Investigation of Phase Imbalance Characteristics and Phase Balancing in Low Voltage Distribution Networks
(Alternate format thesis)

Kong, Wangwei

Award date:
2021

Awarding institution:
University of Bath

Investigation of Phase Imbalance Characteristics and Phase Balancing in Low Voltage Distribution Networks

By

Wangwei Kong

BEng, MSc

The thesis submitted for the degree of

Doctor of Philosophy

The Department of Electronic and Electrical Engineering

University of Bath

March 2021

Signature:…… Date:……..

Abstract

Phase imbalance, also known as phase unbalance, includes phase voltage imbalance and phase current imbalance. Take voltage imbalance as an example; it refers to the fact that either the voltage magnitudes are not the same or their phase angles are not 120° apart from each other, or both. A similar definition applies to current imbalance. Phase voltage and phase current imbalances will lead to phase power imbalance, which means that the three phases' power flows are not equal to each other.

Phase imbalance is a widespread and severe problem in distribution networks, especially in low voltage (415V, LV) distribution networks. It causes energy losses and capacity wastes that lead to high costs. Major causes for this problem are uneven load allocations across the three phases and random load behaviours.

Analysing phase imbalances is difficult as only LV substations are equipped with monitoring devices in existing distribution networks in the UK and the monitored data are only collected once a year. Low carbon technologies (LCTs), active customers and new business models in the electrical distribution system add to the complexity; the increased elements and interactions introduce uncertainties to load behaviour, which affects phase imbalances. Understanding phase imbalance in the distribution system helps the distribution network operators (DNOs) to understand phase balancing business cases and design suitable phase balancing solutions.

This thesis completed the following tasks and delivered contributions:

1) **Developed a new method to decompose the annual three-phase power series into a directional phase imbalance and a non-directional phase imbalance, thus revealing the nature of phase power imbalance.** A phase imbalance direction indicates the phase that is heavier or lighter loaded on average compared to the other two phases. A directional phase imbalance can be addressed by phase swapping, which is a relatively cheap solution. A non-directional phase imbalance can only be addressed by online phase balancing, e.g., demand-side management, which is relatively expensive.
2) **Developed a new data-driven cost-benefit analysis framework of phase balancing solutions for data-scarce LV networks. The framework uses a customised cluster-wise Gaussian process regression (CGPR).** The framework serves as an effective tool to assist DNOs to evaluate the cost-benefit of phase balancing solutions for data-scarce networks with no need to invest in additional monitoring devices.

3) **Explored the impacts of two different low carbon technologies (LCTs) on phase imbalances through a new Monte Carlo simulation framework.** The LCTs considered are single-phase connected electric vehicles and household solar generation. The developed framework helps the DNOs understand the possible imbalance-induced cost for different LCT penetration levels in the LV distribution network.

.

Table of Contents

Acknowledgements

First, I would like to express my deepest gratitude to my supervisors: Dr Kang Ma and Prof Furong Li. They had guided me throughout the past years of my research with patience and valuable research suggestions.

I would like to also extend my thanks to Dr Francis Robinson for his encouragements and kindness.

I would like to say thank you to Dr Qiuyang Ma, Dr Da Huo, Dr Chen Zhao, Dr Minghao Xu, Dr Wei Wei, Dr Heng Shi, Dr Xiaohe Yan, Dr Yuankai Bian and Dr Hantao Wang for guiding me through at the beginning of my research life.

My appreciations also extends to my collegues. Miss Chi Zhang, Mr Xinhe Yang, Miss Lanqing Shan, Dr Zhong Zhang, Dr Ran Li and Dr Chenghong Gu for supporting me in the research projects.

I would also like to express my thanks to Mr Renjie Wei, Mr Lurui Fang, Miss Wenqi Ou and Mr Haiwen Qin for supporting me during collaborating research papers.

Special thanks to all of my friends, Dr Chenning Wu, Miss Yajun Zhang, Miss Catherine Lin, Miss Shuang Cheng, Mr Junlong Li, Mr Jiahang Li, Mr Mike Ndawula, Mr Pengfei Zhao,Miss Jianwei Li, Miss Tian Gan, Miss Yunting Liu, Dr Wenjuan Song, Dr Zhipeng Zhang and Dr Heather Wyman-Pain for supporting and encouraging me along the way.

Last but not least, I would like to express my gratitude to my parents, Ms Xiumin Wang and Mr Dejun Kong; and my partner, Mr Tham Joon Chin, for their love, support and understanding during my PhD career.

List of Figures

List of Tables

List of Abbreviations

AC	Alternative Current
AELC	Additional Energy Loss Cost
ANM	Active Network Management
ARC	Additional Reinforcement Cost
CGPR	Cluster-Wise Gaussian Process Regression
DER	Distributed Energy Resource
DC	Direct Current
DG	Distributed Generation
DNO	Distribution Network Operator
DPIB	Degree of Power Imbalance
DER	Distributed Energy Resource
DSO	Distribution System Operator
D-STATCOM	Distribution Static Compensator
EHP	Electric Heat Pump
ES	Energy Storage
EV	Electric Vehicle
ICT	Information and Communication Technology
LCT	Low Carbon Technology
LV	Low Voltage
MAPE	Mean Absolute Percentage Error
MV	Medium Voltage
Ofgem	Office of Gas and Electricity Markets
PV	Photovoltaic
RIB	Random Imbalance
RMSE	Root-Mean-Square Error
SIB	Systematic Imbalance
TIC	Total Imbalance-Induced Cost
TN-S	Terre-Neutral Separated

TN-C-S	Terre-Neutral-Combined-Separated
VUF	Voltage Unbalance Factor

List of Variables

B^{ds}	Net benefit of applying the phase balancing solution
$B_{PV_N}^{ds}$	Total benefit of phase balancing for the data-scarce networks
σ^2	Noise variance
C_{asset}	Asset capacity
d	Discount rate
DIB_f, DIB_t	Degrees of phase imbalance for main feeders and LV transformers
E_{loss}	Energy loss caused by the neutral line current
E_{trans}, E_i	Energy loss from transformer copper loss under balanced and imbalanced scenarios
E_{t_i}	Imbalance-induced transformer copper loss
E_{total}	Total energy consumption
f_{pb}	Cost of applying a phase balancing solution
f_{PV_N}	Present value of the total imbalance-induced cost
I	Balanced phase current
$\hat{I}$	Annual peak current
I_A, I_B, I_C	Current values for the phases A, B and C
$\bar{I}_a$, $\bar{I}_b$, $\bar{I}_c$	Yearly average phase current values for phases A, B, and C
I_{nlc}	Neutral line current
$\bar{I}_{nlc}$	Virtual neutral line current
$I_\emptyset$	Magnitudes of current on phase $\emptyset$, $\emptyset \in \{A, B, C\}$
K	Kernel matrix
Loss_c , Loss_{c_i}	Transformer copper loss under three-phase balanced scenario and imbalanced scenario
Loss_{nlc}	Energy loss caused by the neutral line current
NRC_B^{PV}, NRC_{IB}^{PV}	Present value for network reinforcement cost under three-phase balanced scenario and imbalanced scenario
$P_\emptyset$	The power on phase $\emptyset$, $\emptyset \in \{A, B, C\}$
$\bar{P}_\emptyset$	The average power of phase $\emptyset$, $\emptyset \in \{A, B, C\}$

P_N	Neutral line power
P_t	Total power of three phases
r	Load growth rate
R_n	Neutral wire resistance
R_w	Transformer winding resistance
t	Time point
$x_r(t), y_r(t), z_r(t)$	The random imbalance component of the three-phase power (phases x, y, and z) at time point t
$x_s(t), y_s(t), z_s(t)$	The systematic imbalance component of three-phase power (phases x, y, and z) at time point t
U_N	Utilization rate
v_f	Feature vector
$V_\emptyset$	Magnitudes of voltage on phase $\emptyset$, $\emptyset \in \{A, B, C\}$
V_1, V_2 , V_0	Positive, negative and zero sequence voltage

Publication List

Journal publications:

1. W. Kong, K. Ma and Q. Wu, "Three-Phase Power Imbalance Decomposition into Systematic Imbalance and Random Imbalance," in *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3001-3012, May 2018, DOI: 10.1109/TPWRS.2017.2751967.
2. W. Kong, K. Ma, L. Fang, R. Wei and F. Li, "Cost-Benefit Analysis of Phase Balancing Solution for Data-Scarce LV Networks by Cluster-Wise Gaussian Process Regression," in *IEEE Transactions on Power Systems*, vol. 35, no. 4, pp. 3170-3180, July 2020, DOI: 10.1109/TPWRS.2020.2966601.
3. K. Ma, L. Fang and W. Kong, "Review of distribution network phase unbalance: Scale, causes, consequences, solutions, and future research direction," in *CSEE Journal of Power and Energy Systems*, DOI: 10.17775/CSEEJPES.2019.03280, 2020.
4. L. Shan, W. Kong, C. Gu and F. Li, ‘‘Roles and Functions of Distribution System Operators in Local Electricity Market Development’’, *Journal of Global Energy Interconnection*, 2020, 3(1):70-78. (in Chinese)

Submitted journal paper:

1. W. Kong, K. Ma and F. Li, " Probabilistic Impact Assessment of Phase Power Imbalance in the LV networks with Increasing Penetrations of Low Carbon Technologies," submitted to *Electric Power Systems Research*, September 2020.

Conference publications:

1. W. Kong, K. Ma, F. Li, E. Thompson and L. Sidebotham, "Impact Assessment Criteria of Distribution System Architecture", *The 25th International Conference and Exhibition on Electricity Distribution (CIRED)*, Madrid, 2019.
2. W. Kong, K. Ma, F. Li and L. Sidebotham, "Future Distribution System Architecture in the UK," *2018 International Conference on Power System Technology (POWERCON)*, Guangzhou, 2018.
3. W. Kong, K. Ma, F. Li and L. Sidebotham, "System Architecture for Customer-Led Distribution System," *2018 15th International Conference on the European Energy Market (EEM)*, Lodz, 2018, pp. 1-5.
4. F. Li, and W. Kong, “Visions, functions and implementation of DSOs – a UK perspective,” in *CIRED Workshop*, Ljubljana, 2018.

Chapter 1. Introduction

This chapter describes the research background, motivation, challenges and contributions of this work. It also provides an overview of the thesis.

1.1. Research Background

1.1.1. Phase Imbalance and Phase Balancing

A balanced three-phase distribution network consists of phase voltages of the same magnitude that are 120° apart from each other; the same applies to currents. Phase imbalance, also known as phase unbalance, includes both voltage imbalance and current imbalance. Take voltage imbalance as an example; it refers to the fact that either the voltage magnitudes are not the same or their phase angles are not 120° apart from each other, or both. A similar definition applies to current imbalance. Voltage and current imbalances will lead to phase power imbalance which means that the power flows on the three phases are not equal to each other. In this thesis, only the magnitudes of phase current and phase voltage are considered to calculate the phase power imbalances.

Phase current imbalance (referred to as current imbalance) causes insufficient use of network capacity. This is because if the network capacities of an LV feeder is used up, the spare network capacities from other phases cannot be transferred to this feeder. An imbalanced network will require early investment as the capacity of the heaviest phase would be used up sooner than if the network was phase balanced [1]. Therefore, current imbalance leads to additional investment cost as a consequence of capacity waste. Besides, current imbalance results in additional energy loss on both feeders [2, 3] and transformers [4]. Additional energy loss is the difference between energy loss with balanced phases and energy loss with imbalanced phases.

Voltage imbalance is defined as the ratio of the negative sequence voltage and the positive sequence voltage [5]. 1% of negative-sequence voltage imbalance results in about 6% of negative-current current imbalance, which increases power losses and causes motor overheating [6, 7]. Therefore, voltage imbalance decreases power quality [8] and causes motor damages [9]. The IEC recommends a limit of 2% for voltage imbalance in LV supply systems [66].

Thus, phase balancing is beneficial for electrical distribution networks by mitigating the consequences mentioned above. In the LV network, phase balancing solutions include but are not limited to phase swapping, demand-side management and deploying phase balancers [10]. Phase swapping moves load from the heavy loaded phase(s) to the light loaded phase(s) to rebalance the three phases [11]. Demand-side management encourages end-users to change usage patterns. The end-users are incentivised by time-of-use electricity price determined to contribute to phase balancing. Phase balancers along with control strategies rebalance the three phases in real-time operation, with an optional function for reactive compensation [10].

In the UK, the monitoring equipment installed at substations mostly depends on the nominal voltage, with the lower voltages having significantly fewer monitoring devices [12]. A typical LV distribution substation monitoring device is a Maximum Demand Indicator (MDI) which records peak phase currents based on the aggregation over half an hour. The MDI does not have any communication options so that its reading is typically manually recorded on an annual basis [43].

The absence of frequent time-series data pose challenges to distribution network operators (DNOs) when estimating load behaviours. It becomes more challenging with the constantly evolving electrical distribution system with low carbon technologies (LCTs), such as electric vehicles (EVs) and Photo-Voltaic (PV) generation.

1.1.2. Changes in the UK's Distribution System

For many decades, electricity has been generated from centrally dispatched power plants, and transmitted through networks to match the demand, as shown in Figure 1-1. However, with the introduction of LCTs, active customers (also known as prosumers) and new business models (such as virtual power plant business model), the electrical distribution system is becoming increasingly complicated.

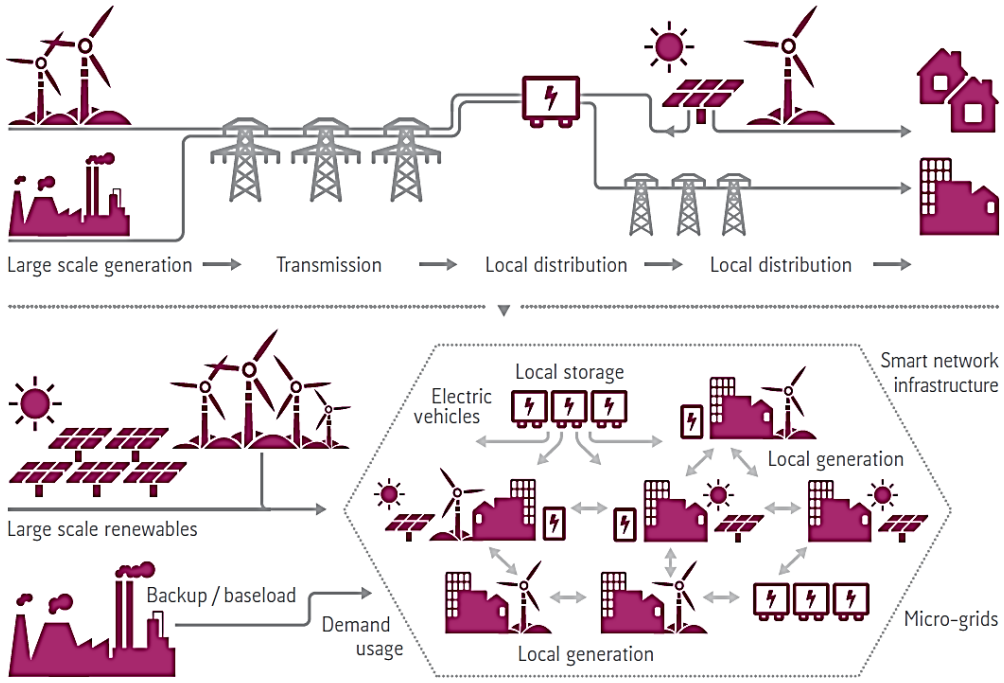


Figure 1-1. Evolution of the UK's power system taken from [13]

The drivers for change are grouped into five aspects by [14]:

1) **Sustainability and decarbonisation** [15, 16]: According to the British government's target on decarbonization, pathways to increase distributed renewable generation, storage, electric heat and electric transport is encouraged. As indicated by the National Grid's ''Two Degrees'' scenarios, the carbon emission should decrease 90% by 2050 compared to 73 megatonnes in 2017 [17].
2) **Cost-effectiveness** [18]: The transition of power system faces growing penetration of LCTs that requires network reinforcement. The investment on network reinforcement would reduce system operation cost and therefore enables a cost-effective integration of LCTs [19]. The cost of transition to a low carbon future needs to be affordable for the businesses. System operator also need smart technologies to actively control and manage the system to secure economic benefits for customers.
3) **Security of supply**: The traditional central generation is replacing by various distributed renewable generations. According to the Future Energy Scenario report from the National Grid [17], the installed capacity of low carbon and renewables reaches 56.6% of the total electricity generation capacity by 2030 with steady progression. At large scales, the mix of generation has low stabilising inertia which could reduce the inherent stability and security of the electrical system. Therefore, the future electrical system needs to find the balance between carbon emission goal and the security of supply [20].
4) **Consumers changing to prosumers**: An increasing amount of household distributed generation (DG) systems are connected to the power grid. This converts the passive consumer to active consumers, also known as prosumers [21-24], capable of injecting power to the main grid. The prosumers' need is changing with time and economic incentives, such as demand-side management [25].
5) **New business models emerging in the power market**: New roles such as aggregator are responsible for organizing prosumers in the future [25]. The new prosumers will also facilitate the development of local markets [26], i.e. peer-to-peer market.

The distribution system is evolving in a decentralised way and becoming increasingly complicated with the massive growth of LCTs and new business models. The customer's electricity usage patterns will be more difficult to be predicted because the combinations of LCTs are able to generate as well as consume electricity. Consequently, the changes in the distribution system will have impacts on network problems, such as phase imbalance. Investigating the possible impacts of the

changing distribution system on phase imbalance will help DNOs manage and operate the distribution networks efficiently and effectively.

1.1.3. Phase Imbalance in the UK's Distribution System

In the United Kingdom (UK), phase imbalance occurs in more than 70% of the UK's low voltage (415V, LV) networks [27]. Data from 800 LV networks show that the majority suffers significant phase power imbalance where the difference between the 'heaviest' loaded phase and the 'lightest' loaded phase is greater than 50% [7]. The TNEI also found that 70.8% of 233 LV feeders suffers severe phase imbalance where the ratio of the phase current to the mean current of the three phases is larger than 1.3 on average [28].

The low carbon transition of the distribution system has uncertain impacts on phase imbalance. The phase imbalance can be aggregated if the connections of LCTs are not properly controlled. However, new controlling and managing strategies can be developed to reduce phase imbalances as well as other network problems.

1.2. Research Motivations and Challenges

1.2.1. Understanding the Characteristics of Phase Imbalances

Two main reasons for phase imbalances are the uneven load allocation and random load behaviour [8, 29, 30]. Uneven load allocation causes systematic imbalance (SIB). SIB means that there is a definite maximum phase (which has the highest power among the three phases), a definite minimum phase (which has the lowest power among the three phases), or ordered three phases (where both definite maximum and minimum phases exit). SIB can be effectively addressed by phase swapping, which has been widely adopted by DNOs. Random load behaviours cause random imbalance (RIB). RIB has neither a definite maximum phase nor a definite minimum phase. RIB requires demand-side management to address and risks non-delivery. The former is relatively cheaper compared to the latter.

Current research mainly focuses on investigating methods of reducing phase voltage, phase current or phase power imbalances [31-38]. These methods lack the decomposition of phase power imbalances into SIB and RIB, which reveals the maximum potential for phase balancing and a

minimum requirement for demand-side management. By decomposing power imbalances, the phase balancing costs can be optimised.

1.2.2. Performing Cost-Benefit Analysis of Phase Balancing

In the LV network, phase swapping [39], demand-side management [40] and deploying phase balancers [10] are the conventional solutions to phase balancing. The phase balancing solutions are able to improve power quality and reduce energy losses with different costs. It is necessary to perform a cost-benefit analysis before applying a solution to the LV networks. Current cost-benefit analysis of phase balancing solutions requires a complete time-series of voltage and current data [2, 3, 41-46]. However, the majority of the UK's LV networks are unmonitored with some exceptions; only substation data are collected once a year. Therefore, it is challenging for DNOs to evaluate the phase balancing solutions before making an investment decision.

As a result, there is a need for a cost-benefit analysis method for phase balancing solutions in existing LV networks that have insufficient data.

1.2.3. Uncovering the Future of Phase Imbalances

To ensure the effective and secure operation of the distribution system, the distribution system is required to coordinate the increasing LCTs with active customers and business models while maximising customers' benefits. Existing research focus on analysing the impacts of LCT penetration, such as EV charging [47-52], PV inverters [35, 52-57] and heat pumps (HPs) [49, 58], on voltage imbalance; and developing strategies for minimising voltage imbalance in the distribution networks [48, 50, 51].

Consequently, there is a need for understanding the possible impacts of LCT penetrations on phase imbalances in the UK's LV distribution networks.

1.3. Research Contributions

This thesis aims to find efficient and cost-effective methodologies to help DNOs analyse the phase imbalances in the distribution system and find optimal phase balancing strategies to minimise the imbalances in the LV networks. The contributions of this work are:

1) **Developed a new method to decompose the annual three-phase power series into a directional phase imbalance and a non-directional phase imbalance, thus revealing the nature of phase imbalance.** A phase imbalance direction indicates the phase that is heavier or lighter on average compared to the other two phases. A directional phase imbalance can be addressed by phase swapping, which is a relatively cheap solution. A non-directional phase imbalance can only be addressed by online phase balancing, e.g., demand-side management, which is relatively expensive.
2) **Developed a new data-driven cost-benefit analysis framework of phase balancing solutions for data-scarce LV networks.** The framework uses a customised cluster-wise Gaussian process regression (CGPR). The framework serves as an effective tool to assist DNOs to evaluate the cost-benefit of phase balancing solutions for data-scarce networks without requiring the investment in additional monitoring devices.
3) **Developed a Monte Carlo simulation analysis to investigate the impacts of LCT penetrations on phase power imbalances and determine the optimal option to balance day-to-day energy loss cost and long-run investment.** The impacts of LCT penetrations on phase imbalances are analysed using Monte Carlo simulation considering LCT uncertainties. Two single-phase connected LCTs are considered for the analysis, i.e. electric vehicles and household solar generation. The developed probabilistic impact assessment framework helps the DNOs understand the potential imbalance-induced costs under different LCT penetration levels.

1.4. Thesis Layout

Chapter 2 presents a comprehensive review of phase imbalance analyses in the distribution system and identifies the research gaps.

Chapter 3 presents the published paper 'Three-phase power imbalance decomposition into systematic imbalance and random imbalance' and additional analysis and discussions.

Chapter 4 presents the published paper 'Cost-benefit analysis of phase balancing solution for data-scarce LV networks by cluster-wise Gaussian process regression' and additional analysis and discussions.

Chapter 5 presents the submitted paper 'probabilistic impact assessment of phase power imbalance in the LV networks with increasing penetrations of low carbon technologies' and additional analysis and discussions.

Chapter 6 concludes the key findings and contributions of the work.

Chapter 7 shows some potential topics for future research.

Chapter 2. Phase Imbalance and Phase Balancing

This chapter presents a comprehensive literature review of existing research on phase balancing. The gaps and limitations of existing research are explained and discussed.

2.1. Introduction

This chapter introduces the review of the definitions, causes and consequences of phase imbalance. Besides, phase balancing solutions are explained in detail and their advantages and limitations are compared. Moreover, published works on the future developments of the system architecture are also reviewed to discover the phase imbalance in the future distribution system.

2.2. LV Distribution Network

Figure 2-1 shows that the general European distribution system, which is a three-phase system. The three-phase power system has an economic advantage in transmitting power compared to a single-phase power system because of lower energy loss and less cost [59].

Two real LV networks from Western Power Distribution's licenced area are given as examples for illustration. Figure 2-1 and Figure 2-2 shows the LV networks of Illminster Avenue and Marwoord Road, respectively. The transformer capacity for Illminster Avenue is 750kVA and its utilization rate is 43%. The transformer capacity for Marwoord Road is 500kVA and its utilization rate is 94.8%.

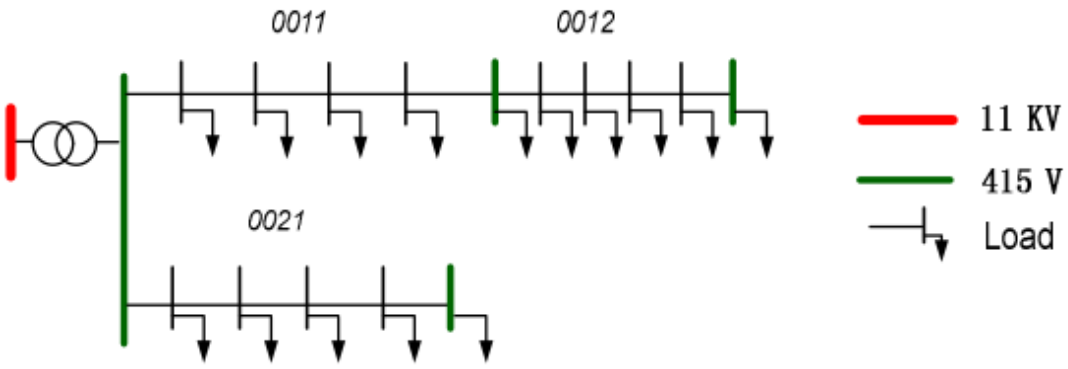


Figure 2-1. Layout of Illminster Avenue taken from [60]

The detailed parameters of the two networks are given in Table 2-1. The Illminister Avenue has a total of 267 customers while the Marwoord Road has a total of 377 customers. Note that the phase allocations of the customers are unknown for Western Power Distribution's networks [61].

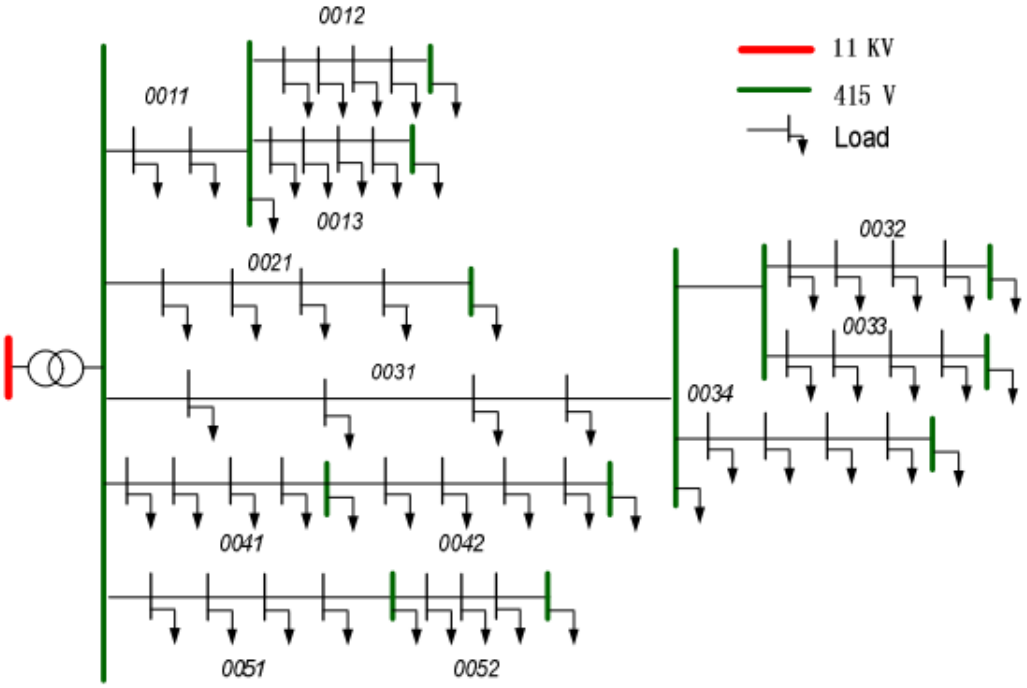


Figure 2-2. Layout of Marwood Road taken from [60]

Table 2-1. Parameters of the two networks taken from [60]

Network name	Feeder name	Number of customers per feeder	Feeder length (m)
Illminster Avenue	Feeder 0011	118	295
	Feeder 0012	28	88
	Feeder 0021	121	287
Marwoord Road	Feeder 0011	3	172
	Feeder 0012	13	109
	Feeder 0013	12	110
	Feeder 0021	67	268
	Feeder 0031	6	170
	Feeder 0032	29	301
	Feeder 0033	42	191
	Feeder 0034	10	101
	Feeder 0041	14	120
	Feeder 0042	52	233
	Feeder 0051	125	345
	Feeder 0052	4	0

2.3. Definition, Causes and Consequences of Phase Imbalance

2.3.1. Definition of Phase Imbalance

Phase imbalance means that, among the three phases, either the voltage (or current) magnitudes are not the same or their phase angles are not 120° apart from each other. Phase imbalance includes voltage and current imbalance. Voltage (or current) imbalance is defined as the ratio of the positive sequence voltage (or current) and the negative sequence voltage (or current) by IEEE [5, 62]. The IEEE true definition of voltage imbalance factor (VUF) is given by [5]:

$$VUF = \frac{V_2}{V_1} \times 100 \quad (2\text{-}1)$$

where V_1 and V_2 represents the positive and negative sequence voltage, respectively.

However, this definition of voltage imbalance ignores the zero sequence component of voltage which is inevitable in the distribution system. A complementary formulation of voltage imbalance ratio ($RMSV$) is given by [63]:

$$RMSV = \frac{\sqrt{|V_0|^2 + |V_2|^2}}{|V_1|} \quad (2\text{-}2)$$

where V_1, V_2 and V_0 represents the positive, negative and zero sequence voltage, respectively.

For example, suppose the three-phase voltages are $v_A = 244\angle 0°$, $v_B = 243\angle 245°$ and $v_C = 242\angle 120°$. Then the magnitudes of symmetrical components are $V_1 = 242.8$, $V_2 = 6.5$ and $V_0 = 7.6$. Therefore, the IEEE true definition gives a value of $\mathrm{VUF} = 2.7\%$ while the complementary formulation gives a value of $RMSV = 4.1\%$. As can be seen from the results, the difference between these two definitions is large. As a result, the zero sequence voltage can not be neglected while analysing phase voltage and current imbalances.

As indicated by Engineering Recommendation P29 [64], voltage imbalance for systems with a nominal voltage below 33kV should not exceed 1.3%; and for systems with a nominal voltage below 132kV, the voltage imbalance should not exceed 1%. The IEC recommends a limit of 2% for voltage imbalance in LV supply systems [65].

For the reason that only the magnitudes of phase voltage are available in the data from 800 LV networks, the phase voltage unbalance rate (PVUR) is adopted to analyse phase voltage imbalance. The PVUR is given by [2]:

$$PVUR = \frac{\text{maximum voltage deviation from the average phase voltage}}{\text{average phase voltage}} \quad (2\text{-}3)$$

The results of the 800 LV networks show that only 11 networks have exceeded the limit of 2% voltage imbalance for only one time point (10-minute resolution) for the whole year. None of the LV networks has a PVUR larger than 3%. Therefore, in this thesis, voltage imbalance of all the networks are assumed to be within the limit.

Phase power ($P_{\emptyset}$) for each phase is defined as:

$$P_{\emptyset} = V_{\emptyset} \times I_{\emptyset} \times cos\varphi_{\emptyset}, \qquad \emptyset \in \{A, B, C\} \tag{2-4}$$

where $V_{\emptyset}$ and $I_{\emptyset}$ represents the magnitudes of phase voltage and phase current, respectively; $cos\varphi_{\emptyset}$ is the power factor.

Voltage and current imbalances will lead to phase power imbalance [8, 66] which means that the power flows on the three phases are not equal to each other.

It should be noted that major difficulty in quantifying the time-varying power factor ($cos\varphi_{\emptyset}$) in the LV distribution network is the lack of phasor measurements. There is hardly any information on power factor in real-time operation. A feasible solution is to assume an average power factor of 0.9 for the three phases. Assuming such an average power factor, if the active power is rebalanced, the reactive power is automatically rebalanced. Hence, only active power is considered for phase balancing in this thesis.

Phase power imbalance is analysed using the degree of power imbalance (DPIB) factor. The DPIB for a main feeder is given by [1]:

$$DPIB_f = \frac{max\{P_{\emptyset}\} - P_t}{P_t} \qquad \emptyset \in \{A, B, C\} \tag{2-5}$$

where P_t is the total power of three phases when the maximum phase power occurs and $P_{\emptyset}$ is the power on phase $\emptyset$. The DPIB for a transformer is given by [1]:

$$DPIB_t = \frac{P_N}{P_t} \qquad \emptyset \in \{A, B, C\} \tag{2-6}$$

where P_N is neutral line power and P_t is the total power of three phases when the maximum phase power occurs

2.3.2. Causes of Phase Imbalance

There are two main causes for phase imbalance in the LV distribution network:

1) Uneven single-phase load allocation among the three phases

The majority of the load in the UK's LV distribution network is single-phase connected [67] and their distribution across the three phases are not even [8, 29, 30, 40, 68]. This will naturally result in uneven demand across the three phases. In other words, it will result in phase imbalances in the LV networks. Data from 800 LV networks show that the majority suffers significant phase power imbalance where the difference between the 'heaviest' loaded phase and the 'lightest' loaded phase is greater than 50% [7]. The TNEI also found that the majority LV feeders suffers severe phase imbalance where the ratio of the phase current to the mean current of the three phases is larger than 1.3 on average [28].

2) Random load behaviours in the distribution system

There are three main types of loads in the LV distribution network: industrial, commercial and residential load. Industrial and commercial load are mostly predictable. However, the residential load has large uncertainties, such as electricity usage patterns and the adoption of low carbon technologies (LCTs). Such random load behaviours also lead to phase imbalances [29, 30].

3) Structural asymmetries

The distribution line segments are inherently asymmetrical [69] because of the connections of single-phase laterals to the three-phase main in the LV distribution network [67]. The structural asymmetries lead to phase imbalances in the distribution networks [58].

2.3.3. Consequences of Phase Imbalance

Phase imbalance causes several consequences, including additional reinforcement cost (ARC) [1, 70], additional energy losses [71] and damages to induction motors [9].

1) Incurred ARC from LV transformer and main feeder

The ARC arises from both LV transformers and main feeder. As for the main feeder, ARC is the consequence of insufficient usage of network asset from phase imbalances [1, 70]. This is because the spare capacities of one phase cannot be transferred to another phase which uses up its capacity. In other words, the usable capacity for the main feeder is decided by the least spare capacity. As for LV transformers, phase imbalance results in the additional power flow in the neutral line and reduces the available capacity [1]. In the UK, the LV side of a three-phase

distribution transformer is star connected [72]. Therefore, only star connection for LV transformers is considered in this thesis.

Figure 2-3 shows that phase imbalance results in a higher utilization rate compared to that of the balanced scenario for both the LV transformer and the main feeder. As a result, phase imbalance leads to an earlier network investment compared to that with balanced three phases.

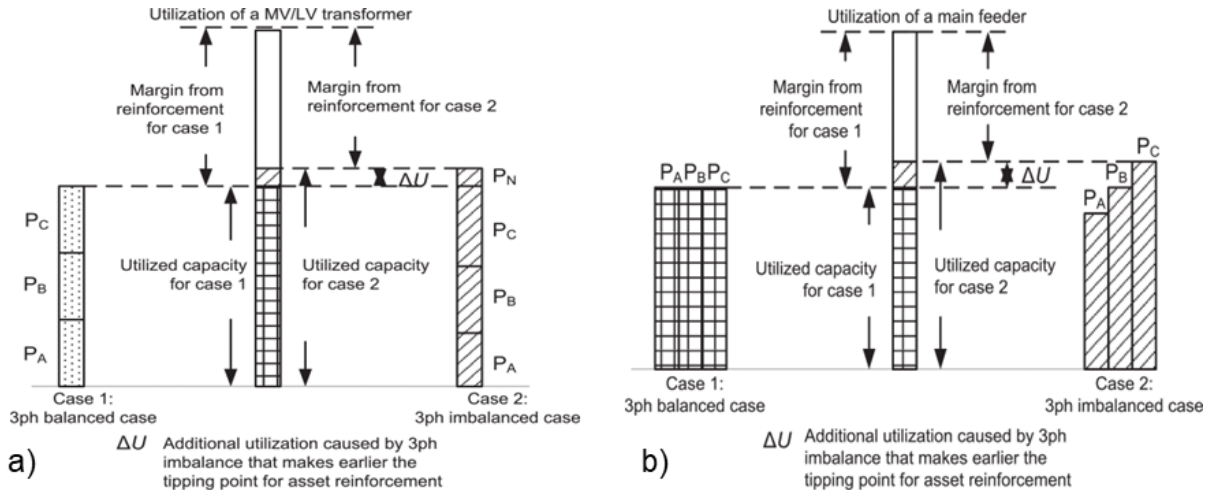


Figure 2-3. The difference between the balanced and imbalanced case in the utilization of a) an MV/LV transformer and b) a main feeder taken from [1]

The calculation of ARC is given by [1]:

$$ARC = NRC_B^{PV} - NRC_{IB}^{PV} \quad (2\text{-}7)$$

where NRC_B^{PV} is the present value for network reinforcement cost under a three-phase balanced scenario and NRC_{IB}^{PV} is the present value for network reinforcement cost under three-phase imbalanced scenario.

Figure 2-4 shows an example of comparing the ARC and NRC_B^{PV} for different groups of LV networks with a fixed utilization rate (90%). With the increase of the degree of phase imbalance, the ARC continue increases to exceed the network reinforcement cost under a three-phase balanced scenario (NRC_B^{PV}).

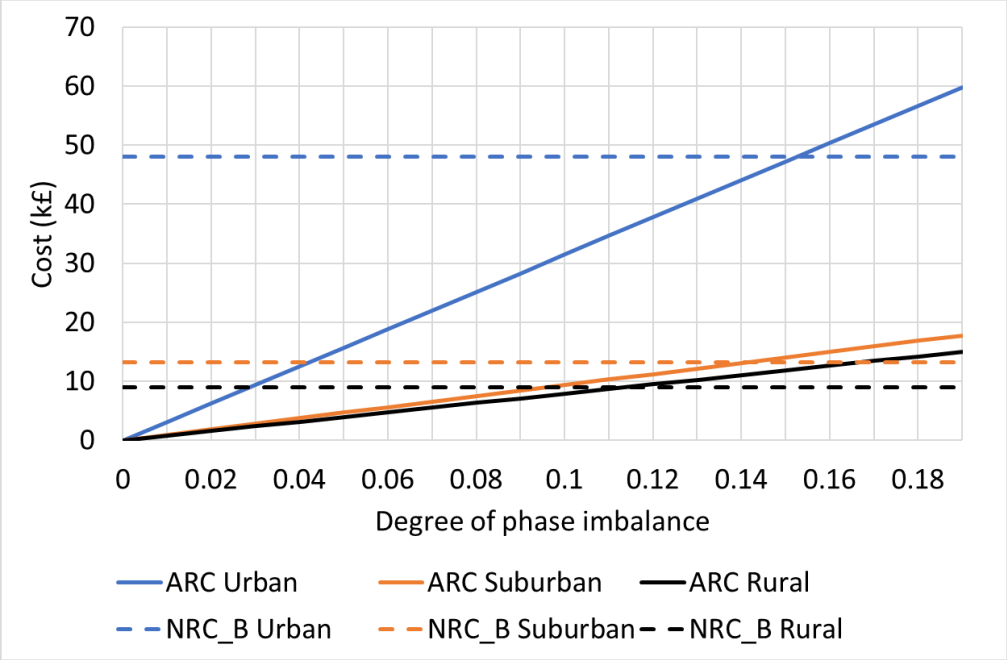


Figure 2-4. Example of ARCs for urban, suburban and rural networks

2) Increased energy losses

It should be noted that energy losses still exist on the phase conductors when three phases are balanced. The energy losses will become larger if the phases are imbalanced. The increased energy loss because of phase imbalance is the additional energy loss. Additional energy loss caused by phase imbalance includes two components: additional transformer copper loss [4, 73], [3, 74] and energy loss caused by neutral line current [75, 76].

a. Transformer copper loss is also known as winding loss [77]. The winding resistance is heated when current flow through the transformer windings; therefore, it generates copper losses. The transformer copper loss is given by [77]

$$\text{Loss}_c = 3I^2R_w \tag{2-8}$$

where I is the balanced phase current and R_w transformer winding resistance.

The transformer copper loss under the imbalanced case is also given in [77]

$$\text{Loss}_{c_i} = \left({I_A}^2 + {I_B}^2 + {I_C}^2\right)R_w \quad (2\text{-}9)$$

where I_A, I_B and I_C are current values for the phases A, B and C respectively.

Therefore, the imbalance-induced transformer copper loss is the difference between Loss_c and Loss_{c_i}.

For example, if the phase currents are$I_A = 100A$, $I_B = 170A$, $I_C = 110A$ and the winding resistance is $R_w = 0.0163\ ohm$. The transformer copper loss under such imbalance is $\text{Loss}_{c_i} = 831.3W$. Balanced phase current is $I_A = I_B = I_C = 126.7A$, the transformer copper loss is $\text{Loss}_c = 784.6\,W$. Thus, the imbalance- induced transformer copper loss is $\text{Loss}_{c_i} - \text{Loss}_c = 46.7W$.

b. In a four-wire-three-phase system, if the three phases are perfectly balanced, the current flowing through the neutral line is zero. However, if phase imbalance exists, there will be current flowing through the neutral line, causing energy losses.

Terre-Neutral Separated (TN-S) systems are adopted by the majority of the UK's LV distribution networks [78] [79]. TN-S was the default earthing system until Terre-Neutral-Combined-Separated (TN-C-S), also known as protective multiple earthing, became common in the 1980s [80]. The difference between these two earthing systems is that neutral and earth conductors are combined in supply-side but separated in customer side in the TN-C-S system; while the neutral and earth conductors are separated throughout in TN-S system [81].

Take TN-S earthing system as an example, neutral line current flows, through the neutral conductor, into the transformer neutral point. Thus, the energy loss caused by the neutral line current is given by [75]

$$\text{Loss}_{nlc} = I_{nlc}^2 R_n \quad (2\text{-}10)$$

where $I_{nlc} = \sqrt{I_A^2 + I_B^2 + I_C^2 - I_A I_B - I_B I_C - I_A I_C}$

where I_{nlc} is the neutral line current; R_n is the neutral wire resistance; and I_A, I_B and I_C are the phase currents for three phases, respectively.

The additional energy losses caused by phase imbalance can be translated into additional energy loss cost (AELC). Together with the ARC, they form the total imbalance-induced cost in the LV networks. In the UK, the phase imbalance in the LV distribution networks increases cost each year [27, 41, 44, 82, 83].

3) Damages to induction motors

Voltage imbalance leads to current imbalance in three-phase load and creates negative-sequence and zero-sequence current flow in the system [84]. Current imbalance results in higher energy losses on the neutral wire. Besides, it also leads to power derating, life-shortening and efficiency reducing of motors [9, 85, 86]. When three-phase induction motors run under unbalanced voltages, the current imbalance increases dramatically and will lead to the malfunction of the protection system [9].

2.4. Phase Balancing Solutions and Limitations

2.4.1. Phase Balancing Solutions

Phase balancing solutions include but are not limited to phase swapping, demand-side management and deploying phase balancers [7].

1) Phase swapping

Phase swapping (as shown in Figure 2-5) is a direct way of rebalancing phases by moving load between the three phases [11]. It includes nodal and lateral phase swapping [41] which means moving a node or lateral from heavily loaded phase(s) lightly loaded phase(s). In real life operations, phase swapping can be static (off-line) or dynamic (on-line) [87]. The dynamic operation has advantages in achieving global optimal strategies while minimising cost [88].

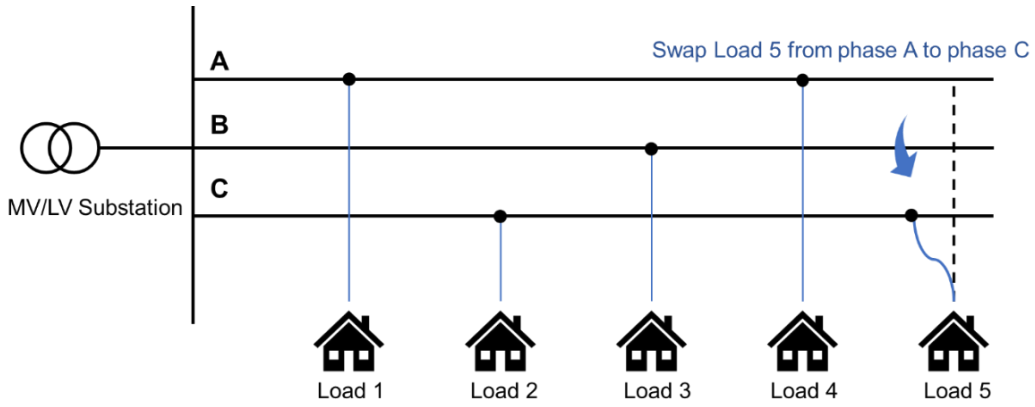


Figure 2-5. Phase swapping in LV distribution network

Various algorithms had been developed based on genetic algorithm (GA) [89, 90], particle swarm optimization (PSO) [84, 91] and shuffled frog leaping algorithm [92] to determine the optimal phase swapping strategies. Six algorithms for phase swapping are compared in [90] and the comparison shows that dynamic programming performs the best among all the algorithms. A statistical approach based on non-negative matrix factorization method is developed in [93] to guide phase swapping for data-scarce networks. Reference [94] decomposes distributed generation (DG) penetrated distribution system into local subsystems and used GA and heuristic algorithm to rebalance the phases for each subsystem. Rephasing and DG sizing are considered simultaneously in [95] to perform effective automatic phase balancing.

The drawbacks of phase swapping are the issues related to switching, i.e. cost and interruptions of supply [96]. It takes about 15 minutes to switch a load to another phase and the overall phase swapping takes about one hour [90]. The total cost, including both labour and preparatory work, may go up to thousands of British pounds [90].

2) Demand-side management

Demand-side management controls LCTs, such as electric vehicles (EVs), photovoltaic (PV) generation and energy storages (ES), by connecting or disconnecting them from the network to reduce phase imbalances in the distribution network. Demand-side management re-distributes the load and LCTs without curtailing load consumptions or wasting energy produced by intermittent renewables [97]. Many efforts were made to analyse the voltage imbalance caused by LCT penetrations, such as EV charging [47-52], PV inverters [35, 52-57] and heat pumps (HPs) [49, 58]. Developed solutions for minimising such imbalances include using control strategies for EV charging [48, 50, 51], using control strategies for on-load tap changer [52],

applying devices to maintain voltage levels [98] and applying generating limits for PV systems [53].

A few research have developed strategies for EV charging to achieve the goal of phase balancing [40, 50, 99]. The strategy developed in [40] focuses on controlling charging spots to detect the phase with the highest phase voltage and connecting EVs to that phase, therefore, reduce voltage imbalance among the three phases. A single-phase EV charging strategy is developed for smart chargers in [50] to mitigate voltage imbalance and increase EV hosting capacity. Reference [99] developed a remuneration scheme for EVs using game theory to minimise phase imbalances. Reference [47] compares the voltage imbalance caused by uncontrolled EV charging and tariff-based EV charging under different penetration levels. Similarly, a central control strategy is developed for EV charging points in [108], and it is applied to two real LV networks to analyse the impacts on voltage imbalance.

PVs are also widely used as a means of phase balancing [100-103]. Reference [100] proposed a local controller to control both EV chargers and PV inverters to reduce voltage imbalance and system losses. Reference [101] controls the reactive power of three-phase PV inverters to reduce the voltage imbalance caused by single-phase PV inverters. A voltage imbalance sensitivity analysis for PV penetrated LV distribution networks is performed in [102] and a method of controlling the reactive power supplied by PV converters is developed. In [103], PV inverters are applied to coordinate with thermostatically controlled loads for relieving voltage imbalance.

Control strategies have also been developed for ES to address the phase imbalance [104, 105]. Reference [104] proposes a real-time algorithm using ES to minimise cost and phase imbalance in the presence of uncertainty. Reference [105] uses controllable ES units to reduce voltage imbalance caused by highly penetrated PVs in LV distribution networks.

The impacts of four type of LCTs, including EVs, PVs and HPs, are analysed in [49] based on two impact factors. The two impact factors are customer side voltage violations and feeder loading levels. These impacts help DNOs estimate the LCT hosting capacities for LV distribution feeders. Reference [58] presented sensitivity analyses of both voltage and current problems on a sample suburban LV network to quantify the impacts of different types of electric HPs.

The drawback of using demand-side management is the high risk of non-delivery because of the uncertainties in LCTs. Besides, additional financial incentives are required to encourage customers to provide demand-side management to the DNOs.

3) Phase balancer

A range of existing works had developed phase balancers with various control strategies. Reference [106] developed a single-stage voltage balancer for bipolar LV DC distribution systems using a three-level dual-active-bridge converter. Closed-loop control is implemented to balance the bipolar voltage levels caused by imbalanced power in the bipolar DC bus. Similarly, a voltage balancer is developed for a bipolar DC microgrid in [107] using Buck/Boost converter with a current control strategy. A simple reactive power control strategy is developed in [10] to address the challenge of large power rating requirement for active load balancers. Besides converters, electric springs and distribution static compensator (D-STATCOM) are also used as phase balancers. A three-phase electric spring circuit, formed by electronic circuits, is deployed to re-allocate the power among non-critical loads to reduce phase imbalance [8]. References [108, 109] propose individual phase control strategies for D-STATCOM to minimise phase imbalance. Reference [110] analysed the voltage imbalance caused by single-phase PVs and used D-STATCOM to reduce such imbalance.

The drawbacks of deploying phase balancer are low efficiency of the power electronics under high switching frequencies and high cost of applying phase balancers to all the LV networks.

2.4.2. Limitations of the Phase Balancing Solutions

Phase balancing solutions face the problems of adaptability and reliability [7]. Phase swapping is suitable for systems where one phase is consistently heavier-loaded or lighter-loaded than the other two. However, the uptake of a large amount of LCTs will cause random changes to the loading situations. Consequently, phase swapping will become less effective in rebalancing the phases. Demand-side management is an economical way of dealing with random imbalances.

