various node points on the low voltage network can rectify the inaccuracies in predicting the peak demands from the customers with missing smart meter data.

Besides the accuracy of the estimates, in practice it is desirable that an approach is:

a) Straightforward to implement.

b) Computationally inexpensive as the objective is to estimate the real-time power flows on the low voltage network.

Methods 1 and 4 are very simple approaches to implement. Hence they are the approaches that seem the most suitable for use in practice. Additionally, these methods improve the previous approaches in Kersting and Philips (2008); Arritt et al. (2012); Vélez et al. (2014) by using actual smart meter data and similar historical times and situations. Extra information such as the half-hourly temperatures can also in theory improve method 5 estimates based on k-nearest weighted average more accurate and more attractive in practice.

7.2 Impact of Smart Meter Time Resolution Intervals on the Estimation of Critical Low Voltage Indicators

Understanding of critical low voltage network indicators such as losses, voltage levels, and cable loading percentages are highly important to the DNOs. A greater understanding of these indicators will enable the DNOs to:

- deliver power to customers and run the low voltage network more efficiently and cost effectively.
- identify the areas of network that need reinforcement.
- identify the network capacities for new connections.
- identify the areas of the network with power quality problems.
- maintain voltage levels within the statutory limits.
- accommodate and manage the integration of embedded generation and low carbon technologies without compromising the quality of power delivered to the customers.
- reconfigure the network arrangements to balance the loads on different phases of the network.
- actively manage the network and smart grid applications such as Demand Side Management.

To this end, smart meter data are deemed to be the most cost effective facilitators of creating bottom-up load flow analysis models that provide the DNOs with a clear picture of their low voltage networks. However, the ability of smart meters to provide accurate information to the DNOs in the UK can be limited due to the fact that the DNOs are only provided with half-hourly averages of customer loads. Therefore, the studies in chapter 5 were designed to investigate the effects of smart meter time resolution intervals on the accuracy of critical low voltage estimates such as losses, voltages, and cable loading percentages.

Initially, fundamental low voltage network information such as network losses, voltage levels, and cable loading percentages and the current practices on estimating such information were discussed in chapter 2. After that the three phase low voltage network model that was created to be populated with smart meter data was described in chapter 3. The sources and characteristics of the data were also described in chapter 3. The time resolution studies in chapter 5 were then carried out using smart meter data from 8 sample dates from two different data sets. The smart meter data time resolutions were averaged from the highest resolution of 1 minute to lower resolutions of 5, 10, 15, 30, 60, and 120 minutes. The data were then added to a balanced and an unbalanced three phase low voltage network model with 100 houses. The total network losses, the minimum voltage levels, and the cable loading percentages are then estimated on the sample dates using the various time resolution intervals mentioned earlier. The loss estimates and voltage levels were also calculated in an unbalanced low voltage network with more customers on the red phases of the network. The importance of the unbalanced network was to represent the unbalanced low voltage networks that are found in real networks. The difference between having 1 minute smart meter data or lower resolution of smart meter data is then expressed in the context of these estimates. The results of our analysis were presented in chapter 5 of this work.

7.2.1 Key Findings

Our analysis showed that as the granularity of smart meter data decreases from 1 minute to 120 minutes, the network loss estimates also drop. The greatest underestimation takes place between 1 minute and half-hourly average time resolution intervals with most of the underestimation occurring between 1 and 15 minute time resolutions. Our sample dates showed that, the underestimation in losses between 1 minute to half-hourly averages range from just under 20% to just above 25%. Our analysis showed that the underestimation of network losses is more severe in unbalanced low voltage networks with a similar correlation between the time resolution of smart meter data and the loss estimates as seen in balanced low voltage models. In the unbalanced network model, the average loss underestimation percentages from 1 minute to half-hourly smart meter time resolution intervals range from just under 30% to just under 35%, which showed about 10% more underestimation compared to the balanced model.

In section 5.3, we examined two methods of predicting the 1 minute loss values based on either half-hourly losses which showed that if the slope of the curves of network loss estimations based on various time granularities of data for weekends and weekdays of a specific month of the year were analysed and a constant for those day types was calculated (α), this constant then can be applied to half-hourly loss estimates to calculate the 1 minute loss values across 52 sample dates. Our analysis showed that using regression models to fit the bust curve to the 120 minute to half-hourly losses in combination with the constant value of $\overline{\alpha}$ led to fairly accurate estimation of 1 minute losses with interquartile range of the APE values between just under 5% and just under 15%.

In terms of voltage levels, our analysis showed that as the time resolution of smart meter data decreases from 1 minute to 120 minutes, the minimum voltage level estimates on the low voltage network rise. The greatest overestimation takes place between 1 minute and half-hourly average intervals, with the highest percentage of underestimation occurring between 1 to 15 minute intervals. Our examples showed that, the average overestimation of minimum voltage estimates between 1 minute to half-hourly averages range from just under 0.17% to just above 0.18%. This underestimation is more severe in the unbalanced network with the average model ranging between 1.12% to 1.24%. This is important when considering the statutory voltage limits of 230V +10%-6%.

The main reason behind these trends is that major peak demand points in data are neglected when the smart meter data are averaged over longer intervals of time. This is confirmed when the estimates of loading percentages of low voltage cables at various time resolutions were analysed in section 5.6. Our analysis in section 5.6 showed that there is a significant underestimation of peak demands on each phase of the low voltage cables as the time resolution of data decreases from 1 minute to 120 minute intervals.

7.2.2 Conclusions

Our analysis shows that having half-hourly smart meter data does not provide the DNOs with accurate low voltage network information and this lack of accuracy can hinder the ability of the DNOs to improve their smart grid applications. The underestimation of losses and the overestimation of minimum voltage levels at half-hourly time resolution intervals can directly and indirectly impact a number of DNO applications in the UK.

The underestimation of losses can lead to the neglect of the areas of the network that need reinforcement and the distortion of the network capacity to accept new connections. It can hide the low voltage networks that need to be reconfigured and balanced. It can also impair the decision making process of the DNOs to design the networks in the most cost effective and efficient way.

The overestimation of minimum voltage levels can mislead the operators and the regulators in determining whether a particular low voltage network is operating within the statutory limits or not. It can also lead to a lack of understanding of the areas of the network where the quality of power is being affected and are likely to cause consumer complaints or experience faults. The overestimation of minimum voltage levels can also affect the low voltage network’s capacity to host embedded generation and low carbon technologies which are closely related.

The underestimation of peak cable capacity percentages can also hinder the ability of the DNOs to effectively reconfigure the existing connections and assign new connections to maintain the network balance. It also affects the ability of the DNOs to run applications such as Active Network Management (ANM) and Demand Side Management (DSM) effectively.

Our analysis used smart meter data from various sample dates and two different data sets and the consistency in the results mean that the findings can be applicable to other similar types of smart meter data time resolution studies. Our models are also improvement on previous studies by Oliveira and Padilha-Feltrin (2009); McKenna et al. (2012); Quiroz et al. (2012); Brandauer et al. (2013); Dashtaki and Haghifam (2013); and Urquhart and Thompson (2015) in a number of ways:

- The number of customers, sample dates, and data sets used.
- The use of a three-phase model instead of a single phase model.
- The time resolution intervals ranging from 1 to 120 minutes.
- The placement of the impact of varying time resolution intervals in the context of the accuracy of three major low voltage network indicators.
- The investigation of the effects on in both a balanced and an unbalanced network model.
- The prediction of 1 minute losses based on lower resolution of smart meter time resolution intervals.

Our analysis in section 5.3 shows that using regression models can to some extent compensate the lack of loss estimate accuracy at half-hourly levels. However, the DNOs and the policy makers need to make a decision to whether it is more effective to have access to higher resolution of smart meter data based on the findings. As highlighted in Northern Powergrid (2016), our findings are one of the main reasons that have led to the Norther Powergrid's decision to invest in installing higher resolution network meters in a number of low voltage substations around the network as well investigating the use of loss forecasting models based on regression models used in section 5.3 (Northern Powergrid 2016). Our findings have also been conveyed to the policy makers such as BEIS and OFGEM. Our work also received a lot of interest from network operators and researchers in the field from different countries when presented at CIRED 2017 in Glasgow.

7.3 Impact of Smart Meter Data Aggregation on the Estimation of Critical Low Voltage Indicators

Another aspect of smart meter implementation programme in the UK that can potentially reduce the smart meter data benefits that can be gained by the DNOs is that the network operators are required to anonymise the customer data as soon as they are received from the DCC. Aggregation of customer data by the DNOs has been recognised to be the most cost effective way of dissociating the individual customer data from the consumers at the low voltage level (OFGEM 2017b). It is also the most appropriate way for the UK electricity market setup that is owned and operated or serviced by various entities such as the DNOs and the suppliers. The optimum level of aggregation is not decided by the regulators to date. Therefore, it is important to quantify the effects of low voltage customer data aggregation by demonstrating the impact on major low voltage network information areas such as estimates of losses and voltage levels. Our study in chapter 6 aimed to achieve this objective.

In the first place, the three phase 100-house low voltage network model was changed to a low voltage network model where the customers on the same phase are grouped together at aggregation points across the low voltage network. The model that consisted of three cables and two branches was also altered to represent an alternative low voltage network model with all the 100 customers on a single straight low voltage cable. In the next step, 5 aggregation levels of 2, 4, 6, 8 and 10 house levels were defined and the models were populated with half-hourly smart meter data, which is the time resolution of the smart meter data transmitted to the DNOs. After that, losses and minimum voltage levels were estimated and compared to the estimates in the balanced and the unbalanced low voltage network models with no aggregation. The effects of aggregation levels on the accuracy of loss and voltage estimates were then expressed compared to the estimates when there was no aggregation.

7.3.1 Key Findings

Our analysis showed that as the aggregation level of smart meter data increases from 1-house to 10-house level, the network loss estimates rise. In the model with two branches, the highest overestimation percentages occurred at 2-house and 6-house levels and the overestimation percentages were more severe in the unbalanced model of the low voltage network. In the balanced model, the average overestimation percentages at

2-house level range from 44.2% to 44.9% and these percentages at 6-house level range from 121.3% to 130.3%. In the unbalanced network model, the average overestimation percentages at 2-house aggregation level increase by approximately 30% and the average overestimation percentage at 6-house aggregation level rises by 30-50%. In the alternative model with no branches however, the highest overestimation percentage takes places at 2-house aggregation level.

In terms of voltage levels, our analysis showed that as the aggregation level of smart meter data increases from 1-house to 10-house level, the minimum voltage level estimates at the end of cables B and C of the network drop. This underestimation was more dramatic on cable C where there are more customer connections. The underestimation of minimum voltage levels was more sever on the red phase in the unbalanced models. In the balanced model, the average underestimation percentages at 2-house level range from 0.14% to 0.76% and the underestimation percentages at 6-house level range from 0.45% to 1.05%. In the unbalanced network model, the average underestimation percentages at each level of aggregation increased compared to similar aggregation level in the balanced network. However, In the alternative model with no branches, the highest underestimation percentage took place at 2-house aggregation level.

7.3.2 Conclusions

The aggregation of customer loads will lead to the provision of less accurate low voltage network information to the DNOs compared to having non-aggregated smart meter data. The use of aggregated customer loads in the load flow models used by the DNOs can cause the overestimation of losses and underestimation of minimum voltage levels as our models demonstrated.

This can be problematic in a number of DNO application areas. Overestimation of losses can falsely identify the areas of the network that are operating efficiently as inefficient and/or faulty. This can lead to the allocation of resources to the wrong areas of the network. The overestimation of losses can also affect the decisions to host new connections, embedded generation, or low carbon technologies. The Overestimation of minimum voltages can also directly impair the judgement of the network operators in managing the embedded generation and low carbon technologies installed at various areas of the low voltage network.

However, maintaining the privacy of customers are also of utmost importance and as EA Technology (2015a) and EA Technology (2015b) have highlighted, the best aggregation level in terms of costs of disaggregation of the customer profiles is the 2-house level. Our studies also proved that the best aggregation level is the 2-house level. However, our studies were carried out in a novel way. The studies by the EA Technology only focus on the costs that the DNOs will incur in disaggregating the customer data aggregated. However, our study is a unique study that shows the effects of 5 data aggregation levels at the low voltage scale on the accuracy of important low voltage network parameters such as loss and voltage level estimates. Also, our models highlighted the importance of the low voltage network topology in the accuracy the estimates at various aggregation levels. Our studies showed that in both network topology models, the best performing aggregation level is the 2-house level. The comparison of results from the network model with two branches and the alternative model with no branches showed that the impact of aggregation on the estimation accuracy of critical low voltage networks is highly related to the placement of aggregation points and the topology of the networks. Our findings in chapter 6 is of huge interest to both the DNOs, the researchers in the field, and the policy makers as the debate over the most appropriate smart meter data aggregation level is still ongoing.

7.4 Limitations and Further Research

Our first set of analyses in chapter 4, was carried out on half-hourly smart meter data to examine two methods of predicting missing or unavailable customer demands half-hourly readings from smart meters using historical smart meter data from neighbouring smart meters. Although our results showed that this can be achieved with a high degree of success in large samples, the absence of more smart meter data with at least 13 months of consumption data limited our sample sizes to only 50. It would be interesting to investigate how the scaling can be improved when the sample of neighbouring smart meters is increased beyond 50 meter. Also, the availability of half-hourly temperature data can potentially improve the statistical methods such as k-nearest weighted averages.

Our studies in chapter 5, investigated the relationship between varying smart meter time resolutions and the estimation accuracy of some of the major low voltage network performance indicators such as network losses, voltage levels, and cable loading percentages. For our analysis, the number of available meters was limited to 150 meters. This limited the possibility of creating larger low voltage network models. It would be interesting to investigate the results on a bigger piece of low voltage network with various cable sizes, branches, more smart meters, and higher penetration of embedded generation. Also, the smart meter data that were used in these models were not accompanied by exact geographical data. This made the substation measurement data that were initially acquired at the beginning of this research redundant, because the smart meter data were not from the connections that are served by that specific low voltage substation data.

Additionally, the inaccuracies in the 1 minute PV generation data restricted our analysis in that the low voltage test models were only populated with net consumption data. It would be interesting to examine the ways in which the addition of Embedded Generation at various points on the network would affect the results.

In the case of the analysis of the effects of smart meter customer aggregation in chapter 6, the network topology dictated the type of analysis that was carried out. It would be interesting to investigate the result on different low voltage models with different network arrangements and characteristics. However, every low voltage network is different and this would be a very difficult task without having a real network model and more information about customer phases.

Also more research should be carried out to identify the ways in which customer smart meter data, low voltage substation data, and performance indicator data such as network loss estimates can be utilised to identify the phases that each customer is connected to. As our results have highlighted the knowledge of customers' will be very important to the DNOs. The effects of having a better picture of the customer connections and the phases on the accuracy of customer aggregation scenarios should also be researched in the future.

Putting it all together, a major research direction that will be of enormous value can be carried out in integrating geographical and weather data with smart meter data to improve the methods of customer load predictions, low voltage network voltage levels and cable loading estimations and the ways in which the results of should be visualised and integrated into the DNOs GIS databases to enhance low voltage network operational applications. Unfortunately, these types of data are usually collected for different purposes by various entities and this limits the possibilities of working with such data sources to conduct practical experiments and studies. For example, smart meter data sets are usually collected by the electricity suppliers in the UK and do not contain any information about the parts of the low voltage network that they come from. Therefore, relating them to appropriate GIS records is not possible. Also, weather data are usually collected at low resolutions, which as we saw earlier is not very helpful. If all these data sources from the same low voltage network area is available studies can be carried out to investigate the ways in which such data can be combined and visualised to inform the network operators and enhance the decision making process of the DNOs. For example, can this data be used to inform the operators the areas of the low voltage that are experiencing high voltage variations or high losses on the GIS map and why? Or where are the areas of the network with high capacity for hosting new connection, embedded generation, or low carbon technologies?

Chapter 8 References

Abusdal, G. M. et al. (2015). Utilization of Advanced Metering Infrastructure in Back-fed Ground Fault Detection. IEEE Power and Energy Society General Meeting, IEEE, DOI: 10.1109/PESGM.2015.7285819

Acharjee, P. and Gunda, J. (2010). Development prospects of smart Grid in India. IEEE International conference on Power and Energy (PECon2010), Kuala Lumpur, pp.953-957.

Ahmadi, H. et al. (2016). A Linear Power Flow Formulation for Three-Phase Distribution Systems. IEEE Transactions on Power Systems, IEEE, Volume 31, pp. 5012-5021.

Al-otaibi, R. et al. (2016). Feature Construction and Calibration for Clustering Daily Load Curves from Smart-Meter Data. IEEE Transactions on Industrial Informatics, IEEE, Volume 12, Issue 2, pp. 645-654.

Alinjak et al. (2017). Improved Three Phase Power Flow Method for Calculation of Power Losses in Unbalanced Radial Distribution Network. 24th International Conference on Electricity Distribution, 2017, Glasgow, UK. CIRED, Paper no. 0489.

Alturki, Y. A. (2011). Investigation of Power Losses in High Loaded Power Networks. 7th International Conference on Electrical and Electronics Engineering, IEEE, pp. I18-I22.

Alzate, C. and Sinn, M. (2013). Improved Electricity Load Forecasting via Kernel Spectral Clustering of Smart Meters. 13th International conference on Data Mining, IEEE, pp. 943-948.

Amin, S.M. and Wollenberg, B.F. (2005). Toward a Smart Grid. IEEE Power and Energy Magazine, Volume 3, Issue 5, pp. 34-41.

Amjadi, N. (2001). Short-Term Hourly Load Forecasting Using Time-Series Modelling with Peak Load Estimation Capability. IEEE Transactions on Power Systems, Volume 16, Issue 3, pp. 798-805.

Apel, R. et al. (2001). Fault management in electrical distribution networks, Electricity Distribution, 2001. Part 1: Contributions. 16th International Conference and Exhibition on (IEE Conf. Pub No. 482), CIRED, Amsterdam, Volume 3, Chapter 3.7.

Armknecht, F. et al. (2008). A Lifetime-optimized End-to-end Encryption Scheme for Sensor Networks Allowing In-network Processing. Computer Communications, Volume 31, pp. 734-749.

Arritt, R. f. et al. (2012). Investigating load estimation methods with the use of AMI metering for distribution system analysis. IEEE Rural electric power conference, B3, 1

9, IEEE xplore conference publication: http://ieeexplore.ieee.org/Xplore/home.jsp, Date accessed: 13th June 2016.

Arritt, R. F., and Dugan, R. C. (2011). Distribution system analysis and the future smart grid. IEEE Transactions on Industry Applications, IEEE, Volume 47, Issue 6, pp. 2343-2350.

Arup et al. (2014). Smart Grid, Smart City: Shaping Australia's Energy Future, Executive Report. ARUP, [online], available from: smartcitiesappg.com/wp-content/uploads/PDF/SmartGrid.pdf

As'ad, M. (2012). Finding the Best ARIMA Model to Forecast Daily Peak Electricity Demand, Proceedings of the Fifth Annual ASEARC Conference - Looking to the future - Programme and Proceedings, University of Wollongong.

Au, M. T. (2008). Simplified Approach in Estimating Technical Losses in Distribution Network Based on Load Profile and Feeder Characteristics. 2nd IEEE International Conference on Power and Energy, IEEE, pp. 1661-1665.

Aung, Z. et al. (2012). Towards Accurate Electricity Load Forecasting in Smart Grids. The Fourth International Conference on Advances in Databases, Knowledge, and Data Applications, DBKDA 2012.

Bala, S.K. et al. (2012). A Review on Development of Smart Grid Technology in India and Its Future Perspectives. Engineering and Systems (SCES), 2012 Students Conference Allahabad, IEEE, pp.1-6.

Bale, P. (2015). Connecting Renewable Energy in Lincolnshire. Western Power Distribution 2015, [online], Available from: http://www.networkrevolution.co.uk/resources/project-library/

Balta-Ozkan et al. (2014). Scenarios for the Development of Smart Grids in the UK. UK Energy Research Centre (UKERC), [online], Available from: www.ukerc.ac.uk/.../scenarios-for-the-development-of-smart-grids-in-the-uk.html

Banyard, J.K. and Bostock, J.W. (1998). Asset Management Investment Planning for Utilities. Proceedings of the Institution of Civil Engineers - Civil Engineering, Volume 126, Issue 2, pp. 65-72.

Baran, M. E. (1993). Data Requirements for Real-time Monitoring and Control of Feeders. Power Engineering Society 1999 Winter Meeting, IEEE, Volume 1, pp. 374-376.

Baran, M. E. and Kelley, A. W., (1994). A Branch-Current-based State Estimation Method for Distribution Systems. IEEE Transactions on Power Systems, IEEE, Volume 10, no. 1, pp. 483-491.

Baran, M. and McDermott, T. E. (2009). Distribution System State Estimation Using AMI Data. Power Systems Conference and Exposition, IEEE, pp. 1-3.

Barbato, A. et al. (2017). Lessons Learnt from Real-time Monitoring of the Low Voltage Distribution Network. Sustainable Energy, Grids and Networks, Elsevier Ltd.

