UNIVERSITY OF LIVERPOOL

Load Pattern Categorisation and Dynamic Pricing Based Demand Response in Smart Grid

Thesis submitted in accordance with the requirements of the University of Liverpool
for the degree of Doctor in Philosophy
by

Xing LUO

Department of Electrical Engineering and Electronics
School of Electrical Engineering and Electronics and Computer Science
University of Liverpool

Supervisor team:
Prof. Eng Gee Lim (Xi'an Jiaotong-Liverpool University)
Dr. Xu Zhu (University of Liverpool)

28 Feb. 2019

Abstract

The Smart Grid is widely regarded as the next generation of the power grid in power system reform. It is the application of digital processing and communications to the power grid, making data flow and information management central to the grid. Demand response (DR) is an essential characteristic of the smart grid and it plays an important role in energy efficiency improvement and wastage reduction by providing encouraging energy-aware consumption. However, efficient DR management involves a variety of challenges including categorisation of load patterns, accurate real-time price (RTP) forecasting, effective DR program designing, *etc.* This thesis extends around these related challenges in the smart grid and presents significant outcomes.

Load pattern categorisation (LPC) plays an important role in DR. However, how to determine a precise cluster number and choose an appropriate clustering algorithm are critical in LPC and remain challenging. In this thesis, as the first contribution, a novel parametric bootstrap (PB) algorithm is proposed, incorporated with a compatible clustering technique to address the cluster number determination problem as well as clustering the load data simultaneously. The PB algorithm is more robust against dimensionality of data and hence applicable to load demand data which is usually of high dimensionality. It is also general and independent of data type, resulting in a more appropriate cluster number determination result than existing methods with little fluctuation. The evaluation results indicate the feasibility and superiority of the proposed approach over others previously published in the literature.

The RTP tariff has become a trend in the smart grid and it is usually utilised as an input control signal to enable efficient load shifting in DR. As the

second contribution of the thesis, a hybrid RTP forecasting model considering deterministic and stochastic features of input data is proposed to forecast short-term electricity prices. The evaluation results clearly demonstrate that the proposed approach is effective in RTP forecasting with a higher accuracy compared with existing models from the literature.

An effective DR strategy is the core of DR. As the third contribution, a number of DR strategies assisted by electric vehicles (EVs) are proposed. Innovative EV assisted DR strategies with the EV as an auxiliary power supply (EV-APS) model and a neighbour energy sharing (NES) model are proposed, to jointly optimise the load distribution for both a single household and multi-household network via vehicle to home (V2H) and vehicle to neighbour (V2N) connections, respectively. The proposed DR strategies take account of the comprehensive impacts of EVs' charging behaviors, user preferences, distributed energy, and load scheduling priorities. The effectiveness of the DR strategies are verified by numerical results in terms of load balancing and cost reduction, and the proposed DR strategies show better performance compared with previously published DR approaches.

Contents

List of Tables

List of Figures

LIST OF ABBREVIATIONS

ACU	Automatic Control Unit
AD	Anderson-Darling
AEMO	Australia Energy Market Operator
AGO	Accumulated Generating Operation
AHC	Agglomerative Hierarchical Clustering
AIC	Akaike Information Criterion
ANN	Artificial Neural Network
APS	Auxiliary Power Supply
AR	Auto Regressive
ARIMA	Auto Regressive Integrated Moving Average
ARMA	Auto Regressive Moving Average
BIC	Bayesian Information Criterion
BP-ANN	Back Propagation-Artificial Neural Network
BS	Bootstrap Simulation
CCU	Centralised Control Unit
CP	Compactness
CPP	Critical Peak Pricing

CR	Charging Rate
CS	Critical Scenario
CVMM	C-Vine Copulas Based Mixture Model
DBI	Davies-Bouldin Index
DCR	Discharging Rate
DHC	Divisive Hierarchical Clustering
DLC	Direct Load Control
DP	Dynamic Pricing
DPMM	Dirichlet Process Mixture Model
DR	Demand Response
DRA	Demand Response Aggregator
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
EISA	Energy Independence and Security Act
EHVAC	Electrical Heating Ventilation Air Conditioning
EM	Expectation Maximisation
EPF	Electricity Price Forecasting
ER	Error Rate
ERoA	Energy Remaining of Arriving Home
ESS	Energy Storage System

EVs	Electric Vehicles
FAN	Field Area Network
FARIMA	Fractional ARIMA
FKM	Fuzzy K-Means
FR	Failure Rate
FS	Flexible Scenario
FTL	Follow the Leader
FWT	Fast Wavelet Transformation
GARCH	Generalised AR Conditional Heteroskedasticity
GMM	Gaussian Mixture Model
GMCM	Gaussian Mixture Copula Model
GP	Grey Prediction
HAN	Home Area Network
HAs	Household Appliances
HEV	Hybrid Electric Vehicle
IAN	Industrial Area Network
IAGO	Inverse Accumulated Generating Operation
IBR	Inclining Block Rate
KKM	Kernel K-Means

KMS	K-Means
KMD	K-Medoids
KMS++	K-Means++
LA	Input Layer
LB	Hidden Layer
LC	Output Layer
LDF	Load Demand Forecasting
LP	Linear Programming
LPC	Load Pattern Categorisation
LS	Least Square
LSC	Load Scheduling Control
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MDP	Markov Decision Process
MILP	Mixed Integer Linear Programming
ML	Machine Learning
MSE	Means Square Error
NAN	Neighbourhood Area Network
NES	Neighbour Energy Sharing

PAM	Partitioning Around Medoids
PB	Parametric Bootstrap
PCC	Pearson Correlation Coefficient
PDF	Probability Density Function
PSO	Particle Swarm Optimisation
RMSE	Root Mean Square Error
RRT	Relative Running Time
RTP	Real Time Price
SARIMA	Seasonal ARIMA
SFE	Supply Function Equilibrium
SOM	Self-Organising Maps
SP	Separation
SPCM	Strategic Production-Cost Model
SSE	Sum of Square Errors
STD	Standard Deviation
SVC	Support Vector Clustering
SVM	Support Vector Machine
TAR	Threshold AR
TC	Total Cost

TLPs	Typical Load Patterns
ToA	Time of Arriving
ToL	Time of Leaving
ToU	Time of Use
V2H	Vehicle to Home
V2N	Vehicle to Neighbour
WAN	Wide Area Network

LIST OF KEY SYMBOLS

Chapter 1

None

Chapter 2

$\mathcal{S}$	load demand dataset
$\mathcal{L}(X\|C)$	log-likelihood of the dataset X according to model C
p	number of parameters in the model C
$f_0(x)$	objective function in convex optimisation
$f_i(x)$	constraint function in convex optimisation
$\mathcal{G}$	game in game theory
$\mathcal{N}$	player in game theory
$\{\mathcal{X}_i\}_{i\in\mathcal{N}}$	strategies in game theory
$\{W_i\}_{i\in\mathcal{N}}$	payoff functions in game theory

Chapter 3

$\mathcal{S}$	sub-cascade
P	feature parameter vector
μ	mean vector
c	covariance

E	covariance matrix
d	dimensionality
k	cluster number
Φ_i^{BS}	sum of square errors of i^{th} BS dataset
Φ^{AC}	sum of square errors of actual dataset
α	threshold in statistical hypothesis
Ω	set of clusters
Ω_k	k^{th} cluster
w	set of centers
w_k	center of k^{th} cluster
m_k	k^{th} medoid
ϕ_i	mixture coefficient of Gaussian distributions
μ	mean vector of a Gaussian distribution
Σ	covariance matrix of a Gaussian distribution
$\mathcal{N}$	Gaussian distribution
θ	parameter set of mixture Gaussian distributions
O_{KMS}	time complexity of K-means algorithm
O_{KMD}	time complexity of K-medoids algorithm

N^{cluster}	number of generated data in a cluster
N^{BS}	number of bootstrap simulations for a hypothesis
N_{failure}	number of failed tests
N_{text}	number of total tests
$\rho(X, Y)$	linear correlation between X and Y

Chapter 4

P_t	forecasting RTP at time t
L_t	estimation of deterministic characteristic of input data
N_t	estimation of stochastic characteristic of input data
E_t^*	error optimisation procedure
H	historical RTP dataset
D_n	RTP data of n^{th} day
J	total square errors
α_d	significance level for fitting degree d
M_t	input dataset of GP model at time t
$X^{(0)}$	primitive sequence dataset
$X^{(1)}$	first accumulated generating dataset
$Z^{(1)}$	background factor
a	development coefficient

u	endogenous control coefficient
a_i	i^{th} node of LA layer
b_r	r^{th} node of LB layer
c_j	j^{th} node of LC layer
W_{ir}	connection weight between a_i and b_r
V_{rj}	connection weight between b_r and c_j
T_r	bias of node b_r of LB layer
θ_j	bias of node c_j of LC layer
d_j	bias of desired value
$\Delta V_{rj}^{'}$	adjusted connection weight between b_r and c_j
$\theta_j^{'}$	adjusted bias of node c_j of LC layer
α	learning ratio
β	momentum factor
e_r	back propagated errors to layer LB nodes
ref	observed true value

Chapter 5

R_t	RTP tariff
E_c	energy consumption of CS appliances
E_f	energy consumption of FS appliances

W_t^{grid}	total energy consumption on grid
P_t^{grid}	total power rate on grid at time t
T_{in}	initial time
T_{term}	terminate time
P_t^{HA}	power rate of household appliances at time t
$P_t^{\text{EV,c}}$	power rate of EV charging at time t
$P_t^{\text{EV,d}}$	power rate of EV discharging at time t
$P_{\text{max}}^{\text{grid}}$	maximal power rate on grid
ε	scheduling priority
$W^{\text{EV,min}}$	minimal battery capacity of EV
$W^{\text{EV,max}}$	maximal battery capacity of EV
$W^{\text{EV,rem}}$	remaining energy of EV
$W^{\text{EV,trip}}$	energy consumption of EV on trip
$W^{\text{EV,(1)}}$	initial energy storage of 1^{st} day
$W^{\text{EV,(2)}}$	initial energy storage of 2^{nd} day
$W^{\text{EV,c}}$	energy of EV charging
$W^{\text{EV,d}}$	energy of EV discharging
η_1	EV battery charging efficiency
η_2	EV battery discharging efficiency

$P_t^{\mathrm{EV,d}}$	rated power of EV discharging
T	time parameter
k	household index
$\hat{k}$	index of household equipped EV
$\tilde{k}$	index of household without EV
α, β	binary parameter to indicate EV status
τ	threshold parameter of EV remaining energy
$P_{\hat{k},t}^{\mathrm{EV,c}}$	power rate of $\hat{k}^{\mathrm{th}}$ EV charging at time t
$P_{\hat{k},t}^{\mathrm{EV,d,v2h}}$	power rate of $\hat{k}^{\mathrm{th}}$ EV discharging at time t via V2H
$P_{\hat{k},t}^{\mathrm{EV,d,v2n}}$	power rate of $\hat{k}^{\mathrm{th}}$ EV discharging at time t via V2N
$\eta_{\hat{k}}^{\mathrm{c}}$	charging efficiency of $\hat{k}^{\mathrm{th}}$ EV
$\eta_{\hat{k}}^{\mathrm{d,v2h}}$	discharging efficiency of $\hat{k}^{\mathrm{th}}$ EV via V2H
$\eta_{\hat{k}}^{\mathrm{d,v2n}}$	discharging efficiency of $\hat{k}^{\mathrm{th}}$ EV via V2N
$P_{\hat{k},t}^{\mathrm{act}}$	actual power demand of household installed EV
$P_{\tilde{k},t}^{\mathrm{act}}$	actual power demand of neighbour household
$B_{\hat{k}}^{\mathrm{NES}}$	obtained benefit of household who sell EV energy
$B_{\tilde{k}}^{\mathrm{NES}}$	obtained benefit of household who buy EV energy
$C_{\tilde{k}}^{\mathrm{dmd}}$	cost for electricity demand without EV sharing

$C_{\hat{k}}^{\mathrm{EV,c}}$	cost for EV charging of EV sharing part
$\theta\%$	profit distribution parameter

Chapter 6

None

Write your injuries in dust,
your benefits in marble.

Benjamin Franklin, Statesman

Acknowledgement

The past four years of the Ph.D. journey is full of obstacles, challenges, happiness, successes, and failures, which will always remind me of the importance of humility, exploration and lifelong learning. The completion of the thesis marks the end of such an eventful experience. I believe that the achievements of my academic research and study during my period to get the Ph.D. certificate can hardly be fulfilled without the help and the support of my supervisors, colleagues, friends, and families.

I would like to first express my heartfelt gratitude to my primary Ph.D. supervisor and one of my close friends, Prof. Eng Gee Lim, in Xi'an Jiaotong -Liverpool University, for giving me a valuable opportunity to start my Ph.D. study, and his kindly supports in my Ph.D. life in Suzhou, China. Although he was very busy as a leader in the department, he always made his best effort to create a comfortable research environment and provide conveniences for me. Besides, the grateful gratitude will be sent to my secondary Ph.D. supervisor, Dr. Xu (Judy) ZHU in the University of Liverpool, for picking me as a Ph.D. candidate. Six years ago, it was my first time to go abroad alone and studied in Liverpool for my Master degree. From that time, I became her student until now. I offer my sincere appreciation and gratitude to her patient advice, warm help, funding, and edits of my thesis as well as academic papers. She tried her best to provide me with both economic and academic support to attend international conferences and other academic activities.

In addition, I would also express my thanks to my colleagues and friends in lab EE511, Xi'an Jiaotong-Liverpool University, Suzhou. They are Mr. Wenfei ZHU, Mr. Haochuang JIANG, Miss. Xiaotong XU, Mr. Jingchen WANG, Mr. Bing

HAN, Miss. Yujie LIU, Miss. Zhenzhen JIANG and Mr. Rui PEI. They offered me intensive and continuous support in my Ph.D. life. Those colleagues and friends are also in the Department of Electric and Electronic Engineering, Xi'an Jiaotong-Liverpool University and the Department of Electrical Engineering and Electronics, University of Liverpool. They are Prof. Yi HUANG for his sincere advice in my academic research, Dr. Yufei JIANG, Dr. Zhongxiang WEI, Dr. Jun YIN, Mr. Boda LIU, Prof. Mark Leach, Prof. Kaizhu HUANG, Dr. Rui LIN, Dr. Shaofeng LU, Ms. Xiaoyi WU, *etc.* Additional thanks will also be sent to my friends at the Technical University of Munich, *e.g.,* Prof. Wolfgang Kellerer who provided me the research visiting opportunity in Munich, Prof. Eckehard Steinbach, Dr. Xiao XU and his wife Mrs. Jie SHEN for leading us hiking around Free State of Bavaria, Prof. Chenglin LI, Miss. Jingyi XU, Miss. Bing FAN, Mr. Kai CUI, Mr. Murat Guersu, Mr. Mu HE, Mr. Patrick Kalmbach, *etc.* I thank all of my teachers, students, friends, relatives and all others who helped and inspired me directly or indirectly, and all who have wished me well in my study, research and other purposes.

This thesis is dedicated to my parents for their unconditional love, and who have always tried to make my life comfortable, always wish me for success and happiness from the moment I came to the world. As the only child of them, I know how hard for them to be apart with me for so many years. Thanks to my parents for consistently encouraging me to be brave, independent and optimistic, and I really appreciate their support in these years.

Last but not least, I would like to thank Mrs. Zhi JIN, who is always troublesome but brave enough to be my wife. Without her, I could complete this thesis faster. However, without this Ph.D. journey, I will not have her in my life.

Chapter 1

Introduction

Chapter 1 provides an overview of the thesis highlighting the aims of the work undertaken and its novelty in the field. Specifically, the motivation behind the work is explained in Section 1.1, the objectives of the thesis are proposed in Section 1.2, the main contributions and organisation of the thesis are presented in Section 1.3. Finally, the publications produced as a result of the work are listed in Section 1.4.

1.1 Motivation

The power grid is a large interconnected infrastructure for delivering electricity from power plants to end consumers. With continuous development and modernisation, great changes have taken place in the last few decades on the information frontier and in relation to control theory and artificial intelligence. The modern power system has not kept pace with the development of these technologies and innovations and is not wholly satisfying the demands of the modern world. Recent electricity blackouts in rural areas of China and other developing countries indicate that there are great challenges, currently and in the future, for conventional power grids. Widely considered to be the next generation of the electricity grid in power system reform, the smart grid has been proposed as the next generation for energy production, transmission, distribution, and consumption.

The smart grid which encompasses advanced power, communications, control, and computing technologies [1], is envisioned as a large-scale cyber-physical

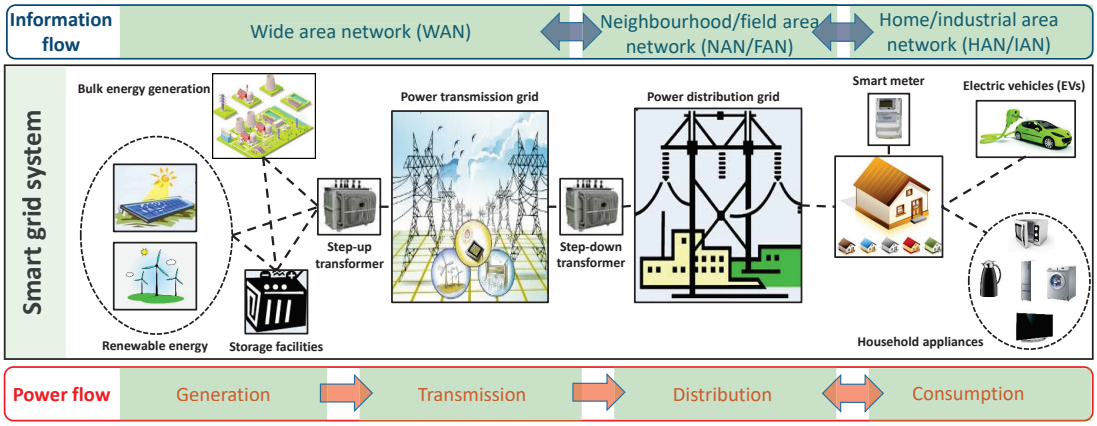


Figure 1.1: Vision of the cyber-physical smart grid system. The cyber part of the proposed system includes wide area network (WAN), neighbourhood/field area network (NAN/FAN), home/industrial area network (HAN/IAN). The physical part of the proposed system consists of energy generation, transmission, distribution and consumption.

system as shown in Figure 1.1. The concept of the smart grid was firstly defined by the Energy Independence and Security Act of 2007 (EISA-2007), which was approved by the US Congress in January 2007, and signed into law by President George W. Bush in December 2007. A common feature to most definitions of the smart grid is the application of digital processing and communications to the power grid, making data flow and information management central to the smart grid. It is a power network composed of intelligent nodes that can operate, communicate, and interact autonomously to efficiently deliver power and electricity to energy consumers. This heterogeneous nature of the smart grid motivates development in a number of advanced techniques to overcome various challenges arising at different levels such as design, control, and implementation. These innovative techniques could be adopted in the smart grid to improve reliability, sustainability, efficiency, security, and environmental friendliness.

With the rapid development of smart metering technologies, bidirectional communications are enabled between power suppliers and users, by having smart meters installed at consumers' premises. Accordingly, demand response (DR) becomes an essential characteristic of the smart grid [2]. The efficient management of DR on the demand side (*e.g.,* residential homes) is widely regarded as an excellent long-term solution to improving energy efficiency

and reducing wastage by providing energy-efficient equipment and encouraging energy-aware consumption. It plays a significant role in balancing energy supply and demand, and enhancing the reliability of the smart grid [3–5].

The basic principle of DR management at a consumers' premises is to reduce or shift the demand for electricity during peak periods in response to a dynamic price (DP) or real-time price (RTP) tariff, or other form of financial incentive in an automated and convenient manner [6, 7], thus achieving the aim of reducing electricity bills for customers. In other ways, it is also beneficial for the power grid as it offers an effective solution to averaging power usage over certain periods alleviating the load burden of the power grid [8–10]. However, designing and implementing efficient DR based energy management solutions involves a variety of challenges:

(1) Clarifying the scheduling object in DR management by deriving precise load consumption patterns from historical load data;

(2) Acquiring a DP or RTP tariff which is can be utilised as the input control signal in DR management to enable efficient load scheduling;

(3) Developing effective DR strategies considering comprehensive external factors (*e.g.,* user preferences, energy storage conditions, scheduling priorities, *etc.*) to satisfy the requirements of different scales of households.

This thesis focuses on the research area of the smart grid and is expanded in accordance with these three interconnected challenges. The relationship between the mentioned smart grid challenges is presented in Figure 1.2.

Load pattern categorisation (LPC) is important to DR management. LPC involves recording users' energy usage over a predetermined time, period i.e. a load pattern, and then categorising users with similar usage into groups. The result of load pattern analysis is used to support assessment of the impact of DR programs [11, 12]. The detailed knowledge of electricity consumption's nature is essential in promoting strategies for peak load reduction, for instance exploiting the customers' willingness to accept price-based demand conditioning on the basis

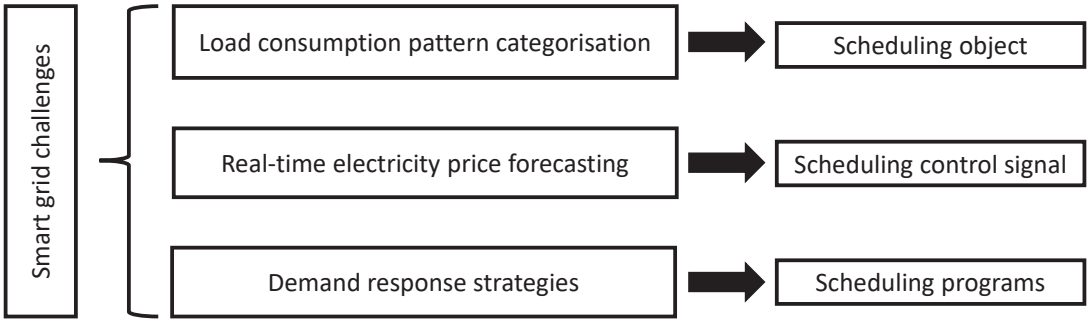


Figure 1.2: Three interconnected smart grid challenges that are presented in the thesis.

of DR programs [13, 14]. Therefore, it is a necessity to study load patterns at consumers' premises. The latest developments in smart metering technologies have given rise to a huge volume of load pattern data, which has greatly benefitted the task of research in LPC. The latest machine learning (ML) technologies are considered to offer potential solutions in performing the classification task.

In addition, implementing DR programs requires a reliable DP/RTP tariff as an input control signal. In general, the RTP tariff is usually provided as an instantaneous property with price rates varying continuously (typically hourly) to reflect wholesale market-price changes, however DR programs must be designed in advance. Hence, it is crucial to obtain a valid RTP tariff ahead of scheduling actions based on efficient forecasting models. The RTP tariff can be obtained by data-driven approaches. Afterwards, the predicted RTP can be utilised as the input control signal in executing DR programs.

As the core part of this thesis, advanced DR strategies considering energy storage facilities (*e.g.,* electric vehicles), vehicle to home (V2H) and vehicle to neighbour (V2N) communications, scale of households (*e.g.,* single household and multi-household network), and other external factors (*e.g.,* user preferences, EVs' charging behaviors, load scheduling priorities, *etc.*) are developed. The importance of DR programs cooperating with EVs increases with the increased population of EVs. Considering the flexible energy storage capacity of EVs, up to date DR strategies that can exploit EV characteristics are required. Compared with a conventional energy storage system (ESS) and other energy production

facilities, the utilisation of the EV as a temporary power source has advantages in employing flexibility and economic efficiency [9]. Use of the EV for this purpose should not require extra infrastructure investment other than the EV. In addition, DR strategies can be expanded beyond the individual consumer, enabling power-sharing via V2N. The available energy in one consumers EV can be shared to neighbours during peak price periods, which can be beneficial to both demand and supply sides. Therefore, EV assisted DR strategies hold wide prospects in practice not only for single households but also for multi-household networks such as small residential communities.

1.2 Objectives

In accordance with the motivation, the main goal of the thesis is to develop dynamic price based DR strategies for load scheduling of domestic appliances to meet the needs of different scales of households. The aim of this work is to balance the energy supply and demand, enhance the reliability of the smart grid, and create economic benefits for users. Since the scheduling object in the DR strategy is the load demand at the consumer side, an investigation of current load consumption patterns at a consumers' premises is explored first. A practical DP forecasting model is then proposed to provide a reliable DP tariff, which is then utilised as the input control signal in executing DR strategies. Finally, a number of DR strategies considering the impacts of EVs are designed. The objectives of the thesis can be summarised as follows:

(1) Investigate consumer load consumption patterns and develop robust ML based algorithms for LPC. The problems of cluster number determination and clustering technique selection will be addressed.

(2) Develop effective and practical RTP/DP forecasting models to acquire short-term electricity prices, which can be used as the input control signal for DR strategies.

(3) Design effective DR strategies and provide advanced DR based energy

management solutions for domestic appliances to meet the requirements of different scales of household:

- A fundamental DR strategy for a single household without EV assistance;
- An EV assisted DR strategy for a single household;
- An EVs assisted DR strategy for a multi-household network.

Comprehensive factors (*e.g.,* user preferences, energy storage conditions, scheduling priorities, *etc.*) have to be considered in developing these strategies.

1.3 Thesis Overview

1.3.1 Thesis Contributions

A number of interconnected challenges in the smart grid are addressed in this thesis. The main contributions of the thesis are summarised in the following:

(1) A parametric bootstrap (PB) algorithm for use in LPC.

An innovative PB algorithm is developed and evaluated in conjunction with compatible clustering techniques to determine the number of clusters within the load data and perform the clustering process.

First of all, the proposed PB algorithm is shown to be more robust against the dimensionality of the data in typical load pattern (TLP) analysis than conventional methods (*e.g.,* G-means [15]). It is able to effectively determine a cluster number for the data with a high dimensional space and therefore is applicable for the load demand data, which is usually of 24- or 48-dimensions.

Secondly, the proposed PB algorithm is general and independent of data type. It is more reliable and stable in cluster number determination than the Akaike Information Criterion (AIC) based algorithm [16], with a much higher probability of successfully finding an appropriate number and lower standard deviation (STD) value.

Thirdly, an effective cascade clustering scheme that classifies the initial load data into a series of sub-cascades according to external features is proposed to reduce clustering errors and improve efficiency over clustering raw data directly. The proposed PB algorithm is evaluated in conjunction with various classifying techniques [17–23], among which K-means++ is demonstrated to offer the best clustering performance in TLP analysis.

This work was published in [24] and is presented in Chapter 3.

(2) A hybrid model for short-term real-time electricity price forecasting.

A hybrid RTP forecasting model which is a consolidation of a least-square (LS) fitting model, grey prediction (GP) model and artificial neural network (ANN), is proposed. The LS fitting model considers the deterministic characteristic of the time series data and the GP model considers the stochastic characteristic. The ANN model is used for error optimisation, the execution of which is dependent on the forecasting performance of the first two stages.

The effectiveness and accuracy of the hybrid forecasting model is verified by numerical results in terms of a number of evaluation criteria and the results indicate that the proposed method is an accurate and efficient tool to predict the day-ahead RTP significantly outperforming previous methods [25–29].

This work was published in [30] and is presented in Chapter 4.

(3) Effective DR strategies and advanced energy management solutions at consumers' premises.

As the core of this thesis, a number of DP based DR strategies for load scheduling of domestic appliances are designed and implemented to different scales of households (*i.e.,* a single household and a multi-household network) in order to alleviate the load burden for the grid and save bills for householders, simultaneously.

For the single household network an EV is utilised as an auxiliary power supply (APS) for the energy consumption of home appliances. An EV-APS

model based DR strategy is proposed in accordance with a fundamental DR strategy.

For the multi-household network, an EVs assisted DR framework including a neighbour energy sharing (NES) model for a residential network with different types of EVs installed at consumers' premises is developed. The available EVs' energy distribution is enabled via vehicle-to-home (V2H) and vehicle-to-neighbour (V2N) connections. The NES based DR framework is valid and effective not only for an independent household but also for a multi-household residential network, which can satisfy broader requirements compared with conventional DR programs [31–35] in literature. The energy trading policy in the neighbourhood is also declared in this work.

Comprehensive affecting factors (*e.g.,* EV behavior, user preference, load scheduling priorities, *etc.*) are considered in scheduling for both EV assisted DR strategies. The effectiveness of the proposed DR strategies is verified by numerical analysis, which demonstrates that the proposed approaches significantly outperform the methods [8, 31] in literature in terms of load balancing and electricity cost reduction.

The related work has been accepted for publication as a book chapter on IntechOpen and was partially published in [9]. It is presented in Chapter 5.

1.3.2 Thesis Organisation

This thesis is organised as follows:

In Chapter 2, general backgrounds on the related topics of the smart grid are proposed. The existing achievements and potential challenges are discussed specifically.

In Chapter 3, an investigation of consumer LPC is performed based on ML algorithms. An innovative and robust PB algorithm is developed to solve the cluster number determination problem and a number of popular clustering techniques are also presented in this chapter.

In Chapter 4, a hybrid RTP forecasting model is illustrated in detail, to

provide an controlling input to DR strategies. Data-driven methods are utilised based on historical data in this chapter.

In Chapter 5, a number of DR strategies for load scheduling of domestic appliances are designed and implemented for different scales of household. Two significant models, an EV-APS model and a NES model are analysed. The effectiveness of the proposed DR strategies is evaluated by case studies.

In Chapter 6, conclusions are drawn and potential future works related to the evolution of the smart grid are discussed.

1.4 List of Publications

Referred Journal Publications

1. Xing Luo, Xu Zhu, and Eng Gee Lim. Parametric Bootstrap Algorithm for Cluster Number Determination of Load Pattern Categorisation [J]. Energy, 180:50-60, Apr. 2019.

2. Xing Luo, Xu Zhu, and Eng Gee Lim. A Hybrid Model for Short Term Real-Time Electricity Price Forecasting in Smart Grid [J]. Journal of Big Data Analytics, 3(1):8, Oct. 2018.

3. Xing Luo, Xu Zhu, Eng Gee Lim, and Yi Huang. A Semi-Blind Model with Parameter Identification for Building Temperature Estimation [J]. Journal of Cognitive Computation, 10(1):105-116, Feb. 2018.

Referred Conference Publications

1. Xing Luo, Xu Zhu, Eng Gee Lim, and Wolfgang Kellerer. Electric Vehicles Assisted Multi-Household Cooperative Demand Response Strategy [C]. In Proc. 2019 IEEE Vehicular Technology Conference (VTC)-Spring, Kuala Lumpur, Apr. 2019.

2. Xing Luo, Xu Zhu, and Eng Gee Lim. Dynamic Pricing Based and Electric Vehicle Assisted Demand Response Strategy [C]. In Proc. 2017 IEEE International Conference on Smart Grid Communications (SmartGridComm), pages 357-362, Dresden, Germany, Oct. 2017.

3. Xing Luo, Xu Zhu, and Eng Gee Lim. Load Scheduling Based on an Advanced Real-Time Price Forecasting Model [C]. In Proc. 2015 IEEE International Conference on Ubiquitous Computing and Communications (IUCC), pages 1252-1257, Liverpool, UK, Oct. 2015.

Book Chapter

1. Xing Luo, Xu Zhu, and Eng Gee Lim. Electric Vehicles Assisted Demand Side Energy Management [B]. Book of IntechOpen (accept for publication).

Chapter 2

Background

In this Chapter the general and theoretical background relating to each of the focussed themes within the thesis are introduced. This begins by looking at the reasons for and methods of load pattern categorisationin Section 2.1, is followed by work on dynamic price forecasting in Section 2.2 and finally addresses the concepts behind demand response management in Section 2.3.

2.1 Load Pattern Categorisation

2.1.1 Introduction to Load Pattern Categorisation

The ongoing development of smart grid technologies for data acquisition and supervision, metering, and communication, has given rise to a huge volume of load data. It is well recognised that this data can offer vast benefits with respect to load pattern analysis. Playing a significant role in load analysis, attention towards load pattern categorisation (LPC) has become increasingly high in recent years. LPC is frequently used in the process of classifying similar electricity consumption data into clusters. The clustering process is an unsupervised learning procedure which has been studied in various contexts and disciplines. It aims to allocate the electricity demand data of numerous consumers into a few homogeneous groups, ensuring that objects in the same group are similar while being dissimilar to objects from other groups. In other words, on the basis of some similarity criteria, the similarity of the objects within the same cluster is maximised and the similarity of objects from different clusters is minimised. Following LPC the

dataset each cluster can be represented by a composite object, which is normally the weighted average (centroid) of the objects within that cluster. Despite some fine details in the data maybe lost in this process, simplification and scalability is achieved.

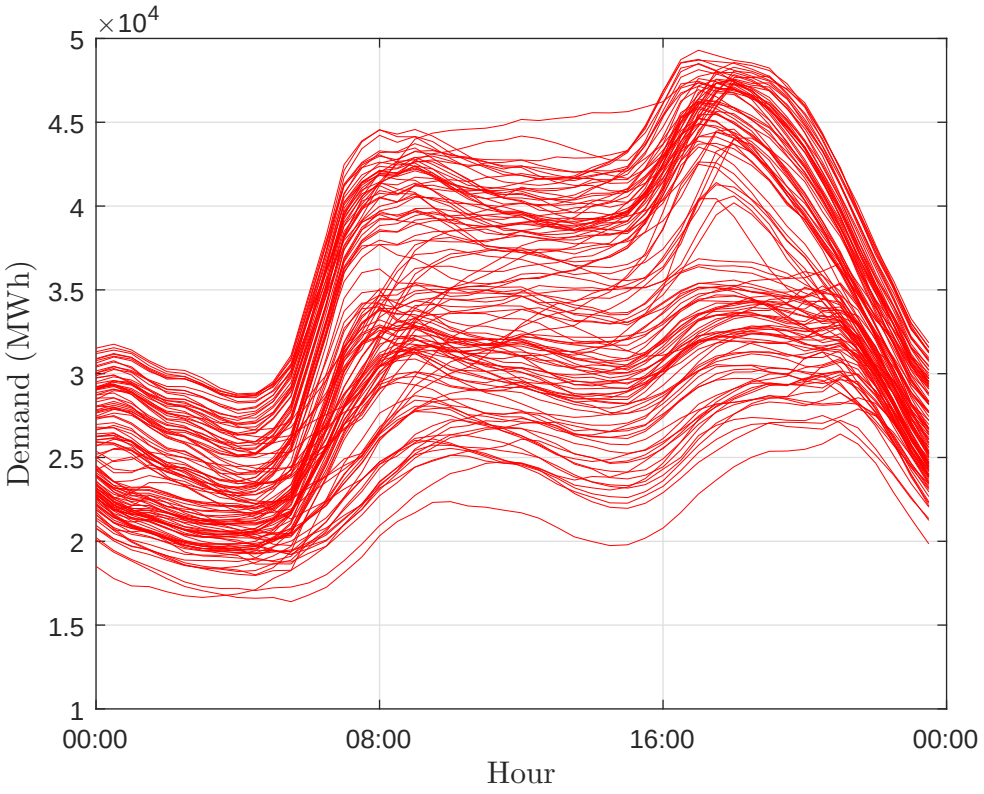


Figure 2.1: Examples of collected electricity demand data in the year 2016.