Nonetheless, it is rather challenging to encourage customers to join the demand-side response and sell their flexibility in energy usage. DNOs need to find effective ways of incentivising customers and protecting customers' private information at the same time. Besides, demand-side management faces a high chance of non-delivery as customers may fail to respond to the DNO or insufficient flexibilities provided by the customers. Phase balancer performs the best in terms of real-time phase balancing among all the solutions. However, the application and maintenance of phase balancers to all the distribution networks will be costly. The phase balancers are composed of power electronics with approximate lifespans of 10~30 years [111, 112] which will incur additional maintenance costs.

Another main limitation of phase balancing is data-scarcity. The phase imbalance needs to be assessed by DNOs before choosing a suitable phase balancing solution for LV distribution networks. The assessment of phase imbalance includes a range of factors, such as the imbalance-induced costs. The assessment will require full time-series data from the network. However, most of the UK's LV networks are data-scarce due to lack of monitoring. The typical monitoring in the UK uses a Maximum Demand Indicator which records peak currents for the three phases [113]. Some references indicate that yearly average data for the three phases are also available [75, 114]. Installing monitoring devices to one network may be affordable for the DNO but scaling up to the mass population of networks will incur extremely high cost.

2.5. Research Gaps

This chapter reviewed existing research and projects about phase imbalance in the current and future distribution system. Four gaps are identified in this chapter:

1) **Decomposing three-phase power series into systematic and random components.**
   Existing research mainly focuses on analysing the causes, consequences and solutions to phase imbalances. There is a gap in decomposing three-phase power imbalances into components. Uneven load distribution and random load behaviours are the main causes of phase power imbalance in the LV networks. Uneven load distribution results in systematic imbalance (SIB) while random load leads to the random imbalance (RIB). Decomposing SIB and RIB of the phase power imbalance helps DNOs minimise the cost and maximise the efficiency of phase balancing.

2) **Estimating the cost-benefit of phase balancing solutions with limited data.**
   While analysing the consequences of phase imbalance, the imbalance-induced costs are discussed. However, the calculation of imbalance-induced costs requires a complete time-series of voltage and current data. This is challenging for the LV distribution networks because only substation data are collected per annum. In other words, only the yearly average, peak and total data are available for the majority of LV distribution networks. Installing monitoring devices to all the LV distribution networks will incur excessively high cost. Consequently, there is a gap in estimating the imbalance-induced cost with limited data. The estimation of the imbalance-induced cost will help DNOs assess the potential phase balancing solutions before investing.

3) **Analysing the potential impacts of LCTs on phase power imbalance.**

The increasing penetrations of LCTs will have uncertain impact on phase imbalance. The phase imbalance could be aggravated or mitigated in the future depending on DNO's system planning and managing. Analysing the potential impacts will help DNOs understand the possible imbalance-induced cost under different LCT penetration levels. The estimated imbalance-induced costs help DNOs analyse the cost-benefits for future phase balancing solutions.

Chapter 3. Characteristics of Phase Imbalances

This chapter introduces a way of decomposing three-phase power imbalance into systematic imbalance and random imbalance to guide phase balancing.

3.1. Introduction

Two primary reasons for phase imbalance in the LV networks are uneven load distribution and arbitrary load behaviours. The uneven load distribution contributes to the systematic part of the total power imbalance, i.e., systematic imbalance (SIB). An effective solution for the SIB is phase swapping. The arbitrary load behaviour is the random component in the total power imbalance, i.e., random imbalance (RIB). Demand-side management shows the advantage of reducing RIB.

This chapter explains how to decompose the three-phase power series into SIB and RIB components to deduce the optimal phase balancing strategy. The strategy aims to maximize phase swapping and minimize demand-side management as the former is relatively cheaper and more reliable.

The content of this chapter is cited from a published article in IEEE Transactions on Power System by the author [66]. This chapter is formed in an alternative-based format. All the indices, figures, tables, equations and references are numbered independently.

The following sections are organised as follows: Section 3.2 presents the published paper which includes the details of gap identification, a priori judgement method, the methodology of decomposition, the numerical results and discussions. Section 3.3 presents additional analysis based on different seasons.

3.2. Three-Phase Power Imbalance Decomposition

This declaration concerns the article entitled:			
Three-phase power imbalance decomposition into systematic imbalance and random imbalance			
Publication status (tick one) **Draft manuscript** ☐ **Submitted** ☐ **In review** ☐ **Accepted** ☐ **Published** ✓			
Publication details (reference)	W. Kong, K. Ma and Q. Wu, "Three-Phase Power Imbalance Decomposition into Systematic Imbalance and Random Imbalance," in IEEE Transactions on Power Systems, vol. 33, no. 3, pp. 3001-3012, May 2018, DOI: 10.1109/TPWRS.2017.2751967.		
Copyright status (tick the appropriate statement) I hold the copyright for this material ☐ Copyright is retained by the publisher, but I have been given permission to replicate the material here ✓			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The candidate contributed to / considerably contributed to / predominantly executed the… Formulation of ideas: • 90% • The use of degree of power imbalance to represent the trend of systematic imbalance is guided by Dr Kang Ma. Design of methodology: • 90% • The use of case IDs in a priori judgement is supported by Dr Kang Ma. Experimental work: • 100% Presentation of data in journal format: • 90% • The changes of data presentation during review process are guided by Dr Kang ma and Dr Qiuwei Wu		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed	Wangwei Kong	**Date**	20/12/2020

Three-Phase Power Imbalance Decomposition into Systematic Imbalance and Random Imbalance

Wangwei Kong, Kang Ma, *Member, IEEE*, and Qiuwei Wu, *Senior Member, IEEE*

***Abstract*— Uneven load allocations and random load behaviors are two major causes for three-phase power imbalance. The former mainly cause systematic imbalance, which can be addressed by low-cost phase swapping; the latter contribute to random imbalance, which requires relatively costly demand-side managements. To reveal the maximum potential of phase swapping and the minimum need for demand-side managements, this paper first proposes a novel a priori judgment to classify any set of three-phase power series into one of four scenarios, depending on whether there is a definite maximum phase, a definite minimum phase, or both. Then, this paper proposes a new method to decompose three-phase power series into a systematic imbalance component and a random imbalance component as the closed-form solutions of quadratic optimization models that minimize random imbalance. A degree of power imbalance is calculated based on the systematic imbalance component to guide phase swapping. Case studies demonstrate that 72.8% of 782 low voltage substations have systematic imbalance components. The degree of power imbalance results reveal the maximum need for phase swapping and the random imbalance components reveal the minimum need for demand side management, if the three phases are to be fully rebalanced.**

***Index Terms*— low voltage distribution network, power imbalance, random imbalance, systematic imbalance, three phase electric power**

I. NOMENCLATURE

$DPIB(t)$	The degree of power imbalance at time point t
N	The total number of time points
$P_{\emptyset}(t)$ where $\emptyset \in \{a, b, c\}$	Phase $\emptyset$ power at time point t
$\bar{P}_{\emptyset}$ where $\emptyset \in \{a, b, c\}$	The average power of phase $\emptyset$ over time
$P_{\emptyset s}(t)$ where $\emptyset \in \{a, b, c\}$	Phase $\emptyset$ power of the systematic imbalance component at time point t

$x_r(t), y_r(t), z_r(t)$	The random imbalance component of the three-phase power (phases x, y, and z) at time point t
$x_s(t), y_s(t), z_s(t)$	The systematic imbalance component of three-phase power (phases x, y, and z) at time point t

II. INTRODUCTION

More than 70% of the UK's low voltage (LV) networks experience observable degrees of three-phase imbalance [1]. Such an imbalance leads to: 1) neutral wire energy losses up to hundreds of millions of British pounds each year in the UK's distribution networks [2], [3]; and 2) additional network investment cost amounting to billions of British pounds each year [4], [5]. Major causes for this issue are uneven load allocations across the three phases and random load behaviors [6], [7], [8].

Uneven load allocations cause systematic imbalance (SIB) where there is a definite maximum phase (a definite phase with the greatest power among the three phases), a definite minimum phase (a definite phase with the least power among the three phases), or both. SIB can be addressed by phase swapping [9], [10], [11], i.e., moving single-phase loads/laterals from one phase to another, which is a relatively cheap and mature technique.

Random load behaviors, on the other hand, are a major contributor to random imbalance (RIB) with neither a definite maximum phase nor a definite minimum phase. RIB requires demand-side managements [12], [13] to address, which incur relatively high implementation and operation costs (including the costs for per-phase monitoring, communication, and control systems) and a risk of non-delivery.

The motivation and objective of this paper is therefore to find a way to decompose any set of time series power data from three phases into a SIB component and a RIB component that reveal the maximum potential for phase swapping and the minimum need for demand-side managements, thus corresponding to the lowest cost to rebalance three-phase supply. This idea is analogous to the decomposition of physics experiment observational errors into systematic errors and random errors [14]: systematic errors result from the non-ideal mechanism (analogous to the non-ideal load allocation across three phases) of the experiment. It has a non-zero mean and is not reduced when observations are averaged [14]. Random errors, on the other hand, result from inherently unpredictable fluctuations [14], which are analogous to the random individual load fluctuations.

Three-phase power imbalance is the direct result of voltage/current imbalances [15]. The majority of publications studied the underlying power imbalance components, i.e., voltage/current imbalances: References [16], [17], [18] estimated voltage imbalance of medium-voltage (MV) distribution networks. Reference [19] assessed the sequence values of imbalanced voltages without phasor measurements. Reference [20] quantified current imbalance on short transmission lines. Reference [21] forecasted voltage imbalance on low voltage feeders with photovoltaic (PV) generation. Reference [22] converts three-phase imbalanced currents into two orthogonal AC currents with equal amplitudes. The above references implicitly decompose three-phase power imbalance into voltage and current imbalances.

A number of publications that focus on power imbalance are about reducing power imbalance [15], [23], [13], rather than on decomposing power imbalance into its underlying components.

The decomposition of three-phase imbalanced power series into a SIB component and a RIB component is a gap. The purpose for the decomposition is mentioned above. To bridge the gap, this paper makes the following contributions:

1) Propose a novel a priori judgment method to classify any set of three-phase power series into one of the following four scenarios: definite-max, definite-order, definite-min, and random imbalance scenarios (their definitions are given in Section III). The judgment method takes into account both the percentage of time when the definite phase occurs and the average power to ensure a robust judgment.

2) Propose a novel three-phase power decomposition method for all scenarios except the random imbalance one to decompose three-phase power series into a SIB component and a RIB component, which are the closed-form solution to a quadratic optimization problem that minimizes the RIB component.

3) Define the degree of power imbalance for each of the definite-max, definite-order, and definite-min scenarios based on the SIB component obtained from 2) and calculate the trend of the degree of power imbalance over time.

The SIB component, as a direct consequence of uneven load allocations, serves as the basis for calculating the degree of power imbalance, which provides a direct guidance for phase swapping; the RIB component, as a result of random individual load behaviors, indicates at least how much power on each phase has to be reduced by demand-side managements, if the three phases are to be fully rebalanced.

Therefore, the research outcome brings three values: 1) the decomposition helps distribution network operators (DNOs) to understand the potential (also the maximum need) of phase swapping to address SIB and how much power on each phase has to be reduced by demand-side managements, if the three phases are to be fully rebalanced; 2) By calculating the degree of power imbalance based on the SIB component, the research also reveals the underlying trend of the SIB over time, reflecting the trend of uneven load allocations – this is particularly useful when increasing single-phase electric vehicles and heat pumps are connected to low voltage networks, causing the SIB to vary over time; 3) the degree of power imbalance also provides a guidance for phase swapping practices.

The remainder of the paper is organized as follows: Section III presents an overview of the methodology; Section IV presents a new a priori judgment method; Section V details the decomposition method; Section VI defines the degree of power imbalance; Section VII performs a case study; and Section VIII concludes the paper.

III. OVERVIEW OF METHODOLOGY

The proposed methodology requires three-phase power series as an input only. Therefore, as a mathematical method, it is applicable to where: 1) there is monitoring of three-phase power (or three-phase voltages and currents which can be used to derive power); and 2) there is three-phase power imbalance. In reality, the methodology is highly suitable for monitored low voltage distribution networks in the UK and the rest of Europe and monitored medium voltage distribution networks in the US, where three-phase power imbalance is obvious.

Figure 3-1 shows an overview of the methodology.

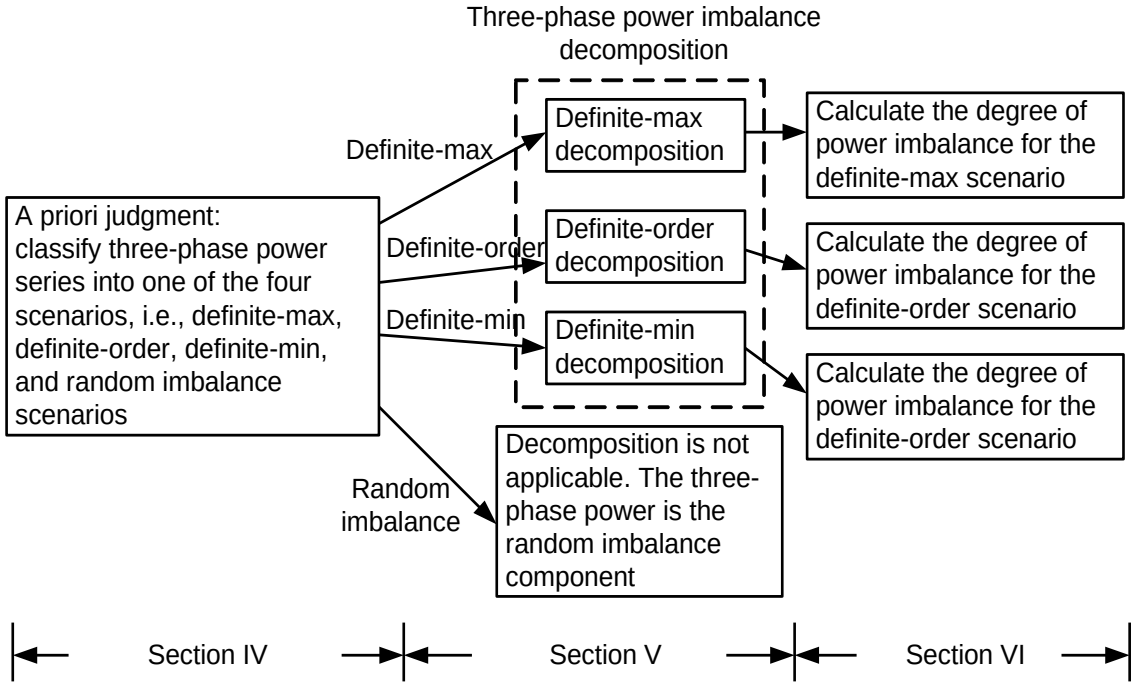


Figure 3-1. An overview of methodology

Each phase has a time series of power (called a power series) monitored at the LV (415V) substation side. The following definitions are used throughout the paper:

1) Three-phase power series: a set of three time series of power data monitored and collected from three phases. The data are normally measured from distribution substations at an interval of T_{int}, e.g., $T_{int} = 10$ min.
2) Definite-max phase: a definite phase with the greatest power among three phases.
3) Definite-max scenario: the scenario where there is a definite-max phase for the majority of time.
4) Definite-min phase: a definite phase with the least power among three phases.
5) Definite-min scenario: the scenario where there is a definite-min phase for the majority of time.
6) Definite order: the existence of both a definite-max phase and a definite-min phase, e.g., 'phase a > phase b > phase c'.
7) Definite-order scenario: the scenario where there are both definite-max and definite-min phases for the majority of time.
8) Random imbalance scenario: the scenario where there is neither a definite-max phase nor a definite-min phase.

9) SIB component: a set of three-phase power series with a definite-max phase, a definite-min phase, or a definite order.
10) RIB component: a set of three power series with neither a definite-max phase nor a definite-min phase.

IV. A PRIORI JUDGMENT

This section presents a new a priori judgment method to classify any set of three-phase power series into one of the four scenarios (definite-max, definite-order, definite-min, and random imbalance scenarios). The judgment considers both the percentage of time when a definite phase/order occurs and the average power. The rationale for this is to ensure robustness: the definite phase/order, if exist, should not only occur for the majority of time but also have the average power showing the same trend. The judgment method consists of three steps:

Step 1): The percentage of time judgment

In principle, Step 1) judgment indicates that:

1) If for the majority of time, phase a is the definite-max phase *and* phase c is the definite-min phase, then this is a definite-order scenario with a definite three-phase order: phase a > phase b > phase c.
2) If condition 1) is *not* met, *and* phase a has the greatest power among the three phases for the majority of time which is no less than the time when any phase has the least power among the three phases, then this is a definite-max scenario where phase a is the definite-max phase.
3) If condition 1) is *not* met, *and* phase c has the least power among the three phases for the majority of time which is more than the time when any phase has the greatest power among the three phases, then this is a definite-min scenario where phase c is the definite-min phase.
4) Any scenario that does not meet conditions 1) – 3) is a random imbalance scenario with neither a definite-max phase nor a definite-min phase.

The percentage of time when phases a is the definite-max phase and phase c is the definite-min phase is given by,

$$C_{ac} = \frac{\sum_{t=1}^{N} \alpha_a(t)\beta_c(t)}{N} \times 100\% \tag{3-1}$$

where N is defined in Section I. $\alpha_a(t)$ is a binary value:

$$\alpha_a(t) = \begin{cases} 1 & \text{when } P_a(t) > (1+\delta_1)P_b(t) \text{ and } P_a(t) > (1+\delta_1)P_c(t) \\ 0 & \text{otherwise} \end{cases} \quad (3\text{-}2)$$

where $P_\emptyset(t)$ is defined in Section I. δ_1 is a threshold to distinguish any two power values, e.g., $\delta_1 = 5\%$. Such a threshold accounts for measurement errors, which arise from monitoring devices, the communication system, and other factors. This value is chosen according to network operator's experience. If the difference between two power values is below this threshold, then the difference is immersed in the measurement error and is not regarded as a credible difference. In this paper, $\delta_1 = 5\%$ by default.

$\beta_c(t)$ is also a binary value:

$$\beta_c(t) = \begin{cases} 1 & \text{when } P_c(t) < (1-\delta_1)P_a(t) \text{ and } P_c(t) < (1-\delta_1)P_b(t) \\ 0 & \text{otherwise} \end{cases} \quad (3\text{-}3)$$

The percentage of time when each phase a has the greatest power is given by,

$$A_a = \frac{\sum_{t=1}^{N} \alpha_a(t)}{N} \times 100\% \quad (3\text{-}4)$$

where $\alpha_a(t)$ is given by (3-2).

Similarly, the percentage of time when each phase c has the least power is given by,

$$B_c = \frac{\sum_{t=1}^{N} \beta_c(t)}{N} \times 100\% \quad (3\text{-}5)$$

where $\beta_c(t)$ is given by (3-3). Based on the results from (3-1), (3-4), and (3-5), Step 1) judgment is listed in Table 3-1.

Table 3-1. Step 1) judgment

Case ID	Condition	Step 1) judgment
1	If ∃phase a, such that $C_{\emptyset_1\emptyset_2} < 50\%$ and $A_a \geq 50\%$ and $A_a \geq B_\emptyset$, where $\emptyset_1, \emptyset_2, \emptyset \in \{a, b, c\}$	Definite-max scenario: phase a is the definite-max phase
2	If ∃phase c, such that $C_{\emptyset_1\emptyset_2} < 50\%$ and $B_c \geq 50\%$ and $B_c \geq A_\emptyset$, where $\emptyset_1, \emptyset_2, \emptyset \in \{a, b, c\}$	Definite-min scenario: phase c is the definite-min phase
3	If ∃phases a and c, such that $C_{\mathrm{ac}} \geq 50\%$	Definite-order scenario: phase a > phase b > phase c
4	Other	Random imbalance scenario

The 50% threshold of time is consistent with the criteria detailed at the beginning of this section, where the term “majority” means a 50% threshold by default.

It should be noted that Step 1) produces preliminary judgment results which are not necessarily the final ones.

Step 2): Calculation of the average power

The average power of each phase ∅ is given by,

$$\bar{P}_\emptyset = \frac{\sum_{t=1}^{N} P_\emptyset(t)}{N} \quad \text{where } \emptyset \in \{a, b, c\} \tag{3-6}$$

where all variables are defined in Section I.

The resultant set of the three average power $\{\bar{P}_a, \bar{P}_b, \bar{P}_c\}$ will be used for judgment in Step 3).

Step 3): Combined judgment

Table 3-2 lists the final judgments of the scenarios (in the right column of Table 3-2) as the combinations of the judgments from Step 1) and Step 2). It should be noted that the logic is 'and' between the conditions in the first (Step 1) and second (Step 2) columns.

Table 3-2. Final judgments

Step 1) case ID	Step 2) condition	Combined judgment
1	If $\bar{P}_a = \max\{\bar{P}_a, (1+\delta_1)\bar{P}_b, (1+\delta_1)\bar{P}_c\}$	Definite-max scenario: phase a is the definite-max phase
	Otherwise	Random imbalance scenario
2	If $\bar{P}_c = \min\{\bar{P}_c, (1-\delta_1)\bar{P}_a, (1-\delta_1)\bar{P}_b\}$	Definite-min scenario: phase c is the definite-min phase
	Otherwise	Random imbalance scenario
3	If $\bar{P}_a > (1+\delta_1)\bar{P}_b$ and $\bar{P}_b > (1+\delta_1)\bar{P}_c$	Definite-order scenario: phase a > phase b > phase c
	If $\bar{P}_c = \min\{(1-\delta_1)\bar{P}_a, (1-\delta_1)\bar{P}_b, \bar{P}_c\}$	Definite-min scenario: phase c is the definite-min phase
	If $\bar{P}_a = \max\{\bar{P}_a, (1+\delta_1)\bar{P}_b, (1+\delta_1)\bar{P}_c\}$	Definite-max scenario: phase a is the definite-max phase
	Otherwise	Random imbalance scenario
4	Any	Random imbalance scenario

δ_1 is the same threshold as appeared in (3-2). The left column, Step 1) case ID, corresponds to the case ID in Table 3-1.

V. POWER IMBALANCE DECOMPOSITION

According to the priori judgment in Section IV, three-phase power series are classified into four scenarios, i.e., definite-max, definite-order, definite-min, and random imbalance scenarios. For the first three scenarios, this section presents three decomposition methods: definite-max decomposition, definite-order decomposition, and definite-min decomposition. Each decomposition

corresponds to a quadratic optimization problem that minimizes the RIB component. The fourth scenario, i.e., the random imbalance scenario, cannot be decomposed.

The quadratic optimization problems have closed-form solutions. Therefore, the decompositions are achieved by directly applying the closed-form solutions without the need for iterations to solve the optimization problems. This significantly simplifies the decomposition process and ensures that the decomposition only has linear complexity (linear to the length of the three-phase power series).

The resultant SIB can be addressed by phase swapping, which is a low frequency, relatively long-lasting, mature solution. However, phase swapping is not suitable for resolving RIB which does not have a particular phase order. The RIB requires solutions such as demand-side managements, which incur higher monitoring, communication, and control costs as well as a risk of non-delivery. Therefore, this justifies the principle of the three decompositions to minimize the RIB component, thus revealing the maximum potential of phase swapping (i.e., the maximum possible reduction in phase imbalance from phase swapping) as well as the minimum need for demand-side managements.

The purposes of the decomposition are twofold: i) to serve as the basis to calculate the degree of power imbalance (in Section VI), which not only reveals the potential of phase swapping to address the SIB but also serves as a guidance for phase swapping; ii) to understand how much power on each phase needs to be reduced by demand-side managements, if the three phases are to be fully rebalanced.

A. Definite-Max Decomposition

The definite-max decomposition decomposes imbalanced three-phase power series into: 1) a SIB component with a definite-max phase; and 2) a RIB component without the definite-max phase.

The definite-max decomposition applies to the definite-max scenario, in which phase a is defined as the definite-max phase. The definite-max decomposition is the solution to the following quadratic optimization problem:

$$\begin{aligned} \min \sum_{t=1}^{N} x_r^2(t) + y_r^2(t) + z_r^2(t) \\ \text{subject to} \quad P_a(t) = x_r(t) + x_s(t); \\ P_b(t) = y_r(t) + y_s(t); \\ P_c(t) = z_r(t) + z_s(t) \\ x_s(t) \geq y_s(t); \\ x_s(t) \geq z_s(t); \\ x_s(t), y_s(t), z_s(t) \geq 0; \\ x_r(t), y_r(t), z_r(t) \geq 0 \end{aligned} \tag{3-7}$$

where all variables are defined in Section I.

The quadratic optimization problem given by (3-7) aims to minimize the RIB component that requires demand-side management to address. This is justified because demand-side management, which addresses RIB, incurs relatively high implementation and operation costs and a risk of non-delivery. On the other hand, phase swapping, which addresses systematic imbalance, is a relatively economic and mature technique. By minimizing RIB (hence maximizing SIB), the quadratic optimization model aims to reveal the minimum need for demand-side management and the maximum need for phase swapping, thus corresponding to the lowest cost. The same justification applies to the optimization problems for the definite-order and definite-min scenarios.

The original problem of (3-7) minimizing the summation of a time series is transformed into a total of N optimizations, each for a time point t. In this way, the summation is removed and the objective function of (3-7) becomes:

$$\forall t \in [1, N] \qquad \min x_r^2(t) + y_r^2(t) + z_r^2(t) \tag{3-8}$$

A closed-form solution exists for the optimization problem in (3-8). The solution includes both the SIB and RIB components, assuming that phase a is the definite-max phase:

$$\text{SIB} = [P_a(t), P_{2s}(t), P_{3s}(t)]^T \tag{3-9}$$

where $P_{2s}(t) = \min\{P_a(t), P_b(t)\};$

$$P_{3s}(t) = \min\{P_a(t), P_c(t)\}.$$

Because

$$\text{SIB} + \text{RIB} = [P_a(t), P_b(t), P_c(t)]^T \tag{3-10}$$

The RIB component is given by

$$\text{RIB} = [0, P_{2r}(t), P_{3r}(t)]^T \tag{3-11}$$

where $P_{2r}(t) = \max\{0,\ P_b(t) - P_a(t)\};$

$$P_{3r}(t) = \max\{0,\ P_c(t) - P_a(t)\}.$$

B. Definite-Order Decomposition

The definite-order decomposition decomposes imbalanced three-phase power series into: 1) a SIB component with a definite-order; and 2) a RIB component without the definite-order.

Suppose that the phase order is ‘a > b > c’. The quadratic optimization model is the same as given by (3-7) except that the first two inequality constraints are replaced by

$$x_s(t) \geq y_s(t); \quad y_s(t) \geq z_s(t)$$

The definite-order decomposition is the closed-form solution to the optimization model. Assuming that the order of the three phases is ‘a > b > c’, the SIB component is given by,

$$\text{SIB} = [P_a(t), P_{2s}(t), P_{3s}(t)]^T \tag{3-12}$$

where $P_{2s}(t) = \min\{P_a(t), P_b(t)\}$, $P_{3s}(t) = \min\{P_a(t), P_b(t), P_c(t)\}$.

Equation (3-10) still holds. The RIB component is given by,

$$\text{RIB} = [0, P_{2r}(t), P_{3r}(t)]^T \tag{3-13}$$

where $P_{2r}(t) = \max\{0,\ P_b(t) - P_a(t)\}$;

$$P_{3r}(t) = \max\{0, P_c(t) - P_a(t), P_c(t) - P_b(t)\}.$$

The definite-order scenario provides more information than the definite-max and definite-min scenarios, because its SIB component gives a definite three-phase order with both definite-max and definite-min phases, whereas the SIB components of the latter two scenarios give only the definite-max or the definite-min phase. On the other hand, the definite-order scenario is more restrictive than the latter two because it requires that a definite three-phase order exists.

C. Definite-Min Decomposition

The definite-min decomposition decomposes imbalanced three-phase power series into: 1) a SIB component with a definite-min phase; and 2) a RIB component without the definite-min phase.

Suppose that the definite-min phase is phase c. The quadratic optimization model is the same as given by (3-7) except that the first two inequality constraints are replaced by

$$x_s(t) \geq z_s(t);\quad y_s(t) \geq z_s(t) \tag{3-14}$$

The definite-min decomposition is the closed-form solution to the optimization problems. Assuming that phase c is the definite-min phase, the SIB component is given by,

$$\text{SIB} = [P_a(t), P_b(t), P_{3s}(t)]^T \tag{3-15}$$

where $P_{3s}(t) = \min\{P_a(t), P_b(t), P_c(t)\}$.

Equation (3-10) still holds. The RIB component is given by,

$$\text{RIB} = [0, 0, P_{3r}(t)]^T \tag{3-16}$$

where $P_{3r}(t) = \max\{0, P_c(t) - P_a(t), P_c(t) - P_b(t)\}$.

For the definite-max, definite-order, and definite-min scenarios, the SIB component is the basis for calculating the degree of power imbalance, which provides a direct guidance for phase swapping as explained in Section VI. The $\text{RIB} = [P_{1r}(t), P_{2r}(t), P_{3r}(t)]^T$ has a clear meaning: for phases a, b, and c, at least $P_{1r}(t)$, $P_{2r}(t)$, and $P_{3r}(t)$ of loads require demand-side managements for phase rebalancing, respectively.

VI. Degree of Power Imbalance

This section presents the definitions for the degree of power imbalance for the definite-max, definite-order, and definite-min scenarios. For all three scenarios, the degree of power imbalance is defined as the deviation of the definite-max/definite-min phase from the average, based on the SIB component. The definition of the degree of power imbalance is to not only reveal the trend of SIB over time but also guide phase swapping (as explained later in this section). Assume that phase a is the definite-max phase for the definite-max scenario; phase c is the definite-min phase for the definite-min scenario; and the phase order is 'a > b > c' for the definite-order scenario. The mathematical definition for the degree of power imbalance for each scenario is given by,

$$\text{Definite-max: } DPIB(t) = \frac{P_{as}(t) - \frac{\sum_{\emptyset\in\{a,b,c\}} P_{\emptyset s}(t)}{3}}{\sum_{\emptyset\in\{a,b,c\}} P_{\emptyset s}(t)} \times 100\% \tag{3-17}$$

$$\text{Definite-min: } DPIB(t) = \frac{\frac{\sum_{\emptyset\in\{a,b,c\}} P_{\emptyset s}(t)}{3} - P_{cs}(t)}{\sum_{\emptyset\in\{a,b,c\}} P_{\emptyset s}(t)} \times 100\% \tag{3-18}$$

Definite-order:

$$DPIB(t) = [D_1 \quad D_2] = \left[\frac{P_{as}(t) - \frac{\sum_{\emptyset \in \{a,b,c\}} P_{\emptyset s}(t)}{3}}{\sum_{\emptyset \in \{a,b,c\}} P_{\emptyset s}(t)} \quad \frac{\frac{\sum_{\emptyset \in \{a,b,c\}} P_{\emptyset s}(t)}{3} - P_{cs}(t)}{\sum_{\emptyset \in \{a,b,c\}} P_{\emptyset s}(t)} \right] \tag{3-19}$$

where all variables are defined in Section I. SIB $= [P_{as}(t), P_{bs}(t), P_{cs}(t)]^T$ as given by (3-9), (3-13), and (3-16). It should be noted that for the definite-max or definite-min scenarios, the degree of power imbalance is a single value; but for the definite-order scenario, the degree of power imbalance is a vector of two values.

The average three-phase power of the SIB component is given by,

$$\bar{P}_s(t) = \frac{\sum_{\emptyset \in \{a,b,c\}} P_{\emptyset s}(t)}{3} \tag{3-20}$$

where $P_{\emptyset s}(t)$ is defined in Section I.

The degree of power imbalance is a time series. It brings three values by: 1) revealing the trend of the SIB over time, i.e., the trend of uneven load allocations – this is particularly useful when increasing single-phase electric vehicles are connected to the network; 2) showing the potential of phase swapping to address SIB; iii) and providing a direct guidance for phase swapping:

i) For the definite-max scenario, the degree of power imbalance suggests the move of loads totalling $3\bar{P}_s(t)DPIB(t)$ from the definite-max phase to the other two phases equally, where $\bar{P}_s(t)$ is given by (3-21).
ii) For the definite-order scenario, the degree of power imbalance suggests the move of loads totaling $3\bar{P}_s(t)D_1$ away from the definite-max phase and the move of $3\bar{P}_s(t)D_2$ to the definite-min phase, where D_1 and D_2 are defined in (3-19).
iii) For the definite-min scenario, the degree of power imbalance suggests the move of loads totaling $3\bar{P}_s(t)DPIB(t)$ to the definite-min phase from the other two phases.

VII. Numerical Results

The input data are three-phase power series for 782 low voltage substations derived from the three-phase voltages and currents monitored at the secondary side of 11kV/415V transformers throughout Western Power Distribution (a UK DNO)'s business area [24]. Therefore, the three-phase power series are the power injected from 11kV networks to 415V networks. The data cover a good mix of geographical characteristics and customer types [24]. Four representative substations are selected to demonstrate the methodology. MATLAB is used for the simulation.

A. Definite-Max Scenario

Substation No. 536,753 is selected to represent the definite-max scenario. The study period is one year, covering five seasons (spring, summer, high summer, autumn, and winter) and different day types (weekday and weekend). Because the original three-phase power series and the SIB component have more than 50,000 time points (one sample every 10 minutes for a year) on the X axis, they are presented in the form of probability density distributions for clarity. This also applies to the definite-order and definite-min scenarios. The probability density functions of the three-phase power series are presented in Figure 3-2.

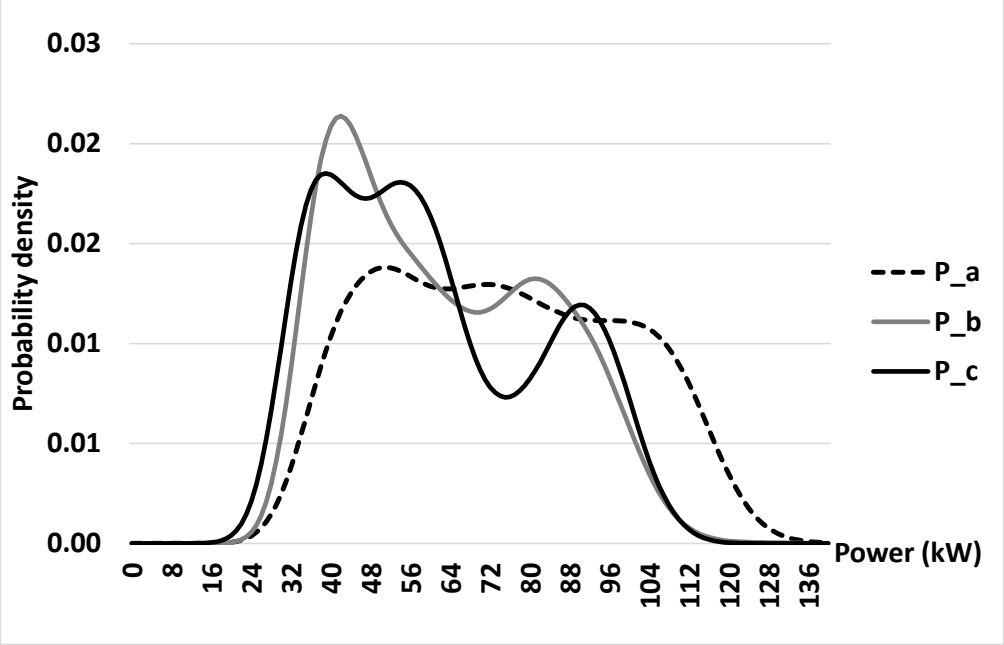


Figure 3-2. The probability density functions of the three-phase power series over a year for definite-max scenario

The a priori judgment process is presented in Table 3-3.

Table 3-3. A priori judgment for the definite-max scenario

Sub No.	Variables	Phase a	Phase b	Phase c
536,753	$A_{\emptyset}$	71.47%	7.15%	2.67%
	$B_{\emptyset}$	1.66%	26.18%	37.88%
	$\bar{P}_{\emptyset}$ (kW)	74.25	62.63	61.03

$A_{\emptyset}$, $B_{\emptyset}$, and $\bar{P}_{\emptyset}$ are given by (3-4), (3-5), and (3-6), respectively.

Phase a is the definite-max phase. Although phase c has the least power among the three phases for the majority of time (as shown in the second row of Table 3-3), its average power is approximately the same as that of phase b (their difference is lower than the threshold δ_1 as defined in (3-2)). Therefore, phase c is not judged as the definite-min phase and only the definite-max phase exists in this case.

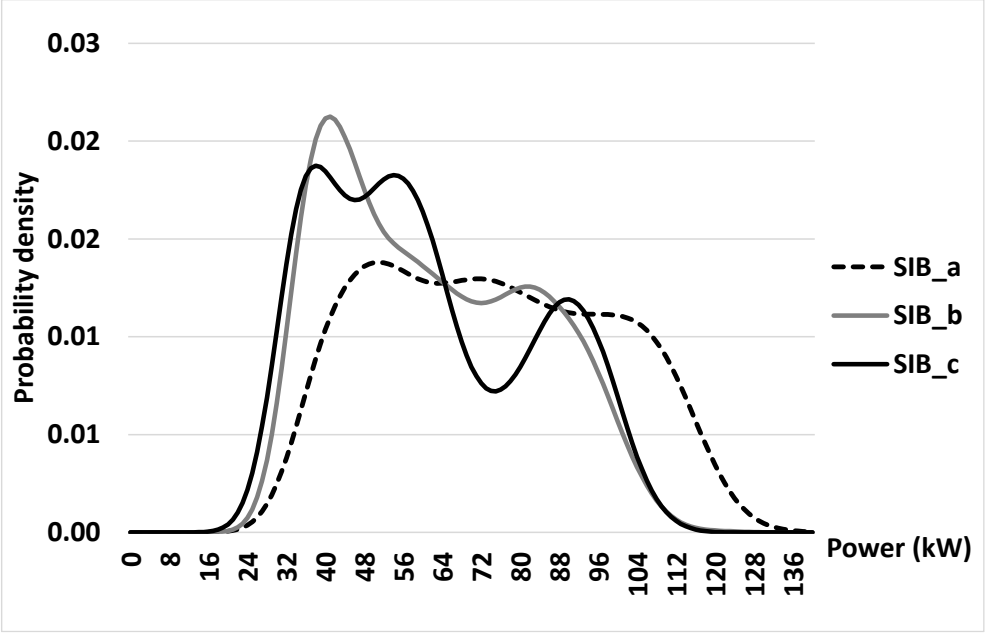


Figure 3-3. The probability density functions of the SIB component over a year for definite-max scenario

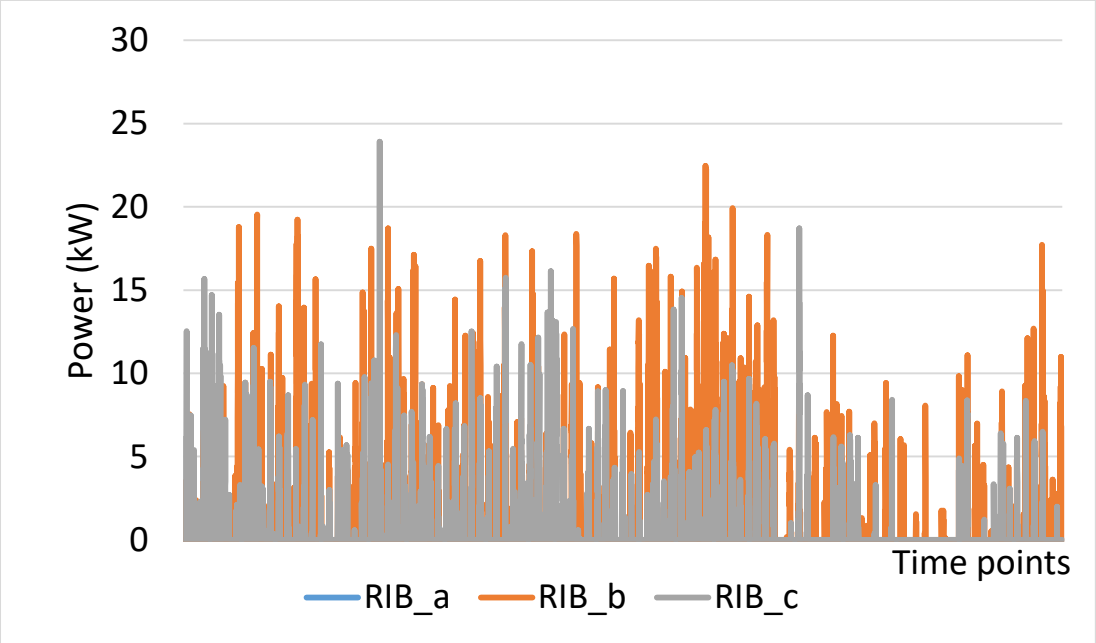


Figure 3-4. The RIB component over a year for definite-max scenario

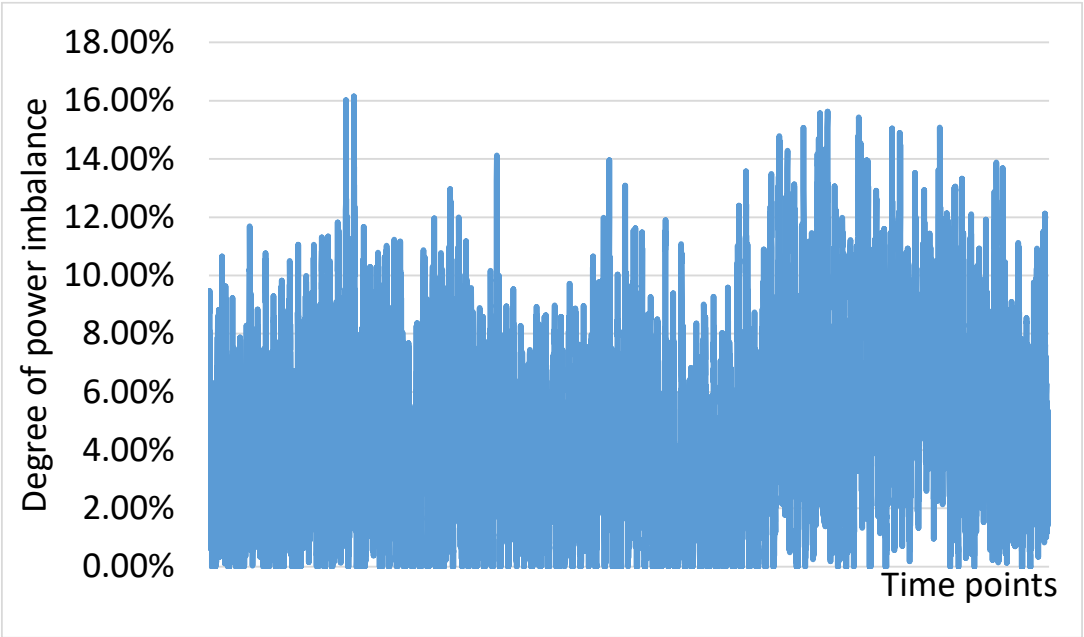


Figure 3-5. The degree of power imbalance over a year for definite-max scenario

Similar as Figure 3-2, the SIB component is also presented in the form of probability density functions in Figure 3-3. It can be seen that the probability density functions in Figure 3-3look similar to those in Figure 3-2 with only slight differences. In the SIB component, phase a has the greatest power

among the three phases for 100% of the time – this is consistent with the definition of the SIB for the definite-max scenario, reflecting that too much load is allocated to phase a.

The RIB component presented in Figure 3-4 shows the anomalies of when either phase b or c overtakes phase a to become the maximum phase – this occurs for 18.40% of the time, reflecting the random load fluctuations on phases b and c.

For the whole year, the degree of power imbalance results in Figure 3-5 provide a guidance for phase swapping: 1) on average, up to 8.61kW of loads can be moved from phase a to the other two phases; 2) at 15:50 on the 80st day in the year (21th March), a maximum of 34.88kW of loads can be moved from phase a to the other two phases; 3) for 4.20% of time, no load needs to be moved from phase a to the other two phases (the degree of power imbalance is zero during this period). If phase swapping is performed to move loads away from phase a, then the loads on the other two phases need to be reduced for phase rebalancing during this minority period.

B. Definite-Order Scenario

Substation No. 512,457 is selected to represent the definite-order scenario. The probability density functions of the three-phase power series are presented in Figure 3-6. The priori judgment process is presented in Table 3-4

Table 3-4. A priori judgment for the definite-order scenario

Sub No.	Variables	Phase a	Phase b	Phase c
512,457	$A_{\emptyset}$	83.35%	5.88%	0
	$B_{\emptyset}$	0	1.10%	96.11%
	$\bar{P}_{\emptyset}$ (kW)	45.36	36.89	25.26

In this case, $C_{\mathrm{ac}} = 79.86\%$. The average power also demonstrates the order of 'a > b > c'. Therefore, phases a and c are the definite-max and definite-min phases, respectively.

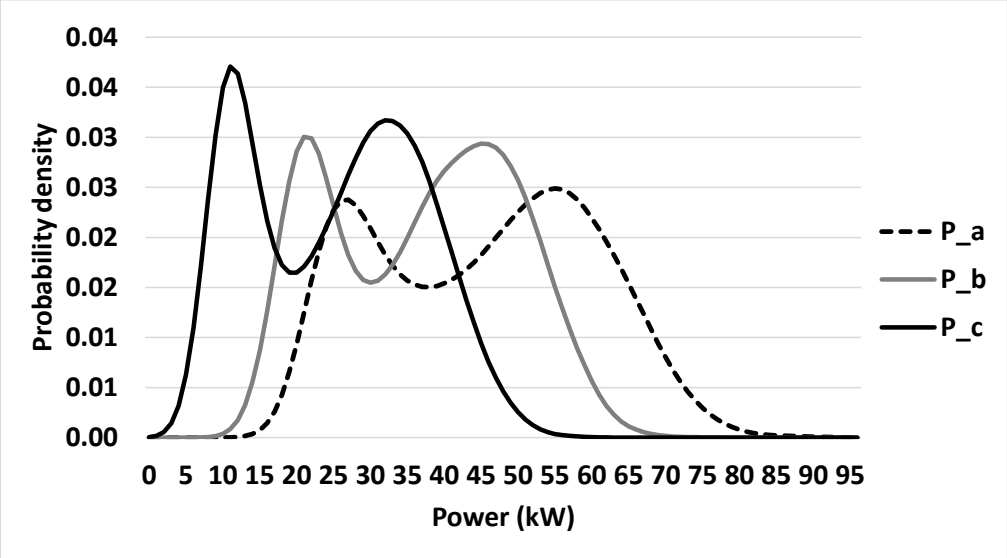


Figure 3-6. The probability density functions of the three-phase power series over a year for definite-order scenario

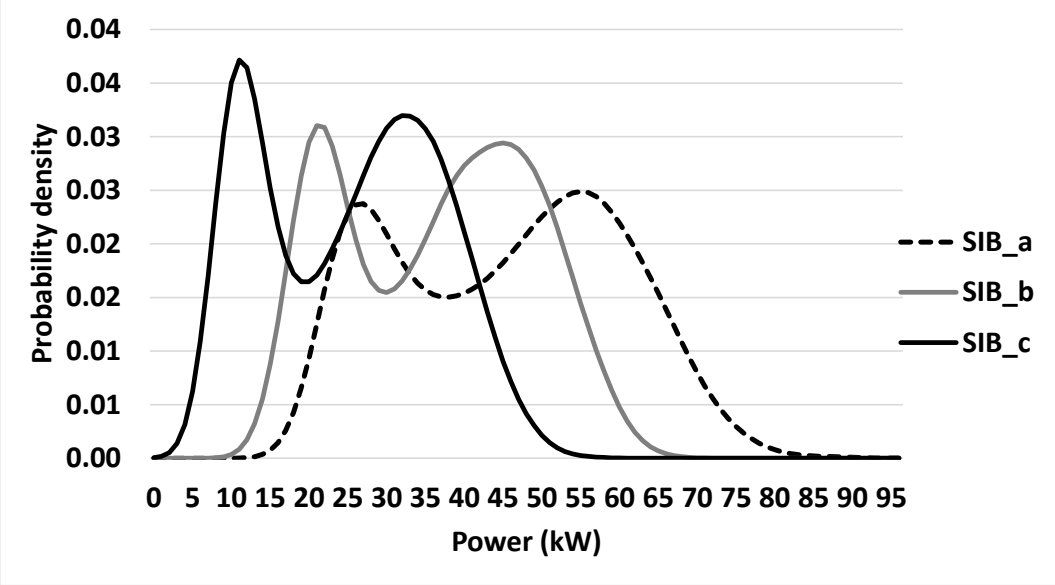


Figure 3-7. The probability density functions of the SIB component over a year for definite-order scenario

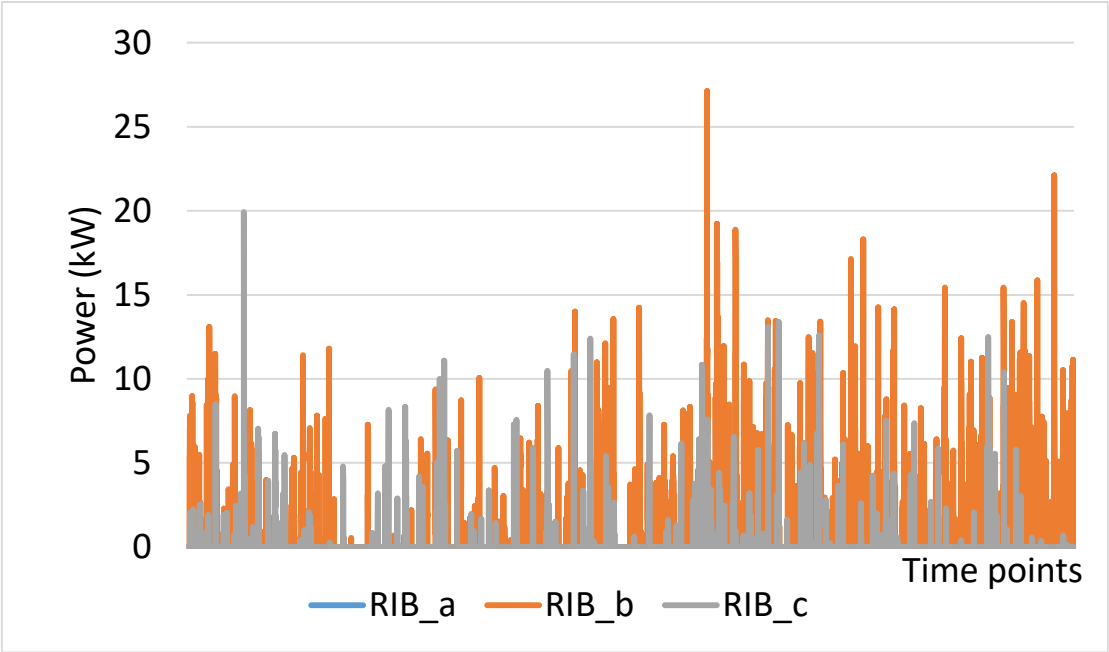


Figure 3-8. The RIB component over a year for definite-order scenario

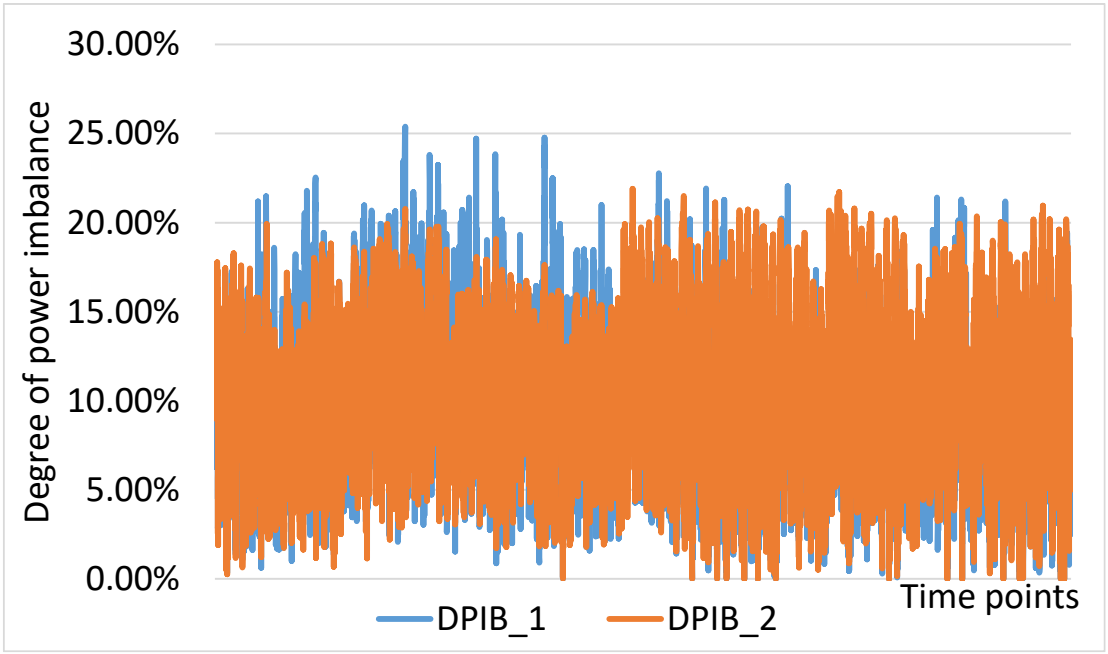


Figure 3-9. The degree of power imbalance over a year for definite-order scenario

For the whole year, the probability density functions of the SIB component is presented in Figure 3-7. According to the SIB component, phase a has the greatest power among the three phases for 100% of the time and that phase c has the least power for 100% of the time – this is consistent with the

definition of the SIB for the definite-order scenario, reflecting the existence of excessive loads on phase a and insufficient loads on phase c.