Barteczko-Hibbert, C. et al. (2015). Insight Report: Baseline Domestic Profile, Test Cell 1a Customer Subgroup Analysis, CLNR, [online], Available from: http://www.networkrevolution.co.uk/resources/project-library/

Beccali, M. et al. (2004). Forecasting Daily Urban Electric Load Profiles Using Artificial Neural Networks. Energy Conversion and Management, Volume 45, pp. 2879-2900.

Beddoes, A. et al. (2007). The Impact of Distributed Generation upon Network losses. 19th International Conference on Electricity Distribution, CIRED, Paper no. 0116.

Bell, K. et al. (2009). Automation to Maximise Distributed Generation Contribution and Reduce Network Losses. 20th International Conference on Electricity Distribution, CIRED, Paper no. 0038.

Bird, J. (2015). Developing the Smarter Grid: The Role of Domestic and Small and Medium Enterprise Customers. CLNR, Document no. CLNR-L246, [online], Available from: http://www.networkrevolution.co.uk/wp-content/uploads/2015/01/small-customers-L2-report-v0-9.pdf

Biselli, A. et al. (2013) Protection of Consumer Data in the Smart Grid Compliant with the German Smart Metering Guideline. SEGS 2013, Proceedings of the first ACM workshop on Smart energy grid security, Berlin, pp. 41-52.

Black, M. (2009). Using Strategic Models in Electricity Distribution Investment Planning. Proceedings of the 20th International Conference on Electricity Distribution, CIRED 526/1- 526/4.

Black, M. et al. (2005). A Semi-Markov Approach for Modelling Asset Deterioration, Journal of Operational Research Society, Volume 56**,** pp. 1241-1249.

Bokhari, A. et al. (2015). Combined Effect of CVR and DG Penetration in the Voltage Profile of Low Voltage Secondary Distribution Networks. IEEE Transactions on Power Delivery, IEEE, Issue 99.

Borozan , V. and Rajakovic, N. (1997). Application Assessments of Distribution Network Minimum Loss Reconfiguration. IEEE Transactions on Power Delivery, IEEE Volume 12, Issue 4, pp. 1786-1792.

Brakken, R. and Ait-Laoussine, T. (2006). Integrating Customer Load and GIS Data for Improved Distribution Planning and Operations. IEEE Transmission and Distribution Conferences and Exhibition, IEEE, pp. 8-14.

Brandauer, W. et al. (2013). Loss Calculation and Optimization in Low Voltage Network. 22nd International Conference on Electricity Distribution, CIRED, Paper no. 1186.

Brint, A.T. et al. (2008). The rise, current position and future direction of Asset Management in utility industries. Journal of Operational Research Society, Volume 60**,** pp. 106-113.

Brint, A.T. and Black, M. (2014). Improving estimates of asset condition using historical data. Journal Operational Research Society 65, Issue 2, pp. 242-251.

Brown, R. (2008). Impact of Smart Grid on Distribution System Design. Power and Energy Society Meeting – Conversion and Delivery of Electrical Energy in the 21st Century, IEEE, pp. 1-4.

Brown, R.E. and Humphrey, G. (2005). Asset Management for Transmission and Distribution. IEEE Power & Energy Magazine, IEEE, volume 3, Issue 3, pp. 39-45.

Bulkeley, H. et al. (2015). High Level Summary of Learning: Domestic Smart Meter Customers. CLNR. [online], Available from: http://www.networkrevolution.co.uk/resources/project-library/

Bulkeley, H. et al. (2016). Making a Smart City for the Smart Grid? The Urban Material Politics of Actualising Smart Electricity Networks. Environment and Planning A: Economy and Space, Sage Journals, Volume 48, Issue 9.

Caire, R. et al. (2002). Impact Assessment of LV Distributed Generation on MV Distribution Network. IEEE Power Engineering Society Summer Meeting, IEEE, Volume 3, pp. 1423-1428.

Caldon, R. et al. (2005). Co-ordinated Voltage Regulation in Distribution Networks with Embedded Generation. 18th International Conference and Exhibition on Electricity Distribution, 2005. CIRED 2005, IET.

Cao, D. M. et al. (2009). Examinations of the Impact of Possible Distribution Network Design on Network Losses. 20th International Conference on Electricity Distribution, CIRED, Paper no. 0950.

Carer, P. (2006). Probabilistic methods used in asset management for MV electrical equipment at EDF. International Conference on probabilistic methods applied to power systems, 2006, PMAPS, pp.1-5.

Carmona, C. et al. (2010). Distribution Transformer Load Allocation from Substation Measurements and Load Patterns. IEEE PES, ISGT Europe, pp. 1-8.

Carpaneto, E. et al. (2006). Branch Current Decomposition Method for Loss Allocation in Radial Distribution Systems with Distribution Generation. IEEE Transactions on Power Systems, IEEE, Volume 21, Issue 3, pp. 1170-1179.

Carson, M. J. and Cornfield, G. (1973) Design of Low-voltage Distribution Networks Interactive Method Based on Calculus of Variations, Proceedings of the Institution of Electrical Engineers, Volume 120, Issue 5, pp. 585-592.

Celik, M. K. and Liu, W. H. E. (1999). A practical Distribution State Calculation Algorithm. IEEE, Proceedings of Engineering Society, Winter Meeting, Volume 1.

Chiandone, M. et al. (2014). Impact of Distributed Generation on Power Losses on an Actual Distribution Network. International Conference on Renewable Energy Research and Application (ICRERA), IEEE.

Chaouch, M. (2014). Clustering-Based Improvement of Nonparametric Functional Time Series Forecasting: Application to Intra-Day Household-Level Load Curves. IEEE Transactions on Smart Grid, IEEE, Volume 5, Issue 1, pp. 411-419.

CLNR (2016). www.networkrevolution.com. Date accessed: July 2016.

Clutterbuck, D. et al. (2005). Development of Technical Asset Lives for HV Transmission Equipment in England and Wales. In: 3rd IEE International Conference on Reliability of Transmission and Distribution Networks. Institution of Electrical Engineers (IEE), London.

Conti, S. et al. (2001). Study of the Impact of PV Generation on Voltage Profile in LV Distribution Networks. IEEE PowerTech Conference, IEEE, Porto.

Costa, P. M. and Mathos, M. A. (2004). Loss Allocation in Distribution Networks with Embedded Generation. IEEE Transactions on Power Systems, IEEE, Volume 19, Issue 1, pp. 384-389.

Covig, C.F. et al. (2014). Smart Grid Projects Outlook 2014. JRC Science and Policy Reports, European Union 2014, Available from: http://ses.jrc.ec.europa.eu/smart-grids-observatory

Cowan, R. K. and Daim, U. T. (2012). Integrated Technology Roadmap Development Process: Creating Smart Grids Roadmaps to Meet Regional Technology Planning Needs in Oregon and the Pacific Northwest, Proceedings of PICMET 2012, Technology management for emerging technologies Vancouver, pp.2871-2885.

Customer-Led Network Revolution (2016). Leading The Way to a Low Carbon World, CLNR, [online], Available from: http://www.networkrevolution.co.uk/

Customer-Led Network Revolution (2017). Project Data, CLNR: Domestic Smart Meters on Time of Use Tariffs, CLNR, Available from: http://www.networkrevolution.co.uk/resources/project-data/

Dashtaki, A. K. and Haghifam, M. R. (2013). A New Loss Estimation Method in Limited Data Electric Distribution Network. IEEE Transactions on Power Delivery, IEEE, Volume 28, Issue 4, pp. 2194-2200.

Datta, E. A. et al. (2014). Accelerated Deployment of Smart Grid Technologies in India – Present Scenario, Challenges and Way Forward ,IEEE PES Innovative Smart Grid Technologies Conference (ISGT), IEEE, pp. 1-5.

Davidson, I. and Ljumba, N. M. (2002). Optimization Model for Loss Minimization in a Deregulated Power Distribution Network. IEEE Africon, 2002, IEEE, Volume 2, pp. 887-894.

DECC (2009). The UK Renewable Strategy, DECC. [online], Available from: https://www.gov.uk/government/publications/the-uk-renewable-energy-strategy

DECC (2010a). Smart Metering Implementation Programme: Prospectus. [online]. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/42718/220-smart-metering-prospectus-condoc.pdf

DECC (2011a). Response to Prospectus Consultation. [online]. Available from: http://www.decc.gov.uk/assets/decc/Consultations/smart-meter-imp-prospectus/1475-smart-metering-imp-response-overview.pdf

DECC (2011b). Central Communications and Data Management. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/42738/1481-central-communications-data-management.pdf

DECC (2011c). Smart Metering Impact Assessment. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/42740/1485-impact-assessment-smart-metering-implementation-p.pdf

DECC (2011d). Functional Requirements Catalogue. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/42737/1480-design-requirement-annex.pdf

DECC (2013). Carbon Emissions Reduction Target (CERT). [online]. Available from: http://www.decc.gov.uk/en/content/cms/funding/funding_ops/cert/cert.aspx

Dede, A. et al. (2016). Toward a New Standard for Secondary Substations: The Viewpoint of a Distribution Utility. IEEE Transactions on Power Delivery, IEEE, Volume PP, Issue 90.

Degefa, M. Z. et al. (2014). Unlocking Distribution Network Capacity Through Real-time Thermal Rating for high Penetration of DGs. Journal of Electric Power Systems Research, Volume 117, pp. 36-46.

Demset, J. et al. (2007). Analysis on the Increased Losses in Supply Systems Due to Voltage Drop and Voltage Distortion. European Conference on Power Electronics and Applications, IEEE, pp. 1-9.

Department for Business, Energy, and Industrial Strategy (2016). Digest of United Kingdom Energy Statistics (DUKEs) 2016, [online], Available from: https://www.gov.uk/government/collections/digest-of-uk-energy-statistics-dukes

Department of Energy and Climate Change (DECC) (2016). Smart Meters: Quarterly Report to end March 2016 Great Britain. DECC, 2016. [online]. Available from: https://www.gov.uk/government/collections/smart-meters-statistics

Diop, A. and Dercle, D. (2005). Losses Modelling for Distribution Networks, 18th International Conference on Electricity Distribution, CIRED, IET.

DNV KEMA (2013). Review of Network Benefits of Smart Meter Message Flows, DNV KEMA Energy and Sustainability, [online], Available from: http://www.energynetworks.org/modx/assets/files/electricity/futures/smart_meters/FINAL%20REPORTS%20from%20consultants/DNV%20KEMA%20DNO%20Reviewing%20Network%20Benefits%20of%20Smart%20Meter%20Message%20Flows%20-%20FINAL%2030-04-13.pdf

Dortolina, C. A. and Nadira, R. (2005). The Loss that is Unknown is No Loss at All: A Top-down/Bottom-up Approach for Estimating Distribution Losses. IEEE Transactions on Power Systems, IEEE, Volume 20, Issue 2, pp. 1119-1125.

Douglas, J. (2011). GB Smart Customer Response Trials Workshop. Western Power Distribution 2011, [online], Available from: http://www.networkrevolution.co.uk/resources/project-library/

Doumbia, M. L. and Agbossou, K. (2007). Voltage Variation Analysis in Interconnected Electrical Network – Distributed Generation. IEEE Canada Electrical Conference, IEEE, pp. 525-530.

Dudeney, C., Ward, J., Sustainability First (2015). Lessons Learnt Report: Customer Trial Equipment Installations. Customer-Led Network Revolution, [online], Available from: http://www.networkrevolution.co.uk/resources/project-library/

EA Technology (2015a). Smart Meter Aggregation Assessment Final Report - Executive Summary. Prepared for the ENA, [online], Available from: www.energynetworks.org

EA Technology (2015b). Smart Meter Aggregation Assessment Final Report – Benefits Reduction. Prepared for the ENA, [online], Available from: www.energynetworks.org

EA Technology (2016). Summary of the Low Carbon Networks Funds Learning. EA Technology 2016 Prepared for Office of Gas and Electricity Markets, Project no. 106080, [online], Available from: https://www.ofgem.gov.uk/ofgem-publications/100314

EDW Technology (2016). Glossary. EDW Technology Integrated Energy Solutions. [online], Available from: http://www.edwt.org/glossary, Date accessed: July 2016.

Electricity Network Strategy Group (ENSG) (2010). A Smart Grid Vision. ENSG, 2009, Available from: http://webarchive.nationalarchives.gov.uk/20100919181607/http:/www.ensg.gov.uk/index.php?article=126

Electricity North West (2014). Low Voltage Network Solutions: Close Down Report. Electricity North West 2014, [online], Available from: http://www.enwl.co.uk/docs/default-source/future-low-voltage/lv-network-solutions-close-down-june-14-v4-1-final.pdf?sfvrsn=2

Electronic Power and Energy Institute (EPRI) (2007). Protecting Investments in Advanced Metering Infrastructure Using IntelliGrid Architecture. EPRI, Available from: www.ferc.gov/CalendarFiles/20070423091937-EPRI%20-%20IntelliGrid.pdf

Emelin, S. et al. (2013). Correlation between Load Density and Voltage Drop. 22nd International Conference and Exhibition on Electricity Distribution, CIRED, IET, pp. 1-4.

Energy Networks Association (ENA) (2011). Developing Networks for Low carbon. Energy Network Association (ENA), [online], Available from: http://www.ofgem.gov.uk/Networks/SGF/Publications/Documents1/Smart%20Grid%20Forum%20Workstream%203%20Report%20071011%20MASTER.pdf

Energy Network Association (ENA) (2012). Analysis of Network Benefits from Smart Meter Message Flows. Energy Network Association (ENA), [online], Available from: http://www.energynetworks.org

Energy Network Association (ENA) (2018). Electricity Distribution Map. Energy Network Association (ENA), [online], Available from: http://www.energynetworks.org

Esri (2007-2010). Working with Geometric Networks, Esri.com: Online course material at www.esri.com

Estebsari, A. et al. (2016). An Improved Fault Location Method for Distribution Networks Exploiting Emerging LV Smart Meters. IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems, IEEE, DOI: 10.1109/EESMS.2016.7504815

ETP Smart Grids (2015). Consolidated View of the ETP SG (European Technology Platform on Smartgrids), [online], Available from: www.smartgrideu.com

EurElectric (2013). Power Distribution in Europe: Facts and Figures. [online]. Available from: http://www.eurelectric.org/powerdistributionineurope/

Fan, J. and Borlase, S. (2009) The Evolution of Distribution. IEEE Power and Energy Magazine, Volume 7, Issue 2, pp. 63-68.

Farhangi, H. (2010). The Path of the Smart Grid. IEEE Power and Energy Magazine, Volume 8, Issue 1, pp. 18-28.

Ferreira, P. M. et al. (2012). Neural Network Based Predictive Control for Thermal Comfort and energy Savings in Public Buildings. Journal of Energy and Buildings, Volume 55, 2012, pp. 238-251.

Fleckenstein, J. E. (2013) Calculating Currents in Balanced and Unbalanced Three Phase Circuits. PDH Course E336, PDH Center, USA.

Fotheringham, A. S. et al. (2002). Geographically Weighted Regression Chichester. Wiley, John Wiley & Sons, Inc.

Fourie, J. W. and Calmeyer, J. E. (2004). A Statistical Method to Minimize Electrical Energy Losses in a Local Electricity Distribution Network. 7th AFRICON Conference in Africa, IEEE, Volume 2, pp. 667-673.

Frame, D. F. et al. (2012). The Uncertainties of Probabilistic LV Network Analysis. IEEE Power and Energy Society General Meeting, IEEE, DOI: 10.1109/PESGM.2012.6344587

Garcia, F. D. and Jacobs, B. (2010). Privacy-friendly Energy-metering via Homomorphic Encryption. Institute for Computing and Information Sciences, Radboud University Nijmegen.

Gono, R. et al. (2007). Reliability Analysis of Distribution Networks. 9th International Conference on Electrical Power Quality and Utilisation, Barcelona, pp.1-5.

Goyal, A. et al. (2016). Asset Health Management Using Predictive and Prescriptive Analytics for the Electric Power Grid. IBM Journal of Research and Development, Volume 60, Issue 1, pp. 4:1-4:14.

Gerbec, D. et al. (2005). Allocation of the Load Profiles to Consumers Using Probabilistic Neural Networks. IEEE Transactions on Power Systems, Volume 20, Issue 2, pp. 548-555.

Gharavi, H., Ghafurian, R. (2011). Smart Grid: The Electric Energy System of the Future. Proceeding of the IEEE, IEEE, Volume 99, Issue 6, pp. 917-921.

Ghofrani, M. et al. (2011). Smart Meter Based Short-Term Load Forecasting for Residential Customers. North American Power Symposium, IEEE, pp. 1-5.

Giordano, V. et al. (2011). Smart Grid Projects in Europe: Lessons Learned and Current Development. JRC Reference Report, European Union, 2011.

GIS (2011). www.esri.com, [online], Date accessed: July 2016.

Ghosh, A. K. et al. (1997a). Load Forecasting for Distribution Circuit State Estimation. IEEE Transactions on Power Delivery, IEEE, Volume 12, Issue 2, pp. 999-1005.

Ghosh, A. K. et al. (1997b). Distribution Circuit State Estimation Using a Probabilistic Approach. IEEE Transactions on Power Delivery, IEEE, Volume 12, Issue 1, pp. 45-51.

Glamocanin, G. et al. (2003). Using a GIS and DLE for Reduction of Outage Time in Distribution Networks. Bologna PowerTech Conference, Bologna, Italy, Volume 2, Available from: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1304332

Gono, R. et al. (2007). Reliability Analysis of Distribution Networks. 9th International Conference on Electrical Power Quality and Utilisation, Barcelona, DOI: 10.1109/EPQU.2007.4424180

Google Street (2016). Google: Image Date 2008, [online], Available from: https://www.instantstreetview.com/@53.418241,-1.186495,-124.87h,-0.63p,1z

Güngör, V. C. et al. (2011). Smart Grid Technologies: Communication Technologies and Standards. IEEE Transactions on Industrial Informatics, IEEE, Volume 7, Issue 4, pp.529-539.

Hahn, H. et al. (2009). Electric Load Forecasting Methods: Tools for Decision Making. European Journal of Operational Research, Volume 199, pp. 902–907.

Hammond, M. and Jones, P. (2000). Effective AM Maintenance & Asset Management, Maintenance and Asset Management Journal, Volume 15, Issue 4, pp. 3-8.

Hattam, L. and Greetham, D. V. (2017). Green Neighbourhoods in Low Voltage Networks: Measuring Impact of Electric Vehicles and Photovoltaics on Load Profiles. Journal of Modern Power Systems and Clean Energy, Volume 5, Issue 1, pp. 105-116.

Hayes, B. and Prodanovic, M. (2014). State Estimation Techniques for Electric Power Distribution Systems. 2014 European Modelling Symposium, IEEE, pp. 303-308.

Heckmann, W. et al. (2013). Detailed Analysis of Network Losses in a Million Customer Distribution Grid with High Penetration of Distributed Generation. 22nd International Conference on Electricity Distribution, CIRED, Paper no. 1478.

Heywood, I. et al. (2006). An Introduction to Geographical Information Systems, 3rd ed., Pearson Education Limited.

Hofstra, N. et al. (2008). Comparison of Six Methods for the Interpolation of Daily. European Climate Data, Journal of Geophysical Research. Atmospheres, Volume 113, Issue D21, pp. 1-19.

Hollingworth, D. et al. (2012). Demonstrating Enhanced Automatic Voltage Control for Today’s Low Carbon Network. CIRED Workshop Lisbon 2012, CIRED, Paper 0319.

Hong, T. and Fan, S. (2016). Probabilistic Electric Load Forecasting: A Tutorial Review. International Journal of Forecasting, Volume 32, Issue 3, pp. 914-938.

Horton, R. M. and Leonard, W. H. (2005). Mathematical Modelling in Science: Using Spreadsheets to Create Mathematical Models and Address Scientific Inquiry. The Science Teacher, volume 72, Issue 5, pp. 40-46.

Howard, R. J. (2001). Infrastructure Asset Management under Australian Accounting Standard. 27 (AAS27) Proceedings of the Institution of Civil Engineers: Municipal Engineer, Volume 145, Issue 4, pp. 305-310.

Hughes, D. (2003). The Use of Health Indices to Determine End of Life and Estimate Remnant life for Distribution Assets. Proceedings of the 17th International Conference on Electricity Distribution, CIRED, pp. 5.15.1-5.15.6.

IEA (2011). Smart Grids Roadmap. International Energy Agency, [online], Available from: https://www.iea.org/publications/freepublications/publication/smartgrids_roadmap.pdf

IEEE (2013). IEEE Grid Vision 2050. IEEE Standard Association, [online], Available from: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6577603&filter=AND(p_Publication_Number:6577601)

Ignatius, K. O. et al. (2015). Analysis of Copper Losses Due to Unbalanced Load in a Transformer. IJRRAS 2015, Volume 23, Issue 1. [online], Available from: www.arpapress.com/Volumes/Vol23Issue1/IJRRAS_23_1_04.pdf

Ijumba, N. M. et al. (1999). Influence of Distribution Generation on Distribution Network Performance. IEEE Xplore Conference: Africon, IEEE, Volume 2.

Ilves, M. (2016). Current Status of WP3: Smart Meters. ESSnet Big Data Meeting, Tallinn, Statistics Estonia.

Infield, D. G. et al. (2007). Potential for Domestic Dynamic Demand-Side Management in the UK. Power Engineering Society General Meeting, IEEE, pp. 1-6.