Figure 2.1 illustrates the electricity demand data of a number of users over a day in the year 2016, each of these datasets is used as an object to be clustered. Mathematically a dataset which consists of N objects as $\mathcal{S}$, $\mathcal{S}$ can be characterised as:

$$\mathcal{S} = \{x_n, n = 1, 2, ..., N, x_n \in \mathbb{R}^d\};$$

where d denotes the dimensionality which indicates the resolution of an individual load curve. Based on the recent metering techniques, the actual load data for a day is usually of 24- or 48-dimensions, which indicates the time interval of a load curve is either 1.0 hr or 0.5 hr, respectively. One of the main objectives of this work is to propose an efficient clustering technique to group the collected

load data into a number of clusters in order to seek typical load patterns (TLPs) from the dataset. The result of LPC is used to support the development of DR programs.

2.1.2 Applications of Load Pattern Categorisation

The attention towards the nature of electricity consumption is becoming increasingly high, since clustering or categorising of load consumption patterns has broad applications in terms of power system planning and operation, demand response (DR) and demand side management, load demand forecasting, electricity tariff design, *etc.* The applications of LPC can be generally summarised by the following aspects.

(1) Power system planning and operation: in most electricity markets, the electricity suppliers are operating in a competitive environment as the electricity distribution and supply services have been unbundled. The electricity suppliers need to get accurate information on the actual load demand of their users to allow the setting up of dedicated commercial offers, thus improving the planning and operation of the power system [36, 37]. Customer grouping on the basis of similar load demand patterns is likely to provide an effective solution.

(2) DR and demand side management: enhanced knowledge of LPC can be useful in supporting DR program development. LPC has been proposed as an effective means to enhance the targeting and tailoring of DR programs as well as providing reasonable load scheduling recommendations, owing to availability of advanced technology for load shifting and to emerging opportunities for flexible demand management, producing incentives and rewards to participating users [38–43].

(3) Load demand forecasting: LPC plays a crucial role in load demand forecasting (LDF) which is an essential part of power generation, distribution and regulation, since LDF is often estimated by the aggregation of typical load patterns (TLPs) which are the outcomes of LPC [44–46]. LDF usually

relies on available data from similar days. Obviously, an effective LPC can provide relevant information to identify such similar days hence improving the performance of LDF [47].

(4) Electricity tariff design: LPC is also proposed for the purpose of designing electricity tariffs. Electricity suppliers now have some degrees of freedom in formulating tariff offers which can meet the requirements set by the regulatory authorities. However, each tariff is formulated with reference to a specific load category, defined by a number of load characteristics [48, 49]. Additionally, LPC also can be used to assist electricity consumers in selecting an appropriate tariff [50].

LPC is an established yet still very active research topic due to its wide implications for the industry. However, the choice of a proper clustering algorithm and determination of an appropriate cluster number remains challenging.

2.1.3 Review on Existing Clustering Techniques

Having extensive applications in the industrial field, a wide variety of clustering techniques have been investigated in previous works and applied to electrical load data. There are many clustering technique classifications. According to different clustering objectives, the clustering technologies can be generally summarised into three categories: partition-based methods, hierarchical methods and model-based methods. A summary of clustering techniques used in various literature papers together with their relevant references, is shown in Table 2.1. On the application side, the mentioned clustering techniques differ in accordance with the principle used in the definition.

To begin with, the basic principle of the partition-based algorithm is to divide the dataset up into groups; one of the most popular representatives is the K-means algorithm. K-means clustering is a method of vector quantisation, originally from signal processing, it is popular for cluster analysis in data mining. It aims to partition N observations into K clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster [17, 18].

This results in a partitioning of the data space into Voronoi cells; however, the weaknesses of K-means are obvious: (i) The number of clusters, K, must be determined as an input parameter before clustering. (ii) K-means is sensitive to initial conditions. Different initial conditions may produce different clustering results and the algorithm may be trapped in a local optimum. (iii) K-means is not robust to outliers. The data points which are far from the centroid may pull the centroid away from the real one. (iv) If the amount of data is not sufficient, the initial grouping will significantly affect the clustering performance. The K-means algorithm is selected as one of the compatible clustering techniques investigated in support of the parametric bootstrap (PB) algorithm, the details of K-means are specified in Subsection 3.3.1.

Table 2.1: Summary of existing data clustering techniques.

Category	Methods	References
Partition-based Methods	K-means (KMS)	[17, 18].
	K-medoids (KMD)	[21].
	K-means++ (KMS++)	[19, 20].
	Fuzzy K-means (FKM)	[51, 52].
	Kernel K-means (KKM)	[53].
	Follow-the-leader (FTL)	[49, 54].
	DBSCAN	[55–58].
Hierarchical Methods	Agglomerative hierarchical clustering (AHC)	[59].
	Divisive hierarchical clustering (DHC)	[60].
Model-based Methods	Gaussian mixture model (GMM)	[22, 23].
	Dirichlet process mixture model (DPMM)	[61].
	Gaussian mixture copula model (GMCM)	[62].
	C-vine copulas based mixture model (CVMM)	[16].
Other Methods	Self-organising maps (SOM)	[63].
	Support vector clustering (SVC)	[64–66].
	Artificial neural networks (ANN)	[67, 68].
	Fast wavelet transformation (FWT)	[15, 69].

On the basis of the standard K-means algorithm, other generations of K-means, such as K-means++ (introduced in Subsection 3.3.2), fuzzy K-means, Kernel K-means, *etc.* have been developed, to improve the clustering

performance. The relevant algorithms have been demonstrated in detail in [19, 20, 51–53].

The K-medoids algorithm is a clustering algorithm related to the K-means algorithm and the medoid-shift algorithm. The K-means and K-medoids algorithms both are partition-based methods and attempting to minimise the distance between points labeled to be in a cluster and a point designated as the centroid of that cluster. In contrast to K-means, the K-medoids algorithm chooses data points as centers and works with a generalisation of the Manhattan Norm to define the distance between data points. It is more robust to noise and outliers compared with K-means because it minimises a sum of pairwise dissimilarities instead of a sum of squared Euclidean distances [21]. The details of the K-medoids algorithm are described in Subsection 3.3.3.

Density-based spatial clustering of applications with noise (DBSCAN) is a data clustering algorithm proposed in 1996 [55]. It is a density-based clustering algorithm and the basic principle can be simply described as: given a set of points in some space, it groups together points that are closely packed together (points with many nearby neighbours), marking as outliers points that lie alone in low-density regions (whose nearest neighbours are too far away). DBSCAN is one of the most common clustering algorithms and also most cited in scientific literature.

Different from the above partition-based clustering algorithms hierarchical clustering methods seek to build a hierarchy of smaller clusters. These methods can be divided into agglomerative hierarchical clustering [59] and divisive hierarchical clustering [60]. Agglomerative clustering starts with small, atomic clusters, that are gradually merged into bigger clusters, while moving up the overall hierarchy. Divisive clustering works the other way round, starting from one big cluster and clusters are stepwise separated into smaller clusters. The results of hierarchical clustering are typically visualised with dendrograms. In order to decide, where to split a cluster or which clusters to merge, a measure of dissimilarity has to be defined. In many hierarchical algorithms, this is done by the use of a distance metric and a linkage criterion.

An alternative approach, beyond partition-based and hierarchical clustering methods, is the use of distribution mixture models. Gaussian mixture models (GMM) is one of the most widely used model-based clustering approaches. GMM is a probability-based unsupervised learning classifier and is often used to classify a wide variety of signals. According to the central limit theorem [70–72], a given set of data is normally hypothesised as a mixture Gaussian distributions, since the mixture Gaussian distribution is able to approximate any probability distribution by increasing the number of the models in theory. In addition, the parameters of GMM are normally estimated by the expectation maximisation (EM) algorithm which is an effective machine learning method for probability density estimation [73–75]. An apparent drawback of GMM is that the convergence speed of using the EM algorithm to achieve the parameters of GMM is quite slow, particularly for large volumes of data with multi-dimensions, due to the seeds which are arbitrarily selected. The principle of GMM is shown in detail in Subsection 3.3.4.

Some other model types that can be used include, a multivariate Dirichlet process mixture model (DPMM) which was proposed to classify electricity profiles in [61]. A Gaussian mixture copula model (GMCM) which was introduced in [62], where a number of multivariate Gaussian copulas are fitted to a range of datasets. A C-vine copulas based mixture model (CVMM) was illustrated in [16] for the clustering of residential load data. Self-organising maps (SOM) [63], support vector clustering (SVC) [64–66], artificial neural networks (ANN) [67, 68] and fast wavelet transformation (FWT) [15, 69]. In Section 3.3, a number of selected compatible clustering techniques that are used to support the PB algorithm will be discussed in detail.

2.1.4 Review on Cluster Number Determination Methods

Clustering algorithms are useful tools for many important data processing tasks such as data mining, compression, probability density estimation, *etc.* However, a common problem in most existing clustering methods is that the number of clusters (called k) is a pre-defined parameter, which is difficult to set in practice. It not always clear what is the best value for k should be and

using an inappropriate cluster number as the input significantly reduces the clustering accuracy and increases complexity [76]. Therefore, reliable cluster number determination is required in advance of clustering.

Several algorithms have been proposed previously to determine k automatically and most of them are developed on the basis of K-means. Other methods that have been applied include X-means [77, 78], G-means [15, 79] and other methods that determine the cluster number by finding an "inflection point" based on specified criteria [80, 81].

A regularisation framework for learning k, which is also called X-means was first proposed in [82]. The algorithm searches over many potential k values and scores each clustering model using the so-called Bayesian information criterion (BIC) which can be defined as: $BIC(C|X) = \mathcal{L}(X|C) - \frac{p}{2}\log n$, where $\mathcal{L}(X|C)$ is the log-likelihood of the dataset X according to model C, $p = k(d+1)$ is the number of parameters in model C with dimensionality d and k cluster centers. Therefore, X-means selects the cluster number with the best BIC score based on the clustering results. The X-means algorithm was adopted for cluster number determination of load profiles based on smart metering data in [77]. Besides BIC, other scoring criteria such as Akaike information criterion (AIC), are also acceptable in the usage of X-means.

The Gaussian-means algorithm, also called G-means, provides another way of determining an appropriate cluster number. G-means starts with a small number of K-means centers and grows the number of centers. The first k value can be initialised to $k = 1$ or it can be selected as a larger value if the range of k is clarified. Each iteration of the algorithm splits into two centers whose data appears not to come from a Gaussian distribution via the Anderson-Darling (AD) test, which is a powerful 1-dimensional statistical test [83]. The splitting continues until the data in all clusters passes the AD test so that the expected cluster number can be obtained. The authors of [15] proposed a load pattern clustering strategy based on the wavelet transformation and G-means to determine the cluster number. In this way, the adopted load data of N-dimensions has to be reduced to a single dimension, as the G-means algorithm is not effective for high dimensional data.

However, actual load data is typically of 24- or 48-dimensions (representing one day), and the dimensional-reduction always gives rise to the risk of information loss. Therefore, the G-means algorithm is not normally suitable for data that is multi-dimensional.

Another popular way to determine an appropriate cluster number is to find an "inflection point" by certain criteria, such as the AIC based method or BIC based method. A mixture model for residential load data clustering was presented in [16]. The Authors selected the optimal cluster number by seeking the first "knee" always at the local maximum of the curve of the AIC or BIC. However, the criteria based algorithm is not reliable and does not guarantee a cluster suitable number will be found at all times, since the estimated "inflection point" normally varies in a range in AIC calculations.

2.2 Dynamic Price Forecasting

2.2.1 Introduction to Dynamic Electricity Price

Dynamic price (DP), also referred to as demand based price, surge price or time-based price, is a pricing strategy in which businesses offer flexible prices for their services or products on the basis of the current market demands. The product providers are able to change prices based on the internal algorithms that consider a variety of factors such as competitor price, demand and supply relationship, as well as other external factors in the market. DP is a common practice in several industrial areas including hospitality, transportation, entertainment, retail, electricity market, *etc.* Each industrial area uses a slightly different set of criteria to reprice their products according to the demand for the market.

In the electricity market, dynamic electricity price, also called dynamic tariff, spot electricity price, or real-time price (RTP) is one critical use case of DP and it was first introduced in the 1980's [84]. Nowadays it is being tentatively applied to the power system in many countries including the US, Australia, *etc.* The RTP tariff is an inexorable trend in the next generation of power

system reformation [85, 86]. Unlike regulated markets in which the companies determine prices independently, electricity prices are significantly dependent on a supply-demand relationship in a deregulated market. Generally speaking, the principle of RTP is that it offers higher prices during peak load demand periods and provides lower prices during off-peak load demand periods [87]. In consideration of the generation cost in different load levels, the DP tariff is a potential load management method for properly allocating incremental prices of electricity consumption to the time of delivery, thus ensuring the overall economic rationality [88].

Considering time scales, electricity price forecasting (EPF) can be classified into ultra-short-term, short-term, medium-term and long-term forecasting horizons [89]. The ultra-short term is from several minutes to 1 hour ahead forecasting. Short-term forecasting generally involves horizons from 1 hour to several hours or a few days ahead, which is important in day-to-day market operations. From a few days to a few months ahead is generally defined as medium-term forecasting. Beyond that, it is the long-term forecasting. It is usually measured in months, quarters or even years and focuses on investment profitability analysis and planning. This thesis focuses on day (24 hours) ahead EPF with a resolution of 0.5 hour, which belongs to short-term forecasting.

2.2.2 Applications of Dynamic Electricity Price Forecasting

DP is regarded as the next generation of electricity tariff. Therefore, the applications of DP/RTP have been widely discussed in recent years.

The DP/RTP tariff can be broadly utilised as a basic control signal to support DR management, which offers an excellent long-term solution to improving energy efficiency and reducing wastage [90]. On the one hand, the DP/RTP tariff is beneficial to the power grid as it offers specific price instructions for participants to average their power usage over different time periods alleviating the load burden of the power grid, especially in peak demand time. On the other hand, such an electricity tariff encourages consumption by price reduction

during periods of abundance and allows customers to have multiple choices in determining the time of electricity consumption. The participants in the electricity market can regulate the operating time of electric devices automatically or manually during high-price periods and gain benefits from low-price periods via DR management, thus achieving the aims of reducing energy usage and saving on their electricity bills [3, 91]. Therefore, the research on the RTP tariff is of interest to researchers, production companies, investors, independent market operators and large industrial consumers [92, 93].

Electricity is a very special commodity, as it is economically non-storable a constant balance between energy production and consumption is required to ensure the stability of the power system. Electricity demand significantly depends on a number of external factors such as weather conditions (*e.g.,* temperature, solar radiation, wind speed, *etc.*) and human activity. As a consequence, these unique characteristics in price dynamics are not easily observed in the market. Therefore, a reliable electricity price forecasting approach is required. In recent years, price forecasting from a few hours to a few months ahead has become of particular interest to power suppliers. On the one hand, a power company able to forecast the volatile wholesale prices with a reasonable level of accuracy and can adjust its bidding strategy and its own production or consumption schedule in order to reduce risk or maximise profits in day-ahead trading [94]. On the other hand, a reliable electricity tariff also can be used to assist consumers in adjusting their schedule in order to reduce electricity cost.

2.2.3 Review on Forecasting Methods

A wide variety of methods have been conducted in previous works on EPF, with varying degrees of success. A taxonomy of electricity price forecasting and modeling approaches in related works is presented in Figure 2.2. According to the figure, the forecasting approaches can be broadly divided into six categories as multi-agent models, fundamental models, statistical models, reduced-form models, computational intelligence models and hybrid models.

First of all, multi-agent models simulate the operations of a system including

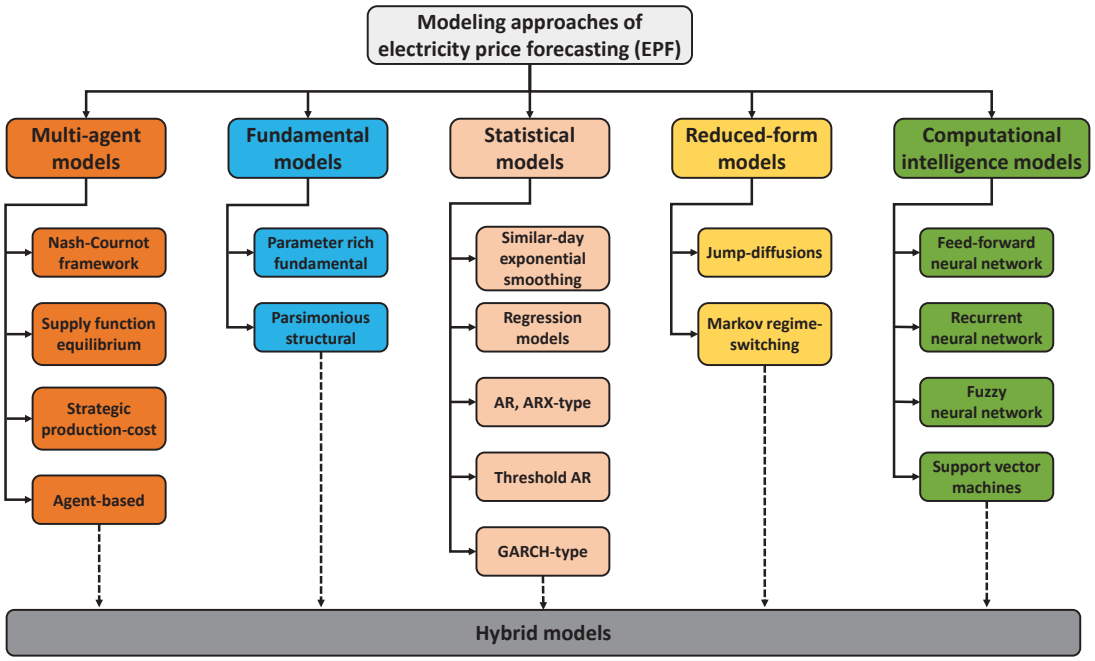


Figure 2.2: A taxonomy of electricity price forecasting and modeling approaches in previous works.

heterogeneous agents (*e.g.,* generating units, companies, *etc.*) interacting with each other, and establish a price model by matching the demand and supply in the market [95]. The members of this category include the Nash-Cournot framework, supply function equilibrium (SFE), strategic production-cost models (SPCM) and agent-based models. The relevant models have been analysed in [96–98].

Secondly, fundamental models try to capture the relationship between the basic physical layer and the economic layer, which represents the production and trading of electricity. The associations between fundamental drivers (*e.g.,* weather conditions, system parameters, load status, *etc.*) are assumed, and the fundamental inputs are modeled and predicted independently, often via statistical, reduced-form or computational intelligence techniques [99]. The parameter-rich models [100] and parsimonious structural models [101] are two subclasses of fundamental models. The main weakness of fundamental models in the practical implementation is that the physical and economic relationships are assumed in establishing models so that the price projections generated by these models are very sensitive to the assumptions.

Reduced-form models try to characterise the statistical properties within electricity prices. The main objective of these kinds of models is not to offer a precise price forecasting result, but rather to replicate the main characteristics such as dynamic regularity of prices, correlations between commodity prices and marginal distributions at future time points. The jump-diffusion model [102–104] and Markov regime-switching model [105] are two popular subclasses of reduced-form models. The main issues with reduced-form models is that the model performance is highly dependent on the input price data. If the chosen price data is not appropriate for capturing the main characteristics of electricity prices, the obtained prices from the models are not likely to be trustable.

The most popular models used for price forecasting are statistical models since the introduction of data metering technologies has provided great convenience in data collecting. Statistical models forecast future prices by using a mathematical combination of the previous prices and previous or current values of exogenous factors, typically consumption and production figures, or weather variables [106]. Statistical models constitute a very rich class which includes similar-day and exponential smoothing models [25], regression models [107], time series models including the auto regressive (AR) model, auto regressive moving average (ARMA) model, auto regressive integrated moving average (ARIMA) model, fractional ARIMA (FARIMA), seasonal ARIMA (SARIMA), threshold AR (TAR) [25, 26, 108–114], and heteroskedastic time series models such as the generalised auto regressive conditional heteroskedasticity (GARCH) model and the AR-GARCH model [115, 116]. Statistical models are very interesting as some physical interpretation may be attached to their components, thus allowing engineers and system operators to understand their behavior. The performance of these models is restricted by their ability to analyse the non-linear characteristics of the electricity prices, in practice they typically do not perform better than non-linear computational intelligence models. [106].

Computational intelligence models for forecasting have been developed in recent years. They are intelligence-based, non-parametric and non-linear statistical techniques that combine elements of machine learning, evolution and

fuzziness to create new approaches. These models are capable of adapting to complex and dynamic systems. The main classes of computational intelligence models for EPF are artificial neural networks (ANN) (*e.g.,* feed-forward neural networks, recurrent neural networks, *etc.*) [117, 118], fuzzy systems [117, 119] and support vector machines (SVMs) [120–123]. The advantage of computational intelligence models is their ability to solve the complex and non-linear systems. In general, these models are better at modeling the features of electricity prices in comparison to the statistical models [117].

From Figure 2.2, the last category of EPF approaches is called hybrid models which combines techniques from two or more of the groups as discussed. However, the shared limitation of most studies mentioned above is that vast amounts of historical RTP data are required for building or training the model. Insufficient historical data causes considerable estimation error. Hence, the research into EPF in this thesis mainly concentrates on building an effective estimation model for EPF in a smart grid with comparatively high accuracy by using limited sets of historical data. The details of the proposed approach are illustrated in Chapter 4.

2.3 Demand Response Management

2.3.1 Introduction to Demand Response

It is widely recognised that a reliable power grid requires a perfect balance between real-time supply and load demand. It is a difficult task in the current system, since both supply and demand change unexpectedly and rapidly for various reasons, such as a power generation unit forced outage, transmission and distribution outage, abrupt load demand change. Demand response (DR) which is an essential characteristic of the smart grid, is regarded as an excellent long-term solution for balancing energy supply and demand, improving energy efficiency, reducing wastage by providing energy-efficient equipment, and encouraging energy-aware consumption.

The concept of DR can be descried as changes in electricity consumption

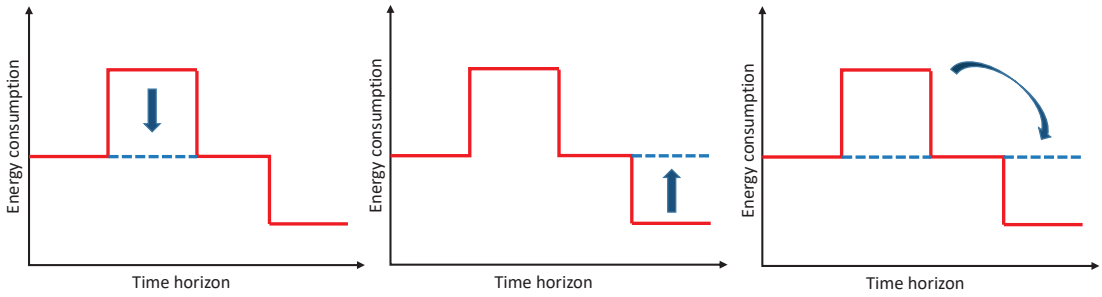
by end-use customers from their normal consumption patterns in response to a DP/RTP tariff over time or other form of financial incentive in an automated and convenient manner. It has also been defined as incentive payments designed to include lower electricity consumption at times of high wholesale market prices or when system reliability is jeopardised [2, 124–126]. From the smart grid perspective, DR management is an effective means of rescheduling users' energy consumption to reduce electricity expense. This innovative technology is able to improve the efficiency of the electricity market, make the power system more reliable and lead to mutual economic benefits for both power utilities and the energy consumers. Last but not least, DR is an eco-friendly technique as it will reduce carbon emissions by enabling efficient utilisation of current grid capacity.

There are three general DR categories, presented as follows (also illustrated in Figure 2.3).

(1) Peak clipping: Reducing users' electricity consumption during critical peak periods, in order to prohibit the load from exceeding the supply capacity of distribution stations and alleviate load burden. However, this action involves the temporary discomfort of users, since peak clipping cuts down some of their demand. An example of this response is achieved when the thermostat setting within a heater or an air-conditioner is temporarily changed.

(2) Valley filling: In contrast to peak clipping, valley filling is to promote off-peak energy consumption through energy storage facilities, such as through use of a solar energy reservoir, rechargeable batteries or plug-in electric vehicles.

(3) Load shifting: Aims to shift energy consumption from peak demand periods to off-peak demand periods, without reducing the total energy consumption within a day. Load shifting can be regarded as the combination of peak clipping and valley filling. An example would be shifting the operating time of some household appliances (*e.g.,* dishwashers, hot water tank, washing machine, *etc.*) to off-peak periods.



(a) (b) (c)

Figure 2.3: Demand response categories. (a) Peak clipping; (b) Valley filling; (c) Load shifting (combination of peak clipping and valley filling).

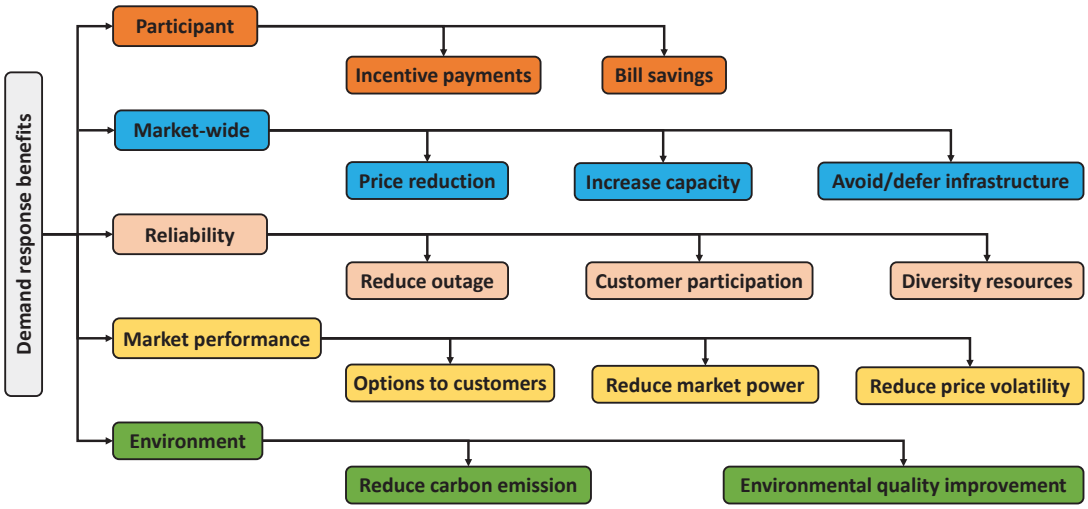


Figure 2.4: Demand response benefits in five aspects: participant, market-wide, reliability and market performance and environmental benefits.

2.3.2 Benefits of Demand Response

Based on the concept of DR, it is obvious that many potential benefits can be expected from DR programs. Those benefits are categorised into five general aspects: participant, market-wide, reliability, market performance and environmental benefits, as illustrated in Figure 2.4.

Firstly, electricity consumers can obtain benefits directly by participating in DR programs. Electricity bills can be reduced if customers are willing to shift the operating time of flexible electrical appliances from peak periods to off-peak periods. In some cases, participants might experience bill savings without changing their consumption patterns. This will be achieved if their normal

consumption during high price periods is lower than average [126]. Meanwhile, some participants are able to increase their total energy consumption without increasing the payment by operating more appliances during off-peak periods. In addition, participants in classical incentive based DR programs can obtain incentive payments for their participation, which is also a great benefit for DR participants.

The benefits of DR programs are not only for participants but also extend market-wide. Based on efficient DR programs, an overall electricity price reduction is expected, which is due to an efficient utilisation of available infrastructures. DR programs (*e.g.,* market based programs) are able to increase the short term capacity, which can avoid or defer capacity costs. All of the avoided or deferred costs will be reflected in the electricity price for all customers.

DR programs are also beneficial for enhancing the reliability of the smart grid. By installing an efficient DR program, on the one hand, participants are involved and have the opportunity to help in reducing the risk of forced outages and interruption. On the other hand, the operator may have more options and resources to maintain the reliability of the system, thus reducing forced outages and improving grid stability.

It is expected that there will be an improvement in electricity market performance in response to the application of DR programs. Participants of DR programs may have more choices in scheduling their home appliances. As a prime driver for many utilities to provide DR programs, the change of consumption behaviors affect the market performance directly, especially for the market-based programs and the dynamic price based programs [127]. In addition, another significant market improvement brought by DR programs is the reduction of price volatility in the spot market. According to the California electricity crisis report [128], a small reduction in electricity demand by 5% will result in a 50% price reduction. This is due to the fact that electricity generation cost increases exponentially near maximum generation capacity. Therefore, a small reduction in demand will lead to a significant reduction in generation cost, affecting the electricity market and reducing the electricity price.

Finally, as an eco-friendly technique, DR will reduce carbon emissions, by enabling a more efficient utilisation of current grid capacity. Other environmental benefits include better land utilisation as a result of avoiding the need for new electricity infrastructures such as generation units and transmission or distribution lines, water and air quality improvement as a result of more efficient utilisation of energy sources, and reduction of natural resource depletion.

2.3.3 Review on Demand Response Programs

DR programs are regarded as the means that the power utility takes to incentivise users to reschedule their energy usage patterns [129]. In other words, an efficient DR program is able to shape consumers' electricity load profiles so as to enhance the reliability, stability, and efficiency of the grid. Having numerous benefits, as discussed above, much research on DR programs has been conducted in literature. DR programs can be mainly divided into the following two branches as described in Figure 2.5.

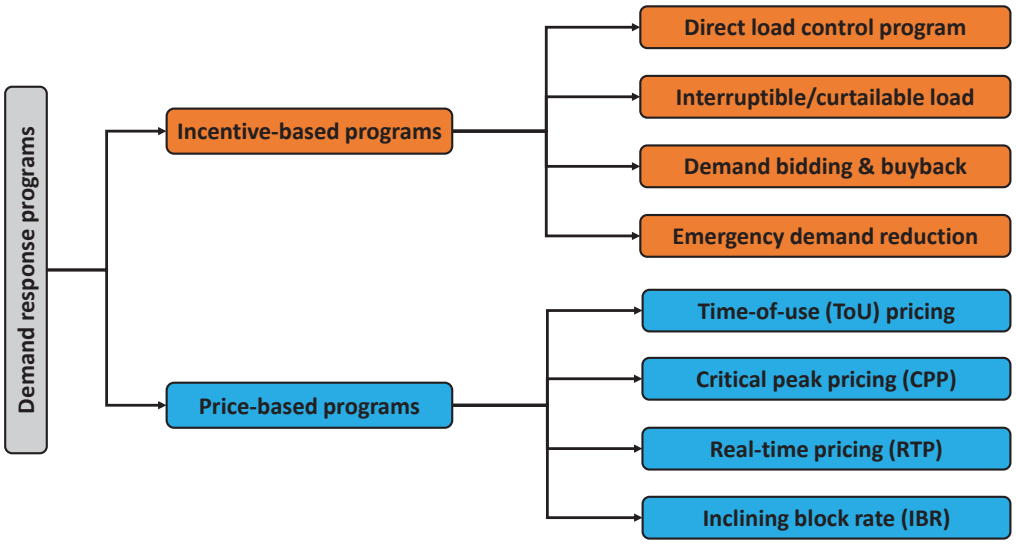


Figure 2.5: Summary of existing demand response programs.

A. Incentive-based programs: Incentive-based DR programs pay participating users for demand reduction, triggered by peak load or system contingencies. Participating users receive incentive payments usually as a discount rate

or a bill credit for their participation in the programs. Several typical incentive-based programs are listed as follows:

(1) Direct load control (DLC): in the DLC program, the power utility is able to remotely control certain home appliances, such as air conditioner and water heater. The participant in the program will obtain incentive payments as a reward [130]. The initial idea of the DLC program is to alleviate load burden during peak hours and it has been offered to residential and small commercial customers for decades.

(2) Interruptible/curtailable load: similar to the DLC program, participating users in interruptible/curtailable load programs will receive a certain incentive discount on electricity bills as a reward if they agree to cut down some portion of their interruptible/curtailable loads when grid reliability is jeopardised [131]. Participants will be asked in advance and participants who do not respond may face penalties depending on the pre-defined terms and conditions.

(3) Demand bidding: in demand bidding (also called a buyback) program, users can benefit from cost reduction if they allow a curtailing of electricity consumption at a specific bid price in the case of peak demand or system contingencies. The accepted bid is always less than the market price [132]. This kind of program is mainly provided to larger users ($\geqslant$ 1 MW) and small users require a third party or agent to unite them to bid.

(4) Emergency demand reduction: in this program, participating users can obtain incentive payments for their load reductions during emergency conditions, such as reliability accidents when the grid is out of reserve [133].

B. Price-based program: the price-based DR program is established on the basis of a dynamic pricing rate in which electricity tariffs are not flat. It provides users with different electricity prices in different time periods as guidance of energy consumption. Based on such information, users will be able to schedule

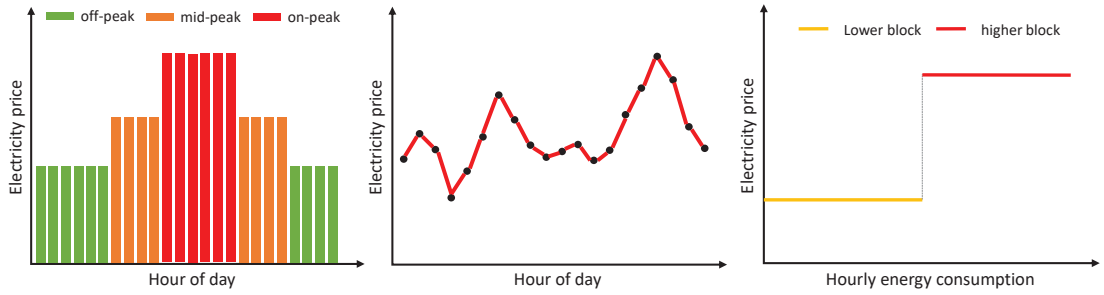
the operating time of their electric appliances and use less electricity when electricity prices are high, thus reducing the demand at peak periods. In other words, the price-based program indirectly induces users to dynamically change their electricity consumption patterns according to the variance of electricity prices, instead of directly controlling their loads. A number of typical pricing rates are as follows (also illustrated in Figure 2.6):

(1) Time-of-use (ToU) pricing: the ToU tariff is the basic type of price-based program. Under ToU pricing, electricity consumers are charged at different prices as they consume electricity at different intervals of a day [134]. The time interval is usually longer than 1 hr. The simplest ToU pricing has two time blocks: on-peak period and off-peak period. A more complicated ToU pricing may include more time blocks as illustrated in Figure 2.6(a). Normally, the electricity price at the on-peak time block is much higher than other time blocks, in order to induce users to reduce their electricity usage at peak demand time. ToU pricing is usually released far in advance and will not be changed for a long time.