The RIB component presented in Figure 3-8 shows the anomalies when the phase order is not 'a > b > c' – this occurs for 12.27% of the time, reflecting the random load fluctuations on each phase.

The degree of power imbalance results in Figure 3-9 provide a guidance for phase swapping: 1) on average, up to 9.66kW of loads can be moved away from phase a and up to 10.50kW of loads can be moved to phase c; 2) at 22:30 on the 32nd day in the year (1st February), a maximum of 28.65kW of loads can be moved from phase a to the other two phases: 9.89kW to phase b and 18.76kW to phase c.

C. Definite-Min Scenario

Substation No. 521,071 is selected to represent the definite-min scenario. The three-phase power series are presented in Figure 3-10. The priori judgment process is presented in Table 3-5.

Phase a is the definite-min phase. Although phase b has a power greater than the other two phases by more than 5% for 67.62% of the time, the order of 'b > c > a' only occurs for 37.79% (< 50%) of the time – it does not meet the criteria for the definite-order scenario. Therefore, only a definite-min phase exists.

Table 3-5. A priori judgment for the definite-min scenario

Sub No.	Variables	Phase a	Phase b	Phase c
521,071	$A_{\emptyset}$	0	67.62%	13.00%
	$B_{\emptyset}$	69.26%	0	15.02%
	$\bar{P}_{\emptyset}$ (kW)	46.92	60.95	54.86

For a whole year, Figure 3-10 – Figure 3-13 depict the probability density functions of the three-phase power series, the probability density functions of the SIB component, the time series of the RIB component, and the degree of power imbalance, respectively.

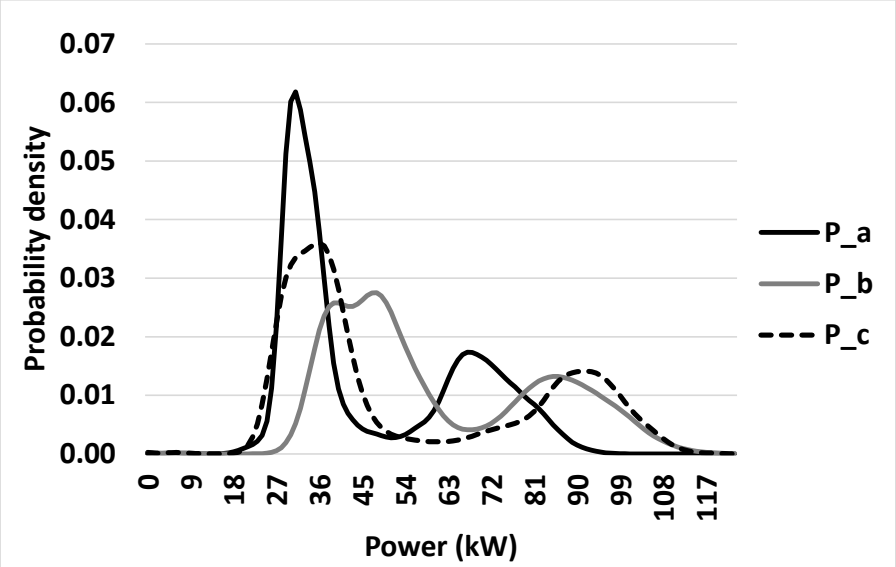


Figure 3-10. The probability density functions of the three-phase power series over a year for definite-min scenario

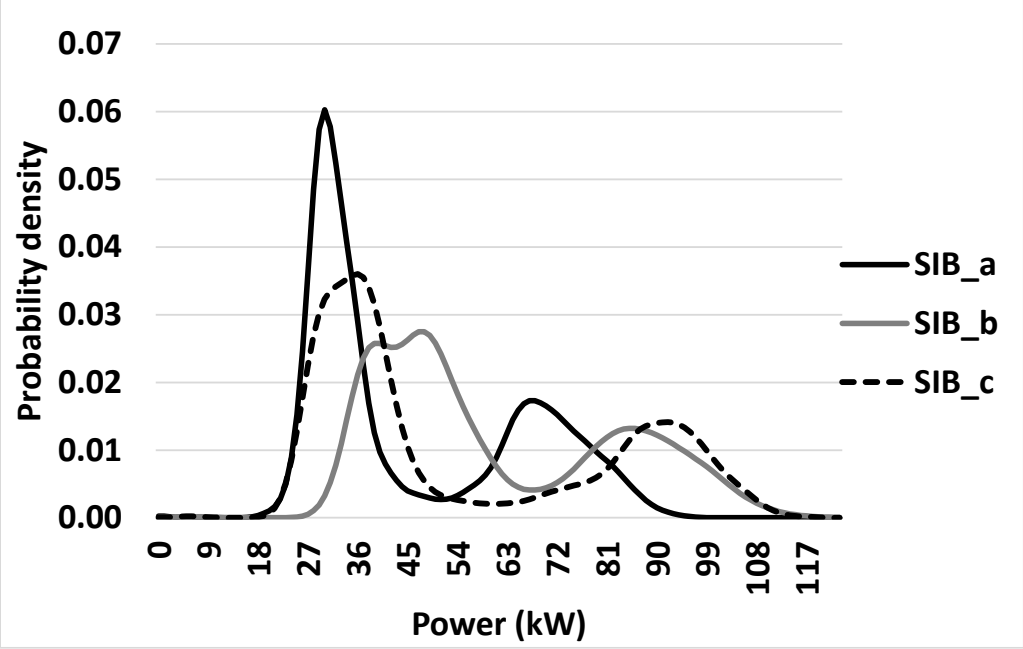


Figure 3-11. The probability density functions of the SIB component over a year for definite-min scenario

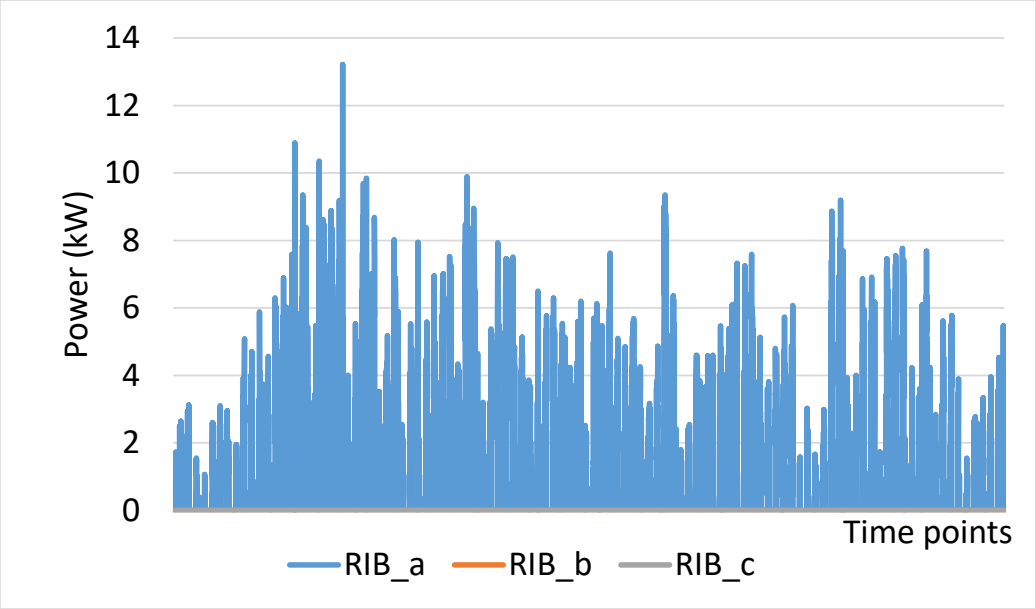


Figure 3-12. The RIB component over a year for definite-min scenario

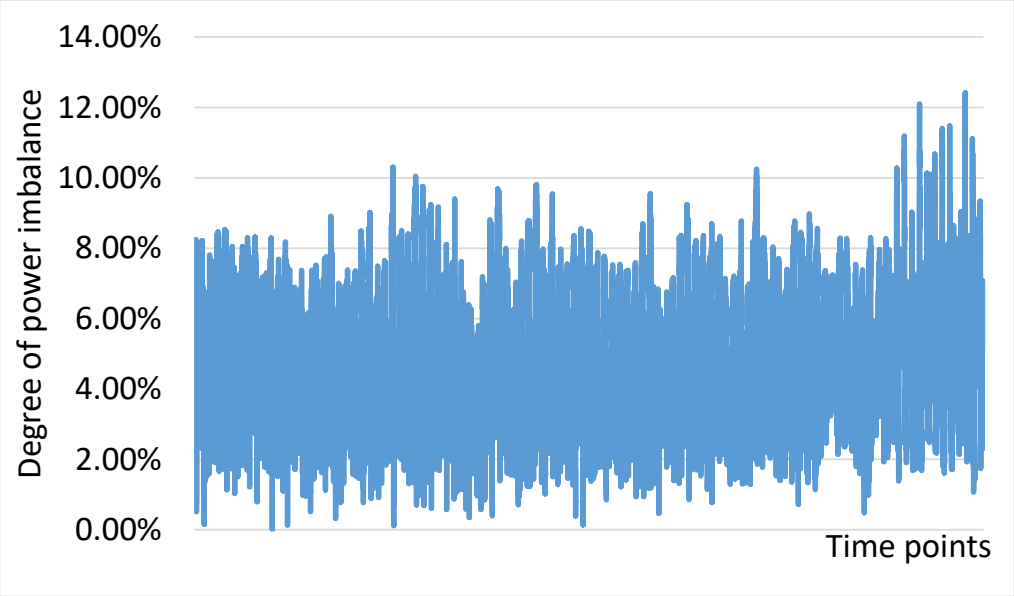


Figure 3-13. The degree of power imbalance over a year for definite-min scenario

The SIB component presented in Figure 3-11 shows that phase a has the least power among the three phases for 100% of the time – this is consistent with the definition of the SIB for the definite-min scenario, reflecting that insufficient loads are allocated to phase a.

The RIB component presented in Figure 3-12 shows that the abnormal cases when phase a is not the minimum phase occur for 21.58% of the time, reflecting the random load fluctuations that breach the majority rule.

The degree of power imbalance results in Figure 3-13provide a guidance for phase swapping: 1) on average, up to 7.75kW of loads can be moved from phases b and c to phase a; 2) At the 14192nd time point (at 13:20, 8th April), a maximum of 21.48kW of loads can be moved from phases b and c to phase a.

D. Random Imbalance Scenario

Substation No. 521,064 is selected to represent the random imbalance scenario. The a priori judgment process is presented in Table 3-6.

Table 3-6. A priori judgment for the random imbalance scenario

Sub No.	Variables	Phase a	Phase b	Phase c
521,064	$A_{\emptyset}$	30.10%	21.03%	26.26%
	$B_{\emptyset}$	25.39%	18.21%	31.99%
	$\bar{P}_{\emptyset}$ (kW)	46.43	45.34	43.52

Although phase c has the least power among the three phases for half of the time, its average power is not lower than that of phase b by more than the threshold δ_1. Therefore, phase c is not judged as the definite-min phase. The three-phase power series then belong to the RIB scenario which cannot be decomposed into SIB and RIB. In this case, the three-phase power series are the RIB component. The probability density functions of the three-phase power series are presented in Figure 3-14.

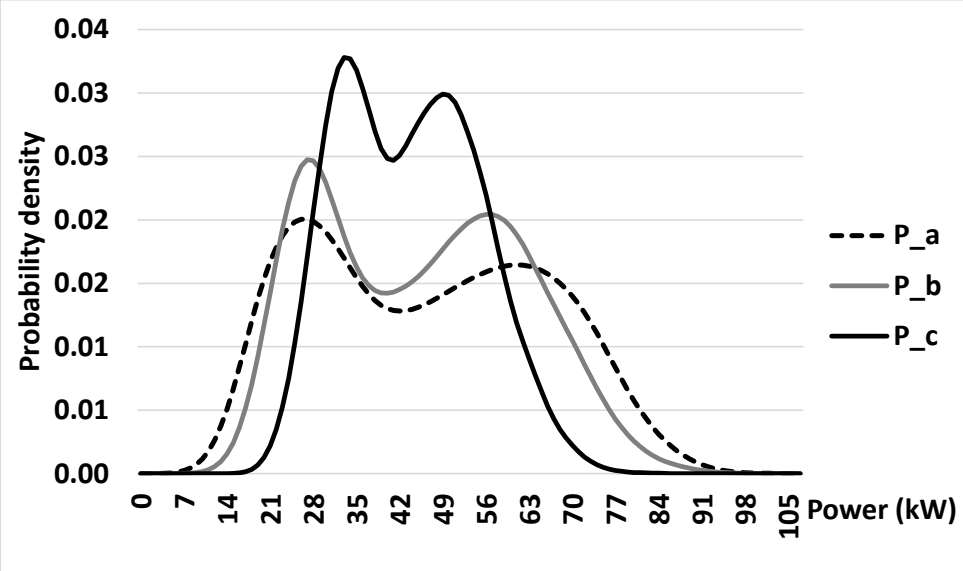


Figure 3-14. The probability density functions of the three-phase power series over a year for random imbalance scenario

E. Impact of Majority Threshold on Decomposition

According to Table 3-1, the majority threshold is the key for the a priori judgment. When the majority threshold is set as 50% (by default), 99.2% of the definite-max cases exhibit this feature: the average power of their definite-max phase is greater than those of the other two phases by more than 5% – this indicates that the majority threshold of 50% is reasonable because the order of the average power is consistent with the percentage of time when the definite-max occurs; similarly, 96.5% of the definite-order cases demonstrate the consistency between the order of the average power and the percentage of time when the order occurs; 97.8% of the definite-min cases demonstrate the consistency between the order of the average power and the percentage of time when the definite-min phase occurs. Therefore, the majority threshold of 50% is judged to be reasonable.

If, for example, the majority threshold is set as 60%, out of 782 substations, 14.07% (110) that were classified as definite-max, definite-order, and definite-min scenarios under the threshold of 50% are now classified as the random imbalance scenario under the new threshold of 60%. The threshold of 60% is not reasonable because those 14.07% (110) of substations actually have a definite-max phase, a definite-min phase, or both in terms of the average power, indicating the existence of systematic imbalance and the potential for phase swapping.

Table Table 3-7 presents the a priori judgment results (i.e, the number of substations belonging to each scenario) under different majority thresholds.

Table 3-7. A priori judgment results under different majority thresholds

Majority threshold	Numbers of definite-max, definite-order, definite-min cases and random imbalance cases, respectively
50%	235, 164, 170, 213
55%	220, 131, 169, 262
60%	205, 101, 153, 323
65%	191, 77, 144, 370

Table Table 3-7 shows that, with the increase of the majority threshold, the numbers of definite-max, definite-order, and definite-min cases all decrease but the number of random imbalance cases increases.

The impact of the majority threshold on three-phase power decomposition is derived from Table Table 3-7: each time the majority threshold increases by 5%, approximately 6.4% of the 'decomposable' cases (i.e., cases that can be decomposed into systematic imbalance and random imbalance) becomes 'non-decomposable' (i.e., belonging to the random imbalance scenario which cannot be decomposed).

However, as mentioned above, the increase of the majority threshold to over 50% masks the existence of systematic imbalance and the potential for phase swapping; the majority threshold of 50% is found to be reasonable.

F. Impact of Measurement Error Threshold on Decomposition

In Equations (3-2) and (3-3), there is a threshold δ_1 that accounts for measurement errors. How this threshold affects the a priori judgment results and consequently the decomposition is investigated in this section. Figure 3-15 depicts the numbers of substations belonging to the four scenarios under different measurement error thresholds.

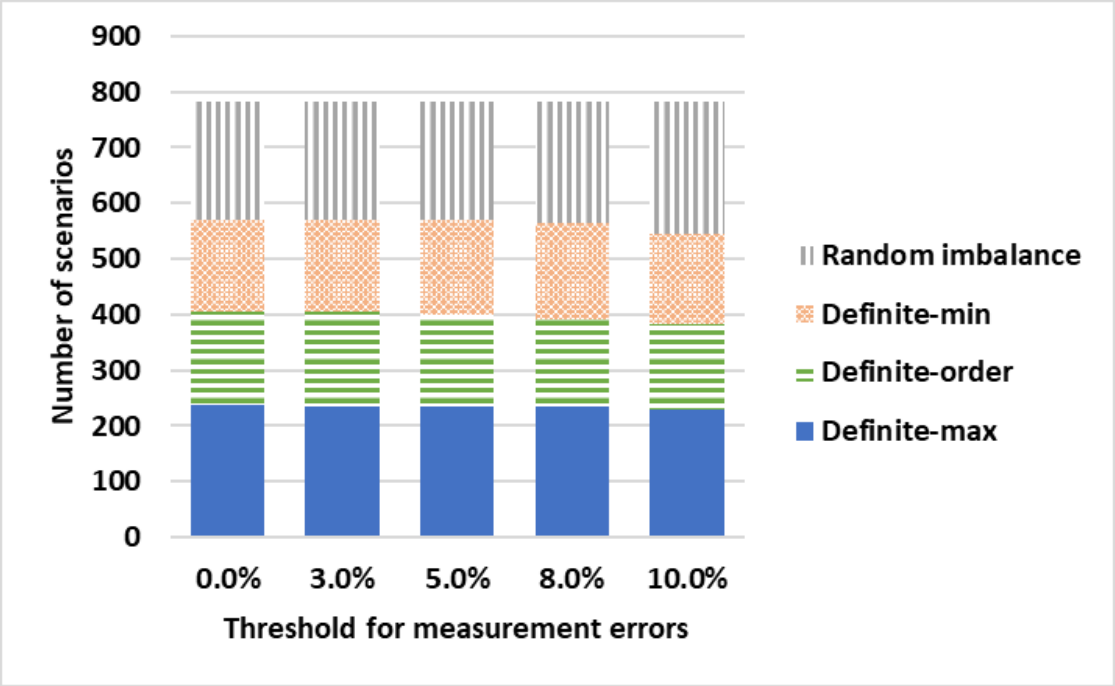


Figure 3-15. The impact of measurement error threshold on judgment results

Figure 3-15 shows that: 1) with the increase of the measurement error threshold from 0 to 10%, the number of random imbalance cases increases from 211 to 237, i.e., 26 more cases become 'non-decomposable' under the threshold of 10% as compared to that under the threshold of 0; 2) when the measurement error threshold is below 5%, the threshold has negligible impact on the a priori judgment results.

G. Validation by Phase Swapping

In this section, preliminary phase swapping is performed under the guidance of the degree of power imbalance to validate the methodology. Take Substation No. 536,753 (belonging to the definite-max scenario) as an example. Before phase swapping, its three-phase power series, the SIB component, the RIB component, and the degree of power imbalance are presented in Figure 3-2– Figure 3-5, respectively.

The degree of power imbalance results suggest that the distribution network operator move an average load of 8.61kW from phase a to the other two phases. Therefore, a preliminary phase swapping strategy is to move 10 single-phase domestic customers from phase a to phases b and c (5 customers to phase b and 5 customers to phase c). Suppose that the total load of these 10

customers follows a normal distribution with an average value of 8kW and a standard deviation of 3kW.

After phase swapping, the three-phase power series then belongs to the random imbalance scenario (the systematic component is zero). The RIB component equals the three-phase power series, the probability density functions of which are presented in Figure 3-16.

.

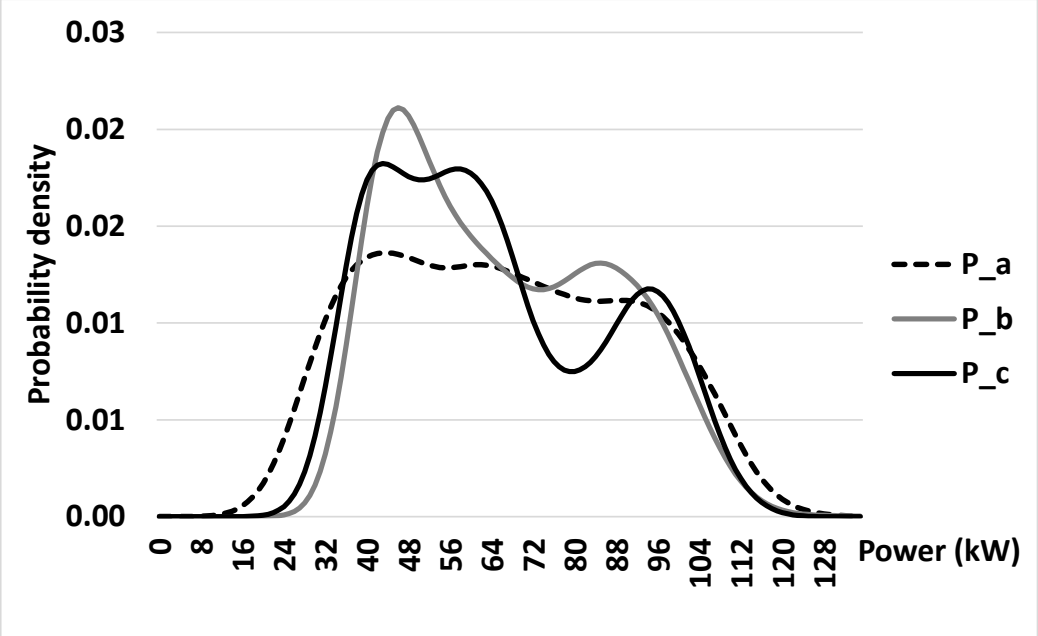


Figure 3-16. Three-phase power series after phase swapping

The phase swapping eliminates systematic imbalance: after phase swapping, there is no phase that exhibits the greatest power among the three phases for more than 50% of the time; neither is there any phase that exhibits the least power among the three phases for more than 50% of the time. Furthermore, after phase swapping, the average power for the three phases are 66.22kW, 66.64kW, and 65.04kW, respectively – the difference is below 2.5%. The remaining random imbalance requires demand-side response to address, if the three phases are to be fully rebalanced.

Based on the above results, it is concluded that the degree of power imbalance provides a useful guidance for phase swapping – this validates the methodology proposed by this paper.

H. Discussions

Among the 782 substations, 235 (30.1%), 164 (21.0%), 170 (21.7%), and 213 (27.2%) of them belong to the definite-max, definite-order, definite-min, and random imbalance scenarios, respectively. This indicates that the majority (72.8%) of the low voltage substations have SIB that can be addressed by phase swapping.

Three-phase power imbalance is directly associated with the costs for distribution network operators, including energy losses along the neutral wire and additional network investment costs. Although there is no regulatory limit on power imbalance, it will save costs for distribution network operators to rebalance three-phase power. A key value of this work is therefore to guide phase rebalancing practices. Furthermore, a few references [13], [15], [23] also focus on three-phase power imbalance.

It is necessary to judge the scenario according to the a priori judgment criteria before performing the decomposition. The necessity of the a priori judgment is because of the nature of three-phase power imbalance, i.e., the fact that any set of three-phase power series belong to one and only one of the four mutually exclusive scenarios. Once the scenario is determined, the three-phase power decomposition is definite as the closed-form solution to the quadratic optimization problem of the scenario.

Phase swapping (also known as rephasing) is a popular technique to rebalance three-phase supply in the medium-to-long term [9], [11], [25]. It requires scheduled outage, the time of which can be carefully chosen to minimize the impact on customers.

The degree of power imbalance based on the SIB component reveals the maximum potential (also the maximum need) of phase swapping. However, it does not mean that phase swapping will always meet the maximum need in practice. Rather, it is common for phase swapping to mitigate the SIB but not completely eliminate it – in this case demand-side managements will be required to resolve the residual SIB and the RIB, if the three phases are to be fully rebalanced; alternatively, phase swapping may deliver more than the maximum need by swapping too much, causing an overcompensation that requires demand-side managements to further rebalance the three phases.

The RIB component reveals the minimum need for demand-side managements. If phase swapping exactly meets the maximum need as indicated by the degree of power imbalance, then demand-side managements only need to reduce the loads equal to the RIB component for each phase. However, because of the imperfect phase swapping in practice (as explained above), the actual need for demand-side managements is likely to be greater than the RIB component.

It should be noted that reactive power also affects network loading and phase imbalance. A major obstacle to quantifying the time-varying power factor (hence the time-varying reactive power) in real-time operation is the lack of phasor measurements in distribution networks, especially in low voltage networks. Therefore, there is hardly any information on power factor (reactive power) in real-time operation. A feasible solution is to assume an average power factor: existing publications [9], [25], [11] on phase rebalancing considers load patterns represented by active power only, based on an implicit assumption of an average power factor. Reference [1] also assumes an average power factor when calculating loading levels. Reference [26] derived an average power factor of 0.9 for residential customers. Assuming such an average power factor, if the active power is rebalanced, the reactive power is automatically rebalanced. Hence, the decomposition method needs to consider three-phase active power only.

The three-phase power decomposition method proposed by this paper is not limited to substations, but is equally applicable to nodes along the feeder with three-phase power measurements. A major obstacle to understanding the phase imbalance along feeders beyond substations is the lack of monitoring along low voltage (415V) feeders. Only a selection of the UK's low voltage substations are monitored [27], because of cost barriers. Furthermore, existing publications [13], [15] focus on phase rebalancing at the substation side to prevent the imbalance from propagating to higher-level networks. Otherwise, three-phase power imbalance will further cause energy losses and increased investment costs in higher-level networks. This research therefore focuses on three-phase power imbalance at the substation side, using the available substation-side data provided by Western Power Distribution.

VIII. Conclusions

This paper identifies the systematic imbalance component and random imbalance component from any set of three-phase power series. The systematic component, as a direct consequence of uneven load allocations, can be addressed by phase swapping; the random imbalance component, as a result of random individual load fluctuations, requires demand-side managements, if the three phases are to be fully rebalanced. A new a priori judgment method is proposed to classify any set of three-phase power series into one of the four scenarios, i.e. definite-max, definite-order, definite-min, and random imbalance scenarios, by judging both the percentage of time and the average power to ensure robustness. For each scenario except the random imbalance one, a novel decomposition method is proposed to decompose three-phase power series into a systematic imbalance component and a random imbalance component, which are the closed-form solution to a quadratic

optimization problem that minimizes the random imbalance component. The degree of power imbalance is defined for each scenario based on the systematic imbalance component.

Case studies demonstrate that 30.1%, 21.0%, 21.7%, and 27.2% of 782 low voltage substations belong to the definite-max, definite-order, definite-min, and random imbalance scenarios, respectively. Decompositions are applied to the first three groups and the degree of power imbalance values are calculated based on the systematic imbalance component. The effectiveness of the degree of power imbalance as a guidance for phase swapping is validated by preliminary phase swapping.

The methodology is highly suitable for monitored low voltage distribution networks in the UK and the rest of Europe and monitored medium voltage distribution networks in the US. Distribution network operators can use the results to find out the maximum potential of phase swapping to address systematic imbalance and the minimum need for demand-side managements to address random imbalance, if the three phases are to be fully rebalanced. In addition, the degree of power imbalance not only reveals the underlying trend of systematic imbalance over time but also provides a guidance for phase swapping practices.

REFERENCES

[1] K. Ma, R. Li, and F. Li, "Utility-Scale Estimation of Additional Reinforcement Cost from 3-Phase Imbalance Considering Thermal Constraints," IEEE Transactions on Power Systems, vol. PP, pp. 1-1, 2016.

[2] (2010). Electricity distribution units and loss percentages summary. Available: https://www.ofgem.gov.uk/sites/default/files/docs/2010/08/distribution-units-and-loss-percentages-summary.pdf

[3] (2009). Electricity Distribution Price Control Review Final Proposals - Incentives and Obligations. Available: https://www.ofgem.gov.uk/ofgem-publications/46748/fp2incentives-and-obligations-final.pdf

[4] K. Ma, R. Li, and F. Li, "Quantification of Additional Asset Reinforcement Cost From 3-Phase Imbalance," IEEE Transactions on Power Systems, vol. 31, pp. 2885-2891, 2016.

[5] J. Zhu, M. Y. Chow, and F. Zhang, "Phase balancing using mixed-integer programming4" Power Systems, IEEE Transactions on, vol. 13, pp. 1487-1492, 1998.

[6] G. Mokryani, A. Majumdar, and B. C. Pal, "Probabilistic method for the operation of three-phase unbalanced active distribution networks," IET Renewable Power Generation, vol. 10, pp. 944-954, 2016.

[7] M. W. Siti, D. V. Nicolae, A. A. Jimoh, and A. Ukil, "Reconfiguration and Load Balancing in the LV and MV Distribution Networks for Optimal Performance," Power Delivery, IEEE Transactions on, vol. 22, pp. 2534-2540, 2007.

[8] S. Yan, S. C. Tan, C. K. Lee, B. Chaudhuri, and S. Y. R. Hui, "Electric Springs for Reducing Power Imbalance in Three-Phase Power Systems," IEEE Transactions on Power Electronics, vol. 30, pp. 3601-3609, 2015.

[9] C. H. Lin, C. S. Chen, H. J. Chuang, M. Y. Huang, and C. W. Huang, "An Expert System for Three-Phase Balancing of Distribution Feeders," IEEE Transactions on Power Systems, vol. 23, pp. 1488-1496, 2008.

[10] R. A. Hooshmand and S. Soltani, "Fuzzy Optimal Phase Balancing of Radial and Meshed Distribution Networks Using BF-PSO Algorithm," IEEE Transactions on Power Systems, vol. 27, pp. 47-57, 2012.

[11] M. Y. Huang, C. S. Chen, C. H. Lin, M. s. Kang, H. J. Chuang, and C. W. Huang, "Three-phase balancing of distribution feeders using immune algorithm," IET Generation, Transmission & Distribution, vol. 2, pp. 383-392, 2008.

[12] P. Lico, M. Marinelli, K. Knezović, and S. Grillo, "Phase balancing by means of electric vehicles single-phase connection shifting in a low voltage Danish grid," in 2015 50th International Universities Power Engineering Conference (UPEC), 2015, pp. 1-5.

[13] S. Sun, B. Liang, M. Dong, and J. A. Taylor, "Phase Balancing Using Energy Storage in Power Grids Under Uncertainty," IEEE Transactions on Power Systems, vol. 31, pp. 3891-3903, 2016.

[14] Random vs Systematic Error. Available:https://www.physics.umd.edu/courses/Phys276/Hill/Information/Notes/ErrorAnalysis.html

[15] S. Yan, S.-C. Tan, C.-K. Lee, B. Chaudhuri, and S. Y. R. Hui, "Electric Springs for Reducing Power Imbalance in Three-Phase Power Systems," IEEE Transactions on Power Electronics, vol. 30, pp. 3601-3609, 2015.

[16] N. C. Woolley and J. V. Milanovic, "Statistical Estimation of the Source and Level of Voltage Unbalance in Distribution Networks," Power Delivery, IEEE Transactions on, vol. 27, pp. 1450-1460, 2012.

[17] Z. Liu and J. V. Milanović, "Probabilistic Estimation of Voltage Unbalance in MV Distribution Networks With Unbalanced Load," IEEE Transactions on Power Delivery, vol. 30, pp. 693-703, 2015.

[18] W. Yaw-Juen and L. Pierrat, "A method integrating deterministic and stochastic approaches for the simulation of voltage unbalance in electric power distribution systems," IEEE Transactions on Power Systems, vol. 16, pp. 241-246, 2001.

[19] J. A. L. Ghijselen and A. P. M. V. d. Bossche, "Exact voltage unbalance assessment without phase measurements," IEEE Transactions on Power Systems, vol. 20, pp. 519-520, 2005.

[20] A. Kalyuzhny and G. Kushnir, "Analysis of Current Unbalance In Transmission Systems With Short Lines," IEEE Transactions on Power Delivery, vol. 22, pp. 1040-1048, 2007.

[21] D. Schwanz, F. Moller, S. K. Ronnberg, J. Meyer, and M. H. J. Bollen, "Stochastic Assessment of Voltage Unbalance due to Single-Phase-Connected Solar Power," IEEE Transactions on Power Delivery, vol. PP, pp. 1-1, 2016.

[22] G. Tan, J. Cheng, and X. Sun, "Tan-Sun Coordinate Transformation System Theory and Applications for Three-Phase Unbalanced Power Systems," IEEE Transactions on Power Electronics, vol. 32, pp. 7352-7380, 2017.

[23] V. B. Bhavaraju and P. N. Enjeti, "Analysis and design of an active power filter for balancing unbalanced loads," IEEE Transactions on Power Electronics, vol. 8, pp. 640-647, 1993.

[24] R. Li, C. Gu, F. Li, G. Shaddick, and M. Dale, "Development of Low Voltage Network Templates Part I: Substation Clustering and Classification," Power Systems, IEEE Transactions on, vol. 30, pp. 3036-3044, 2015.

[25] L. Chia-Hung, C. Chao-Shun, C. Hui-Jen, and H. Cheng-Yu, "Heuristic rule-based phase balancing of distribution systems by considering customer load patterns," IEEE Transactions on Power Systems, vol. 20, pp. 709-716, 2005.

[26] A. J. Collin, G. Tsagarakis, A. E. Kiprakis, and S. McLaughlin, "Development of Low-Voltage Load Models for the Residential Load Sector," Power Systems, IEEE Transactions on, vol. 29, pp. 2180-2188, 2014.

[27] M. Dale. (2013). LV Network Templates for a Low-Carbon Future. Available: https://www.westernpowerinnovation.co.uk/Document-library/2016/LVT-Closedown-Final.aspx

3.3. Additional Analysis and Discussions

For this thesis, only phase voltage and current data measured from distribution substations are available. Consequently, the voltage imbalance cannot be calculated as it requires the phase angle measurement data. Therefore, this thesis only focuses on phase power imbalance. Phase power imbalance means that the power magnitudes of the three phases are not equal to each other.

To further investigate the nature of phase imbalance, an additional analysis based on two different seasons, i.e., summer and winter, is presented in this subsection. The same substations as discussed in Section 3.2-VII are used for demonstration.

A. Definite-Max Scenario

A priori judgment process for both summer and winter is presented in Table 3-8. It can be seen that $A_{\emptyset}$ is larger than 50% for both the seasons which indicates that phase a is the definite-max phase for the majority of the time. Compared to summer, phase c shows 15.28% more time of having the least power. However, the $B_{\emptyset}$ for phase c in winter is still lower than 50%. Besides, the average power $\bar{P}_{\emptyset}$ of phase c is approximately the same as that of phase b in both summer and winter. Consequently, phase c is not judged as the definite-min phase and only the definite-max phase exists in this case.

Table 3-8. A priori judgment for the definite-max scenario

Sub No.	Season	Variables	Phase a	Phase b	Phase c
536,753	Summer	$A_{\emptyset}$	66.58%	7.15%	4.15%
		$B_{\emptyset}$	2.07%	34.17%	29.83%
		$\bar{P}_{\emptyset}$ (kW)	71.98	61.59	61.64
	Winter	$A_{\emptyset}$	85.19%	2.13%	0.61%
		$B_{\emptyset}$	0.29%	22.80%	45.11%
		$\bar{P}_{\emptyset}$ (kW)	77.93	62.36	60.18

The SIB components for summer and winter are presented in the form of probability density functions in Figure 3-17 and Figure 3-18. It can be seen that in the SIB component of phase a has the greatest power among the three phases. It means that too much load is allocated to phase a in both seasons.

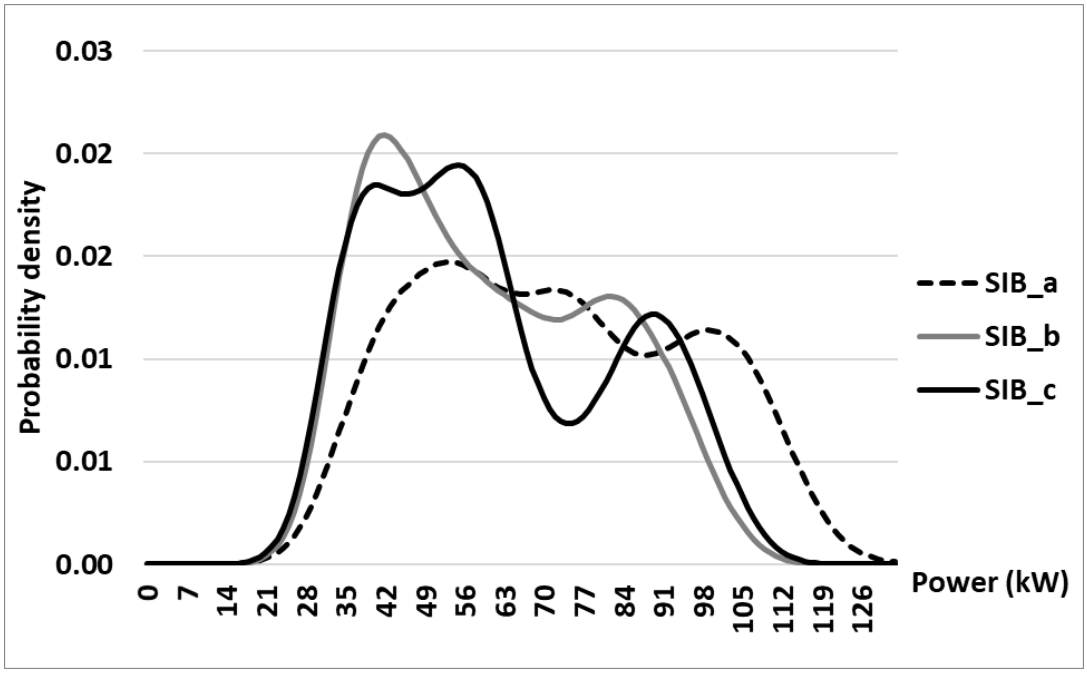


Figure 3-17. The probability density functions of the SIB component over summer for definite-max scenario

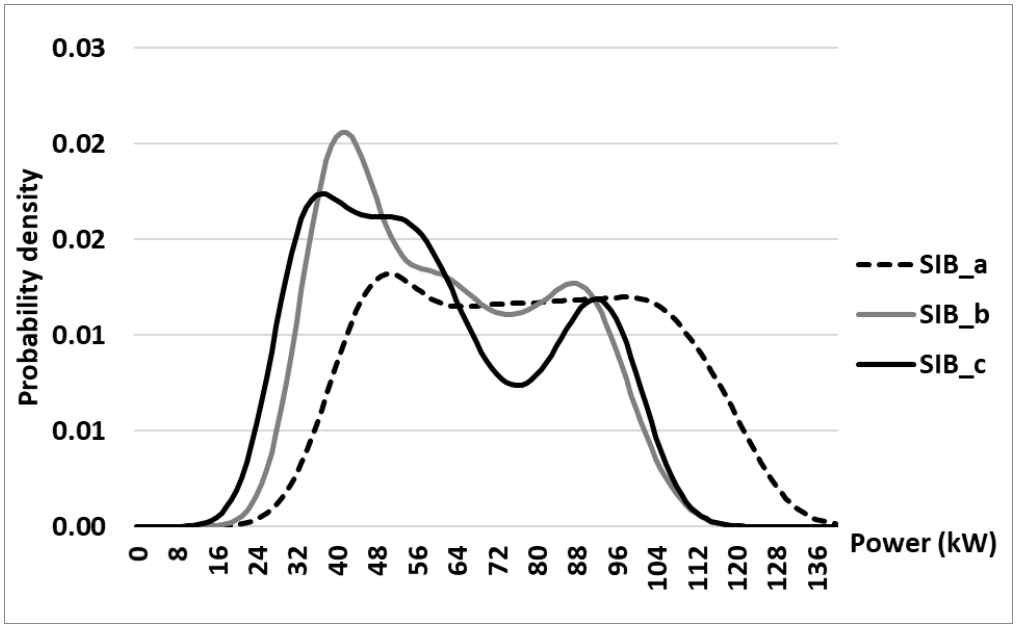


Figure 3-18. The probability density functions of the SIB component over winter for definite-max scenario

The results in Figure 3-19 shows the degree of power imbalance in winter is about 1.6 times larger than that of summer. This is because electricity demand is larger in the winter for the purpose of heating. Therefore, on average, up to 8.12kW of loads can be moved from phase a to phase b and

phase c in summer. In winter, up to 13.68kW of loads can be moved from phase a to phase b and phase c.

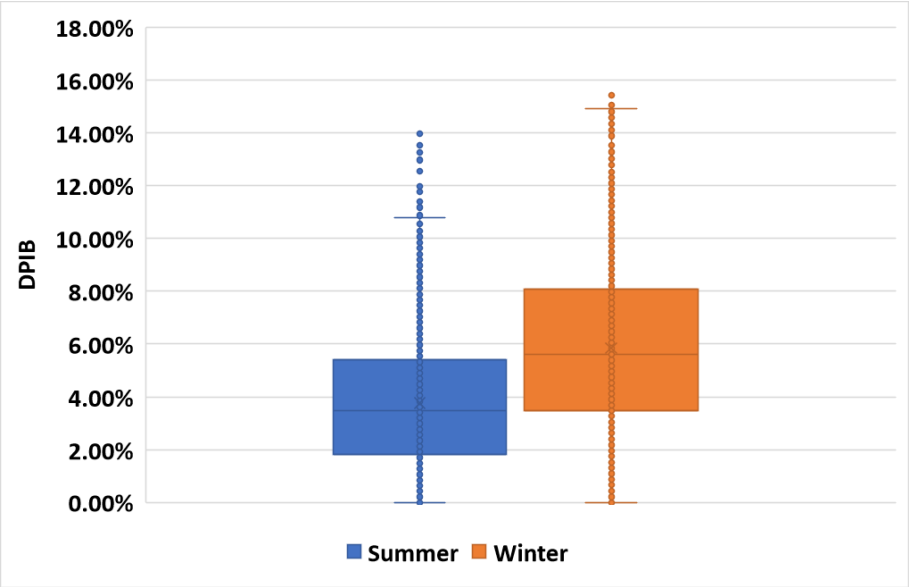


Figure 3-19. The degree of power imbalance over summer for definite-max scenario

B. Definite-Order Scenario

A priori judgment process for both summer and winter is presented in Table 3-9.. It can be seen that phase a has the largest power and phase c has the least power for the majority of time in both summer and winter. The average power also demonstrates the order of 'a > b > c'. As a result, phase a is the definite-max phase and phase c is the definite-min phase.

Table 3-9. A priori judgment for the definite-order scenario

Sub No.	Season	Variables	Phase a	Phase b	Phase c
512,457	Summer	$A_{\emptyset}$	73.74%	9.83%	0.05%
		$B_{\emptyset}$	0.05%	0.77%	96.78%
		$\overline{P}_{\emptyset}$ (kW)	42.18	36.26	23.89
	Winter	$A_{\emptyset}$	91.92%	1.99%	0.08%
		$B_{\emptyset}$	0.00%	0.98%	96.57%
		$\overline{P}_{\emptyset}$ (kW)	47.58	36.99	25.94

For both seasons, the probability density functions of the SIB component is shown in Figure 3-20 and Figure 3-21. According to the SIB component, phase a has the greatest power among the three phases for 100% of time; phase c has the least power among the three phases for 100% of time. This is consistent with the definition of the SIB for the definite-order scenario. It indicates that there exist excessive loads on phase a and insufficient loads on phase c for both seasons.