Ivy Link (2015). Smart Meters Explained. Ivy Link Partnership. [online]. Available from: http://www.ivylink.co.uk/smart-meters-explained/, Date accessed: 1st August 2016.

Jadhav, N.G and Dharme, A.A. (2012). Technical Challenges for Development of Smart Grid in India, Technical Challenges for Development of Smart Grid in India, IEEE-International Conference on Advances in Engineering, Science and Management (ICAESM -2012), IEEE, Nagapattinam, pp.784-788.

Jagtap, K. M. and Khatod, D. J. (2015a). Loss Allocation in Distribution Network with Distributed Generations. IET Journal of Generation, Transmission, and Distribution, IET, Volume 9, Issue 13, pp. 1628-1641.

Jagtap, K. M. and Khatod, D. J. (2015b). Loss Allocation in Radial Distribution Networks with Different Load Models and Distributed Generation. IET Journal of Generation, Transmission, and Distribution, IET, Volume 9, Issue 12, pp. 1275-1291.

Jain, A.K. (2010). Data Clustering 50 Years Beyond K-means. Michigan State University, [online], Available from: https://www.cs.utah.edu/~piyush/teaching/kmeans50yrs.pdf, Date accessed: April 2017.

Jamali, S. and Bahmanyar, A. (2016). A New Fault Location Method for Distribution Networks Using Sparse Measurements. Electrical Power and Energy Systems, Volume 81, pp. 459-468.

Jiang Y. et al. (2016). Outage Management of Distribution Systems Incorporating Information from Smart Meters. IEEE Transactions on Power Systems, IEEE, Volume 31, Issue 5, pp. 4144-4154.

Jirapornanan, A. (2010). Study of Smart Grid for Thailand and Identification of the Required Research and Development, Technology Management for Global Economic Growth (PICMET) Phuket, Proceedings of PICMET'10, IEEE, pp. 1-6.

JRC (2016). Smart Grid Project Map. JRC Website, [online], Available from: http://ses.jrc.ec.europa.eu/sites/ses.jrc.ec.europa.eu/files/u24/2014/project_maps_28_april_2014.html, Date accessed: 1st August 2016.

Kadar, D. (2011). Ashton Hayes Smart Village Project. SP Energy Networks 2011, [online], Available from: http://www.networkrevolution.co.uk/resources/project-library/

Karimi, M. et al. (2013) Distribution State Estimation for Smart Distribution System. IEEE PowerTech Grenoble, IEEE, pp. 1-6.

Kaspirek, M. et al. (2013). Problems of Voltage Stabilization in MV and LV Distribution Grids Due to the Operation of Renewable Energy Sources. 22nd International Conference on Electricity Distribution, CIRED, Paper no. 0062.

Kersting, W. H. and Philips, W. H. (2008). Load Allocation Based Upon Automatic Meter Readings. IEEE Transmission and Distribution Conference and Exposition, 1-8 (IEEE xplore conference publication, IEEE, Available from: http://ieeexplore.ieee.org/Xplore/home.jsp, Date accessed: 13th July 2016.

Klonari, V. et al. (2016). Probabilistic Simulation Framework for Balanced and Unbalanced Low Voltage Networks. Electrical Power and Energy Systems, Elsevier, Volume 82, pp. 439-451.

Konstantelos, I. et al. (2017). Strategic Valuation of Smart Grid Technology Options in Distribution Networks. IEEE Transactions on Power Systems, IEEE, Volume 32, Issue 2.

Labeeuw, W. and Deconinck, G. (2013). Residential Electrical Load Model Based on Mixture Model Clustering and Markov Models. IEEE Transactions on Industrial Informatics, IEEE, Volume 9, Issue 3, pp. 1561-1569.

Lauret, P. et al. (2007). Bayesian Neural Network Approach to Short Time Load Forecasting. Journal of Energy Conversion and Management, Volume 49, pp. 1156–1166.

Lees, M. (2014). Enhanced Network Monitoring Report. Customer-Led Network Monitoring Report. [online], Available from:

http://www.networkrevolution.co.uk/project-library/insight-report-enhanced-domestic-monitoring/

Lee, Y. C. and Etezadi-Amoli (1993). An Improved Modelling Technique for Distribution Feeders with Incomplete Information. IEEE Transactions on Power Delivery, IEEE, Volume 8, Issue 4, pp. 1966-1972.

Ledolter, J. (2013). Data Mining and Business Analytics with R New Jersey, Wiley. John Wiley & Sons, Inc.

Lehtonen, M. et al. (2000). An Advanced Model for Automatic Fault Management in Distribution Networks, Power Engineering Society Winter Meeting, IEEE, Volume 3, pp.1900-1904.

Li, S. et al. (2011). Discussion on Smart Grid Development in China. Power and Energy Engineering Conference (APPEEC) 2011 Asia-Pacific, IEEE, Wuhan, pp.1-4.

Li, W. et al. (2012). The Study for GIS Based Distribution Network Monitoring and Control Area Fault Locations Methods, IEEE Computer Society, Computer Science & Service System (CSSS), 2012 International Conference, IEEE, Nanjing, pp.1758-1761.

Lin, L. et al. (2012) Effect of Load Power Factor on Voltage Stability of Distribution Substation. IEEE Power and Energy Society General Meeting, IEEE, pp. 1-4.

Ling, A. P. A. L. et al. (2012). The Japanese Smart Grid Initiatives, Investments, and Collaborations, International Journal of Advanced Computer Science and Applications, Volume 3, No. 7, 2012.

Linnenluecke, M. K. (2017). Resilience in Business and Management Research: A Review of Influential Publications and a Research Agenda. International Journal of Management Reviews, Volume 19, pp. 4-30.

Liu, Y. and Schulz, N. N. (2002). Knowledge-Based System for Distribution System Outage Locating Using Comprehensive Information. IEEE Transactions on Power Systems, IEEE, Volume 17, Issue 2, pp. 451-456.

Losa, I.(2013). Analysis of the On-going Research and Demonstration Efforts on Smart Grids in Europe. 22nd International Conference on Electricity Distribution, CIRED, Paper no. 0836.

Ma, K. et al. (2016). Estimation of Voltage-driven Reinforcement Cost for LV Feeders Under 3-phase Imbalance. IEEE Access 4 2016, IEEE pp. 1345-1354.

Mahmud, M. A et al. (2014). Voltage Variation on Distribution Networks with Distributed Generation: Worst Case Scenario. IEEE Systems Journal, IEEE, Volume 8, Issue 4, December 2014.

Makinen, A. et al. (2013). Neutral Fault Management in LV Network Operation Supported by AMR System. 22nd International Conference on Electricity Distribution, CIRED, Paper no. 1214.

Marcoci, A. et al. (2013). The Meter-On project: How to Support the Deployment of Advanced Metering Infrastructures in Europe, 22nd International Conference on Electricity Distribution, CIRED, Paper no. 1261.

Marmol, F. G. et al. (2012a). Do Not Snoop My Habits: Preserving Privacy in The Smart Grid. IEEE Communications Magazine, IEEE, Volume 50, Issue 5, DOI: 10.1109/MCOM.2012.6194398

Marmol, F. G. et al. (2012b). Privacy-enhanced Architecture for Smart Metering. International of Information Security, Springer-Verlag Berlin Heidelberg 2012, DOI: 10.1007/s10207-012-0181-6

Matos, P. G. (2013). InovGrid, A smart vision for next generation distribution system. 22nd International Conference on Electricity Distribution, CIRED. P. 1-4, IEEE.

McDonald, L. et al. (2010). Minimisation of Distribution Network Real Power Losses using a Smart Grid Active Network Management System. 45th Universities Power Engineering Conference (UPEC), IEEE.

McGranaghan, M. (2017). Overview of Advanced DMS Applications and Requirements for Integrating Distributed Resources. CIRED, Glasgow, Tutorial 3, EPRI.

McKenna, E. et al. (2011). Smart Meter Data: Balancing Consumer Privacy Concerns with Legitimate Applications. Loughborough University Institutional Repository, Elsevier Ltd.

McKenna, E. et al. (2012). Smart Meter Data: Balancing Consumer Privacy Concerns with Legitimate Applications. Energy Policy, Volume 41, pp. 807-814, Available from: https://doi.org/10.1016/j.enpol.2011.11.049

McKenna, E. and Thompson, M. (2015). High-resolution Stochastic Integrated Thermal-electrical Domestic Demand Model. Journal of Applied Energy, Volume 165, pp. 445-461.

Mcloughlin et al. (2015). A Clustering Approach to Domestic Electricity Load Profile Characterisation Using Smart Metering Data. Applied Energy, Elsevier, Volume 141, pp. 190-199.

McQueen, D. H. O. et al. (2004). Monte Carlo Simulation of Residential Electricity Demand for Forecasting Maximum Demand on Distribution Networks. IEEE Transactions on Power Systems, IEEE, Volume 19, Issue 3, pp. 1685-1689.

Meliopoulos, S. et al. (2011). Smart Grid Infrastructure for Distribution Systems and Applications. Proceedings of the 44th Hawaii International Conference on System Sciences, 2011, pp. 1-11.

Miller, D. (2015). Optimal Solutions for Smarter Distribution Businesses. Northern Powergrid, CLNR, [online]. Available from: http://www.networkrevolution.co.uk/project-library/optimal-solutions-smarter-network-businesses/

Mirowski, P. et al. (2014). Demand Forecasting in Smart Grids. Bell Labs Technical Journal, Volume 18, Issue 4, pp. 135-158.

Mohsenzadeh, A. et al. (2016). Impact of Smart Home Management Strategies on Expected Lifetime of Transformer. IEEE/PES Transmission and Distribution Conference and Exposition, IEEE, DOI: 10.1109/TDC.2016.7519945

Momoh, J.A. (2009). Smart Grid Design for Efficient and Flexible Power Networks Operation and Control. IEEE, Power Systems Conference and Exposition, pp. 1-8.

Morton, K. (1999). Asset Management in the Electricity Supply Industry. IEEE Power Engineering Journal, IEEE, Volume 13, Issue 5, pp. 233–240.

Mutanen, A. et al. (2011). Customer Classification and Load Profiling Method for Distribution Systems, IEEE Transaction on Power Delivery, IEEE, Volume 26, Issue 3, pp. 1755-1763.

Nationalgrid (2016). Towards Smart Energy Systems. [online]. Available from: www.nationalgridconnecting.com, Date accessed: 1st August 2016.

Navarro-Espinoza, A. and Ochoa, L. F. (2016). Probabilistic Impact Assessment of Low Carbon Technologies in LV Distribution Systems. IEEE Transactions on Power Systems, Volume 31, Issue 3, pp. 2192-2203.

Neaimeh, M. et al. (2015). A Probabilistic Approach to Combining Smart Meter and Electric Vehicle Charging Data to Investigate Distribution Network Impacts. Applied Energy, Volume 157, pp. 688-698.

Nijhuis, M. et al. (2017). Valuation of Measurement Data for Low Voltage Network Expansion. Electric Power Systems Research, Volume 151, pp. 59-67.

Northern Powergrid (2016). Delivering Enhanced Understanding of Electricity Losses. Northern Powergrid, January 2016, [online], Available from: https://www.ofgem.gov.uk/ofgem-publications/100439

Npower (2011). Smart Metering Trials Feedback and Learning. Npower 2011, [online], Available from: http://www.networkrevolution.co.uk/resources/project-library/

OFGEM, (2004). Electricity Distribution Price Control Review: Final Proposals. Office of Gas and Electricity Markets, OFGEM 2004.

OFGEM, (2010). Handbook for Implementing the RIIO Model. [online], Available from: https://www.ofgem.gov.uk/ofgem-publications/51871/riio-handbook.pdf

OFGEM (2015). Creating Britain's Low Carbon Future. OFGEM 2016, [online], Available from: https://www.ofgem.gov.uk/ofgem.../84840/2013innovationcompetitionbrochure-pdf

OFGEM (2016a). Losses Incentive Mechanism. OFGEM 2016, [online], Available from: https://www.ofgem.gov.uk/electricity/distribution-networks/losses-incentive-mechanism, Date accessed: July 2016.

OFGEM (2017a). Distributed Generation. OFGEM, [online], Available from: https://www.ofgem.gov.uk/electricity/distribution-networks/connections-and-competition/distributed-generation, Date accessed: April 2017.

OFGEM (2017b). License Guide: Smart Metering. OFGEM, [online], Available from: https://www.ofgem.gov.uk/publications-and-updates/licence-guide-smart-metering, Date accessed: November 2017.

Oliver, M. A. and Webster, R. (1990). Kriging: A Method of Interpolation for Geographical Information Systems. International Journal of Geographical Information Systems, Volume 4, Issue 3, pp. 313-332.

Oliveira, M. E. and Padilha-Feltrin, A. (2009). A Top-down Approach for Distribution Loss Evaluation. IEEE Transactions on Power Delivery, IEEE, Volume 24, Issue 4, pp. 2117-2124.

Onen, M. and Molva, R. (2012). Secure Data Aggregation with Multiple Encryption. Proceedings of the 4th European conference on Wireless sensor networks, EWSN'07, pp. 117-132, Berlin, Heidelberg, 2007, Springer-Verlag.

Open Meter (2011). Second Report on Status Standardization. The Open Meter Consortium, 2011. [online]. Available from: www.metersandmore.com/wp-content/uploads/2013/09/D58.pdf

Parsons, D. (2006). Regulating Asset Management through serviceability and a common framework for investment planning. Proceedings of the 1st World Congress on Engineering Asset Management Gold Coast, Australia 11-14th July, 2006, pp. 794-805.

Paterakis, N. G. et al. (2016). Coordinated Operation of Neighbourhood of Smart Households Comprising Electric Vehicles. Energy Storage and Distributed Generation, IEEE Transactions on Smart Grids, IEEE, Volume 7, Issue 6, pp. 2736-2747.

Pezeshki, H. and Wolfs, P. (2012). Correlation Based Method for Phase Identification in a Three Phase LV Distribution Network. 22nd Australasian Universities Power Engineering Conference (AUPEC), Indonesia, IEEE.

Phoghosyan, A. et. al (2016). Long Term Individual Load Forecast Under Different Electrical Vehicles Uptake Scenarios. Journal of Applied Energy, Volume 157, pp. 699-709.

Pompodakis, E. E. et al. (2016). Photovoltaic Systems in Low-Voltage Networks and Overvoltage Correlation with Reactive Power Control. IET Renewable Generation, Volume 10, Issue 3, pp. 410-417.

Poursharif, G. et al. (2015). Smarter Business Processes Resulting from Smart Data. 23rd International Conference on Electricity Distribution Lyon, CIRED, Paper no. 0520.

Poursharif, G. et al. (2017). Analysing the Ability of Smart Meter Data to Provide Accurate Information to the UK DNOs, 24th International Conference on Electricity Distribution, CIRED, Paper no. 0654.

Queiroz, L. M. O. et al. (2012). Energy Losses Estimation in Power Distribution Systems. IEEE Transactions on Power Systems, IEEE, Volume 27, Issue 4, pp. 1879-1887.

Quilumba, F. L. et al. (2015). Using Smart Meter Data to Improve the Accuracy of Intraday Load Forecasting Considering Customer Behaviour Similarities. IEEE Transactions on Smart Grid, IEEE, Volume. 6, no. 2, 2015. pp. 911-918.

Rao, M. V. et al. (2008). Experiences on Implementation of GIS Based Tools for Analysis, Planning, and Design of Distribution systems. Power and Energy Society General Meeting – Conversion and Delivery of Electrical energy in the 21st Century 2008, IEEE, pp. 1-8.

Reid, S. (2011). Learning from NINES. SSE 2011, [online], Available from: http://www.networkrevolution.co.uk/resources/project-library/

Reinders, J. et al. (2017). Utilizing Residential Flexibility in the Planning of LV Network. 24th International Conference on Electricity Distribution, CIRED, Paper no. 0381.

Repo, S. et al. (2013). Active Distribution Network Concept for Distributed Management of Low Voltage Networks. 4th IEEE/PES Innovative Smart Grid Technologies Europe, IEEE, DOI: 10.1109/ISGTEurope.2013.6695428

Richardson, I. et al. (2010). Domestic Electricity Use: A High-Resolution Energy Demand Model. Journal of Energy and Buildings, Volume 42, pp. 1878-1887.

Richardson, I. (2011). Integrated High-resolution Modelling of Domestic Electricity Demand and Low Voltage Electricity Distribution Networks. Ph.D. dissertation, Centre for Renewable Energy Systems Technology, Loughborough University, Loughborough, U.K., 2011.

Richardson, I. and Thomson, M. (2010). One-Minute Resolution Domestic Electricity Data, 2008-2009, SN: 6583, UK Data Archive.

Roupioz, G. et al. (2013). First Use of Smart Grid Data in Distribution Network Planning. 22nd International Conference on Electricity Distribution, CIRED, Paper no. 0609.

Ruihua, Z. et al. (2010). New Challenges to Power System Planning and Operation of Smart Grid Development in China. 2010 International Conference on Power System Technology, IEEE, Hangzhou, pp.1-8.

Salim, A. A. and Malik, O. P. (2011). Electric Distribution Systems. IEEE and Wiley, Wiley and Sons Inc. Publications.

Salman, S. K. (1996). The Impact of Embedded Generation on Voltage Regulation and Losses of Distribution Network. IEE Colloquium on the Impact of Embedded Generation on Distribution Networks, IEEE Digest no. 1996/194, pp. 2/1-2/5.

Sand, K. et al. (2013). Experiences from Norwegian Smart Grid pilot projects, 22nd International Conference on Electricity Distribution, CIRED. Paper no. 1389.

Scheppe, F. C and Wildes, J. (1970). Power System Static-State Estimation, Part I: Exact Model. IEEE Transactions on Power Apparatus and Systems, IEEE, Issue 1, pp. 120-125.

Sanduleac, M. et al. (2015). Power Quality Assessment in LV Networks Using New Smart Meters Design. 9th International Conference on Compatibility and Power Electronics, CPE, DOI: 10.1109/CPE.2015.7231057

Saputro, N. and Akkaya, K. (2013). On Preserving User Privacy in Smart Grid Advanced Metering Infrastructure Applications. Wiley Online Library (wileyonlinelibrary.com), DOI: 10.1002/sec.706

Scott, N. C. et al. (2002). Use of Load Control to Regulate Voltage on Distribution Network with Embedded Generation. IEEE Transactions on Power Systems, IEEE, Volume 17, Issue 2, pp. 510-515.

Senjyu, T. (2008). Optimal Distribution Voltage Control and Coordinated with Distributed Generation. IEEE Transactions on Power Delivery, IEEE, Volume 23, Issue 2, pp. 1236-1242.

Sexauer, J. M. and Mohagheghi, S. (2013) Voltage Quality Assessment in a Distribution System with Distributed Generation-A Probabilistic Load Flow Approach. IEEE Transactions on Power Delivery, IEEE, Volume 28, Issue 3.

Sharifian, R. et al. (2012). State Estimation for Load Allocation in Distribution Power Systems. IJTPE Journal, 2012, Issue 10, Volume 4, no. 1, pp. 86-90.

Silva, N. et al. (2011). Dynamic SCADA/DMS Data Model – Plug and Play Smart Grid Solution. 21st International Conference on Electricity Distribution, CIRED, Paper no. 1022.

Sirto, O. et al. (2015). Improving Reliability in an Urban Network. Electric Power Systems Research, Volume 120, pp. 47-55.

Slavickas, R. A. (2000). Allocation of Network Losses to Variable Electrical Loads. Proceedings of Ninth International Conference on Harmonics and Quality of Power, IEEE, Volume 2, pp. 391-396.

Smart Energy GB (2017). About the Smart Meter Roll-out. Smart Energy GB. [online]. Available from: https://www.smartenergygb.org, Date accessed: 20th November 2017.

Smart Grids Forum (2012). Smart Grid Forum Work Stream 6 Report. [online]. Available from: https://www.ofgem.gov.uk/ofgem-publications/56835/ws6-report-aug12.pdf

Smart Grid Forum (2014). Smart Grid Vision and Routemap. Department of Energy and Climate Change (DECC), 2014, [online], Available from: https://www.gov.uk/government/publications/smart-grid-forums-smart-grid-vision-and-routemap

Soepboer, P. et al. (2013). SMARTLIFE2: Improving Cable Systems Specifications, 22nd International Conference on Electricity Distribution, CIRED. Paper no. 0584.

Sohn Associates (2009). Electricity Distribution Systems Losses. Sohn Associates. [online]. Available from: https://www.ofgem.gov.uk/ofgem.../sohn-overview-losses-final-internet-version.pdf

Sohn Associates (2013). Is There an Incentive to Reduce GB Electricity Distribution Losses? [online]. Available from: www.sohnassociates.com/Distribution_Network_Incentives_Final_v2_3Nov2014.pdf

SP Energy Networks (2015). Flexible Networks for a Low Carbon Future. SP Energy Networks, [online]. Available from: http://www.spenergynetworks.co.uk/pages/flexible_networks_for_a_low_carbon_future.asp

Spencer, A. (2015). Developing the Smarter Grid: The Role of Industrial & Commercial and Distributed Generation Customers. Customer-Led Network Revolution, CLNR, [online]. Available from: http://www.networkrevolution.co.uk/resources/project-library/

SSE (2014). About Electricity Distribution Networks: Now and In the Future. Scottish and Southern Energy (SSE) 2014, [online], Available from: http://docplayer.net/2664237-About-electricity-distribution-networks-now-and-in-the-future.html

Stein, D. et al. (2013). Enel's Large Scale Demonstration Project Inside Grid4EU: The Challenges of RES Integration in the MV Network. 22nd International Conference on Electricity Distribution, CIRED. Paper no. 1127.