(2) Critical peak pricing (CPP): the CPP tariff, also called contingency electricity pricing, is designed based on ToU pricing. It includes a pre-defined electricity price which is much higher than normal flat rates. The design purpose of CPP is to guarantee reliability for the power system and balance demand with supply. Therefore, the CPP tariff is employed only for a limited number of hours or days, when grid reliability is jeopardised [135].

(3) Real-time pricing (RTP): the RTP tariff, also referred to as dynamic tariff, where the electricity price usually varies at different time intervals of a day as illustrated in 2.6(b). The RTP tariff is typically released on a day-ahead or an hour-ahead basis. The RTP tariff is widely considered to be the most direct and efficient price-based program in the current competitive electricity market [136, 137]. The proposed DR programs in this thesis are based on the RTP tariff.

(4) Inclining block rate (IBR): IBR is designed with a two-level rate structure (lower rate and higher rate) as shown in 2.6(c). The electricity price is significantly related to energy consumption. If the user's hourly/daily/monthly energy usage exceeds a certain threshold, the electricity price will climb up to a higher value, which means the more electricity a user consumes, the more money paid per kWh [138]. IBR encourages energy consumers to reduce their electricity consumption. This tariff has been broadly adopted in some areas of many countries including the USA, UK, and Canada since the 1980s [138].



(a) (b) (c)

Figure 2.6: Illustration of time-based pricing tariffs. (a) Time of use (ToU) pricing; (b) Real-time pricing (RTP); (c) Inclining block rate (IBR).

2.3.4 Review on Demand Response Approaches

DR management is usually formulated as optimisation problem of utility maximisation or cost minimisation, which can be solved by various approaches. Recently investigated DR approaches include convex optimisation, game theory, dynamic programming, Markov decision process, stochastic programming, and particle swarm optimisation. However, the selection of the DR approach is dependent on how optimisation problem will be formulated. In this thesis, the optimisation problem in DR management is formulated as a convex problem and mixed-integer linear programming (MILP) is used as the optimisation technique.

Convex Optimisation

Convex optimisation is currently the most popular and widely applied DR approach. It is the problem whose objective and constraint functions are convex. A convex problem means an optimum solution for the objective function can be found for certain. Mathematically, it is usually defined as $\min_x f_0(x)$ under constraints $f_i(x) \leqslant b_i$, $i = 1, ..., m$, and $f_0, ..., f_m : \mathbb{R}^n \to \mathbb{R}$ are all convex functions. As mentioned above, the DR is usually formulated as cost minimisation or utility maximisation where the cost function is convex and the utility function is concave. However, the problem of maximising a concave function f can be reformulated equivalently as minimising the function $-f$, which is convex.

In addition, the constraint functions of DR are always convex, for example, the energy demand of a single household is constrained by lower and upper bounds [139–141] where the minimum energy consumption level denotes the baseline demand from the must-run electric appliances and the maximum energy consumption level usually represents the total energy demand if all home appliances are turned on. The householder may be concerned about whether specific tasks need to be finished within a specific time period. For example, the dishwasher after lunch should complete the task of washing dishes before dinner or the drying machine should finish the task of drying after clothes have been washed. Further, the total energy demand of all households in the grid network is limited to a maximum supply threshold for safety consideration. In addition to these, considering the energy storage facilities (*e.g.,* conventional battery, electric vehicle, *etc.*) and renewable energy resources (*e.g.,* solar energy, wind energy, *etc.*) into the DR programs, more constraints should be accounted for. All of these constraints are linear problems, therefore, DR can be formulated as a convex optimisation problem.

With the development of computing and optimisation theory, convex optimisation can be solved by linear programming (LP) straightforwardly. Linear programming is a mathematical modeling technique in which a linear function is maximised or minimised when subjected to various constraints. LP is applicable

in decision situations where quantities (variables) can take any real values only restricted by linear equalities (*e.g.,* for representing capacity constraints). LP techniques have been widely adopted in solving convex optimisation problems. For example, in [142], the authors proposed an optimisation method to the demand side energy management of a given consumer, such as an industrial compound or a university campus. The objective of the proposed method is to maximise the utilisation of the cluster of demands. In order to solve this optimisation problem, an LP approach which allows the cluster of demand to buy, store and sell energy at appropriate times to adjust the load levels, has been implemented. In addition, a $\{0, 1\}$ LP method for fixed-profile load scheduling and demand management is presented in [143]. The problem in this work involves organising the requested load over time, without modifying the load profile of individual requests, while satisfying constraints on the transient response of the system to changes in the load. Moreover, in [144], the profit-maximising DR of an energy load in the real-time electricity market was considered. To tackle the high computational complexity, the authors proposed a dual approximate approach that transforms the optimisation problem into an LP problem by exploiting the threshold structure of the optimal solution.

In summary, convex optimisation is regarded as one of the most relevant approaches to the DR problem. However, there are other methodologies in addition to convex optimisation can be used in DR and they are briefly discussed in the following subsections.

Other Approaches

(1) Game Theory. Game theory is another popular approach used in solving the DR problem. Game theory is a study of mathematical models of strategic interaction between rational decision-makers [1, 145]. It has applications in all fields of social science, as well as in logic and computer science. A game $\mathcal{G}$ includes three fundamental components: players $\mathcal{N}$, strategies $\{\mathcal{X}_i\}_{i\in\mathcal{N}}$ and payoff functions $\{W_i\}_{i\in\mathcal{N}}$. Each player $i \in \mathcal{N}$ will select a strategy $x_i \in \mathcal{X}_i$ to maximise the payoff $W_i(x_i, \boldsymbol{x}_{-i})$, which is dependent on both the strategy

x_i and other players' strategies $\boldsymbol{x}_{-i}$. In DR, the game participants involve the power utilities and energy consumers. Related works were presented in [146–148].

(2) Dynamic Programming. Dynamic programming decomposes a complicated problem into a number of subproblems, and each subproblem can be solved backward over the stage. The method was developed by Richard Bellman in the 1950s and has found applications in numerous fields, from aerospace engineering to economics. The advantage of this method is that less time is consumed compared with heuristical methods, particularly for some subproblems with overlapping characteristics. The related works on the dynamic programming approach have been performed in [149–153].

(3) Markov Decision Process. A Markov decision process (MDP) is a discrete time stochastic control process and it is an extension of Markov chains. It provides a mathematical framework for modeling decision-making in situations where outcomes are partly random and partly under the control of a decision maker. In DR management, the exact future electricity prices (adopted as a control signal) are usually unknown, but the prices can be obtained from a large number of historical price data. According to the uncertainty of prices, the problem of home appliance scheduling for minimising the total electricity cost over a typical day can be naturally cast as a Markov decision process [154]. The related works have been stated in [155–157].

(4) Stochastic Programming. In the field of mathematical optimisation, stochastic programming is a framework for modeling optimisation problems that involve uncertainty, taking advantage of the fact that probability distributions are known or can be estimated. It has been utilised in many previous papers related to DR management. Stochastic programming can be regarded as a special type of dynamic programming. It is able to address the time-varying parameters in DR whose probability distributions are known or can be estimated. As one of the effective approaches in solving DR problems,

stochastic programming can make wise decisions on the basis of statistic knowledge about random parameters. Examples are presented in [158–160]

(5) Particle Swarm Optimisation. The last approach that has been considered for DR is particle swarm optimisation (PSO). PSO is originally attributed to Kennedy, Eberhart, and Shi [161, 162] and was first intended for simulating social behavior [163], as a stylised representation of the movement of organisms in a bird flock or fish school. PSO optimises a problem by iteratively trying to improve a candidate solution (particle) with regard to a given measure of quality. Each candidate solution is driven to move around, with its trajectory affected by its experience and other candidate solutions, towards the best solution. Each particle can be dynamically adjusted by inertia, personal best and group best. The related works have been performed in [164–168].

The objectives of this work involve reducing consumer electric bills and alleviating grid load burdens during peak demand periods. It is therefore reasonable for the objective function of the DR design to be formulated as a cost minimisation where the cost function is convex. Additionally, the relevant constraints (*e.g.,* maximum energy consumption level, capacity of EV battery, *etc.*) are also convex. Based on this, and its relative high efficiency and ease of implementation in comparison to the other methods convex optimisation is adopted in this work.

2.3.5 EV Impacts on Demand Response

Electric vehicles (EVs) are a growing trend for next generation transportation, due to their economic and environmental benefits, and the rapid advance of rechargeable battery technology [169–171]. Along with the worldwide application of DP, an increasing adoption of EVs in residences brings about both opportunities and challenges for the smart grid. Residences with EVs consume more electricity and react more elastically to electricity price changes [172]. According to the report provided by the U.S. Energy Information Association

[173], the fast charging of an EV is equivalent to about 120 houses coming on line for half an hour, which is a severe issue for the power grid. However, the usage of EVs as energy storage units via vehicle-to-home (V2H) offers an effective solution to load shaping at the demand side. In addition to this, the energy available from an EV can be delivered to a neighbour via a vehicle-to-neighbour (V2N) network if it is enabled. Hence, householders are able to participate in load scheduling and may have multiple options in energy allocation.

The importance of ensuring DR programs are developed to cooperate with EVs has increased together with the recent prevalence of EVs in society . Considering the flexible energy storage capacity of EVs, up to date DR strategies which take the behaviors of EVs into account are required. The implementation of DR with EVs requires efficient energy distribution management and high-performance batteries as a basis. Moreover, DP provides a basic control signal to optimally schedule the charging and discharging of EVs, by minimising the overall cost.

Compared with the conventional energy storage system (ESS) and other energy production facilities, the utilisation of an EV as a temporary power source has advantages in employing flexibility and economic efficiency [9]. It should not require extra investment besides the daily used EV. Meanwhile, power-sharing can be enabled through a V2N connection. The energy available from EVs could be shared to a neighbour during peak price times and benefit both sides. Therefore, EV assisted DR strategies hold strong prospects in practice not only for a single household but also for a multi-household network.

Much research has been conducted on DR and there are many popular DR strategies considering EV impacts being presented in the literature. For example, in [31], an optimisation framework based DR program was proposed, with high penetration of EVs and storage systems from a residential customer's perspective as well as the utility company's perspective. In [32], the authors focused on EVs' charging behaviors based on the data collected from EV charging sessions and different types of charging behaviors were derived. Nonetheless, specific DR programs with proposed charging profiles have not yet been published. To analyse

the potential usage of EVs in the power grid, the optimal time of EVs' charging and discharging was explored in [33]. However, all the mentioned studies are limited to the operation of a single user and fail to attempt the scheduling of EVs among a group of households in the DR program.

Moreover, [34] proposed an algorithm for EVs' scheduling in DR to optimise the peak demand. The optimisation problems are studied in a game framework. However, other electric appliances have not been considered in this work. In [35], an intelligent preemptive DR management scheme using a building energy management system was proposed to better schedule the energy consumption within a building. In this work, dynamic EV charging scheduling, priority-based load shedding, and air conditioning system were considered. Paper [174] presented an optimal behavior of plug-in EV parking lots in the energy and reserve market. Both price-based and incentive-based DR programs were developed and uncertainties of plug-in EVs were also considered by using the stochastic programming approach. In addition to these, [175–177] also described a number of interesting DR programs coordinating with EVs.

Although much research has been conducted on DR management, considering the impact of the EV on DR programs still remains challenging and worth studying.

Chapter 3

A Parametric Bootstrap Algorithm for Load Pattern Categorisation in Smart Grid

In this Chapter the scheduling object in demand response (DR) is clarified by deriving precise load consumption patterns. An innovative parametric bootstrap (PB) algorithm incorporated with a compatible clustering technique is proposed, to address the cluster number determination problem, as well as simultaneously classifying load demand data in load pattern categorisation (LPC). Typical load patterns (TLPs) are extracted and the categorising performance is evaluated. The main contributions of this work can be summarised as follows.

(1) The proposed PB algorithm is more robust against dimensionality of the data in LPC than conventional methods (*e.g.,* G-means [15]). It is able to effectively determine the cluster number for data in a high dimensional space and therefore is applicable for the load demand data, which is usually of 24- or 48-dimensions.

(2) The proposed PB algorithm is general and independent of the data type. It is more reliable and stable in cluster number determination than the AIC based algorithm [16], with a much higher probability of successfully finding an appropriate cluster number and lower standard deviation (STD) value.

(3) An effective cascade clustering scheme that classifies the initial load data into a series of sub-cascades according to external features is proposed to

reduce clustering errors and improve efficiency over clustering the raw data directly. Besides, the proposed PB algorithm is incorporated with various classifying techniques [17–23], among which K-means++ demonstrates the best clustering performance in LPC.

The rest of this chapter is organised as follows. In Section 3.1, the cascade clustering scheme is presented specifically. In Section 3.2, a parametric bootstrap algorithm is proposed in detail for the determination of the cluster number. Four compatible clustering techniques are presented and compared in Section 3.3. Additionally, the proposed approach is assessed in Section 3.4. The verified approach is applied to the actual load data to address the cluster number determination problem and the objective TLPs are obtained in Section 3.5. Finally, this work is summarised concisely in Section 3.6.

3.1 Cascade Clustering Scheme for Load Data Processing

A cascade clustering scheme that comprises two major stages for load data processing is illustrated in Figure 3.1.

Stage 1: A set of load demand data is classified into a series of sub-cascades based on the external features including seasonality and day characteristics.

Stage 2: Objects in an individual sub-cascade are further clustered into numbers of clusters based on the PB algorithm incorporated with a compatible clustering technique.

Based on the observations of load demand at different time periods in the UK (as shown in Figure 3.2), the total load demand of electricity consumers is significantly influenced by the following apparent external factors:

(1) The load demand for weekends is evidently less than that of weekdays, even though the trends are similar.

(2) The weekly periodicity of the load series is broken by the occurrence of a bank holiday, as shown in Figure 3.2(b).

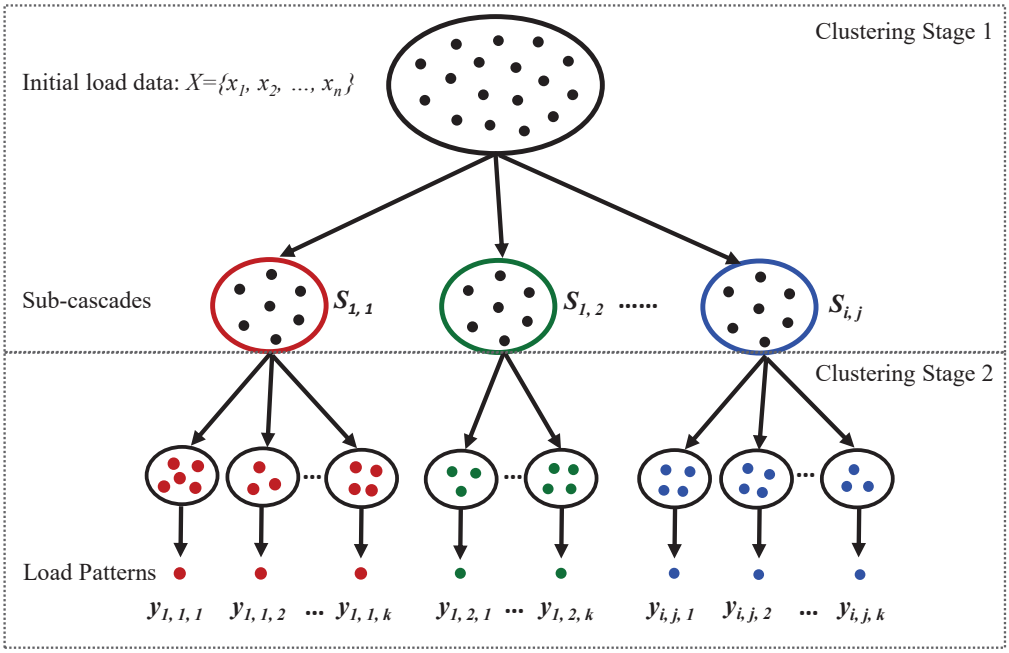


Figure 3.1: Cascade clustering scheme for load pattern categorisation and typical load pattern recognition.

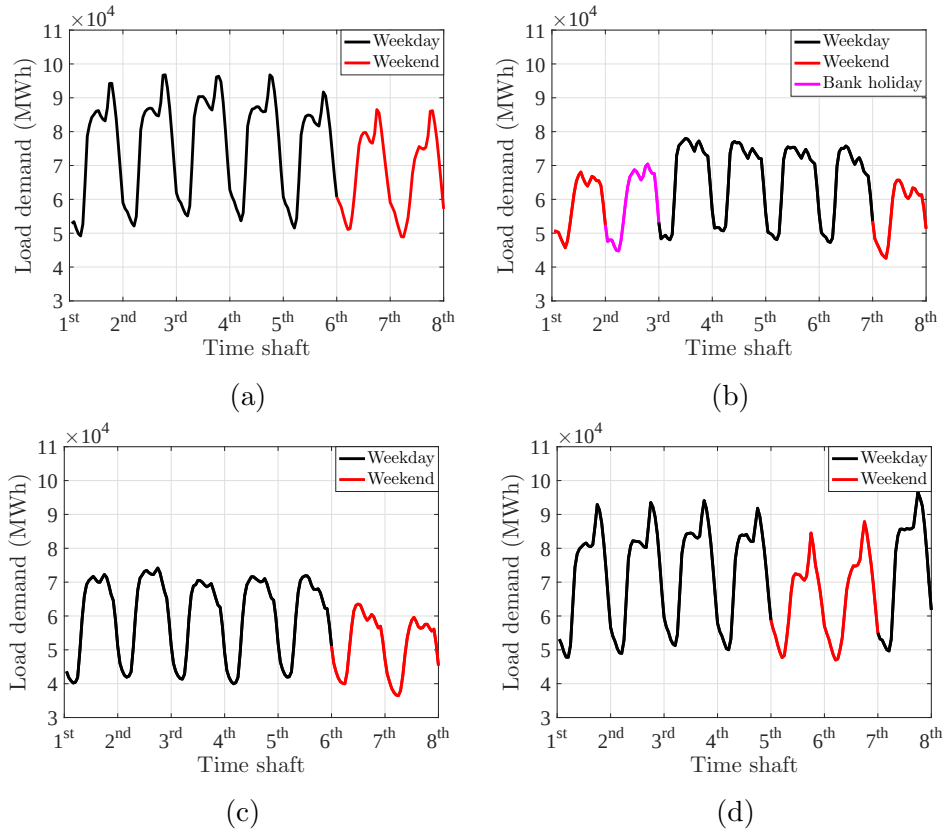


Figure 3.2: Load demand observations of different time in the year 2016. (a) Data of February; (b) Data of May; (c) Data of August; (d) Data of November.

(3) The load shapes of bank holidays are also dissimilar to both weekdays and weekends.

(4) The load demand apparently varies with seasonality, as shown in Figure 3.2(a) and Figure 3.2(c).

According to these factors, pre-clustering the initial load demand data into a series of sub-cascades at the first stage of the proposed cascade clustering scheme is significant, it is capable of reducing clustering errors and improving efficiency when compared with clustering the raw data directly. As a result, the initial load demand data is divided into $i \times j$ sub-cascades, where i and j are referred to as the day characteristic and the seasonality, respectively. In this work, the day types are considered to be "working day" consisting of weekdays excluding bank holidays, "non-working day" consisting of weekends excluding bank holidays and "bank holiday", *i.e.,* Good Friday, Easter Monday, Christmas Day, *etc.* Meanwhile, the seasonality is divided by month.

Following this, the second stage of the cascade scheme focuses on finding the internal relationships between objects within the same sub-cascade and assigns the objects into refined clusters, ensuring that the objects within the same cluster are similar. In terms of sub-cascade clustering, an appropriate cluster number k, has to be determined at first. Thus, a robust parametric bootstrap (PB) algorithm is proposed to resolve the cluster number determination problem and it is incorporated with a compatible clustering technique to cluster load data simultaneously. This process allows typical load patterns (TLPs) for each sub-cascade to be extracted as $y_{i,j,k}$, as shown in Figure 3.1. Note that in a practical implementation, sub-cascade clustering can be executed in parallel, making this a potentially very fast scheme.

3.2 Parametric Bootstrap Algorithm for Cluster Number Determination

The PB algorithm is a kind of re-sampling based algorithm. The basic principle of the PB algorithm is to use feature parameters that are obtained from the

adopted data to generate a number of synthetic data to testing the hypothesis. The hypothesis in this work is the hypothesis of the cluster number. It estimates the number of components by incrementally testing the hypothesis that there are $k+1$ components against the null hypothesis that there are k components via parametric bootstrap. An accepted k value is determined based on the significance level of the hypotheses. Figure 3.3 illustrates the flow chart of determining an appropriate cluster number based on the PB algorithm.

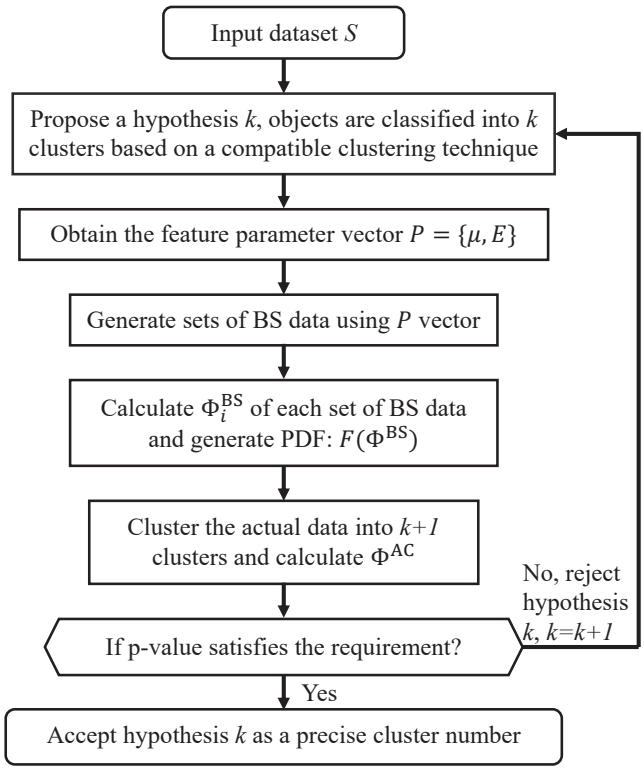


Figure 3.3: Flow chart of determining a precise cluster number using the parametric bootstrap (PB) algorithm.

Specifically, given an input dataset of a random sub-cascade $\mathcal{S}$ with N objects, it can be mathematically characterised as:

$$\mathcal{S} = \{x_n, n = 1, 2, ..., N, x_n \in \mathbb{R}^d\};$$

where d denotes the dimensionality which indicates the resolution of a load curve. To begin with, the input data is firstly hypothesised as consisting of k clusters and it is classified into k groups using a compatible clustering technique.

In addition, the feature parameter vector of a cluster, $P = \{\mu, E\}$, consisting of mean vector μ and a $d \times d$ covariance matrix E, can be obtained from Equations (3.1)-(3.2), respectively.

Mean vector μ:

$$\mu = \frac{\sum_{n=1}^{N_k} x_n}{N_k} \tag{3.1}$$

Covariance matrix E:

$$E = \begin{bmatrix} c(x^{(1)}, x^{(1)}) & c(x^{(1)}, x^{(2)}) & \cdots & c(x^{(1)}, x^{(d)}) \\ c(x^{(2)}, x^{(1)}) & c(x^{(2)}, x^{(2)}) & \cdots & c(x^{(2)}, x^{(d)}) \\ \vdots & \vdots & \ddots & \vdots \\ c(x^{(d)}, x^{(1)}) & c(x^{(d)}, x^{(2)}) & \cdots & c(x^{(d)}, x^{(d)}) \end{bmatrix} \tag{3.2}$$

where

$$c(x^{(a)}, x^{(b)}) = \frac{\sum_{n=1}^{N_k} (x_n^{(a)} - \mu^{(a)})(x_n^{(b)} - \mu^{(b)})}{N_k - 1} \tag{3.3}$$

In Equation (3.1)-(3.3), N_k is the number of total objects within the k^{th} cluster among K clusters and $a, b \in [1, d]$. Afterwards, sets of synthetic data with the same size as the real data can be generated periodically based on the obtained feature parameters. The process of generating synthetic data is the core of the algorithm and it is called "bootstrap simulation" (BS).

Moreover, the sum of square errors (SSE, denoted as Φ) which is the summation of the squared distance of each point within a cluster from the cluster center, is proposed to evaluate the clustering quality. The SSE of the i^{th} BS dataset is described by Equation (3.4). The probability density function (PDF) of Φ^{BS}, denoted as $F(\Phi^{\text{BS}})$, also can be calculated.

$$\Phi_i^{\text{BS}} = \sum_{k=1}^{K} \sum_{n=1}^{N_k} \sum_{d=1}^{D} (x_{k,n}^{\text{BS},(d)} - \mu_k^{\text{BS},(d)})^2 \tag{3.4}$$

Further, to assess the hypothesis that a dataset is composed of $k+1$ clusters against the null hypothesis that it has only k clusters, the actual data is clustered into $k+1$ clusters, the SSE of the actual dataset on the hypothesis k can be obtained from Equation (3.5).

$$\Phi^{\text{AC}} = \sum_{k=1}^{K+1} \sum_{n=1}^{N_k} \sum_{d=1}^{D} (x_{k,n}^{(d)} - \mu_k^{(d)})^2 \tag{3.5}$$

Although the real dataset actually consists of k clusters, the performance of $k+1$ clusters (or more) is normally "better" than simulations, since the data is grouped into smaller and tighter clusters. Thus, the objective is to find the rate at which the SSE decreases will slow down for k beyond the objective cluster number. Accordingly, the p value (proposed in Equation (3.6)) which is a widely used parameter in a statistical hypothesis [178] is adopted to determine an appropriate k.

$$p = \int_{-\infty}^{\Phi^{\mathrm{AC}}} F\left(\Phi^{\mathrm{BS}}\right) \cdot d\Phi^{\mathrm{BS}} \tag{3.6}$$

In the proposed PB algorithm, the requirement of an acceptable hypothesis k is defined as: if $p < \alpha$, the hypothesis of using k clusters is rejected and it tends to hypothesise that the data has at least $k+1$ clusters. The evaluation process continues with increasing values of k until it satisfies the condition $p \geqslant \alpha$, where α is an acceptable threshold, which is general is equal to 0.01 or 0.05 [179]. In order to execute the PB algorithm, a compatible clustering technique is required. In this work, K-means++ which is an improved version of the K-means algorithm is selected as the classifier due to its higher efficiency and improved robustness compared with others (*e.g.,* standard K-means, K-medoids, Gaussian mixture models, *etc.* [17–23]). The principles of the selected compatible clustering techniques are introduced specifically in the next section.

3.3 Compatible Clustering Algorithms for Load Data Classification

3.3.1 Standard K-Means Clustering Algorithm

Given a dataset $\mathcal{S} = \{x_n, n = 1, 2, ..., N, x_n \in \mathbb{R}^d\}$, the standard K-means algorithm divides $\mathcal{S}$ into k exhaustive clusters $\Omega = \{\Omega_k, k = 1, 2, ..., K\}$, $\bigcup_{i=1}^{k} \Omega_i = \mathcal{S}$, $\Omega_i \cap \Omega_j = \varnothing$ for $1 \leqslant i \neq j \leqslant K$. For a cluster, its centroid is given by:

$$\omega_i = \frac{1}{|\Omega_i|} \sum_{x \in \Omega_i} x \tag{3.7}$$

Let $\omega = \{\omega_1, ..., \omega_k\}$ be a set of centers and $\|x_i - x_j\|$ represent the Euclidean distance between x_i and x_j. The objective of K-means is to find an optimal ω to

minimise:

$$\arg\min_{\Omega} \sum_{k=1}^{K} \sum_{x \in \Omega_k} \|x - \omega_k\|_2 \tag{3.8}$$

The standard K-means algorithm tries to put object x into a cluster Ω_k to be similar to each other whilst being dissimilar to objects in other clusters. It takes the cluster number (hypothesised in Section 3.2) as the input parameter and k mean values are randomly selected as the initial centers of clusters. Afterwards, the remaining objects are assigned to the clusters with the closest centers according to the similarity. The algorithm continues to update the means of clusters until the means converge and become stable. Hence, the classical K-means algorithm proceeds by alternating two steps, the assignment step and the update step, and it can be illustrated concisely in Algorithm 1.

Algorithm 1 Standard K-means

Input: k, $\mathcal{S} = \{x_n, n = 1, 2, ..., N, x_n \in \mathbb{R}^d\}$.
Output: $\omega = \{\omega_1, \omega_2, ..., \omega_k\}$.
1: Arbitrarily initialise k centers $\omega = \{\omega_1^{(1)}, ..., \omega_k^{(1)}\}$.
2: Assign each object x_p to a cluster $\Omega^{(t)}$ according to:

$$\Omega_i^{(t)} = \{x_p : \|x_p - \omega_i^{(t)}\|_2 \leqslant \|x_p - \omega_j^{(t)}\|_2 \ \forall j, 1 \leqslant j \leqslant k\}.$$

3: Update centers of clusters according to:

$$\omega_i^{(t+1)} = \frac{1}{|\Omega_i^{(t)}|} \sum_{x_j \in \Omega_i^{(t)}} x_j.$$

4: Repeat step 2 and 3 until ω becomes stable.

3.3.2 K-Means++ Clustering Algorithm

A good clustering result satisfies the condition that the distance between two arbitrary clusters should be as large as possible [180]. Intuitively, it is a wise choice to choose initial centers that are far away from each other in the beginning. The K-means++ algorithm follows this idea, but the farthest point is not always chosen to be a center. Actually, except the first center, which is chosen uniformly and randomly from the data points, each subsequent center is chosen from the remaining data points with the probability proportional to its squared distance from the existing cluster center that is closest to the point. Let $D(x)$ be the

Euclidean distance between x and the nearest center that has already been chosen. The K-means++ algorithm is presented as follows.

Algorithm 2 K-means++

Input: k, $\mathcal{S} = \{x_n, n = 1, 2, ..., N, x_n \in \mathbb{R}^d\}$.
Output: $\omega = \{\omega_1, \omega_2, ..., \omega_k\}$.
1: $\omega \leftarrow \varnothing$.
2: Choose one center x from $\mathcal{S}$ at random, $\omega = \omega \cup x$.
3: Choose $x \in \mathcal{S}$ with probability: $\frac{D(x)^2}{\sum_{x \in \mathcal{S}} D(x)^2}$, $\omega = \omega \cup x$.
4: Repeat Step 3 until k centers are chosen.
5: Proceed as with the standard K-means algorithm.

Compared with the standard K-means algorithm, the K-means++ algorithm guarantees to find a solution that is $O(logk)$ competitive to the optimal K-means solution, which means that K-means++ has better effectiveness in clustering than the standard K-means algorithm. Although the special seeding in K-means++ takes extra time, its K-means part converges fast so that the clustering efficiency is significantly improved.

3.3.3 K-Medoids Clustering Algorithm

The classical K-means algorithm is sensitive to outliers since the least squared Euclidean distance is used. Unlike K-means which uses mean value as a centroid of a cluster, the K-medoids algorithm utilises a representative object within a cluster as the centroid, thus being less sensitive to outliers and extreme values [181]. This way also minimises the sum of distances for arbitrary distance functions. Similar to K-means, given a set of observations, $\mathcal{S}$, the objective of the K-medoids algorithm is to find:

$$\arg\min_{\Omega} \sum_{i=1}^{k} \sum_{x \in \Omega_i} \|x - m_i\|_2 \tag{3.9}$$

where the centroid $m_i \in \Omega_i$. The most common realisation of K-medoids clustering is the partitioning around medoids (PAM) algorithm. PAM uses a greedy search in finding the optimal solution, which is faster than the exhaustive search. It is executed by arbitrarily selecting k objects within the dataset as the initial medoids or seeds. The remaining objects are then assigned to the clusters with

the closest centers based on the absolute distance with the medoids. The medoids are continuously swapped with the other objects in an order in the corresponding cluster until the swapped medoid has the minimum sum of distances of objects to their medoid in a cluster. The algorithm finally stops when the medoids converge. The K-medoids based PAM algorithm can be summarised by Algorithm 3.

Algorithm 3 K-medoids based partitioning around medoids (PAM)

Input: k, $\mathcal{S} = \{x_n, n = 1, 2, ..., N, x_n \in \mathbb{R}^d\}$.
Output: $m = \{m_1, m_2, ..., m_k\}$.

1: Arbitrarily select an initial set of k objects as medoids, $m = \{m_1^{(1)}, ..., m_k^{(1)}\}$.
2: Assign each object x_p to a cluster $\Omega^{(t)}$ according to:

$$\Omega_i^{(t)} = \{x_p : \|x_p - m_i^{(t)}\| \leqslant \|x_p - m_j^{(t)}\| \ \forall j, 1 \leqslant j \leqslant k\}.$$

3: Update medoids of clusters.
4: **for** $i = 1, ..., k, q = 1, ..., n$
5: Swap medoid and objects in order:

$$m_i^{(t+1)} = x_q.$$

6: Recompute the sum of distances of objects to their medoid:

$$C_i = \sum_{x \in \Omega_i} \|x - m_i\|_2.$$

7: **if** C_i increased in step 6
8: **undo** step 5.
9: **end for**
10: Repeat step 2 - step 9 until m becomes stable.

3.3.4 Gaussian Mixture Models for Clustering

Different from K-means and K-medoids, Gaussian mixture models (GMM) are a probability based unsupervised learning classifier and often used to classify a wide variety of signals. According to the central limit theorem [70, 72, 182], a given set of data is normally hypothesised as a mixture Gaussian distributions, since a mixture Gaussian distributions is in theory able to approximate to any probability distribution by increasing the number of models. Thus, it has broad applications in data clustering. In addition, the parameters of GMM are normally estimated by an expectation maximisation (EM) algorithm which is an effective machine learning method for probability density estimation [73–75]. However, an apparent drawback of GMM is that the convergence speed of using the EM algorithm to achieve the parameters of GMM is quite slow, particularly to large

volumes of data with multi-dimensions, due to the seeds which are arbitrarily selected. In order to resolve this issue, a heuristic method that is using the classical K-means algorithm to initialise the data first and then defining weights based on K-means memberships as the initial setting is applied in the initialisation of GMM. This action is able to reduce the computation due to the fast convergence of K-means [183–185] and significantly decreases the probability of the local minimum appearing.