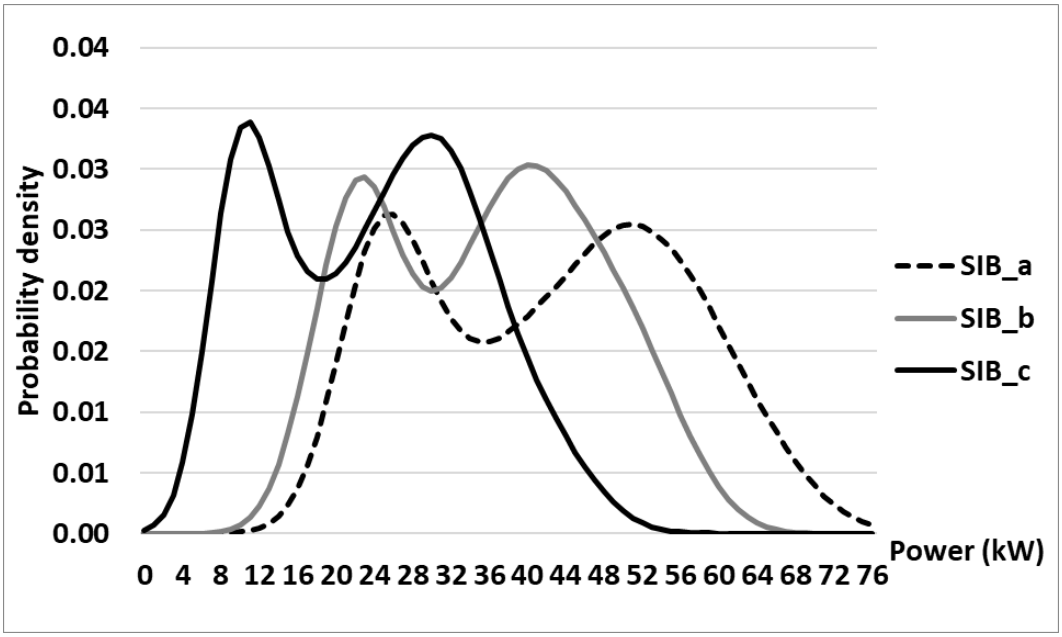


Figure 3-20. The probability density functions of the SIB component over summer for definite-order scenario

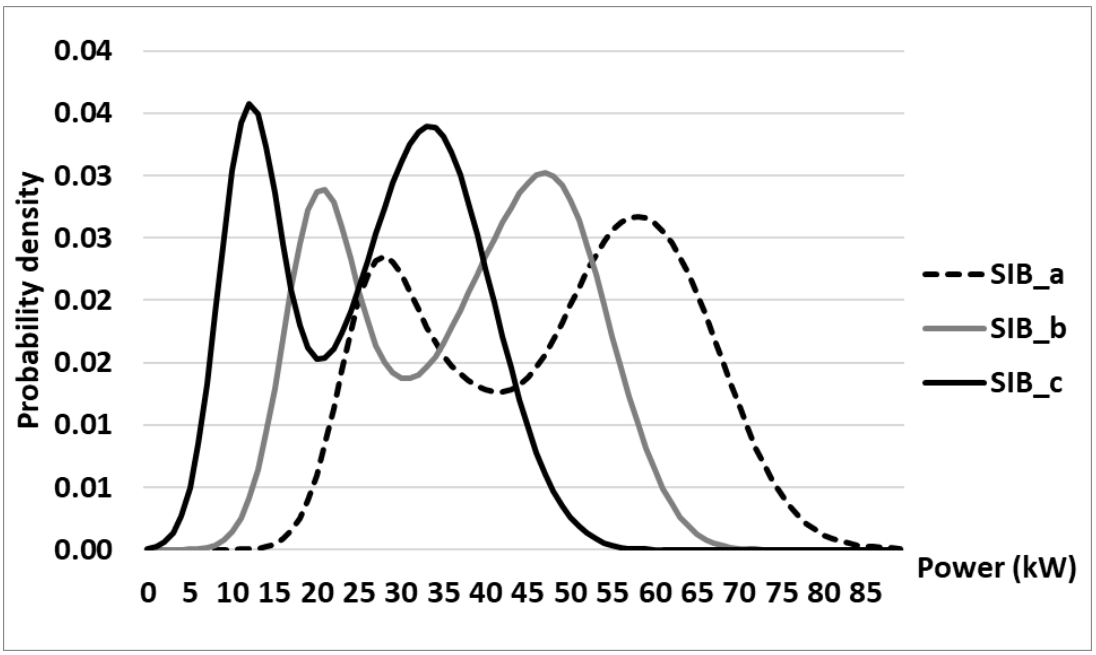


Figure 3-21. The probability density functions of the SIB component over winter for definite-order scenario

The results in Figure 3-22 shows that the two values of degree of power imbalance in summer are similar, i.e., both are approximately 10.0% on average. However, the two values of the degree of power imbalance in winter has a difference of 1.6%. Therefore, on average, up to 13.31kW of loads can be moved away from phase a and up to 7.62kW of loads can be moved to phase c in summer. In winter, up to 12.24kW of loads can be moved away from phase a and up to 8.41kW of loads can be moved to phase c.

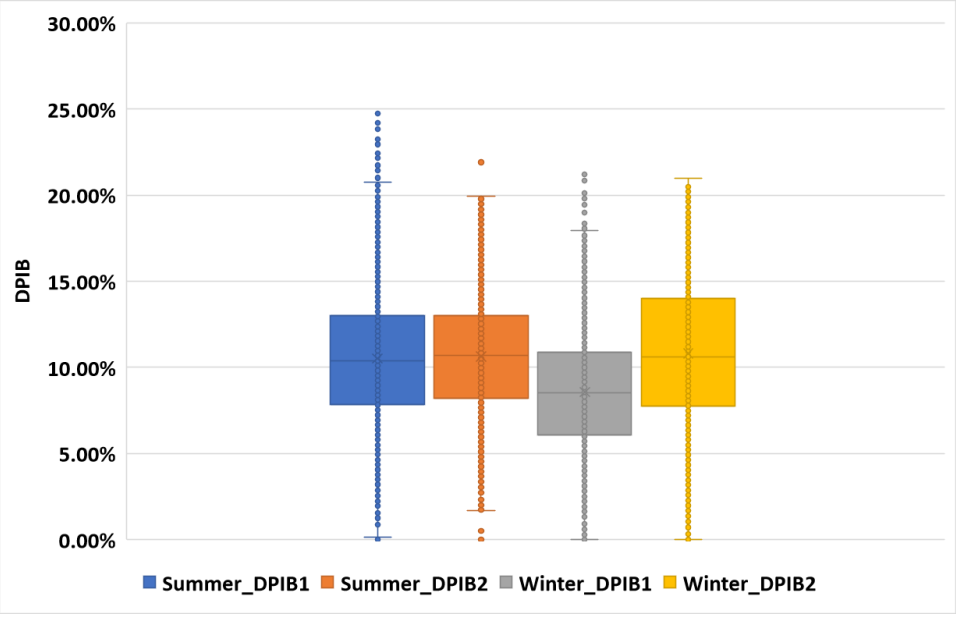


Figure 3-22. The degree of power imbalance over summer for definite-order scenario

C. Definite-Min Scenario

A priori judgment process is presented in Table 3-10. Although phase b has the largest power among and phase a has the least power among the majority of the time, the order of 'b>c>a' occurs less than 50% of the time. Consequently, it does not meet the criteria for the definite -order scenarios as indicated in the judgement process. Thus, only a definite-min phase (i.e., phase a) exist for both summer and winter.

Table 3-10. A priori judgment for the definite-min scenario

Sub No.	Season	Variables	Phase a	Phase b	Phase c
521,071	Summer	$A_{\emptyset}$	0.03%	67.35%	7.95%
		$B_{\emptyset}$	69.37%	0.00%	17.10%
		$\overline{P}_{\emptyset}$ (kW)	46.61	60.36	54.02
	Winter	$A_{\emptyset}$	0.00%	65.89%	17.90%
		$B_{\emptyset}$	70.73%	0.00%	11.98%
		$\overline{P}_{\emptyset}$ (kW)	45.99	60.85	55.30

For both seasons, the probability density functions of the SIB component is shown in Figure 3-23 and Figure 3-24. According to the SIB component, phase a has the least power among the three phases for 100% of the time, reflecting that more load can be moved to phase a from phase b and phase c.

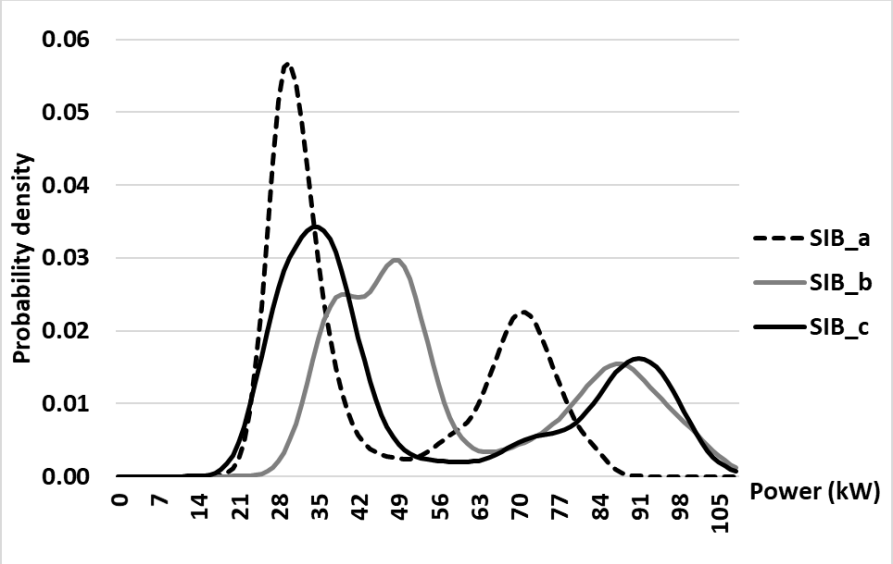


Figure 3-23. The probability density functions of the SIB component over summer for definite-min scenario

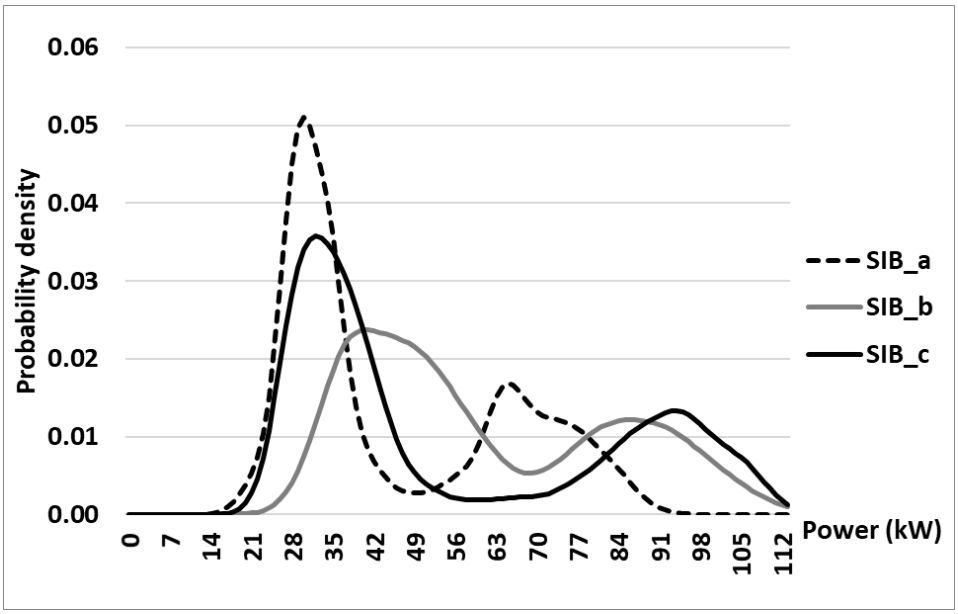


Figure 3-24. The probability density functions of the SIB component over winter for definite-min scenario

The results in Figure 3-25 shows that the degree of power imbalance in winter is larger than that of summer on average. This is because of more heating demand is required in winter. Therefore, in summer, up to 6.60kW of loads can be moved from phase b and phase c to phase a on average. In winter, up to 7.06kW of loads can be moved from phase b and phase c to phase a on average.

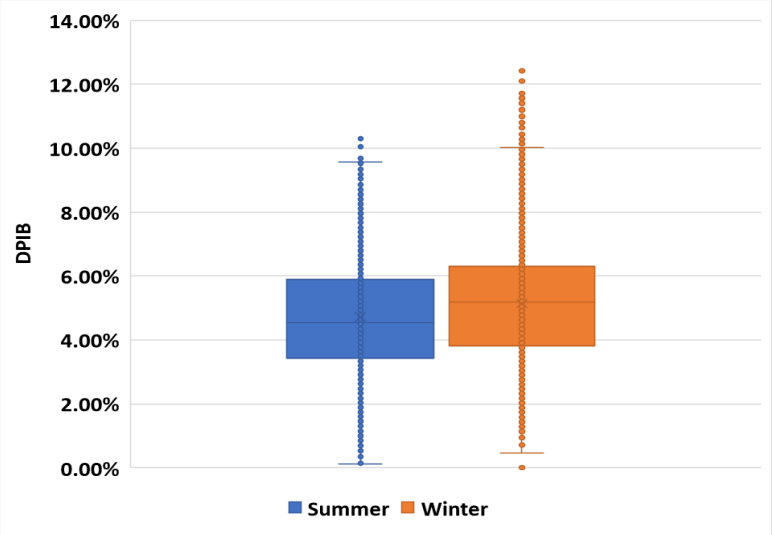


Figure 3-25. The degree of power imbalance over summer for definite-min scenario

D. Random Imbalance Scenario

A priori judgment process for summer and winter is presented in Table 3-11. It can be seen that the differences among the average powers of the three phases are within the threshold for both seasons. In other words, they are treated as equal to each other within each season. Thus, the three-phase power series belong to the random imbalance scenario which cannot be decomposed into SIB and RIB.

Table 3-11. A priori judgment for the random imbalance scenario

Sub No.	Season	Variables	Phase a	Phase b	Phase c
521,064	Summer	$A_{\emptyset}$	31.83%	13.88%	31.61%
		$B_{\emptyset}$	21.51%	20.71%	31.53%
		$\overline{P}_{\emptyset}$ (kW)	46.35	43.95	43.37
	Winter	$A_{\emptyset}$	24.98%	29.35%	21.28%
		$B_{\emptyset}$	28.15%	15.05%	33.85%
		$\overline{P}_{\emptyset}$ (kW)	45.88	46.44	43.21

Chapter 4. Cost-Benefit Analysis of Phase Balancing

This chapter proposes a new cost-benefit analysis framework to identify networks that are worth phase balancing for data-scarce LV networks.

4.1. Introduction

Existing phase balancing solutions are highly suitable for monitored LV distribution networks. However, the majority of UK's LV networks are either unmonitored or have minimal data, such as annual peak current or total energy consumption from the substations. The data-scarcity brings challenges for DNOs to analyse the costs and evaluate the benefits of phase balancing solutions.

Therefore, this chapter introduces a new cost-benefit analysis framework to identify data-scarce LV networks that are worth phase balancing. The developed approach supports the DNOs in deciding whether phase balancing is economically feasible and which phase balancing solution yields the greatest net benefit compared to alternatives.

The content of this chapter is cited from a published article in IEEE Transactions on Power System by the author [76]. This chapter is formed in an alternative-based format. All the indexes, figures, tables, equations and references are numbered independently.

The following sections are organised as follows: Section 4.2 presents the published paper, including the details of estimation of imbalance-induced costs, the net benefits calculations from adopting phase balancing solutions, the numerical results and discussions. Section 4.3 present additional analysis based on changing load growth rate.

4.2. Cost-Benefit Analysis of Phase Balancing Solution for Data-Scarce LV Networks

<table>
<tr><th colspan="4">This declaration concerns the article entitled:</th></tr>
<tr><td colspan="4">Cost-benefit analysis of phase balancing solution for data-scarce LV networks by cluster-wise Gaussian process regression</td></tr>
<tr><td colspan="4">Publication status (tick one)
Draft manuscript ☐ Submitted ☐ In review ☐ Accepted ☐ Published ✓</td></tr>
<tr><td>Publication details (reference)</td><td colspan="3">W. Kong, K. Ma, L. Fang, R. Wei and F. Li, "Cost-Benefit Analysis of Phase Balancing Solution for Data-Scarce LV Networks by Cluster-Wise Gaussian Process Regression," in IEEE Transactions on Power Systems, vol. 35, no. 4, pp. 3170-3180, July 2020, DOI: 10.1109/TPWRS.2020.2966601.</td></tr>
<tr><td colspan="4">Copyright status (tick the appropriate statement)
I hold the copyright for this material ☐ Copyright is retained by the publisher, but I have been given permission to replicate the material here ✓</td></tr>
<tr><td>Candidate's contribution to the paper (provide details, and also indicate as a percentage)</td><td colspan="3">The candidate contributed to / considerably contributed to / predominantly executed the…

Formulation of ideas:
• 90%
• The idea of performing cost-benefit analysis for phase imbalance is guided by Dr Kang Ma and Prof Furong Li.

Design of methodology:
• 90%
• The use of cluster-wised method is guided by Dr Kang Ma.

Experimental work:
• 100%

Presentation of data in journal format:
• 90%
• The changes of data presentation during review process are done with help from Luri Fang and Renjie Wei.</td></tr>
<tr><td>Statement from Candidate</td><td colspan="3">This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.</td></tr>
<tr><td>Signed</td><td>Wangwei Kong</td><td>Date</td><td>20/12/2020</td></tr>
</table>

Cost-benefit analysis of phase balancing solution for data-scarce LV networks by cluster-wise Gaussian process regression

Wangwei Kong, Kang Ma, *Member, IEEE*, Lurui Fang, Student Member, Renjie Wei, Student Member, and Furong Li, Senior Member, IEEE

***Abstract*—Phase imbalance widely exists in the UK's low voltage (415V, LV) distribution networks. The imbalances not only lead to insufficient use of LV network assets but also cause energy losses. They lead to hundreds of millions of British pounds each year in the UK. The cost-benefit analyses of phase balancing solutions remained an unresolved question for the majority of the LV networks. The main challenge is data-scarcity – these networks only have peak current and total energy consumption that are collected once a year. To perform a cost-benefit analysis of phase balancing for data-scarce LV networks, this paper develops a customized cluster-wise Gaussian process regression (CGPR) approach. The approach estimates the total cost of phase imbalance for any data-scarce LV network by extracting knowledge from a set of representative data-rich LV networks and extrapolating the knowledge to any data-scarce network. The imbalance-induced cost is then translated into the benefit from phase balancing and this is compared against the costs of phase balancing solutions, e.g. deploying phase balancers. The developed CGPR approach assists distribution network operators (DNOs) to evaluate the cost-benefit of phase balancing solutions for data-scarce networks without the need to invest in additional monitoring devices.**

Index Terms—cost-benefit analysis, Gaussian process regression, low voltage, phase balancing, phase imbalance, power distribution, three-phase system

I. INTRODUCTION

THREE-PHASE imbalance exists in the majority (>70%) of UK's low voltage (415V, LV) networks [1] because of the uneven load allocation and random load behavior [2], [3], [4]. Phase imbalance causes additional energy losses [5], [6] and extra network investment costs [7], [8]. The additional energy losses include losses caused by neutral line currents and imbalance-induced transformer copper losses. The additional network investment costs include the additional investments on both LV transformers and network feeders, because phase imbalance wastes network capacity.

Phase balancing solutions include phase swapping [9], [10], demand-side management [11] and deploying phase balancers based on power electronics [12]. To justify any phase balancing solution, it is important to perform a cost-benefit analysis of the solution before making any investment decision. However, up until now, no published work performs a cost-benefit analysis of phase balancing solutions for the majority of the UK's LV networks that only have a minimal amount of data, e.g. data collected only once a year. These networks are referred to as data-scarce LV networks.

A number of references investigate imbalance-induced energy loss, which is a key input for the cost-benefit analysis. Reference [6] improves the backward-forward sweeping method to calculate the power loss in an imbalanced distribution network. Reference [13] introduces an imbalance factor to evaluate line losses under the imbalanced situation. Reference [14] performs a loss analysis based on a power flow algorithm for imbalanced radial distribution networks. References [15] and [16] perform power loss analysis for PV penetrated systems with full data of the network topology, load and generation. Reference [17] developed a statistical approach as a combination of clustering, classification and range estimation to estimate imbalance-induced energy losses for data-scarce networks.

This paper addresses a different problem from [17]: Reference [17] estimates the imbalance-induced energy loss only, whereas this paper performs a cost-benefit analysis of any phase balancing solution on data-scarce networks. This paper significantly extends [17] by considering a comprehensive range of imbalance-induced costs, including the ARC, the imbalance-induced energy losses caused by neutral line currents, and the imbalance-induced transformer copper losses. Furthermore, this paper develops a completely different methodology from [17]: Reference [17] develops a combined approach of clustering, classification and range estimation, whereas this paper develops a regression methodology tailored for the cost-benefit analysis of phase balancing solutions.

This paper addresses a real need for the UK industries: to identify, among a mass population of LV networks, a subset of networks that are worth phase balancing, i.e. where the benefit from phase balancing outweighs its cost [18], [19]. However, existing solutions require full data from distribution networks. There is a gap in performing cost-benefit analyses of phase balancing on data-scarce LV networks. This paper directly addresses the industrial need by bridging the gap. This paper for the first time performs a cost-benefit analysis of phase balancing for any data-scarce LV network. To this end, this paper develops a new cost-benefit analysis framework for phase balancing on data-scarce LV networks. The core of the framework is a customized cluster-wise Gaussian process regression (CGPR) approach, which accounts for a full range of imbalance-induced costs. The approach estimates the total cost of phase imbalance for any data-scarce LV network by extracting

knowledge from a set of representative data-rich LV networks and extrapolating the knowledge to any data-scarce LV network. The imbalance-induced cost is then translated into the benefit from phase balancing and is compared against the costs of candidate phase balancing solutions, e.g. deploying phase balancers.

The CGPR approach supports the distribution network operators (DNOs) to perform cost-benefit analyses of phase balancing solutions on data-scarce LV networks. In this way, DNOs can decide whether phase balancing is economically feasible and which phase balancing solution yields the greatest net benefit compared to alternatives.

The remainder of the paper is organized as follows: Section II presents an overview of the methodology; Section III introduces the formulas for calculating imbalance-induced costs; Section IV presents the cost-benefit analysis framework, including the CGPR approach; Section V performs a case study and Section VI concludes the paper.

II. Overview of Methodology

To perform an accurate cost-benefit analysis of a phase balancing solution, full time-series of phase voltage and current data are required as the input data. However, these data are not available from the majority of UK's LV networks. In this paper, we have the time-series of phase current and voltage data of 800 representative data-rich LV networks throughout a year. These networks are located within the business area of a UK DNO and the data are the deliverables of the "Low Voltage Network Templates" project [20]. When conducting the trial project and collecting network data, Western Power Distribution specifically chose networks of a diverse and heterogeneous nature so that the dataset is representative. These 800 networks cover various customer types (domestic, commercial and industrial customers) and geographical areas (urban, suburban, and rural areas). For example, Cardiff contains a large number of commercial customers and load; Monmouthshire is a representative for the rural area [20].

Figure 4-1 presents an overview of the CGPR approach. The key to this approach is to evaluate the imbalance-induced cost (including the cost of additional energy losses and the cost of additional network investment) for data-scarce LV networks. The approach consists of three stages:

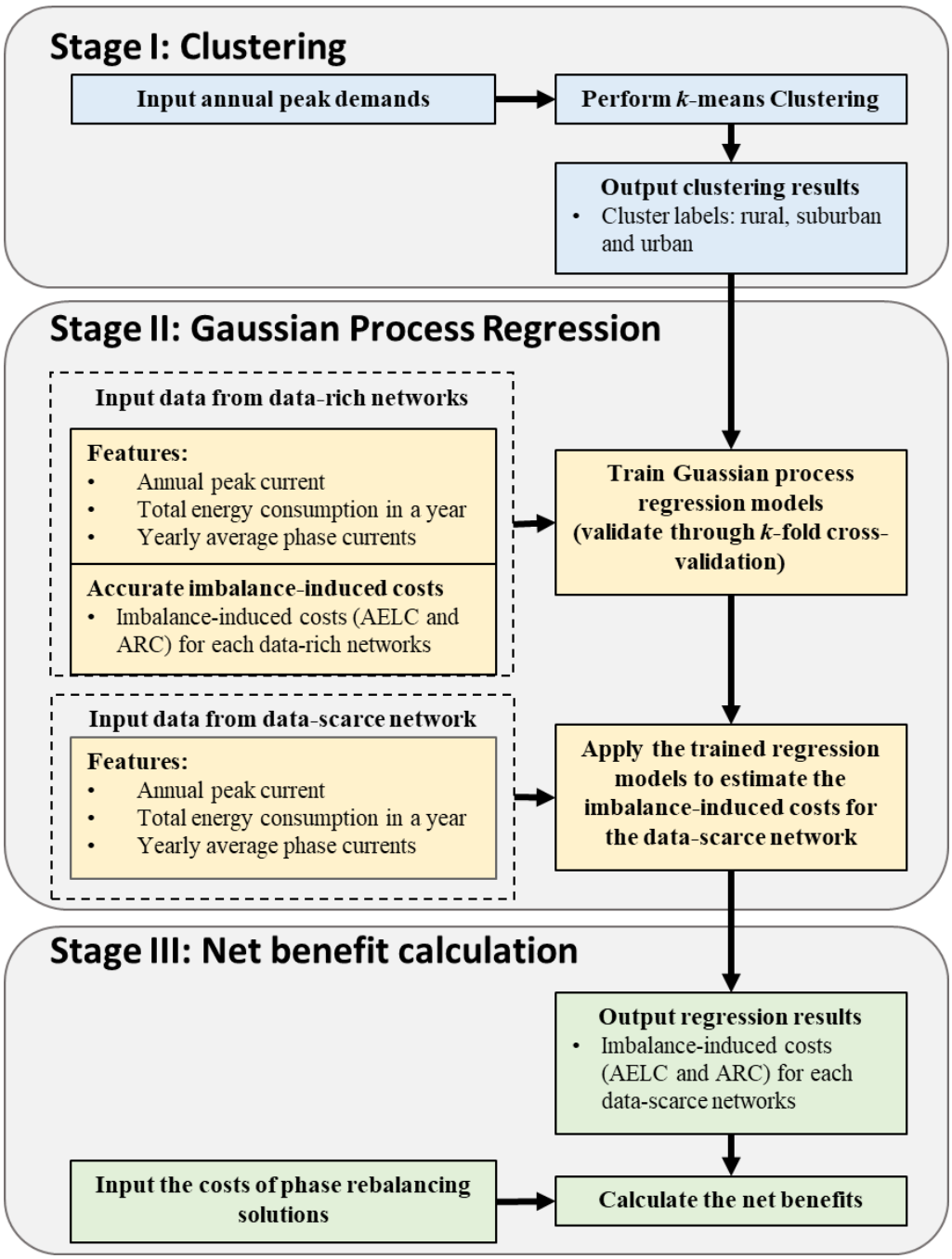


Figure 4-1. Overview of the CGPR approach

Stage I: The 800 data-rich networks are clustered into three groups, i.e., urban, suburban and rural, by applying the k-means clustering method.

Stage II: Input features are selected for regression and these features are available from data-scarce LV networks. Then, utilizing the data-rich LV networks, Gaussian process regression (GPR) models are trained for each cluster of the LV networks to model the relationship between the selected

features and the two imbalance-induced costs, i.e. the ARC and additional energy loss cost (AELC). The trained models are applied to data-scarce networks that only have the aforementioned features to estimate the imbalance-induced costs. An advantage of the approach is that it only requires features that are available from the majority of UK's data-scarce LV networks. Cross-validations are performed to validate the estimated imbalance-induced costs.

Stage III: The total imbalance-induced cost is calculated based on the estimations of the ARC and AELC. The imbalance-induced cost is then translated into the potential benefit from phase balancing. This benefit is compared to the cost of the phase balancing solution. This leads to a conclusion of whether the phase balancing solution is economically feasible or not as well as which phase balancing solution yields the greatest net benefit compared to alternatives.

III. Imbalance-Induced Cost for Individual Data-Rich Networks

This section presents the methods to calculate the components of the imbalance-induced cost for LV networks. The imbalance-induced cost consists of the ARC and the AELC. The AELC is broken down into the cost of energy losses caused by neutral line currents and the cost of transformer copper losses. The future cost is discounted back to form the present value. Then, the cost-benefit analysis is performed based on present values.

The present value of the ARC is detailed in [7], it represents the difference between the present value of reinforcement cost with phase imbalance and the present value of reinforcement cost without phase imbalance.

$$ARC = f_{PV}(DIB) \approx 3k_f DIB_f + k_t DIB_t$$

$$\text{Subject to } k_\chi = Asset_\chi \cdot (1+d)^{\frac{logU_N}{\log(1+r)}} \cdot \frac{\log(1+d)}{\log(1+r)} \qquad (4\text{-}1)$$

$$\chi \in \{f, t\}$$

where DIB_f and DIB_t are the degrees of phase imbalance for main feeders and LV transformers (%), respectively. The mathematical definitions of DIB_f and DIB_t are given by (4-3) and (4-4), respectively. $Asset_\chi$ is the future asset reinforcement cost (£), where subscript χ can be either f (feeder) or t (transformer); d is the discount rate (%); U_N is the asset utilization rate (%) and r is the load growth rate (%).

The factors U_N, DIB_f and DIB_t are given by (4-2), (4-3) and (4-4), respectively [7].

$$U_N = \frac{3 \cdot max\{P_\emptyset\}}{C_{asset}} \qquad \emptyset \in \{A, B, C\} \tag{4-2}$$

where $P_\emptyset$ is the power on phase ∅ (kW) and C_{asset} is the asset capacity (kVA).

$$DIB_f = \frac{max\{P_\emptyset\} - P_t}{P_t} \qquad \emptyset \in \{A, B, C\} \tag{4-3}$$

where P_t is the total power of three phases (kW) when the maximum phase power occurs. $P_\emptyset$ is defined in (4-2).

$$DIB_t = \frac{P_N}{P_t} \tag{4-4}$$

where P_N is neutral line power (kW). P_t is defined in (4-3).

A. Imbalance-induced energy loss

The imbalance-induced energy loss contains two components: the energy loss caused by a neutral line current [17] and the transformer copper loss.

1) Energy loss caused by neutral current

The energy loss caused by neutral line current is calculated considering different earthing systems [21], e.g., Terre-Neutral-Combined (TN-C) and Terre-Neutral (TN-S) systems [22]. The majority of the UK's LV distribution networks follow the TN-S earthing system [22]. Therefore, this paper considers the TN-S earthing system.

The estimation of energy loss caused by the neutral current is given in [17]

$$E_{loss} \approx \sum_{t=1}^{N_t} {I_{nlc}}^2(t) \cdot R_n \cdot \Delta t \tag{4-5}$$

where $I_{nlc}(t) = [I_A^2(t) + I_B^2(t) + I_C^2(t) - I_A(t)I_B(t) - I_B(t)I_C(t) - I_A(t)I_C(t)]^{\frac{1}{2}}$

where $I_A(t), I_B(t) and\ I_C(t)$ are current values (A) for the phases A, B and C at time t, respectively; $I_{nlc}(t)$ denotes the neutral line current (A) at time t; R_n denotes the neutral wire resistance (Ω). N_t is the number of hours within the year.

The neutral line energy loss for the Nth year is

$$E_{loss_N} = E_{loss} \cdot (1 + r)^{2(N-1)} \tag{4-6}$$

where N represent the Nth year; r is defined in (4-1); and E_{loss} is defined in (4-5).

2) Transformer copper loss cost

Phase imbalance increases the transformer copper loss beyond that under the phase balanced scenario. The transformer copper loss under the balanced case is given in [23]:

$$E_{trans} = 3\sum_{t=1}^{N_t} I^2(t) \cdot R_w \cdot \Delta t$$

$$\text{where } I = \frac{1}{3}(I_A + I_B + I_C) \tag{4-7}$$

$$I_A = I_B = I_C$$

where $I(t)$ is the balanced phase current (A) at time t and R_w is the resistance of the transformer winding (Ω); N_t is the number of hours within a year.

The transformer copper loss under the imbalanced case is also given in [23]

$$E_i = \sum_{t=1}^{N_t} \left({I_A}^2(t) + {I_B}^2(t) + {I_C}^2(t) \right) \cdot R_w \cdot \Delta t \tag{4-8}$$

where $I_A(t), I_B(t) and\ I_C(t)$ are current values for the phases A, B and C at time t, respectively; R_w and N_t are defined in (4-7).

As a result, the imbalance-induced transformer copper loss is:

$$E_{t_i} = E_i - E_{trans} \tag{4-9}$$

where all variables are defined in (4-7) and (4-8).

The transformer copper loss for the Nth year is

$$E_{t_{iN}} = E_{t_i} \cdot (1 + r)^{2(N-1)} \tag{4-10}$$

where r is the load growth rate (%); all other variables are defined in (4-7), (4-8) and (4-9).

B. The present value of the total imbalance-induced cost

As stated above, the total additional energy loss is the sum of losses caused by neutral line current and transformer copper. Therefore, the total imbalance-induced energy loss in year N is given by

$$E_{tot_N} = E_{loss_N} + E_{t_{iN}} \tag{4-11}$$

where E_{loss_N} and $E_{t_{iN}}$ are defined in (4-6) and (4-10), respectively.

The total AELC of the Nth year is transferred to the present value.

$$AELC = f_{PV}\left(E_{tot_N}\right) = \frac{E_{tot_N} \cdot \pi}{(1 + d)^N} \tag{4-12}$$

where π is the energy price (£); d is the discount rate (%); and E_{tot_N} is defined in (4-11).

The imbalance-induced energy losses incur costs every year until the three phases are rebalanced. In contrast, the ARC is a one-off investment when the asset capacity is reached. Therefore, the present value of the total imbalance-induced cost is given by

$$f_{PV_N} = f_{PV}(DIB) + \sum_{n=1}^{N} f_{PV}\left(E_{tot_n}\right) \tag{4-13}$$

where the function $f_{PV}(DIB)$ is defined in (4-1); the function $f_{PV}\left(E_{tot_n}\right)$ is defined in (4-12).

In this paper, the present value of the total imbalance-induced cost is referred to as the imbalance-induced cost for simplicity.

IV. Methodology

A. Clustering

In this section, a CGPR approach is presented as a combination of clustering and a Gaussian process regression (GPR). As mentioned in the previous section, the imbalance-induced cost includes two parts: ARC and AELC. Figure 4-2 shows the relation between the annual peak currents and the ARCs for the 800 LV networks. It can be seen that three distinctive relationships exist. The clustering is based on transformer capacities of each LV distribution substation [37].

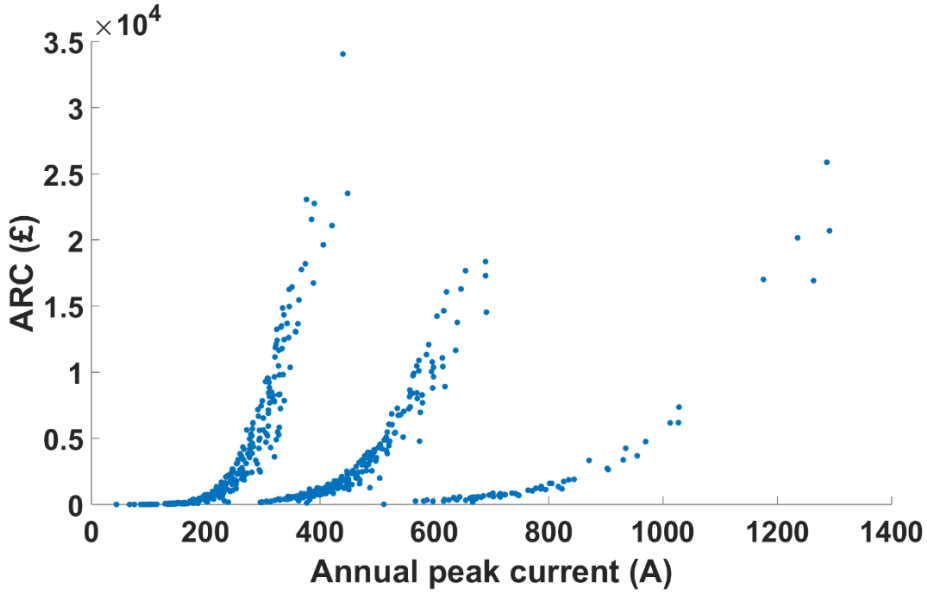


Figure 4-2. The relationship between annual peak current and ARC

The underlying reason is that the ARCs are strongly correlated to the type of the LV networks, i.e. urban, suburban, and rural types. The three different relationships justify the development of a

cluster-wise regression as opposed to a simple regression. Cluster-wise regression is an effective way of addressing problems with multiple regression models [24], [25].

As shown in Figure 4-1, k-means clustering is used to cluster the networks into 3 groups (rural, suburban and urban) by the annual peak demands. This corresponds to Stage I in Figure 4-1. The direct output of the clustering is which cluster each LV network belongs to (i.e. the cluster label for each LV network). From the outputs, it is straightforward to derive the range of annual peak currents for each cluster of the LV networks. In this way, given any LV network, determine which range its annual peak current falls into. This reveals the cluster to which the network in question belongs, i.e. whether the network is an urban, suburban, or rural one.

B. Gaussian process regression

The output of Stage I is used to train Gaussian process regression (GPR) models to model the relationship between the selected features and the imbalance-induced costs (i.e., AELC and ARC). The imbalance-induced costs are calculated using data from data-rich networks.

Then, the networks are treated as data-scarce networks and the selected features are used as the input to the trained GPR models. The GPR models output estimated imbalance-induced costs.

The regression process consists of the following steps:

1) Feature selection

Typical LV distribution substation monitoring is a Maximum Demand Indicator (MDI) which records peak phase currents based on the aggregation over half an hour. The MDI does not have any communication options so that its reading is typically manually recorded on an annual basis [43].

For the majority of the UK's LV networks, the annual peak current ($\hat{I}$) and annual total energy consumption (E_{total}) are readily available. According to [75], the average phase current values can be obtained with minimal efforts from either the per-phase energy meters or the protection system for data-scarce networks. The average phase current values are transformed into a virtual neutral line current:

$$\bar{I}_{nlc} = \sqrt{\bar{I}_a{}^2 + \bar{I}_b{}^2 + \bar{I}_c{}^2 - \bar{I}_a\bar{I}_b - \bar{I}_b\bar{I}_c - \bar{I}_a\bar{I}_c} \tag{4-14}$$

where $\bar{I}_a$, $\bar{I}_b$ and $\bar{I}_c$ are the yearly average phase current values for phases A, B, and C, respectively.

Two input feature vectors are defined to suit different levels of data availability in data-scarce networks. The first feature vector (v_{f1}) contains two features ($\hat{I}$ and E_{total}):

$$v_{f1} = \left[\hat{I}\,, E_{total}\right] \tag{4-15}$$

This feature vector is applicable in the absence of the average phase current values. The second feature vector (v_{f2}) contains three features ($\hat{I}$, E_{total} and $\bar{I}_{prc}$):

$$v_{f2} = \left[\hat{I}, E_{total}, \bar{I}_{nlc}\right] \tag{4-16}$$

This feature vector requires that the data-scarce network have the average phase current data.

2) Gaussian Process Regression model training

In this step, regression models are trained for each cluster of LV networks. The regression models map the feature vectors defined in step 1) to the ARC and AELC (the ARC and AELC are calculated in Section III) separately. In this paper, the Gaussian process regression (GPR) is adopted. The reasons why the GPR is adopted are: 1) Gaussian process models allow the quantification of uncertainty, considering both intrinsic noises in the problem and parameter errors in estimation [26]; 2) the case studies confirm that the GPR achieves the best performance among classical regression models.

Take the GPR that maps the feature vectors to the ARC as an example. The GPR model is given by

$$p\left(ARC_* \mid ARC, v_f, v_{f_*}\right) \sim \mathcal{N}(\mu^*, \Sigma^*)$$

$$\text{where} \quad \mu^* = K\left(v_f, v_f\right)\left(K\left(v_f, v_f\right) + \sigma^2 I\right)^{-1} ARC \tag{4-17}$$

$$\Sigma^* = K\left(v_{f_*}, v_{f_*}\right) + \sigma^2 I - K\left(v_{f_*}, v_f\right)\left(K\left(v_f, v_f\right) + \sigma^2 I\right)^{-1} K\left(v_f, v_{f_*}\right)$$

where $p\left(ARC_* \mid ARC, v_f, v_{f_*}\right)$ is the probability distribution for ARC estimation; v_f and ARC are the feature vector and the ARC for the data-rich networks, respectively; v_{f_*} and ARC_* are the feature vector and the predicted ARC for the data-scarce network, respectively; the ARC is given by (1);

$\mathcal{N}(\mu^*, \Sigma^*)$ denotes a Gaussian distribution with the mean μ^* and covariance Σ^*; K is a kernel matrix given by the squared exponential kernel function [26]; σ^2 is the noise variance; and I is the identity matrix. The feature vector v_f could be v_{f1} and v_{f2} as given in (4-15) and (4-16), depending on the choice of features.

The GPR is detailed in [26]. The above GPR model is developed for each cluster of the LV networks. The GPR is detailed in [26]. The GPR model for the AELC estimation is the same as that for the ARC estimation as shown in **Error! Reference source not found.**, except that the ARC is replaced b y the AELC. The results are compared with linear regression, which is detailed in [27] and which is not repeated in this paper.

C. Cross-validation

The CGPR approach is validated through *k*-fold cross-validation. This is a popular validation method as explained in [28]. The cross-validation is detailed as follows: the full dataset of 800 data-rich LV networks, including the features and the accurate ARC and AELC results, are randomly separated into *k* (*k*=10 in this paper) equal-sized groups. In each iteration of the *k*-fold cross-validation, one group of the LV networks are reserved as the validation set, whereas the remaining nine groups serve as the training set. The CGPR model is trained using the training set only. Then, the trained CGPR model predicts the imbalance-induced costs on the validation set, which are treated as if they were data-scarce. The outputs are estimated imbalance-induced costs for the LV networks in the validation set. These results are compared against the accurate imbalance-induced cost (the calculated costs from data-rich networks) results so that the CGPR model is validated. Each group is selected as the validation set once and there are ten iterations. It should be emphasized that throughout the process, the validation set and the training set are strictly separated from each other and the validation set is not used for training. The *k*-fold cross-validation is detailed in Figure 4-3.

D. Removal of outliers

Following the cross-validation, 11% of the networks are identified as the outliers and are removed. This percentage is derived by using Chebyshev's inequality. Chebyshev's inequality is a widely adopted method for removing outliers [29]. When the distribution of the data is unknown, the Chebyshev's inequality is given by:

$$P(|X-\mu| \leq k\sigma) \geq 1-\frac{1}{k^2} \tag{4-18}$$

where X is the set of sample data, μ is the mean of the sample data, σ is the standard deviation and k is a factor.

It is common practice to regard data samples that occur beyond 3σ (i.e., $k=3$) from the mean as outliers [30], [31]. Therefore, the outliers account for approximately 11% of the whole population of networks. Note that outliers are an objective existence and they can be identified and removed from consideration for better performance. In this paper, all the LV networks are clustered into three groups, i.e., urban, suburban and rural. However, there are some networks have distinctive characteristics that does not fall into any of the groups. These networks are identified as outliers and further investigations are required to analyse these networks.

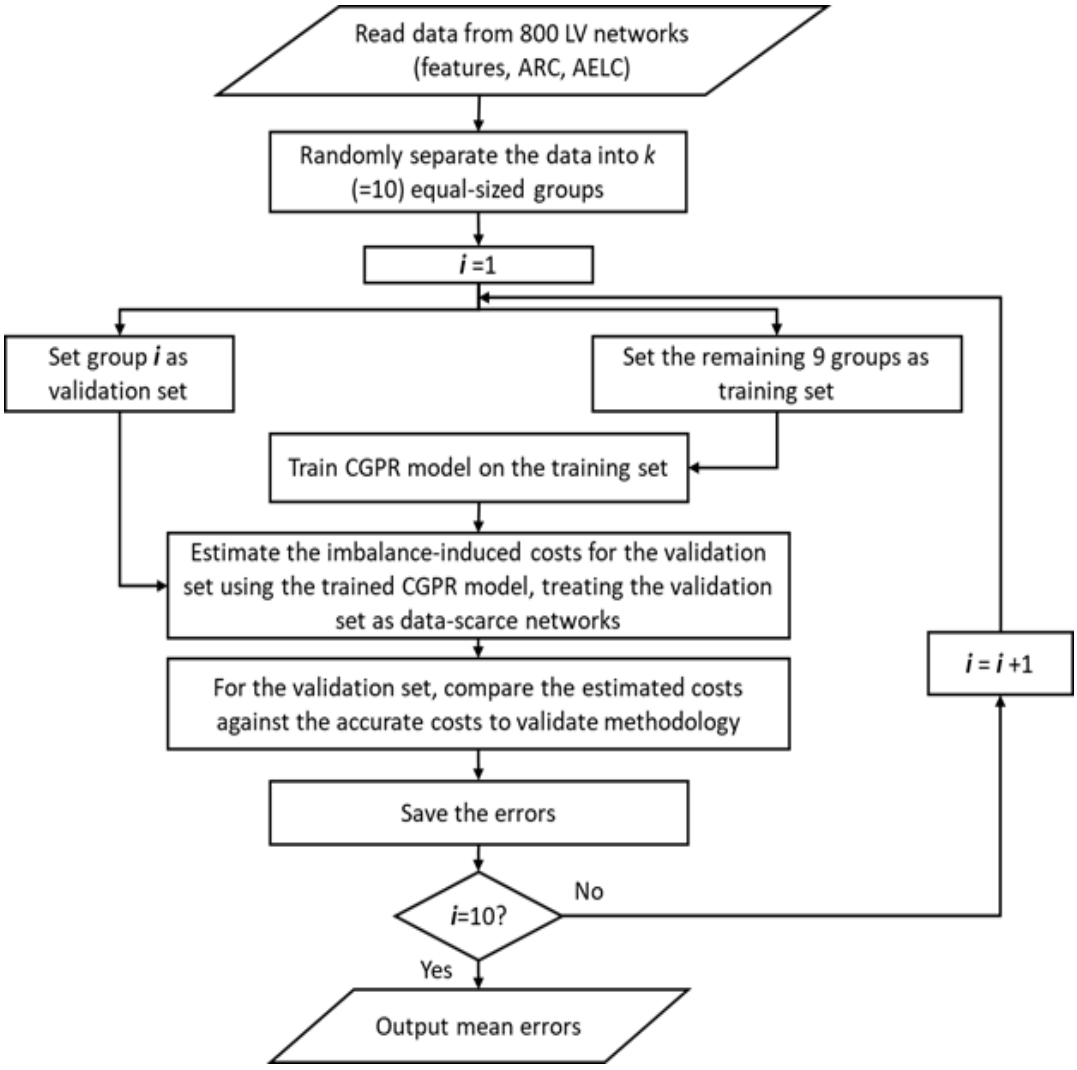


Figure 4-3. The flow chart of k-fold cross-validation

E. Net benefit calculation

The trained CGPR model takes the features of any given data-scarce network as the input and outputs the estimated imbalance-induced cost.

Note that the phase balancing solutions may not be able to fully rebalance the three phases. Therefore, the benefit from phase balancing is given by the difference of the total imbalance-induced costs before and after phase balancing

$$B_{PV_N}^{ds} \approx f_{PV_N}^{before} - f_{PV_N}^{after} \tag{4-19}$$

where $f_{PV_N}^{before}$ and $f_{PV_N}^{after}$ are the estimated total imbalance-induced cost before and after phase balancing, respectively; the superscript ds means data-scarce; the subscript PV_N represents present value for N years.

Then, the benefit is compared with the cost of the phase balancing solution to determine whether it is beneficial to apply the phase balancing solution in question. Hence, the net benefit of applying the phase balancing solution is given by

$$B^{ds} \approx B_{PV_N}^{ds} - f_{pb} \tag{4-20}$$

where $B_{PV_N}^{ds}$ is the total benefit of phase balancing for the data-scarce networks; f_{pb} is the cost of applying a phase balancing solution.

Note that the net benefit B^{ds} can be negative, which means that it is not economically feasible to deploy the phase balancing solution.

V. Case Studies

This section presents case studies. The input data are shown in Section V-A. The results from the cluster-wise regression model are presented in Section V-B. Section V-C gives the discussions. Section V-D gives the cost-benefits analysis for two phase balancers (ZM-SPC [32] and EQU18 [33])

and active network management (ANM) scheme, respectively. The case study is based on the time-series phase current and phase voltage data from the 800 data-rich LV networks throughout a year.

A. Imbalance-induced cost for data-rich networks

This sub-section presents the calculation of imbalance-induced cost for data-rich networks. To derive the additional energy losses (defined **Error! Reference source not found.** - **Error! Reference source not found.**), the neutral wire resistance (R_n) is set as 0.244 Ω/km [17]. The winding resistances (R_w) are calculated from [34] and presented in Table 4-1.

To derive the additional reinforcement costs (defined in (4-1) – (4-4)), the investment costs of the feeder and transformer are given in Table 4-1. The discount value (d) is set as 5.0% [1] and [35]. The load growth rate (r) is set as 0.82% [36].

Table 4-1. Parameters for different areas [34], [37]

Assets Area	Urban	Suburban	Rural
Transformer investment cost (k£)	26.4	16.1	5.8
Main feeder investment cost (k£/km)	67.2	16.4	15.0
Main feeder length (km)	0.2	0.3	0.4
No. of feeders connected from transformers	5	3.5	1.5
Winding resistance (Ω)	0.0163	0.0265	0.0413

A limitation of the work is that this paper assumes that the phase currents are 120° apart from each other. This is because there is hardly any LV network that has phasor measurements, as distribution network operators cannot justify the investment in phasor measurements in terms of the return on investment. Therefore, it is valid to assume that the phase currents are 120° apart from each other while phasor measurements are absent.

The neutral line current is the minimum, under the assumption that the phase currents are 120° apart from each other. Therefore, this assumption corresponds to a conservative cost-benefit analysis. If the actual phase currents are not 120° apart, the neutral line current will increase, so will the imbalance-induced energy losses and the associated cost. This means that the potential benefit from phase balancing will also increase, hence the net benefit will increase.

In this paper, a power factor of 0.9 is assumed and the harmonic distortion is not considered. The harmonic distortion results in the decrease of power factor and eventually increases the ARC. Besides, the harmonic currents cause additional energy losses which lead to higher AELC. Therefore, it shows that the estimation of the imbalance-induced costs is conservative, resulting in conservative net benefits, i.e. the lower bounds of the net benefits. The actual net benefits can be higher than the estimated value.

Figure 4-4 shows the present values of AELC and ARC for urban, suburban and rural networks. The ARC is the present value of the future cost while the AELC is calculated for 10 years for each of the network. The average AELC is approximately twice as much as the average ARC. The rural networks correspond to the least AELC and the greatest ARC among all three types of networks. In contrast, the urban networks correspond to the greatest AELC and the least ARC.

The reason for this is that the rural networks have the largest DIB (degree of imbalance) values, which causes the greatest ARC, in both LV transformers and main feeders among the three types of networks. However, the rural networks have the lowest loading levels, which lead to the lowest energy losses on the neutral lines and LV transformers. As a result, the rural networks have the largest average ARC but least average AELC. On the contrary, urban networks have the lowest DIB, which leads to the lowest ARC. They have the highest energy loss because of their high loading levels. Therefore, the urban network has the least AELC but largest ARC.