Stephen, B. et al. (2014). Enhanced Load Profiling for Residential Network Customers. IEEE Transactions on Power Delivery, Volume 29, Issue 1, pp. 88-96.

Strbac, G. (2008). Demand Side Management: Benefits and Challenges. Energy Policy, Volume 36, pp. 4419–4426.

Strbac, G. et al. (2010). Benefits of Advanced Smart Metering for Demand Response based Control of Distribution Networks. Energy Network Association (ENA), Available from: http://www.energynetworks.org

Sun, Q. (2010). Comparison of the Development of Smart Grids in China and the United Kingdom. IEEE PES Conference on Innovative Smart Grid Technologies Europe, IEEE, Göteborg, pp.1-6.

Taylor, J. and McSharry, P. E. (2007). Short-Term Load Forecasting Methods: An Evaluation Based on European Data. IEEE Transactions on Power Systems, IEEE, Volume 22, Issue 4, pp. 2213-2219.

Tellechea, D. J. M. (2013). Meter and More Technology: Key Figures of a Success Story. European Utility Week 2013, Amsterdam, Meter and More Open Technologies.

Thomas et al. (2012). Control of Electricity Network Using Smart Meter Data. CIGRE 2012. [online]. Available from: www.cigre.org/content/download/16860/679895/version/1/file/C6_114_2012.pdf

Tonkoski, R. and Lopes, L. A. C. (2008). Voltage Regulation in Radial Distribution Feeders with High Penetration of Photovoltaic. IEEE Energy 2030 Conference, IEEE, pp. 1-7.

Tor, O. and Shahidehpour, M. (2006). Power Distribution Asset Management. Power Engineering Society General Meeting, 2006, IEEE.

Tranfield, D. et al. (2003). Towards a Methodology for Developing Evidence-informed Management Knowledge by Means of Systematic Review. British Journal of Management, Volume 14, pp. 207-222.

UK Power Networks (2013). Smart Urban Low Voltage Networks. UK Power Networks, [online], Available from: http://innovation.ukpowernetworks.co.uk/innovation/en/Projects/tier-1-projects/smart-urban-low-voltage-network/

UK Power Networks (2015). Engineering Design Standards, EDS 08-0136, LV Network Design. UK Power Networks 2015, Document no. EDS 08-0136.

Uribe-Perez, N. et al. (2016). State of The Art and Trends Review of Smart Metering in Electricity Grids. Journal of Applied Sciences, Volume 6, pp. 68.

Urquhart, A. J. and Thomson, M. (2015). Impacts of Demand Data Time Resolution on Estimates of Distribution System Energy Losses. IEEE Transactions on Power Systems, IEEE, Volume 30, Issue 3. pp. 1483-1491.

Urquhart, A.J. et al. (2017). Accurate Determination of Distribution Network Losses. 24th International Conference on Electricity Distribution, CIRED, Paper no. 1076.

Utility Week Intelligence (2015). Smart Meter Quarterly Report. Utility Week Intelligence, July 2015. www.utilityweekintelligence.co.uk

Vadacchino, V., Denda, R., Fantini, G. (2013). First Results in Endesa Smart Metering Roll-out. 22nd International Conference on Electricity Distribution, CIRED, IET, DOI: 10.1049/cp.2013.0850

Valgaev, O. et al. (2016). Low-voltage Power Demand Forecasting Using K-nearest Neighbours Approach. IEEE Innovative Smart Grid Technologies – Asia, DOI: 10.1109/ISGT-Asia.2016.7796525

Vélez, V. M. et al. (2014). Low Voltage Distribution System Planning Using Diversified Demand Curves. International Journal of Electrical Power and Energy Systems, Volume 61, pp. 691-700.

Verho, P. et al. (2004). Aspects on Development of Distribution Network Fault Management Location and Management. Development in Power System Protection, 2004. 8th IEE International Conference, IEEE, Volume, pp. 784-787.

Vujošević, I. et al. (2002). One Method for the Estimation of Voltage Drop in distribution Systems. IEEE Power Engineering Society Summer Meeting, IEEE, Volume 1, pp. 566-569.

Wang, Z. et al. (2004). Interpreting GIS Data for Operation and Control of Distribution Networks. Power Systems Conference and Exposition, 2004, IEEE PES, IEEE Volume 2, pp. 907-912.

Wang, P. et al. (2012). Customer Led Network Revolution-Integrating Renewable Energy into LV Networks Using Energy Storage. CIRED Workshop, Lisbon, CIRED, Paper no. 0223.

Wang, Y. et al. (2016). A Voltage Regulation Method Using Distributed Energy Storage Systems in LV Distribution Networks. IEEE International Energy conference, IEEE, pp. 1-6.

Watson, J. D. et al. (2016). Impact of Solar Photovoltaics on The Low-voltage Distribution Network in New Zealand. IET Journal on Generation, Transmission and Distribution, IET, Volume 10, Issue 1, PP. 1-9.

Weatheronline (2016). www.weatheronline.co.uk, [online], available from: http://www.weatheronline.co.uk/weather/maps/forecastmaps?LANG=en&CONT=ukuk, Date accessed: April 2017.

Western Power Distribution (2013). RIIO-ED1 Business Plan. Western Power Distribution 2013, [online], Available from: https://www.westernpower.co.uk/docs/About-us/Stakeholder-information/Our-future-business-plan/WPD-RII-ED1-Business-Plan/WPD-RIIO-ED1-Business-Plan.aspx

Western Power Distribution (2014). Low Voltage Network Templates Summary Report. Western Power Distribution and Accenture, [online], Available from: https://www.westernpowerinnovation.co.uk/Document-library/2014/WPD-Industry-Report_English.aspx

Wong, K.P. and Cheung, H.N. (2015). Artificial Intelligence Approach to Load Allocation in distribution Substations. IEEE Conference Proceedings-Generation, Transmission and Distribution, IEEE, Volume 134, Issue 5, pp. 357-365.

Wright, A. and Firth, S. (2006) The Nature of Domestic Electricity-Loads and Effects of Time Averaging on Statistics and On-site Generation Calculations. Journal of Applied Energy, Volume 84, pp. 389-403.

Wu, S. (2014). Improved K-nn for Short-Term Traffic Forecasting Using Temporal and Spatial Information, Journal of Transportation Engineering, Volume 140, Issue 7, pp. 1-9.

Wuergler, E. and Vanhemelryck, C. (2012). Use of CIM for Workflows Across Network Operation, Asset Management and Network Planning Systems at a Distribution Utility. Power and Energy Society general meeting, 2012 IEEE San Diego, IEEE, pp.1-4.

Ying-Feng, H. et al. (2012). State Estimation in Electric Power Grids. IEEE Signal Processing Magazine, IEEE, Volume 29, Issue 5, pp. 33-43.

You, S. et al. (2014). An Overview of Trends in Distribution Network Planning. IEEE/PES T&D Conference and Exposition, IEEE, Chicago, USA, DOI: 10.1109/TDC.2014.6863446

Yu, X. et al. (2011). The New Frontier of Smart Grids. IEEE Industrial Electronics Magazine, Volume 5, Issue 3, pp. 49-63.

Yuan, X. et al. (2012). Discussion on The Development Trend of Smart Grid and Its Key Technology. IEEE China International Conference on Electricity Distribution (CICED) Shanghai, IEEE, pp.1-8.

Zhang, J. and Bo, Z. Q. (2010) Research of the Impact of Distribution Generation on Distribution Network Loss. 45th International Universities Power Engineering Conference, IEEE, pp. 1-4.

Zhao, Y. et al. (2013). Outage Detection in Power Distribution Networks with Optimally-Deployed Power Flow Sensors. Proceedings of IEEE Power and Energy Society General Meeting, IEEE.

Zhou, L. et al. (2014). Cost/Benefit Assessment of a Smart Distribution System with Intelligent Electric Vehicle Charging. IEEE Transactions on Smart Grid, IEEE, Volume 5, Issue 2.

Appendix A: Python Code for Extraction of the CLNR Data Files

The code in this section can used in Python to extract the smart meter data for each meter ID as CSV separate files.

```
# Start of code

# CLNR_filename = "zati.txt"

CLNR_filenameA = "TrialMonitoringDataT5.csv"

OutputFilestem = "chunk_"

MaximumNumberOfFiles = 200

MaximumNumberOfCustomersPerFile = 1

# ##########################################################

CLNR_file = open(CLNR_filenameA,"r")

try:

    print("Opened file ", CLNR_filenameA)

    # Get rid of header line
```

```
CLNR_file.readline()


OLD_CustomerID = "Empty"
NumberOfCustomers = 0
OutputFileNumber = 1
OutputFilename = OutputFilestem + str(OutputFileNumber) + ".csv"
Output_file = open(OutputFilename,"w")
try:
    ContinueReading = True
except:
    ContinueReading = False
    print("FAILED to open ", OutputFilename)


while ContinueReading == True:
    CLNR_row = CLNR_file.readline()
    try:
        CustomerID = CLNR_row.split(",", 4)
        #
        ReadingType = " "
        if CustomerID[1] == "solar power":
            ReadingType = "s"
        if CustomerID[1] == "whole home power import":
            ReadingType = "w"
```

```
# Date is followed by time with only a space as separator
old_date_time = CustomerID[3].split(" ", 2)
new_date_time = old_date_time[0] + "," + old_date_time[1]
if OLD_CustomerID == CustomerID[0]:
    new_CLNR_row = CustomerID[0] + "," + ReadingType + "," + new_date_time + "," + CustomerID[4]
    Output_file.write(new_CLNR_row)
    # Output_file.write(CLNR_row)
else:
    NumberOfCustomers = NumberOfCustomers + 1
    if NumberOfCustomers > MaximumNumberOfCustomersPerFile:
        NumberOfCustomers = 1
        OutputFileNumber = OutputFileNumber + 1
        if OutputFileNumber > MaximumNumberOfFiles:
            ContinueReading = False
        else:
            Output_file.close()
            OutputFilename = OutputFilestem + str(OutputFileNumber) + ".csv"
            Output_file = open(OutputFilename,"w")
            new_CLNR_row = CustomerID[0] + "," + ReadingType + "," + new_date_time + "," + CustomerID[4]
            Output_file.write(new_CLNR_row)
    else:
```

```
                new_CLNR_row = CustomerID[0] + "," + ReadingType + "," +
new_date_time + "," + CustomerID[4]

                Output_file.write(new_CLNR_row)

            OLD_CustomerID = CustomerID[0]

        except:

            ContinueReading = False



    Output_file.close()

    CLNR_file.close()

except:

    print("FAILED to open ", CLNR_filename)



# Tidy up
```

Appendix B: Characteristics of Northern PowerGrid Low Voltage Cables

HELLENIC CABLES S.A.

HELLENIC CABLE INDUSTRY S.A.

CABLEL®

Electrical Data:		
Frequency:	50	Hz
Maximum conductor's temperature at continuous operation:	70	°C
Maximum conductor DC resistance at 20°C:	0.164	Ω/km
Calculated conductor AC resistance at maximum operating temperature:	0.199	Ω/km
Calculated inductive reactance:	0.08	Ω/km
Calculated phase capacitance:	0.723	μF/km
Calculated charging current: *Based on the calculated phase capacitance and operating phase-to-ground voltage*	0.14	mA/m/phase
Zero sequence impedance: *Return through copper wire screen only, resistance calculated at maximum operation temperature*	0.81 + j0.106	Ω/km
Continuous current carrying capacity of cables laid directly in ground - *Soil thermal resistivity: 1.2 K.m/W* - *Depth of laying: 0.45 m* - *Ground temperature: 15°C* - *Load factor: 1.0* - *One circuit* Continuous current currying capacity of cables laid in air (not exposed in sunlight) - *Air temperature: 25°C* - *Load factor: 1.0* - *One circuit*		
A *Buried in ground*		
Current:	304	A, for each phase
B *Laid in air*		
Current:	295	A, for each phase
Installation Data:		
Maximum pulling force with pulling eye:	5.55 (for each conductor)	kN
Minimum installation temperature:	0	°C
Minimum bending radius during installation directly in ground:	12 x cable diameter	m

Υ.Σ.:	557/2010	Design & Development Department	
T.M.K.:	258/2010	Detailed by:	K. Tastavridis
Date – Revision:	21/05/2010 – 0	Checked by:	D. Tsiavos
Client – Destination country:	UNITED UTILITIES	Approved by:	G. Georgallis

Figure B-1: Characteristics of 185 mm^2 LV cable (cables B and C)

HELLENIC CABLES S.A.
HELLENIC CABLE INDUSTRY S.A.

CABLEL®

Electrical Data:		
Frequency:	50	Hz
Maximum conductor's temperature at continuous operation:	70	°C
Maximum conductor DC resistance at 20°C:	0.100	Ω/km
Calculated conductor AC resistance at maximum operating temperature:	0.123	Ω/km
Calculated inductive reactance:	0.079	Ω/km
Calculated phase capacitance:	0.816	μF/km
Calculated charging current: *Based on the calculated phase capacitance and operating phase-to-ground voltage*	0.15	mA/m/phase
Zero sequence impedance: *Return through copper wire screen only, resistance calculated at maximum operation temperature*	0.73 + j0.098	Ω/km
Continuous current carrying capacity of cables laid directly in ground *- Soil thermal resistivity: 1.2 K.m/W* *- Depth of laying: 0.45 m* *- Ground temperature: 15°C* *- Load factor: 1.0* *- One circuit* Continuous current currying capacity of cables laid in air (not exposed in sunlight) *- Air temperature: 25°C* *- Load factor: 1.0* *- One circuit*		
A — *Buried in ground*		
Current:	402	A, for each phase
B — *Laid in air*		
Current:	407	A, for each phase
Installation Data:		
Maximum pulling force with pulling eye:	9 (for each conductor)	kN
Minimum installation temperature:	0	°C
Minimum bending radius during installation directly in ground:	12 x cable diameter	m

Υ.Σ.:	557/2010	Design & Development Department	
T.M.K.:	258/2010	Detailed by:	K. Tastavridis
Date – Revision:	21/05/2010 – 0	Checked by:	D. Tsiavos
Client – Destination country:	UNITED UTILITIES	Approved by:	G. Georgallis

Figure B-2: Characteristics of 300 mm2 LV cable (cable A)

Appendix C: Minimum Voltage Level Estimates on the Yellow and Blue Phases at Various Time Resolution Intervals-Balanced Network

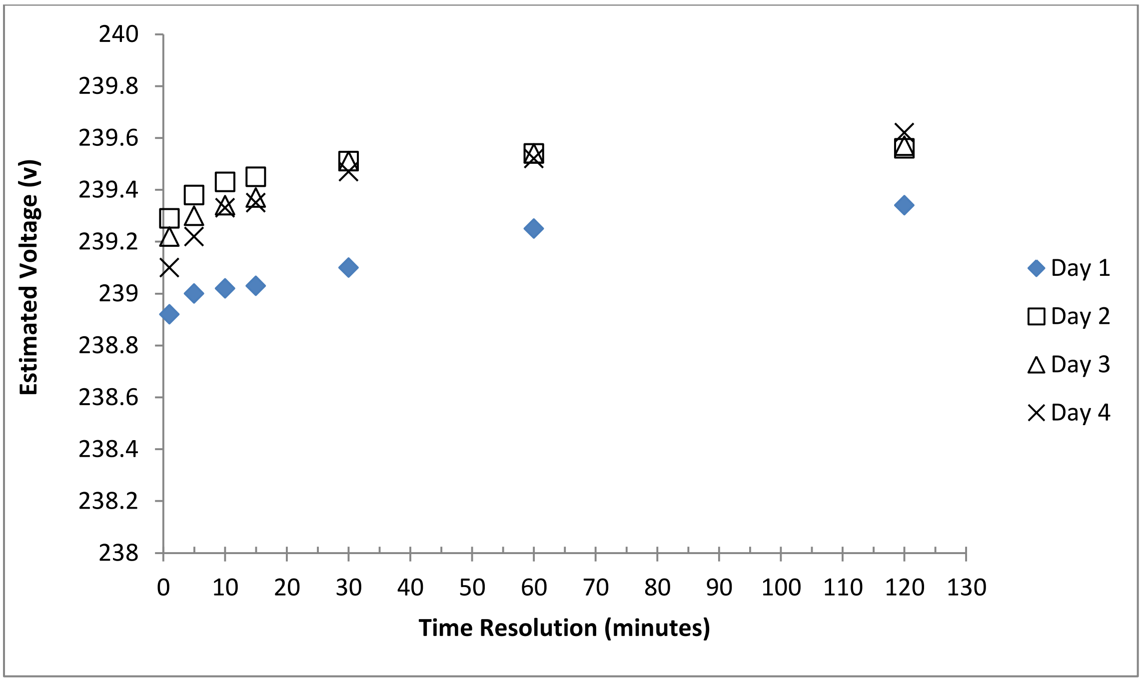


Figure C-1: Minimum voltage levels on the yellow phase estimated using various smart meter time resolution intervals (Loughborough data set)

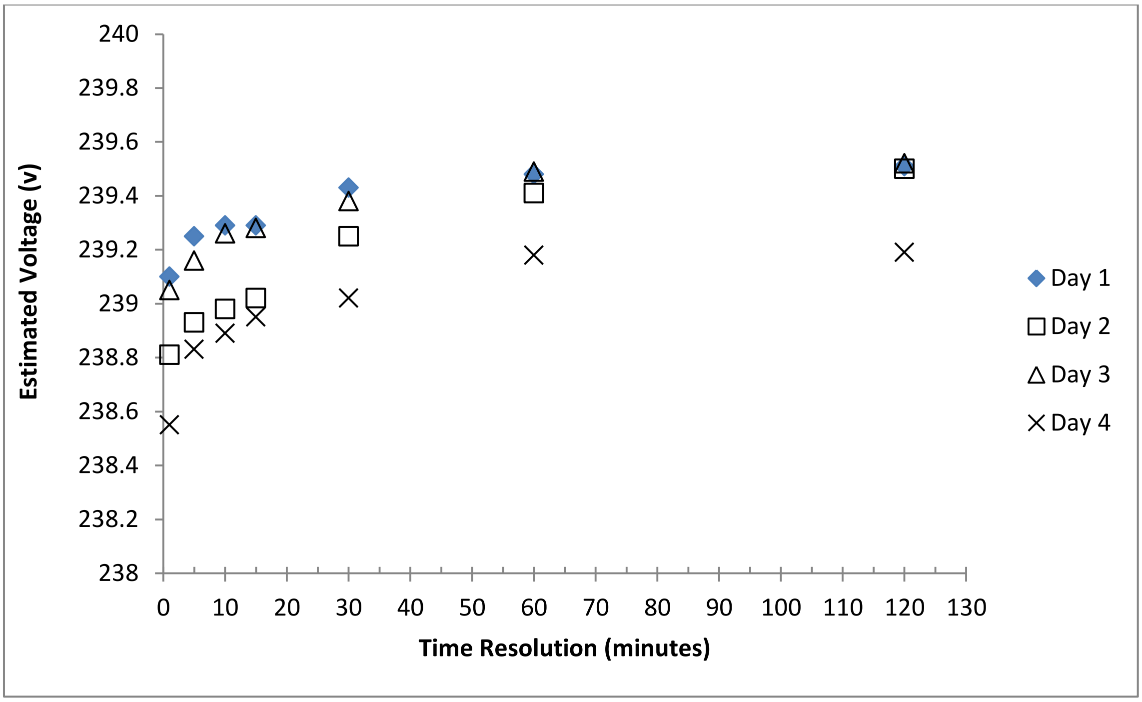


Figure C-2: Minimum voltage levels on the yellow phase estimated using various smart meter time resolution intervals (CLNR data set no.8)