To begin with, the Gaussian mixture models are a model composed of multivariate Gaussian distributions. Following this, a multivariate Gaussian distribution can be given by Equation (3.10).

$$p(x) = \sum_{i=1}^{k} \phi_i \mathcal{N}(x|\mu_i, \Sigma_i) \tag{3.10}$$

where μ_i is the mean vector and Σ_i is the covariance matrix of a cluster Ω_i. The mixture coefficient ϕ_i indicates the weight of each distribution in the model, and $\sum_{i=1}^{k} \phi_i = 1$. However, for a d-dimensions Gaussian distribution, it can be expanded as:

$$\mathcal{N}(x|\mu, \Sigma) = \frac{1}{(2\pi)^{d/2}} \frac{1}{|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right\} \tag{3.11}$$

Defining a parameter set of the mixture Gaussian distribution, $\theta = \{\phi_i, \mu_i, \Sigma_i, i = 1, ..., k\}$, for a series of objects within $\mathcal{S}$, the objective function can be formulated as maximising the log-likelihood function of $\mathcal{S}$ as shown in Equation (3.12).

$$\max_{\theta} \ \ln p(\mathcal{S}|\theta) = \sum_{j=1}^{n} \ln \sum_{i=1}^{k} \phi_i \mathcal{N}(x_j|\mu_i, \Sigma_i) \tag{3.12}$$

where $p(\mathcal{S}|\theta) = \prod_{j=1}^{n} p(x_j|\theta)$. With this object, the EM algorithm has to be used in resolving this density estimation problem. The process of the EM algorithm estimating the parameters of mixture Gaussian distributions can be summarised by Algorithm 4.

The EM algorithm is an iterative algorithm that includes two main steps: expectation-step (E-step) and maximisation-step (M-step). Applied in this estimation problem, it randomly hypothesises which Gaussian distribution a

sample x_i belongs to in the E-step. In the M-step, the parameters of the model are updated based on the previous hypotheses. Since in the M-step, it is pretending that the hypotheses in the E-step were correct, the maximisation becomes possible.

Algorithm 4 Expectation maximisation (EM) algorithm for parameters estimation of mixture Gaussian distrubutions

Input: k, $\mathcal{S} = \{x_n, n = 1, 2, ..., N, x_n \in \mathbb{R}^d\}$.
Output: θ.

1: Select an initial setting for the parameter $\bar{\theta}$ by using the classical K-means clustering algorithm.
2: Expectation-step (E-step):
3: **for** $j = 1, ..., n,\ i = 1, ..., k$

$$\gamma_{j,i} = p(k|x_j, \bar{\theta}) := \frac{\bar{\phi}_i \mathcal{N}(x\mathrm{j}|\bar{\mu}_i, \bar{\Sigma}_i)}{\sum_{i=1}^{k} \bar{\phi}_i \mathcal{N}(x_j|\bar{\mu}_i, \bar{\Sigma}_i)}.$$

4: **end for**
5: Maximisation-step (M-step):
6: **for** $i = 1, ..., k$

$$\phi_i = \frac{n_i}{n};$$
$$\mu_i = \frac{1}{n_i} \sum_{j=1}^{n} \gamma\mathrm{j,i} x_j;$$
$$\Sigma_i = \frac{1}{n_i} \sum_{j=1}^{n} \gamma_{j,i}(x_j - \mu_i)(x_j - \mu_i)^T;$$
$$\text{where } n_i = \sum_{j=1}^{n} \gamma_{j,i}.$$

7: **end for**
8: $\bar{\theta} \leftarrow \theta$.
9: Repeat E-step and M-step until $\bar{\theta}$ converges.

3.3.5 Comparison of Compatible Classifiers

K-means and K-medoids are both partition-based and widely used algorithms, and they are related to each other. The objectives of both algorithms are to minimise the distance between points labeled to be in a cluster and the corresponding centroid. The main difference between K-means and K-medoids is that the K-means algorithm takes the mean value of all labeled data within a cluster as the center, but K-medoids chooses data points as centers of clusters. Thus, the K-medoids algorithm is more robust to noise and outliers as compared to K-means, since K-medoids minimises a sum of pairwise dissimilarities instead

of a sum of squared Euclidean distance. However, the time complexity of K-means is less than that of K-medoids [21], since the K-means algorithm only computes the mean values in centroid updating. The time complexities of K-means and K-medoids are $O_{\mathrm{KMS}} = n \times k \times t$ and $O_{\mathrm{KMD}} = n^2 \times k \times t$, respectively, where t is the number of iterations.

On the other hand, compared with the EM based Gaussian mixture models, the K-means and K-medoids algorithms are classified into "hard" cluster methods, while the GMM algorithm is a "soft" clustering method. The soft clustering methods assign a score to a data point for each cluster. The value of the score indicates the association strength of the data point to the cluster. As opposed to the hard clustering methods, soft clustering methods are flexible in that they can assign a data point to more than one cluster. When clustering with the EM based GMM algorithm, the term "soft" refers to the hypotheses being probability and taking values in $[0, 1]$. In contrast, a "hard" hypothesis is one that represents a single best guess, such as taking values in $\{0, 1\}$. Moreover, EM based GMM clustering can accommodate clusters that have different sizes and correlation structures within them. However, the drawbacks of the EM based GMM algorithm are also obvious. According to the evaluation results in Subsection 3.5.5, it is found that the efficiency of the EM based GMM algorithm is highly dependent on the dimensional conditions and the actual distribution of the data. The efficiency of the EM based GMM algorithm drops dramatically when high dimensional data is considered (*e.g.,* $d = 48$). Additionally, it does not perform well when the data does not follow a Gaussian distribution.

In fact, for a given set of data, there are several clustering techniques that can be used. Each technique is able to complete the task of identifying groups, however, all such compatible clustering techniques suffer from the inherent limitations of obtaining a local optimal caused by the arbitrarily selected seeds and the determination of an appropriate cluster number. To reduce the risk of ending up in a local optimum, the algorithms can be executed multiple times from which the best solution is selected. To obtain an appropriate cluster number, the proposed PB algorithm might be a great solution.

The proposed PB algorithm and the compatible clustering techniques are verified in Section 3.4 in terms of the clustering accuracy and the precise cluster number determination.

3.4 Algorithms Verification

It is impossible to evaluate the feasibility of algorithms on the actual load data since the data is not labeled by groups. Therefore, this section proposes unsupervised examples to assess the effectiveness of the proposed PB algorithm. The test dataset is generated based on the data features of actual load data in different months. The actual load data is utilised as the centroid of a certain cluster, and $N^{\text{cluster}} = 10^2$ objects in 24-dimensional space are generated for each cluster based on the centroid using a Gaussian distribution. In this case, the test dataset actually consists of 4 certain clusters and the hypotheses of $k = 2$ to 5 are presented. For each hypothesis, $N^{\text{BS}} = 5 \times 10^5$ bootstrap simulations are generated. The amount of BS data is usually set according to the requirements that an accurate PDF of Φ^{BS} can be obtained. Additionally, the significance thresholds $\alpha_1 = 0.01$ and $\alpha_2 = 0.05$ are both applied in the evaluation. The proposed PB algorithm is shown to be valid and effective when the evaluation result corresponds to the initial setup, *i.e.,* the obtained cluster number is equal to 4.

Figure 3.4 presents the cluster number determination result of the proposed test data based on the PB algorithm incorporated with the K-means++ technique, where red curves denote the SSE of the actual data (Φ^{AC}) in $k + 1$ clustering and black curves represent the PDF of the total square errors of the bootstrap simulation data (Φ^{BS}). Specifically, the result shows that the p-values of the cases $k = 2$ and $k = 3$ in Figure 3.4(a) and Figure 3.4(b), respectively, are equal to 0, which indicates that the hypotheses are rejected and the dataset includes at least $k + 1$ clusters. In addition, when $k = 4$ in Figure 3.4(c), the p-value increases to 0.194, which is greater than the pre-defined significance threshold α. Therefore, the hypothesis $k = 4$ is accepted as an appropriate cluster

number according to the cluster number determination requirement in Section 3.2. Moreover, it can be seen that the p-value (0.251) of hypothesis $k = 5$ in Figure 3.4(d) is greater than the p-value (0.194) of hypothesis $k = 4$ as well as greater than α. The result conforms to the expectation that the rate of Φ decreases with an increasing value of k. Therefore, the hypothesis of $k = 4$ is regarded as the appropriate cluster number, which also corresponds to the initial setup in this case.

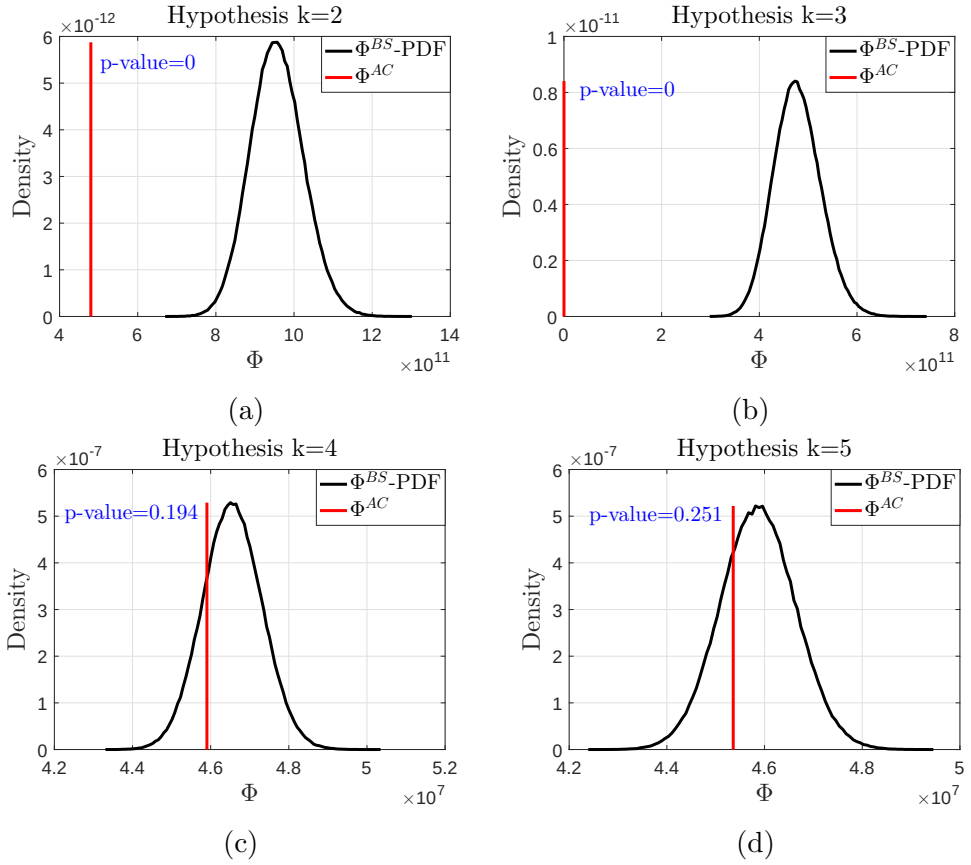


Figure 3.4: Cluster number determination for a 24-dimensional space dataset based on the parametric bootstrap algorithm incorporated with K-means++. Hypotheses of $k = 2$ to 5 are evaluated. (a) Hypothesis $k = 2$; (b) Hypothesis $k = 3$; (c) Hypothesis $k = 4$; (d) Hypothesis $k = 5$.

Compared with the most popular methods, G-means [15] and the AIC based algorithm [16], which also can be used in the cluster number determination, the PB algorithm is more robust and reliable, particularly in processing the high dimensional space data. Specifically, the G-means algorithm is not effective for high dimensional data, such as the load data, since it cannot ensure the

data within all dimensions simultaneously passes the Anderson-Darling (AD) statistical test. One solution is to reduce the data of N-dimensions to a single dimension [15]. However, the dimensional-reduction always gives rise to the risk of information loss and often leads to a failure to obtain an appropriate cluster number.

The AIC based algorithm determines the precise cluster number by seeking the first knee always at the local maximum of the curve of AIC. However, the AIC based algorithm also cannot guarantee to find an appropriate cluster number as the estimated inflection point normally varies in a range in AIC calculations.

In order to evaluate the reliability of an algorithm, a performance metric of failure rate (FR) is defined as:

$$FR = \frac{N_{\text{failure}}}{N_{\text{test}}} \tag{3.13}$$

where N_{failure} and N_{test} represent the numbers of failed tests and total tests, respectively.

Table 3.1: Performance comparison of cluster number determination between algorithms

Criterion \ Algorithm	PB	G-means	AIC based algorithm
Probability	0.98	0.42	0.32
STD	0.14	-	0.99
FR	0	0.26	0

The comparison results of algorithms in terms of the probability of finding an actual cluster number, the standard deviation (STD) and FR, over $N_{\text{test}} = 10^2$ tests are presented in Table 3.1.

The results show that the proposed PB algorithm is more effective in cluster number determination than the G-means and AIC based algorithms [15, 16], with a much higher probability (0.98) of successfully finding an actual cluster number. In addition, the PB algorithm is more robust against dimensionality of the data than G-means ($\text{FR}_{\text{G-means}} = 0.26$). However, the dimensional conditions have little affect ($\text{FR}_{\text{PB}} = 0$) on the outcome of cluster number determination

using the PB algorithm. Moreover, the PB algorithm is also more stable than the AIC based algorithm with a lower STD value ($\text{STD}_{\text{PB}} = 0.14$) (as a number of tests by G-means failed to obtain a cluster number, the G-means algorithm is not considered in the STD evaluation). In summary, although G-means and AIC based algorithms can achieve the appropriate cluster numbers in some cases, both algorithms cannot guarantee to find a precise cluster number at all times due to their inherent defects.

The verified PB algorithm incorporated with K-means++ is specifically applied to the actual load data in next section. The initial load dataset is classified into a number of sub-cascades at first according to the proposed scheme described in Section 3.1. Afterwards, the cluster numbers of selected sub-cascades are determined and typical load patterns (TLPs) are derived.

3.5 Case Study

3.5.1 Case Descriptions

In this study, a large set of historical load demand data on the national level provided by National Grid Ltd, UK [186] is adopted. The dataset includes 3653 days' load demand data with time intervals 2 hrs ($d = 12$) and 0.5 hr ($d = 48$) from 2007 to 2016.

As objects in the initial dataset are classified into a series of sub-cascades, mathematical numbers are simply utilised to label the seasonal features and the day characteristics of each sub-cascade. Specifically, the numbers "1 - 12" are used to label the 12 months (for consideration of the different seasons) and "1 - 3" are used to indicate the day types ("1"→ "working day", "2" → "none-working day" and "3" → "bank holiday"). For instance, $\mathcal{S}_{5,1}$ denotes the sub-cascade of a working day in May. The rest of the possible sub-cascades can be deduced by analogy. Additionally, due to limited data samples on special events days such as UK bank holidays that is available (8 UK bank holidays per year and 82 UK bank holidays in total), the exclusive sub-cascade $\mathcal{S}_{\forall,3}$ consists of all load information of UK bank holidays. Other types of special events data can be analysed in a

similar way.

Due to the limited space, 9 typical sub-cascades (*i.e.,* $\mathcal{S}_{2,1}$, $\mathcal{S}_{5,1}$, $\mathcal{S}_{8,1}$, $\mathcal{S}_{11,1}$, $\mathcal{S}_{2,2}$, $\mathcal{S}_{5,2}$, $\mathcal{S}_{8,2}$, $\mathcal{S}_{11,2}$ and $\mathcal{S}_{\forall,3}$) covering various scenarios are selected as examples to perform the analysis.

3.5.2 Evaluation Criteria

In order to evaluate the similarity of TLPs between a variety of sub-cascades, the Pearson correlation coefficient (PCC, denoted as ρ) which is a measure of the linear correlation between two variables X and Y in statistics, is proposed in Equation (3.14).

$$\rho(X,Y) = \frac{\sum_{d=1}^{D}(x_d - \overline{x}) \cdot (y_d - \overline{y})}{\sqrt{\sum_{d=1}^{D}(x_d - \overline{x})^2} \cdot \sqrt{\sum_{d=1}^{D}(y_d - \overline{y})^2}} \tag{3.14}$$

where D is the sample size of the compared time series TLPs. Variables x_d, y_d are the individual sample points indexed with d. $\overline{x} = \sum_{d=1}^{D} x_d/D$ and analogously for $\overline{y}$. PCC has a value between +1 and -1, where +1 indicates the total positive linear correlation, 0 represents no linear correlation, and -1 denotes the total negative linear correlation between X and Y.

In terms of assessing the clustering performance between a variety of algorithms, a number of evaluation metrics [15, 44] are introduced. The proposed metrics which are object distance based, mainly evaluate the compactness quality between objects within the same cluster and the separation quality between disparate clusters.

Given a sub-cascade $\mathcal{S}$, it is assumed that the objects in $\mathcal{S}$ are classified into K clusters. For an individual cluster Ω_i, $x_{i,n}$ and ω_i represent an object within Ω_i and the center of Ω_i, respectively. Hence, the evaluation metrics including compactness (CP), separation (SP) and Davies-Bouldin index (DBI) can be presented in Equations (3.15)-(3.16), (3.17) and (3.18), respectively, as follows.

Compactness (CP) metric:

$$CP = \frac{1}{K}\sum_{i=1}^{K}\overline{CP}_i \tag{3.15}$$

$$\overline{CP}_i = \frac{1}{|\Omega_i|}\sum_{x_{i,n}\epsilon\Omega_i}\|x_{i,n} - \omega_i\| \tag{3.16}$$

where $\overline{\mathrm{CP}}_i$ is the averaged taxicab distance between objects $x_{i,n}$ and the cluster center ω_i within the cluster Ω_i. CP shows the compactness or homogeneity between objects within clusters. The smaller the value of CP the more compact objects are within a cluster.

Separation (SP) metric:

$$SP = \frac{2}{K^2 - K}\sum_{i=1}^{K}\sum_{j=i+1}^{K}\|\omega_i - \omega_j\|_2 \tag{3.17}$$

SP describes the separation or the distance between clusters. The larger the value of SP the greater the distance between the centroids of the clusters.

Davies-Bouldin index (DBI) metric:

$$DBI = \frac{1}{K}\sum_{i=1}^{K}\max_{j\neq i}\left(\frac{\overline{CP}_i + \overline{CP}_j}{\|\omega_i - \omega_j\|_2}\right) \tag{3.18}$$

DBI represents the system-wide average of the similarity measures of each cluster with its most similar cluster. The smaller the value of DBI, the better the performance.

In addition, the sum of square errors (Φ) which is the most significant metric, is also considered in the evaluation of accurate clustering. As the processing time of algorithms executed on various platforms is different, the relative running time (RRT) which is defined in Equation (3.19), is taken to evaluate the efficiency between algorithms.

$$RRT = \frac{CT_i}{CT_{\min}} \tag{3.19}$$

where CT_i denotes the running time of the i^{th} algorithm and $\mathrm{CT}_{\min}$ is the minimal running time among all the algorithms of comparison.

3.5.3 Cluster Number Determination for Load Demand Data

In this section, the cluster number determination program for each sub-cascade has been run multiple times, to ensure a reliable result. The significance levels α_1 = 0.01 and α_2 = 0.05 are both taken into account. Based on the verified PB algorithm incorporated with K-means++, the cluster number determination results for the load data of 9 selected sub-cascades are presented in Table 3.2.

Table 3.2: Results of the cluster number determination with p-values for the typical sub-cascades using the PB algorithm incorporated with K-means++.

Dimension	Sub-cascade / SL	$\mathcal{S}_{2,1}$	$\mathcal{S}_{5,1}$	$\mathcal{S}_{8,1}$	$\mathcal{S}_{11,1}$	$\mathcal{S}_{2,2}$	$\mathcal{S}_{5,2}$	$\mathcal{S}_{8,2}$	$\mathcal{S}_{11,2}$	$\mathcal{S}_{\forall,3}$
d=12	α_1 = 0.01	k=4 p=0.017	k=4 p=0.040	k=4 p=0.039	k=4 p=0.036	k=3 p=0.057	k=3 p=0.038	k=3 p=0.027	k=3 p=0.042	k=3 p=0.110
	α_1 = 0.05	k=5 p=0.064	k=6 p=0.073	k=5 p=0.066	k=5 p=0.053	k=3 p=0.057	k=4 p=0.051	k=4 p=0.060	k=4 p=0.064	k=3 p=0.110
d=48	α_1 = 0.01	k=4 p=0.017	k=4 p=0.030	k=4 p=0.012	k=4 p=0.019	k=3 p=0.037	k=3 p=0.037	k=3 p=0.018	k=3 p=0.049	k=3 p=0.133
	α_1 = 0.05	k=5 p=0.068	k=5 p=0.086	k=5 p=0.064	k=5 p=0.120	k=4 p=0.115	k=4 p=0.070	k=4 p=0.105	k=4 p=0.090	k=3 p=0.133

Specifically, the results illustrate that the cluster numbers of the "working day" in sub-cascades $\mathcal{S}_{2,1}$, $\mathcal{S}_{5,1}$, $\mathcal{S}_{8,1}$ and $\mathcal{S}_{11,1}$ are normally equal to 4 for α_1 = 0.01, and 5 for α_2 = 0.05. However, the cluster numbers of the "none-working day" in sub-cascades $\mathcal{S}_{2,2}$, $\mathcal{S}_{5,2}$, $\mathcal{S}_{8,2}$ and $\mathcal{S}_{11,2}$ are usually equal to 3 for α_1 = 0.01, and 4 for α_2 = 0.05. In addition to these, the exclusive sub-cascade $\mathcal{S}_{\forall,3}$ is suggested as being classified into 3 clusters. The obtained cluster numbers are used as the input parameters for data classification and the extraction of TLPs in LPC.

3.5.4 Load Pattern Categorisation

In order to reduce the risk of ending up in a local optimum, the clustering process has been executed multiple times with random initialisations. The cluster numbers adopted for the categorisation process are those obtained for a significance level of α_1 = 0.01. The center of a cluster which leads to a minimal SSE of the cluster is regarded as one of the TLPs within a sub-cascade. While the final TLPs are regarded as the aggregation of TLPs within each individual

sub-cascade under the cascade clustering scheme.

The obtained TLPs of nine typical sub-cascades are proposed as examples in Figure 3.5. Specifically, the light grey curves represent the actual load demand data that is required to be clustered and the colorised curves are the objective TLPs. Based on the results, it can be seen that the extracted TLPs follow the shape of actual load curves and cover the most actual load curves within the same sub-cascade.

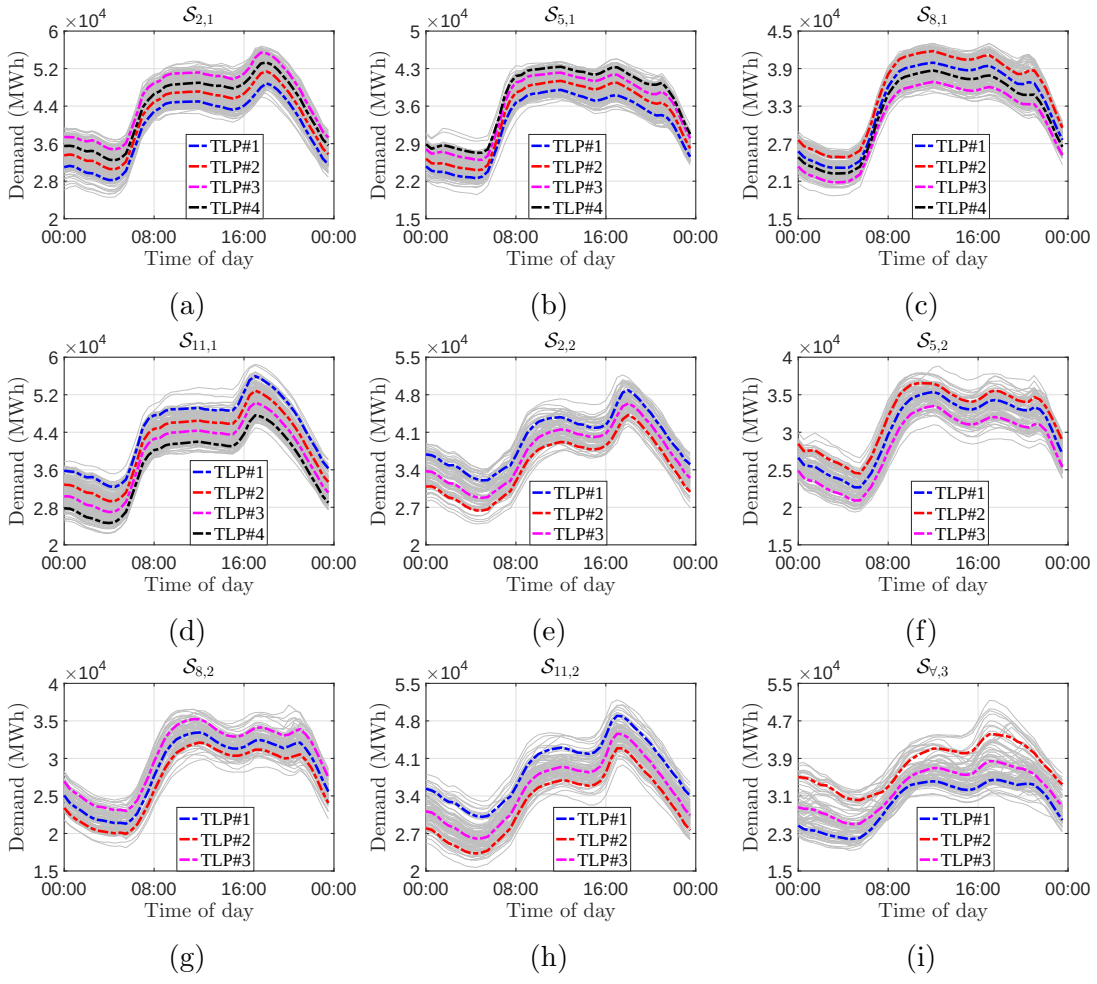


Figure 3.5: Extracted typical load patterns (TLPs) of selected sub-cascades. The objective TLPs are colorised. (a) Sub-cascade $\mathcal{S}_{2,1}$; (b) Sub-cascade $\mathcal{S}_{5,1}$; (c) Sub-cascade $\mathcal{S}_{8,1}$; (d) Sub-cascade $\mathcal{S}_{11,1}$; (e) Sub-cascade $\mathcal{S}_{2,2}$; (f) Sub-cascade $\mathcal{S}_{5,2}$; (g) Sub-cascade $\mathcal{S}_{8,2}$; (h) Sub-cascade $\mathcal{S}_{11,2}$; (i) Sub-cascade $\mathcal{S}_{\forall,3}$.

In addition, Table 3.3 illustrates the evaluation of similarity between TLPs of three sub-cascades based on PCC. It can be seen that the achieved coefficients of the compared TLPs are very high ($\geqslant 0.837$) in general, which means that

there are positive relationships among TLPs. Moreover, the coefficients between TLPs within the same sub-cascades are extremely high ($\geqslant 0.97$). On the contrary, the coefficients between TLPs from different sub-cascades are relatively small for most cases. The data indicates that higher correlations between TLPs within the same sub-cascade and relatively weaker relationships between TLPs from different sub-cascades. It is also in line with the analysis in Section 3.1.

Table 3.3: Similarity evaluation between TLPs of 3 typical sub-cascades based on the metric of PCC.

TLPs		$\mathcal{S}_{2,1}$				$\mathcal{S}_{8,2}$			$\mathcal{S}_{\forall,3}$		
		TLP#1	TLP#2	TLP#3	TLP#4	TLP#1	TLP#2	TLP#3	TLP#1	TLP#2	TLP#3
$\mathcal{S}_{2,1}$	TLP#1	**1.000**	**0.999**	**0.998**	**0.998**	0.911	0.908	0.908	0.934	0.837	0.910
	TLP#2	**0.999**	**1.000**	**0.999**	**0.999**	0.910	0.908	0.907	0.934	0.848	0.915
	TLP#3	**0.998**	**0.999**	**1.000**	**0.999**	0.906	0.905	0.902	0.931	0.848	0.914
	TLP#4	**0.998**	**0.999**	**0.999**	**1.000**	0.915	0.914	0.911	0.939	0.858	0.923
$\mathcal{S}_{8,2}$	TLP#1	0.911	0.910	0.906	0.914	**1.000**	**0.998**	**0.999**	**0.994**	0.906	**0.973**
	TLP#2	0.908	0.908	0.905	0.914	**0.998**	**1.000**	**0.997**	**0.994**	0.919	**0.980**
	TLP#3	0.908	0.906	0.902	0.911	**0.999**	**0.997**	**1.000**	**0.993**	0.899	0.968
$\mathcal{S}_{\forall,3}$	TLP#1	0.934	0.934	0.931	0.939	**0.994**	**0.994**	**0.993**	**1.000**	0.923	**0.983**
	TLP#2	0.837	0.847	0.848	0.858	0.906	0.919	0.899	0.923	**1.000**	**0.976**
	TLP#3	0.910	0.915	0.913	0.922	**0.972**	**0.980**	0.968	**0.983**	**0.976**	**1.000**

3.5.5 Categorising Performance Comparison

Categorising performance is another issue of interest in the research community. The performance comparison of different clustering techniques incorporated with the PB algorithm is presented in Table 3.4. The best performance metrics are marked in bold in the table. Obviously, the K-means++ algorithm outperforms in most of the metrics as an overall comparison. Specifically, K-means++ performs the best in Φ and CP evaluations, which means the objects clustered by K-means++ algorithm are more compact and have less square errors. In addition, in terms of SP and DBI evaluations, K-means++ also performs the best in some cases, such as $\mathcal{S}_{5,1}$, $\mathcal{S}_{8,1}$, $\mathcal{S}_{11,1}$ and $\mathcal{S}_{2,2}$. On the contrary, the GMM algorithm has the worst clustering performance in the evaluation as it is not effective in clustering the data with a high dimensional space, such as $d = 48$.

Moreover, the adopted K-means++ algorithm is also the most efficient

clustering algorithm as it spends the minimal amount of running time among all compared algorithms. In accordance with the RRT results in Table 3.4, K-means++ is 2.11, 3.93 and 381 times faster than K-means, K-medoids, and GMM, respectively on average.

Table 3.4: Performance comparison of different clustering techniques incorporated with the PB algorithm (best performance metrics are marked in bold).

Metric	Algorithms	Selected Sub-Cascade Index																		
		d = 12									d = 48									
		$\mathcal{S}_{2,1}$	$\mathcal{S}_{5,1}$	$\mathcal{S}_{8,1}$	$\mathcal{S}_{11,1}$	$\mathcal{S}_{2,2}$	$\mathcal{S}_{5,2}$	$\mathcal{S}_{8,2}$	$\mathcal{S}_{11,2}$	$\mathcal{S}_{\forall,3}$	$\mathcal{S}_{2,1}$	$\mathcal{S}_{5,1}$	$\mathcal{S}_{8,1}$	$\mathcal{S}_{11,1}$	$\mathcal{S}_{2,2}$	$\mathcal{S}_{5,2}$	$\mathcal{S}_{8,2}$	$\mathcal{S}_{11,2}$	$\mathcal{S}_{\forall,3}$
$\Phi^{\times 10^{10}}$	K-means++	**4.16**	**2.45**	**1.85**	**3.81**	**2.73**	**1.65**	**1.25**	**2.95**	**4.69**	**1.10**	**0.648**	**0.496**	**1.00**	**0.712**	**0.434**	**0.331**	**0.770**	**0.123**
	K-means	4.16	2.45	1.85	3.81	2.73	1.65	1.25	2.95	4.69	1.10	0.648	0.496	1.00	0.712	0.434	0.331	0.770	0.123
	K-medoids	4.61	2.76	1.96	4.12	3.29	1.93	1.42	3.31	5.58	1.26	0.739	0.536	1.09	0.863	0.513	0.383	0.879	0.149
	GMM	4.69	2.91	2.11	5.95	2.76	1.70	-	-	5.34	-	-	-	-	-	-	-	-	-
$\mathrm{CP}^{\times 10^4}$	K-means++	**3.91**	**2.93**	**2.64**	**3.72**	**5.10**	**3.68**	**3.15**	**5.13**	**6.61**	**4.01**	**3.00**	**2.70**	**3.80**	**5.19**	**3.73**	**3.24**	**5.23**	**6.75**
	K-means	3.91	2.93	2.64	3.72	5.10	3.68	3.15	5.13	6.61	4.02	3.00	2.70	3.80	5.19	3.74	3.24	5.23	6.75
	K-medoids	4.10	3.04	2.69	3.78	5.42	3.80	3.26	5.39	6.86	4.24	3.14	2.76	3.88	5.52	3.89	3.35	5.50	7.04
	GMM	4.15	3.22	2.78	4.39	5.15	3.75	-	-	7.92	-	-	-	-	-	-	-	-	-
$\mathrm{SP}^{\times 10^4}$	K-means++	4.88	**4.48**	**3.73**	**6.87**	**4.99**	3.38	3.18	6.62	7.45	2.44	**2.25**	**1.87**	**3.45**	**2.50**	1.70	1.59	3.31	3.73
	K-means	4.88	4.48	3.73	6.87	4.99	3.37	3.18	6.62	7.45	2.44	2.25	1.87	2.95	2.50	1.69	1.60	3.31	3.73
	K-medoids	**4.95**	3.60	3.64	5.96	4.97	**3.67**	**3.35**	**6.63**	**8.44**	**2.49**	1.82	1.81	2.98	2.42	**1.85**	**1.69**	**3.33**	**4.24**
	GMM	4.74	3.69	3.50	5.41	4.77	3.28	-	-	7.12	-	-	-	-	-	-	-	-	-
DBI	K-means++	2.80	**2.61**	2.62	**2.22**	**2.83**	3.10	2.69	2.54	2.46	5.75	**5.29**	**5.41**	**4.55**	**5.78**	6.31	5.51	5.16	5.06
	K-means	2.80	2.61	2.62	2.22	2.83	3.10	2.69	**2.51**	2.46	5.75	5.29	5.41	4.55	5.78	6.31	5.51	**5.12**	5.06
	K-medoids	**2.79**	3.11	**2.41**	2.40	2.97	**2.94**	**2.59**	2.55	**2.38**	**5.70**	5.93	5.44	5.01	6.04	**6.06**	**5.28**	5.18	**4.87**
	GMM	3.20	3.59	2.95	3.10	3.05	3.42	-	-	3.16	-	-	-	-	-	-	-	-	-
RRT	K-means++	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
	K-means	2.29	2.01	2.13	2.16	2.79	1.74	2.50	2.99	3.03	1.48	1.34	1.45	1.31	1.95	2.25	2.10	2.27	2.23
	K-medoids	4.86	4.83	4.41	4.60	4.56	4.66	4.42	5.32	5.44	2.70	2.91	2.72	2.48	3.07	3.45	3.40	3.50	3.49
	GMM	583	581	446	703	120	113	-	-	122	-	-	-	-	-	-	-	-	-

Much more interesting, it is noticed that the clustering performance of the K-means algorithm is similar to the K-means++ algorithm, except the RRT evaluation. This is because the K-means++ algorithm is an improved version of the standard K-means algorithm. The difference is that the special seeding in K-means++ significantly accelerates the convergence process of K-means. Besides, the multiple simulations which are taken in this work significantly avoid the risk of ending in a local optimum for the K-means++ algorithm.