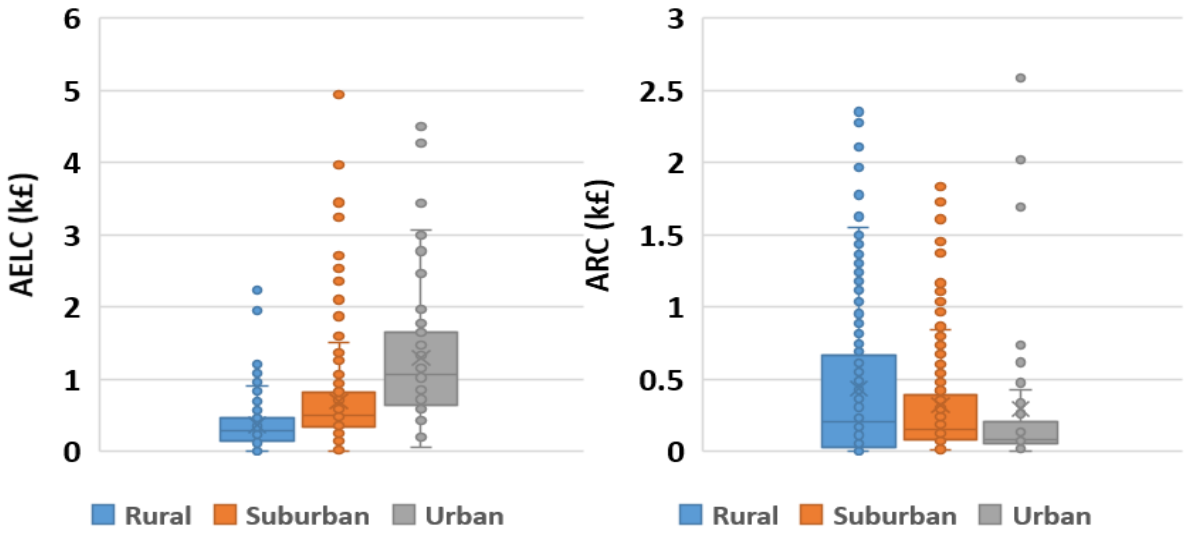


Figure 4-4. The AELC and ARC for the 800 LV networks

B. Cluster-wise Gaussian Process Regression

In this section, the CGPR results are shown, where the cost-benefit analyses are performed over a time horizon of 10 years.

The ARC and AELC estimation are calculated using four regression methods: linear regression (LR), cluster-wise LR (CLR), GPR and CGPR. Results from all methods are validated through 10-fold cross-validations. The results obtained by applying these four regression methods are compared with each other in terms of the root mean squared error (RMSE). The reference cost for comparison is calculated using the time-series voltage and current data of the networks. As mentioned in Section IV-A, two feature vectors are used as input, the first vector v_{f1} contains two features ($\hat{I}$ and E_{total}), while the second vector v_{f2} contains three features ($\hat{I}$, E_{total} and $\bar{I}_{prc}$). Therefore, the performances of different regression methods are compared with each other.

Figure 4-5 presents the RMSE values of using LR, CLR, GPR and CGPR with two and three features. In Figure 4-5 - a) (i.e., the ARC estimation using two features), the GPR model performs better than the LR model and the CGPR model performs better than the CLR model in terms of RMSE. The RMSE of CLR is 2,537.94, while the RMSE of CGPR is 1,443.24.

In Figure 4-5 - b) (i.e., the AELC estimation using two features), the GPR model has a similar performance to the LR model and the CGPR model also has a similar performance to the CLR model. The RMSE of CLR is 4,885.80 while the RMSE of CGPR is 4,752.92.

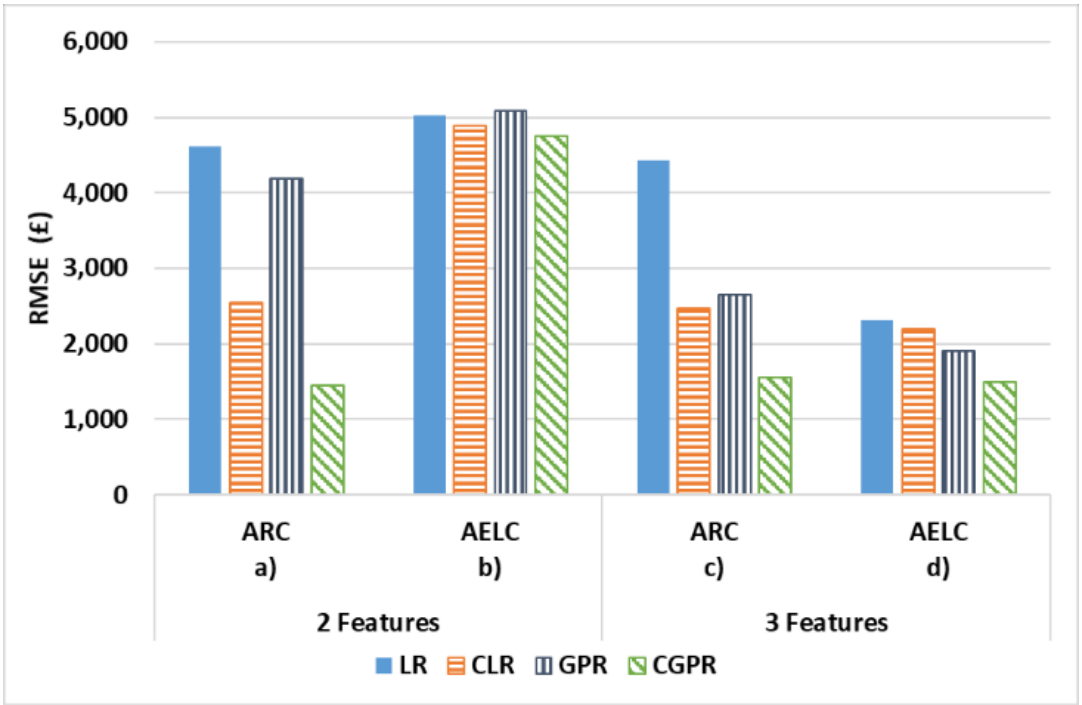


Figure 4-5. Comparison of RMSEs of ARC and AELC estimation with different regression methods

In Figure 4-5 - c) (i.e., the ARC estimation using three features), the GPR method performs better than LR; the CGPR method performs better than the CLR method. The RMSE of the CLR is 2,466.06, while the RMSE of CGPR is 1,554.89.

In Figure 4-5 - d) (i.e., the AELC estimation using three features), the GPR method performs better than the LR; the CGPR method performs better than the CLR method. The RMSE of CLR is 2,199.55, while the RMSE of CGPR is 1,487.71. As a result, CGPR has the best performance among all methods.

For the CGPR model with three features as input, with 95% confidence, the range of RMSEs are [910.84, 1,309.20], [913.59, 1,184.83] and [1,916.03, 3,291.87] for rural, suburban and urban networks, respectively. The suburban networks have the smallest range of the RMSE while the urban networks have the largest range of the RMSE. Therefore, the GPR model performs the best on the imbalance-induced cost estimation for suburban networks among the three types of networks.

C. Discussions

Using Chebyshev's inequality, 11% of the networks are identified as outliers. Figure 4-6shows the comparison of the mean average percentage error (MAPE) before and after the removal of outliers. When using two features, the MAPE of the ARC drops from 29.95% to 23.76% and the MAPE of AELC decreases from 53.86% to 40.75%. When using three features, the MAPE of the ARC drops from 30.06% to 23.32% and the MAPE of AELC decreases from 53.87% to 21.33%.

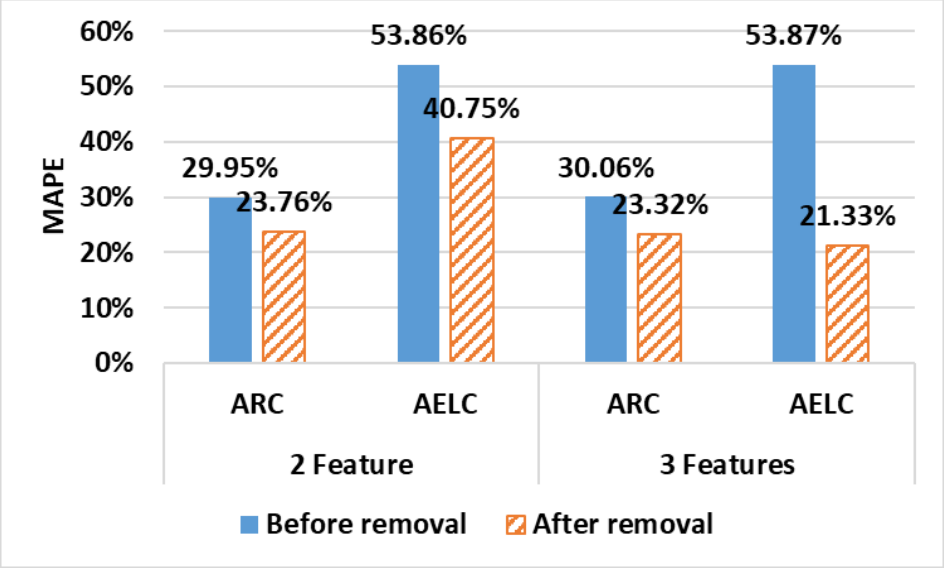


Figure 4-6. Comparison of results before and after removing outliers

One of the main reasons why the MAPE is approximately 23% is that the CGPR approach only requires two or three features from data-scarce LV networks. Another reason is that only one year's

data is used to estimate the imbalance-induced costs over the future 10 years (or 30 years), resulting in an accumulation of errors over the years. Among the three types of networks, the MAPE values for suburban networks are the lowest. In other words, the cost estimations for suburban networks demonstrate the best performance among the three types of networks. On the other hand, the cost estimations for rural networks demonstrate the worst performance among the three types of networks.

In general, there is a lack of monitoring in the UK's millions of LV networks. The two sets of features are chosen in this paper because they are either routinely collected by distribution network operators or are readily available to be collected. Using these features leads to a feasible cost for data collections and the feasibility of the cost-benefit analyses, if scaled up from individual networks to a mass population of networks. Therefore, the features are chosen to best suit the existing level of monitoring in the UK's LV networks and making the methodology scalable to the whole LV networks.

Utilities use load factors to estimate loss factors, which are then used to determine the energy losses of the system. Reference [38] discussed the ways of determining the energy losses using load factor and loss factor. The equations of calculating loss factor is given by [38]:

$$Loss\ Factor = a \times Load\ Factor + (1 - a) \times Load\ Factor^2 \tag{4-21}$$

where $Load\ Factor = \frac{Load_{ave}}{Load_{peak}}$, $\quad Loss\ Factor = \frac{Loss_{ave}}{Loss_{peak}}$

where a is coefficient; $Load_{ave}$ is the average load; $Load_{peak}$ is the peak load; $Loss_{ave}$ is the average loss and $Loss_{peak}$ is the peak loss.

Two values for the coefficient 'a' are suggested by [38], i.e., a = 0.16 and a = 0.3. Both the values are adopted and the lower error of using this method to the estimate energy loss cost is 67.09%. The reason for the large error is that it is difficult to determine the values of 'a' for a data-scarce distribution system. Besides, the distributions system has multiple branches connected to the main feeder which results in a higher estimation error. However, the developed CGPR approach only incurs an error of 21.33% when estimating the AELC. The developed CGPR approach performs better than the method adopted by utilities.

The estimated total imbalance-induced cost using CGPR are compared with the actual values for validation. The actual values are calculated costs using the 10-min resolution time-series data from a year. A random selection of the comparison results (20 networks for each group of networks) are

presented. As shown in **Error! Reference source not found.**, Figure 4-8 and Figure 4-9, the e stimation results follow a similar trend to the actual results.

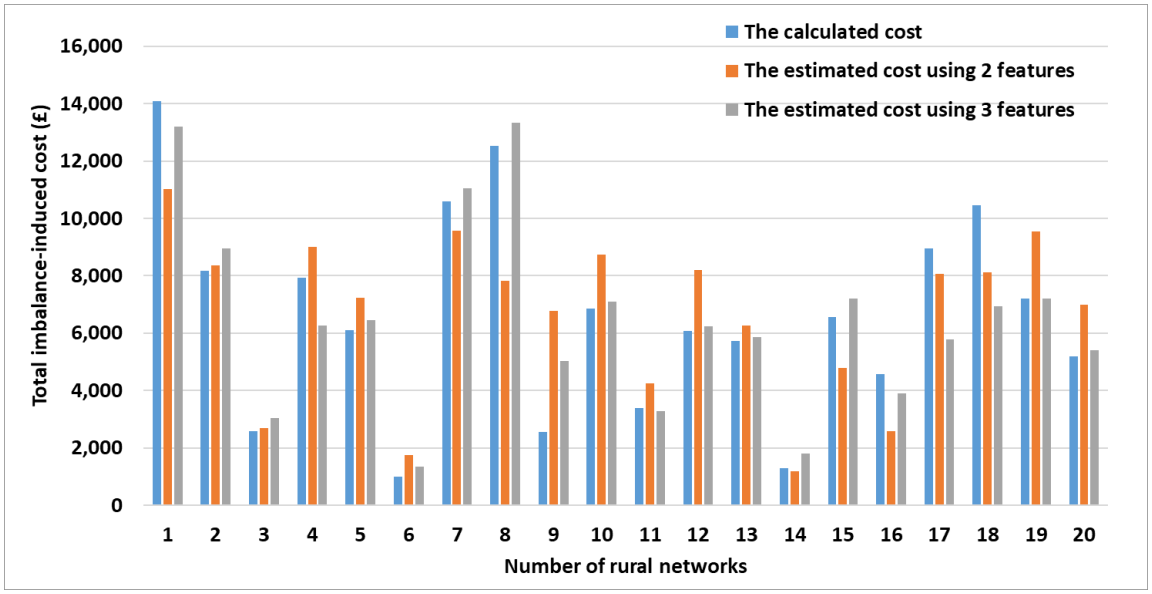


Figure 4-7. Comparison of calculated and estimated total imbalance-induced cost of rural networks

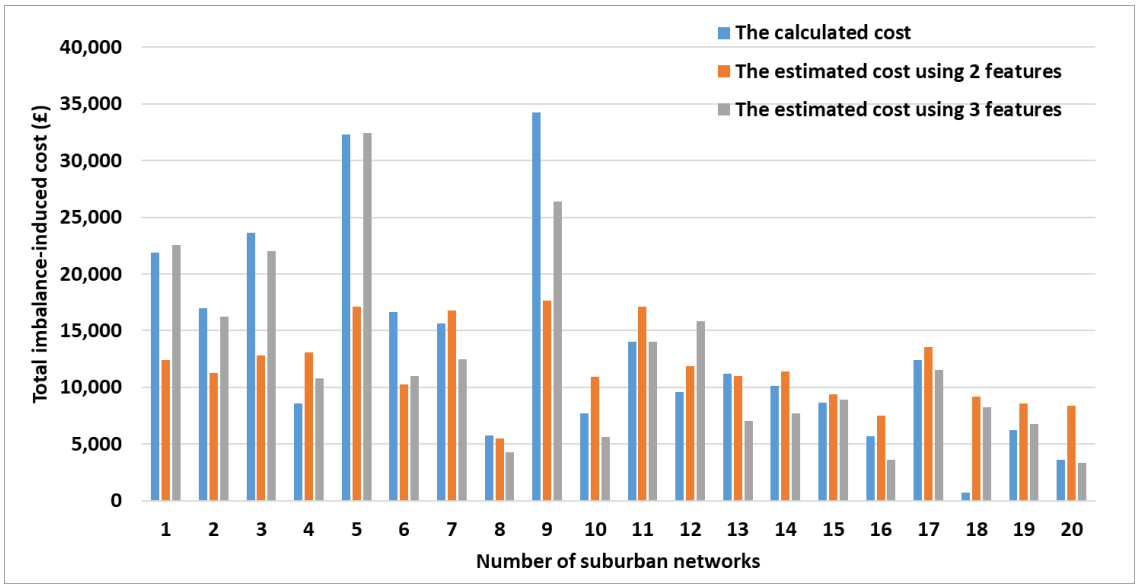


Figure 4-8. Comparison of calculated and estimated total imbalance-induced cost of suburban networks

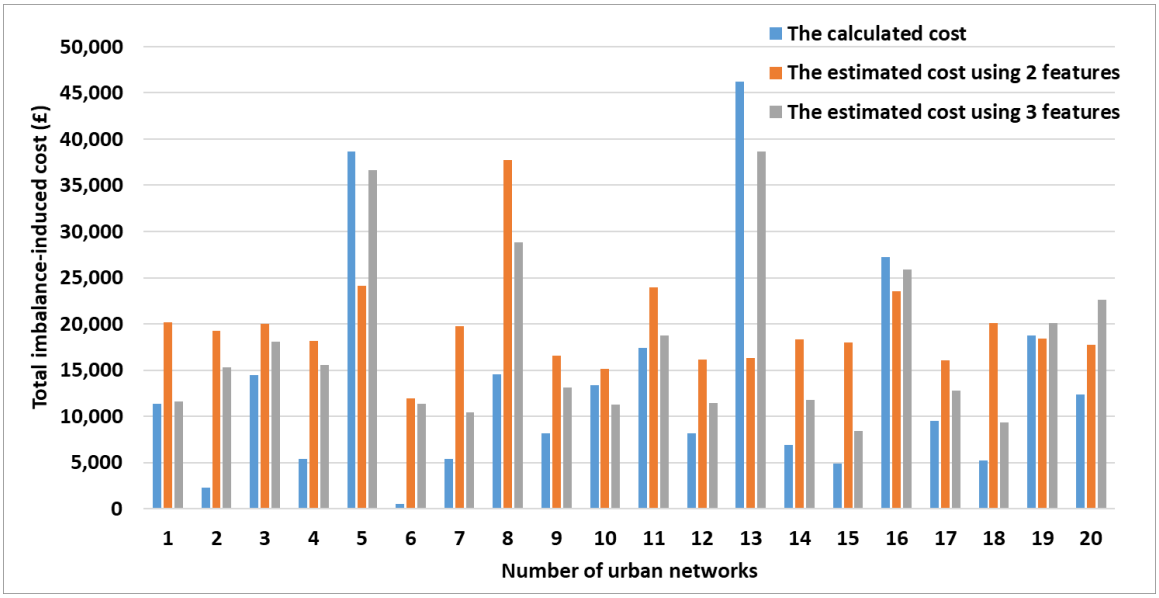


Figure 4-9. Comparison of calculated and estimated total imbalance-induced cost of urban networks

The use of two features and three features are compared with each other. The latter is highly recommended as it incurs a much lower error for the majority of the networks. However, the use of two features still has its value just in case some LV networks do not have three features (i.e. they only have yearly peak current and total energy consumption). In the absence of the third feature (i.e. the yearly average phase currents), one way to perform cost-benefit analyses is to use the two-feature-version of the methodology; an alternative way is to collect the third feature from the networks, but this incurs a data collection cost. This cost can be prohibitively high when the cost-benefit analyses are to be scaled up to a mass population of networks. Therefore, a trade-off should be made between the data collection cost and the accuracy of the methodology for cost-benefit analyses.

Within the dataset of 800 LV networks, 11.2%, 44.4%, and 44.4% are urban, suburban, and rural networks, respectively. The same dataset was used to: 1) develop 11 representative LV substation load profiles [20], [25]; 2) classify four types of phase imbalance in terms of the imbalance direction [39]; 3) estimate the imbalance-induced energy losses in the neutral and ground for data-scarce LV networks [17]. These publications prove the diversity and heterogeneity within the dataset. Furthermore, the dataset corresponds to a geographical area of a similar size and is of a similar nature (a mixture of urban, suburban, and rural networks) to that used in [40].

Given that the model is trained on the dataset from South Wales, UK, the model is applicable to networks within the region of a similar nature to South Wales (a mixture of urban areas like Cardiff, suburban and rural areas like Momonthshire). Caution has to be exercised when applying the trained model on substantially different areas, e.g. central London which is extremely urban and which is unlike anywhere else in the UK. The CGPR methodology is generic. If it is to be applied to other countries or the central London area, it should be trained on the dataset representative of the area in question.

D. Net benefit calculation

Given any data-scarce network, its imbalance-induced costs calculated through CGPR are used for a net benefit calculation. These costs are translated into the benefits of phase balancing for the data-scarce network using (4-13).

Table 4-2 shows the two selected types of phase balancers, along with their costs and lifetimes. The net benefits by applying two phase balancers are calculated using (4-20).

Table 4-2. Costs of phase balancers

Type	ZM-SPC [111]	EQU18 [33]	ANM [41], [42]
Lifetime (Years)	>10	>30	>20
Total costs (£)	4,890	2,381	73,600

The net benefits from phase balancing for data-scarce networks are estimated over the respective lifetime of the two phase balancers and the ANM scheme, i.e. 10 years for ZM-SPC, 30 years for EQU18 and 20 years for the ANM scheme. This paper assumes that power-electronics-based phase balancers and the ANM scheme can achieve full phase balancing because they can perform high-resolution real-time balancing. Detailed information of ZM-SPC and EQU18 are provided in the Appendix. ANM scheme monitors the distribution system and controls the connection of renewable generation and energy storages to minimise phase imbalances.

As stated in the previous section, it is highly recommended using three features as the input for the proposed CGPR approach. In this section, the net benefits are estimated using three features Figure

4-10, Figure 4-11 and Figure 4-12 show the distribution of the estimated net benefits using three features from phase balancing by ZM-SPC for the rural, suburban and urban networks, respectively. Results show that approximately 70% of rural networks, 80% of suburban networks and 90% of urban networks benefit from ZM-SPC.

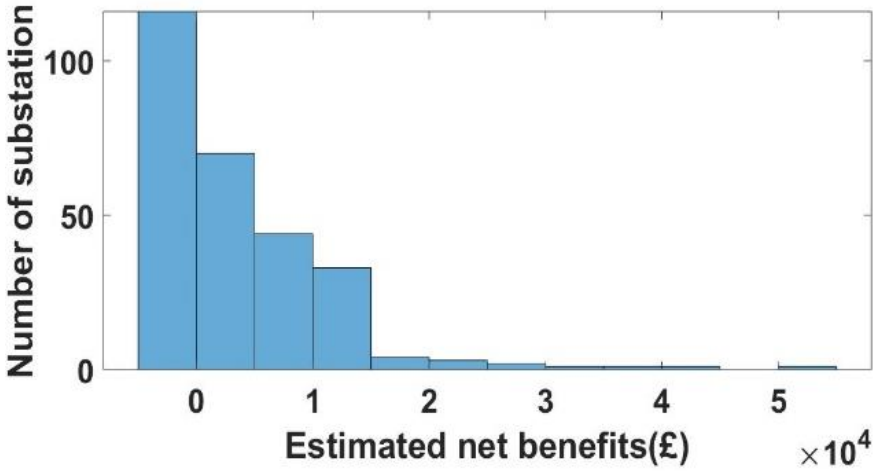


Figure 4-10. The distribution of mean net benefits for rural networks from phase balancing by ZM-SPC

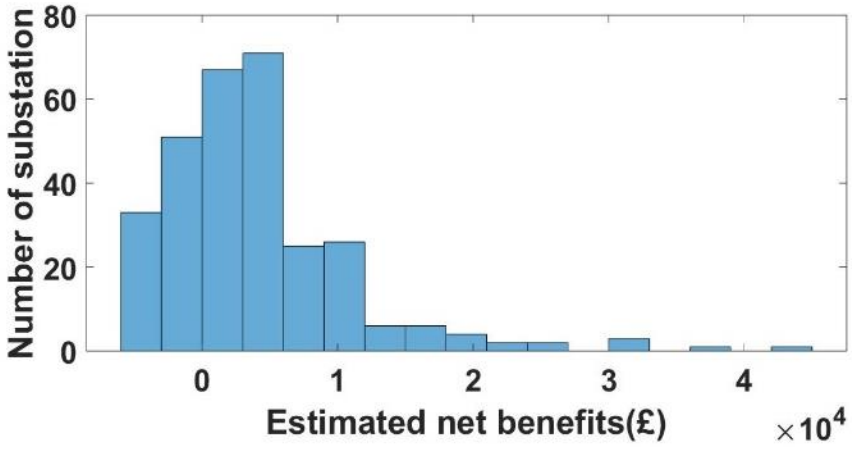


Figure 4-11. The distribution of mean net benefits for suburban networks from phase balancing by ZM-SPC

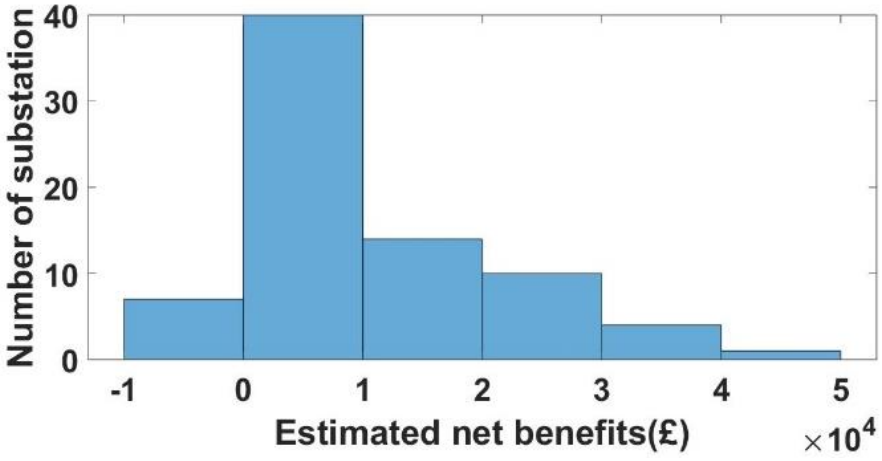


Figure 4-12. The distribution of mean net benefits for urban networks from phase balancing by ZM-SPC

With 95% confidence, the range of net benefits from ZM-SPC for rural, suburban and urban networks are [£2,814.66, £5,106.51], [£3,461.50, £5,346.27] and [£7,591.93, £12,977.50], respectively. The percentage of benefits from ZM-SPC for rural, suburban and urban networks are [36.53%, 51.08%], [41.45%, 52.23%] and [60.82%, 72.63%], respectively.

Figure 4-13, Figure 4-14 and Figure 4-15 show the distribution of the estimated net benefits using three features from phase balancing by EQU18 for the rural, suburban and urban networks, respectively. Results show that approximately 94% of rural networks, 97% of suburban networks and 99% of urban networks benefit from EQU18.

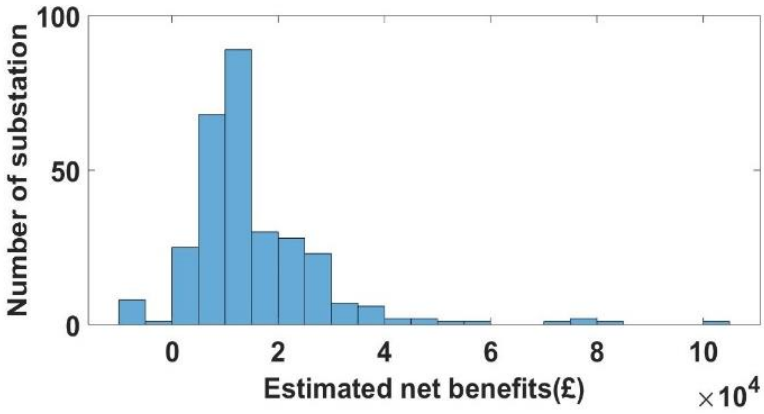


Figure 4-13. The distribution of mean net benefits for rural networks from phase balancing by EQU18

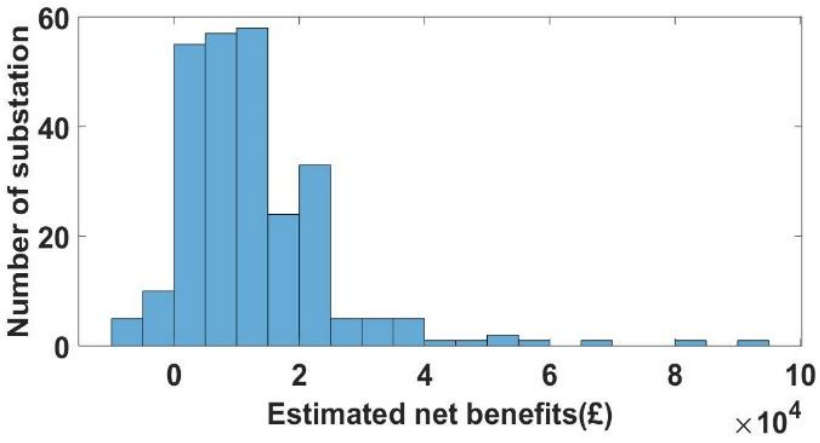


Figure 4-14. The distribution of mean net benefits for suburban networks from phase balancing by EQU18

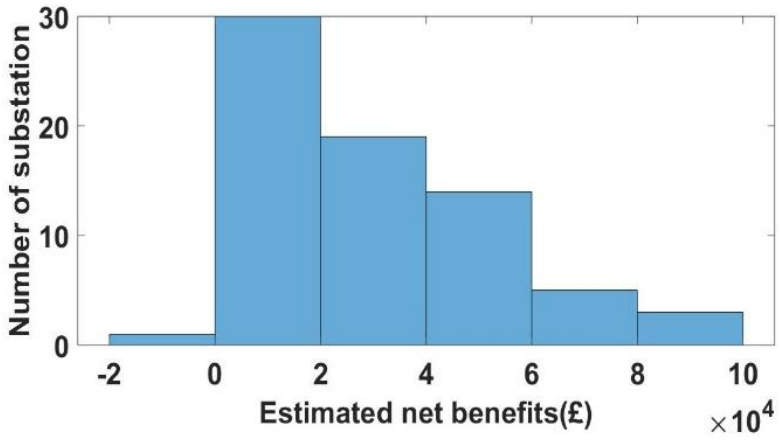


Figure 4-15. The distribution of mean net benefits for urban networks from phase balancing by EQU18

With 95% confidence, the range of net benefits from EQU18 for rural, suburban and urban networks are [£11,153.87, £14,975.80], [£15,218.09, £18,974.98] and [£26,926.63, £39,441.18], respectively. The percentage of benefits from EQU18 for rural, suburban and urban networks are [82.41%, 86.28%], [86.47%, 88.85%] and [91.88%, 94.31%], respectively.

Figure 4-16, Figure 4-17 and Figure 4-18 show the distribution of the estimated net benefits using three features from phase balancing using ANM for the rural, suburban and urban networks, respectively. Results show that approximately 1% of rural networks, 1% of suburban networks and no urban network benefit from the ANM scheme.

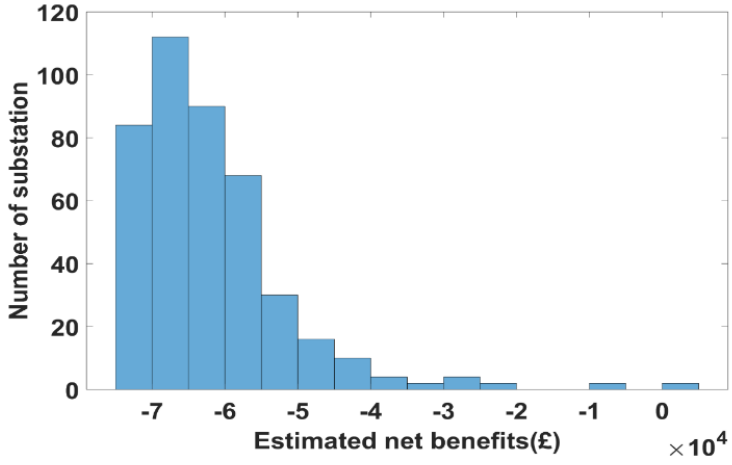


Figure 4-16. The distribution of mean net benefits for rural networks from phase balancing by the ANM scheme

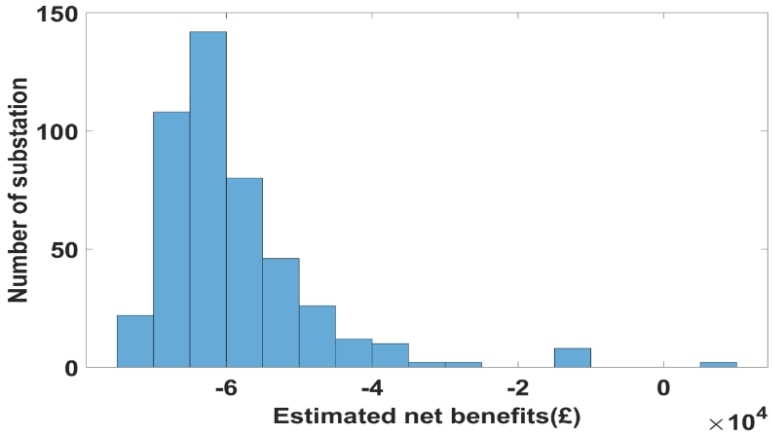


Figure 4-17. The distribution of mean net benefits for suburban networks from phase balancing by the ANM scheme

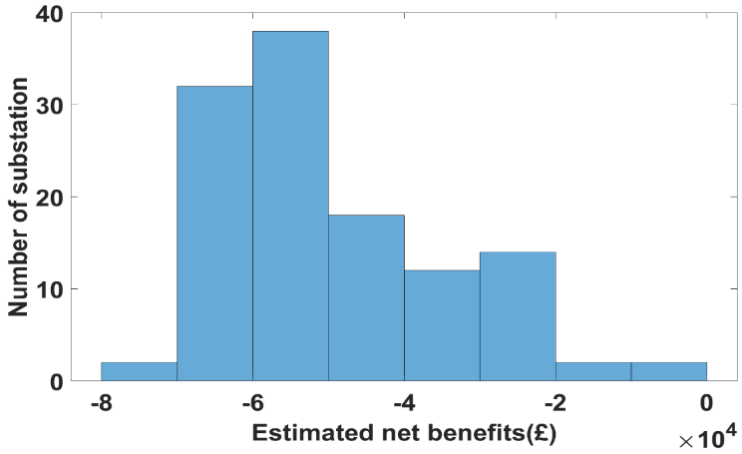


Figure 4-18. The distribution of mean net benefits for urban networks from phase balancing by the ANM scheme

With 95% confidence, the range of net benefits from applying the ANM scheme for rural, suburban and urban networks are [£-63,127.49, £-60,249.45], [£-60,396.22, £-57,564.91] and [£-53,102.38, £-45,313.21], respectively. The net benefits are negative, meaning that adopting the ANM scheme for phase balancing is not cost-effective. However, it is worth mentioning that the ANM scheme typically brings other benefits such as relieving thermal overloads and voltage violations, apart from phase balancing. Therefore, an additional analysis is performed for estimating the benefits of ANM and presented in Section 4.3.

Comparing the RMSEs (given in Section V-B) with the net benefits from phase balancing, it can be found that the RMSEs are insignificant.

Figure 4-19 and Figure 4-20 show the probability that the phase balancing solutions by ZM-SPC and EQU18 would produce a positive net benefit for any data-scarce LV network with 95% confidence, respectively. The probability of having positive net benefit assist DNOs to make the decision on whether to invest in phase balancing.

For example, the CGPR is used to estimate the net benefit for a data-scarce network 10036 from ZM-SPC. The network 10036 is a rural network and its estimated net benefit is £5001. Thus, with 95% confidence, the corresponded probability of network 10036 having a positive net benefit is 96.6%. If the DNO set the acceptable probability as 90%, the network 10036 is therefore identified as worth for phase balancing.

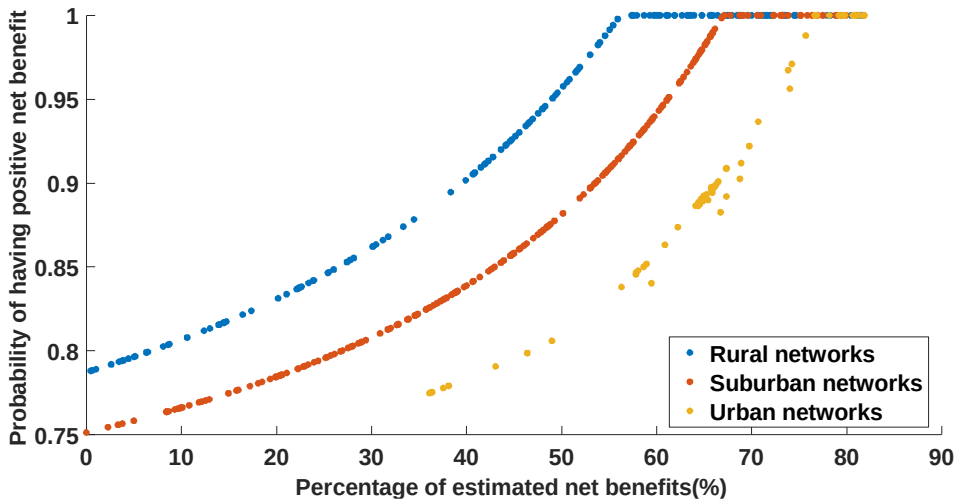


Figure 4-19. The probability of having positive net benefits from phase balancing by ZM-SPC

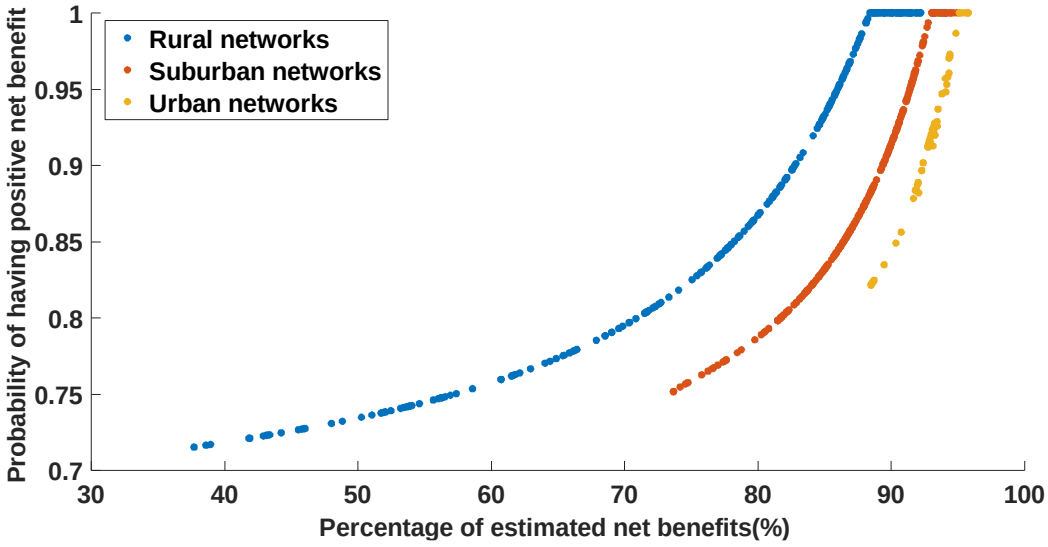


Figure 4-20. The probability of having positive net benefits from phase balancing by EQU18

There is a way to strengthen the robustness of the CGPR model. The CGPR model already outputs the data-scarce LV networks where it is highly likely that a given phase balancing solution will deliver more benefit than cost. In this way, the CGPR model serves as a filter. For these networks (which are a subset of the whole population of networks) that the CGPR model identifies as being worthy of phase balancing, the DNO can further check the cost-benefit of phase balancing on these networks by collecting time-series data from these networks and performing accurate cost-benefit analysis.

VI. Conclusions

This paper addresses a previously unresolved problem faced by the distribution network operators (DNOs), i.e., the cost-benefit analysis of phase balancing solutions for the vast majority of the low voltage (LV) networks that are data-scarce. To this end, this paper develops a new cluster-wise Gaussian process regression (CGPR) approach.

The approach is validated by the case studies considering two types of phase balancers and the active network management (ANM) scheme. The phase balancers are ZM-SPC and EQU18 with different costs and lifetime. The maximum potential net benefits for all types of LV networks are calculated for each phase balancer. Given any data-scarce network and phase balancing solution, the probability that the solution will produce a positive net benefit is quantified.

A major advantage of the approach is that it only requires the annual peak current and the total energy consumption throughout a year – these data are collected only once a year. The developed

approach offers a cost-effective and efficient way to help DNOs understand: 1) whether a phase balancing solution is economically feasible for any data-scarce network; 2) if yes, the maximum potential net benefit from the solution.

REFERENCES

[1] K. Ma, R. Li, and F. Li, “Utility-Scale Estimation of Additional Reinforcement Cost From Three-Phase Imbalance Considering Thermal Constraints,” IEEE Transactions on Power Systems, vol. 32, no. 5, pp. 3912-3923, 2017.

[2] G. Mokryani, A. Majumdar, and B. C. Pal, “Probabilistic method for the operation of three-phase unbalanced active distribution networks,” IET Renewable Power Generation, vol. 10, no. 7, pp. 944-954, 2016.

[3] M. W. Siti, D. V. Nicolae, A. A. Jimoh, and A. Ukil, “Reconfiguration and Load Balancing in the LV and MV Distribution Networks for Optimal Performance,” Power Delivery, IEEE Transactions on, vol. 22, no. 4, pp. 2534-2540, 2007.

[4] S. Yan, S. C. Tan, C. K. Lee, B. Chaudhuri, and S. Y. R. Hui, “Electric Springs for Reducing Power Imbalance in Three-Phase Power Systems,” IEEE Transactions on Power Electronics, vol. 30, no. 7, pp. 3601-3609, 2015.

[5] J. D. Watson, N. R. Watson and I. Lestas, "Optimized Dispatch of Energy Storage Systems in Unbalanced Distribution Networks," in IEEE Transactions on Sustainable Energy, vol. 9, no. 2, pp. 639-650, April 2018, doi: 10.1109/TSTE.2017.2752964.

[6] T. Alinjak, I. Pavic, and K. Trupinic, “Improved three-phase power flow method for calculation of power losses in unbalanced radial distribution network,” CIRED - Open Access Proceedings Journal, vol. 2017, no. 1, pp. 2361-2365, 2017.

[7] K. Ma, R. Li, and F. Li, “Quantification of Additional Asset Reinforcement Cost From 3-Phase Imbalance,” IEEE Transactions on Power Systems, vol. 31, no. 4, pp. 2885-2891, 2016.

[8] J. Zhu, M. Y. Chow, and F. Zhang, “Phase balancing using mixed-integer programming [distribution feeders],” Power Systems, IEEE Transactions on, vol. 13, no. 4, pp. 1487-1492, 1998.

[9] A. Kavousi-Fard, T. Niknam, and M. Fotuhi-Firuzabad, “A Novel Stochastic Framework Based on Cloud Theory and Modified Bat Algorithm to Solve the Distribution Feeder Reconfiguration,” IEEE Transactions on Smart Grid, vol. 7, no. 2, pp. 740-750, 2016.

[10] C. H. Lin, C. S. Chen, H. J. Chuang, M. Y. Huang, and C. W. Huang, “An Expert System for Three-Phase Balancing of Distribution Feeders,” IEEE Transactions on Power Systems, vol. 23, no. 3, pp. 1488-1496, 2008.

[11] P. Lico, M. Marinelli, K. Knezović, and S. Grillo, "Phase balancing by means of electric vehicles single-phase connection shifting in a low voltage Danish grid." pp. 1-5.

[12] T. S. Win, Y. Hisada, T. Tanaka, E. Hiraki, M. Okamoto, and S. R. Lee, “Novel Simple Reactive Power Control Strategy With DC Capacitor Voltage Control for Active Load Balancer in Three-Phase Four-Wire Distribution Systems,” IEEE Transactions on Industry Applications, vol. 51, no. 5, pp. 4091-4099, 2015.

[13] T. Chen, “Evaluation of line loss under load unbalance using the complex unbalance factor,” IEE Proceedings - Generation, Transmission and Distribution, vol. 142, no. 2, pp. 173-178, 1995.

[14] L. Ochoa, R. Ciric, A. Padilha-Feltrin, and G. Harrison, Evaluation of distribution system losses due to load unbalance, 2005.

[15] M. Baggu, J. Giraldez, T. Harris, N. Brunhart-Lupo, L. Lisell, and D. Narang, "Interconnection assessment methodology and cost benefit analysis for high-penetration PV deployment in the Arizona Public Service system." pp. 1-6.

[16] F. M. Camilo, R. Castro, M. E. Almeida, and V. F. Pires, “Assessment of overvoltage mitigation techniques in low-voltage distribution networks with high penetration of photovoltaic microgeneration,” IET Renewable Power Generation, vol. 12, no. 6, pp. 649-656, 2018.

[17] L. Fang, K. Ma, R. Li, Z. Wang, and H. Shi, “A Statistical Approach to Estimate Imbalance-Induced Energy Losses for Data-Scarce Low Voltage Networks,” IEEE Transactions on Power Systems, pp. 1-1, 2019.

[18] U. P. Networks. "Phase Switch System," https://innovation.ukpowernetworks.co.uk/projects/phase-switch-system.

[19] S. E. Networks, HV and LV Phase Imbalance Assessment 7640–07–D4, September 2015.

[20] R. Li, C. Gu, F. Li, G. Shaddick, and M. Dale, “Development of Low Voltage Network Templates Part I: Substation Clustering and Classification,” Power Systems, IEEE Transactions on, vol. 30, no. 6, pp. 3036-3044, 2015.

[21] A. A. Sallam, and O. P. Malik, Electric Distribution Systems, Hoboken, NJ, USA: Hoboken, NJ, USA: John Wiley & Sons, Inc., 2012.

[22] G. Cronshaw, EARTHING: YOUR QUESTIONS ANSWERED, IEE Wiring Matters, 2005.

[23] E. J. O, O. S.O, and I. S. A;, “Evaluation Of Distribution System Losses Due To Unbalanced Load In Transformers A Case Study Of Guinness 15MVA, 33/11KV, Injection Substation And Its Associated 11/0.415kv Transformers In Benin City, Nigeria,” International Journal of Engineering Research & Technology, vol. 2, no. 3, March, 2013.

[24] H. Ying-Yi, and C. Zuei-Tien, “Development of energy loss formula for distribution systems using FCN algorithm and cluster-wise fuzzy regression,” IEEE Transactions on Power Delivery, vol. 17, no. 3, pp. 794-799, 2002.

[25] R. Li, C. Gu, F. Li, G. Shaddick, and M. Dale, “Development of Low Voltage Network Templates—Part II: Peak Load Estimation by Clusterwise Regression,” IEEE Transactions on Power Systems, vol. 30, no. 6, pp. 3045-3052, 2015.

[26] C. B. Do, “Gaussian processes,” 2008.

[27] H. J. Seltman, Experimental Design and Analysis, 2018.

[28] T. Wong, and N. Yang, “Dependency Analysis of Accuracy Estimates in k-Fold Cross Validation,” IEEE Transactions on Knowledge and Data Engineering, vol. 29, no. 11, pp. 2417-2427, 2017.

[29] B. G. A. T. A. F. S. K. Cooley, “Data outlier detection using the Chebyshev theorem,” in 2005 IEEE Aerospace Conference, Big Sky, MT, USA, 2005, pp. 3814-3819.

[30] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, “Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median,” Journal of Experimental Social Psychology, vol. 49, no. 4, pp. 764-766, 2013.

[31] Y. Zhang, J. Liu, and H. Li, “An Outlier Detection Algorithm Based on Clustering Analysis,” in 2010 First International Conference on Pervasive Computing, Signal Processing and Applications, 2010, pp. 1126-1128.

[32] "ZM-SPC Series Three phase unbalanced device," March, 2019; http://zmgs.com/en/index.php?r=article/Content/index&content_id=596.

[33] "EQUI8 : THREE-PHASE LOW VOLTAGE NETWORK BALANCER," March, 2019; http://cmetransformateur.com/equi8-en/.

[34] S. Electric. "HV/LV distribution transformers," http://mt.schneider-electric.be/main/tfo/catalogue/an_iec.pdf.

[35] A. S. Sidhu, M. G. Pollitt, and K. L. Anaya, “A social cost benefit analysis of grid-scale electrical energy storage projects: A case study,” Applied Energy, vol. 212, pp. 881-894, 2018/02/15/, 2018.

[36] "Pathways for the GB Electricity Sector to 2030," https://www.energy-uk.org.uk/publication.html?task=file.download&id=5722.

[37] Y. Zhang, F. Li, Z. Hu, and G. Shaddick, “Quantification of low voltage network reinforcement costs: A statistical approach,” IEEE Transactions on Power Systems, vol. 28, no. 2, pp. 810-818, 2013.

[38] K. M. P. K. Sen, “A Better Understanding of Load and Loss Factors ” in 2008 IEEE Industry Applications Society Annual Meeting, Edmonton, AB, Canada, 2008.

[39] W. Kong, K. Ma, and Q. Wu, “Three-Phase Power Imbalance Decomposition Into Systematic Imbalance and Random Imbalance,” IEEE Transactions on Power Systems, vol. 33, no. 3, pp. 3001-3012, 2018.

[40] V. Rigoni, L. F. Ochoa, G. Chicco, A. Navarro-Espinosa, and T. Gozel, “Representative Residential LV Feeders: A Case Study for the North West of England,” IEEE Transactions on Power Systems, vol. 31, no. 1, pp. 348-360, 2016.

[41] Low Carbon London Project Closedown Report, UK Power Networks, 2015.

[42] Z. Hu, and F. Li, “Cost-Benefit Analyses of Active Distribution Network Management, Part II: Investment Reduction Analysis,” IEEE Transactions on Smart Grid, vol. 3, no. 3, pp. 1075-1081, 2012.

[43] *Demonstrating the Benefits of Monitoring LV Networks with embedded PV Panels and EV Charging Point*, Scottish and Southern Energy Power Distribution, 2013.

4.3. Additional Analysis and Discussion

The ARC and AELC are the differences between the three-phase imbalanced scenario and a balanced scenario. Figure 4-21 shows the percentage of AELC compared to the energy loss cost with balanced three phases; and the percentage of ARC compared to the present value of network reinforcement cost with balanced three phases. The load growth rate is set to be 0.82% which is the same as previous case studies. As can be seen, the suburban networks have the largest AELC, while the rural networks have the least AELC compared to the balanced scenario. The average AELCs for rural, suburban and urban networks are 4%, 18% and 5%, respectively.