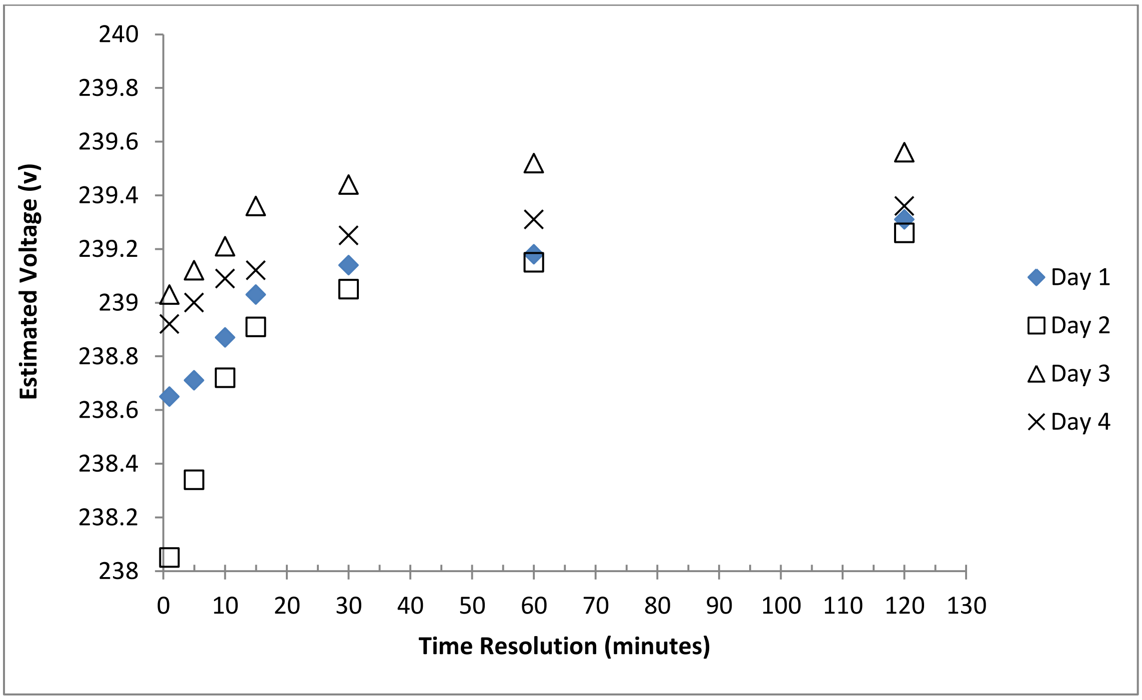


Figure C-3: Minimum voltage levels on the blue phase estimated using various smart meter time resolution intervals (Loughborough data set)

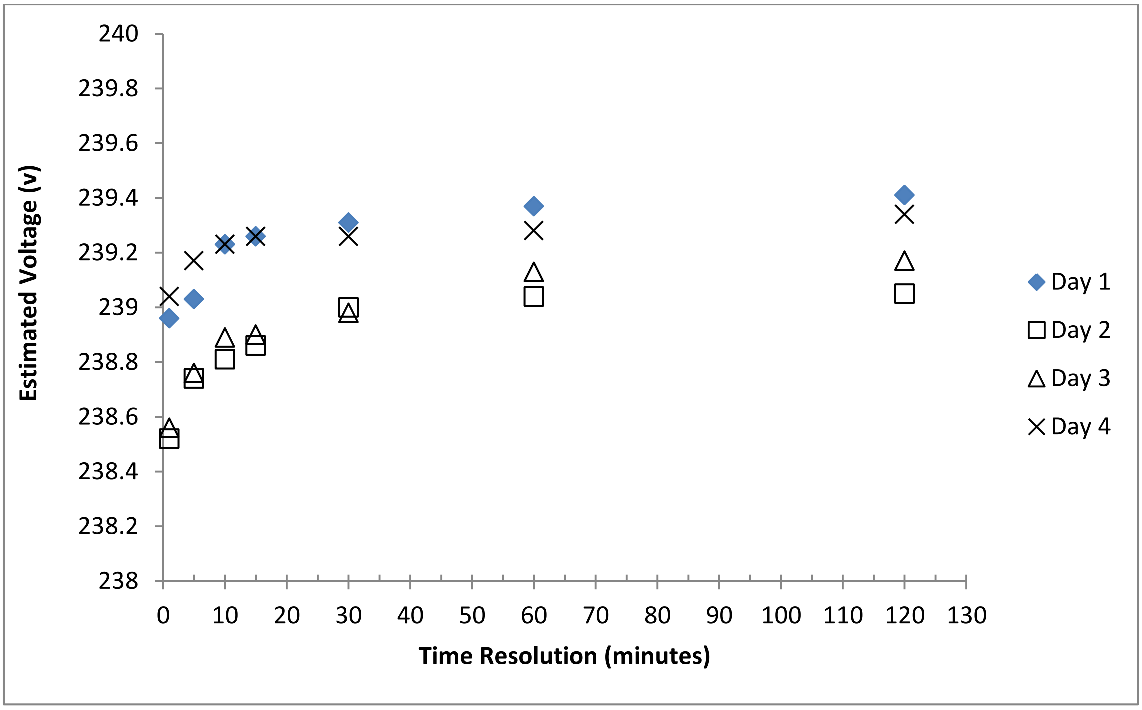


Figure C-4: Minimum voltage levels on the blue phase estimated using various smart meter time resolution intervals (CLNR data set no.8)

Appendix D: Minimum Voltage Level Estimates on the Yellow and Blue Phases at Various Time Resolution Intervals-Unbalanced Network

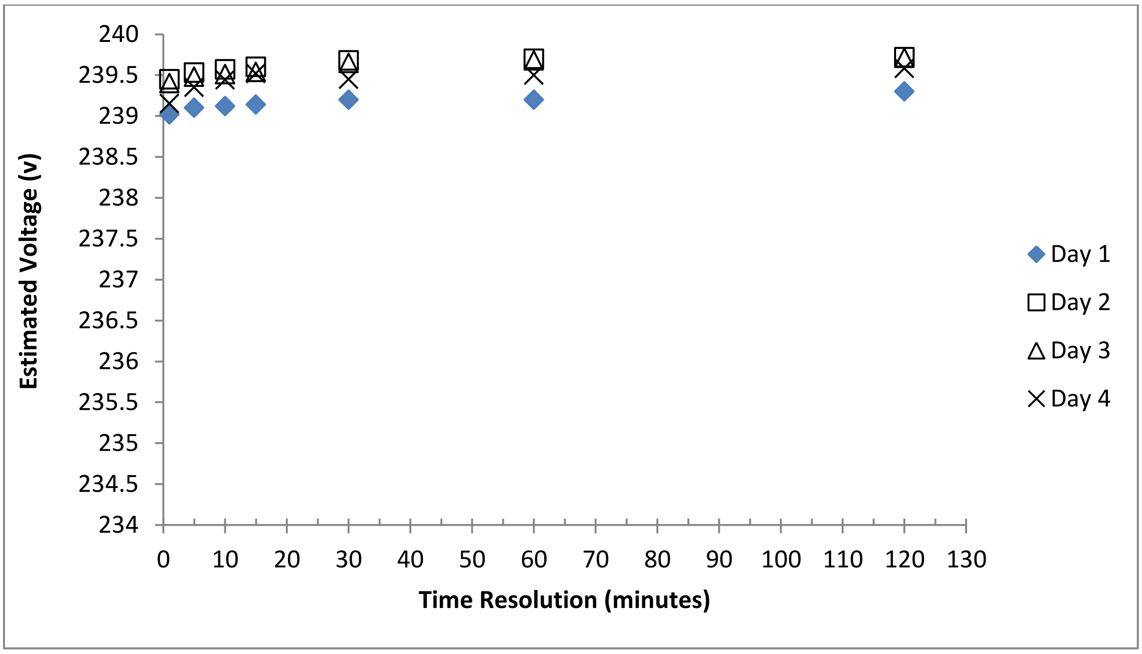


Figure D-1: Minimum voltage levels on the blue phase estimated using various smart meter time resolution intervals in an unbalanced network (Loughborough data set)

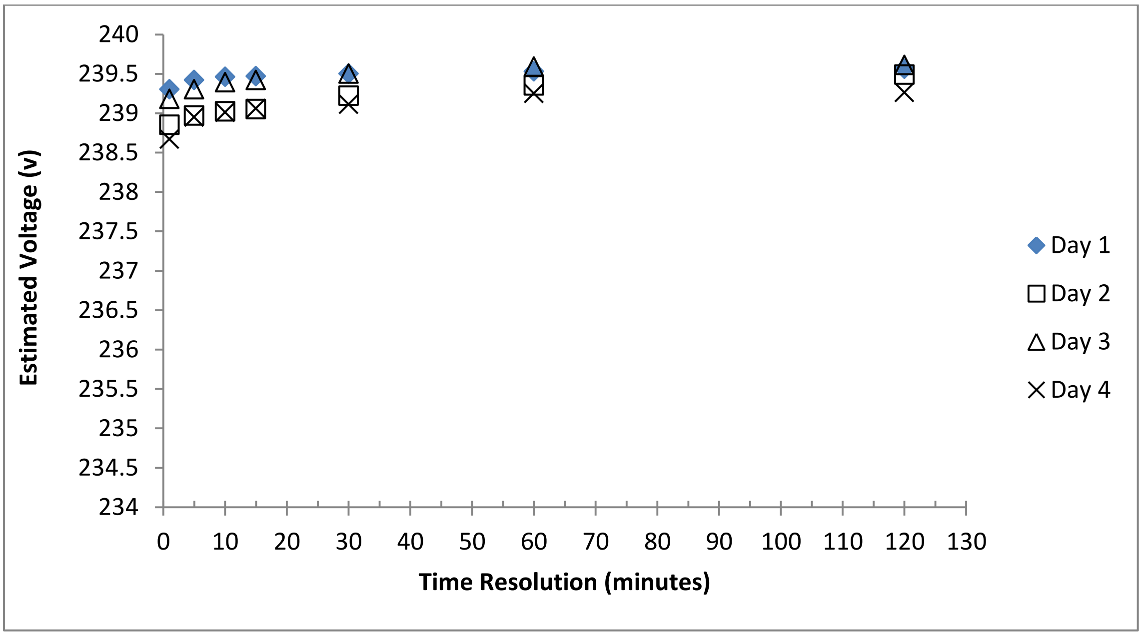


Figure D-2: Minimum voltage levels on the blue phase estimated using various smart meter time resolution intervals in an unbalanced network (CLNR data set no.8)

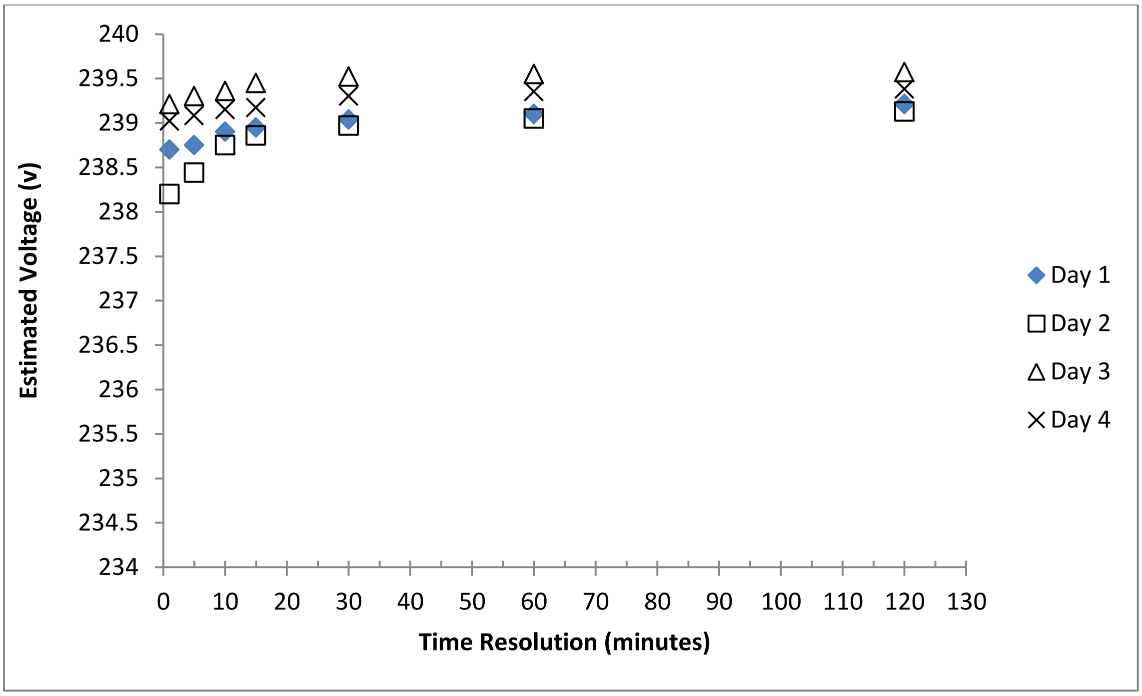


Figure D-3: Minimum voltage levels on the blue phase estimated using various smart meter time resolution intervals in an unbalanced network (Loughborough data set)

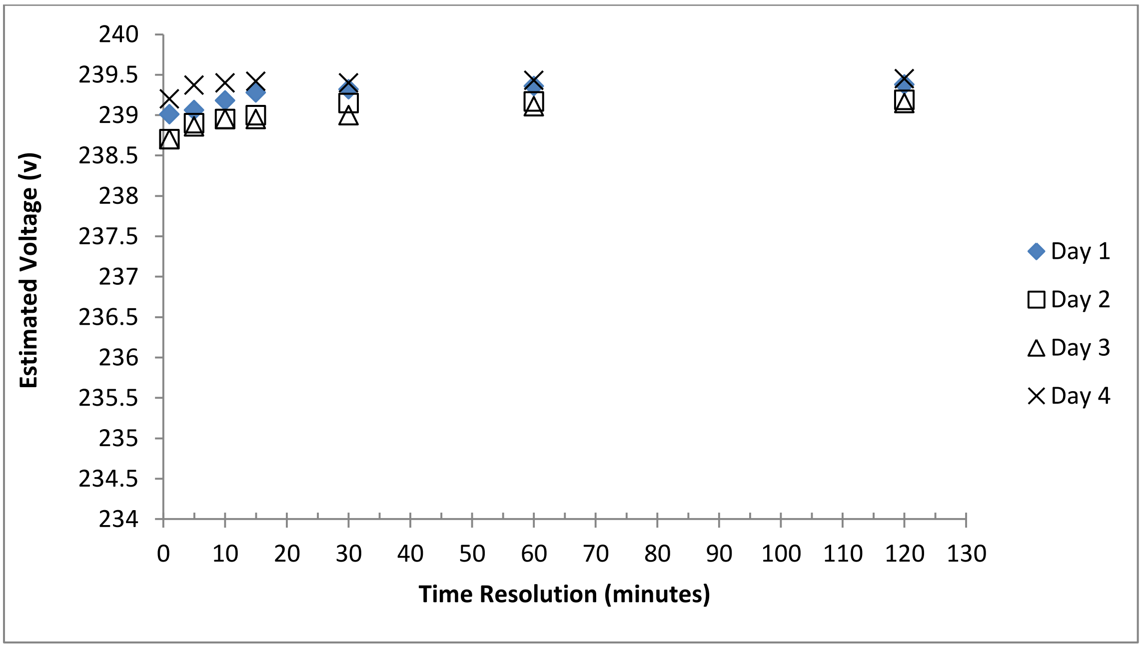


Figure D-4: Minimum voltage levels on the blue phase estimated using various smart meter time resolution intervals in an unbalanced network (CLNR data set no.8)

Appendix E: Cable Capacity Percentage Estimation on Various Phases of Cables B and C at Different Time Resolutions

Figure E-1: Red phase cable loading percentages on cable B (16/01/2008)

Figure E-2: Yellow phase cable loading percentages on cable B (16/01/2008)

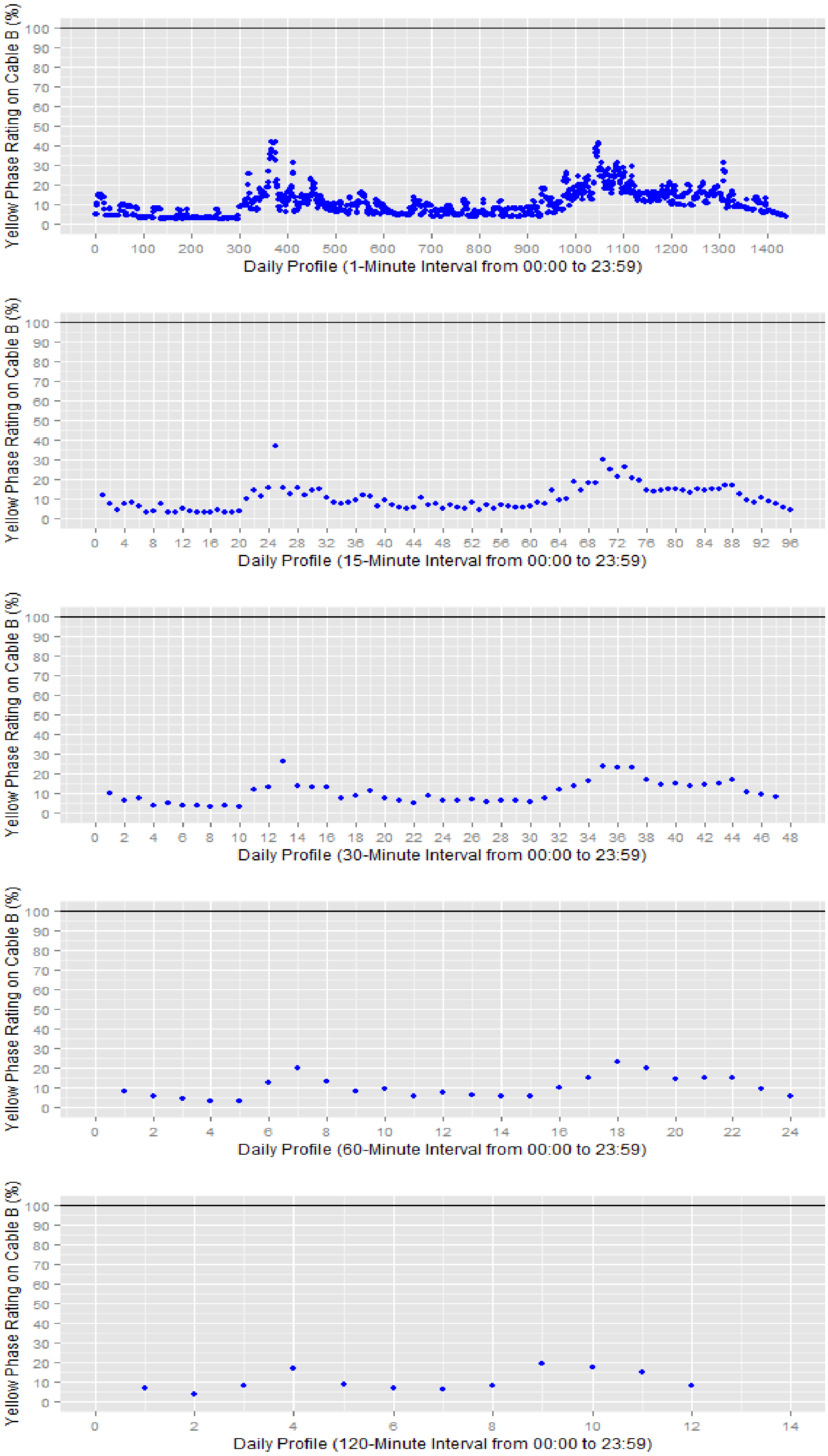


Figure E-3: Blue phase cable loading percentages on cable B (16/01/2008)

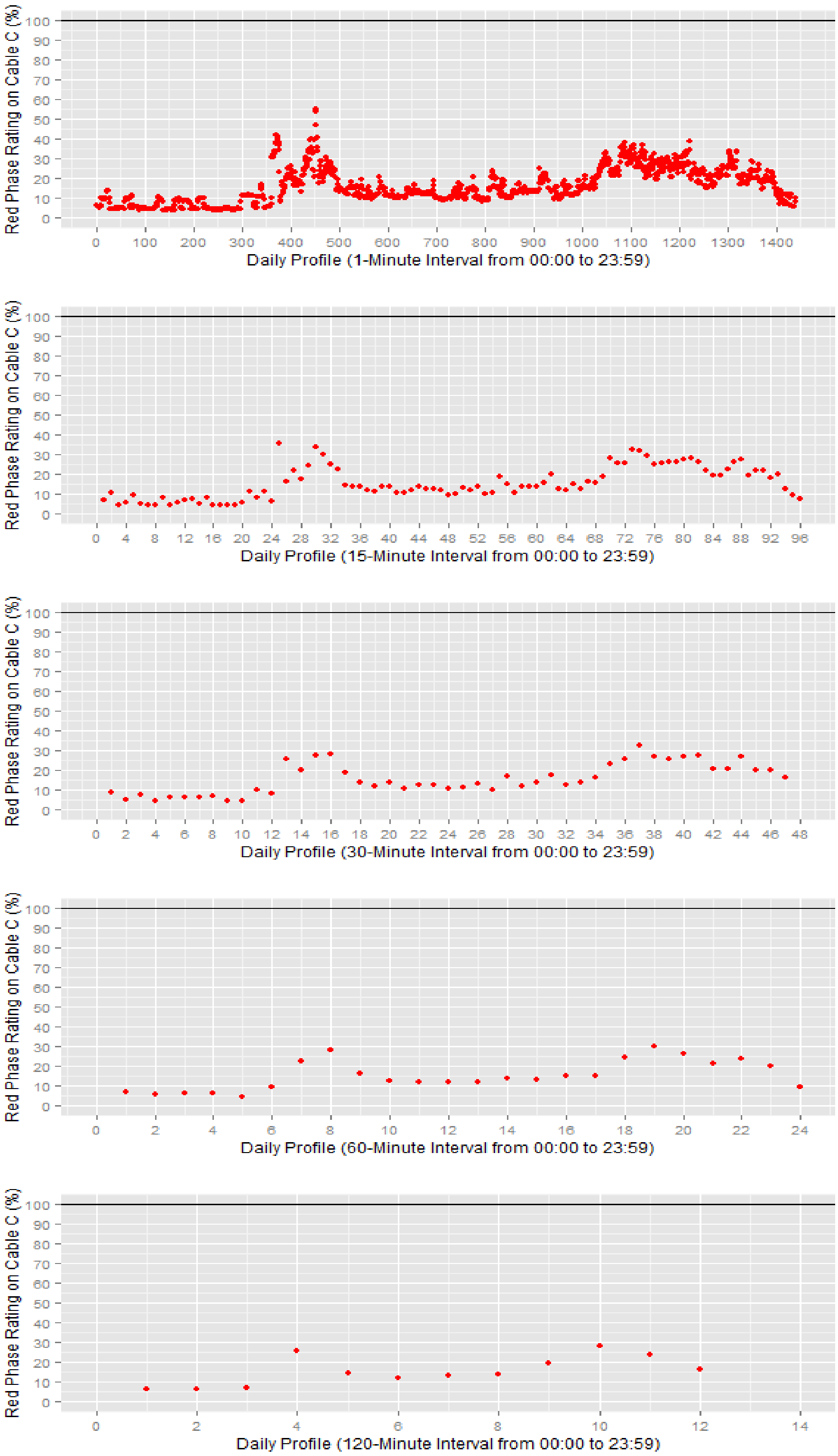


Figure E-4: Red phase cable loading percentages on cable C (16/01/2008)