3.6 Summary

An innovative parametric bootstrap algorithm incorporated with K-means++ which is a comparatively robust clustering technique has been proposed in this chapter, to address the cluster number determination problem as well as cluster the load data simultaneously in LPC analysis, under a cascade clustering scheme. A number of evaluations have been presented in this work, which indicate that the proposed approach is more robust and reliable in finding an appropriate cluster number than conventional methods [15, 16] and it also has a better performance in load pattern clustering in comparison to results published in the literature [17–23]. In fact, for a given dataset, the data clustering performance is largely influenced by the choice of the clustering algorithm, the input cluster number and the characteristics of the data. Several clustering algorithms can be applied to a certain dataset, however, there is no single best method for all datasets. Therefore, a comparatively robust clustering technique that is able to complete the load data classification task efficiently has been used, instead of developing a new method in this work. On the other hand, this work focused on proposing a more effective algorithm incorporated with the adopted clustering technique to address the cluster number determination problem.

The research on LPC is important to DR management. The obtained TLPs in this work are used to support the assessment of the impact of a DR program. The detailed knowledge of electricity consumption's nature is essential to promoting strategies for peak load reduction, for instance exploiting the customers' willingness to accept price-based demand conditioning on the basis of DR programs. In addition, LPC has been proposed as effective means for enhancing targeting and tailoring DR programs as well as providing reasonable load scheduling recommendations, owing to the availability of advanced technology for load shifting and to emerging opportunities for flexible demand management, providing incentives and rewards to participating users. Therefore, the study of LPC is essential in advance of designing DR programs.

After gaining a comprehensive understanding of the load consumption

patterns on the demand side, a hybrid real-time electricity price forecasting model considering deterministic and stochastic characteristics of input data is proposed to forecast the short-term electricity prices which are usually used as an input control signal for DR management in the next chapter.

Chapter 4

A Hybrid Model for Real-time Electricity Price Forecasting of Smart Grid

The research in this chapter mainly concentrates on building an effective, high accuracy, estimation model for electricity price forecasting in the smart grid using limited sets of historical data. The obtained real-time price (RTP) can be used as an input control signal in DR management. The main contributions of this work can be summarised as follows.

(1) A hybrid RTP forecasting model which is a consolidation of the least-square (LS) fitting model, the grey prediction (GP) model and an artificial neural network (ANN), is proposed. The LS fitting model tracks the deterministic characteristics of the time series data where as the GP model tracks the stochastic characteristics. The ANN is used for error optimisation; execution of the ANN is dependent on the forecasting performance of the first two stages. To the authors best knowledge, this is the first time work combining the above three processes together for short-term RTP forecasting has been presented and tested.

(2) The accuracy of time series RTP forecasting is shown to improve through application of this hybrid model. The effectiveness of the model is verified and the results indicate that the proposed method is an accurate and efficient tool in the forecasting of day-ahead RTP and it also significantly outperforms the previous methods [25–29].

The rest of this chapter is organised as follows. In Section 4.1, the structure of the hybrid forecasting model is proposed and the three individual model components are analysed in detail in Section 4.2. A case study is then performed in Section 4.3, to evaluate the feasibility of the proposed approach, and the forecasting performance is tested against a variety of methods from the literature. Finally, this work is summarised in Section 4.4.

4.1 Structure of the Forecasting Model

The structure of the proposed hybrid forecasting model is presented in this section. Figure 4.1 shows the historical RTP data over 5 historical day samples as provided by the Australian Energy Market Operator (AEMO) [187]. The time series dynamic electricity prices vary dependent on load demand at different time periods. Based on the variations of historical RTP samples in Figure 4.1, electricity prices exhibit a prominent regularity and consist of deterministic and stochastic information along with the varying prices. It is therefore sensible that the characteristics of the deterministic and stochastic properties of time series data have to be incorporated into the forecasting model. The proposed forecasting model is formulated as:

$$P_t = L_t + N_t + E_t^* \tag{4.1}$$

where P_t is the forecasting RTP at time t. L_t and N_t represent the estimations of deterministic and stochastic characteristics of the input data, respectively. Additionally, E_t^* denotes the error optimisation procedure and the execution of E_t^* is dependent on the forecasting performance of the first two stages.

In order to present the principle of the proposed hybrid forecasting model, Figure 4.2 illustrates a flow chart of forecasting day-ahead real-time electricity prices based on 5 days' historical RTP data. Specifically, the historical data is input as the basis to establish the model. Then, the deterministic characteristic of the data is estimated using the LS fitting model. Following this, the GP model is applied to the estimation of the stochastic characteristic within the data. Then it will be determined whether or not the ANN model based error optimisation

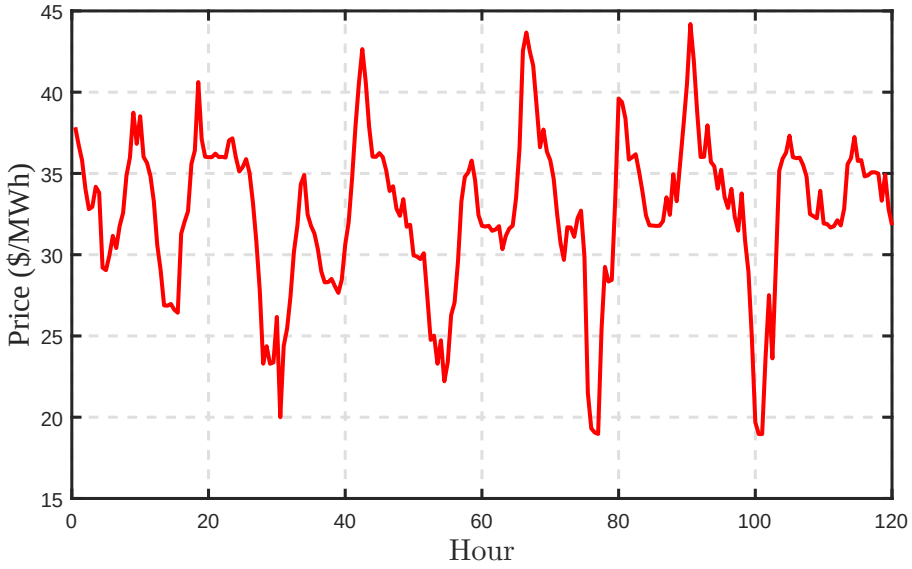


Figure 4.1: Historical RTP over 5 historical day samples (120 hours).

procedure is necessary for this stage in accordance with the spot error rate (ER) of the initial forecasting result. The ANN model will be executed to improve the specific forecasting accuracy if the spot ER exceeds the maximum tolerable ER. Finally, the objective day-ahead RTP can be obtained from the integrated model.

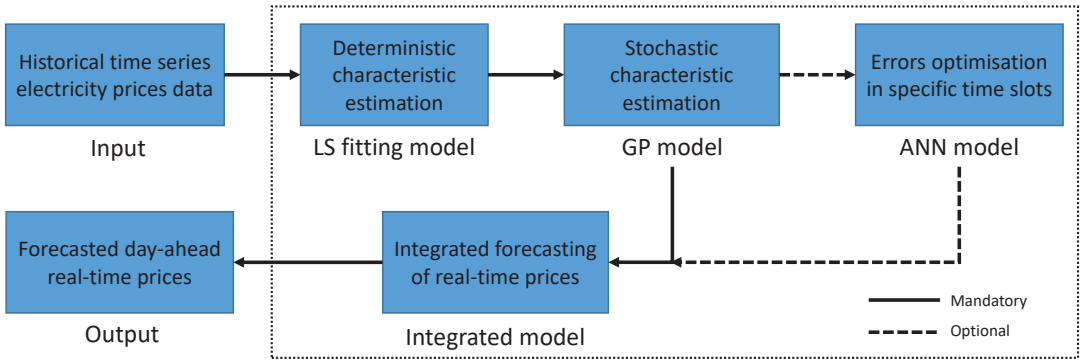


Figure 4.2: Flow chart of forecasting day-ahead real-time electricity prices.

In the following chapter, the specific descriptions of the relevant forecasting components in the hybrid model are presented in detail.

4.2 Hybrid Forecasting Model

4.2.1 Least Square Fitting Model for Deterministic Characteristic Forecasting

Based on Equation (4.1), the LS fitting model is employed to obtain the deterministic characteristic L_t within the input data. The least square fitting for data is a standard approach in regression analysis to the approximate solution of over determined systems. It is one of the fitting algorithms [27, 188–190]. On the stage of deterministic characteristic forecasting, the LS fitting model can be used to build a fitting function to express the mainstream variation among the historical data. Assume the input dataset H consists of n days' historical RTP data, H can be formulated as:

$$H = \{D_1, D_2, ..., D_n\} \tag{4.2}$$

The historical RTP of a day can be treated as a number of discrete values with an interval. In this study, the time interval is set as 0.5 hour, which means $t = 48$ fixed values are included in a day. Hence, D_n is represented as:

$$D_n = [y_{n,1}, y_{n,2}, ..., y_{n,t}] \tag{4.3}$$

In addition, the fitting function $L(t)$ is taken to model the mainstream variation in the deterministic characteristic estimation. However, the general formats of the fitting function include Fourier, Gaussian, polynomial, sum of sine, *etc.* and they can be formulated in Equation (4.4)- (4.7), respectively, as follows.

$$\text{Gaussian format:} \quad f(x) = \sum_{i=1}^{d} a_i \cdot e^{\left(-\left(\frac{x-b_i}{c_i}\right)^2\right)} \tag{4.4}$$

$$\text{Fourier format:} \quad f(x) = \sum_{i=0}^{d} a_i \cdot \cos(i \cdot \omega \cdot x) + b_i \cdot \sin(i \cdot \omega \cdot x) \tag{4.5}$$

$$\text{Polynomial format:} \quad f(x) = \sum_{i=0}^{d} p_i \cdot x^i \tag{4.6}$$

$$\text{Sum of sine format:} \quad f(x) = \sum_{i=1}^{d} a_i \cdot \sin(b_i \cdot x + c_i) \tag{4.7}$$

where $d \in N^+$ is the degree of the adopted function. Additionally, a_i, b_i, c_i, ω_i and p_i are undetermined constant parameters in the model. Although all the proposed fitting function formats are effective in modeling the deterministic characteristic within the data, the Fourier format is adopted in this study due to its better fitting performance. Therefore, the objective function on this stage can be formulated as determining a group of appropriate parameters (a_i, b_i and ω_i) to minimise the total square errors J. The objective function is presented in Equation (4.8).

$$\arg\min_{a,b,\omega} \; J = \sum_{n=1}^{N} \sum_{t=1}^{48} (L_t - y_{n,t})^2 \tag{4.8}$$

On the one side, a higher value of the fitting degree d leads to a better performance of the estimation when J is in a reasonable range. On the other hand, it results in more complexity of the calculation and more CPU wastage. Therefore, selecting an appropriate fitting degree in the fitting model is significant and may lead to a better deterministic characteristic estimation performance. In order to obtain an appropriate fitting degree, a significance level for d, (α_d), is defined in Equation (4.9).

$$\alpha_d = \frac{|J_d - J_{d+1}|}{J_d} \tag{4.9}$$

where J_d denotes the square errors when the fitting degree is d. If $\alpha_{d-1} > 0.05$ and $\alpha_d \leqslant 0.05$, the value of d can be regarded as an appropriate fitting degree.

Table 4.1 presents the total square errors J with different values of fitting degree which range from 1 to 7. It can be seen that when $d \in [1, 3]$, J decreases quickly with increasing fitting degree. For example, there is only $|J_{d=5} - J_{d=6}| = |2.995 - 2.981| = 0.015$ difference between cases $d = 5$ and $d = 6$. When $d = 4$, $J_{d=4} \leqslant 0.05$ and $J_{d=3} > 0.05$. Thus, $d = 4$ is taken as a proper fitting degree in this study. Afterwards, the relevant parameters can be determined as in Table 4.2.

Table 4.1: Total square errors with different values of fitting degree, $d = 1$ to 7 are evaluated.

d	1	2	3	4	5	6	7
$J^{\times 10^2}$	5.601	4.152	3.223	2.988	2.995	2.981	2.950

Table 4.2: Parameter values in L_t. Fourier format is selected and $d = 4$.

Parameters	a_0	a_1	a_2	a_3	a_4	b_1	b_2	b_3	b_4	ω
Values	32.640	0.570	2.153	0.022	-0.831	-2.950	-4.158	1.738	1.185	0.289

4.2.2 Grey Prediction Model for Stochastic Characteristic Forecasting

The second stage in the proposed hybrid RTP forecasting model is to estimate N_t which denotes the stochastic characteristic within the input data. Obviously, the stochastic information within the data is included in the forecasting errors after using the LS fitting model. Thus, the stochastic characteristic within the historical data at time t can be expressed by Equation (4.10) and the initial records within M_t are used to estimate the next record $(D_{n+1,t} - L_t)$ by using the GP model.

$$M_t = \{D_{1,t} - L_t, D_{2,t} - L_t, ..., D_{n,t} - L_t\} \tag{4.10}$$

The GP model or GM(1,1) was first proposed to deal with the data in grey systems. It is able to analyse a system that includes insufficient information and unapparent relationships [28, 191, 192]. Hence, the GP model is often used in predicting data in a stochastic systems based on limited information. It transforms the forms of the irregular discrete sequences and displays the potential regularities within the sequences. Transforming the forms of the sequences can make the properties of stochastic and randomness get weaker thereby turning irregular sequences to regular ones [193–195]. Since only a few stochastic data proceeded from LS fitting model, it is quite appropriate to employ the GP model to estimate the stochastic characteristic within the input data on this stage.

The GP model is established by using generalised series data. The primitive sequence dataset is defined as $X^{(0)}$ and it can be presented as:

$$X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\} \tag{4.11}$$

where $x^{(0)}(n) = D_{n,t} - L_t$ and $x^{(0)}(n) \geqslant 0$. However, if there are any $x^{(0)}(n) < 0$ in the primitive sequence, all the candidates in the sequence have to be improved until $\forall x^{(0)}(n) \geqslant 0$. Afterwards, the first accumulated generating dataset $X^{(1)}$ can

be obtained from Equation (4.12).

$$X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), ..., x^{(1)}(n)\} \tag{4.12}$$

where

$$x^{(1)}(n) = \sum_{i=1}^{n} x^{(0)}(i) \tag{4.13}$$

In addition, $Z^{(1)}$ which is determined by $X^{(1)}$, is defined as the background factor and $Z^{(1)}$ can be represented as:

$$Z^{(1)} = \{z^{(1)}(1), z^{(1)}(2), ..., z^{(1)}(n)\} \tag{4.14}$$

where

$$z^{(1)}(n) = \frac{1}{2}[x^{(1)}(n) + x^{(1)}(n-1)] \tag{4.15}$$

For example, when $t = 26$, the initial sequence (after data preprocessing) can be represented as $X^{(0)} = \{6.5, 7.9, 11.1, 11.4, 13.5\}$ as shown in Figure 4.3(a). Then, the first accumulated generating data $X^{(1)}$ can be calculated as $X^{(1)} = \{6.5, 14.4, 25.5, 36.9, 50.4\}$ as shown in Figure 4.3(b). Obviously, there are not any prominent regularities between the numbers in $X^{(0)}$. However, after the first accumulated generating operation (AGO) towards $X^{(0)}$, the new sequence $X^{(1)}$ is provided with the quasi-exponential property (otherwise, the second AGO will be executed towards $X^{(1)}$). Therefore, $X^{(1)}$ can be regarded as being satisfied with the first order ordinary differential equation which is shown in Equation (4.16).

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = u \tag{4.16}$$

where a is treated as a development coefficient that describes the increasing speed of numbers in $X^{(0)}$ and u is an endogenous control coefficient in the system. The parameters $U = [a, u]^T$ can be determined by Equation (4.17).

$$U = (B^T B)^{-1} B^T Y \tag{4.17}$$

where Y is a $(n-1) \times 1$ matrix and B is a $(n-1) \times 2$ matrix. Y and B can be written as:

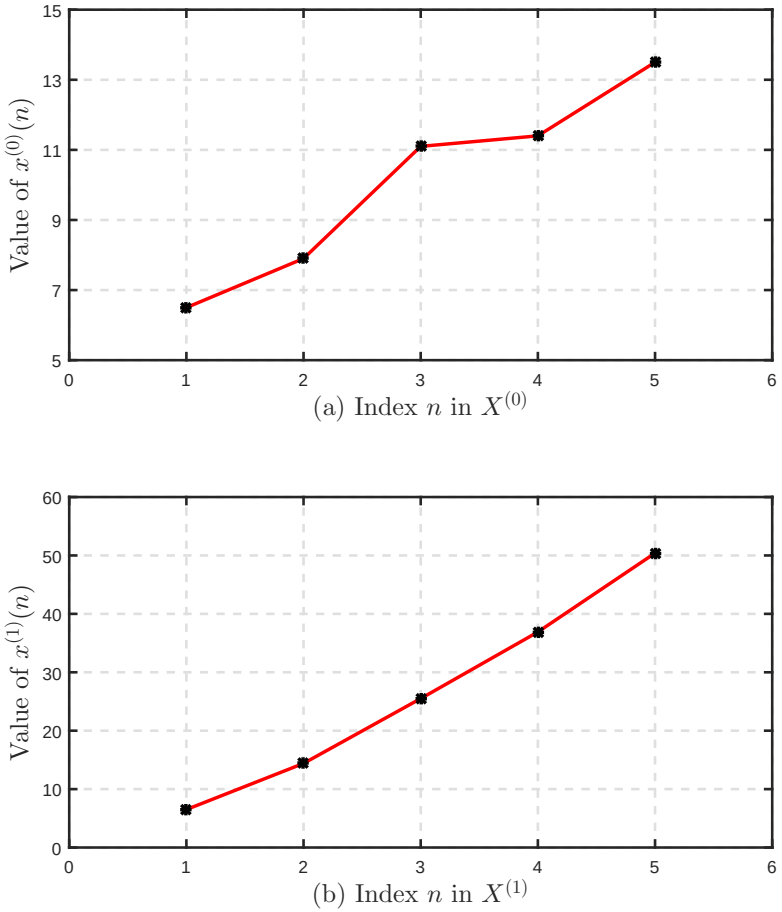


Figure 4.3: Examples of $X^{(0)}$ and $X^{(1)}$. Case $t = 26$ is adopted as an example. (a) Example data in $X^{(0)}$; (b) Example data in $X^{(1)}$.

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} \text{ and } B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}.$$

According to these, Equation (4.16) can be resolved using the obtained parameters a and u, so that the forecasting formula of $\hat{X}^{(1)}$ can be denoted as shown in Equation (4.18).

$$\hat{x}^{(1)}(n+1) = \left(x^{(1)}(1) - \frac{u}{a}\right) e^{-ak} + \frac{u}{a} \tag{4.18}$$

Based on Equation (4.18), when $n = 1, 2, ..., N-1$, $\hat{x}^{(1)}(n+1)$ is a fixed value. When $n \geqslant N$, $\hat{x}^{(1)}(n+1)$ is a predicted value of $X^{(1)}$. Afterwards, the predicted formula of $\hat{X}^{(0)}$ can be achieved through the inverse accumulated generating

operation (IAGO) as shown in Equation (4.19).

$$\begin{aligned}\hat{x}^{(0)}(n+1) =& \hat{x}^{(1)}(n+1) - \hat{x}^{(1)}(n) \\ =& (1-e^{a})\left[x^{(0)}(1) - \frac{u}{a}\right]e^{-ak}\end{aligned} \tag{4.19}$$

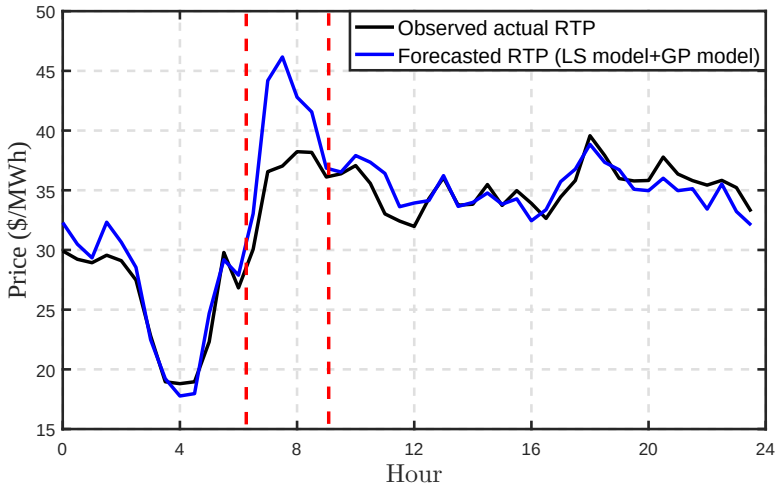


Figure 4.4: Real-time electricity prices forecasting result based on LS model+GP model.

According to Equation (4.19), when $n = N$, $\hat{x}^{(0)}(n+1)$ is the objective forecasting value. Following this, the GP model can be utilised periodically to achieve forecasting values of all time spots. Furthermore, the forecasting result is the combination of the deterministic characteristic and the stochastic characteristic within the input data, which is described by Figure 4.4. In general, the variation of the obtained RTP by using the combined LS+GP model is in line with actual RTP overall. The result indicates that the ERs are lower than 10% most of the time, particularly between 12:30 - 22:00 (lower than 5%). However, the ERs are unexpectedly higher than 10% during the time period 6:30 - 8:30. These unexpected errors may be caused by defects in the current forecasting models. Since limited datasets were used in the hybrid model, the random error increases when there are great differences among the input datasets.

Therefore, in order to improve the forecasting accuracy during specific time slots based on the initial forecasting result, the ANN model based error optimisation procedure is required. If the ERs for a time period $[t_a, t_b]$ are higher than 10%, the ANN based error optimisation procedure will be executed.

4.2.3 Artificial Neural Network Model for Error Optimisation

The artificial neural network (ANN) is a non-linear modeling where any prior knowledge of the relationship between input and output is needed [196]. It gives great results for forecasting problems [29]. To establish the model, only a sufficient amount of data is required to assimilate the connection between inputs and outputs. The main parameters of the ANN model are the number of the input vectors, the number of layers and the number of neurons in each layer [197–199]. However, large and sudden spikes in the input data will lead to less accuracy in the output using ANN [67, 68]. In this study, the back propagation (BP) algorithm is utilised to train the ANN model.

Figure 4.5 shows the topography of a typical 3-layer back propagation neural network [200, 201]. A 3-layer back propagation neural network is a typical multiple-layer network and it includes the input layer (LA), hidden layer (LB), and output layer (LC). There are no connections between nodes that belong to the same layer. LA has m nodes that correspond to the m inputs of the network. LC consists of n nodes that correspond to the n outputs of the network. The node number of LB can be varied to fit the task.

Define W_{ir} as the connection weight between the node a_i of the LA layer and the node b_r of the LB layer. Similarly, let V_{rj} be the connection weight between the node b_r of the LB layer and the node c_j of the LC layer. Set T_r and θ_j as the bias of the node b_r of the LB layer and the bias of the node c_j of the LC layer, respectively. Then the output function of the LB layer node should be:

$$b_r = f\left(\sum_{i=1}^{m} W_{ir} \cdot a_i - T_r\right) \ (r = 1, 2, ..., u) \tag{4.20}$$

The output function of the LC layer node should be:

$$c_j = f\left(\sum_{r=1}^{u} V_{rj} \cdot b_r - \theta_j\right) \ (j = 1, 2, ..., n) \tag{4.21}$$

where $f(x)$ is a sigmoid function and it can be expressed as:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{4.22}$$

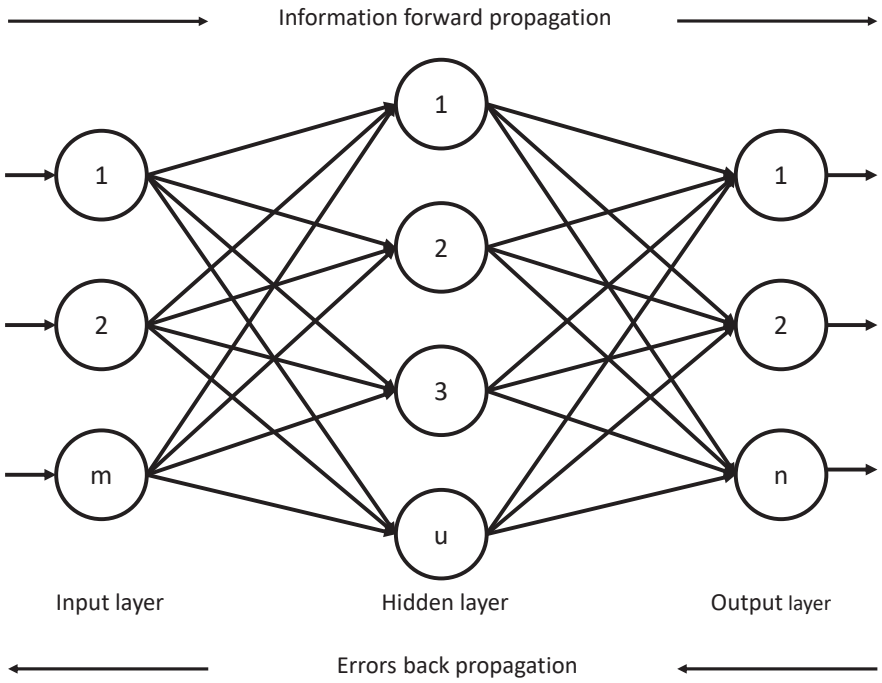


Figure 4.5: Topography of a typical 3-layer back propagation neural network.

In addition, the BP learning algorithm which is a typical error-revised learning algorithm is used to learn and store knowledge in 3-layer back propagation neural networks. The learning procedures can be illustrated as follows.

(1) Initialise the variables W_{ir}, T_r, V_{rj} and θ_j with small random values.

(2) For each model pair $(A^{(k)}, C^{(k)})$ $(k = 1, 2, ..., p)$, take the following steps.

- Input the values of $A^{(k)}$ at layer LA, then calculate b_r and c_j by Equation (4.20) and (4.21).
- Calculate the bias d_j of the desired value and calculate the value c_j of the layer LC nodes and let

$$d_j = c_j \cdot (1 - c_j) \cdot (c_j^{(k)} - c_j) \tag{4.23}$$

- Back propagate the errors to the layer LB nodes and let

$$e_r = b_r \cdot (1 - b_r) \cdot \left(\sum_{j=1}^{n} V_{rj} \cdot d_j \right) \tag{4.24}$$

- Adjust the connection weights V_{rj} and the bias of the layer LC nodes θ_j:

$$V_{rj} = V_{rj} + \alpha \cdot d_j + \beta \cdot \Delta V_{rj}^{'} \tag{4.25}$$

$$\theta_j = \theta_j + \alpha \cdot d_j + \beta \cdot \Delta \theta_j^{'} \tag{4.26}$$

  where $\Delta V_{rj}^{'}$ and $\theta_j^{'}$ are the adjusting values of the previous learning loop. α is the learning ratio and $0 < \alpha < 1$. β is the momentum factor.

- Adjust the connection weights W_{ir} and the bias of the layer LB nodes T_r:

$$W_{ir} = W_{ir} + \alpha \cdot e_r + \beta \cdot \Delta W_{ir}^{'} \tag{4.27}$$

$$T_r = T_r + \alpha \cdot e_r + \beta \cdot \Delta T_r^{'} \tag{4.28}$$

  where $\Delta W_{ir}^{'}$ and $\Delta T_r^{'}$ are the adjusting values of the previous learning loop.

(3) Repeat step (2), until d_j becomes adequately small.

In accordance with the analysis in previous chapters, the ANN model is used on this stage to improve the accuracy of the RTP forecasting further in particular time slots, such as between 6:30 - 8:30 as shown in Figure 4.4. In this case, 2 hidden layers with 20 and 40 neurons are designed and 10-day historical data is adopted. In Section 4.3, a number of simulations are carried out to prove the effectiveness of the proposed hybrid model and the forecasting quality is also evaluated in terms of a number of evaluation criteria.

In Section 4.3, a case study is proposed to evaluate the hybrid forecasting model and the forecasting results are also discussed.

4.3 Case Study

4.3.1 Short-Term RTP Forecasting Results

The real-time electricity price forecasting results using the proposed hybrid forecasting model are presented in this section. Limited datasets (5 days) of the historical RTP data with a time interval of 0.5 hour from Australia are adopted.

The obtained results are also compared with other methods (*e.g.,* ARIMA model [25, 26], independent back propagation-artificial neural network (BP-ANN) [196], *etc.*) in this work.

A number of evaluation criteria including mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE) and mean absolute percentage error (MAPE) [202–204] are proposed for the evaluation of the forecasting performance. To begin with, x is defined as the forecasting value of a model, ref as the observed true values and $t = 48$ which means 48 points are required to predict a days pricing. Hence, the mathematical formulations of MAE, MSE, RMSE, and MAPE can be expressed by Equation (4.29)-(4.32), respectively.

$$MAE = \frac{1}{t} \cdot \sum_{i=1}^{t} |ref_i - x_i| \tag{4.29}$$

$$MSE = \frac{1}{t} \cdot \sum_{i=1}^{t} (ref_i - x_i)^2 \tag{4.30}$$

$$RMSE = \sqrt{\frac{1}{t} \cdot \sum_{i=1}^{t} (ref_i - x_i)^2} \tag{4.31}$$

$$MAPE = \frac{1}{t} \cdot \sum_{i=1}^{t} \left| \frac{ref_i - x_i}{ref_i} \right| \tag{4.32}$$

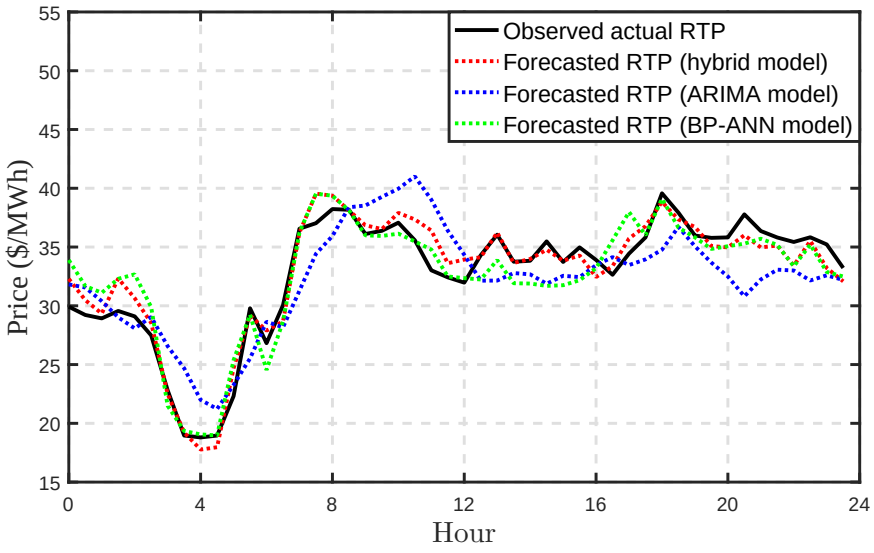


Figure 4.6: RTP forecasting results comparison between models.

Figure 4.6 shows the RTP forecasting results comparison between the proposed hybrid model and some typical models. According to the obtained results, it can be seen that all three compared models, *i.e.,* the hybrid model, the ARIMA model and, the BP-ANN model are able to accomplish the task of forecasting the RTP in advance. The forecasting price variations are in line with the observed actual RTP in general. The RTP forecasting by the hybrid model is slightly better than the other two models on the basis of the results. However, comparing the RTP variations in Figure 4.4 and Figure 4.6, it is apparent that the forecasting errors are significantly reduced during the time period 6:30 - 8:30, due to the contribution of the BP-ANN model in the error optimisation procedure.

Table 4.3: RTP Forecasting quality evaluation comparison between models (best performance is marked in bold).

Models	Evaluation Criteria			
	MAE ($/MWh)	MSE ($/MWh)	RMSE ($/MWh)$^{\frac{1}{2}}$	MAPE (%)
ARIMA Model	2.61	9.25	3.04	8.29
LS Model	2.52	9.41	3.07	8.51
GP Model	1.35	3.01	1.74	4.29
LS+GP Models	1.53	5.33	2.31	4.65
BP-ANN Model	1.49	3.50	1.87	4.69
Hybrid Model	**1.06**	**1.72**	**1.31**	**3.38**

The RTP forecasting evaluation comparison between the hybrid model and other state-of-the-art models [27, 28, 188, 191, 196] are also presented in Table 4.3. Based on the obtained results, the proposed hybrid model performs best in forecasting quality evaluation in an overview, which confirms the advantages of the proposed approach. Specifically, the MAE, MSE, RMSE, and MAPE of the hybrid model are 1.06, 1.72, 1.31 and 3.38%, respectively, which are the lowest among all models. The ARIMA model performs the worst in the MAE evaluation and the independent LS model performs worst in MSE, RMSE and MAPE evaluations.

4.3.2 Discussions

The hybrid model analyses the input data in views of deterministic characteristics, stochastic characteristics, and error optimisation within the data. The advantage of the proposed approach is that the hybrid model is more robust in dealing with forecasting tasks based on insufficient data compared with the traditional models such as ARIMA [25] which needs a large number amount of historical data for training. The RTP forecasting quality evaluation results in Table 4.3 also indicate that the ARIMA model is not effective in cases where there is limited input data. In addition, the individual LS model did not perform well in this case, as the LS model only extracts the mainstream within the input data. Therefore, using an LS model independently to forecast the RTP will lead to considerable errors as expected. It is much more interesting to see that the LS model in cooperation with the GP model performs a bit worse than the independent GP model in the overall evaluation. This is because the combined model (LS model + GP model) performs not well in a specific time period, *i.e.,* 6:30 to 8:30 in this case, so that the errors increased significantly, although it has higher forecasting accuracies in other time periods compared with the GP model.

Given a group of historical data, actually, there are several forecasting models that can be used and each able to complete the task of forecasting but with varying degrees of accuracies. After a great number of tests, it is realised that the forecasting performance is crucially dependent on both selecting an appropriate model and the data correlations. A forecasting model works well to one group of data, but maybe not effective for another group of data. Therefore, a hybrid forecasting model is generally more efficient than an independent model.

4.4 Summary

In this chapter, a hybrid model consisting of the LS model, GP model, and BP-ANN model, has been proposed to forecast the day-ahead real-time prices based on limited historical data. The evaluation results have clearly demonstrated that the proposed approach is an accurate and efficient tool in RTP forecasting

and it also significantly outperforms the previous forecasting models. As the RTP tariff is a trend for the smart grid, the theories in this chapter have bright prospects not only in RTP forecasting but also in applications in other industrial fields, such load forecasting, wind forecasting, GDP forecasting, *etc.*

In DR management, a valid RTP is usually utilised as an input control signal for DR to enable efficient load shifting. In the next chapter, a number of DR strategies are proposed based on the RTP tariff.