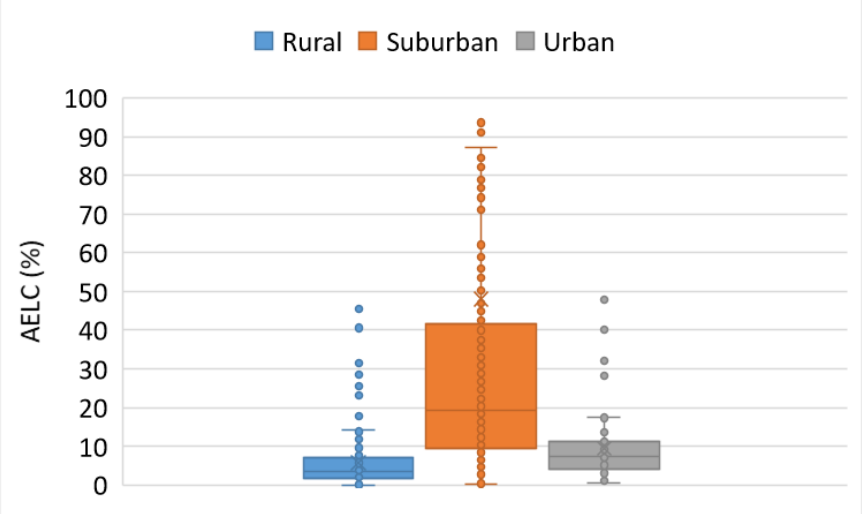


Figure 4-21. The percentages of AELC for LV networks compared to the balanced scenario

Figure 4-22 also shows that the rural networks have the largest ARC while the urban networks have the least ARC. Comparing to the balanced scenario, the average ARCs for rural, suburban and urban networks are 202%, 116% and 78%, respectively.

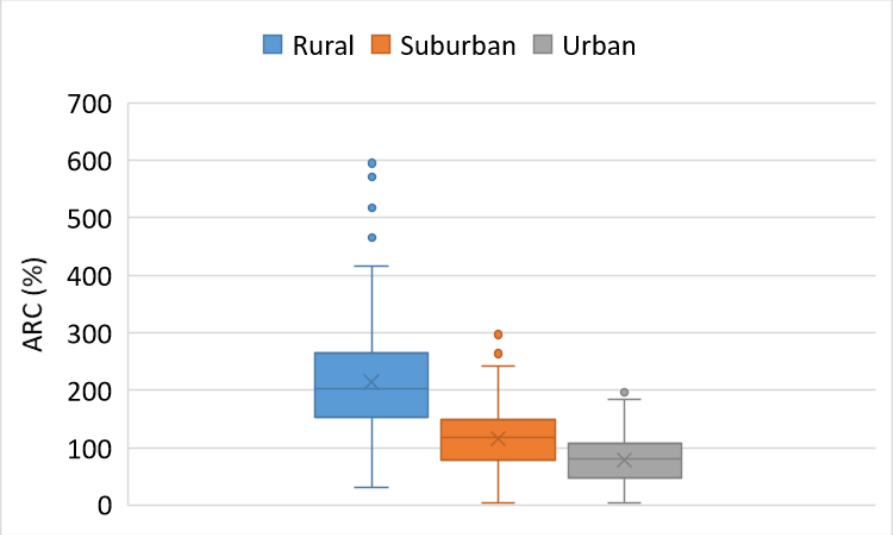


Figure 4-22. The percentages of ARC for LV networks compared to the balanced scenario

Figure 4-23, Figure 4-24 and Figure 4-25 present the ARC with changing load growth rate for different group of networks. It can be seen that the percentage of ARC decreases with increasing load growth rate. It means that with large load growth, the impact of phase imbalance on LV distribution network reinforcement cost becomes low. This is because a large increase in load demand will lead to high network reinforcement cost. Therefore, the differences in network reinforcement between balanced and imbalanced scenarios, i.e., the ARC, become smaller. When the load growth rate changes from 0.5% to 2.0%, the percentage of ARC for the LV networks drop rapidly. When the load growth rate increases from 2.0% to 4.0%, the percentage of ARC for the LV networks decreases slowly.

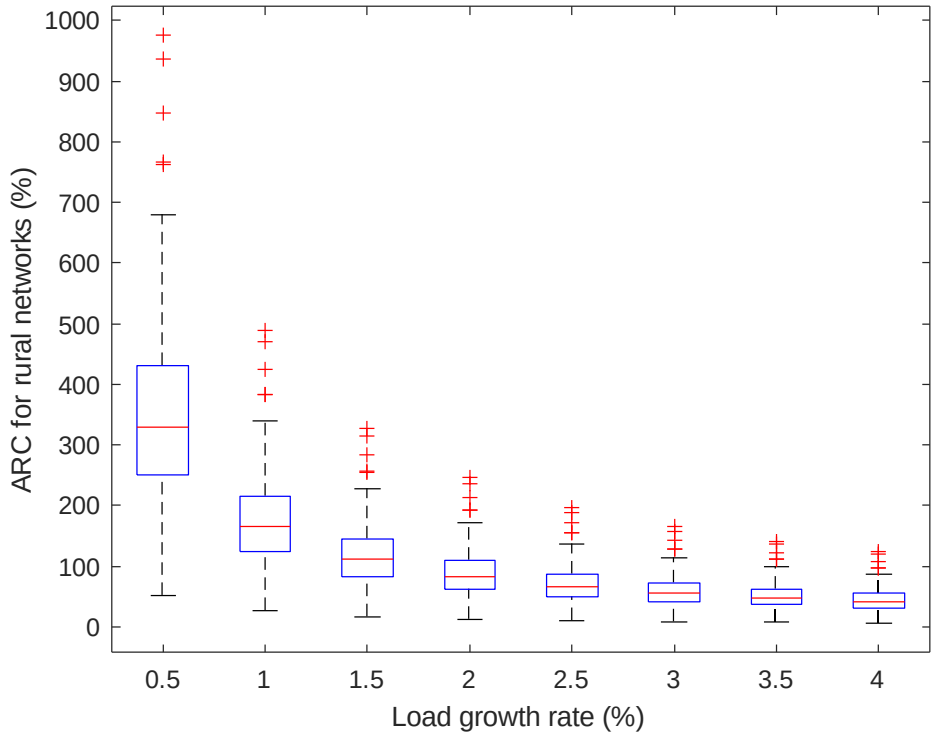


Figure 4-23. The ARC for rural networks with changing load growth rate compared to the balanced scenario

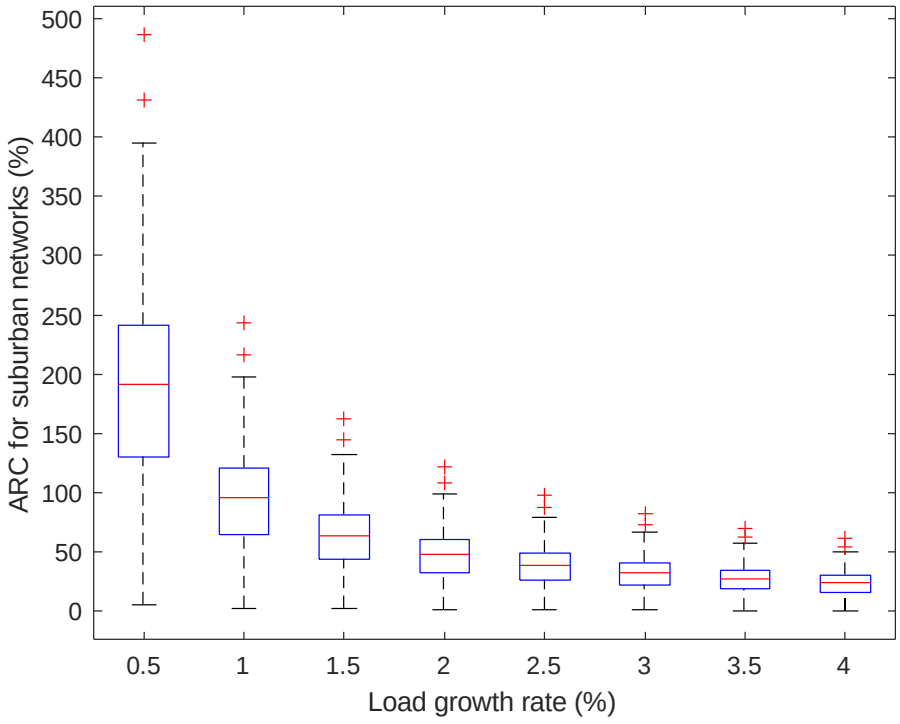


Figure 4-24. The ARC for suburban networks with changing load growth rate compared to the balanced scenario

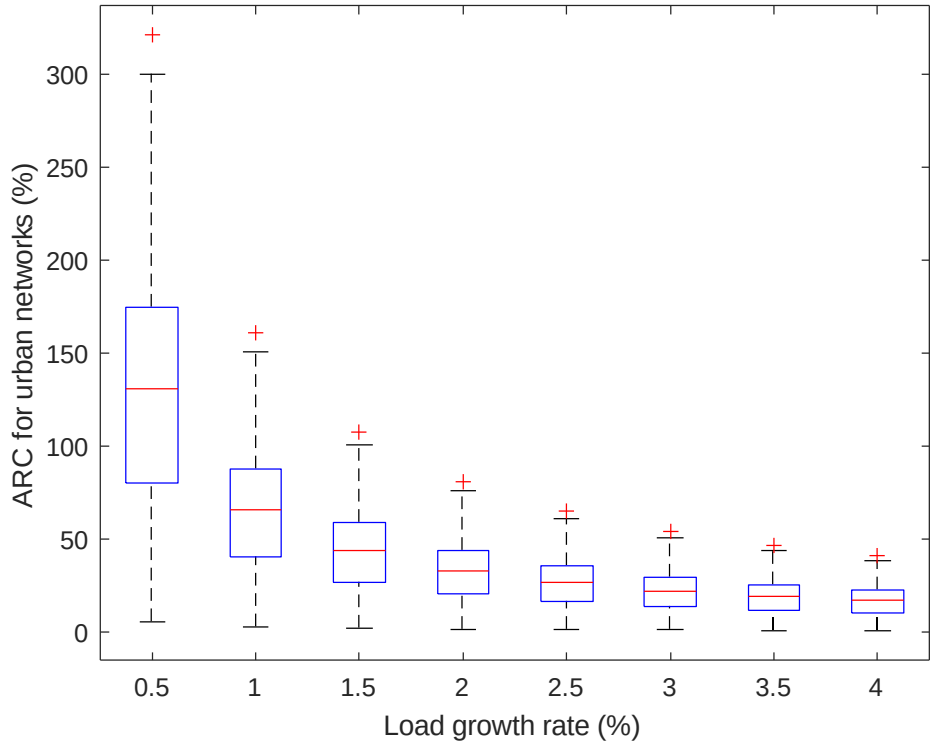


Figure 4-25. The ARC for urban networks with changing load growth rate compared to the balanced scenario

The ANM scheme typically brings other benefits such as relieving thermal overloads and voltage violations, apart from phase balancing. Therefore, it is more accurate to include a factor to reflects the percentage cost of ANM applied for phase balancing. However, up to now, there is no research analysing such factor. As a result, the factor is assumed to be 5% in this thesis.

Figure 4-26, Figure 4-27 and Figure 4-28 show the distribution of the estimated net benefits using three features from phase balancing using ANM for the rural, suburban and urban networks, respectively. Results show that approximately 93.27% of rural networks, 83.19% of suburban networks and 100% urban network benefit from the ANM scheme.

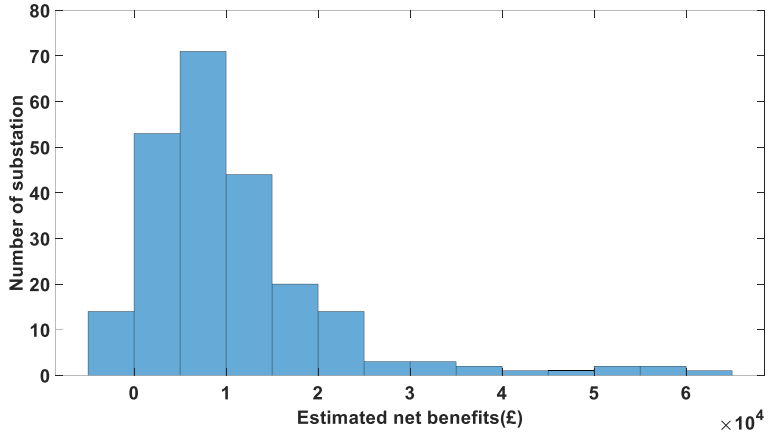


Figure 4-26. The distribution of mean net benefits for rural networks from phase balancing by the ANM scheme

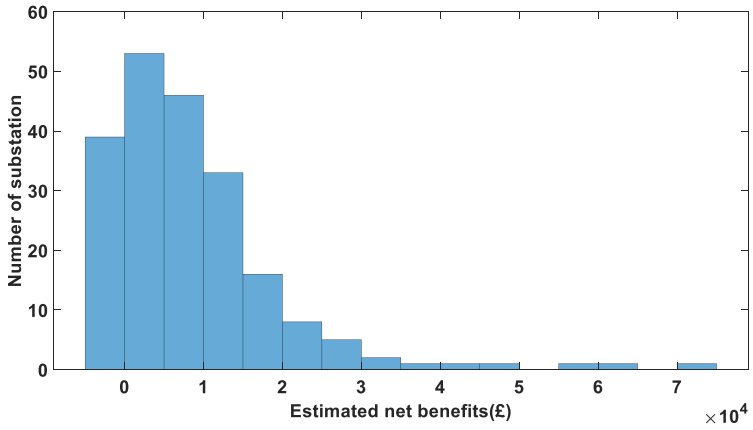


Figure 4-27. The distribution of mean net benefits for suburban networks from phase balancing by the ANM scheme

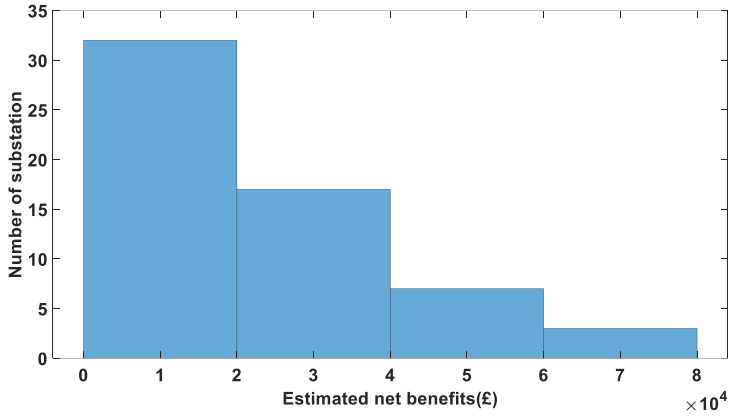


Figure 4-28. The distribution of mean net benefits for urban networks from phase balancing by the ANM scheme

Figure 4-29 shows the probability that the phase balancing solution by ANM would produce a positive net benefit for any data-scarce LV network with 95% confidence. Compared to ZM-SPC, ANM has lower probability of bringing positive benefits with similar percentage of estimated benefits. Compared to EQU18, ANM has a lower percentage of estimated benefits for a similar probability of bringing positive benefits. However, it is worthy to note that changing the factor, i.e., the cost share for phase balancing, of the ANM cost will have impacts on the results. Further investigations could be done to achieve a more accurate estimation of the factor.

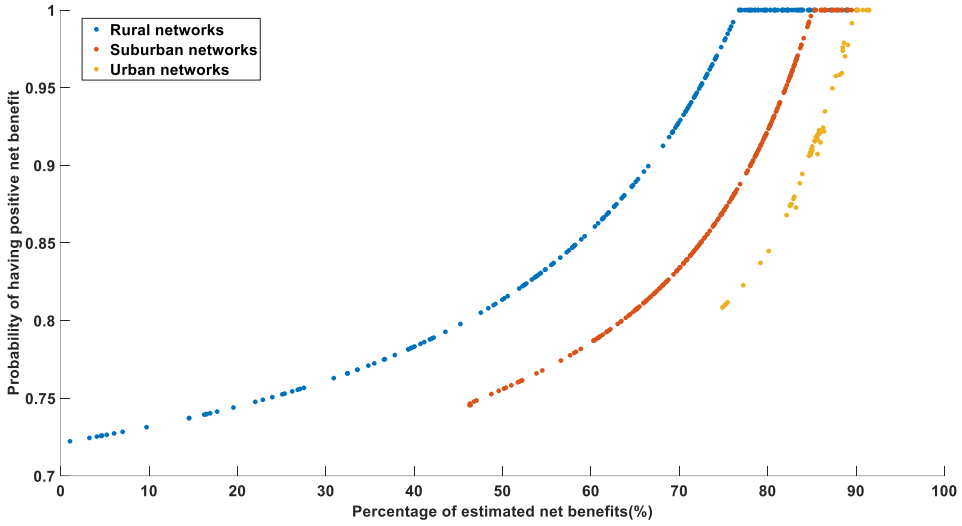


Figure 4-29. The probability of having positive net benefits from phase balancing by ANM

Chapter 5. Phase Imbalances in The Future Distribution System

This chapter analyses the impacts of future distribution system architectures on phase imbalances using Monte Carlo simulation. The imbalance-induced costs are calculated and analysed.

5.1. Introduction

The distribution system is evolving because of increasing decarbonisation, digitalisation and decentralisation. The phase imbalance problem is further complicated by growing uptake of single-phase connected low carbon technologies (LCTs), including photovoltaic (PV) systems and electric vehicles (EVs) in the distribution system [115], [116]. The National Grid estimates that the UK's number of EVs on the road could reach 36 million by 2040 and the capacity of PV units could reach 38GW by 2050 [17]. The increasing LCTs cause phase power imbalance to change randomly and therefore change the consequences of phase power imbalance. As a result, it is important to estimate the consequences of phase power imbalance under the increasing penetration of LCTs.

Therefore, this chapter develops a new probabilistic analysis to investigate the impact of increasing LCT penetration on phase power imbalance in the UK's LV distribution networks. The LCT considered in this chapter is residential PV generation and one type of EV battery with slow charging. Monte Carlo simulations are performed to account for the uncertainties in LCT sizes, connection locations and connection time. The developed probabilistic impact assessment helps the distribution network operators (DNOs) understand the possible imbalance-induced costs under different penetration levels of LCTs. The estimated imbalance-induced costs help DNOs analyse the cost-benefits for future phase balancing in light of the future LCT growth.

The content of this chapter is cited from a submitted article in Electric Power System Research by the author. This chapter is formed in an alternative-based format. All the indexes, figures, tables, equations and references are numbered independently.

The following sections are organised as follows: Section 5.2 presents the probabilistic impact assessment; Section 5.3 discusses the impacts of EV and PV penetrations on three-phase power imbalance decomposition.

5.2. Probabilistic Impact Assessment of Phase Power Imbalance

<table>
<tr><th colspan="4">This declaration concerns the article entitled:</th></tr>
<tr><td colspan="4">Probabilistic impact assessment of phase power imbalance in the LV networks with increasing penetrations of low carbon technologies</td></tr>
<tr><td colspan="4">Publication status (tick one)
Draft manuscript ☐ Submitted ☐ In review ☑ Accepted ☐ Published ☐</td></tr>
<tr><td>Publication details (reference)</td><td colspan="3">W. Kong, K. Ma and F. Li, " Probabilistic Impact Assessment of Phase Power Imbalance in the LV networks with Increasing Penetrations of Low Carbon Technologies," submitted to Electric Power Systems Research, September 2020.</td></tr>
<tr><td colspan="4">Copyright status (tick the appropriate statement)
I hold the copyright for this material ☑ Copyright is retained by the publisher, but I have been given permission to replicate the material here ☐</td></tr>
<tr><td>Candidate's contribution to the paper (provide details, and also indicate as a percentage)</td><td colspan="3">The candidate contributed to / considerably contributed to / predominantly executed the…

Formulation of ideas:
• 90%
• The idea of performing impact analysis of LCTs on phase imbalance is guided by Dr Kang Ma and Prof Furong Li.

Design of methodology:
• 100%

Experimental work:
• 100%

Presentation of data in journal format:
• 100%</td></tr>
<tr><td>Statement from Candidate</td><td colspan="3">This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.</td></tr>
<tr><td>Signed</td><td>Wangwei Kong</td><td>Date</td><td>20/12/2020</td></tr>
</table>

Probabilistic Impact Assessment of Phase Power Imbalance in the LV networks with Increasing Penetrations of Low Carbon Technologies

Wangwei Kong, Kang Ma, *Member, IEEE*, and Furong Li, *Senior Member, IEEE*

***Abstract*: Phase imbalances cause a range of network issue, from day-to-day energy losses to long-run capacity wastes that increase investment costs. The impact on low voltage (LV) network from phase imbalance has been investigated independently for losses and investment. However, no research was carried out on the total imbalance-induced cost (TIC) that includes both day-to-day energy losses and long-run capacity wastes, and how the relationship between the two may change with the increasing penetrations of single-phase low carbon technologies (LCTs). Analysing the TIC is important for distribution network operators (DNOs) as the day-to-day energy loss cost cannot be ignored as it may exceed the long-run network investment cost. This paper develops a new probabilistic analysis to investigate the impact of increasing LCT penetration on TIC in the UK's LV distribution networks. Monte Carlo simulations are performed to account for the uncertainties in LCT sizes, connection locations and connection time. Case studies show that the additional energy loss cost exceeds the additional reinforcement cost in urban networks when the LCT penetration level reaches 70%. The key findings will help the DNOs understand the range of TIC and the relationship between imbalance-induced energy losses and capacity wastes under increasing LCT penetrations.**

1. Introduction

Phase imbalance means either the magnitudes of the three phases are not the same, or their phase angles are not 120° apart from each other. Phase imbalance is a widespread problem in the UK. More than 70% of LV networks [1] suffer severe phase power imbalances, mainly caused by uneven load allocation [2], [3] and random load behaviours [3], [4]. Phase imbalance causes two consequences to distribution networks: energy losses [2], [3] and capacity wastes (that are translated into additional investment costs [4], [5]). The phase imbalance problem is further complicated by growing uptake of low carbon technologies (LCTs), including photovoltaic (PV) systems and electric vehicles (EVs) in the distribution system [6], [7]. The National Grid estimates that the UK's number of EVs on the road could reach 36m by 2040 and the capacity of PV units could reach 38GW by 2050 [8]. The increasing LCTs cause phase power imbalance to change randomly and therefore

change the relationship between the above two imbalance-induced consequences. As a result, it is important to quantify the consequences of phase power imbalance under increasing penetration of LCTs. This is the focus of the paper.

Much effort is made to analyse the voltage imbalance caused by LCT penetrations, such as uncoordinated EV charging [9-14], PV inverters [14-20] and heat pumps (HPs) [11, 21] in the distribution networks. However, none of the research discussed the impacts of increasing LCT penetration on the phase power imbalance. The phase power imbalance is a direct consequence of voltage and current imbalances [22]. It incurs additional long-run network investment and day-to-day energy loss costs to the distribution networks. These imbalance-induced costs had also been investigated previously in [23], [24]. Reference [23] presented a way to estimate the additional reinforcement cost (ARC) for both LV transformers and main feeders using the degree of power imbalance. Our previous work [24] proposed a method to estimate the ARC and the additional energy loss cost (AELC) for data-scarce LV networks. The AELC includes the transformer copper loss cost and the costs caused by the neutral line current [24]. Nonetheless, these research works only focus on the LV networks with traditional passive loads, rather considering the increasing penetrations of LCTs.

Therefore, there is a gap in assessing the total imbalance-induced cost (TIC), which includes both day-to-day energy loss cost and long-run network investment cost (the latter is caused by imbalance-induced capacity wastes), under increasing penetrations of LCTs.

This paper addresses a different problem from [11]: Reference [11] analysed the impacts of four type of LCTs on customer voltage violations and feeder loading levels to help DNOs estimate the LCT hosting capacities for LV distribution feeders. However, this paper focuses on the impacts of LCTs on phase power imbalance and the corresponding TIC. Furthermore, this paper helps DNO find the balance between the day-to-day and long-run costs under increasing penetrations LCTs. The LCTs considered in this paper are EV and PV units because they are expected to rapidly increase in the near future [8]. However, this framework can also be extended to other LCTs.

This paper performs a new analysis to quantify the impacts of LCTs on phase power imbalance, including how LCTs affect the TIC, considering both the day-to-day energy losses and long-term investment costs induced by phase imbalance. In this framework, Monte-Carlo simulations are adopted to account for the LCT uncertainties within sizes, connection locations and connection time. The LCT considered are EV with one type battery and slow charging and residential PV generation (<4kW).

Analysing the TIC helps distribution network operators (DNOs) understand how the relationship between the two consequences may change with the increasing penetrations of single-phase LCTs. The methodology considers three scenarios: EV only scenario, PV only scenario and both EV and PV scenario, given the fact that both PV and EV grow rapidly in the foreseeable future.

The developed framework has values: 1) the probabilistic impact assessment helps the DNOs understand the possible impacts of LCTs on power imbalances in the LV distribution networks; 2) the estimated TIC help DNOs understand the relationship between imbalance-induced energy losses and capacity wastes as well as when the TIC reaches the minimum, under increasing LCT penetrations.

The remainder of this paper is organized as follows: Section 2 presents the overview of the developed methodology; Section 3 shows the network, LCTs profiles and the calculation of the imbalance-induced costs; Section 4 performs the case studies of the probabilistic impact assessment, Section 5 discuss the results, and Section 6 concludes the paper.

2. Overview of methodology

To perform an accurate impact assessment, full time-series of phase voltage and current data and LCT generation/ consumption data are required as the input data. However, the majority of UK's LV networks are unmonitored and there are significant uncertainties of the LCT's consumption or generation. Full data from 800 representative LV networks throughout a year are used in this paper. Details of the LV networks are explained in Section III-A.

Figure 5-1 shows an overview of the approach. The key to this approach is using Monte Carlo simulations to represent the uncertainties of LCTs and calculate the imbalance-induced costs for all the LV networks with changing LCT penetrations. The approach consists of three stages:

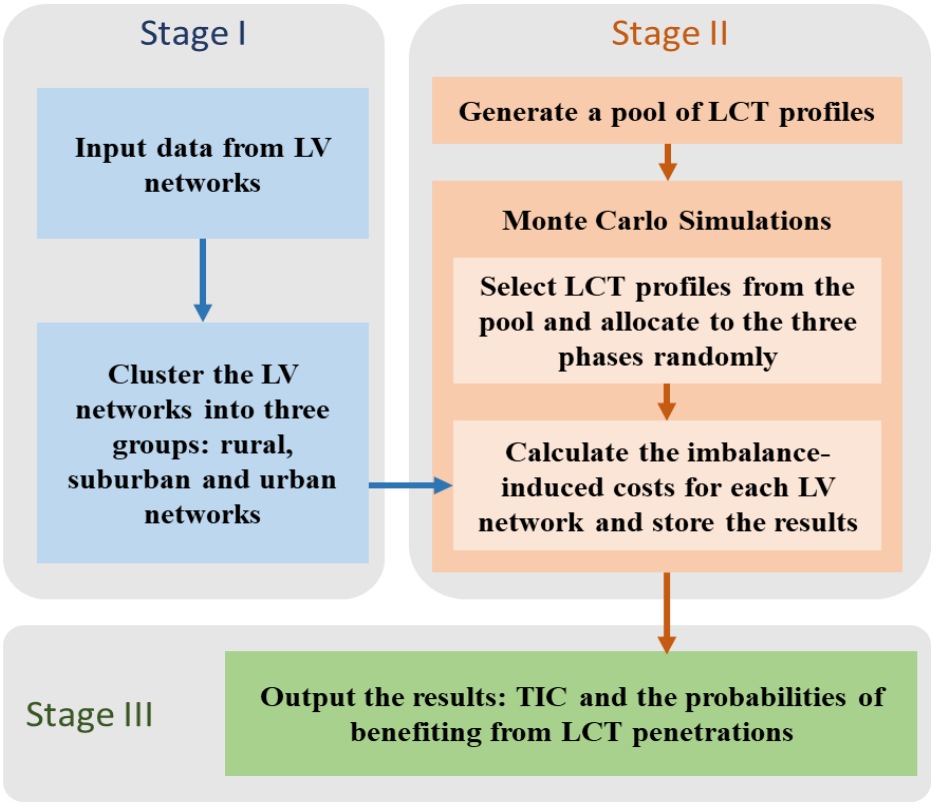


Figure 5-1. Overview of the methodology

Stage I: Applying k-means clustering method to group the 800 data-rich networks into three clusters, i.e., urban, suburban and rural.

Stage II: A pool of 1000 EV charging profiles and 1000 PV generating profiles are generated (detailed in Section 3.2). For each penetration level, the LCT profiles are randomly selected from the pool and randomly allocated to the three phases using Monte Carlo simulation. Finally, the imbalance-induced costs are calculated for each LV network under every LCT penetration level. This process iterates for 100 times to perform Monte Carlo analysis.

Stage III: The outputs from Stage II are ARC and AELC for each network. They work as input for Stage III. The TIC is calculated for each network and the probabilities of being beneficial from LCT penetration are analysed. If one network benefits from LCT penetration, it means that this network has lower TIC with LCT penetration compared to that without LCT penetration. The TICs under each LCT penetration level are compared and the conclusions of which LCT penetration level has higher probabilities of being beneficial can be drawn.

3. LV networks and LCT profiles

3.1. LV networks

In this paper, 800 representative data-rich LV networks from the "Low Voltage Network Templates" project [25] are used. These networks are located within the business area of a UK DNO (Western Power Distribution) and cover various geographical areas with different customer types, i.e., domestic, commercial and industrial customers. For example, Cardiff is representative of urban areas that contain large amounts of commercial load; Monmouthshire is a representative for the rural area [25]. These 800 networks cover various customer types and geographical areas (urban, suburban, and rural areas) [24].

3.2. LCT profiles

A pool of 1000 slow charging residential EV profiles is created considering the battery and the probability distributions of connection times and energy requirements [26]. The highest probability of connecting time happens at 6.30 p.m. and 10.30 p.m. [26]. The highest probability of energy requirement is 8-9 kWh [26]. Slow charging (3kW) is a popular type of charging for UK residential customers. According to [27], 75% of total annual EV demand is charged at the residential side. Therefore, all the EV batteries are assumed to be a common type, i.e., Nissan Leaf (3kW and 24kWh) [11].

A pool of 1000 residential PV generating profiles is generated considering various installation sizes of PV systems and the sun irradiances. It is assumed that all the PV systems receive the same sun irradiances. According to [28], the residential PV systems have seven different sizes and the size of 4 kW is the most popular choice (37% of the total installation). Therefore, the probabilities of PV system sizes for the pool are shown in Table 5-1.

Table 5-1. Probabilities of PV system sizes [28]

Size (kW)	1	1.5	2	2.5	3	3.5	4
Probability	0.01	0.08	0.13	0.14	0.14	0.12	0.37

Figure 5-2 demonstrates loads of a urban substation 513503 with 20% LCT penetration for one typical day (24 hours). It shows the total traditional load, EV load and PV generation for the three phases. The PV generations are shown as negative values.

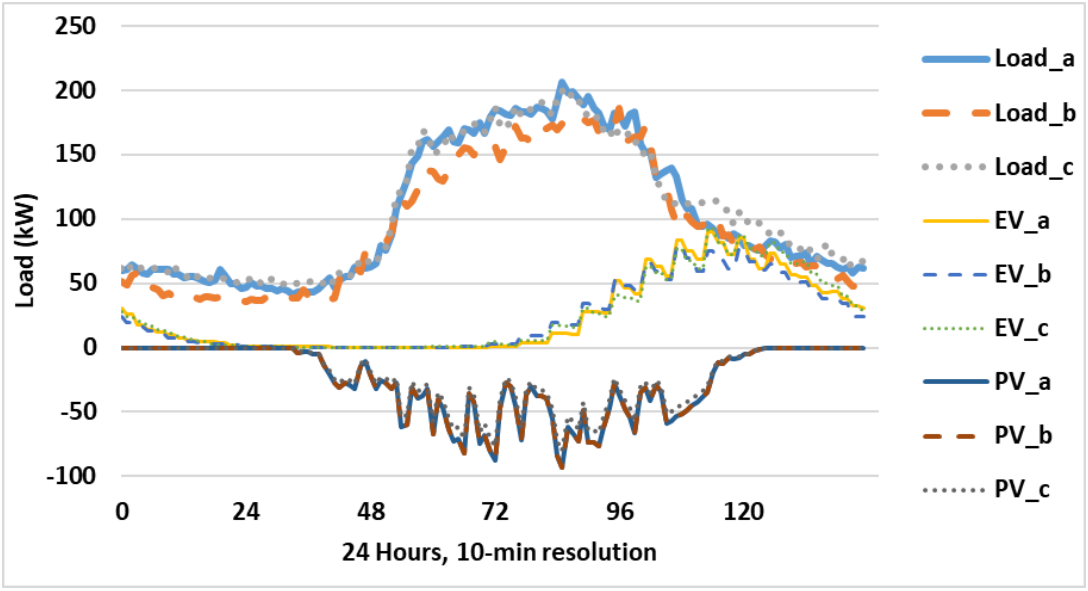


Figure 5-2. Loads of substation 513503 with 20% LCT penetration

3.3. The imbalance-induced costs

3.3.1. The additional reinforcement cost (ARC):

Degree of power imbalances (DPIB) is a common factor for measuring the severity of power imbalances. It is used as a guidance for phase swapping [22]. The DPIB is the main components of the ARC. The present value of the ARC is detailed in [4]

$$ARC \approx 3k_f DPIB_f + k_t DPIB_t \tag{5-1}$$

where $k_\chi = Asset_\chi \cdot (1+d)^{\frac{logU_N}{\log(1+r)}} \cdot \frac{\log(1+d)}{\log(1+r)}$, $\chi \in \{f, t\}$

$$DPIB_f = \frac{max\{P_\emptyset\} - \frac{P_t}{3}}{P_t} \quad \emptyset \in \{A, B, C\}, \qquad DPIB_t = \frac{P_N}{P_t}$$

$$DPIB_f = \frac{max\{P_\emptyset\} - \frac{P_t}{3}}{P_t} \quad \emptyset \in \{A, B, C\}, \qquad DPIB_t = \frac{P_N}{P_t}$$

where $DPIB_f$ and $DPIB_t$ are the degree of phase imbalance for main feeders and LV transformers, respectively. $Asset_\chi$ is the future asset reinforcement cost, where subscript χ can be either f (feeder) or t (transformer); d is the discount rate; U_N is the asset utilization rate and r is the load growth rate; P_t is the total power of three phases when the maximum phase power occurs and $P_\emptyset$ is the power on phase $\emptyset$; P_N is neutral line power.

3.3.2. The additional energy loss cost (AELC):

The imbalanced-induced energy loss contains two components: the energy loss caused by the neutral line current [29] and the additional transformer copper loss [30]. Therefore, the AELC is the sum of these two components:

$$AELC = \left(E_{loss} + E_{t_i}\right) \times \pi \tag{5-2}$$

where E_{t_i} is the energy loss caused by the additional transformer copper loss; E_{loss} is the energy loss and π is the energy price. E_{t_i} and E_{loss} will be explained in (5-3) and (5-4), respectively.

The majority of the UK's LV distribution networks follow the Terre-Neutral (TN-S) systems [31]. Therefore, the energy loss caused by the neutral line current is calculated considering the TN-S earthing system [31] in this paper.

The estimation of energy loss caused by the neutral line current is given in [29]

$$E_{loss} = \sum_{t=1}^{N_t} {I_{nlc}}^2(t) \cdot R_n \cdot \Delta t \tag{5-3}$$

$$\text{where } I_{prc}(t) = \sqrt{I_A^2(t) + I_B^2(t) + I_C^2(t) - I_A(t)I_B(t) - I_B(t)I_C(t) - I_A(t)I_C(t)}$$

where $I_A(t), I_B(t) and\ I_C(t)$ are current values for the phases A, B and C at time t, respectively; $I_{nlc}(t)$ denotes the neutral line current at time t; R_n denotes the neutral wire resistance. N_t is the number of hours within the year.

Phase imbalance increases the transformer copper loss compared to phase balanced scenario. The transformer copper loss under the balanced case is given in [30]

$$E_{trans} = 3\sum_{t=1}^{N_t} I^2(t) \cdot R_w \cdot \Delta t \tag{5-4}$$

where $I(t)$ is the balanced phase current at time t and R_w is the resistance of the transformer winding; N_t is the number of hours within a year.

The transformer copper loss under the imbalanced case is also given in [30]

$$E_i = \sum_{t=1}^{N_t} \left({I_A}^2(t) + {I_B}^2(t) + {I_C}^2(t) \right) \cdot R_w \cdot \Delta t \tag{5-5}$$

where $I_A(t), I_B(t) and\ I_C(t)$ are current values for the phases A, B and C at time t, respectively; R_w and N_t are defined in (5-4).

As a result, the imbalance-induced transformer copper loss is:

$$E_{t_i} = E_i - E_{trans} \tag{5-6}$$

where all variables are defined in(5-4) and (5-5).

3.3.3. The total imbalance-induced cost (TIC):

The TIC is a sum of the ARC and AELC:

$$TIC = ARC + AELC \quad (5\text{-}7)$$

where ARC and AELC are explained in (5-1) and (522), respectively.

The ARC is a present value for the long-term network investment while the AELC is the sum of day-to-day energy loss cost for a year. Considering TIC instead of ARC only helps DNOs avoid excessive energy losses caused by LCTs effectively.

4. Probabilistic impact assessment

4.1. Methodology

The developed methodology, as shown in Figure 5-3, analyses the probabilistic impacts of LCT penetration on imbalance-induced costs. It considers the uncertainties of EV charging energy requirement, PV system size, connection time and connection location through Monte Carlo simulations under different LCT penetration levels.

It is worthy to note that the substation monitors the total output from the transformer, which is the accumulated load consumption of the whole LV network. The imbalance-induced costs are calculated from the voltage and current data monitored by the substation. The network topology and load distribution are not necessary for this analysis as only the accumulated data from the substation side is required by this analysis. Therefore, the LCT penetration for a network is considered as the accumulated generation or consumption patterns of all the LCTs in the network. The main steps are:

1) Input data from 800 LV networks and cluster them into three groups, i.e., urban, suburban and rural. K-means clustering is used to group the networks by their annual peak current. This clustering process is done to analyse the impacts on different groups of networks.
2) Generate a pool of 1000 EV charging profiles and 1000 PV generating profiles. The pool of EV charging profiles follows the probability distributions of connection times and energy requirements [26]. The pool of PV generation profiles considers the installed sizes and sun irradiances [28]. The detailed process of generating the pools for LCT profiles is explained in Section IV-B.
3) Increase the LCT penetration level from 0% to 100% with a step of 10%. For the both PV and EV scenario, the penetration levels of these two LCTs increase at the same time. For example, if EV penetration level is 20%, the PV penetration level is 20% as well. The LCT penetration level is defined as the percentage of energy required or generated by LCT over the total traditional

passive load consumption. Increasing the LCT penetration level from 0% to 100% aims to cover a wide range of possible situations for the future. Although the 100% LCT penetration is very unlikely for the near future, it can be used as an extreme scenario for DNOs to analyze the impacts on phase power imbalance.

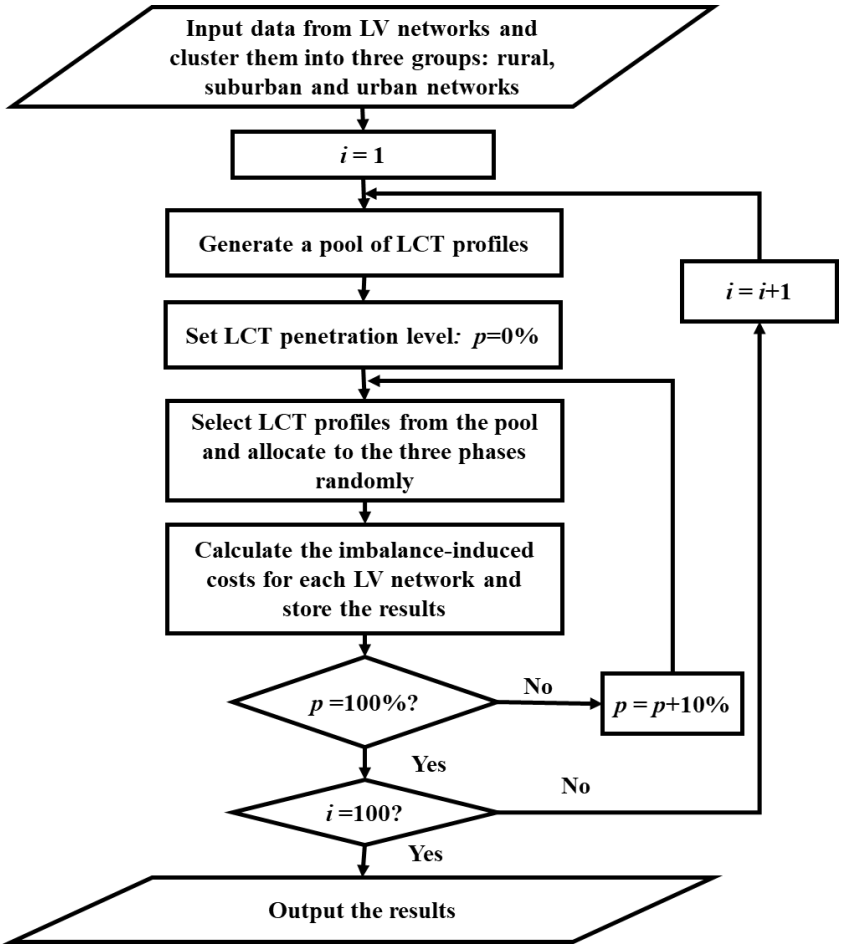


Figure 5-3. Flow chart of the methodology

4) Select LCT profiles from the pool according to each penetration level and allocate them to the three phases. Both the selection and allocation processes use the Monte Carlo method to embed uncertainties of LCTs.

5) Calculate ARC and AELC for each network and store the results. The AELC includes both energy loss cost caused by neutral line current and the energy loss cost caused by transformer copper loss. Note that the ARC is a long-run cost while the AELC is a day-to-day cost. Thus, the calculated ARC is a present value discounted from the future while the AELC represents the total

energy loss cost for a year. Besides, the TIC is calculated from ARC and AELC (using Equation 5-7) and it is used to evaluate the probability of a network to benefit from LCT penetrations.

6) Repeat the steps 100 times to account for uncertainties. Note that 10, 50, 200, 500 and 1000 times of Monte Carlo simulations had been run. However, the 10 and 50 times of simulation cannot cover the whole possible impacts. The rest of the simulation times have a very similar result. Thus, 100 times of simulation is chosen to show better results with shorter programming running time.

4.2. Probabilistic study

The imbalance-induced costs are calculated for each LV network under different LCT penetration levels. The 800 LV networks consist of urban (11.2%), suburban (44.4%), and rural (44.4%) networks. The average imbalance-induced costs for each group of networks are shown in the case study. A 95% confidence interval is considered while estimating the average costs in this analysis. The probabilistic study considers three scenarios, i.e., EV only, PV only and both EV and PV.

In this paper, the neutral wire resistance (R_n) is set as 0.244 Ω/km [29]. The winding resistances (R_w) are calculated from [32] and presented in Table 5-2.

Table 5-2. Parameters for different areas [32], [33]

Assets Area	Urban	Suburban	Rural
Transformer investment cost (k£)	26.4	16.1	5.8
Main feeder investment cost (k£/km)	67.2	16.4	15.0
Main feeder length (km)	0.2	0.3	0.4
No. of feeders connected from transformers	5	3.5	1.5
Winding resistance (Ω)	0.0163	0.0265	0.0413

To derive the ARC, the investment costs of the feeder and transformer are given in Table II. The discount value (d) is set as 5.0% [23] and [34]. The load growth rate (r) is set as 0.82% [35].

4.2.1. EV only scenario

Figure 5-4 shows how the ARC and AELC change with increasing EV penetrations for urban, suburban and rural networks. It can be seen that without EV penetration (i.e., EV penetration level is 0%), the rural networks have the largest ARC but least AELC. The reason is that the ARC is proportional to the DPIB while the AELC is influenced by loading level (as shown in equation (5-1) and (5-2)). The rural networks have the largest DPIB, but the lowest loading levels compared to suburban and urban networks.

Figure 5-4 shows that the ARC decreases with EV penetration while the AELC increases with EV penetration. For urban and suburban networks, the ARC decreases gradually. The ARC for rural networks decreases rapidly compared to other networks. It shows that EV penetration reduces the DPIB for all the networks. The DPIB in rural networks has the largest drop compared to suburban and urban networks.

It also shows that the AELC increases with EV penetration. The EV penetration level is defined as the percentage of energy required by EV over the total traditional passive load consumption. Therefore, the loading level is increased proportionally to the EV penetration. As discussed above, the AELC increases with loading level. The urban networks have the largest passive load consumption compared to rural and suburban networks. Thus, the urban networks have the most significant increase in loading level. Consequently, the AELC of urban networks increases dramatically while increasing EV penetration.

When the EV penetration level exceeds 60%, the urban networks' AELC becomes higher than the ARC. When the EV penetration level reaches 100%, the suburban networks' AELC catches up with the ARC and has the trend to keep increasing to exceed the ARC.

Figure 5-5 shows the average of the TIC for rural, suburban and urban networks. In rural networks, the TIC decreases as the EV penetration level increases. In suburban networks, the TIC reduces as EV penetration level increases up to 50% and stabilises after 50%. In urban networks, the TIC decreases as EV penetration level increases up to 50% and increases after 50%.

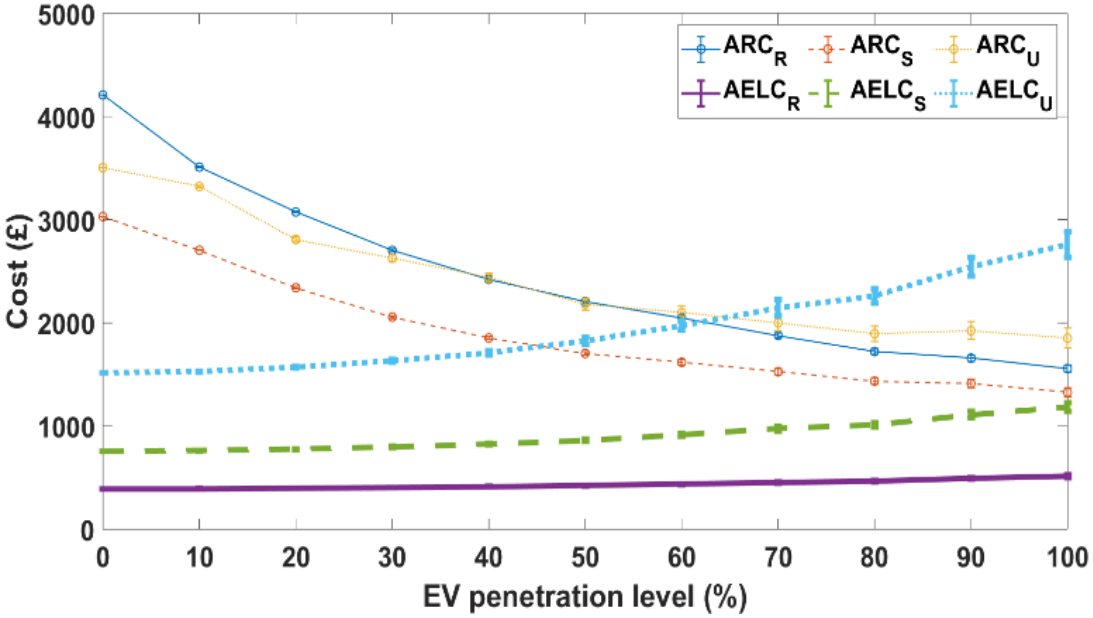


Figure 5-4. Variation of the average ARC and AELC of urban, suburban and rural networks with EV penetration

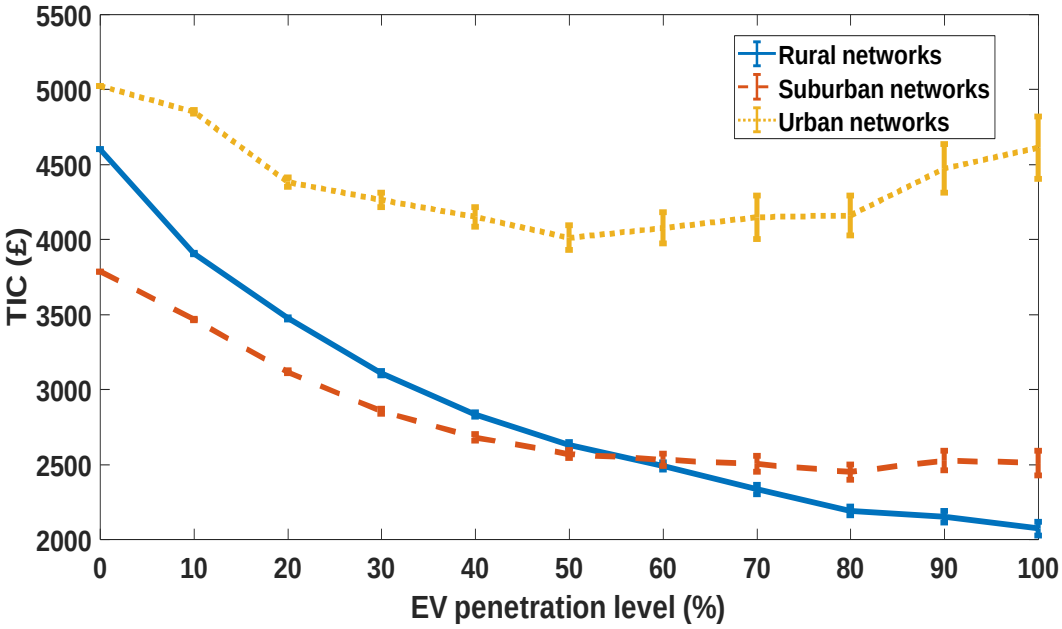


Figure 5-5. Variation of average TIC of urban, suburban and rural networks with EV penetration

It indicates that considering the full imbalance-induced cost, i.e., ARC and AELC, 50% of EV penetration brings the maximum benefits for the urban networks. The suburban networks gain more benefit from EV penetration that is larger than 50%. The benefits for rural networks increase with the EV penetration level.