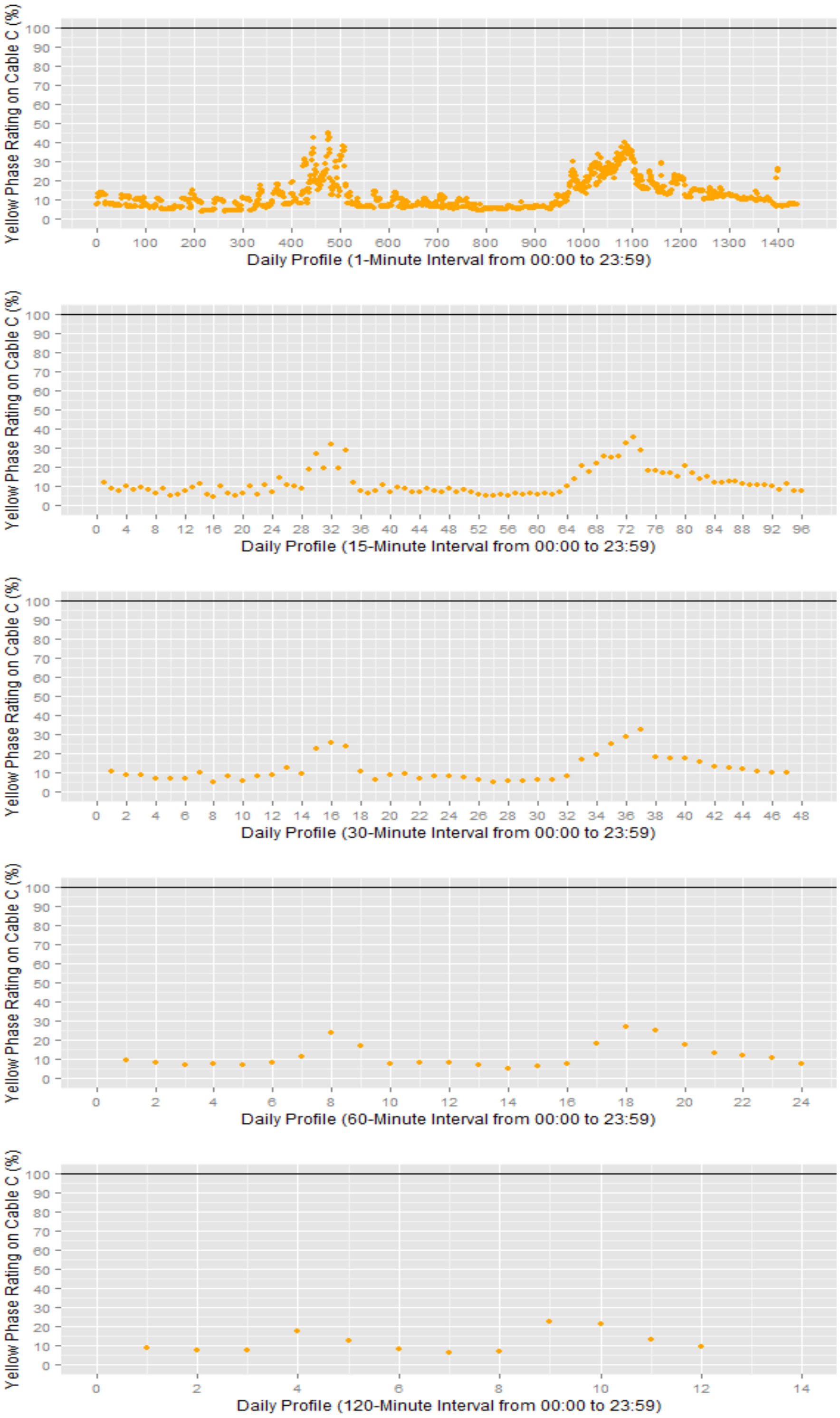


Figure E-5: Yellow phase cable loading percentages on cable C (16/01/2008)

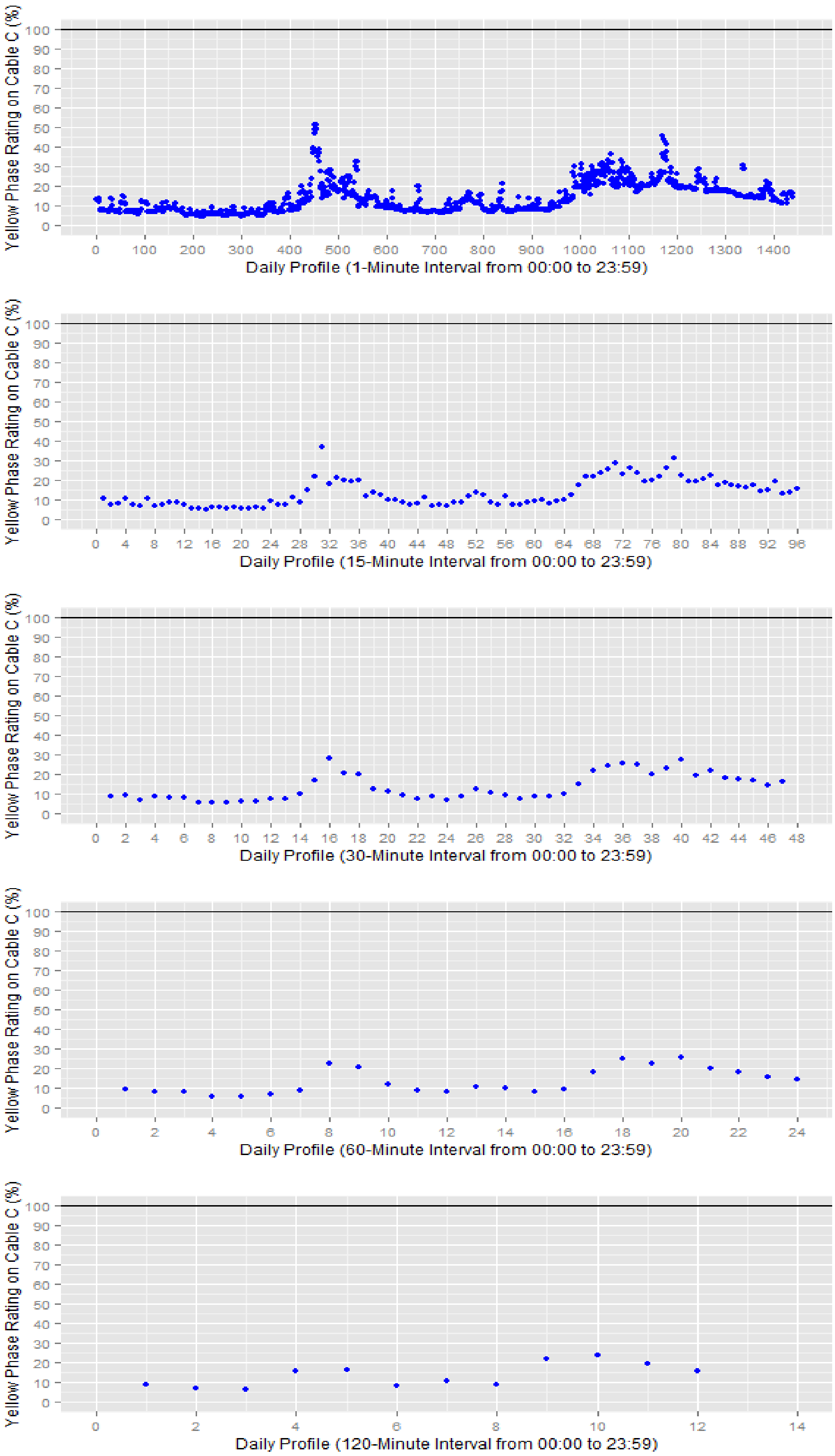


Figure E-6: Blue phase cable loading percentages on cable C (16/01/2008)

Appendix F: Density of Load Percentage Estimates on Cables B and C at Various Time Resolutions

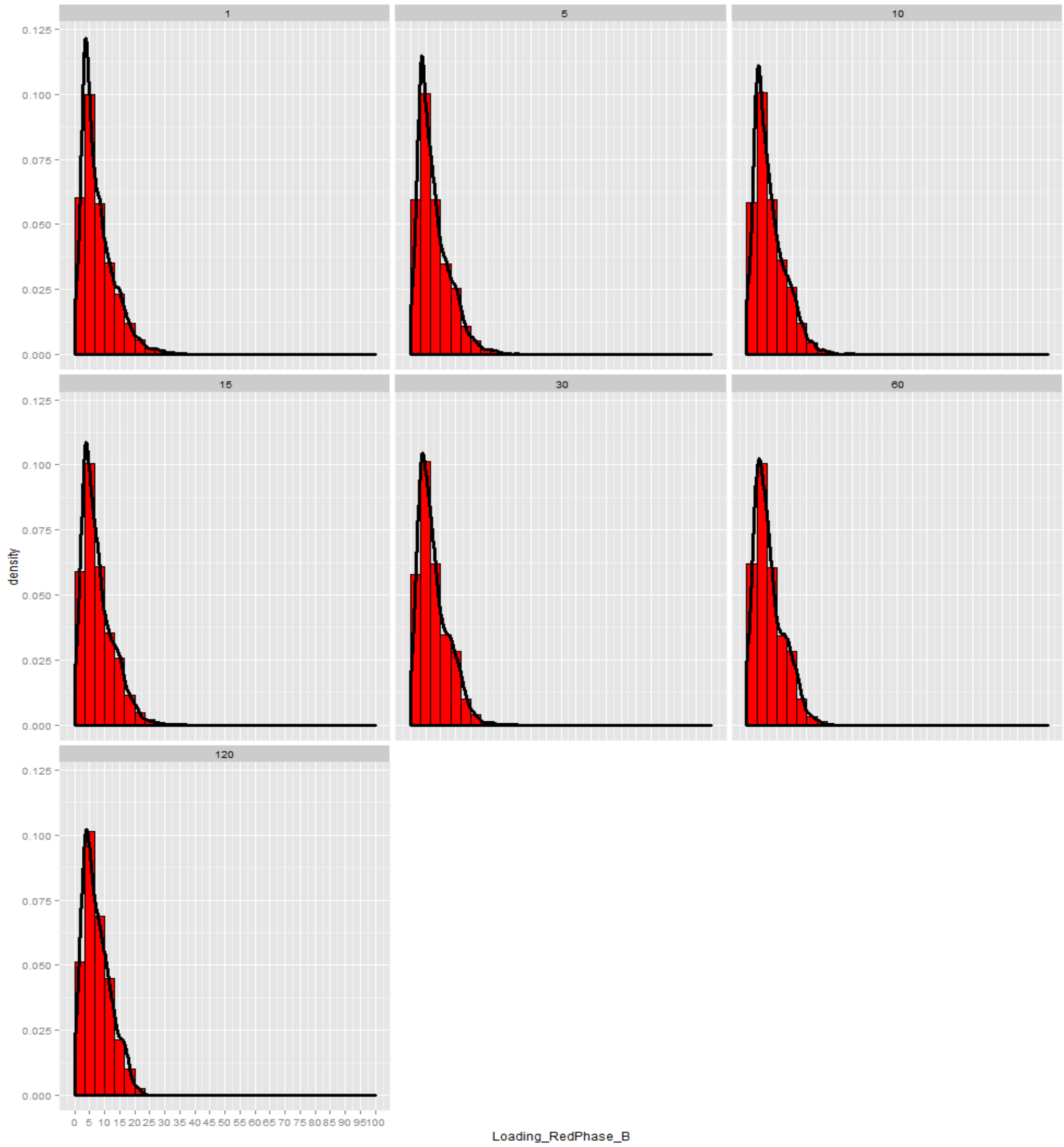


Figure F-1: Density plots of loading percentages frequency at each time resolution interval on cable B (Red phase)

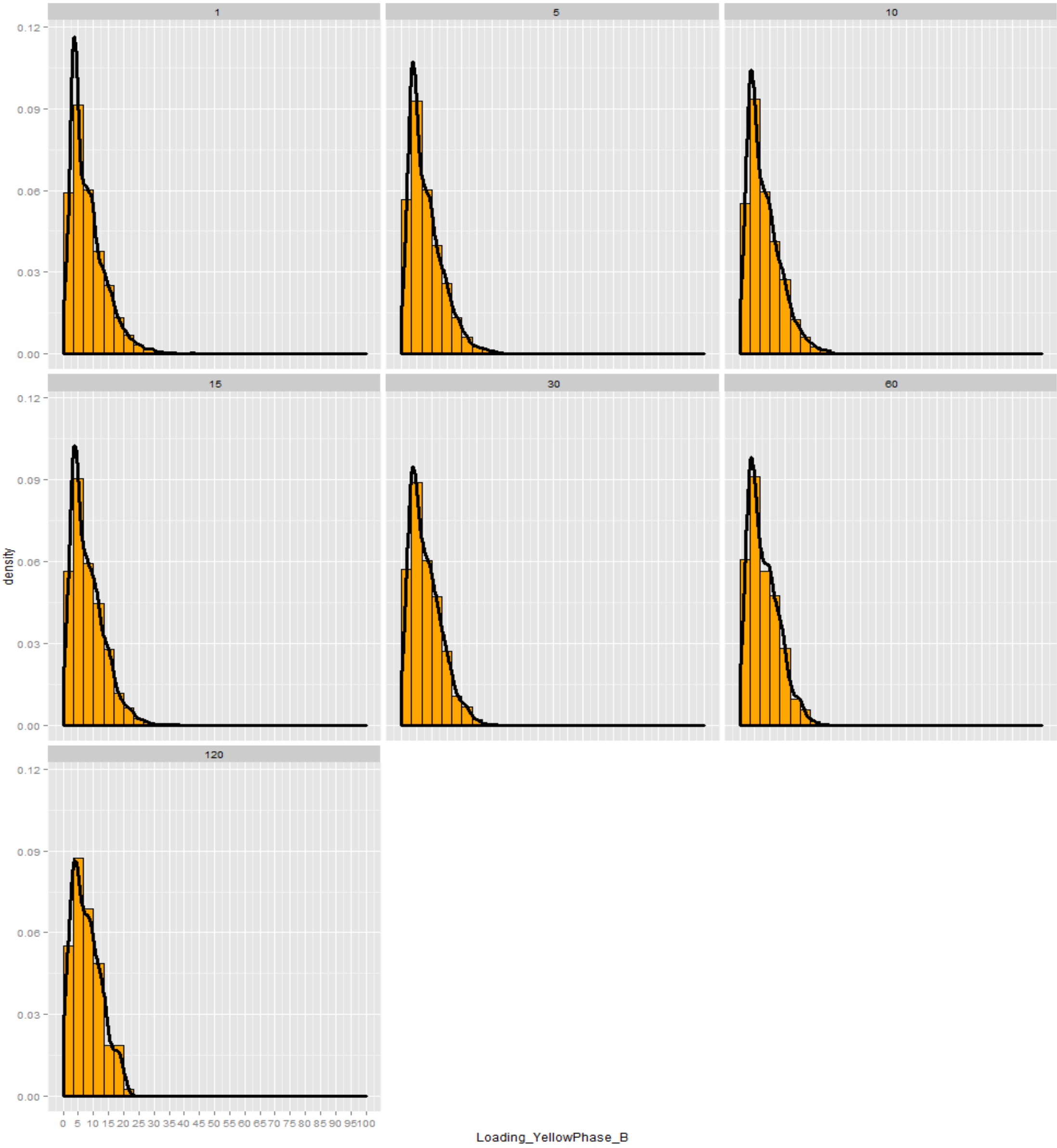


Figure F-2: Density plots of loading percentages frequency at each time resolution interval on cable B (Yellow phase)

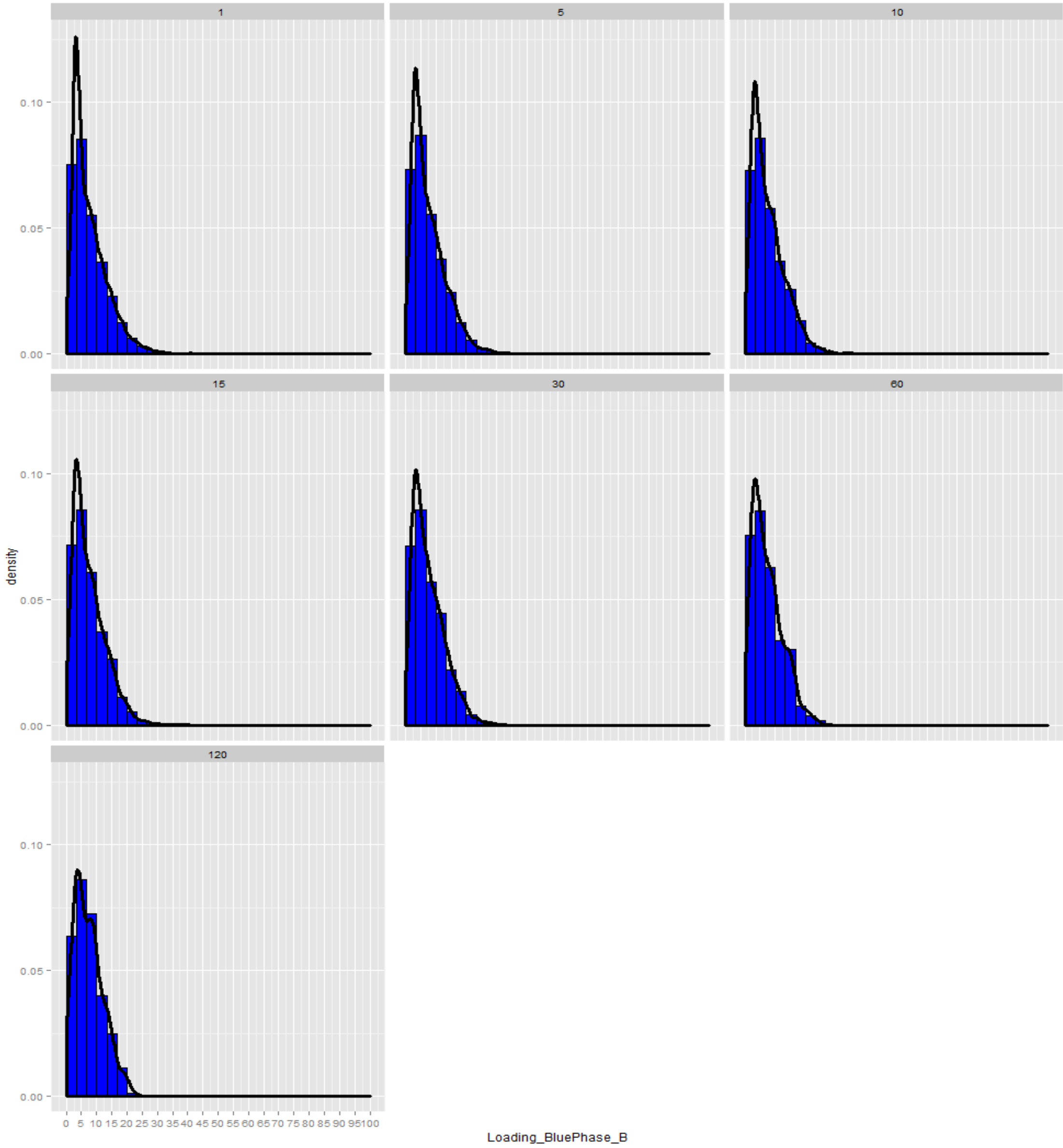


Figure F-3: Density plots of loading percentages frequency at each time resolution interval on cable B (Blue phase)