Chapter 5

Electric Vehicles Assisted Demand Response in Smart Grid

Demand response (DR) is an essential characteristic of the smart grid and it plays an important role in energy efficiency improvement and wastage reduction, by providing encouraging energy-aware consumption. In this chapter, A number of DR strategies considering the impacts of using electric vehicles (EVs) as flexible storage and supply mechanisms are proposed. Compared with the previous research, the main contributions of this work are:

(1) A number of DR strategies for domestic appliance scheduling are designed and implemented to different scales of household (*i.e.,* a single household and a multi-household network) in order to simultaneously alleviate the load burden for the grid and reduce bills for householders. For the single household network, the EV is utilised as an auxiliary power supply (APS) for energy consumption of home appliances. An EV-APS model based DR strategy is proposed in accordance with a fundamental DR strategy. For the multi-household network, an EVs assisted DR framework including a neighbour energy sharing (NES) model for a residential network with different types of EVs installed at consumers' premises is developed.

(2) The available EVs' energy distribution is enabled via vehicle-to-home (V2H) and vehicle-to-neighbour (V2N) connections in this work. The NES based DR framework is valid and effective not only for an independent household but also for a multi-household residential network, which can satisfy broader

requirements compared with conventional DR programs in literature. The energy trading policy in the neighbourhood is also declared.

(3) Comprehensive affecting factors (*e.g.,* EVs' behaviors, user preferences, load scheduling priorities, *etc.*) are considered in scheduling for both EV assisted DR strategies. The effectiveness of the proposed DR strategies is verified by numerical results, which demonstrate that the proposed approaches achieve great performances in terms of load balancing and electricity cost reduction.

The rest of this chapter is organised as follows. Firstly, a fundamental DR strategy is proposed to manage home appliances intelligently in Section 5.1. Secondly, a dynamic price (DP) based and EV assisted DR strategy for appliance scheduling for a single household is developed in Section 5.2. Third, an EVs assisted DR framework for a multi-household network is proposed in Section 5.3. Two case studies are carried out to evaluate the feasibility and effectiveness of the proposed DR strategies in Section 5.4. Finally, this work is summarised in Section 5.5.

5.1 Fundamental Demand Response for a Single Household

5.1.1 Appliances Classification

In general, more than 15 types of household appliances will be used in domestic homes every day. Considering the operating characteristic of each appliance, it is not necessary to schedule all of them via DR programs. Hence, in accordance with the device operating characteristics, household appliances can be classified into different scenarios. As a result, the appliances are sorted into two main scenarios in this study.

(1) Critical scenario (CS): CS contains the appliances that have to be used at a specific time or cannot be scheduled. Examples include lighting, TV, laptop, *etc.*

(2) Flexible scenario (FS): FS contains the appliances that can be powered on with a tolerable delay and have a flexible operating time. Hot water tank and washer are typical representatives in FS.

According to the sorting scheme above, 16 frequently used appliances are listed and classified as follows:

(1) CS appliances: refrigerator, water dispenser, toaster, microwave oven, lights, electric cooker, electric kettle, TV, PC, hair drier, cleaner.

(2) FS appliances: dish-washing machine, hot water tank, washer, drying machine, EVs.

Therefore, the scheduled appliances in DR programs are from FS appliances.

5.1.2 Fundamental Demand Response

In accordance with the principle of DR and the appliances classification above, the objective of the proposed fundamental DR program is to reduce the total cost (TC) which consists of two major parts: the cost of critical appliances $R_t \cdot E_c$ and the cost of flexible appliances $R_t \cdot E_f$, where R_t represents the real-time electricity price (RTP) as detailed in Chapter 4. In addition, E_c and E_f denote the total energy consumption of appliances in CS and FS, respectively. The value of E_c is equivalent to the summation of energy consumption of each individual appliance in CS. For the energy usage of the i^{th} appliance, E_i, it can be denoted as the integration of the appliances power rating and the corresponding service time. The time factors α_i and β_i describe the start and the end of the appliance i operating time. Variable E_f is similar to E_c. Therefore, the load scheduling control (LSC) process can be formulated as:

$$\min TC = R_t(E_c + E_f) \tag{5.1}$$

Subject to:

$$E_c = \sum_{i=1}^{m} E_i, \;\; E_f = \sum_{j=1}^{n} E_j \tag{5.2}$$

$$E_i = \int_{\alpha_i}^{\beta_i} p_i d(t), \; E_j = \int_{\alpha_j}^{\beta_j} p_j d(t) \tag{5.3}$$

$$[\alpha_j, \beta_j] \subseteq [\alpha_e, \beta_e] \tag{5.4}$$

$$\forall t_0 \subseteq [\alpha_e, \beta_e], \; p_{\max}(t_0) \geqslant \sum p_i(t_0) + \sum p_j(t_0) \tag{5.5}$$

The objective function in Equation (5.1) is subject to the constraints in Equation (5.2)-(5.5). Specifically, considering customers' preferences, the service time of the schedualable appliances, $[\alpha_j, \beta_j]$, should be subject to users' expected operating time, $[\alpha_e, \beta_e]$. On the other hand, $p_{\max}(t_0)$ is proposed to limit the maximum load on the grid at time t_0 for safety considerations.

5.2 Electric Vehicle Assisted Demand Response for a Single Household

5.2.1 EV-APS Demand Response Network

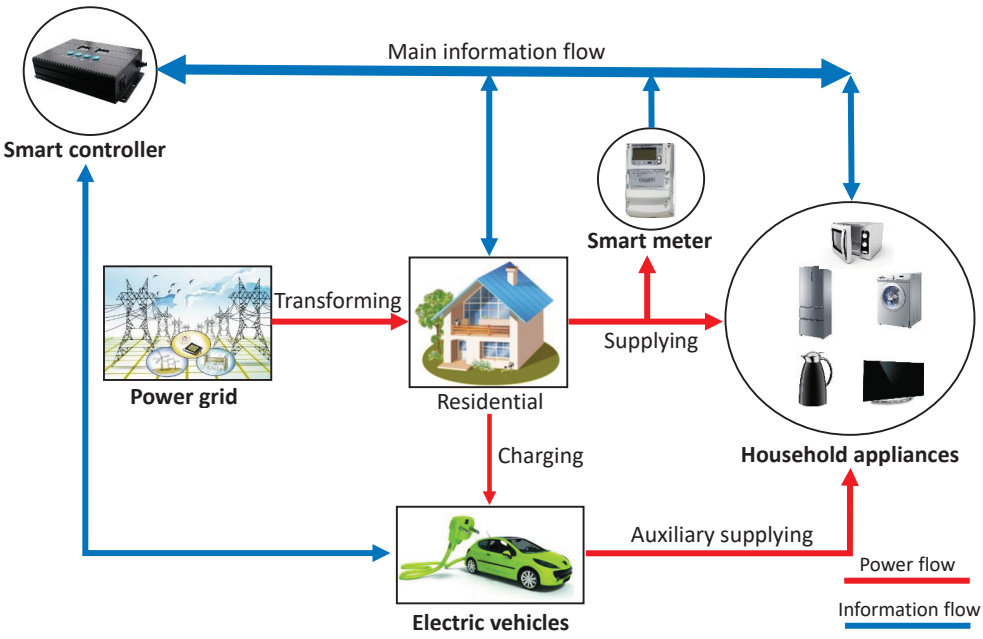


Figure 5.1: Schematic diagram of a DR strategy with EV-APS model for a single household.

The schematic diagram of the proposed DR strategy with EV-APS model is shown in Figure 5.1. Specifically, householders buy electricity from the power grid for daily usage including EV charging, under the dynamic price. Normally,

domestic appliances are directly powered by the main power grid. However, as an interim energy storage unit, the EV is able to supply power for the household appliances in auxiliary capacity on appropriate occasions, especially during high price periods. The time for activation of the EV-APS is dependent on the instructions from the smart controller.

In addition, the smart controller plays the role of supervisor in the network. It regulates the energy source supply and the operating time of the household appliances based on real-time load information which is received from the smart meter, and other signals (*e.g.,* DP, EV status, load priority, *etc.*).

5.2.2 System Models

In Subsection 5.2.2, the formulation of the EV-APS DR strategy consisting of the main power supply model and EV auxiliary power supply model is thoroughly analysed.

Main Power Supply Model

Variables W_t^{grid} and P_t^{grid} are defined as total energy consumption and total power rate on the grid, respectively, at time t. The main power supply model including the corresponding constraints can be represented by:

$$W_t^{\text{grid}} = \int_{T_{\text{in}}}^{T_{\text{term}}} P_t^{\text{grid}} \cdot d(t) \tag{5.6}$$

$$P_t^{\text{grid}} = P_t^{\text{HA}} + P_t^{\text{EV,c}} - P_t^{\text{EV,d}} \tag{5.7}$$

$$P_t^{\text{HA}} = \sum_{j=1}^{n} P_{t,j}^{\text{CS}} + \varepsilon_i \cdot \sum_{i=1}^{m} P_{t,i}^{\text{FS}} \tag{5.8}$$

Subject to:

$$\forall t \in [T_{\text{in}}, T_{\text{term}}], \ P_t^{\text{grid}} \leqslant P_{\text{max}}^{\text{grid}} \tag{5.9}$$

$$P_t^{\text{EV,c}} = 0, \text{ if } P_t^{\text{EV,d}} > 0 \tag{5.10}$$

$$P_t^{\text{EV,d}} = 0, \text{ if } P_t^{\text{EV,c}} > 0 \tag{5.11}$$

Equation (5.6) indicates that the total energy consumption (W_t^{grid}) is equal to the integral of total power (P_t^{grid}) through time. that is between the initial time

T_{in} and the termination time T_{term}. Equation (5.7) illustrates the relationships between the total power and each power consuming component. P_t^{HA} is the load power consumed by the household appliances at time t. $P_t^{\text{EV,c}}$ and $P_t^{\text{EV,d}}$ represents the power rates of the EV charging and discharging, respectively.

Additionally, as shown in Equation (5.8), P_t^{HA} consists of the power cost by CS appliances ($P_{t,j}^{\text{CS}}$) and FS ($P_{t,i}^{\text{FS}}$) appliances, where j and i are the indexes of the appliances. The ε parameters have small positive values (*e.g.,* $1+e^{-8}$, $1+2e^{-8}$ and $1+3e^{-8}$) that are determined by assumption. Thus, the total power of appliances is not affected. This setting meets the requirement of having a priority according to user preferences in scheduling the FS appliances. The smaller the value of ε indicates a higher priority in the scheduling process by the DR program.

$P_{\text{max}}^{\text{grid}}$ is proposed as being constrained (5.9) to limit the maximum power rate on the grid at time t for safety considerations. Further, constraints (5.10) and (5.11) expressed that battery charging and discharging cannot execute simultaneously, otherwise, the battery will be damaged.

EV-APS Model

Determining the EV-APS model requires sufficient knowledge from previous research. From investigation of the current EV market, Table 5.1 illustrates the core parameters of five major brands of EVs around the world [205–207]. The parameters include the maximal battery capacity $W^{\text{EV,max}}$, the discharging power $P_t^{\text{EV,d}}$ and the driving range per charge.

Table 5.1: Major brands of EVs in current market.

Manufacturer and Model	Battery Capacity	Discharging Power	Driving Range per Charge
Tesla, Model-S (EV)	60 kWh	3.0 kW	273 miles
BYD, Tang-100 (HEV)	23 kWh	3.3 kW	63 miles
BMW, i3 (EV/HEV)	33 kWh	2.5 kW	114 miles
GM, Chevrolet Bolt (EV)	60 kWh	-	283 miles
Nissan, Leaf (EV)	30 kWh	-	107 miles

Moreover, multiple charging schemes are provided for each EV. In Section 5.4, the Tesla-Model-S is taken as an example in this study and Table 5.2 shows the

relevant charging schemes that will be considered in the DR strategy. It can be seen that the charging power $P^{\mathrm{EV,c}}$ plays an important role for the grid due to the high power rate of battery charging.

Table 5.2: Tesla-Model-S charging schemes

Charging Circuit	Charging Power	Charging Speed	Charging Time Cost per 100 Miles
Wall connector (1-phase grid)	7.4 kW	22 miles/hr	4.5 hrs
Wall connector (3-phase grid)	11 kW	34 miles/hr	2.9 hrs
High power charger upgrade	16.5 kW	51 miles/hr	2.0 hrs
3-pin domestic adapter	2.3 kW	6.8 miles/hr	14.7 hrs

Further, $W^{\mathrm{EV},(1)}$ and $W^{\mathrm{EV},(2)}$ are defined as the initial EV energy storage when people leave home in the morning of the 1st day and the 2nd day, respectively. Therefore, the EV model can be proposed as below:

$$W^{\mathrm{EV,rem}} = W^{\mathrm{EV},(1)} - W^{\mathrm{EV,trip}} \tag{5.12}$$

$$W^{\mathrm{EV,trip}} = \frac{D^{\mathrm{trip}}}{D^{\max}} \cdot W^{\mathrm{EV,max}} \tag{5.13}$$

$$W^{\mathrm{EV},(2)} = W^{\mathrm{EV,rem}} + W^{\mathrm{EV,c}} - W^{\mathrm{EV,d}} \tag{5.14}$$

$$W^{\mathrm{EV,c}} = \eta_1 \cdot \int_{T_{c,b}}^{T_{c,e}} P_t^{\mathrm{EV,c}} \cdot d(t) \tag{5.15}$$

$$W^{\mathrm{EV,d}} = \eta_2 \cdot \int_{T_{d,b}}^{T_{d,e}} P_t^{\mathrm{EV,d}} \cdot d(t) \tag{5.16}$$

Subject to:

$$\forall t,\ W^{\mathrm{EV,min}} \leqslant W^{\mathrm{EV,rem}} \leqslant W^{\mathrm{EV,max}} \tag{5.17}$$

$$\forall t \in [T_{d,b}, T_{d,e}],\ P_t^{\mathrm{EV,d}} \leqslant P_t^{\mathrm{EV,d,rated}} \tag{5.18}$$

$$\varnothing = [T_{c,b}, T_{c,e}] \cap [T_{d,b}, T_{d,e}] \tag{5.19}$$

Equation (5.12)-(5.13) indicate the state relationships between the initial energy of the 1st day ($W^{\mathrm{EV},(1)}$), the remaining energy ($W^{\mathrm{EV,rem}}$) and the energy consumption on a trip ($W^{\mathrm{EV,trip}}$). It is apparent that $W^{\mathrm{EV,trip}}$ is directly proportional to the driving distance. Additionally, Equation (5.14) implies that the remaining EV energy can be used to cover a portion of household appliance energy usage via battery discharging ($W^{\mathrm{EV,d}}$). The EV will then be charged to an appropriate level for usage on the 2nd day.

Moreover, Equation (5.15) explains that the relationship between the total energy charging ($W^{\mathrm{EV,c}}$) and the charging power rate ($P_t^{\mathrm{EV,c}}$). η_1 is the battery charging efficiency. Time parameters $T_{c,b}$ and $T_{c,e}$ denote the begin time and end time of the charging operation. Meanwhile, the battery discharging occasion which is described in Equation (5.16), is similar to Equation (5.15).

The constraint (5.17) presents a limit on the the actual amount energy for the EV battery. It cannot drop below the minimum allowed battery capacity ($W^{\mathrm{EV,min}}$) or exceed the maximum allowed battery capacity ($W^{\mathrm{EV,max}}$). Constraint (5.18) limits the actual discharging power rate ($P_t^{\mathrm{EV,d}}$) to be less than the rated power of the EV discharging. Additionally, since battery damage will be caused by simultaneous charging and discharging, constraint (5.19) restricts the operation time of battery charging and discharging.

5.2.3 Problem Formulation and Optimisation

According to the previous analysis, the problem in this study can be formulated as minimising the total cost (TC) by scheduling the operating time of the household appliances. Hence, the objective function can be proposed as:

$$\min TC = \int_{T_{\mathrm{in}}}^{T_{\mathrm{term}}} W_t^{\mathrm{grid}} \cdot R_t \cdot d(t) \tag{5.20}$$

where symbol W_t^{grid} represents the total energy bought from the power grid in time period $[T_{\mathrm{in}}, T_{\mathrm{term}}]$. Additionally, the price variable R_t is time dependent and varies hourly depending on the total load demand [208]. The DP tariff that is used in the simulation is given in Figure 5.3 in Section 5.4.

In order to obtain the optimal solution and reduce cost to a minimum, the exhaustive search technique can be used on the basis of the established models. The description of the technique is not the focus of this work and is not emphasised here.

Under the given constraints, the program continuously searches the solutions of appliances allocation by minimising the global cost according to the DP signals. The objective appliances in FS are scheduled in sequence based on the pre-set

priority. Meanwhile, as the auxiliary power source, the operating time of EV discharging is dependent on the EV status and the load demand. In this study, the remaining EV energy is assumed to be firstly consumed in high price hours to ensure maximum utilisation of the available EV energy.

5.3 Electric Vehicles Assisted Demand Response for a Multi-Household Network

5.3.1 EVs Assisted NES Demand Response Network

An overview of the proposed demand response framework for a residential community with multiple households is illustrated in Subsection 5.3.1.

The block diagram of the proposed DR framework with EVs assisted NES model is shown in Figure 5.2. In this study, it is assumed that each household in the community is registered in the network and controlled by the corresponding automatic control unit (ACU) which plays the role as an instructor for each household. The ACU regulates the power supply and the operating time of the household appliances (HAs, *e.g.,* flexible appliances and critical appliance [8, 177]) based on the dynamic load information which is received from the smart meter, and other request signals (*e.g.,* EV status, scheduling priority, DP, *etc.*). In addition, the centralised control unit (CCU) that is the highest controller in the network globally monitors the status of the ACUs and optimally manages the EVs assisted NES model through information flows. In the proposed DR framework, customers in the network are registered for two types of connection: V2H connection and V2N connection.

Specifically, the householders buy electricity from the power grid for daily consumption including supplying HAs and EVs charging, under the DP tariff. On the one hand, the domestic appliances are directly powered by the public power grid in general. However, households which are outfitted with an EV are able to provide power from the EV battery to their HAs on appropriate occasions, such as peak demand periods or power grid outage, via V2H connection. On the other hand, since a limited number of households are equipped with an

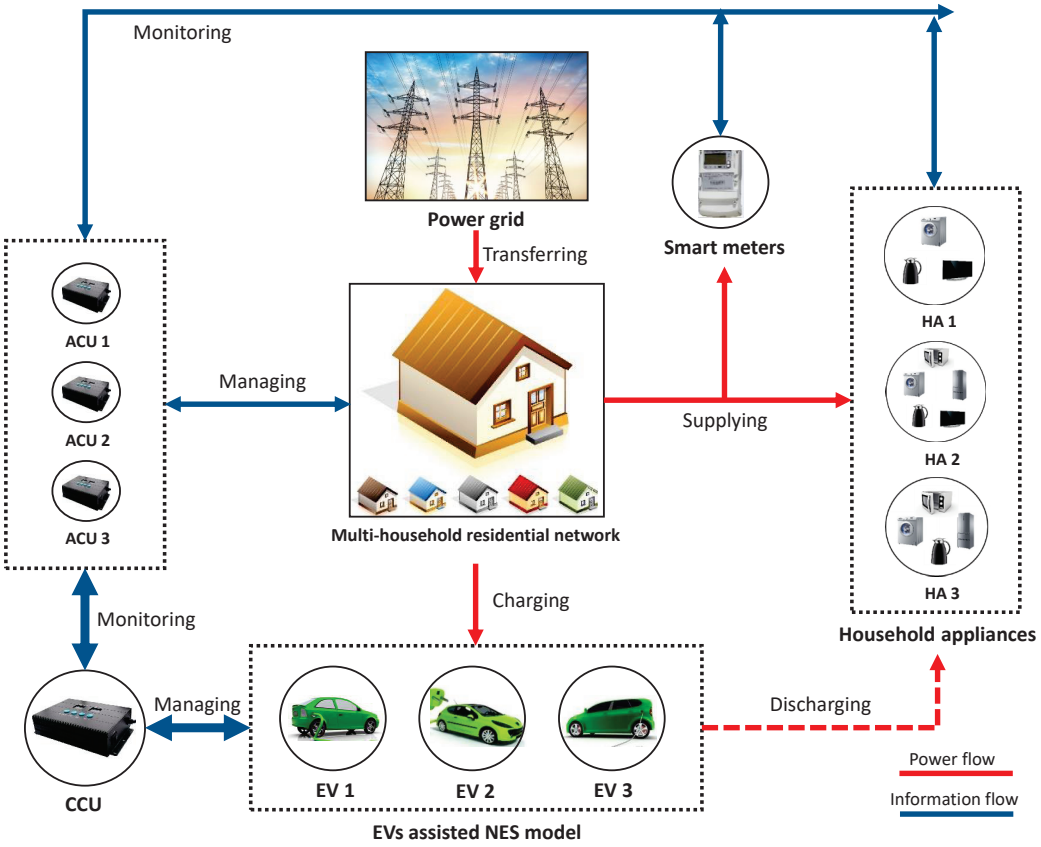


Figure 5.2: Schematic diagram of a DR framework for a multi-household network.

EV, the households without the energy storage unit may need power assistance from the NES model via a V2N connection, particularly during high price times. When there is energy available in EVs, the CCU determines when and how to allocate the available energy to the personal house or the neighbours' houses who have the energy assistance requirements. Ideally, the EV energy will satisfy the demand of the EV owner in priority. The energy transaction in the neighbourhood happens when the power grid is not able to fulfill the demand or serve the load in peak demand and cost periods. Thus, a customer can receive the power from a neighbour at comparatively lower prices.

5.3.2 System Models

EVs are utilised as flexible energy storage units to enable energy trading in a neighbourhood. The following section presents the mathematical modeling of the

system components in detail.

Global Energy Balance Model

In order to precisely present the energy transactions between each component in the network with K households, W_t^{grid} and $W_{k,t}^{\text{grid}}$ are defined as the total energy consumption of the entire network and the k^{th} household, respectively, in a time period $[T_{\text{in}}, T_{\text{term}}]$. Afterwards, the global energy model can be proposed as:

$$W_t^{\text{grid}} = \sum_{k=1}^{K} W_{k,t}^{\text{grid}} \tag{5.21}$$

where

$$W_{k,t}^{\text{grid}} = \int_{T_{\text{in}}}^{T_{\text{term}}} P_{k,t}^{\text{grid}} \cdot d(t) \tag{5.22}$$

Moreover, considering the specific power including CS appliances ($P_{k,t}^{\text{CS}}$), FS appliances ($P_{k,t}^{\text{FS}}$), EV charging ($P_{k,t}^{\text{EV,c}}$) and EV discharging ($P_{k,t}^{\text{EV,d}}$) in the network, $P_{k,t}^{\text{grid}}$ can be extended as in Equation (5.23)-(5.24).

$$P_{k,t}^{\text{grid}} = P_{k,t}^{\text{HA}} + \alpha \cdot \left(\beta \cdot P_{k,t}^{\text{EV,c}} - (1-\beta) \cdot P_{k,t}^{\text{EV,d}}\right) \tag{5.23}$$

$$P_{k,t}^{\text{HA}} = \sum_{j=1}^{m} P_{k,t,j}^{\text{CS}} + \varepsilon_i \cdot \sum_{i=1}^{n} P_{k,t,i}^{\text{FS}} \tag{5.24}$$

Subject to:

$$\forall t \in [T_{\text{in}}, T_{\text{term}}], \ P_{k,t}^{\text{grid}} \leqslant P_{k,\max}^{\text{grid}} \tag{5.25}$$

$$\forall t, \ \sum_{k=1}^{K} P_{k,t}^{\text{grid}} \leqslant P_{\max}^{\text{grid}} \tag{5.26}$$

Binary parameters α and β in Equation (5.23) are both used to indicate the EV status that is given as:

$$\text{EV Status} = \left\{ \begin{array}{ll} \text{Disabled}, & \text{if } \alpha = 0, \ \beta = \forall \\ \text{Charging}, & \text{if } \alpha = 1, \ \beta = 1 \\ \text{Discharging}, & \text{if } \alpha = 1, \ \beta = 0 \end{array} \right\} \tag{5.27}$$

Furthermore, $P_{k,t}^{\text{HA}}$ in Equation (5.24) denotes the load of electric appliances consisting of CS load $P_{k,t,j}^{\text{CS}}$ and FS load $P_{k,t,i}^{\text{FS}}$ at time t, where j and i are the indexes of the appliances. The parameter ε_i indicates the scheduling priority of the shiftable appliances, it is a small positive value (*e.g.,* $1+e^{-8}$, $1+2e^{-8}$ and $1+3e^{-8}$) determined by a householder according to their scheduling preferences.

Besides, the maximum power rate of a single household $P^{\text{grid}}_{k,\max}$ and the maximum power rate of the network $P^{\text{grid}}_{\max}$ are proposed in Equation (5.25) and (5.26), respectively, to limit the real-time load for safety considerations.

EVs Assisted NES Model

In a residential community, different classes of customers exist. It is not possible for every household to purchase an EV. Thus, it is assumed that only some houses have a EV and are indexed as $\hat{k}$, and the houses without an EV are indexed as $\tilde{k}$. Then, $W^{\text{EV},(1)}_{\hat{k}}$ and $W^{\text{EV},(2)}_{\hat{k}}$ are defined as the initial energy within an EV battery when the EV leaves the home on the 1st day and the 2nd day, respectively. $W^{\text{EV,rem}}_{\hat{k}}$ represents the remaining energy within an EV. The energy cost on a daily trip is proposed as $W^{\text{EV,trip}}_{\hat{k}}$. Additionally, $D^{\text{trip}}_{\hat{k}}$ and $D^{\max}_{\hat{k}}$ are proposed to indicate the actual travel distance of the vehicle and the maximum travel distance with a fully charged EV. Moreover, the energy charging to the EV and discharging from the EV are assumed as $W^{\text{EV,c}}_{\hat{k}}$ and $W^{\text{EV,d}}_{\hat{k}}$, respectively. The EV balance model with the relevant constraints for a single house are proposed as:

$$W^{\text{EV,rem}}_{\hat{k}} = W^{\text{EV},(1)}_{\hat{k}} - W^{\text{EV,trip}}_{\hat{k}} \tag{5.28}$$

$$W^{\text{EV,trip}}_{\hat{k}} = \frac{D^{\text{trip}}_{\hat{k}}}{D^{\max}_{\hat{k}}} \cdot W^{\text{EV,max}}_{\hat{k}} \tag{5.29}$$

$$W^{\text{EV},(2)}_{\hat{k}} = W^{\text{EV,rem}}_{\hat{k}} + W^{\text{EV,c}}_{\hat{k}} - W^{\text{EV,d}}_{\hat{k}} \tag{5.30}$$

Subject to:

$$\forall t,\ W^{\text{EV,min}}_{\hat{k}} \leqslant W^{\text{EV,rem}}_{\hat{k}} \leqslant W^{\text{EV,max}}_{\hat{k}} \tag{5.31}$$

$$\tau \cdot W^{\text{EV,max}}_{\hat{k}} \leqslant W^{\text{EV},(1)}_{\hat{k}} \approx W^{\text{EV},(2)}_{\hat{k}} \leqslant W^{\text{EV,max}}_{\hat{k}} \tag{5.32}$$

where $W^{\text{EV,min}}_{\hat{k}}$ and $W^{\text{EV,max}}_{\hat{k}}$ in (5.31) represent the minimum and the maximum allowed EV battery capacity, respectively. However, constraint (5.32) is proposed to ensure that the EV leaves home with an appropriate energy storage level, where τ is a threshold parameter.

Moreover, considering the power impacts on the grid, $P^{\text{EV,c}}_{\hat{k},t}$, $P^{\text{EV,d,v2h}}_{\hat{k},t}$ and $P^{\text{EV,d,v2n}}_{\hat{k},t}$ are utilised to describe the power rates of EV charging, EV discharging

via V2H and EV discharging via V2N at time t, respectively. Therefore, $W_{\hat{k}}^{\mathrm{EV,c}}$ and $W_{\hat{k}}^{\mathrm{EV,d}}$ can be extended to:

$$W_{\hat{k}}^{\mathrm{EV,c}} = \eta_{\hat{k}}^{\mathrm{c}} \cdot \left\{ \sum_{l=1}^{L} \int_{T_{\hat{k},l}^{\mathrm{c},1}}^{T_{\hat{k},l}^{\mathrm{c},2}} P_{\hat{k},t}^{\mathrm{EV,c}} \cdot d(t) \right\} \tag{5.33}$$

$$\begin{aligned} W_{\hat{k}}^{\mathrm{EV,d}} = \frac{1}{\eta_{\hat{k}}^{\mathrm{d,v2h}}} \cdot & \left\{ \sum_{m=1}^{M} \int_{T_{\hat{k},m}^{\mathrm{d},1}}^{T_{\hat{k},m}^{\mathrm{d},2}} P_{\hat{k},t}^{\mathrm{EV,d,v2h}} \cdot d(t) \right\} \\ & + \frac{1}{\eta_{\hat{k}}^{\mathrm{d,v2n}}} \cdot \left\{ \sum_{n=1}^{N} \int_{T_{\hat{k},n}^{\mathrm{d},1}}^{T_{\hat{k},n}^{\mathrm{d},2}} P_{\hat{k},t}^{\mathrm{EV,d,v2n}} \cdot d(t) \right\} \end{aligned} \tag{5.34}$$

Subject to:

$$\eta_{\hat{k}}^{\mathrm{c}},\ \eta_{\hat{k}}^{\mathrm{d,v2h}} \text{ and } \eta_{\hat{k}}^{\mathrm{d,v2n}} \in (0,1) \tag{5.35}$$

$$\forall t \in [T_{\hat{k},m}^{\mathrm{d},1}, T_{\hat{k},m}^{\mathrm{d},2}],\ P_{\hat{k},t}^{\mathrm{EV,d,v2h}} \leqslant P_{\hat{k}}^{\mathrm{EV,rated}},\ P_{\hat{k},t}^{\mathrm{EV,d,v2h}} \leqslant P_{\hat{k},t}^{\mathrm{act}} \tag{5.36}$$

$$\forall t \in [T_{\hat{k},n}^{\mathrm{d,n}}, T_{\hat{k},n}^{\mathrm{d},2}],\ P_{\hat{k},t}^{\mathrm{EV,d,v2n}} \leqslant P_{\hat{k}}^{\mathrm{EV,rated}},\ P_{\hat{k},t}^{\mathrm{EV,d,v2n}} \leqslant P_{\tilde{k},t}^{\mathrm{act}} \tag{5.37}$$

$$\forall [T_{\hat{k},l}^{\mathrm{c},1}, T_{\hat{k},l}^{\mathrm{c},2}] \cap \forall \left\{ [T_{\hat{k},m}^{\mathrm{d},1}, T_{\hat{k},m}^{\mathrm{d},2}] \cup [T_{\hat{k},n}^{\mathrm{d},1}, T_{\hat{k},n}^{\mathrm{d},2}] \right\} = \varnothing \tag{5.38}$$

where $\eta_{\hat{k}}^{\mathrm{c}}$, $\eta_{\hat{k}}^{\mathrm{d,v2h}}$ and $\eta_{\hat{k}}^{\mathrm{d,v2n}}$ denote the efficiencies of the corresponding EV behaviors. Specifically, $\eta_{\hat{k}}^{\mathrm{c}}$ represents the charging efficiency of $\hat{k}^{\mathrm{th}}$ EV. Symbols $\eta_{\hat{k}}^{\mathrm{d,v2h}}$ and $\eta_{\hat{k}}^{\mathrm{d,v2n}}$ are the discharging efficiencies of $\hat{k}^{\mathrm{th}}$ EV via V2H and V2N connections, respectively. Since the EV behaviors are discontinuous and may execute at different periods, different time labels are proposed. For example, time parameters $T_{\hat{k},l}^{\mathrm{c},1}$ and $T_{\hat{k},l}^{\mathrm{c},2}$ in Equation (5.33) represent the start time and the end time of the l^{th} charging period. The definitions of the time parameters in EV discharging periods are similar to Equation (5.33).

Furthermore, the discharging power via V2H connection ($P_{\hat{k},t}^{\mathrm{EV,d,v2h}}$) cannot exceed the rated power ($P_{\hat{k}}^{\mathrm{EV,rated}}$) nor the actual power required by the household ($P_{\hat{k},t}^{\mathrm{act}}$) as it is shown in constraint (5.36). Constraint (5.37) is similar to (5.36). Variable $P_{\tilde{k},t}^{\mathrm{act}}$ represents the actual load demand of the neighbours household which receives the power assistance from the EV household via V2N connection. As shown in constraint (5.38), EV charging and discharging are not allowed to operate simultaneously for the purpose of protecting the EV battery from damage.

Energy Trading Model in Neighbourhood

The proposed EVs assisted NES model ensures energy trading in the neighbourhood via V2H and V2N connections. However, it is necessary to declare the trading policy in the neighbourhood in advance as follows.

(i) The EV energy will be provided in priority to satisfy the load demand of the household which owns the EV.

(ii) After (i), the available EV energy will be used in priority to supply the households which are not equipped with any energy storage units (*e.g.,* EVs.).

(iii) If multiple EVs have available energy, the EV with the most energy reserve will be adopted in priority to assist a neighbours' load demand.

(iv) If multiple households require energy assistance, the household which requires the largest load demand during a high pricing period will receive the energy sharing in priority and each house can obtain energy assistance from only one EV energy provider.

(v) The distribution of the EV energy will follow the principle of maximising the benefits of the EV provider.

In addition to these, $B_{\hat{k}}^{\mathrm{NES}}$ and $B_{\tilde{k}}^{\mathrm{NES}}$ are proposed to describe the obtained benefits of the households who sold EV energy and received energy assistance, respectively, via the NES model. Hence, $B_{\hat{k}}^{\mathrm{NES}}$ and $B_{\tilde{k}}^{\mathrm{NES}}$ can be formulated as follows.

$$B_{\hat{k}}^{\mathrm{NES}} = \theta\% \cdot \left(C_{\tilde{k}}^{\mathrm{dmd}} - C_{\hat{k}}^{\mathrm{EV,c}}\right) \tag{5.39}$$

$$B_{\tilde{k}}^{\mathrm{NES}} = \left(1 - \theta\%\right) \cdot \left(C_{\tilde{k}}^{\mathrm{dmd}} - C_{\hat{k}}^{\mathrm{EV,c}}\right) \tag{5.40}$$

subject to:

$$C_{\tilde{k}}^{\mathrm{dmd}} - C_{\hat{k}}^{\mathrm{EV,c}} > 0 \tag{5.41}$$

where $C_{\tilde{k}}^{\mathrm{dmd}}$ is the cost for electricity demand without EV sharing within a none EV household and $C_{\hat{k}}^{\mathrm{EV,c}}$ is the cost for EV charging of the energy sharing part. θ is a profit distribution parameter and normally $\theta\% = 0.5$, which means the participants in energy trading share the profits equally. However, the energy transaction via NES model occurs only when it is profitable as shown in Equation

(5.41). Obviously, this type of EV based energy sharing model is beneficial for the trading participants on both sides.