4.2.2. PV only scenario

Figure 5-6 shows how the ARC and AELC change with increasing PV penetrations for urban, suburban and rural networks. As discussed above, the rural networks have the largest ARC because they have the largest DPIB compared to other networks. The loading level of rural networks is the lowest, which leads to the lowest AELC. The PV penetration has minor influences on both ARC and AELC as the values of ARC and AELC have only increased slightly with PV penetration.

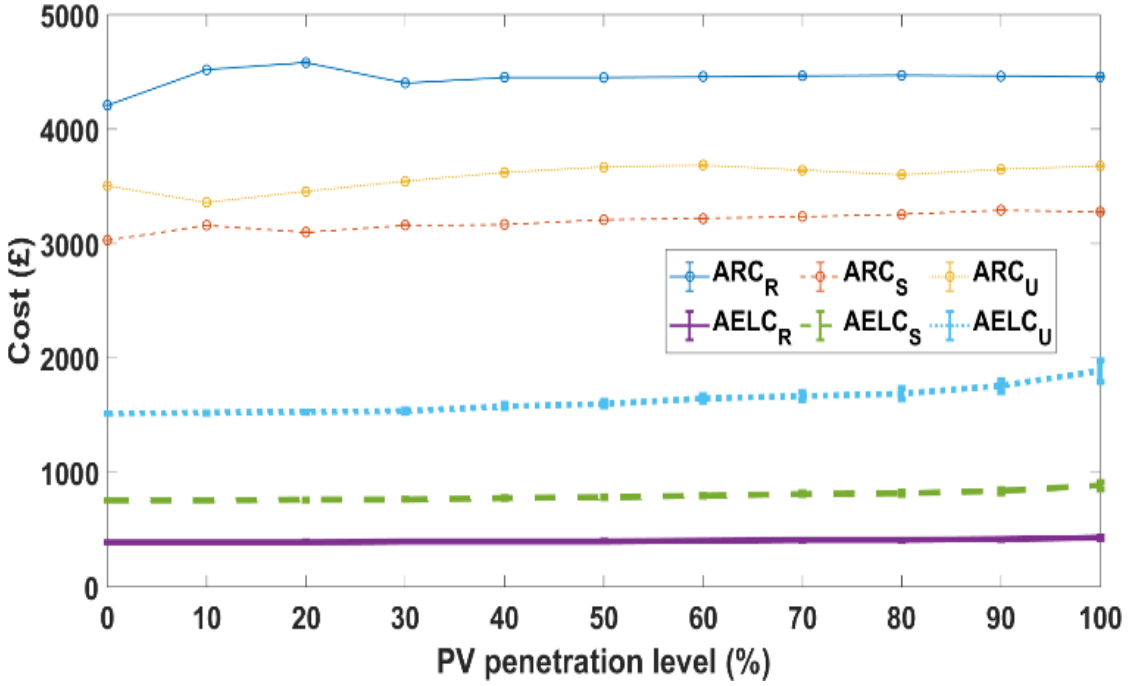


Figure 5-6. Variation of average ARC and AELC of urban, suburban and rural networks with PV penetration

The reason for this phenomenon is that PV generation mainly changes the DPIB in the noontime because of the nature of the solar system. However, the ARC is decided by the maximum DPIB throughout the whole year. Thus, the impacts of PV penetration on the ARC is insignificant. A detailed discussion of different impacts on DPIB is given in Section 6.5.

The increase of PV generation only reduces the loading level in the noontime. However, the AELC is an accumulated value of a whole year. Thus, the impacts of PV penetration on the AELC is insignificant.

Because of the minor changes in both ARC and AELC, there are insignificant increases of TIC for all the networks (as shown in Figure 5-7). However, such an increase of TIC is negligible comparing to other network operations.

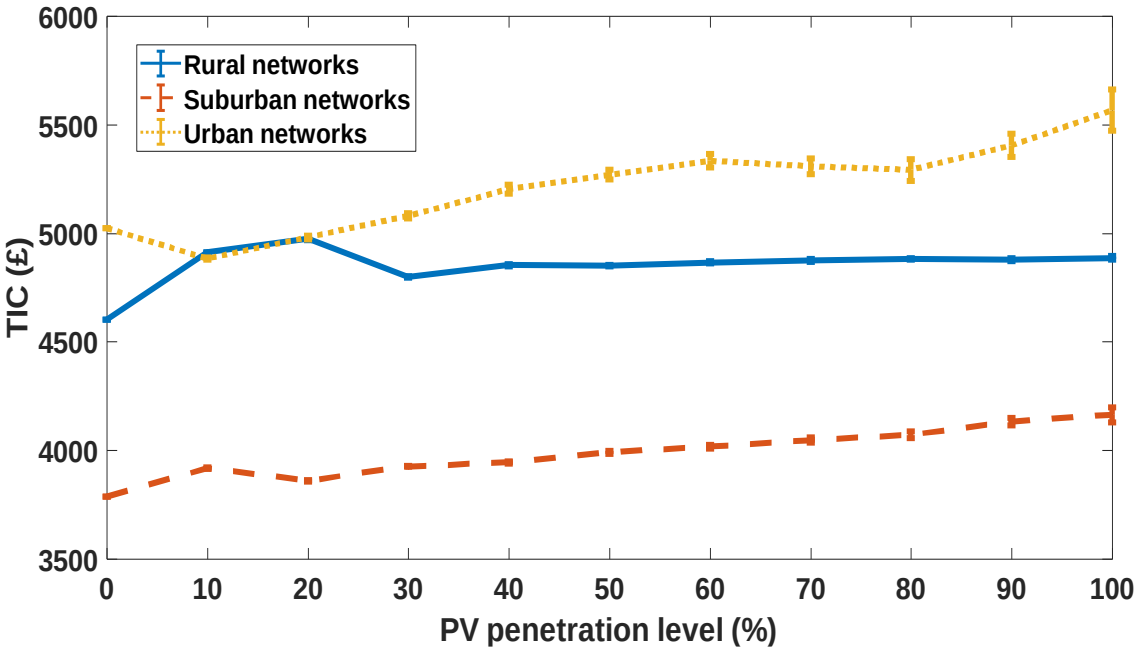


Figure 5-7. Variation of average TIC of urban, suburban and rural networks with PV penetration

4.2.3. Both EV and PV scenario

The third scenario considers both EV and PV at the same time. In the following content, 'EV and PV' are referred to as 'LCT' for simplicity.

For this scenario, EV and PV are considering having the same penetration level, i.e., if the LCT penetration level is 10%, it means that both the EV and PV have a penetration level of 10%. The LCTs are randomly selected from the pool using Monte Carlo and randomly allocated to the three phases using norm distribution.

Figure 5-8 shows that the ARC decreases with LCT penetration while the AELC increases with LCT penetration. For urban and suburban networks, the ARC decreases gradually. The ARC for rural networks decreases rapidly compared to other networks. It shows that EV penetration reduces the DPIB for all the networks. The DPIB in rural networks has the largest drop compared to suburban and urban networks.

It also shows that the AELC increases with LCT penetration. Though the total amount of EV consumption equals the PV generation, PV generation mainly reduces the loading level in the noontime. In contrast, EV consumption has possibilities to increase the loading level at any time of the day. Therefore, the AELC has raised because of the increasing of LCT connections. Among all

the networks, the urban networks have the largest increase in loading level. Consequently, the AELC of urban networks increases with increasing LCT penetration.

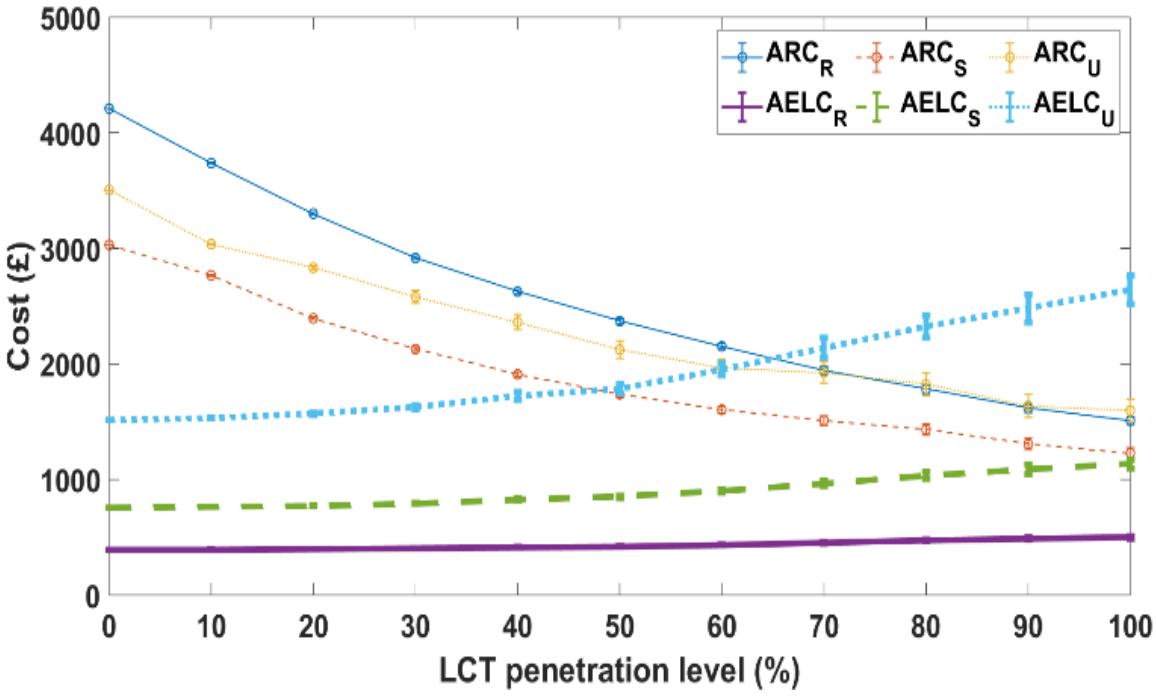


Figure 5-8. Variation of average ARC and AELC of urban, suburban and rural networks with LCT penetration

When the EV penetration level exceeds 65%, the urban networks' AELC becomes higher than the ARC. When the LCT penetration level reaches 100%, the suburban networks' AELC catches up with the ARC and has the trend to keep increasing to exceed the ARC.

Figure 5-9 shows that, in urban networks, the TIC decreases as the LCT penetration level increases up to 50% and decreases after 50%. In suburban networks, the TIC reduces as LCT penetration level increases up to 60% and stabilises after 60%. In rural networks, the TIC decreases as the LCT penetration level increases.

Therefore, to balance the long-run investment cost and day-to-day energy loss cost, 50% - 60% of LCT penetration brings the maximum benefits for the urban networks. The suburban networks will gain more benefits from LCT penetration that is larger than 50%. The rural networks will always benefit from LCT penetration.

To further understand the possible influences of LCT penetration. The probability (with 95% confidence) of networks to have reduced TIC with LCT penetration is calculated for each penetration level. This demonstrates the benefits from LCT penetration.

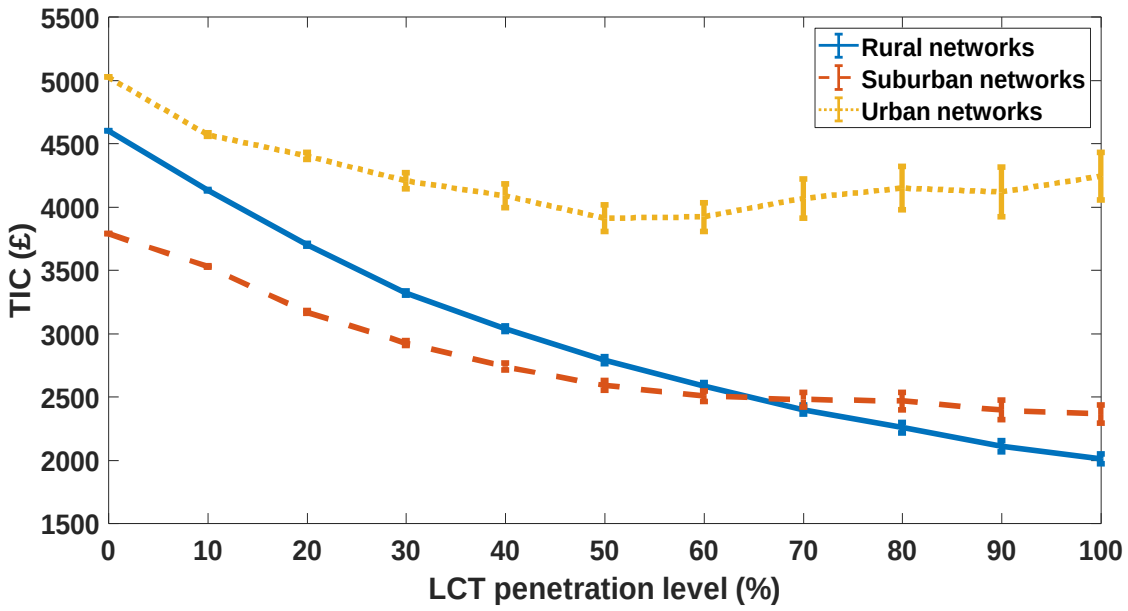


Figure 5-9. Variation of average TIC of urban, suburban and rural networks with LCT penetration

Table 5-3, Table 5-4 and Table 5-5 show the probabilities of having reduced TIC with LCT penetration from the Monte Carlo simulations in rural, suburban and urban networks, respectively. The colourmap indicates the ratio of networks applicable to each scenario. For example, with 20% LCT penetration, 40% of rural networks, 43% of suburban networks and 40% of urban networks have more than 0.5 probability to benefit from EV penetration. It is also shown that 60% of LCT penetrations have higher probabilities of bringing benefits for the majority of LV networks compared to that of 50% of LCT penetration.

Table 5-3. The probability of being beneficial from LCT penetration for rural networks

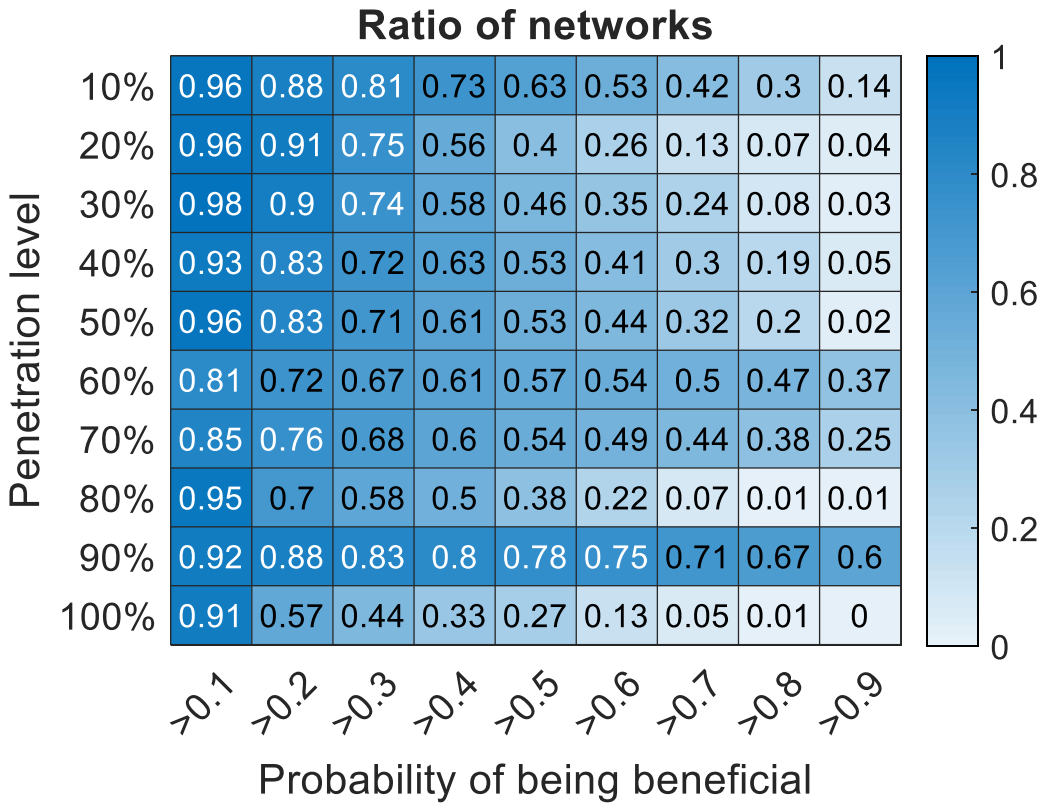


Table 5-4. The probability of being beneficial from LCT penetration for suburban networks

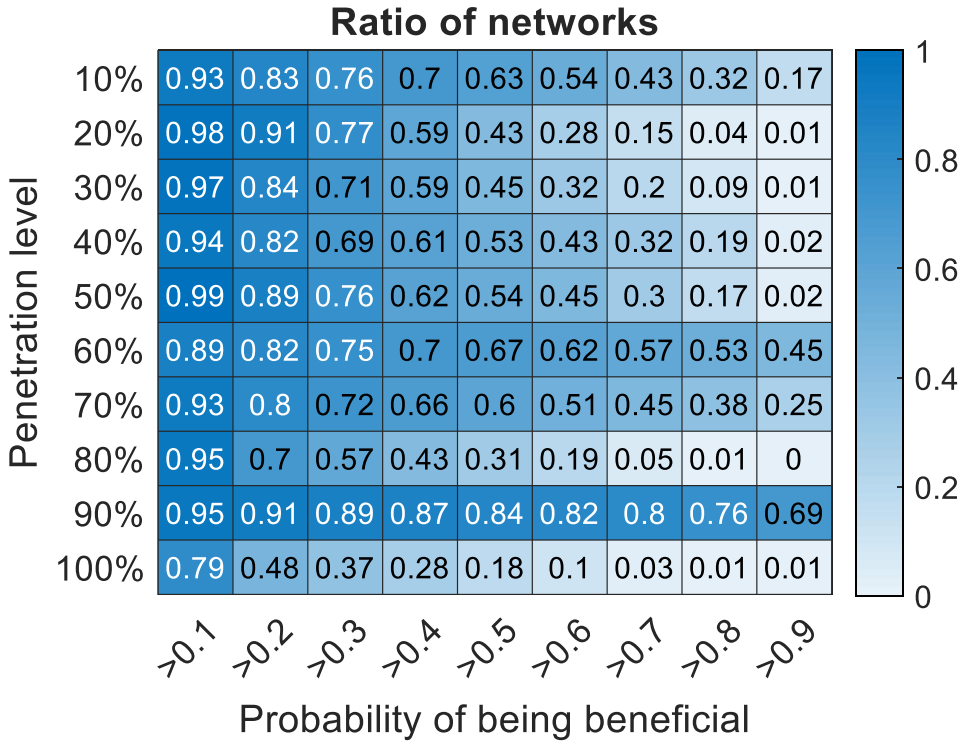
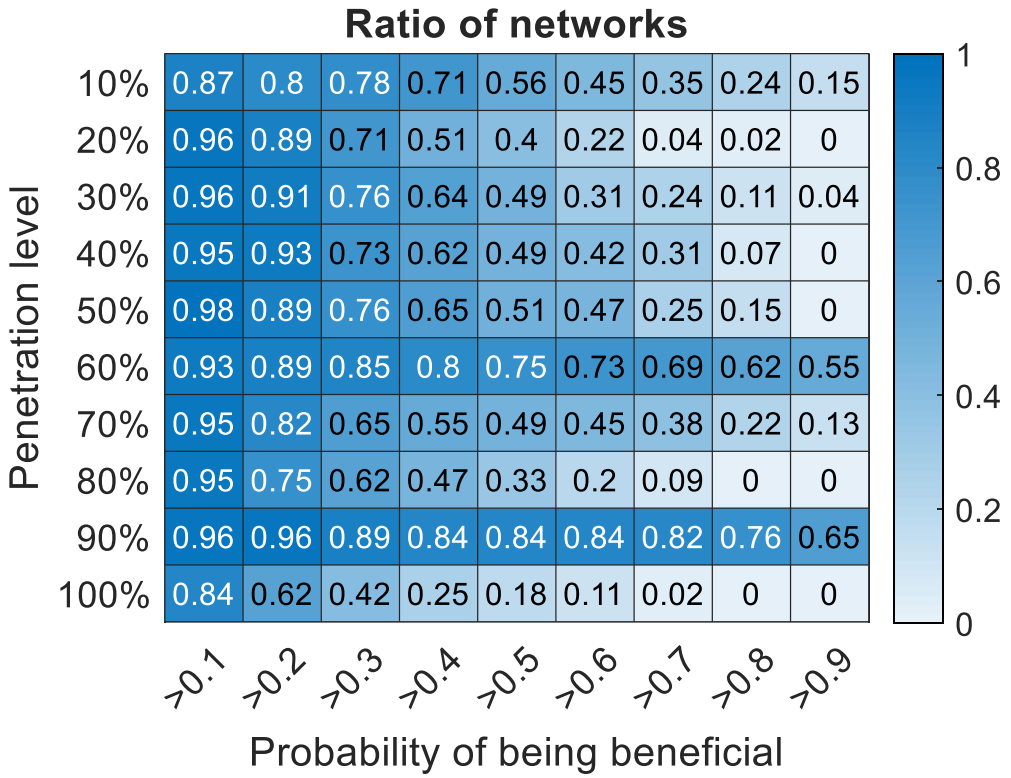


Table 5-5. The probability of being beneficial from LCT penetration for urban networks



4.2.4. Worst-case scenario

A worst-case scenario of considering both EV and PV at the same time is simulated. In the worst-case scenario, all the EVs are connected to the phase which has the largest total annual demand while all the PV generations are connected to the phase which has the smallest total annual demand. Again, for this scenario, EV and PV are considering having the same penetration level.

Figure 5-10 shows that both ARC and AELC increase with LCT penetration for the reason that the worst-case scenario increases the DPIB of the LV networks by increasing LCT penetration levels. Urban networks show the largest increase of both ARC and AELC, while rural networks show the smallest increases. The AELC increases dramatically under the worst-case scenario. When the LCT penetration level is 100%, the AELC for urban networks is about 2.6 times and 10 times of that for suburban and rural networks. Though the increase of ARC is smaller than that of AELC, the ARC for the urban network is about 1.7 times that for rural networks with 100% LCT penetration.

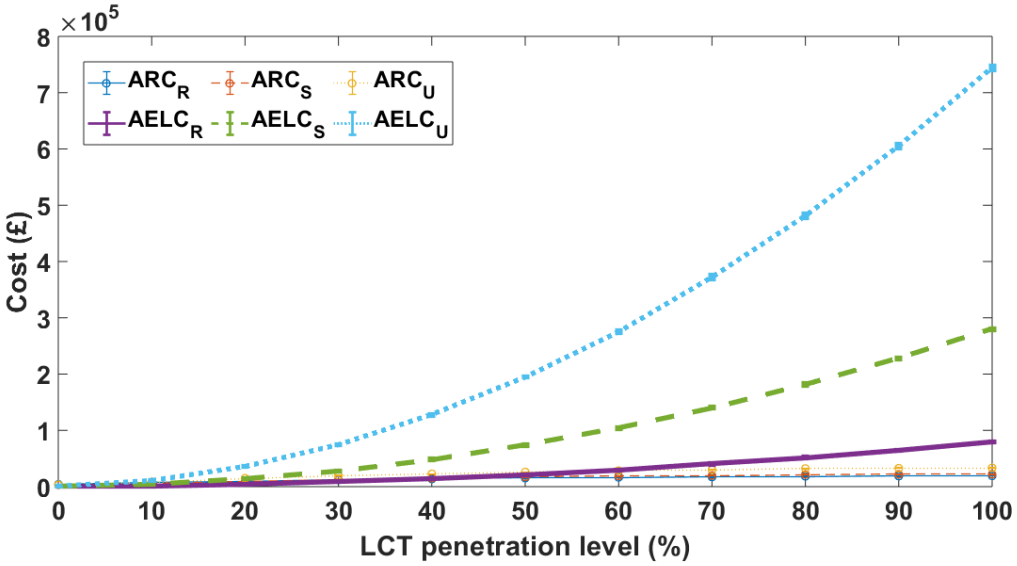


Figure 5-10. Variation of average ARC and AELC of urban, suburban and rural networks under the worst scenario

Figure 5-13 shows that the TIC increases significantly with the LCT penetration under the worst-case scenario. Comparing to the TIC results of the random scenario (as shown in Figure 5-9), it can be seen that the worst-case scenario TIC is about 160 times larger with 100% LCT penetrations. While the LCT penetration is 50%, the worst-case scenario TIC is about 50 times larger than that of the random scenario.

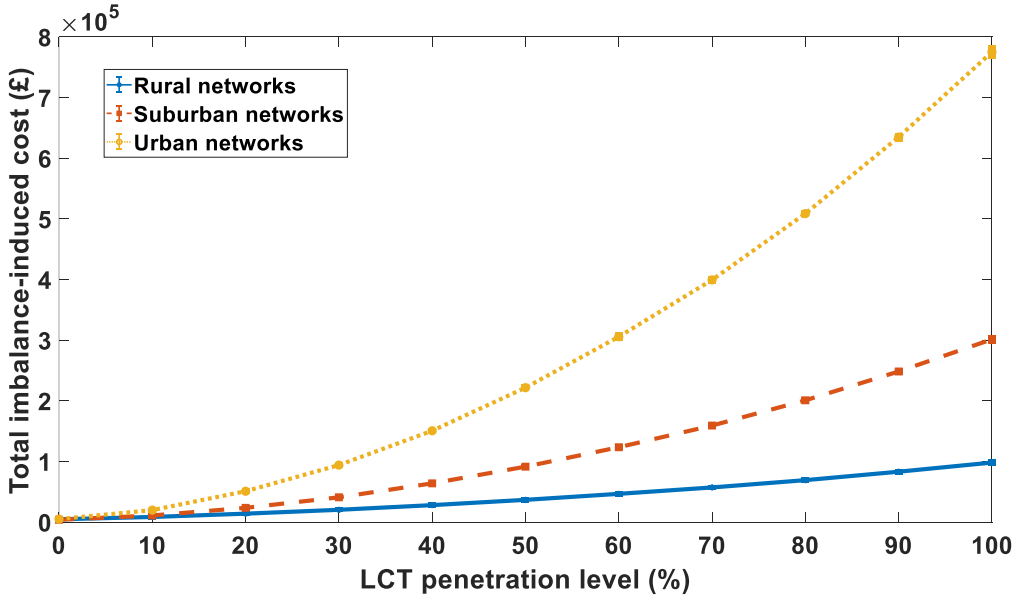


Figure 5-11. Variation of average TIC of urban, suburban and rural networks under the worst scenario

5. Discussion

As discussed previously, DPIB contributes to the ARC while neutral line current contributes to the AELC. Figure 5-12 shows the changes of DPIB in different scenarios for all the networks, i.e., no LCT scenario, EV only scenario, PV only scenario and both EV and PV scenario. The penetration level is 100% for all the scenarios. Rural networks have the largest DPIB while urban networks have the smallest DPIB. Compared to the scenario of no LCT penetration, the EV penetration scenario and LCT penetration scenario both reduce DPIB significantly. In contrast, the PV penetration scenario has very small impacts on the DPIB. The DPIB in LCT penetration scenario is slightly smaller than that of the EV penetration scenario.

The reason is that DPIB is defined as the ratio of the deviation of the maximum power ($max\{P_{\emptyset}\}$) from the average power ($\frac{P_t}{3}$) to the total power (P_t) along the whole year (as explained in Chapter 2). In LV distribution networks, the maximum load demand mainly occurs in the evening time around 6:30 p.m., as shown in Figure 5-2. This is also the time that majority EVs are connected to the grid to begin charging. Consequently, the EV connections have great possibilities of increasing the $\frac{P_t}{3}$ and as a result, reduce the DPIB. Oppositely, PV generations mainly generate energy during the noontime. As a result, the PV penetration has low possibilities of reducing the DPIB.

It can be seen that the DPIB in the LCT scenario is slightly smaller than that of the EV scenario. This is because all the EVs and PV generation are random allocated to the three phases, there exits situations that the EVs are connected to the lightest loaded phase while the PV generation is connected to the heaviest loaded phase. As a result, the DPIB of LCT scenario becomes lower than that of the EV only scenario.

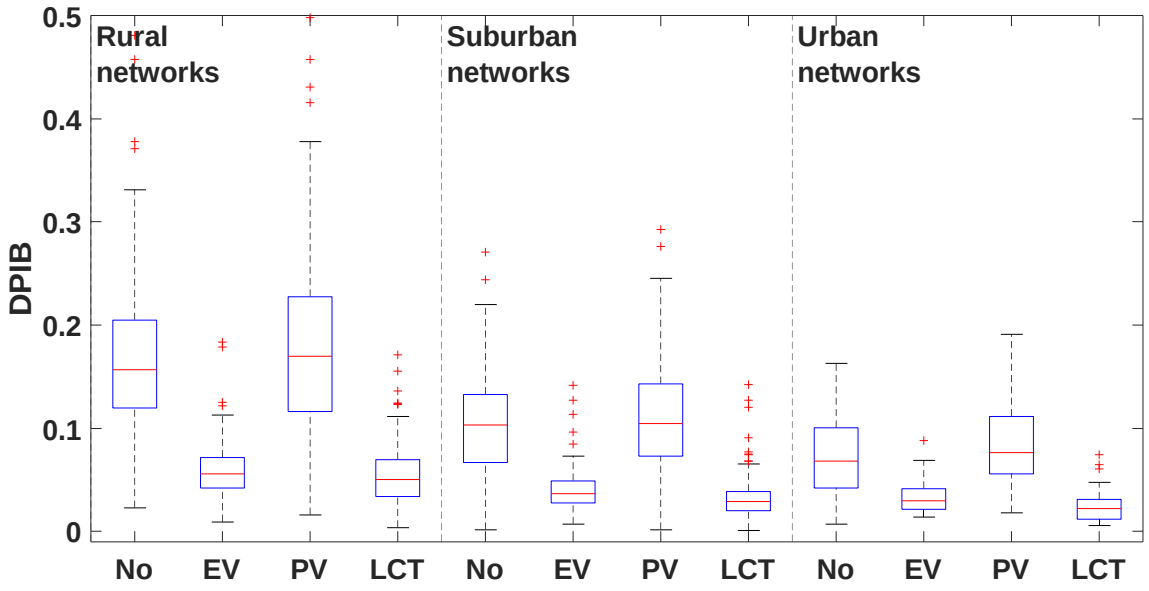


Figure 5-12. DPIB in different scenarios

Figure 5-13 shows the changes of the neutral line current (I_{nlc}) in different LCT scenarios for all the networks. The penetration level is 100% for all the scenarios. Urban networks have the largest I_{nlc} while rural networks have the smallest I_{nlc}. Compared to the scenario of no LCT penetration, the EV penetration scenario increases I_{nlc} relatively significantly compared to the PV and LCT penetration scenario. The I_{nlc} in EV penetration scenario is the largest among all the scenarios.

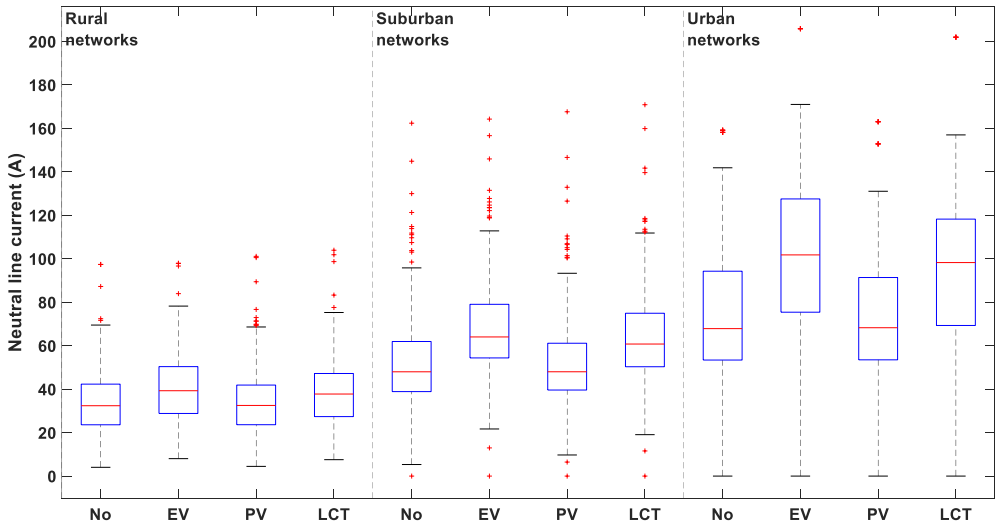


Figure 5-13. Neutral line current (I_{nlc}) in different scenarios

Consequently, the EV penetration scenario has larger impacts on the AELC compared to other scenarios. Oppositely, PV and LCT penetration scenarios have minor impacts on the AELC.

Again, there exits situations that the EVs are connected to the lightest loaded phase while the PV generation is connected to the heaviest loaded phase. As a result, the I_{nlc} is slightly lower of the LCT scenario than that of the EV only scenario.

This section assumes that the power factors for the three are the same with the value of 0.9. Assuming such a power factor, if the active power is rebalanced, the reactive power is automatically rebalanced. Hence, the impact assessment needs to consider three-phase active power only.

6. Conclusions

A probabilistic impact analysis is performed to analyse the total imbalance-induced cost (TIC) in the low voltage (LV) networks with increasing penetrations of low carbon technologies (LCTs). The TIC includes both day-to-day energy loss cost and long-run network investment cost. The framework uses Monte Carlo simulations to account for the uncertainties associated with the LCTs. Full time-series data from 800 LV substations are used for the case studies.

The results show that the energy loss cost may exceed the network investment with penetration of single-phase LCTs. To balance the long-run investment cost and day-to-day energy loss cost, 60% of LCT penetration has the highest probability to bring the maximum benefits for the majority of the LV networks.

The developed impact assessment framework help DNOs understand the potential of benefits that LV networks can obtain from LCTs penetrations. Moreover, the developed framework can be used as a tool to perform a cost-benefit analysis for phase balancing solutions. Therefore, it guides the DNOs in investing phase balancing solutions to cope with the increasing LCT penetrations.

7. References

[1] K. Ma, R. Li, and F. Li, “Utility-Scale Estimation of Additional Reinforcement Cost from 3-Phase Imbalance Considering Thermal Constraints,” IEEE Transactions on Power Systems, vol. PP, no. 99, pp. 1-1, 2016.

[2] "Electricity distribution units and loss percentages summary," https://www.ofgem.gov.uk/sites/default/files/docs/2010/08/distribution-units-and-loss-percentages-summary.pdf.

[3] "Electricity Distribution Price Control Review Final Proposals - Incentives and Obligations," https://www.ofgem.gov.uk/ofgem-publications/46748/fp2incentives-and-obligations-final.pdf.

[4] K. Ma, R. Li, and F. Li, “Quantification of Additional Asset Reinforcement Cost From 3-Phase Imbalance,” IEEE Transactions on Power Systems, vol. 31, no. 4, pp. 2885-2891, 2016.

[5] Zhu, M. Y. Chow, and F. Zhang, “Phase balancing using mixed-integer programming,” IEEE Transactions on Power Systems, vol. 13, no. 4, pp. 1487-1492, 1998.

[6] Future Power System Architecture, Institution of Engineering and Technology, 2016.

[7] Open Networks Future Worlds, Energy Networks Association, 2018.

[8] "Future Energy Scenarios," 2019; http://fes.nationalgrid.com/media/1363/fes-interactive-version-final.pdf.

[9] Ul-Haq, C. Cecati, K. Strunz, and E. Abbasi, “Impact of Electric Vehicle Charging on Voltage Unbalance in an Urban Distribution Network,” Intelligent Industrial Systems, vol. 1, no. 1, pp. 51-60, 2015.

[10] Clement-Nyns, E. Haesen, and J. Driesen, “The Impact of Charging Plug-In Hybrid Electric Vehicles on a Residential Distribution Grid,” IEEE Transactions on Power Systems, vol. 25, no. 1, pp. 371-380, 2010.

[11] Navarro-Espinosa, and L. F. Ochoa, “Probabilistic Impact Assessment of Low Carbon Technologies in LV Distribution Systems,” IEEE Transactions on Power Systems, vol. 31, no. 3, pp. 2192-2203, 2016.

[12] Vega-Fuentes, and M. Denai, “Enhanced Electric Vehicle Integration in the UK Low-Voltage Networks With Distributed Phase Shifting Control,” IEEE Access, vol. 7, pp. 46796-46807, 2019.

[13] Quirós-Tortós, L. F. Ochoa, S. W. Alnaser, and T. Butler, “Control of EV Charging Points for Thermal and Voltage Management of LV Networks,” IEEE Transactions on Power Systems, vol. 31, no. 4, pp. 3028-3039, 2016.

[14] J. F. Franco, A. T. Procopiou, J. Quirós-Tortós, and L. F. Ochoa, “Advanced control of OLTC-enabled LV networks with PV systems and EVs,” IET Generation, Transmission & Distribution, vol. 13, no. 14, pp. 2967-2975, 2019.

[15] T. R. Ricciardi, K. Petrou, J. F. Franco, and L. F. Ochoa, “Defining Customer Export Limits in PV-Rich Low Voltage Networks,” IEEE Transactions on Power Systems, vol. 34, no. 1, pp. 87-97, 2019.

[16] J. D. Watson, N. R. Watson, D. Santos-Martin, A. R. Wood, S. Lemon, and A. J. V. Miller, “Impact of solar photovoltaics on the low-voltage distribution network in New Zealand,” IET Generation, Transmission & Distribution, vol. 10, no. 1, pp. 1-9, 2016.

[17] S. Hashemi, and J. Østergaard, “Methods and strategies for overvoltage prevention in low voltage distribution systems with PV,” IET Renewable Power Generation, vol. 11, no. 2, pp. 205-214, 2017.

[18] Schwanz, F. Möller, S. K. Rönnberg, J. Meyer, and M. H. J. Bollen, “Stochastic Assessment of Voltage Unbalance Due to Single-Phase-Connected Solar Power,” IEEE Transactions on Power Delivery, vol. 32, no. 2, pp. 852-861, 2017.

[19] R. Torquato, D. Salles, C. O. Pereira, P. C. Meira, and W. Freitas, “A Comprehensive Assessment of PV Hosting Capacity on Low-Voltage Distribution Systems,” IEEE Transactions on Power Delivery, vol. 33, no. 2, pp. 1002-1012, 2018.

[20] Dubey, and S. Santoso, “On Estimation and Sensitivity Analysis of Distribution Circuit's Photovoltaic Hosting Capacity,” IEEE Transactions on Power Systems, vol. 32, no. 4, pp. 2779-2789, 2017.

[21] Navarro-Espinosa, and P. Mancarella, “Probabilistic modeling and assessment of the impact of electric heat pumps on low voltage distribution networks,” Applied Energy, vol. 127, pp. 249-266, 2014.

[22] W. Kong, K. Ma, and Q. Wu, “Three-Phase Power Imbalance Decomposition Into Systematic Imbalance and Random Imbalance,” IEEE Transactions on Power Systems, vol. 33, no. 3, pp. 3001-3012, 2018.

[23] Ma, R. Li, and F. Li, “Utility-Scale Estimation of Additional Reinforcement Cost From Three-Phase Imbalance Considering Thermal Constraints,” IEEE Transactions on Power Systems, vol. 32, no. 5, pp. 3912-3923, 2017.

[24] W. Kong, K. Ma, L. Fang, R. Wei, and F. Li, “Cost-Benefit Analysis of Phase Balancing Solution for Data-scarce LV Networks by Cluster-Wise Gaussian Process Regression,” IEEE Transactions on Power Systems, pp. 1-1, 2020.

[25] R. Li, C. Gu, F. Li, G. Shaddick, and M. Dale, “Development of Low Voltage Network Templates Part I: Substation Clustering and Classification,” Power Systems, IEEE Transactions on, vol. 30, no. 6, pp. 3036-3044, 2015.

[26] P. Richardson, M. Moran, J. Taylor, A. Maitra, and A. Keane, Impact of electric vehicle charging on residential distribution networks: An Irish demonstration initiative, 2013.

[27] T. Dodson, and S. Slater, Electric vehicle charging behaviour study, National Grid ESO, 2019.

[28] Navarro, L. F. Ochoa, and D. Randles, “Monte Carlo-based assessment of PV impacts on real UK low voltage networks,” in 2013 IEEE Power & Energy Society General Meeting, 2013, pp. 1-5.

[29] Fang, K. Ma, R. Li, Z. Wang, and H. Shi, “A Statistical Approach to Estimate Imbalance-Induced Energy Losses for Data-Scarce Low Voltage Networks,” IEEE Transactions on Power Systems, pp. 1-1, 2019.

[30] O, O. O, and I. A;, “Evaluation Of Distribution System Losses Due To Unbalanced Load In Transformers A Case Study Of Guinness 15MVA, 33/11KV, Injection Substation And Its Associated 11/0.415kv Transformers In Benin City, Nigeria,” International Journal of Engineering Research & Technology, vol. 2, no. 3, March, 2013.

[31] Cronshaw, EARTHING: YOUR QUESTIONS ANSWERED, IEE Wiring Matters, 2005.

[32] S. Electric. "HV/LV distribution transformers," http://mt.schneider-electric.be/main/tfo/catalogue/an_iec.pdf.

[33] Y. Zhang, F. Li, Z. Hu, and G. Shaddick, “Quantification of low voltage network reinforcement costs: A statistical approach,” IEEE Transactions on Power Systems, vol. 28, no. 2, pp. 810-818, 2013.

[34] S. Sidhu, M. G. Pollitt, and K. L. Anaya, “A social cost benefit analysis of grid-scale electrical energy storage projects: A case study,” Applied Energy, vol. 212, pp. 881-894, 2018/02/15/, 2018.

[35] "Pathways for the GB Electricity Sector to 2030," https://www.energy-uk.org.uk/publication.html?task=file.download&id=5722

5.3. Additional Analysis and Discussions

This additional analysis discusses the impacts of LCT penetration on phase imbalance decomposition (as discussed in Chapter 3). The probabilities of having a different decomposition judgement results for all the networks with 100 Monte Carlo simulations are shown in Figure 5-14. The percentage of change means that with the penetration of LCTs, the probability that a network is classified as a different group compared to the case of without LCT penetration. For examples, network 536753 is classified as a definite-max scenario with no LCT penetration (as shown in Chapter 2). With 50% of LCT penetration, the probability of the network not being classified as a definite-max scenario is 7%. With 100% of LCT penetration, the probability goes to 29%. As can be seen from Figure 5-14, with the increase of LCT penetration, there is a higher probability of changing network scenarios, i.e., definite-max, definite-min, definite-order and random imbalance. Consequently, the corresponding decomposition results of systematic component and the random component will change.

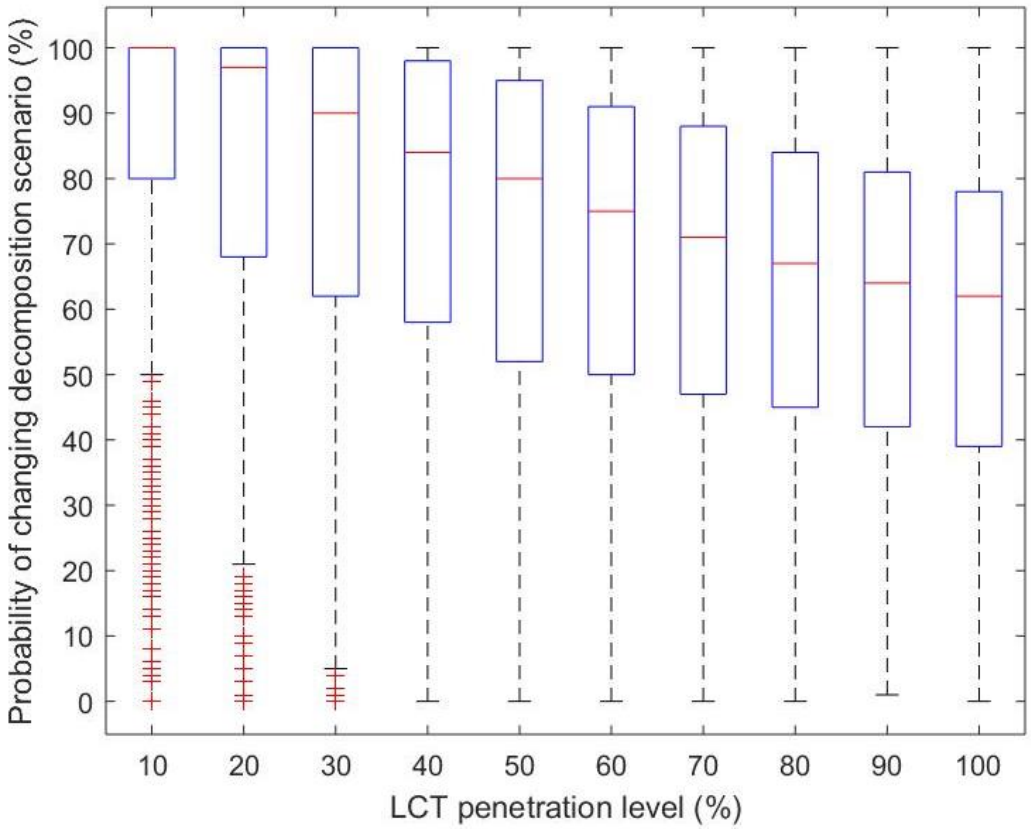


Figure 5-14. Probability of changing decomposition judgement results with different LCT penetration levels

Further investigations on the change of decomposition judgement results for each scenario are made. The change of scenario means the change of systematic and random imbalance components.

Figure 5-15 shows the probabilities for networks changing from the definite-max scenario to random imbalance scenario with changing LCT penetration levels. As can be seen, on average, the probability for networks changing from definite-max scenario to random imbalance scenario is 15%. The LCT penetration level has insignificant influences on the probability.

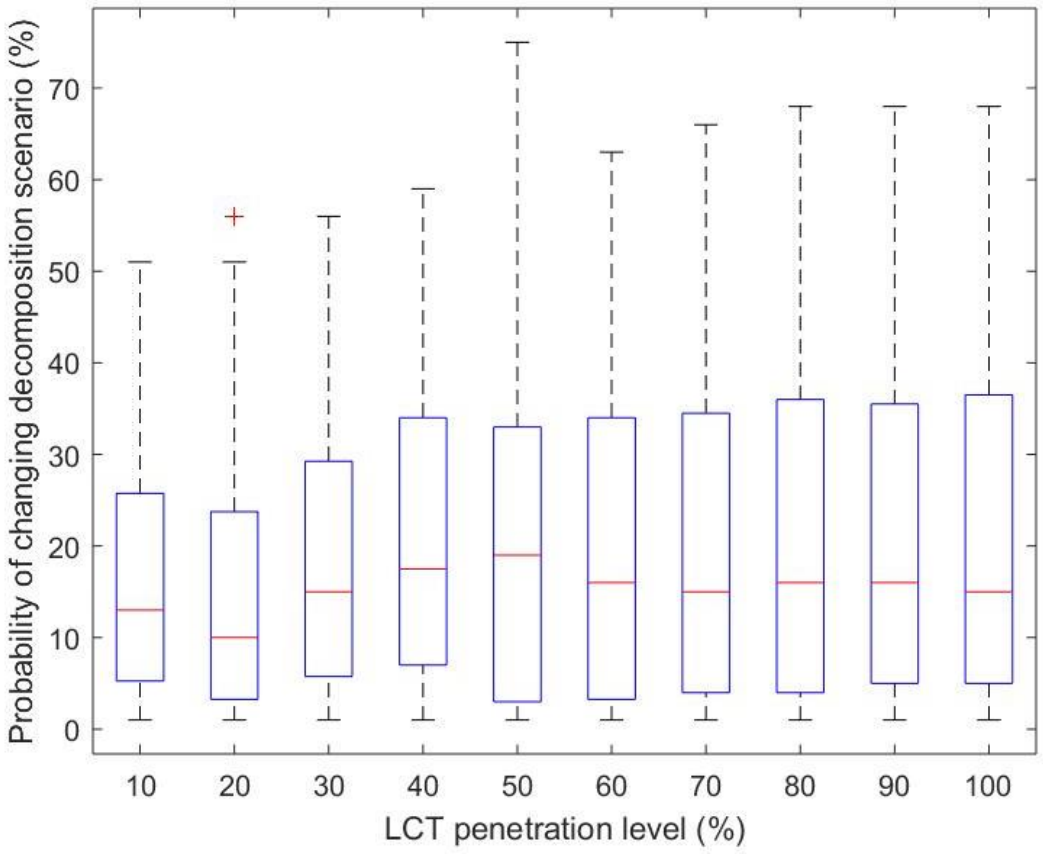


Figure 5-15. Probability of changing from definite-max scenario to random imbalance scenario with different LCT penetration levels

Figure 5-16 shows the probabilities for networks changing from the definite-min scenario to random imbalance scenario with changing LCT penetration levels. As can be seen, on average, the probability for networks changing from definite-min scenario to random imbalance scenario is 15%. The probability increases steadily with the LCT penetration level. The impacts of LCT penetration are similar for definite-max and definite-min scenarios.