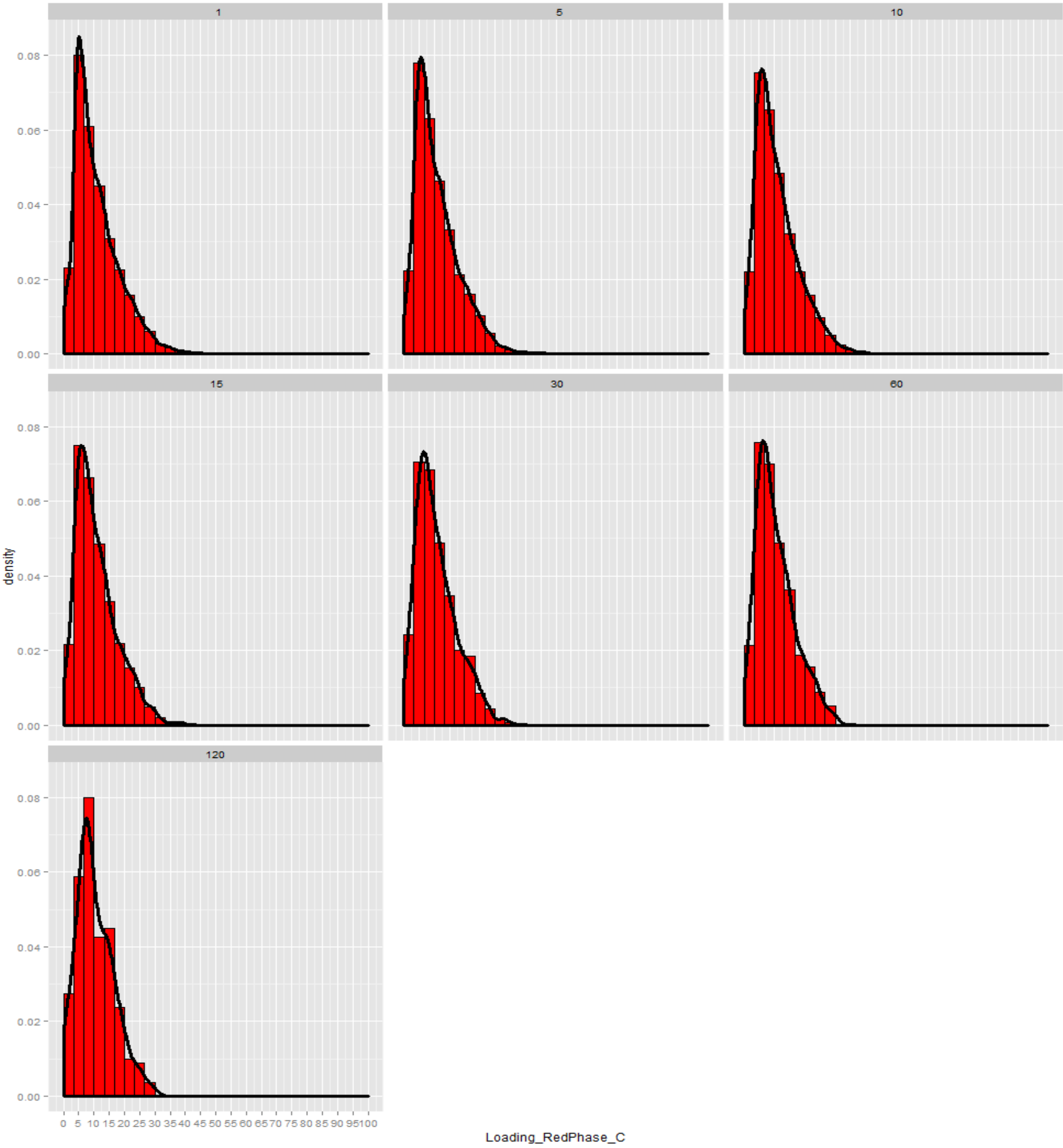


Figure F-4: Density plots of loading percentages frequency at each time resolution interval on cable C (Red phase)

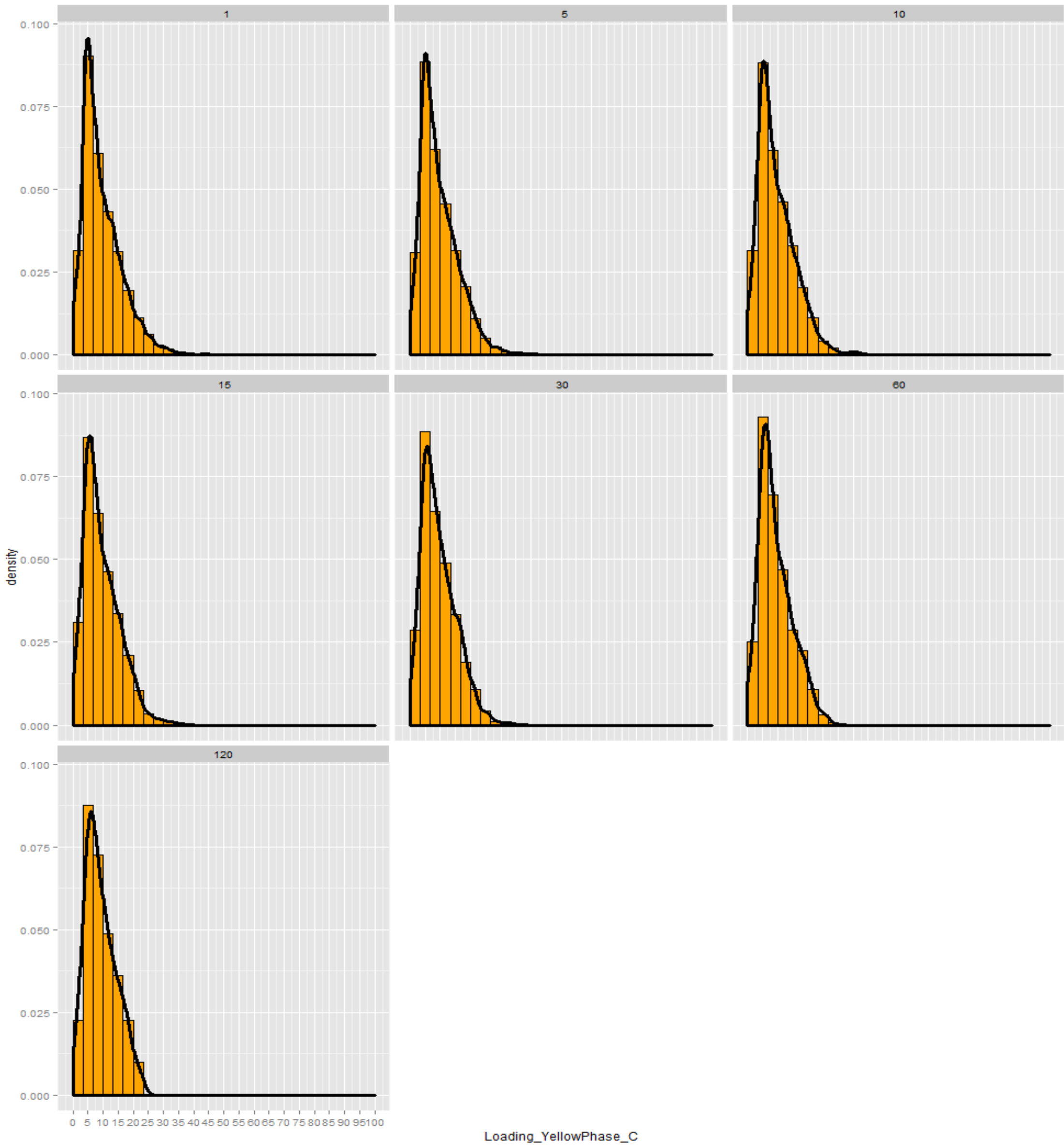


Figure F-5: Density plots of loading percentages frequency at each time resolution interval on cable C (Yellow phase)

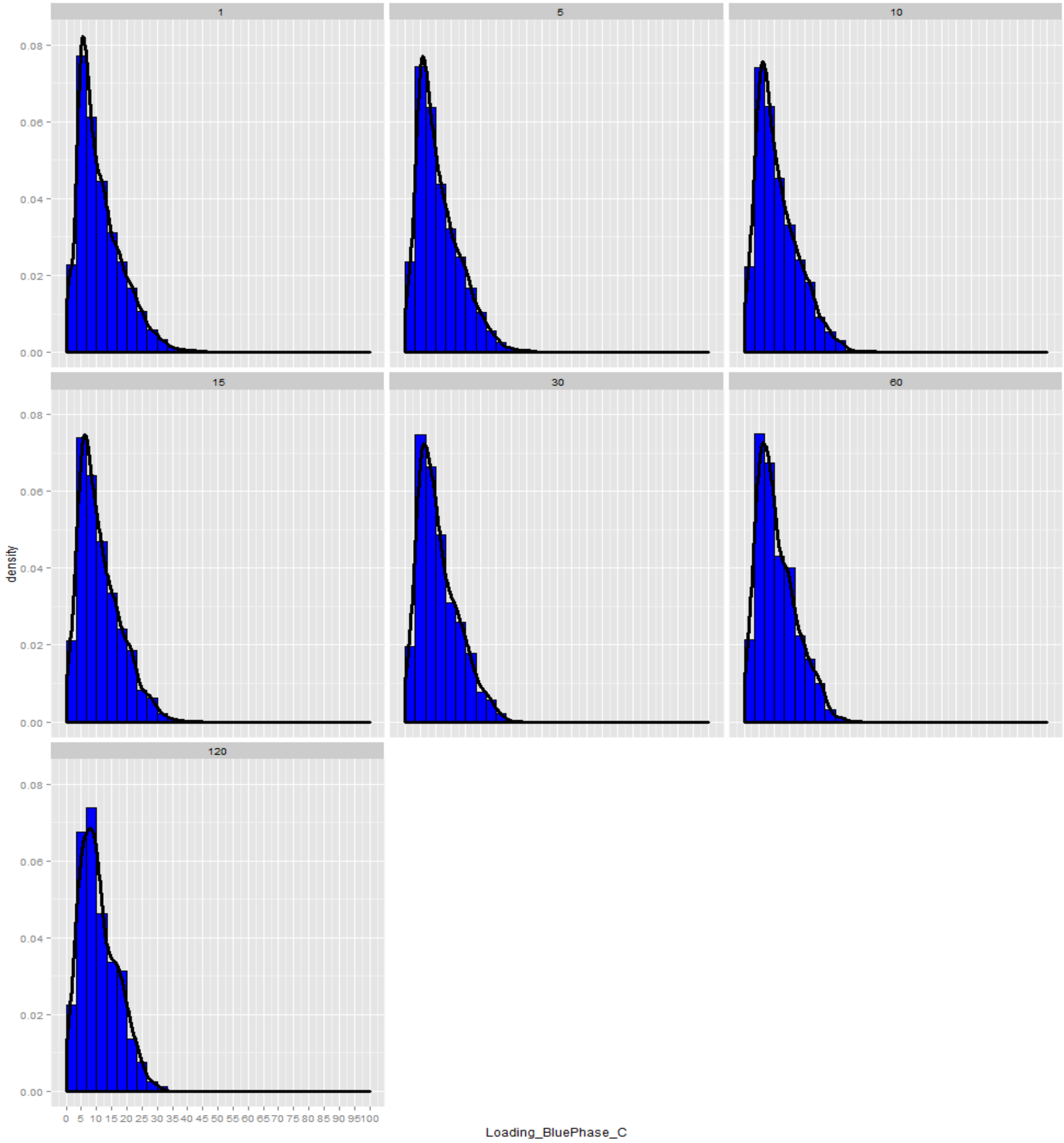


Figure F-6: Density plots of loading percentages frequency at each time resolution interval on cable C (Blue phase)

Appendix G: Minimum Voltage Level Estimates on the Yellow and Blue Phases at Various Aggregation Levels-Balanced Network

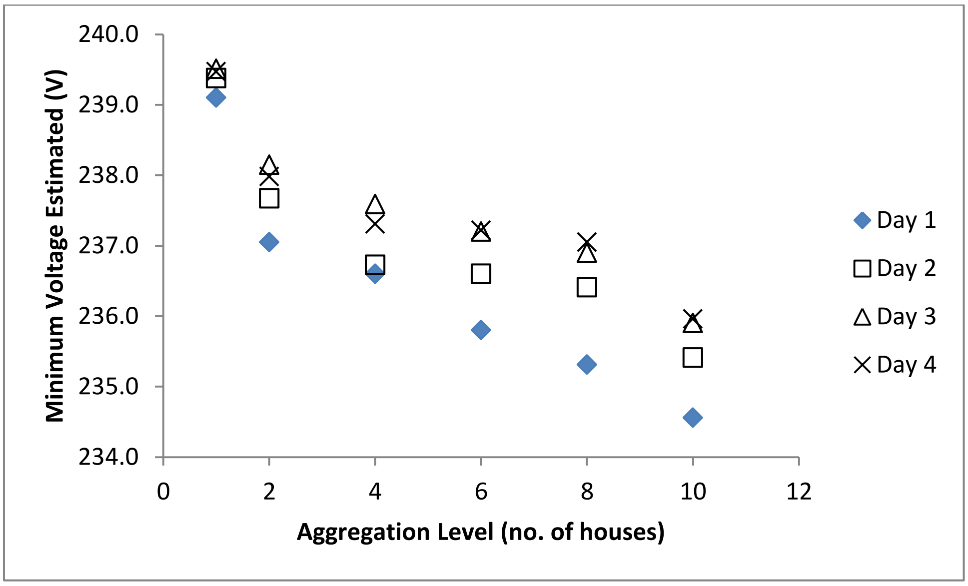


Figure G-1: Minimum voltage levels at various aggregation levels on the yellow phase estimated at various aggregation levels (Loughborough data set)

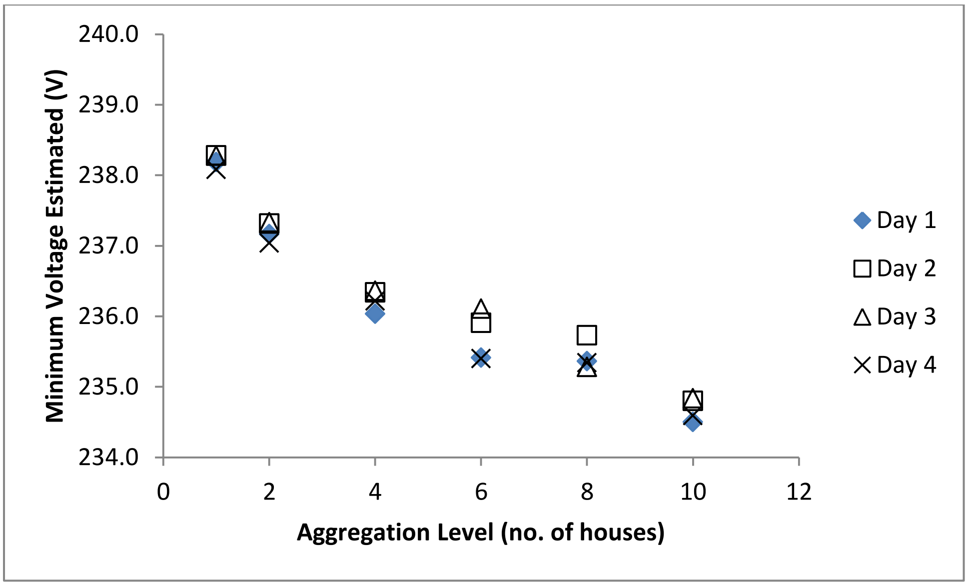


Figure G-2: Minimum voltage levels at various aggregation levels on the yellow phase estimated at various aggregation levels (CLNR data set no.8)

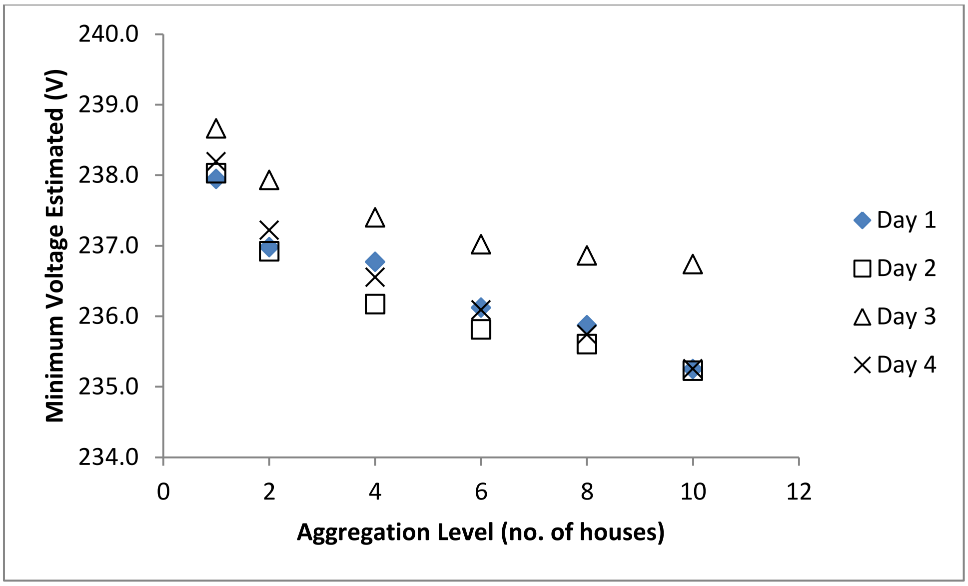


Figure G-3: Minimum voltage levels at various aggregation levels on the blue phase estimated at various aggregation levels (Loughborough data set)

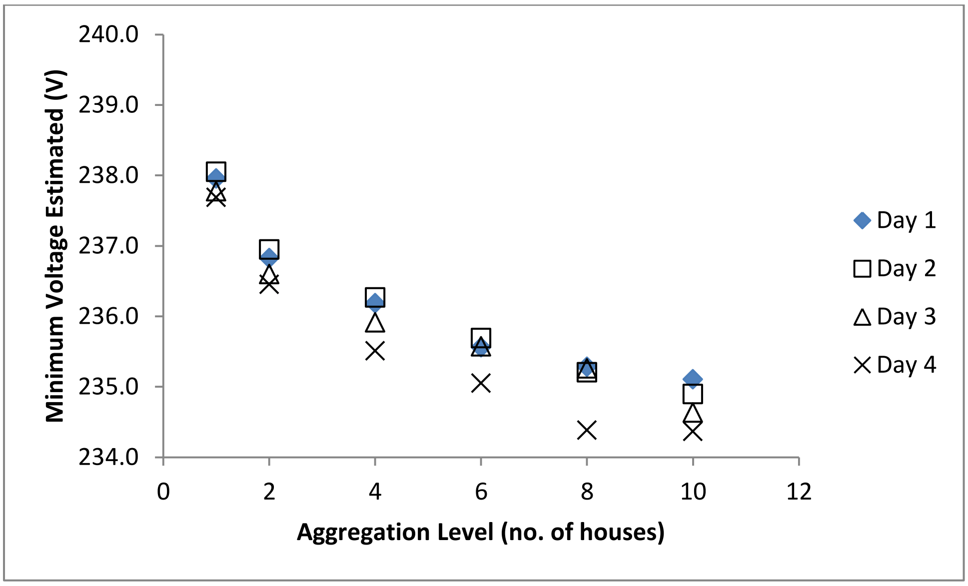


Figure G-4: Minimum voltage levels at various aggregation levels on the blue phase estimated at various aggregation levels (CLNR data set no.8)

Alternative Model

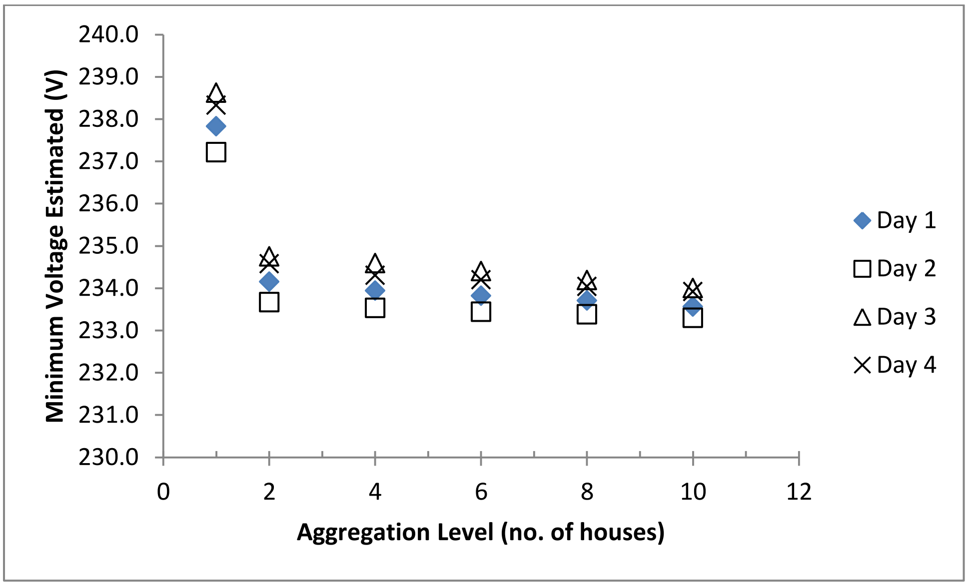


Figure G-5: Minimum voltage levels at various aggregation levels on the yellow phase estimated at various aggregation levels-alternative topology (Loughborough data set)