5.3.3 Problem Formulation and Optimisation

The objective of this work is to minimise the total daily cost of energy usage for the residential network with K households as well as shaping the load to a proper level in peak demand time. To begin with, the day is split into equal time divisions with a time interval and indexed as t. The total cost function is given in Equation (5.42).

$$\min \text{ TC} = \sum_{k=1}^{K} \left\{ \sum_{t=1}^{24} (R_t \cdot W_{k,t}^{\text{grid}}) - B_k^{\text{NES}} \right\} \tag{5.42}$$

where R_t is the real-time electricity price, $W_{k,t}^{\text{grid}}$ is the energy consumed on the grid of the k^{th} household and B_k^{NES} represents the cost benefit that the householder can obtain in energy trading in the neighbourhood by using the NES model.

However, in order to minimise TC, it has to minimise each TC_k which denotes the total cost of the k^{th} household in the network. Therefore, the objective function can be represented as:

$$\begin{aligned} \min \text{ TC} &= \sum_{k=1}^{K} \min \ \{\text{TC}_k\} \\ &= \sum_{k=1}^{K} \left\{ \min \ \left\{ \sum_{t=1}^{24} (R_t \cdot W_{k,t}^{\text{grid}}) \right\} - \max \ \left\{ B_k^{\text{NES}} \right\} \right\} \end{aligned} \tag{5.43}$$

According to the objective function in Equation (5.43) and the energy trading policy in 5.3.2, the optimisation process can be executed in two stages. First, it minimises the total cost for electricity bills by optimally allocating the EV energy via V2H connection. Second, it maximises the benefit from energy transaction in the neighbourhood by optimally distributing the available energy via V2N connection. Based on the previous model descriptions, both optimisation stages are linear problems. Therefore, mixed-integer linear programming (MILP) has been used to obtain the optimal solution. However, the description of the technique is not the focus of this study and so it is not emphasised here.

Under the given models and the relevant constraints, the proposed DR strategy is able to optimally schedule appliances within the multi-household network in accordance with the comprehensive affecting factors, such as EVs' behaviors, user preferences, and load scheduling priorities. The maintenance cost for EVs and home appliances is neglected in this work.

5.4 Case Study

5.4.1 Case 1 - EV Assisted DR Strategy for a Single Household

Case 1 demonstrates how the proposed EV-APS DR strategy can be applied to a household to alleviate the load burden in peak demand periods and save electricity bills.

Case Description

First of all, the selected time interval for the optimisation is set as 3 minutes (0.05 hr). The households comprise over 15 types of commonly used loads covering both CS and FS appliances. The rated power and the pre-set operating time of the corresponding appliances are given in Table 5.3. The EV and four other commonly used appliances, hot water tank, dishwashing machine, washer and drying machine, are considered as the flexible loads in this study.

In addition, the ε parameters are given to indicate the priorities of the related loads. According to user preferences, they are assumed as, $\varepsilon_0 < \varepsilon_1 < \varepsilon_2 < \varepsilon_3 < \varepsilon_4$, which means EV charging is assigned the highest priority in scheduling among all FS loads. Besides, in accordance with the operating habits, the objective scheduling time for these appliances are set randomly, such as EV charging, [0:00-8:00]; hot water tank, [17:00-22:00]; dishwashing machine, [18:30-24:00]; washer, [17:00-24:00]; drying machine, [0:00-8:00].

Moreover, the Tesla-Model-S (EV) with a battery rating of 30 kWh (up to 60 kWh) is employed in the case study. On the one hand, it is provided with a charging wall connector (1-phase grid) limited to a charging power of 7.4 kW. On the other hand, the discharging power for household appliances is up to 3.0 kW

as shown in Table 5.1. The charging and discharging efficiencies are considered as η_1=η_2=0.95. It is also considered that the householder always arrives home at 5:00 p.m. with 18 kWh (60%) remaining energy in the EV battery and leaves home at 8:00 a.m. the next morning with a fully charged EV battery (100±5%, 30±1.5 kWh). The minimum remaining energy in the EV is restricted to 7.5 kWh (25±5%) to avoid deep discharge as this may cause damage to the battery and reduce battery life [209].

Table 5.3: Pre-set household appliances information

CS Appliances	Power (kW)	Operating Time
Refrigerator	0.1	0:00-24:00
Water Dispenser	0.1	0:00-24:00
Toaster	0.6	7:30-7:45
Microwave Oven1	2.4	7:30-8:00
Lights	0.4	17:00-24:00
Electric Cooker	0.5	17:00-17:45
Electric Kettle	2.0	17:15-17:30
Microwave Oven2	2.4	17:30-18:00
Television	0.2	18:00-23:00
Cleaner	0.9	19:00-19:30
Laptop	0.4	21:00-23:30
Hair-Drier	2.0	22:30-23:00
FS Appliances	Power (kW)	Operating Time
EV, ε_0	7.4	18:00-20:00
Hot Water Tank, ε_1	2.5	20:00-22:00
Dish Machine, ε_2	0.5	18:30-19:15
Washer, ε_3	0.6	19:00-20:15
Drying Machine, ε_4	2.5	20:30-21:30

Furthermore, the UK dynamic pricing data of a typical day [210] used in this case is presented in Figure 5.3.

Simulation Results

It is assumed that the threshold of household demand is 8 kW in this study. Figure 5.4 presents the overall load shaping results of the household appliances. Specifically, Figure 5.4(a) shows the original load profile without DR. It can be seen that the peak demand time occurs between 6:00 p.m. and 8:10 p.m. The

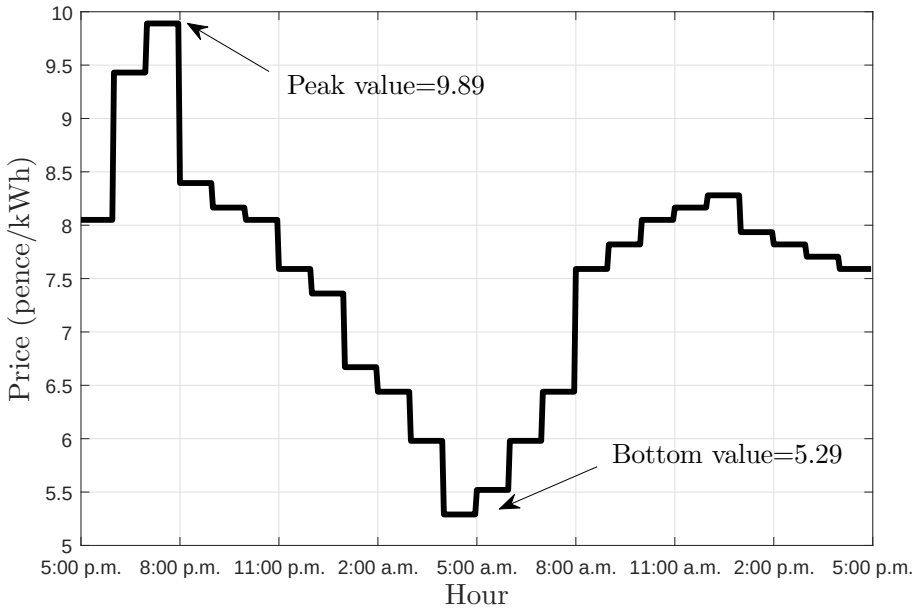
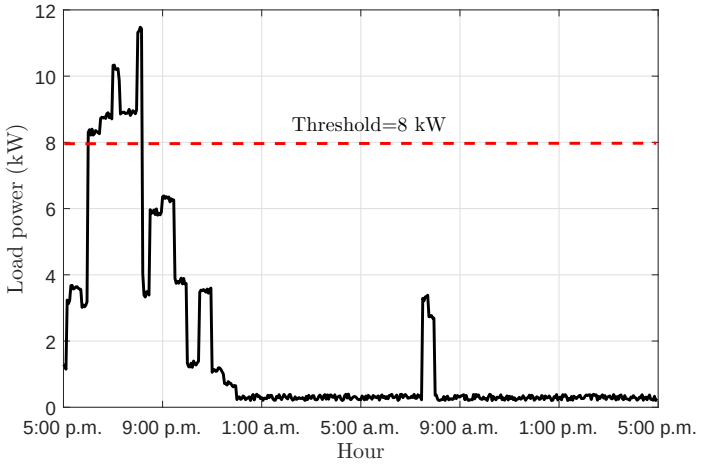


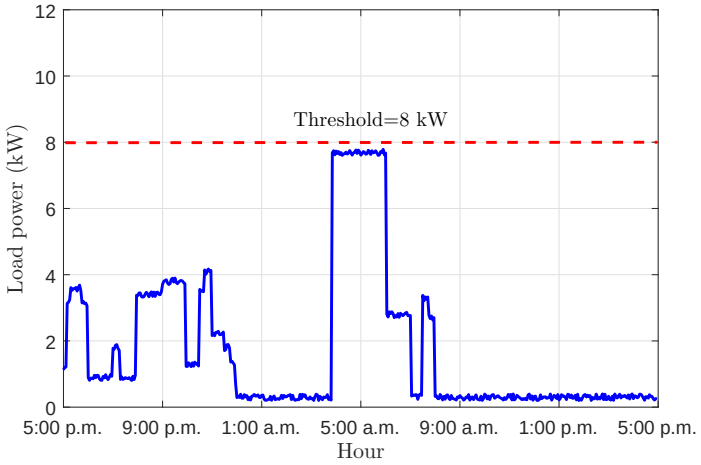
Figure 5.3: UK real-time pricing data

total house load exceeds the 8 kW limit during this period and the maximum load demand is 11.5 kW which occurs at around 8:00 p.m. Additionally, (b) and (c) present the load profiles after scheduling by using the LSC DR strategy (presented in Section 5.1) and the proposed EV-APS DR strategy, respectively. It is seen that the load burden is alleviated and the load decreases to an appropriate level in both (b) and (c). Nonetheless, compared with the results in (b), the load demand in (c) between 6:00 p.m. and 9:40 p.m. approaches to a very low level since the EV discharging is activated during this time. As a consequence, the EV takes 3.2 hours to charge as it is shown in (c), which is longer than the charging time (2.1 hours) in (b).

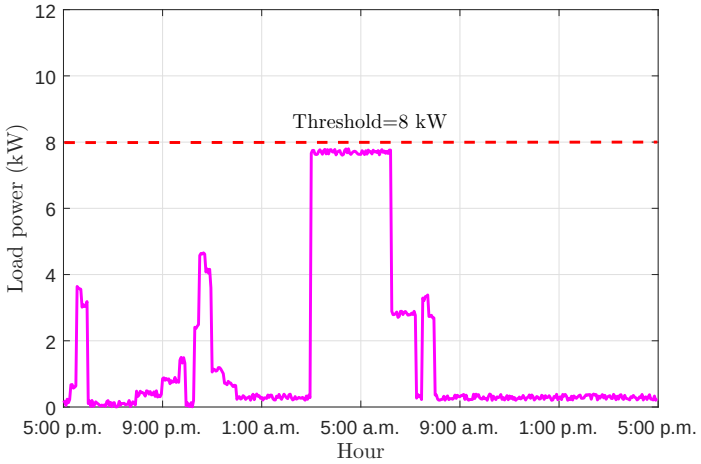
Moreover, since the EV plays a great role in supplying power in the model, the real-time EV remaining energy variation at household parking station when using the proposed EV-APS DR strategy is illustrated in Figure 5.5. Specifically, the EV arrives at home at 5:00 p.m. and between 5:00 p.m. and 10:18 p.m., EV discharging is activated and some of the household appliances are powered by the EV until the amount of EV remaining energy reaches the minimum threshold (7.5 kWh). However, the EV is charged from 3:00 a.m. to 6:18 a.m. the next morning



(a)



(b)



(c)

Figure 5.4: Overall load shaping results. The load profiles of (a) without DR; (b) by the LSC DR; (c) by the proposed EV-APS DR

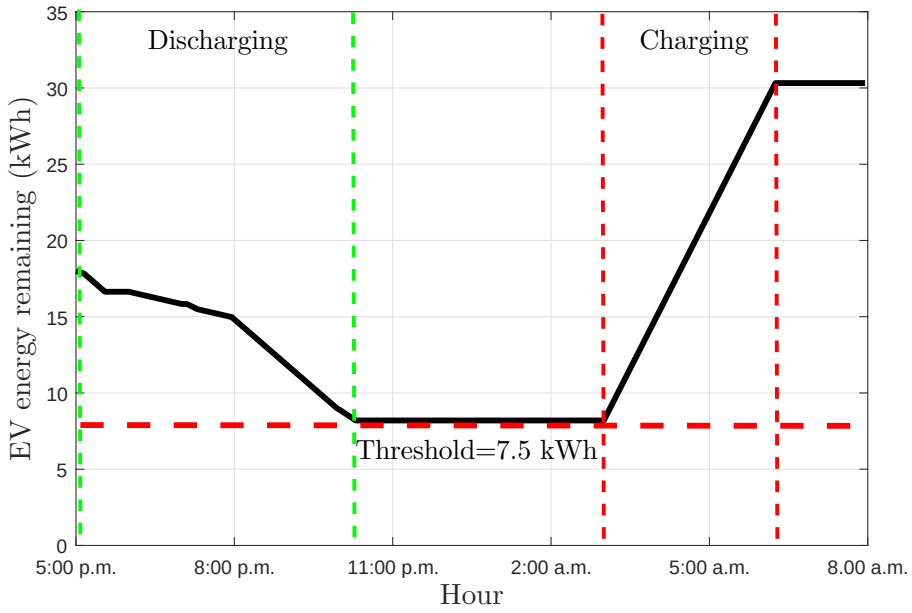


Figure 5.5: Real-time EV remaining energy variation at parking station

to enable the EV to leave with a fully charged battery at 8:00 a.m. According to the results, it can be seen that the remaining EV energy variation directly corresponds to the load curve in Figure 5.4(c), indicating that the proposed EV-APS DR is correct and feasible.

Furthermore, Figure 5.6 shows the accumulative probabilities of the reshaped load distributions by DR strategies during the peak load demand period, which is between 5:00 p.m. and 12:00 midnight. Based on the figure, it can be seen that the probabilities for the case $P_{\text{grid}} < 1$ kW of the original load profile without DR, the LSC DR shaping profile and the EV-APS DR shaping profile are 7.1%, 24.3%, and 72.9%, respectively. For the case $P_{\text{grid}} < 3$ kW, the probabilities are 23.6%, 53.1% and 86.4%, respectively. The results indicate that the load shaping performance of the EV-APS DR strategy is the best as a higher percentage load is shaped to a low level compared with the others, proving that the proposed method is an effective tool in load balancing.

The total cost is another issue of concern for customers. On the basis of the DP tariff, the daily electric cost can be obtained. Figure 5.7 presents the accumulative cost comparison between different DR strategies. Based on the results, it can

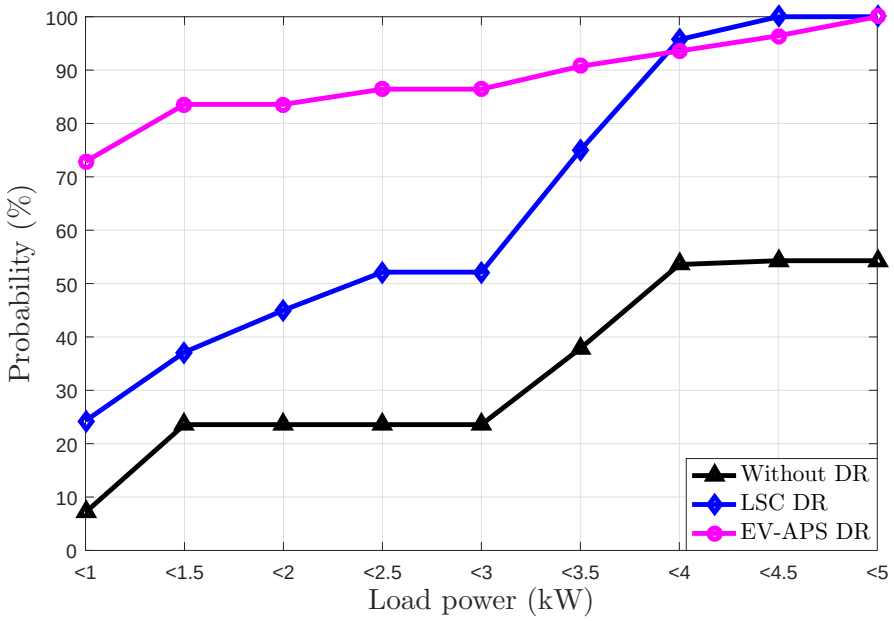


Figure 5.6: Accumulative probability of the load distribution during peak load demand hours

be seen that the proposed EV-APS DR strategy performs superior to other approaches in comparison. The original total cost for the pre-set household appliances in a day is about £3.60. However, it decreases to £2.90 and £2.50 by using the LSC DR and the EV-APS DR, respectively. The total cost savings are about £0.70 and £1.10, which are equivalent to 19.4% and 30.6%, respectively. Compared with the LSC DR strategy in literature, the proposed DR strategy in this work has a better performance in load shaping and a higher cost saving percentage (11.2% improved compared with LSC DR).

5.4.2 Case 2 - EVs Assisted DR Strategy for a Multi-Household Network

A case study to demonstrate how the DR strategy with the EVs assisted NES model can be implemented at the side of a residential community, to simultaneously save on electricity bills and alleviate the load burden during peak demand time, is proposed here.

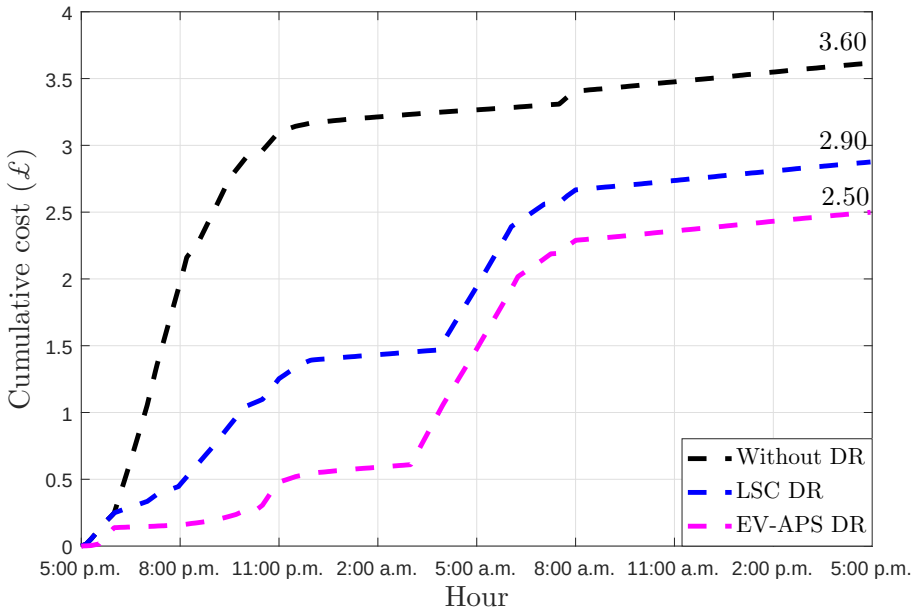


Figure 5.7: Accumulative cost comparison results between DR strategies

Case Description

The optimisation problem for the total cost minimisation is formulated as a linear programming problem aimed to reduce the daily bill of each household as much as possible. In the case study, the selected time interval for the optimisation is set at 3 minutes. The adopted multi-household network is assumed to comprise 5 households. For each household, over 15 types of commonly used domestic appliance covering both FS and CS are considered.

In addition to these, the ε parameter which is used to indicate the scheduling priority of FS appliances (*e.g.,* hot water tank, dishwashing machine, washer, *etc.*) is randomly selected to simulate various circumstances. The objective scheduling time for FS appliances is set between 5 p.m. to 12 midnight. according to users' preferences.

Moreover, as not all users are able to purchase an electric vehicle, only 3/5 of the households are assumed to be equipped with EVs to support neighbour energy sharing. For each EV device, a battery capacity of 35 kWh is used. The charging and discharging (via V2H and V2N) efficiencies are all considered to be 0.95 for convenience. The minimum remaining energy in an EV is restricted

to 10% (τ = 0.1) of the battery capacity to avoid the deep discharging. Also, the parameters of EV status, time of arriving (ToA), time of leaving (ToL), charging rate (CR), discharging rate (DCR) and energy remaining on arriving home (ERoA) of the specific EV within each household are given in Table 5.4.

Table 5.4: Electric vehicle parameter specification

Parameters	House #1	House #2	House #3	house #4	House #5
EV Status	Active	Active	Active	Disable	Disable
ToA (1st day)	5 p.m.	6 p.m.	7 p.m.	-	-
ToL (2nd day)	8 a.m.	9 a.m.	10 a.m.	-	-
CR (kW)	7.5	6.5	5.5	-	-
DCR (kW)	3.5	3	2.5	-	-
ERoA (kWh)	26	24	22	-	-

Simulation Results

Figure 5.8 presents the overall load shaping results obtained using different DR programs. It is assumed that the threshold of the overall load demand is 25 kW. Specifically, it can be seen that the LSC DR program can slightly alleviate the load burden, particularly around 9 p.m. This is because limited appliances are scheduled and none of the EVs are adopted in the LSC DR program. However, the load shaping performances of using EVs without NES and EV assisted NES in (c) and (d), respectively, are much better than the results in (a) and (b). The load demand of the entire network in both (c) and (d) has remained below the threshold apparently due to the EVs discharging contributions.

In addition, compared with the load distribution in (c), the load demand in (d) approaches to a lower level during the peak time period around 7 p.m. to 9 p.m. This is because the households without EVs received energy assistance from neighbours via the V2N so that the overall load demand on the grid decreases. As a consequence, it is clear that the EVs take more time to charge their batteries at the off-peak time for usage the 2nd day. Moreover, since the EVs play a great role in power transaction within the network, the real-time energy remaining variations of EVs (#1, #2 and #3) at the parking station are illustrated in Figure 5.9.

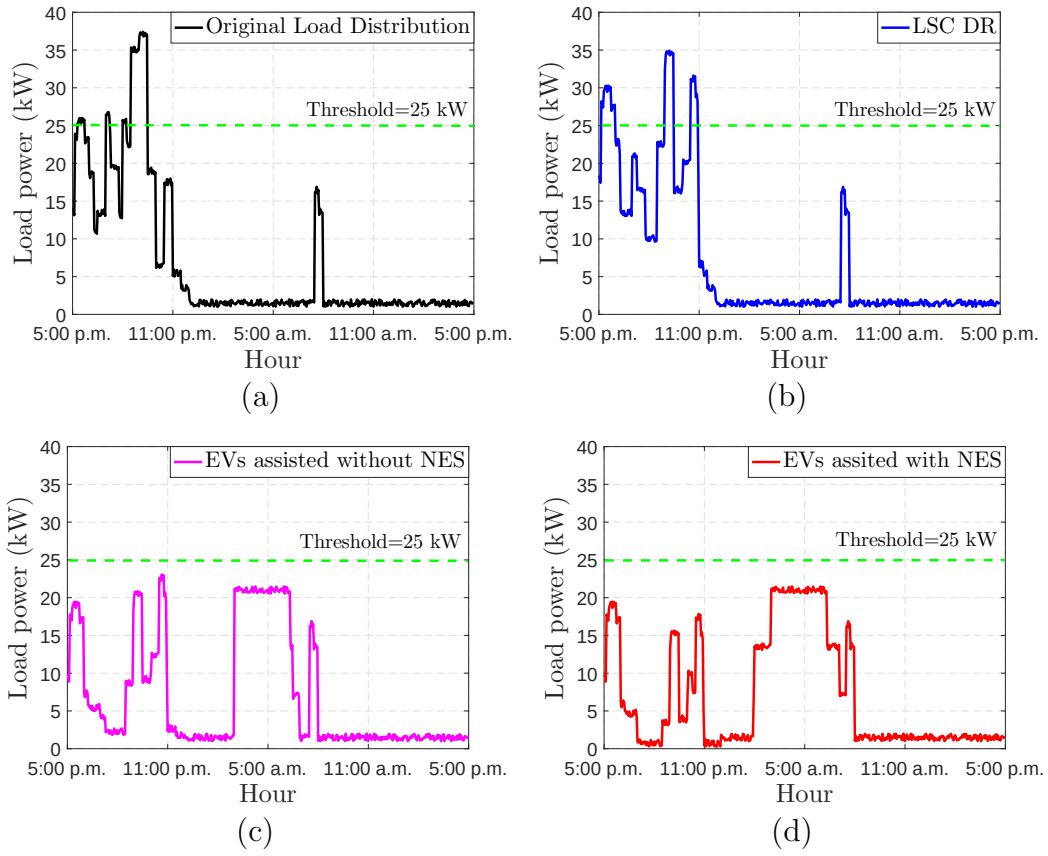


Figure 5.8: Overall load shaping results by using different DR programs. The load profiles of (a) without DR; (b) by LSC DR; (c) by EV without NES DR; (d) by EV assisted NES DR.

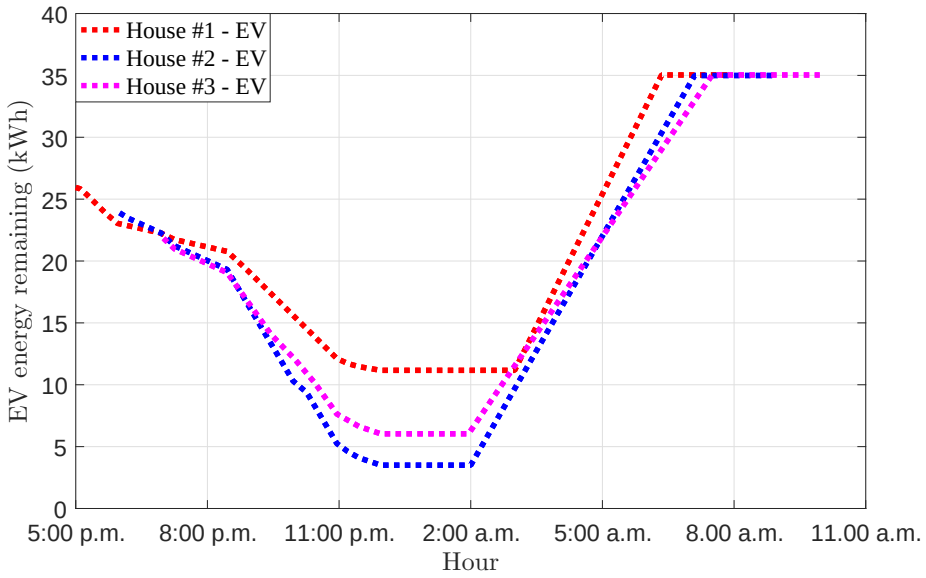


Figure 5.9: Real-time energy remaining variations of EVs at parking station.

In terms of the daily electricity cost, the proposed approach offers more benefits when compared with DR programs presented in the literature [8, 9] as shown in Table 5.5. According to the cost results, the proposed DR with an EVs assisted NES model performs the best with the lowest cost comparison for all cases. Specifically, as house #1 does not participate in energy trading in the neighbourhood due to its lower distributing priority, there is not any cost difference between using EVs with and without NES. Nonetheless, as the energy providers in the transaction, the costs of house #2 and house #3 are reduced by 47.3% and 46.1%, respectively, by adopting the EVs assisted NES model compared with the original cost. Additionally, about 10.5% and 7.4% cost reduction can be achieved compared with the method of using EVs without NES. On the other side of the trading, house #4 and house #5 that are not equipped with EVs also obtain the benefits from energy sharing. About £0.24 and £0.32 which are equivalent to 10.4% and 13.7% cost savings can be gained during the transaction for house #4 and house #5, respectively.

Table 5.5: Daily cost (£) comparison by adopting different DR programs (best performance is marked in bold).

Methods	House #1	House #2	House #3	House #4	House #5
Original	3.15	3.55	3.88	2.31	2.34
LSC	3.09	3.49	3.73	2.24	2.26
EVs Without NES	1.83	2.09	2.30	2.31	2.34
EVs Assisted NES	**1.83**	**1.87**	**2.13**	**2.07**	**2.02**

In overview, for this selected residential community including 5 individual households, the total payment saving is about £5.31 which is equivalent to 34.9% in this case. Obviously, the adopted EVs based NES model is beneficial for the energy trading participants on both sides and significant improvements can be achieved compared with DR programs in the literature.

5.5 Summary

The aim of this chapter was to develop efficient DR strategies with EVs, to jointly optimise the household appliance scheduling and economic cost based on DP for

different scales of households. An EV-APS based DR strategy has been proposed first. Afterwards, a DR strategy considering an EVs assisted NES model has been studied as an extension. The numerical results demonstrated that: for the single household network, 86.4% of the load can be shaped to a low level in peak demand hours and that the daily electricity cost can be reduced by 30.6% by using the EV-APS based DR strategy. For the multi-household network, the load can be significantly shaped to an appropriate level and the daily electricity cost of the entire network can be reduced by 34.9%. On the basis of the achieved results, it can be summarised that the proposed DR strategies are energy-efficient tools and can fulfill the tasks of load balancing and saving bills for the smart grid and DR participants simultaneously.

Chapter 6

Conclusions and Future Work

In this chapter the thesis is concluded in consideration of the instal aims and objectives of the work, highlighting the novelty and discussing specific outcomes; potential directions for future work are also discussed.

6.1 Conclusions

The smart grid is widely considered to be the next generation of the electricity grid in power system reform. Among a variety of advanced technologies in the smart grid, demand response (DR) has become an essential characteristic of the smart grid. In this thesis, a number of significant and interconnected challenges related to demand side management have been addressed. A comprehensive investigation of load pattern categorisation (LPC) based on the latest machine learning algorithms has been proposed and a number of the dynamic price (DP) based DR strategies have been designed.

Specifically, the challenge of clarifying the scheduling object has been resolved in Chapter 3. An innovative parametric bootstrap (PB) algorithm incorporated with K-means++, a comparatively robust clustering technique, has been proposed to address the cluster number determination problem as well as clustering the load data simultaneously, under a cascade clustering scheme to achieve LPC. A number of typical load patterns (TLPs) have been extracted as scheduling objects for demand side management. The evaluation results indicate that the proposed approach is more robust and reliable in finding an appropriate cluster number than conventional methods and it also has a better performance

in load data clustering compared with methods presented in literature. As the proposed PB algorithm is entirely unsupervised, it can be widely adopted to solve the cluster number determination problem for a variety of data processing tasks.

In Chapter 4, a hybrid model consisting of LS model, GP model, and BP-ANN model has been proposed, to forecast the day-ahead real-time prices as an input control signal in executing DR programs. The achieved forecasting results clearly demonstrated that the proposed approach is an accurate and efficient tool in RTP forecasting and it also significantly outperforms previously published forecasting models.

A number of DR strategies with EVs have been developed in Chapter 5, to jointly optimise household appliance scheduling and economic cost reduction based on DP for different scales of households. An EV-APS based DR strategy has been proposed first. Afterwards, a DR strategy considering an EVs assisted NES model has been studied as an extension. The numerical results demonstrated that for the single household network, 86.4% of the load can be shaped to a low level in peak demand hours and that daily electricity costs can be reduced by 30.6% by using the EV-APS based DR strategy. For the multi-household network, the load can be significantly shaped to an appropriate level and the daily electric cost of the entire network can be reduced by 34.9%. On the basis of the achieved results, it can be concluded that the proposed DR strategies are energy-efficient tools and can simultaneously fulfill the tasks of load balancing for the smart grid and bill reduction for DR participants.

In conclusion, this thesis extends around these three challenges related to demand side management in the smart grid. The results of LPC are used to support the DR management and the predicted RTP plays a role as an input control signal in load shifting. Meanwhile, the DR strategies are the core of DR management. According to the results presented in the thesis, it can be concluded that significant outcomes in DR management have been achieved.

6.2 Future Work

There are a number of significant issues worthy of further study in this field. Based on the existing achievements, future research can be focused on the following aspects:

- Consider the bias of data in data pre-processing of load pattern categorisation.

  With the development of the era, the electricity consumption pattern of today is different from the pattern of a decade ago. Therefore, in order to obtain more accurate load pattern results compared with the recent results, it is necessary to consider the bias of data into the data pre-processing based on the proposed algorithms.

- Propose an investigation of the relationship between the number of households in a neighbourhood and the number of EVs, and their effects on the DR management.

  In recent work, only a few households and EVs are considered in the network. It is interesting to investigate the relationship between the number of households in a neighbourhood and the number of electric vehicles, and how the numbers of households and EVs can affect DR management.

- Develop an EV charging strategy for a large quantity of EVs.

  With the population of EVs increasing, more and more EVs will be connected to the power grid in the future. EV charging accounts for a large proportion of total electricity consumption and it is also a severe issue for the power grid. Hence, an EV charging strategy for a large number of EVs is required.

- Involve electrical heating ventilation air conditioning (EHVAC) system in the DR problem.

  The EHVAC system is a particular home appliance within a household and typically accounts for a significant proportion of electricity usage, especially

during winter time. The operating time of the EHVAC system is affected by a number of factors, such as weather conditions, house structure, human activities, thermal dynamics, *etc.* Therefore, better and smarter control of EHVAC system in a DR program is beneficial for both customers and the smart grid.

Bibliography

[1] W. Saad, Z. Han, H. V. Poor, and T. Basar. Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications. *IEEE Signal Processing Magazine*, 29(5):86–105, Sep. 2012.

[2] J. S. Vardakas, N. Zorba, and C. V. Verikoukis. A survey on demand response programs in smart grids: Pricing methods and optimization algorithms. *IEEE Communications Surveys and Tutorials*, 17(1):152–178, Firstquarter 2015.

[3] L. Yao, W. H. Lim, and T. S. Tsai. A real-time charging scheme for demand response in electric vehicle parking station. *IEEE Transactions on Smart Grid*, 8(1):52–62, Jan. 2017.

[4] Q. Chen, F. Wang, B. M. Hodge, J. Zhang, Z. Li, M. Shafie-Khah, and J. P. S. Catalo. Dynamic price vector formation model-based automatic demand response strategy for pv-assisted ev charging stations. *IEEE Transactions on Smart Grid*, 8(6):2903–2915, Nov. 2017.

[5] A. Brooks, E. Lu, D. Reicher, C. Spirakis, and B. Weihl. Demand dispatch. *IEEE Power and Energy Magazine*, 8(3):20–29, May 2010.