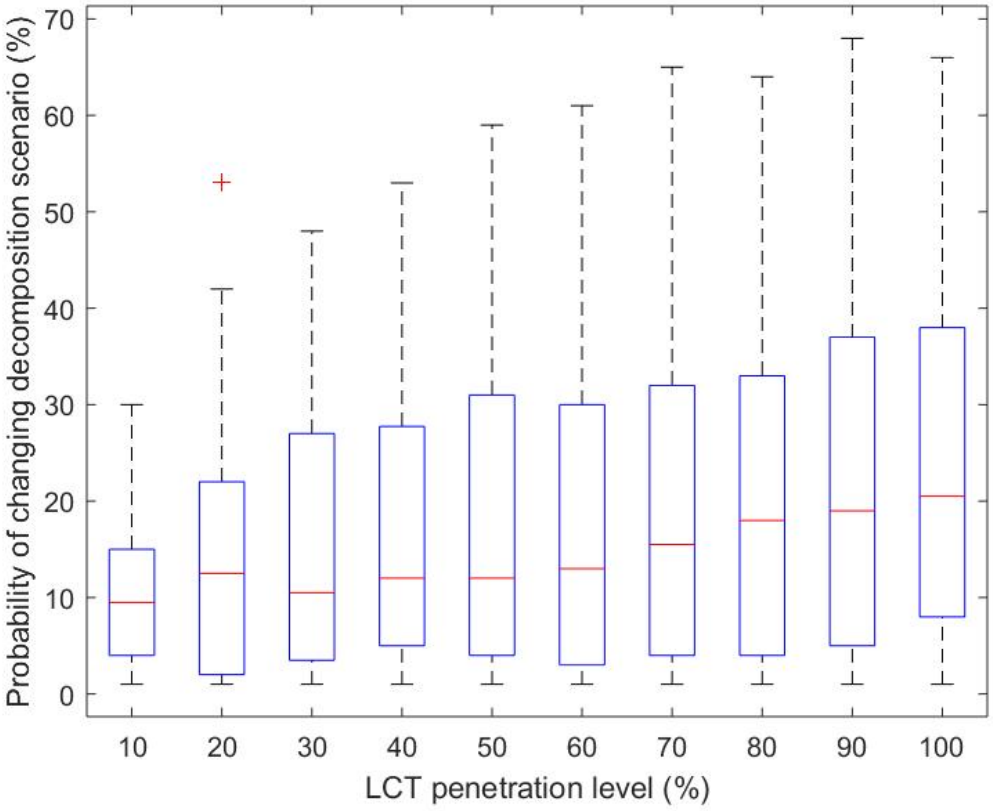


Figure 5-16. Probability of changing from definite-min scenario to random imbalance scenario with different LCT penetration levels

Figure 5-17 shows the probabilities for networks changing from the definite-order scenario to random imbalance scenario with changing LCT penetration levels. As can be seen, when the LCT penetration level is below 50%, no network has been changed to random imbalance scenario. The average probability is 3% for networks change to random imbalance scenario. The probability is much lower compared to that of the definite-max and definite-min scenario. This is because definite-order is a more restrict scenario compared to the other two. When LCT penetration increases, the decomposition judgement will identify the changing of the definite-order scenario to definite-max or definite-min scenario before random imbalance. The reason is that the aim of the decomposition is to find the maximum systematic component in the phase imbalance.

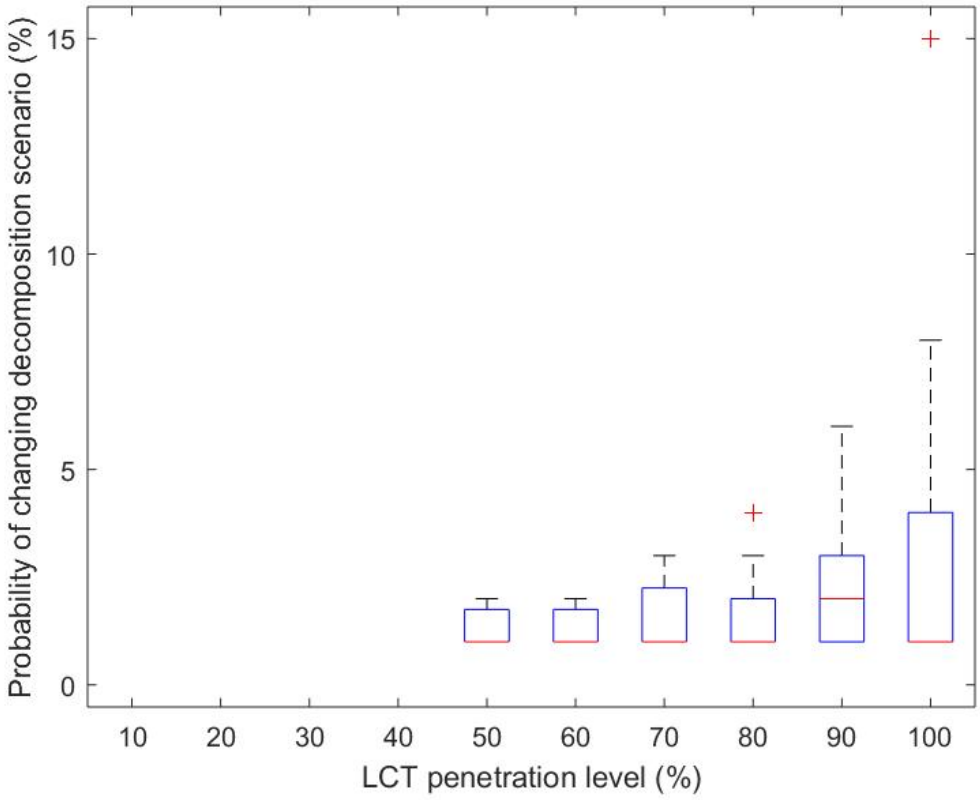


Figure 5-17. Probability of changing from definite-order scenario to random imbalance scenario with different LCT penetration levels

Figure 5-18 shows the probabilities for networks changing from the random imbalance scenario to other scenarios with changing LCT penetration levels. As can be seen, on average, the probability for networks changing from definite-max scenario to random imbalance scenario is 40%. The 30% of LCT penetration level brings the lowest probability of changing while 10% of LCT penetration brings the highest probability of changing. This means that when the LCT penetration is 10%, more than 50% of random imbalance scenarios will change to scenarios that have a systematic imbalance component.

The average probability of changing for random imbalance scenario is 40%, which is higher than that of the other scenarios, i.e., 15%, 15% and 3% for definite-max, definite-min and definite-order scenarios. It indicates that the LCT penetration reduces phase imbalances in the LV networks and this also confirms the findings in Chapter5.

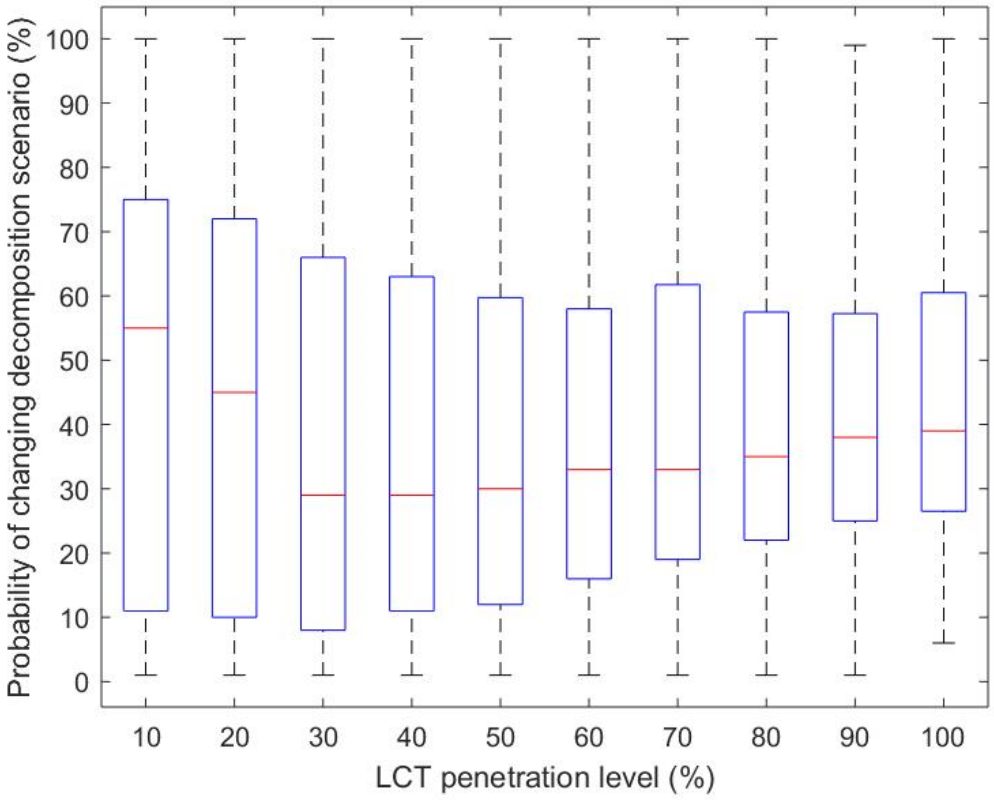


Figure 5-18. Probability of changing random imbalance scenario to other scenarios with different LCT penetration levels

Chapter 6. Conclusions

This chapter draws the conclusions to the thesis based on the research findings.

In summary, this thesis provides efficient methodologies to help distribution network operators (DNOs) analyse the phase power imbalances in the distribution system. A framework is developed for DNOs to perform a cost-benefit analysis on phase balancing solutions for data-scarce low voltage (LV) networks. The impacts of two single-phase connected low carbon technologies (LCTs), i.e., PV generation and EVs, on phase power imbalance are analysed and discussed. Specifically, the thesis addresses two major challenges:

1) Data-scarcity in the LV distribution networks for analysing network problems.
2) Uncertainty of the future of phase imbalance with the increasing penetration of EVs and PV generation.

In detail, the conclusions and contributions of this thesis are grouped into three aspects:

6.1. Characteristics of Phase Imbalances

To reveal the nature of phase imbalance, this thesis developed a new method to decompose the annual three-phase power series into a directional phase imbalance and a non-directional phase imbalance. A priori judgement is developed to classify the three-phase power series. The three-phase power series are classified into one of the four scenarios, i.e., definite-max, definite-order, definite-min, and random imbalance scenarios. The first three scenarios are decomposed into a systematic component (SIB) and a random component (RIB). These two components are the direct consequences of the two major cause of phase imbalances in the LV networks. The SIB and RIB can be effectively addressed by phase swapping and demand-side management respectively. Moreover, the degree of power imbalance (DPIB) is discovered to guide phase swapping.

Results indicated that 72.8% of 782 LV substations have SIB that can be addressed by phase swapping. The results show that 30.1%, 21.0%, 21.7%, and 27.2% of the 782 LV substations belong to the definite-max, definite-order, definite-min, and random imbalance scenarios, respectively. The methodology is highly suitable for monitored LV distribution networks in the UK and the rest of Europe. The DNOs can use the developed methodology to fully rebalance the phases by determining the maximum potential of phase swapping and the need for demand-side management. By calculating the DPIB based on the SIB component, the thesis reveals the underlying trend of SIB over time. Understanding the trend of SIB helps DNOs in system planning for future adoption of LCTs.

The decomposition analysis of phase power imbalance is also performed for two different seasons, i.e., summer and winter. The results show that although the value of DPIB varies in different seasons, the classification of network scenarios remains the same for different seasons.

6.2. Cost-Benefit Analysis of Phase Balancing

Among a mass population of LV networks, identifying the networks that are worth phase balancing, i.e., where the benefit from phase balancing outweighs its cost, is a real need for the UK industries. Existing cost-benefit analyses for phase balancing solutions require full data from distribution networks. However, the majority of LV distribution networks are data-scarce in the UK. There is a gap in evaluating the phase balancing solutions for data-scarce networks.

This thesis developed a new cost-benefit analysis framework for phase balancing on data-scarce LV networks to address the challenge of data-scarcity. The core of the framework is using customised cluster-wise Gaussian process regression (CGPR) to extract knowledge from data-rich LV networks and extrapolate the knowledge to data-scarce LV networks. The CGPR approach accounts for a full range of imbalance-induced costs, including both additional reinforcement cost (ARC) and additional energy loss cost (AELC).

The estimated net benefits of applying two different power-electronics-based phase balances (ZM-SPC and EQU18) are compared for each data-scarce LV network. Results show that approximately 70% of rural networks, 80% of suburban networks and 90% of urban networks benefit from ZM-SPC; approximately 94% of rural networks, 97% of suburban networks and 99% of urban networks benefit from EQU18. The probability that a phase balancing solution will produce a positive net benefit is also quantified.

The developed cost-benefit analysis if performed for different load growth rates and the corresponding ARC is compared with the network reinforcement cost when three phases are perfectly balanced. The results show that with large load growth, the network reinforcement cost increases rapidly for a balanced network. As a result, the percentage of ARC compared to the network reinforcement cost decreases.

The developed CGPR approach helps DNOs judge whether a phase balancing solution is economically feasible and the maximum potential net benefit from phase balancing for any data-scarce network before making any investment.

6.3. Impacts of EVs and PV Generation Penetrations on Phase Imbalances

A probabilistic impact assessment framework is developed to analyse the phase imbalances in the future distribution system with increasing penetration of EVs and PV generation. One type of EV and slow charging during the night are considered for EV penetrations; residential size PV generations (<4kW) is considered for PV penetrations.

The developed framework uses Monte Carlo simulations to include the uncertainties of the LCTs. Three LCT scenarios were considered for the probabilistic study, i.e., EV only scenario, PV only scenario and both EV and PV scenario.

The results show that the energy loss cost may exceed the network investment with penetration of single-phase LCTs. To balance the long-run investment cost and day-to-day energy loss cost, 60% of LCT penetration has the highest probability to bring the maximum benefits for the majority of the LV networks.

An analysis of the impacts of LCTs on phase power imbalance decomposition is also performed and discussed. The result shows that definite-max scenario and definite-min scenario networks have a probability of 15% to change to other scenarios, definite-order scenario networks have a probability of 3% of changing to other scenarios, and the random imbalance scenario networks have about 40% of probability of changing to other scenarios. Consequently, the penetration of LCTs reduces the phase power imbalances in the LV networks.

The developed impact assessment framework help DNOs understand the potential of benefits that LV networks can obtain from LCTs penetrations. Moreover, the developed framework can be used as a tool to perform a cost-benefit analysis for phase balancing solutions. Therefore, it guides the DNOs in investing phase balancing solutions to cope with the increasing LCT penetrations.

Chapter 7. Future Work

This chapter presents some future research topics for improving phase balancing through various approaches.

7.1. Using Smart Meter Data for Phase Balancing

The UK's government has planned to roll out smart meters to all the customers by the end of 2024 [117]. Smart meter benefits both the customers and the future distribution system operators (DSOs) at the same time. It provides customers with real-time tariffs so that customers could adjust their energy using behaviours to save money. At the same time, it also provides a large amount of customer-side data for the DSOs. These data could be processed by DSOs to be used on improving energy efficiency by realising real-time phase balancing.

Moreover, the smart meter data could also be used to forecast the phase imbalances in the future. This thesis developed a customized cluster-wise Gaussian process regression (CGPR) approach which estimates the changes of phase imbalance based on historical data from data-rich networks. The estimating accuracy will increase if more real-time data are available. Therefore, future research could be investigating the data availability from smart meters and improving methodologies of phase imbalance forecasting for the future distribution system. Increasing the forecasting accuracy will not only help DNOs better identify the future imbalance-induced costs but also assist DNOs in assessing the phase balancing solutions in terms of its capability of phase balancing and the potential benefits from phase balancing. As a result, the phase balancing in LV distribution networks can be achieved more efficiently and effectively.

7.2. Using Structural Approaches for Phase Balancing

The distribution networks face structural changes because of decentralisation. For example, the Renewable Energy Association proposed to install three-phase power supplies to all new housing developments [118]. The move aims to allow customers to have higher PV generation discharging and quicker EV charging compared to traditional single-phase connections [118]. The replacement of three-phase power supplies has uncertain impacts of phase imbalance. The impacts are yet to be investigated. The increase of three-phase load demand is expected to have the result of reducing the degree of phase imbalance. Therefore, the imbalance-induced reinforcement cost will be reduced.

As a result, further investigations on the detailed structural design of the future distribution networks and the application of structural approaches to reduce phase imbalances could be done. Analysing

the structural changes will help DNO understand the impacts on network reinforcements cost and perform a more accurate cost-benefit analysis for phase balancing solutions.

7.3. Using Market-Driven Solutions for Phase Balancing

The evolution of the distribution system facilitates the development of both the local energy market and the local service market. The local energy market is an efficient way of balancing load and generation while local service market acts as a solution for network problems. The development of local markets will effectively reduce energy losses and extracts the maximum values from LCTs by avoiding long-distance energy transmission and providing services locally. Existing research on local markets focuses on designing the trading mechanism and analysing the potential benefits for the customers. The impacts of local markets on phase imbalance and the use of the market approach for phase balancing are not discovered. Moreover, energy trading in local markets could happen within one phase or across the three phases. Different ways of local trading will have different influences on phase imbalance.

Therefore, future work could be done to analyse the detailed market structure design for the future distribution system and assess the impacts of local markets with different ways of trading on phase imbalance. Understanding such impacts will increase the accuracy of forecasting phase imbalance and allow DNO to perform phase balancing in a more cost-effective way.

References

[1] K. Ma, R. Li, and F. Li, “Quantification of Additional Asset Reinforcement Cost from Three-Phase Imbalance,” *IEEE Transactions on Power Systems,* vol. 31, no. 4, pp. 2885-2891, 2016.

[2] T. Chen, “Evaluation of line loss under load unbalance using the complex unbalance factor,” *IEE Proceedings - Generation, Transmission and Distribution,* vol. 142, no. 2, pp. 173-178, 1995.

[3] L. Ochoa, R. Ciric, A. Padilha-Feltrin, and G. Harrison, *Evaluation of distribution system losses due to load unbalance*, 2005.

[4] A. Njafi, I. Ires, and G. Naci, “Evaluating and derating of three-phase distribution transformer under unbalanced voltage and unbalance load using finite element method,” *IEEE 8th International Power Engineering and Optimization Conference*, pp. 160-165, 2014.

[5] P. Pillay, and M. Manyage, “Definitions of Voltage Unbalance,” *IEEE Power Engineering Review,* vol. 21, no. 5, pp. 49-51, 2001.

[6] "Eliminate Voltage Unbalance," https://www.nrel.gov/docs/fy00osti/27832.pdf, 2000].

[7] K. Ma, L. Fang, and W. Kong, “Review of Distribution Network Phase Imbalance: Scale, Causes, Consequences, Solutions, and Future Research Direction,” *CSEE Journal of Power and Energy Systems*, February, 2020.

[8] S. Yan, S. Tan, C. Lee, B. Chaudhuri, and S. Hui, “Electric Springs for Reducing Power Imbalance in Three-Phase Power Systems,” *IEEE Transactions on Power Electronics,* vol. 30, no. 7, pp. 3601-3609, 2015.

[9] A. Jalilian, and R. Roshanfekr, “Analysis of Three-phase Induction Motor Performance under Different Voltage Unbalance Conditions Using Simulation and Experimental Results,” *Electric Power Components and Systems,* vol. 37, no. 3, pp. 300-319, 2009.

[10] T. S. Win, Y. Hisada, T. Tanaka, E. Hiraki, M. Okamoto, and S. R. Lee, “Novel Simple Reactive Power Control Strategy With DC Capacitor Voltage Control for Active Load Balancer in Three-Phase Four-Wire Distribution Systems,” *IEEE Transactions on Industry Applications,* vol. 51, no. 5, pp. 4091-4099, 2015.

[11] L. Chia-Hung, C. Chao-Shun, C. Hui-Jen, and H. Cheng-Yu, “Heuristic rule-based phase balancing of distribution systems by considering customer load patterns,” *IEEE Transactions on Power Systems,* vol. 20, no. 2, pp. 709-716, 2005.

[12] *Sensors and Measurement Strategy*, Western Power Distribution, 2020.

[13] *The future role of network operators: The emerging active DSO model*, Baringa, 2016.

[14] *Future Power System Architecture Phase 2*, The Institution of Engineering and Technology, 2017.

[15] "2010 to 2015 government policy: low carbon technologies," Department of Energy & Climate Change, 2015.

[16] "Solar photovoltaics deployment," Department for Business, Energy & Industrial Strategy, 2017.

[17] "Future Energy Scenarios," 2020; http://fes.nationalgrid.com/media/1363/fes-interactive-version-final.pdf.

[18] "Building our Industrial Strategy Green Paper," HM Government, 2017.

[19] G. Strbac, D. Pudjianto, M. Aunedi, P. Djapic, F. Teng, X. Zhang, H. Ameli, R. Moreira, and N. Brandon, "Role and value of flexibility in facilitating cost-effective energy system decarbonisation," *Progress in Energy,* vol. 2, no. 4, pp. 042001, 2020/09/24, 2020.

[20] K. Bell, and G. Hawker, *Security of electricity supply in a low-carbon Britain*, UK Energy Research Centre, 2016.

[21] P. G. Silva, D. Ilic, and S. Karnouskos, "The impact of smart grid prosumer grouping on forecasting accuracy and its benefits for local electricity market trading," *IEEE Transactions on Smart Grid* vol. 5, no. 1, pp. 402-410, Jan, 2014.

[22] S. Grijalva, and M. U. Tariq, "Prosumer-based smart grid architecture enables a flat, sustainable electricity industry," in IEEE PES Innovative Smart Grid Technologies (ISGT), Anaheim, CA, USA, 2011.

[23] A. J. Rathnayaka, V. M. Potdar, O. Hussain, and T. Dillon, "Identifying prosumer's energy sharing behaviours for forming optimal prosumer-communities," in 2011 International Conference on Cloud and Service Computing (CSC), Hong Kong, China, 2011.

[24] R. Pal, C. Chelmis, M. Frincu, and V. Prasanna, "MATCH for the Prosumer Smart Grid The Algorithmics of Real-Time Power Balance," *IEEE Transactions on Parallel and Distributed Systems,* vol. 27, no. 12, pp. 3532-3546, 2016.

[25] M. Curtis, "Demand side response aggregators: How they decide customer suitability," in 14th International Conference on the European Energy Market (EEM), Dresden, Germany, 2017.

[26] Y. Parag, and B. Sovacool, "Electricity market design for the prosumer era," *Nature Energy*, May 31, 2016.

[27] K. Ma, R. Li, and F. Li, "Utility-Scale Estimation of Additional Reinforcement Cost from 3-Phase Imbalance Considering Thermal Constraints," *IEEE Transactions on Power Systems,* vol. 32, no. 5, pp. 3912-3923, 2016.

[28] *HV and LV Phase Imbalance Assessment* 7640–07–D4, SP Energy Networks, 2015.

[29] G. Mokryani, A. Majumdar, and B. C. Pal, “Probabilistic method for the operation of three-phase unbalanced active distribution networks,” *IET Renewable Power Generation,* vol. 10, no. 7, pp. 944-954, 2016.

[30] M. W. Siti, D. V. Nicolae, A. A. Jimoh, and A. Ukil, “Reconfiguration and Load Balancing in the LV and MV Distribution Networks for Optimal Performance,” *Power Delivery, IEEE Transactions on,* vol. 22, no. 4, pp. 2534-2540, 2007.

[31] Z. Liu, and J. V. Milanović, “Probabilistic Estimation of Voltage Unbalance in MV Distribution Networks With Unbalanced Load,” *IEEE Transactions on Power Delivery,* vol. 30, no. 2, pp. 693-703, 2015.

[32] W. Yaw-Juen, and L. Pierrat, “A method integrating deterministic and stochastic approaches for the simulation of voltage unbalance in electric power distribution systems,” *IEEE Transactions on Power Systems,* vol. 16, no. 2, pp. 241-246, 2001.

[33] J. A. Ghijselen, and A. P. Bossche, “Exact voltage unbalance assessment without phase measurements,” *IEEE Transactions on Power Systems,* vol. 20, no. 1, pp. 519-520, 2005.

[34] A. Kalyuzhny, and G. Kushnir, “Analysis of Current Unbalance In Transmission Systems With Short Lines,” *IEEE Transactions on Power Delivery,* vol. 22, no. 2, pp. 1040-1048, 2007.

[35] D. Schwanz, F. Moller, S. K. Ronnberg, J. Meyer, and M. Bollen, “Stochastic Assessment of Voltage Unbalance due to Single-Phase-Connected Solar Power,” *IEEE Transactions on Power Delivery,* vol. PP, no. 99, pp. 1-1, 2016.

[36] G. Tan, J. Cheng, and X. Sun, “Tan-Sun Coordinate Transformation System Theory and Applications for Three-Phase Unbalanced Power Systems,” *IEEE Transactions on Power Electronics,* vol. 32, no. 9, pp. 7352-7380, 2017.

[37] V. B. Bhavaraju, and P. N. Enjeti, “Analysis and design of an active power filter for balancing unbalanced loads,” *IEEE Transactions on Power Electronics,* vol. 8, no. 4, pp. 640-647, 1993.

[38] N. C. Woolley, and J. V. Milanovic, “Statistical Estimation of the Source and Level of Voltage Unbalance in Distribution Networks,” *IEEE Transactions on Power Delivery,* vol. 27, no. 3, pp. 1450-1460, 2012.

[39] A. Kavousi-Fard, T. Niknam, and M. Fotuhi-Firuzabad, “A Novel Stochastic Framework Based on Cloud Theory and Modified Bat Algorithm to Solve the Distribution Feeder Reconfiguration,” *IEEE Transactions on Smart Grid,* vol. 7, no. 2, pp. 740-750, 2016.

[40] P. Lico, M. Marinelli, K. Knezović, and S. Grillo, “Phase balancing by means of electric vehicles single-phase connection shifting in a low voltage Danish grid,” in 50th International Universities Power Engineering Conference (UPEC), 2015, pp. 1-5.

[41] J. Zhu, M. Y. Chow, and F. Zhang, “Phase balancing using mixed-integer programming,” *IEEE Transactions on Power Systems,* vol. 13, no. 4, pp. 1487-1492, 1998.

[42] M. Dilek, R. P. Broadwater, J. C. Thompson, and R. Seqiun, “Simultaneous phase balancing at substations and switches with time-varying load patterns,” *IEEE Transactions on Power Systems,* vol. 16, no. 4, pp. 922-928, 2001.

[43] L. Wang, W. Liu, C. Yeh, C. Yu, X. Lu, B. Kuan, and A. V. Prokhorov, “Reduction of three-phase voltage unbalance subject to special winding connections of two single-phase distribution transformers of a microgrid system using a designed D-STATCOM controller,” in IEEE Industry Applications Society Annual Meeting, 2017, pp. 1-9.

[44] T. Alinjak, I. Pavic, and K. Trupinic, “Improved three-phase power flow method for calculation of power losses in unbalanced radial distribution network,” *CIRED - Open Access Proceedings Journal,* vol. 2017, no. 1, pp. 2361-2365, 2017.

[45] M. Baggu, J. Giraldez, T. Harris, N. Brunhart-Lupo, L. Lisell, and D. Narang, “Interconnection assessment methodology and cost benefit analysis for high-penetration PV deployment in the Arizona Public Service system,” in IEEE 42nd Photovoltaic Specialist Conference (PVSC), 2015, pp. 1-6.

[46] F. M. Camilo, R. Castro, M. E. Almeida, and V. F. Pires, “Assessment of overvoltage mitigation techniques in low-voltage distribution networks with high penetration of photovoltaic microgeneration,” *IET Renewable Power Generation,* vol. 12, no. 6, pp. 649-656, 2018.

[47] A. Ul-Haq, C. Cecati, K. Strunz, and E. Abbasi, “Impact of Electric Vehicle Charging on Voltage Unbalance in an Urban Distribution Network,” *Intelligent Industrial Systems,* vol. 1, no. 1, pp. 51-60, 2015.

[48] K. Clement-Nyns, E. Haesen, and J. Driesen, “The Impact of Charging Plug-In Hybrid Electric Vehicles on a Residential Distribution Grid,” *IEEE Transactions on Power Systems,* vol. 25, no. 1, pp. 371-380, 2010.

[49] A. Navarro-Espinosa, and L. F. Ochoa, “Probabilistic Impact Assessment of Low Carbon Technologies in LV Distribution Systems,” *IEEE Transactions on Power Systems,* vol. 31, no. 3, pp. 2192-2203, 2016.

[50] E. Vega-Fuentes, and M. Denai, “Enhanced Electric Vehicle Integration in the UK Low-Voltage Networks With Distributed Phase Shifting Control,” *IEEE Access,* vol. 7, pp. 46796-46807, 2019.

[51] J. Quirós-Tortós, L. F. Ochoa, S. W. Alnaser, and T. Butler, “Control of EV Charging Points for Thermal and Voltage Management of LV Networks,” *IEEE Transactions on Power Systems,* vol. 31, no. 4, pp. 3028-3039, 2016.

[52] J. F. Franco, A. T. Procopiou, J. Quirós-Tortós, and L. F. Ochoa, “Advanced control of OLTC-enabled LV networks with PV systems and EVs,” *IET Generation, Transmission & Distribution,* vol. 13, no. 14, pp. 2967-2975, 2019.

[53] T. R. Ricciardi, K. Petrou, J. F. Franco, and L. F. Ochoa, “Defining Customer Export Limits in PV-Rich Low Voltage Networks,” *IEEE Transactions on Power Systems,* vol. 34, no. 1, pp. 87-97, 2019.

[54] J. D. Watson, N. R. Watson, D. Santos-Martin, A. R. Wood, S. Lemon, and A. J. V. Miller, “Impact of solar photovoltaics on the low-voltage distribution network in New Zealand,” *IET Generation, Transmission & Distribution,* vol. 10, no. 1, pp. 1-9, 2016.

[55] S. Hashemi, and J. Østergaard, “Methods and strategies for overvoltage prevention in low voltage distribution systems with PV,” *IET Renewable Power Generation,* vol. 11, no. 2, pp. 205-214, 2017.

[56] R. Torquato, D. Salles, C. O. Pereira, P. C. Meira, and W. Freitas, “A Comprehensive Assessment of PV Hosting Capacity on Low-Voltage Distribution Systems,” *IEEE Transactions on Power Delivery,* vol. 33, no. 2, pp. 1002-1012, 2018.

[57] A. Dubey, and S. Santoso, “On Estimation and Sensitivity Analysis of Distribution Circuit's Photovoltaic Hosting Capacity,” *IEEE Transactions on Power Systems,* vol. 32, no. 4, pp. 2779-2789, 2017.

[58] A. Navarro-Espinosa, and P. Mancarella, “Probabilistic modeling and assessment of the impact of electric heat pumps on low voltage distribution networks,” *Applied Energy,* vol. 127, pp. 249-266, 2014.

[59] A. J. Conejo, and L. Baringo, "Power System Fundamentals: Balanced Three-Phase Circuits," *Power System Operations*, A. J. Conejo and L. Baringo, eds., pp. 17-54, Cham: Springer International Publishing, 2018.

[60] Z. Wang, “Smart Pricing for Smart Grid,” The Department of Electronic and Electrical Engineering, University of Bath, 2014.

[61] *Network Assessment Tool: Interim Development Report*, Western Power Distribution Innovation, 2019.

[62] A. K. Singh, G. K. Singh, and R. Mitra, “Some Observations on Definitions of Voltage Unbalance,” in 39th North American Power Symposium, 2007, pp. 473-479.

[63] M. T. Bina, and A. Kashefi, “Three-phase unbalance of distribution systems: Complementary analysis and experimental case study,” *International Journal of Electrical Power & Energy Systems,* vol. 33, no. 4, pp. 817-826, 2011/05/01/, 2011.

[64] E. n. association, "Engineering Recommendation P29," *Planning Limits for Voltage Unbalance in the United Kingdom*, 1990.

[65] *Electromagnetic compatibility (EMC)-limits-assessment of emission limits for the connection of unbalanced installations to MY, HY and EHV power systems*, International Electrotechnical Commision, 2008.

[66] W. Kong, K. Ma, and Q. Wu, “Three-Phase Power Imbalance Decomposition Into Systematic Imbalance and Random Imbalance,” *IEEE Transactions on Power Systems,* vol. 33, no. 3, pp. 3001-3012, 2018.

[67] V. Rigoni, L. F. Ochoa, G. Chicco, A. Navarro-Espinosa, and T. Gozel, “Representative Residential LV Feeders: A Case Study for the North West of England,” *IEEE Transactions on Power Systems,* vol. 31, no. 1, pp. 348-360, 2016.

[68] J. N. Fidalgo, C. Moreira, and R. Cavalheiro, “Impact of Load Unbalance on Low Voltage Network Losses,” in IEEE Milan PowerTech, 2019, pp. 1-5.

[69] W. H. Kersting, "The computation of neutral and dirt currents and power losses." pp. 213-218 vol.1.

[70] J. Zhu, G. Bilbro, and C. Mo-Yuen, “Phase balancing using simulated annealing,” *IEEE Transactions on Power Systems,* vol. 14, no. 4, pp. 1508-1513, 1999.

[71] S. Pajic, and A. Emanuel, “Effect of neutral path power losses on the apparent power definitions: A preliminary study,” in IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008, pp. 1-1.

[72] N. Johnson, *Central Networks Earthing Manual*, E.ON UK 2007.

[73] R. Salustiano, E. Tavares, and M. Martinez, “The unbalanced load cost on transformer losses at a distribution system,” in 22nd International Conference and Exhibition on Electricity Distribution (CIRED), Stockholm, Sweden, 2013, pp. 1-3.

[74] J. D. Watson, N. R. Watson, and I. Lestas, “Optimized Dispatch of Energy Storage Systems in Unbalanced Distribution Networks,” *IEEE Transactions on Sustainable Energy,* vol. 9, no. 2, pp. 639-650, 2018.

[75] L. Fang, K. Ma, R. Li, Z. Wang, and H. Shi, “A Statistical Approach to Estimate Imbalance-Induced Energy Losses for Data-Scarce Low Voltage Networks,” *IEEE Transactions on Power Systems*, pp. 1-1, 2019.

[76] W. Kong, K. Ma, L. Fang, R. Wei, and F. Li, “Cost-Benefit Analysis of Phase Balancing Solution for Data-scarce LV Networks by Cluster-Wise Gaussian Process Regression,” *IEEE Transactions on Power Systems*, pp. 1-1, 2020.

[77] T. A. Short, *Electric power distribution handbook*, Boca Raton: CRC Press, 2004.

[78] B. J. Elliott, *Designing a Structured Cabling System to ISO 11801 (2nd Edition)*: Woodhead Publishing.

[79] G. Cronshaw, *EARTHING: YOUR QUESTIONS ANSWERED*, IEE Wiring Matters, 2005.

[80] U. P. Networks, "Engineering Design Stardard," *Customer LV Installation Earthing Design*, 2018.

[81] A. Boardman, *LV NETWORK EARTHING DESIGN*, UK Power Network, 2015.

[82] "Electricity distribution units and loss percentages summary," 2019; https://www.ofgem.gov.uk/sites/default/files/docs/2010/08/distribution-units-and-loss-percentages-summary.pdf.

[83] K. Ma, F. Li, and R. Aggarwal, “Quantification of Additional Reinforcement Cost Driven by Voltage Constraint Under Three-Phase Imbalance,” *IEEE Transactions on Power Systems,* vol. 31, no. 6, pp. 5126-5134, 2016.

[84] D. Singh, R. K. Misra, and S. Mishra, “Distribution system feeder re-phasing considering voltage-dependency of loads,” *International Journal of Electrical Power & Energy Systems,* vol. 76, pp. 107-119, 2016/03/01/, 2016.

[85] P. F. Ryff, T. H. Fawzi, and A. M. Hussein, “Enclosure losses in high power systems,” *IEEE Transactions on Magnetics,* vol. 30, no. 6, pp. 4338-4340, 1994.

[86] Z. Liu, Z. Zheng, L. Xu, K. Wang, and Y. Li, “Current Balance Control for Symmetrical Multiphase Inverters,” *IEEE Transactions on Power Electronics,* vol. 31, no. 6, pp. 4005-4012, 2016.

[87] C. Peng, L. Xu, X. Gong, H. Sun, and L. Pan, “Molecular Evolution Based Dynamic Reconfiguration of Distribution Networks With DGs Considering Three-Phase Balance and Switching Times,” *IEEE Transactions on Industrial Informatics,* vol. 15, no. 4, pp. 1866-1876, 2019.

[88] S. Dimitrijevic, and N. Rajakovic, “Service Restoration of Distribution Networks Considering Switching Operation Costs and Actual Status of the Switching Equipment,” *IEEE Transactions on Smart Grid,* vol. 6, no. 3, pp. 1227-1232, 2015.

[89] C. Tsai-Hsiang, and C. Jeng-Tyan, “Optimal phase arrangement of distribution transformers connected to a primary feeder for system unbalance improvement and loss reduction using a genetic algorithm,” in Proceedings of the 21st International Conference on Power Industry Computer Applications. Connecting Utilities. PICA 99. To the Millennium and Beyond (Cat. No.99CH36351), 1999, pp. 145-151.

[90] K. Wang, S. Skiena, and T. G. Robertazzi, “Phase balancing algorithms,” *Electric Power Systems Research,* vol. 96, pp. 218-224, 2013/03/01/, 2013.

[91] R. A. Hooshmand, and S. Soltani, “Fuzzy Optimal Phase Balancing of Radial and Meshed Distribution Networks Using BF-PSO Algorithm,” *IEEE Transactions on Power Systems,* vol. 27, no. 1, pp. 47-57, 2012.

[92] S. H. Soltani, M. Rashidinejad, and A. Abdollahi, “Dynamic phase balancing in the smart distribution networks,” *International Journal of Electrical Power & Energy Systems,* vol. 93, pp. 374-383, 2017.

[93] L. Fang, K. Ma, and X. Zhang, “A Statistical Approach to Guide Phase Swapping for Data-Scarce Low Voltage Networks,” *IEEE Transactions on Power Systems,* vol. 35, no. 1, pp. 751-761, 2020.

[94] F. Ding, and K. A. Loparo, “Feeder Reconfiguration for Unbalanced Distribution Systems With Distributed Generation: A Hierarchical Decentralized Approach,” *IEEE Transactions on Power Systems,* vol. 31, no. 2, pp. 1633-1642, 2016.

[95] S. Soltani, M. Rashidinejad, and A. Abdollahi, “Stochastic Multiobjective Distribution Systems Phase Balancing Considering Distributed Energy Resources,” *IEEE Systems Journal,* vol. 12, no. 3, pp. 2866-2877, 2018.

[96] N. Gupta, A. Swarnkar, and K. R. Niazi, “A novel method for simultaneous phase balancing and mitigation of neutral current harmonics in secondary distribution systems,” *International Journal of Electrical Power & Energy Systems,* vol. 55, pp. 645-656, 2014.

[97] G. Strbac, “Demand side management: Benefits and challenges,” *Energy Policy,* vol. 36, no. 12, pp. 4419-4426, 2008/12/01/, 2008.

[98] A. H. Zaidi, K. Sunderland, and M. Conlon, “Role of reactive power (STATCOM) in the planning of distribution network with higher EV charging level,” *IET Generation, Transmission & Distribution,* vol. 13, no. 7, pp. 951-959, 2019.

[99] S. Chen, Z. Guo, Z. Yang, Y. Xu, and R. S. Cheng, “A Game Theoretic Approach to Phase Balancing by Plug-in Electric Vehicles in the Smart Grid,” *IEEE Transactions on Power Systems*, pp. 1-1, 2019.

[100] S. Weckx, and J. Driesen, “Load Balancing With EV Chargers and PV Inverters in Unbalanced Distribution Grids,” *IEEE Transactions on Sustainable Energy,* vol. 6, no. 2, pp. 635-643, 2015.

[101] C. G. Bajo, S. Hashemi, S. B. Kjsær, G. Yang, and J. Østergaard, “Voltage unbalance mitigation in LV networks using three-phase PV systems,” in 2015 IEEE International Conference on Industrial Technology (ICIT), 2015, pp. 2875-2879.

[102] F. Shahnia, R. Majumder, A. Ghosh, G. Ledwich, and F. Zare, “Voltage imbalance analysis in residential low voltage distribution networks with rooftop PVs,” *Electric Power Systems Research,* vol. 81, no. 9, pp. 1805-1814, 2011/09/01/, 2011.

[103] S. Acharya, M. S. El-Moursi, A. Al-Hinai, A. S. Al-Sumaiti, and H. H. Zeineldin, “A Control Strategy for Voltage Unbalance Mitigation in an Islanded Microgrid Considering Demand Side Management Capability,” *IEEE Transactions on Smart Grid,* vol. 10, no. 3, pp. 2558-2568, 2019.

[104] S. Sun, B. Liang, M. Dong, and J. A. Taylor, “Phase Balancing Using Energy Storagein Power Grids Under Uncertainty,” *IEEE Transactions on Power Systems,* vol. 31, no. 5, pp. 3891-3903, 2016.

[105] K. H. Chua, J. Wong, Y. S. Lim, P. Taylor, E. Morris, and S. Morris, “Mitigation of Voltage Unbalance in Low Voltage Distribution Network with High Level of Photovoltaic System,” *Energy Procedia,* vol. 12, pp. 495-501, 2011.

[106] J. Lee, Y. Cho, and J. Jung, “Single-Stage Voltage Balancer With High-Frequency Isolation for Bipolar LVDC Distribution System,” *IEEE Transactions on Industrial Electronics,* vol. 67, no. 5, pp. 3596-3606, 2020.

[107] F. Wang, Z. Lei, X. Xu, and X. Shu, “Topology Deduction and Analysis of Voltage Balancers for DC Microgrid,” *IEEE Journal of Emerging and Selected Topics in Power Electronics,* vol. 5, no. 2, pp. 672-680, 2017.

[108] S. M. Fazeli, H. W. Ping, N. B. Rahim, and B. T. Ooi, “Individual-phase control of 3-phase 4-wire voltage-source converter,” *IET Power Electronics,* vol. 7, no. 9, pp. 2354-2364, 2014.

[109] S. M. Fazeli, H. W. Ping, N. B. Rahim, and B. T. Ooi, “Individual-phase decoupled P-Q control of three-phase voltage source converter,” *IET Generation, Transmission & Distribution,* vol. 7, no. 11, pp. 1219-1228, 2013.

[110] F. Shahnia, A. Ghosh, G. Ledwich, and F. Zare, “Voltage Unbalance reduction in low voltage distribution networks with rooftop PVs,” in 20th Australasian Universities Power Engineering Conference, 2010, pp. 1-5.

[111] "ZM-SPC Series Three phase unbalanced device," March, 2019; http://zmgs.com/en/index.php?r=article/Content/index&content_id=596.

[112] "EQUI8 : THREE-PHASE LOW VOLTAGE NETWORK BALANCER," March, 2019; http://cmetransformateur.com/equi8-en/.

[113] G. Evans, and D. MacLeman, *Demonstrating the Benefits of Monitoring LV Networks with embedded PV Panels and EV Charging Point*, Southern Electric Power Distribution 2013.

[114] "Sepam™ Series 20 Protective Relays User’s Manual," S. Electric, ed., 2007.

[115] *Future Power System Architecture*, Institution of Engineering and Technology, 2016.

[116] *Open Networks Future Worlds*, Energy Networks Association, 2018.

[117] S. Gompertz. "Smart meter rollout delayed for four years," March, 2020; https://www.bbc.co.uk/news/business-49721436.

[118] "The feasibility, costs and benefits of three phase power supplies in new homes," https://www.r-e-a.net/news/new-homes-shouldnt-be-held-back-by-pre-wwii-electrical-standards, [August, 2018].

[119] M. Dale, *LV Network Templated for A Low-Carbon Future*, Western Power Distribution, 2013.

Appendices

Appendix-A EQU18 Three-Phase LV Network Balancer

Table A-1. Detailed technical specifications of EQUI8 [112]

Product Name	**EQUI8**
Nominal Current ($I_{neutral}$) ***Short duration Overload***	45 Amps *60 Amps 3 hours / 80 Amps 15 min.*
Dimension (D x W x H)	21x54x55 cm
Weight	53 kg
Network	4-Wire, 3-Phase, 400 V between lines
Frequency	50Hz
Enclosure Ingress Protection	IP34D
Inner components Ingress Protection	IP2X
Insulation Class	Class II
Efficiency	97.7% under full load Idle power < 25 W
Enclosure mechanical Protection	IK10
Working Temperature	-40 to 70 °C
Humidity	0 to 100%RH
Salt Mist Withstand	Essai Ka (EN 60068-2-11)
Expected lifespan	30 years

Figure A-1. EQU18 taken from [112]

Appendix-B ZM-SPC Three-Phase LV Network Balancer

Table A-2. Detailed technical specifications of ZM-SPC [111]

System data	ZM-SPC30	ZM-SPC50	ZM-SPC100
Rated voltage rating	400V		
Input phase voltage range	40%~20%		
Grid frequency	50Hz/60Hz(Range45Hz-63Hz)		
Efficiency	>97%		
Network structure	Three-phase three-wire, three-phase four-wire		
Circuit topology	Three levels		
Performance			
Three-phase balance compensation	Unbalance compensation to 50% or less		
Rated current (or maximum unbalanced compensation current)	45A	75A	150A
Reactive compensation range	-1~1(Adjustable)		
Reactive power compensation capacity	30kvar	50kvar	100kvar
Reactive power compensation rate	>99%		
Response time	15ms		
Protective function	Overvoltage protection, undervoltage protection, short circuit protection, reverse bridge reverse protection, over compensation protection, lightning protection double layer		
Power distribution function	Single-channel power analysis, with C-class lightning protection function		Double-channel power analysis, with C-class lightning protection function
Communication monitoring function			
Display content	Voltage, current, frequency, power factor, operating temperature and other real-time operating information		
Communication Interface	RS485 / CAN / network port		
Protocol	Modbus protocol / power master protocol		
Mechanical characteristic			
Dimensions	W820xD350xH1180mm		W820xD400xH1180mm
	Stainless steel		
	58kg	72kg	135kg
Installation	The holding pole，H Rod，F Rod		
Mounting angle	<5degree		
Environment			
Operating temperature	-10°C~55°C		
Altitude	<1500, 1500 meters or more in accordance with GB / T3859.2 please use the derating, or contact the dealer		
Degree of protection	IP44		
Seismic capacity	7Class		

Figure A-2. ZM-SPC taken from [111]

Appendix-C Data of LV Networks

In this thesis, time-series of phase current and voltage magnitudes data of 800 representative data-rich LV networks throughout a year is used. These networks are located within the business area of a UK DNO and the data are the deliverables of the “Low Voltage Network Templates” project . When conducting the trial project and collecting network data, Western Power Distribution specifically chose networks of a diverse and heterogeneous nature so that the dataset is representative. These 800 networks cover various customer types (domestic, commercial and industrial customers) and geographical areas (urban, suburban, and rural areas). For example, Cardiff contains a large number of commercial customers and load; Monmouthshire is a representative for the rural area.

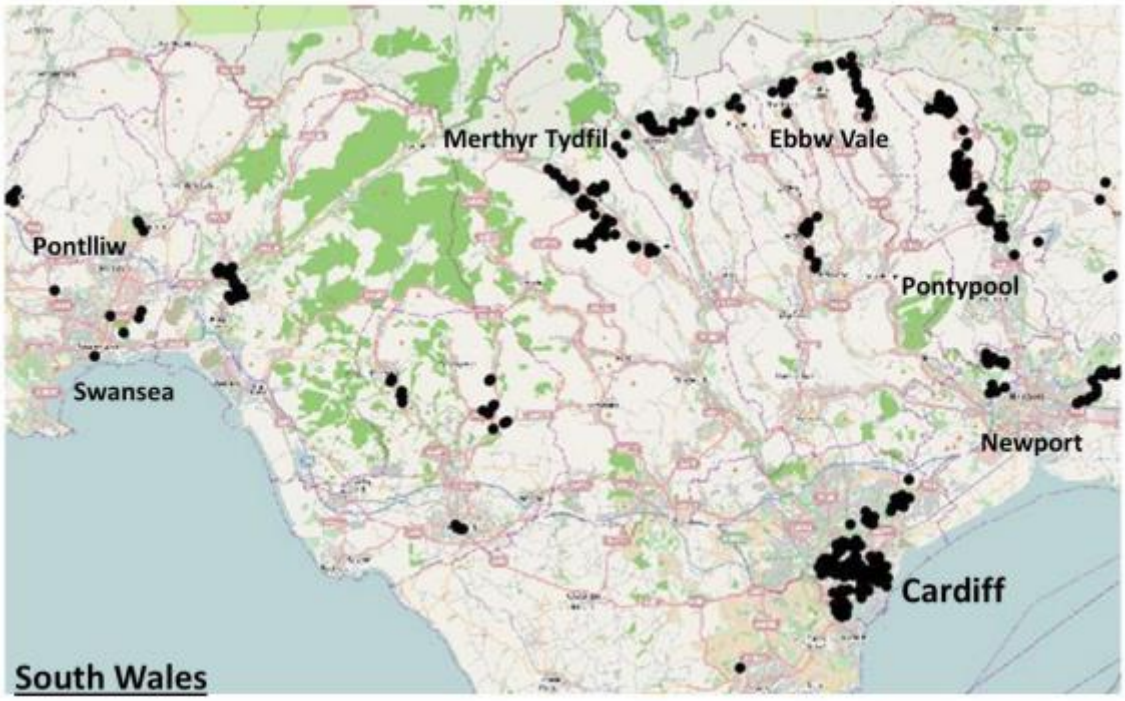


Figure C-1, Map of locations of monitored substations within the South Wales Study Area taken from [119]