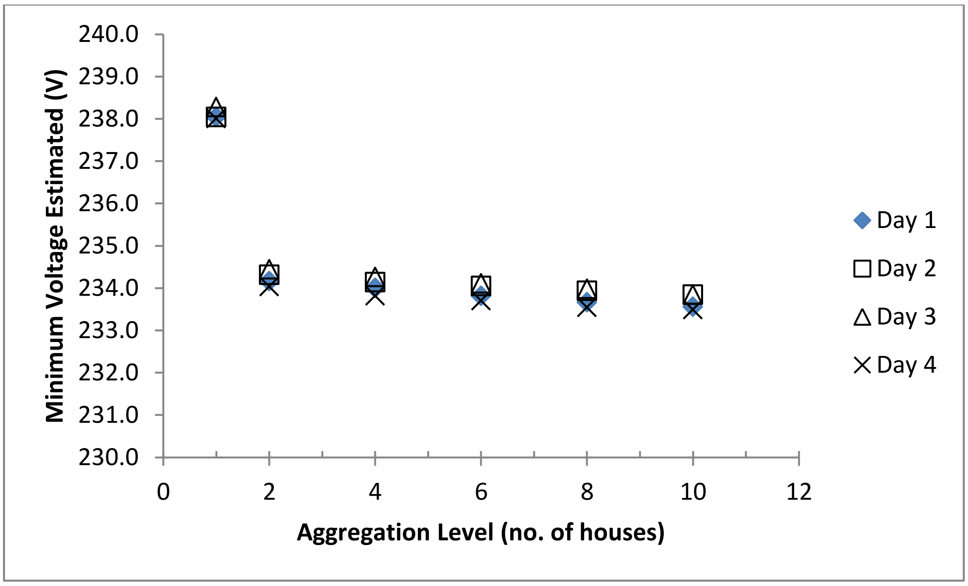


Figure G-6: Minimum voltage levels at various aggregation levels on the yellow phase estimated at various aggregation levels-alternative topology (CLNR data set no.8)

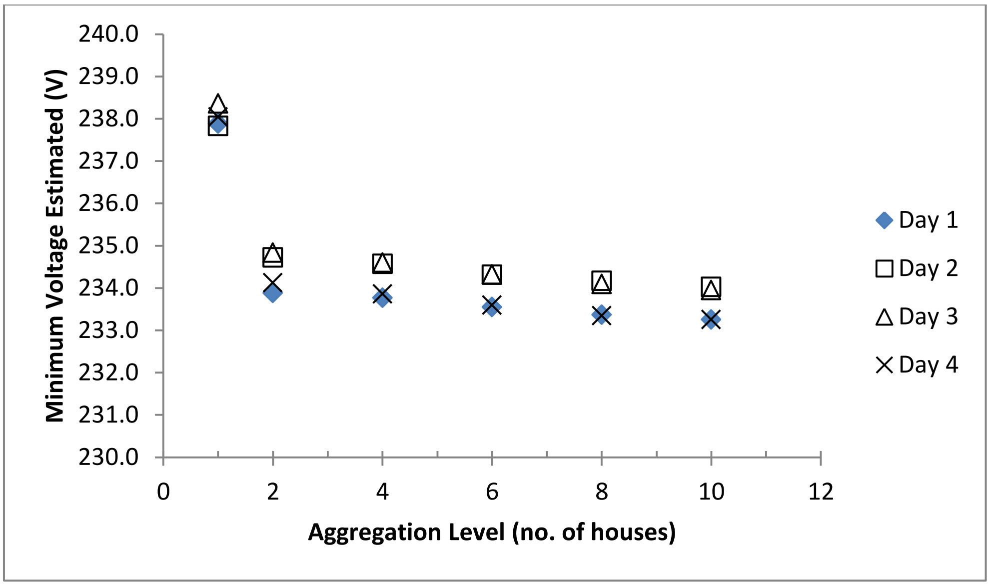


Figure G-7: Minimum voltage levels at various aggregation levels on the blue phase estimated at various aggregation levels-alternative topology (Loughborough data set)

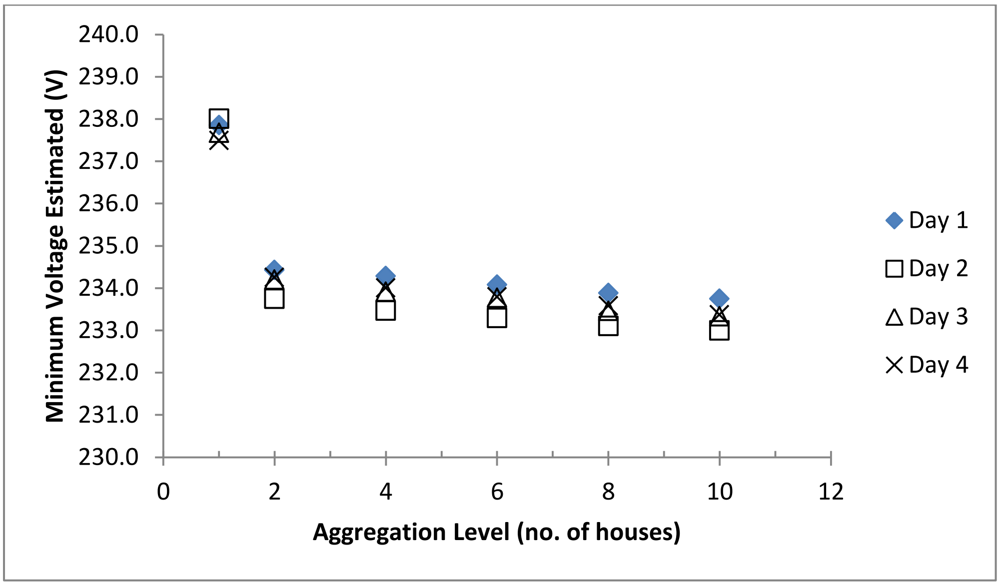


Figure G-8: Minimum voltage levels at various aggregation levels on the blue phase estimated at various aggregation levels-alternative topology (CLNR data set no.8)

Appendix H: Minimum Voltage Level Estimates on the Yellow and Blue Phases at Various Aggregation Levels-Unbalanced Network

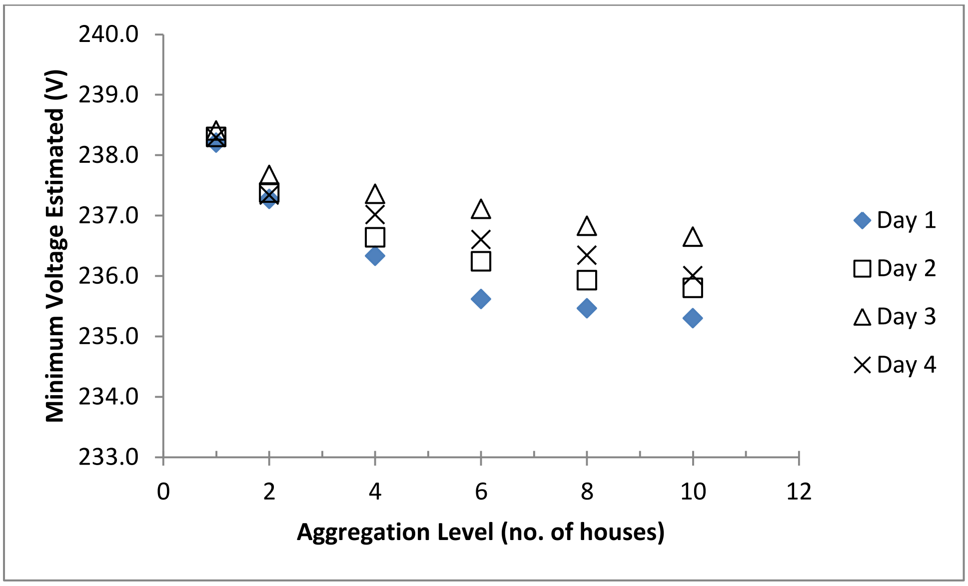


Figure H-1: Estimated minimum voltage levels on the yellow phase at various aggregation levels-unbalanced network (Loughborough data set)

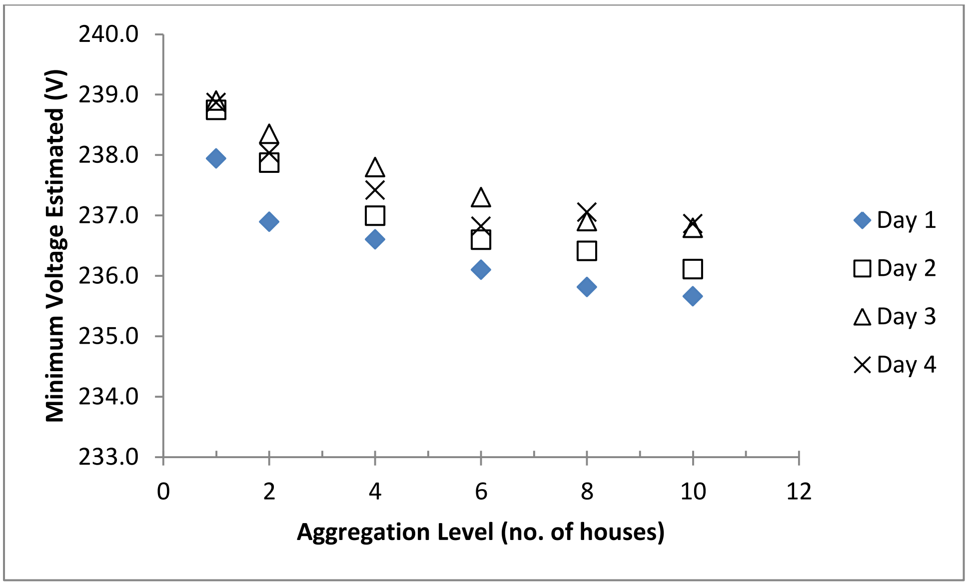


Figure H-2: Estimated minimum voltage levels on the yellow phase at various aggregation levels-unbalanced network (CLNR data set no.8)

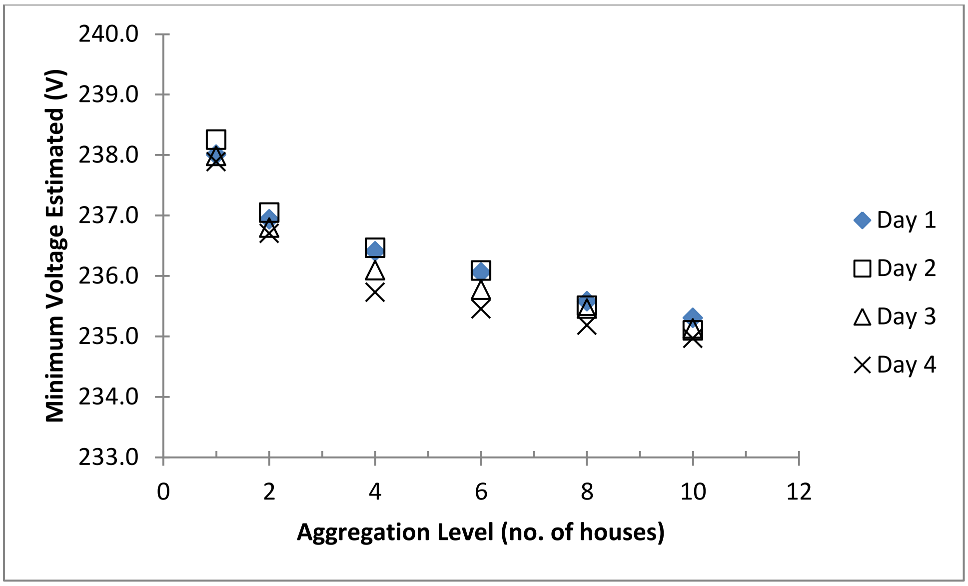


Figure H-3: Estimated minimum voltage levels on the blue phase at various aggregation levels-unbalanced network (Loughborough data set)

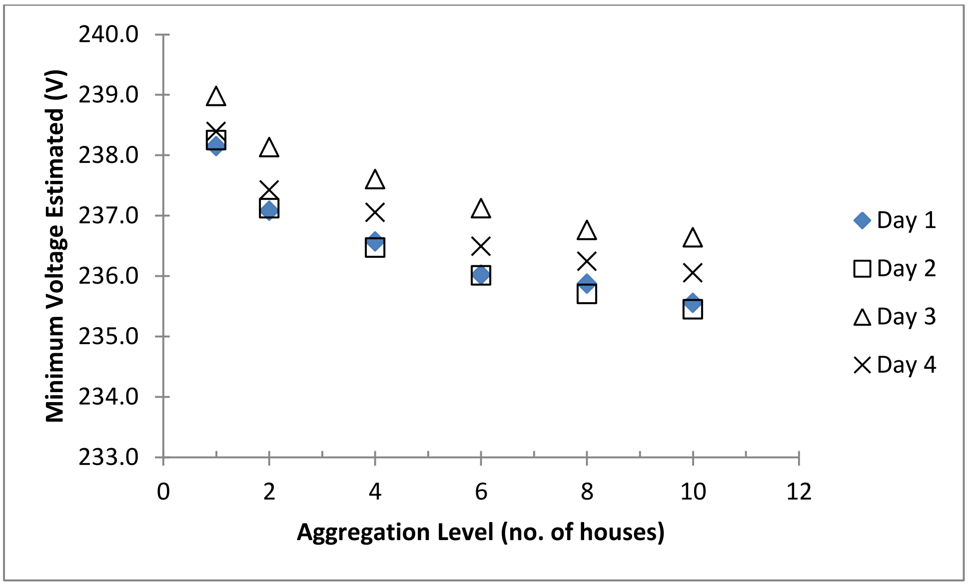


Figure H-4: Estimated minimum voltage levels on the blue phase at various aggregation levels-unbalanced network (CLNR data set no.8)

Alternative Modelling

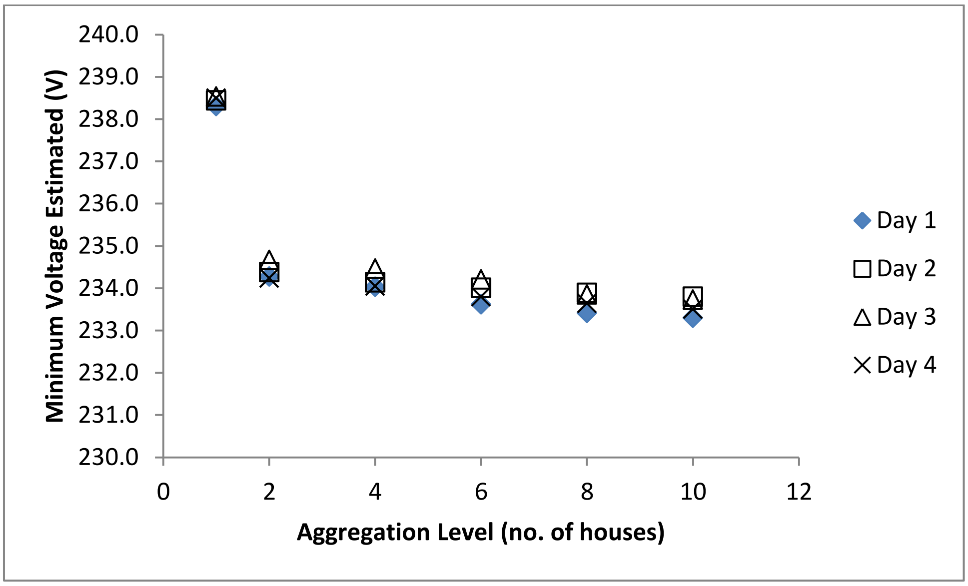


Figure H-5: Estimated minimum voltage levels on the yellow phase at various aggregation levels-unbalanced network-alternative topology (Loughborough data set)

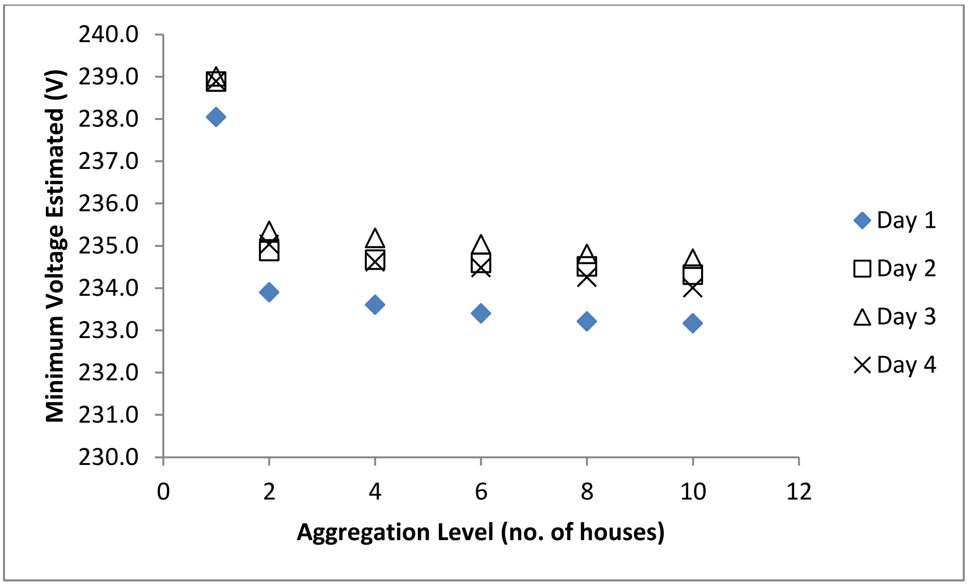


Figure H-6: Estimated minimum voltage levels on the yellow phase at various aggregation levels-unbalanced network-alternative topology (CLNR data set no.8)

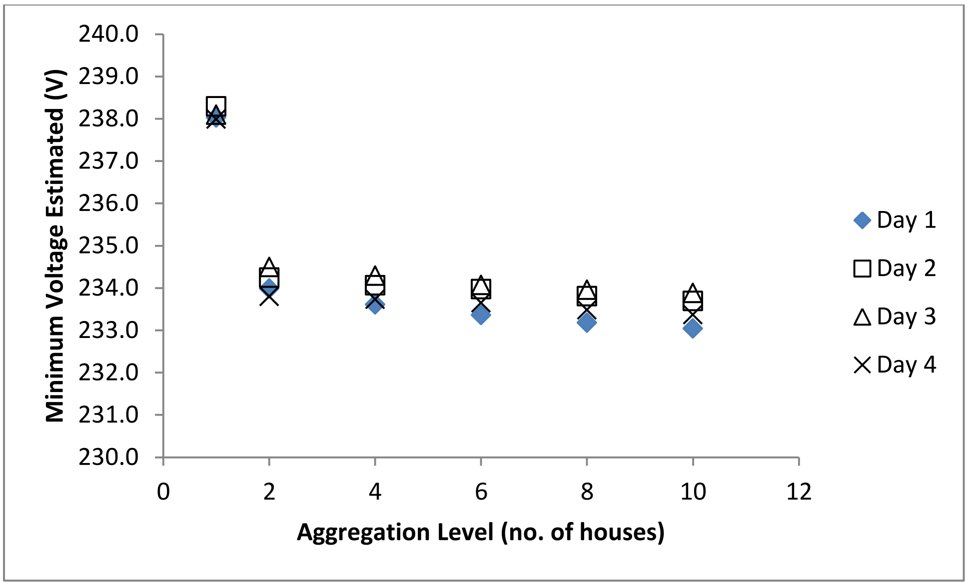


Figure H-7: Estimated minimum voltage levels on the blue phase at various aggregation levels-unbalanced network-alternative topology (Loughborough data set)

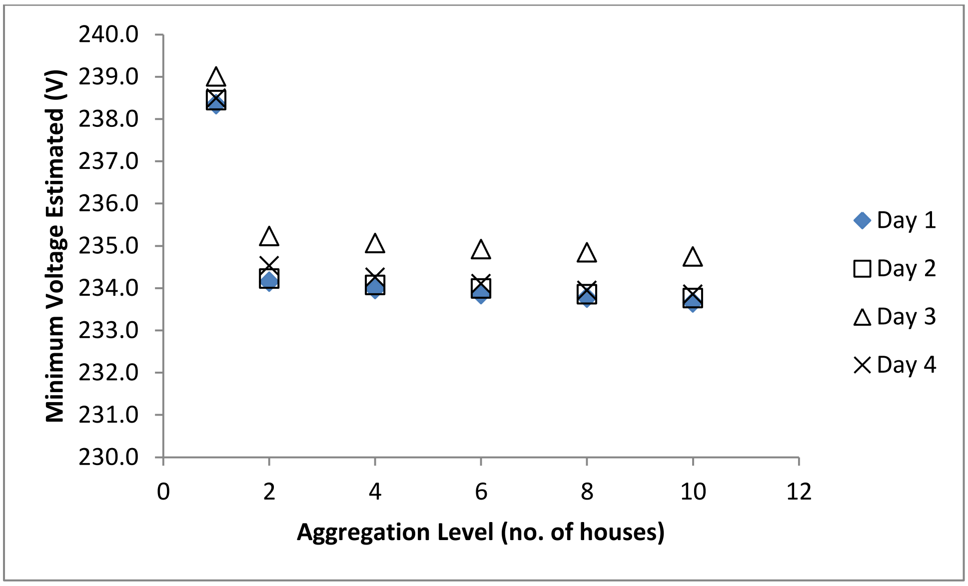


Figure H-8: Estimated minimum voltage levels on the blue phase at various aggregation levels-unbalanced network-alternative topology (CLNR data set no.8)