[6] J. M. Guerrero, M. Chandorkar, T. Lee, and P. C. Loh. Advanced control architectures for intelligent microgridspart i: Decentralized and hierarchical control. *IEEE Transactions on Industrial Electronics*, 60(4):1254–1262, Apr. 2013.

[7] J. M. Guerrero, P. C. Loh, T. Lee, and M. Chandorkar. Advanced control architectures for intelligent microgridspart ii: Power quality, energy storage, and ac/dc microgrids. *IEEE Transactions on Industrial Electronics*, 60(4):1263–1270, Apr. 2013.

[8] X. Luo, X. Zhu, and E. G. Lim. Load scheduling based on an advanced real-time price forecasting model. In *Proc. 2015 IEEE International Conference on Ubiquitous Computing and Communications*, pages 1252–1257, Liverpool, UK, Oct. 2015.

[9] X. Luo, X. Zhu, and E. G. Lim. Dynamic pricing based and electric vehicle assisted demand response strategy. In *Proc. 2017 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 357–362, Dresden, Germany, Oct. 2017.

[10] C. O'Dwyer, R. Duignan, and M. O'Malley. Modeling demand response in the residential sector for the provision of reserves. In *2012 IEEE Power and Energy Society General Meeting*, pages 1–8, San Diego, USA, Jul. 2012.

[11] A. Safdarian, M. Fotuhi-Firuzabad, and M. Lehtonen. A distributed algorithm for managing residential demand response in smart grids. *IEEE Transactions on Industrial Informatics*, 10(4):2385–2393, Nov. 2014.

[12] N. G. Paterakis, A. Tascikaraoglu, O. Erdinc, A. G. Bakirtzis, and J. P. S. Catalao. Assessment of demand-response-driven load pattern elasticity using a combined approach for smart households. *IEEE Transactions on Industrial Informatics*, 12(4):1529–1539, Aug. 2016.

[13] C. Su and D. Kirschen. Quantifying the effect of demand response on electricity markets. *IEEE Transactions on Power Systems*, 24(3):1199–1207, Aug. 2009.

[14] A. J. Conejo, J. M. Morales, and L. Baringo. Real-time demand response model. *IEEE Transactions on Smart Grid*, 1(3):236–242, Dec. 2010.

[15] K. Mets, F. Depuydt, and C. Develder. Two-stage load pattern clustering using fast wavelet transformation. *IEEE Transactions on Smart Grid*, 7(5):2250–2259, Sep. 2016.

[16] M. Sun, I. Konstantelos, and G. Strbac. C-vine copula mixture model for clustering of residential electrical load pattern data. *IEEE Transactions on Power Systems*, 32(3):2382–2393, May 2017.

[17] F. L. Quilumba, W. J. Lee, H. Huang, D. Y. Wang, and R. L. Szabados. Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities. *IEEE Transactions on Smart Grid*, 6(2):911–918, Mar. 2015.

[18] K. Xing, C. Hu, J. Yu, X. Cheng, and F. Zhang. Mutual privacy preserving k-means clustering in social participatory sensing. *IEEE Transactions on Industrial Informatics*, 13(4):2066–2076, Aug. 2017.

[19] R. M. Esteves, T. Hacker, and C. Rong. Cluster analysis for the cloud: Parallel competitive fitness and parallel k-means++ for large dataset analysis. In *Proc. 4th IEEE International Conference on Cloud Computing Technology and Science Proceedings*, pages 177–184, Beijing, China, Dec. 2012.

[20] S. Agarwal, S. Yadav, and K. Singh. Notice of violation of ieee publication principles lt;br gt;k-means versus k-means ++ clustering technique. In *Proc. 2012 Students Conference on Engineering and Systems*, pages 1–6, Uttar Pradesh, India, Mar. 2012.

[21] A. Onan. A k-medoids based clustering scheme with an application to document clustering. In *Proc. 2017 International Conference on Computer Science and Engineering (UBMK)*, pages 354–359, Antalya, Turkey, Oct. 2017.

[22] A. Ganjavi, E. Christopher, C. M. Johnson, and J. Clare. A study on probability of distribution loads based on expectation maximization

algorithm. In *Proc. 2017 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5, Washington, USA, Apr. 2017.

[23] J. Zhang, Z. Yin, and R. Wang. Pattern classification of instantaneous cognitive task-load through gmm clustering, laplacian eigenmap, and ensemble svms. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 14(4):947–965, Jul. 2017.

[24] Xing Luo, Xu Zhu, and Eng Gee Lim. A parametric bootstrap algorithm for cluster number determination of load pattern categorization. *Energy*, 180:50–60, 2019.

[25] T. Jonsson, P. Pinson, H. A. Nielsen, H. Madsen, and T. S. Nielsen. Forecasting electricity spot prices accounting for wind power predictions. *IEEE Transactions on Sustainable Energy*, 4(1):210–218, Jan. 2013.

[26] M. Zhou, Z. Yan, Y. X. Ni, G. Li, and Y. Nie. Electricity price forecasting with confidence-interval estimation through an extended arima approach. *IEE Proceedings-Generation, Transmission and Distribution*, 153(2):187–195, Mar. 2006.

[27] Z. Miao, L. Fan, H. G. Aghamolki, and B. Zeng. Least squares estimation based sdp cuts for socp relaxation of ac opf. *IEEE Transactions on Automatic Control*, 63(1):241–248, Jan. 2018.

[28] H. Haiping, H. Shichao, C. Jiutian, and W. Ruchuan. Image hiding algorithm in discrete cosine transform domain based on grey prediction and grey relational analysis. *China Communications*, 10(7):57–70, Jul. 2013.

[29] P. Mandal, A. K. Srivastava, and J. W. Park. An effort to optimize similar days parameters for ann-based electricity price forecasting. *IEEE Transactions on Industry Applications*, 45(5):1888–1896, Sep. 2009.

[30] Xing Luo, Xu Zhu, and Eng Gee Lim. A hybrid model for short term real-time electricity price forecasting in smart grid. *Big Data Analytics*, 3(1):8, Oct. 2018.

[31] Jiayun Zhao, Sadik Kucuksari, Esfandyar Mazhari, and Young-Jun Son. Integrated analysis of high-penetration pv and phev with energy storage and demand response. *Applied Energy*, 112:35–51, Dec. 2013.

[32] C. Develder, N. Sadeghianpourhamami, M. Strobbe, and N. Refa. Quantifying flexibility in ev charging as dr potential: Analysis of two real-world data sets. In *Proc. 2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 600–605, Beijing, China, Nov. 2016.

[33] S. Althaher, P. Mancarella, and J. Mutale. Automated demand response from home energy management system under dynamic pricing and power and comfort constraints. *IEEE Transactions on Smart Grid*, 6(4):1874–1883, Jul. 2015.

[34] F. Rassaei, W. Soh, and K. Chua. Demand response for residential electric vehicles with random usage patterns in smart grids. *IEEE Transactions on Sustainable Energy*, 6(4):1367–1376, Oct. 2015.

[35] B. Sivaneasan, K. Nandha Kumar, K. T. Tan, and P. L. So. Preemptive demand response management for buildings. *IEEE Transactions on Sustainable Energy*, 6(2):346–356, Apr. 2015.

[36] Toshihiko Handa, Akihiro Oda, Tomokazu Tachikawa, Junichi Ichimura, Yuji Watanabe, and Hiroaki Nishi. Knives: A distributed demand side management system-integration with zigbee wireless sensor network and application. In *Proc. 2008 6th IEEE International Conference on Industrial Informatics*, pages 324–329, Daejeon, Korea, Jul. 2008.

[37] Y. Suhara, T. Nakabe, G. Mine, and H. Nishi. Distributed demand side management system for home energy management. In *Proc. IECON 2010-*

36th Annual Conference on IEEE Industrial Electronics Society, pages 2430–2435, Glendale, USA, Nov. 2010.

[38] H. Cao, C. Beckel, and T. Staake. Are domestic load profiles stable over time? an attempt to identify target households for demand side management campaigns. In *Proc. IECON 2013-39th Annual Conference of the IEEE Industrial Electronics Society*, pages 4733–4738, Vienna, Austria, Nov. 2013.

[39] J. Kwac, J. Flora, and R. Rajagopal. Household energy consumption segmentation using hourly data. *IEEE Transactions on Smart Grid*, 5(1):420–430, Jan. 2014.

[40] Y. Li and P.C. Flynn. Electricity deregulation, spot price patterns and demand-side management. *Energy*, 31(6):908–922, 2006. Electricity Market Reform and Deregulation.

[41] Jacopo Torriti, Mohamed G. Hassan, and Matthew Leach. Demand response experience in europe: Policies, programmes and implementation. *Energy*, 35(4):1575–1583, 2010. Demand Response Resources: the US and International Experience.

[42] Lorna A. Greening. Demand response resources: Who is responsible for implementation in a deregulated market? *Energy*, 35(4):1518–1525, 2010. Demand Response Resources: the US and International Experience.

[43] Jin-Ho Kim and Anastasia Shcherbakova. Common failures of demand response. *Energy*, 36(2):873–880, 2011.

[44] M. Chaouch. Clustering based improvement of nonparametric functional time series forecasting: Application to intra-day household-level load curves. *IEEE Transactions on Smart Grid*, 5(1):411–419, Jan. 2014.

[45] K. H. S. V. S. Nunna and S. Doolla. Responsive end-user-based demand side management in multimicrogrid environment. *IEEE Transactions on Industrial Informatics*, 10(2):1262–1272, May 2014.

[46] Fintan McLoughlin, Aidan Duffy, and Michael Conlon. A clustering approach to domestic electricity load profile characterisation using smart metering data. *Applied Energy*, 141(Supplement C):190–199, Mar. 2015.

[47] Y. Chen, P. B. Luh, C. Guan, Y. Zhao, L. D. Michel, M. A. Coolbeth, P. B. Friedland, and S. J. Rourke. Short-term load forecasting: Similar day-based wavelet neural networks. *IEEE Transactions on Power Systems*, 25(1):322–330, Feb. 2010.

[48] C. S. Chen, J. C. Hwang, and C. W. Huang. Application of load survey systems to proper tariff design. *IEEE Transactions on Power Systems*, 12(4):1746–1751, Nov. 1997.

[49] G. Chicco, R. Napoli, P. Postolache, M. Scutariu, and C. Toader. Customer characterization options for improving the tariff offer. *IEEE Transactions on Power Systems*, 18(1):381–387, Feb. 2003.

[50] D. Fischer, B. Stephen, A. Flunk, N. Kreifels, K. B. Lindberg, B. Wille-Haussmann, and E. H. Owens. Modeling the effects of variable tariffs on domestic electric load profiles by use of occupant behavior submodels. *IEEE Transactions on Smart Grid*, 8(6):2685–2693, Nov. 2017.

[51] J. Qin, W. Fu, H. Gao, and W. X. Zheng. Distributed k-means algorithm and fuzzy c-means algorithm for sensor networks based on multiagent consensus theory. *IEEE Transactions on Cybernetics*, 47(3):772–783, Mar. 2017.

[52] S. N. Sulaiman and N. A. Mat Isa. Adaptive fuzzy-k-means clustering algorithm for image segmentation. *IEEE Transactions on Consumer Electronics*, 56(4):2661–2668, Nov. 2010.

[53] J. Santarcangelo and X. P. Zhang. Dynamic time-alignment k-means kernel clustering for time sequence clustering. In *Proc. 2015 IEEE International Conference on Image Processing (ICIP)*, pages 2532–2536, Quebec, Canada, Sep. 2015.

[54] G. Chicco, R. Napoli, and F. Piglione. Comparisons among clustering techniques for electricity customer classification. *IEEE Transactions on Power Systems*, 21(2):933–940, May 2006.

[55] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise. In *Proc. 2nd International Conference on Knowledge Discovery and Data Mining*, KDD'96, pages 226–231, Portland, Oregon, 1996. AAAI Press.

[56] J. Shen, X. Hao, Z. Liang, Y. Liu, W. Wang, and L. Shao. Real-time superpixel segmentation by dbscan clustering algorithm. *IEEE Transactions on Image Processing*, 25(12):5933–5942, Dec. 2016.

[57] L. Yang, X. Guangqiang, L. Xiaomei, and L. Hua. Dependent function interval parameters training algorithm based on dbscan clustering. In *Proc. 31st Chinese Control Conference*, pages 7709–7712, Hefei, China, Jul. 2012.

[58] L. Zhang, S. Deng, and S. Li. Analysis of power consumer behavior based on the complementation of k-means and dbscan. In *Proc. 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2)*, pages 1–5, Beijing, China, Nov. 2017.

[59] S. Zhou, Z. Xu, and F. Liu. Method for determining the optimal number of clusters based on agglomerative hierarchical clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 28(12):3007–3017, Dec. 2017.

[60] Y. De Smet. An extension of promethee to divisive hierarchical multicriteria clustering. In *Proc. 2014 IEEE International Conference on Industrial Engineering and Engineering Management*, pages 555–558, Selangor, Malaysia, Dec. 2014.

[61] Ramon Granell, Colin J. Axon, and David C.H. Wallom. Clustering disaggregated load profiles using a dirichlet process mixture model. *Energy Conversion and Management*, 92:507–516, 2015.

[62] A. Tewari, M. J. Giering, and A. Raghunathan. Parametric characterization of multimodal distributions with non-gaussian modes. In *Proc. 2011 IEEE 11th International Conference on Data Mining Workshops*, pages 286–292, Washington, USA, Dec. 2011.

[63] S. V. Verdu, M. O. Garcia, C. Senabre, A. G. Marin, and F. J. G. Franco. Classification, filtering, and identification of electrical customer load patterns through the use of self-organizing maps. *IEEE Transactions on Power Systems*, 21(4):1672–1682, Nov. 2006.

[64] Jaewook Lee and Daewon Lee. An improved cluster labeling method for support vector clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3):461–464, Mar. 2005.

[65] H. Elaidi, Y. Elhaddar, Z. Benabbou, and H. Abbar. An idea of a clustering algorithm using support vector machines based on binary decision tree. In *Proc. 2018 International Conference on Intelligent Systems and Computer Vision (ISCV)*, pages 1–5, Fez, Morocco, Apr. 2018.

[66] Qiu-Huan Zhao, Ming-Hu Ha, Gui-Bing Peng, and X. Zhang. Support vector machine based on half-suppressed fuzzy c-means clustering. In *Proc. 2009 International Conference on Machine Learning and Cybernetics*, volume 2, pages 1236–1240, Baoding, China, Jul. 2009.

[67] W. Wu and M. Peng. A data mining approach combining k-means clustering with bagging neural network for short-term wind power forecasting. *IEEE Internet of Things Journal*, 4(4):979–986, Aug. 2017.

[68] M. Ghofrani, D. Carson, and M. Ghayekhloo. Hybrid clustering-time series-bayesian neural network short-term load forecasting method. In *Proc. 2016 North American Power Symposium (NAPS)*, pages 1–5, Denver, USA, Sep. 2016.

[69] Z. Jiang, R. Lin, F. Yang, and B. Wu. A fused load curve clustering

algorithm based on wavelet transform. *IEEE Transactions on Industrial Informatics*, 14(5):1856–1865, May 2018.

[70] M. Blum. On the central limit theorem for correlated random variables. *Proceedings of the IEEE*, 52(3):308–309, Mar. 1964.

[71] J. S. Malik, A. Hemani, J. N. Malik, B. Silmane, and N. D. Gohar. Revisiting central limit theorem: Accurate gaussian random number generation in vlsi. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 23(5):842–855, May 2015.

[72] D. Dehay, J. Leskow, and A. Napolitano. Central limit theorem in the functional approach. *IEEE Transactions on Signal Processing*, 61(16):4025–4037, Aug. 2013.

[73] N. Wu, H. Wang, and J. M. Kuang. Code-aided snr estimation based on expectation maximisation algorithm. *Electronics Letters*, 44(15):924–925, Jul. 2008.

[74] S. Z. Xia and H. W. Liu. Bayesian track-before-detect algorithm with target amplitude fluctuation based on expectation-maximisation estimation. *IET Radar, Sonar Navigation*, 6(8):719–728, Oct. 2012.

[75] H. Watanabe, S. Muramatsu, and H. Kikuchi. Interval calculation of em algorithm for gmm parameter estimation. In *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, pages 2686–2689, May 2010.

[76] G. Guo, L. Chen, Y. Ye, and Q. Jiang. Cluster validation method for determining the number of clusters in categorical sequences. *IEEE Transactions on Neural Networks and Learning Systems*, 28(12):2936–2948, Dec. 2017.

[77] R. Li, F. Li, and N. D. Smith. Multi-resolution load profile clustering for smart metering data. *IEEE Transactions on Power Systems*, 31(6):4473–4482, Nov. 2016.

[78] S. Bogucharskiy and V. Mashtalir. Image segmentation via x-means under overlapping classes. In *Proc. 2015 International Scientific and Technical Conference "Computer Sciences and Information Technologies" (CSIT)*, pages 45–47, Yerevan, Armenia, Sep. 2015.

[79] T. Erdelic, S. Vrbancic, and L. Rosic. A model of speed profiles for urban road networks using g-means clustering. In *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pages 1081–1086, May 2015.

[80] T. Zhang, G. Zhang, J. Lu, X. Feng, and W. Yang. A new index and classification approach for load pattern analysis of large electricity customers. *IEEE Transactions on Power Systems*, 27(1):153–160, Feb. 2012.

[81] I. P. Panapakidis, M. C. Alexiadis, and G. K. Papagiannis. Enhancing the clustering process in the category model load profiling. *IET Generation, Transmission Distribution*, 9(7):655–665, Apr. 2015.

[82] Dan Pelleg and Andrew Moore. X-means: Extending k-means with efficient estimation of the number of clusters. *Machine Learning*, Jan. 2002.

[83] B. Scheers, D. Teguig, and V. Le Nir. Modified andersoncdarling detector for spectrum sensing. *Electronics Letters*, 51(25):2156–2158, 2015.

[84] E. V. Garcia and J. E. Runnels. The utility perspective of spot pricing. *IEEE Power Engineering Review*, PER-5(6):45–45, Jun. 1985.

[85] H. Saele and O.S. Grande. Demand response from household customers: Experiences from a pilot study in norway. *IEEE Transactions on Smart Grid*, 2(1):102–109, Mar. 2011.

[86] A.G. Vlachos and P.N. Biskas. Simultaneous clearing of energy and reserves in multi-area markets under mixed pricing rules. *IEEE Transactions on Power Systems*, 26(4):2460–2471, Nov. 2011.

[87] Rongshan Yu, Wenxian Yang, and S. Rahardja. A statistical demand-price model with its application in optimal real-time price. *IEEE Transactions on Smart Grid*, 3(4):1734–1742, Dec. 2012.

[88] C. S. Chang and Minjun Yi. Real-time pricing related short-term load forecasting. In *Proc. 1998 International Conference on Energy Management and Power Delivery*, volume 2, pages 411–416 vol.2, Singapore, Mar. 1998.

[89] K. R. Nair, V. Vanitha, and M. Jisma. Forecasting of wind speed using ann, arima and hybrid models. In *Proc. 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, pages 170–175, Kannur, India, Jul. 2017.

[90] K. Qian, C. Zhou, M. Allan, and Y. Yuan. Modeling of load demand due to ev battery charging in distribution systems. *IEEE Transactions on Power Systems*, 26(2):802–810, May 2011.

[91] K.M. Tsui and S.C. Chan. Demand response optimization for smart home scheduling under real-time pricing. *IEEE Transactions on Smart Grid*, 3(4):1812–1821, Dec. 2012.

[92] D. H. Mazengia and Le Anh Tuan. Forecasting spot electricity market prices using time series models. In *Proc. 2008 IEEE International Conference on Sustainable Energy Technologies*, pages 1256–1261, Singapore, Nov. 2008.

[93] R. Tahmasebifar, M. K. Sheikh-El-Eslami, and R. Kheirollahi. Point and interval forecasting of real-time and day-ahead electricity prices by a novel hybrid approach. *IET Generation, Transmission Distribution*, 11(9):2173–2183, 2017.

[94] Rafal Weron. *Modelling Prices in Competitive Electricity Markets*. Wiley, 2006.

[95] Mariano Ventosa, Alvaro Baillo, Andres Ramos, and Michel Rivier. Electricity market modeling trends. *Energy Policy*, 33(7):897–913, 2005.

[96] Eric Guerci, Stefano Ivaldi, and Silvano Cincotti. Learning Agents in an Artificial Power Exchange: Tacit Collusion, Market Power and Efficiency of Two Double-auction Mechanisms. *Computational Economics*, 32(1):73–98, Sep. 2008.

[97] C. Batlle and J. Barquin. A strategic production costing model for electricity market price analysis. *IEEE Transactions on Power Systems*, 20(1):67–74, Feb. 2005.

[98] C. M. Ruibal and M. Mazumdar. Forecasting the mean and the variance of electricity prices in deregulated markets. *IEEE Transactions on Power Systems*, 23(1):25–32, Feb. 2008.

[99] B. Burger, M. Graeber and G. Schindlmayr. *Managing Energy Risk: An Integrated View on Power and Other Energy Markets*. Wiley, 2007.

[100] Krzysztof Eydeland, Alexander; Wolyniec. *Energy and Power Risk Management: New Developments in Modeling, Pricing, and Hedging*. Wiley, 2003.

[101] Rene Carmona and Michael Coulon. A survey of commodity markets and structural models for electricity prices. 2012.

[102] Fred Espen Benth, Rudiger Kiesel, and Anna Nazarova. A critical empirical study of three electricity spot price models. *Energy Economics*, 34(5):1589–1616, 2012.

[103] Rafal Weron. Market price of risk implied by asian-style electricity options and futures. *Energy Economics*, 30(3):1098–1115, 2008.

[104] S. Deng. Pricing electricity derivatives under alternative stochastic spot price models. In *Proc. 33rd Annual Hawaii International Conference on System Sciences*, pages 10–15, Hawaii, USA, Jan. 2000.

[105] Joanna Janczura and Rafal Weron. An empirical comparison of alternate

regime-switching models for electricity spot prices. *Energy Economics*, 32(5):1059–1073, 2010.

[106] Rafal Weron. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30(4):1030–1081, 2014.

[107] Nektaria V. Karakatsani and Derek W. Bunn. Forecasting electricity prices: The impact of fundamentals and time-varying coefficients. *International Journal of Forecasting*, 24(4):764–785, 2008. Energy Forecasting.

[108] A. A. El Desouky and M. M. Elkateb. Hybrid adaptive techniques for electric-load forecast using ann and arima. *IEE Proceedings-Generation, Transmission and Distribution*, 147(4):213–217, Jul. 2000.

[109] M. Coulon and A. Swami. Least squares detection of multiple changes in fractional arima processes. In *Proc. 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 5, pages 3177–3180 vol.5, Salt Lake City, USA, May 2001.

[110] N. Sadek, A. Khotanzad, and T. Chen. Atm dynamic bandwidth allocation using f-arima prediction model. In *Proc. 12th International Conference on Computer Communications and Networks*, pages 359–363, Dallas, USA, Oct. 2003.

[111] Y. Yang, C. Liu, and F. Guo. Forecasting method of aero-material consumption rate based on seasonal arima model. In *Proc. 2017 3rd IEEE International Conference on Computer and Communications (ICCC)*, pages 2899–2903, Chengdu, China, Dec. 2017.

[112] M. Xie, C. Sandels, K. Zhu, and L. Nordstrom. A seasonal arima model with exogenous variables for elspot electricity prices in sweden. In *Proc. 2013 10th International Conference on the European Energy Market (EEM)*, pages 1–4, Stockholm, Sweden, May 2013.

[113] Z. Xinxiang, Z. Bo, and F. Huijuan. A comparison study of outpatient visits forecasting effect between arima with seasonal index and sarima. In *Proc. 2017 International Conference on Progress in Informatics and Computing (PIC)*, pages 362–366, Nanjing, China, Dec. 2017.

[114] Z. Zhao, X. Wang, J. Qiao, and H. Sun. Wind speed prediction based on improved self excitation threshold auto regressive model. In *Proc. 2018 37th Chinese Control Conference (CCC)*, pages 1498–1503, Wuhan, China, Jul. 2018.

[115] H. Chen, F. Li, Q. Wan, and Y. Wang. Short term load forecasting using regime-switching garch models. In *2011 IEEE Power and Energy Society General Meeting*, pages 1–6, Detroit, USA, Jul. 2011.

[116] G. Rasool, N. Bouaynaya, and K. Iqbal. Muscle activity detection from myoelectric signals based on the ar-garch model. In *2012 IEEE Statistical Signal Processing Workshop (SSP)*, pages 420–423, Aug. 2012.

[117] N. Amjady. Day-ahead price forecasting of electricity markets by a new fuzzy neural network. *IEEE Transactions on Power Systems*, 21(2):887–896, May 2006.

[118] Dogan Keles, Jonathan Scelle, Florentina Paraschiv, and Wolf Fichtner. Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks. *Applied Energy*, 162:218–230, 2016.

[119] C. P. Rodriguez and G. J. Anders. Energy price forecasting in the ontario competitive power system market. *IEEE Transactions on Power Systems*, 19(1):366–374, Feb. 2004.

[120] L. Yang, M. He, J. Zhang, and V. Vittal. Support-vector-machine-enhanced markov model for short-term wind power forecast. *IEEE Transactions on Sustainable Energy*, 6(3):791–799, Jul. 2015.

[121] E. E. Elattar, J. Goulermas, and Q. H. Wu. Electric load forecasting based on locally weighted support vector regression. *IEEE Transactions*

on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40(4):438–447, Jul. 2010.

[122] J. Shi, W. Lee, Y. Liu, Y. Yang, and P. Wang. Forecasting power output of photovoltaic systems based on weather classification and support vector machines. *IEEE Transactions on Industry Applications*, 48(3):1064–1069, May 2012.

[123] Bo-Juen Chen, Ming-Wei Chang, and Chih-Jen lin. Load forecasting using support vector machines: a study on eunite competition 2001. *IEEE Transactions on Power Systems*, 19(4):1821–1830, Nov. 2004.

[124] P. L. Langbein. Demand response participation in pjm wholesale markets. In *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)*, pages 1–3, Jan. 2012.

[125] S. N. Singh and J. Ostergaard. Use of demand response in electricity markets: An overview and key issues. In *Proc. 2010 7th International Conference on the European Energy Market*, pages 1–6, Madrid, Spain, Jun. 2010.

[126] M. H. Albadi and E. F. El-Saadany. Demand response in electricity markets: An overview. In *2007 IEEE Power Engineering Society General Meeting*, pages 1–5, Tampa, USA, Jun. 2007.

[127] Galen Barbose, Charles Goldman, and Bernie Neenan. A survey of utility experience with real time pricing. *Lawrence Berkeley National Laboratory*, Dec. 2004.

[128] S. Braithwait and K. Eakin. The role of demand response in electric power market design. *Prepared for Edison Electric Institute, Madison*, Oct. 2002.

[129] M.H. Albadi and E.F. El-Saadany. A summary of demand response in electricity markets. *Electric Power Systems Research*, 78(11):1989–1996, 2008.

[130] Torgeir Ericson. Direct load control of residential water heaters. *Energy Policy*, 37(9):3502–3512, 2009.

[131] H.A. Aalami, M. Parsa Moghaddam, and G.R. Yousefi. Demand response modeling considering interruptible/curtailable loads and capacity market programs. *Applied Energy*, 87(1):243–250, 2010.

[132] J. Saebi, H. Taheri, J. Mohammadi, and S. S. Nayer. Demand bidding/buyback modeling and its impact on market clearing price. In *Proc. 2010 IEEE International Energy Conference*, pages 791–796, Dec. 2010.

[133] R. Tyagi and J. W. Black. Emergency demand response for distribution system contingencies. In *IEEE PES T and D 2010*, pages 1–4, Apr. 2010.

[134] S. Datchanamoorthy, S. Kumar, Y. Ozturk, and G. Lee. Optimal time-of-use pricing for residential load control. In *Proc. 2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 375–380, Brussels, Belgium, Oct. 2011.

[135] Karen Herter. Residential implementation of critical-peak pricing of electricity. *Energy Policy*, 35(4):2121–2130, 2007.

[136] Q. Pang, P. Su, and B. Sun. Real-time price based home appliances intelligent control. In *Proc. 2012 Third International Conference on Digital Manufacturing Automation*, pages 634–637, Guilin, China, Jul. 2012.

[137] A. S. M. A. Mahmud and P. Sant. Real-time price savings through price suggestions for the smart grid demand response model. In *2017 5th International Istanbul Smart Grid and Cities Congress and Fair (ICSG)*, pages 65–69, Apr. 2017.

[138] Severin Borenstein. Equity effects of increasing-block electricity pricing. *Center for the Study of Energy Markets*, 2008.

[139] S. Pal and R. K. Malviya. Dsm scheduling mechanism with pricing schemes using integer linear programming. In *Proc. 2015 39th National Systems Conference (NSC)*, pages 1–6, Greater Noida, India, Dec. 2015.

[140] N. Li, L. Chen, and S. H. Low. Optimal demand response based on utility maximization in power networks. In *2011 IEEE Power and Energy Society General Meeting*, pages 1–8, Jul. 2011.

[141] P. Samadi, A. Mohsenian-Rad, R. Schober, V. W. S. Wong, and J. Jatskevich. Optimal real-time pricing algorithm based on utility maximization for smart grid. In *Proc. 2010 First IEEE International Conference on Smart Grid Communications*, pages 415–420, Gaithersburg, USA, Oct. 2010.

[142] N. Loganathan and K. Lakshmi. Demand side energy management system using ann based linear programming approach. In *Proc. 2014 IEEE International Conference on Computational Intelligence and Computing Research*, pages 1–5, Tamil Nadu, India, Dec. 2014.

[143] J. Alende, Y. Li, and M. Cantoni. A 0, 1 linear program for fixed-profile load scheduling and demand management in automated irrigation channels. In *Proc. 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pages 597–602, Beijing, China, Dec. 2009.

[144] S. Wang, S. Bi, and Y. A. Zhang. Demand response management for profit maximizing energy loads in real-time electricity market. *IEEE Transactions on Power Systems*, 33(6):6387–6396, Nov. 2018.

[145] Z. M. Fadlullah, D. M. Quan, N. Kato, and I. Stojmenovic. Gtes: An optimized game-theoretic demand-side management scheme for smart grid. *IEEE Systems Journal*, 8(2):588–597, Jun. 2014.

[146] E. Nekouei, T. Alpcan, and D. Chattopadhyay. Game-theoretic frameworks

for demand response in electricity markets. *IEEE Transactions on Smart Grid*, 6(2):748–758, Mar. 2015.

[147] N. H. Tran, D. H. Tran, S. Ren, Z. Han, E. Huh, and C. S. Hong. How geo-distributed data centers do demand response: A game-theoretic approach. *IEEE Transactions on Smart Grid*, 7(2):937–947, Mar. 2016.

[148] B. Chai, J. Chen, Z. Yang, and Y. Zhang. Demand response management with multiple utility companies: A two-level game approach. *IEEE Transactions on Smart Grid*, 5(2):722–731, Mar. 2014.

[149] L. Jiang and S. Low. Real-time demand response with uncertain renewable energy in smart grid. In *Proc. 2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1334–1341, Monticello, USA, Sep. 2011.

[150] Q. Huang, M. Roozbehani, and M. A. Dahleh. Efficiency-risk tradeoffs in electricity markets with dynamic demand response. *IEEE Transactions on Smart Grid*, 6(1):279–290, Jan. 2015.

[151] L. M. Costa and G. Kariniotakis. A stochastic dynamic programming model for optimal use of local energy resources in a market environment. In *2007 IEEE Lausanne Power Tech*, pages 449–454, Jul. 2007.

[152] S. Kishore and L. V. Snyder. Control mechanisms for residential electricity demand in smartgrids. In *Proc. 2010 First IEEE International Conference on Smart Grid Communications*, pages 443–448, Gaithersburg, USA, Oct. 2010.

[153] R. N. Anderson, A. Boulanger, W. B. Powell, and W. Scott. Adaptive stochastic control for the smart grid. *Proceedings of the IEEE*, 99(6):1098–1115, Jun. 2011.

[154] T. T. Kim and H. V. Poor. Scheduling power consumption with price uncertainty. *IEEE Transactions on Smart Grid*, 2(3):519–527, Sep. 2011.

[155] N. Al Khafaf, M. Jalili, and P. Sokolowski. Demand response planning tool using markov decision process. In *Proc. 2018 IEEE 16th International Conference on Industrial Informatics (INDIN)*, pages 484–489, Porto, Portugal, Jul. 2018.

[156] F. Ruelens, B. J. Claessens, R. Belmans, and G. Deconinck. Sequential decision-making strategy for a demand response aggregator in a two-settlement electricity market. In *Proc. 2016 European Control Conference (ECC)*, pages 1229–1235, Aalborg, Denmark, Jun. 2016.

[157] D. O'Neill, M. Levorato, A. Goldsmith, and U. Mitra. Residential demand response using reinforcement learning. In *Proc. 2010 1st IEEE International Conference on Smart Grid Communications*, pages 409–414, Gaithersburg, USA, Oct. 2010.

[158] X. Liu. Economic load dispatch constrained by wind power availability: A wait-and-see approach. *IEEE Transactions on Smart Grid*, 1(3):347–355, Dec. 2010.

[159] M. Parvania and M. Fotuhi-Firuzabad. Demand response scheduling by stochastic scuc. *IEEE Transactions on Smart Grid*, 1(1):89–98, Jun. 2010.

[160] J. Zhang, J. D. Fuller, and S. Elhedhli. A stochastic programming model for a day-ahead electricity market with real-time reserve shortage pricing. *IEEE Transactions on Power Systems*, 25(2):703–713, May 2010.

[161] J. Kennedy and R. Eberhart. Particle swarm optimization. In *Proc. ICNN'95-International Conference on Neural Networks*, volume 4, pages 1942–1948 vol.4, Los Angeles, USA, Nov. 1995.

[162] Y. Shi and R. Eberhart. A modified particle swarm optimizer. In *Proc. 1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence*, pages 69–73, Anchorage, Alaska, May 1998.

[163] J. Kennedy. The particle swarm: social adaptation of knowledge. In *Proc. 1997 IEEE International Conference on Evolutionary Computation (ICEC '97)*, pages 303–308, Indianapolis, USA, Apr. 1997.

[164] P. Faria, Z. A. Vale, J. Soares, and J. Ferreira. Particle swarm optimization applied to integrated demand response resources scheduling. In *2011 IEEE Symposium on Computational Intelligence Applications In Smart Grid (CIASG)*, pages 1–8, Apr. 2011.

[165] M. A. A. Pedrasa, T. D. Spooner, and I. F. MacGill. Coordinated scheduling of residential distributed energy resources to optimize smart home energy services. *IEEE Transactions on Smart Grid*, 1(2):134–143, Sep. 2010.

[166] N. Kinhekar, N. P. Padhy, and H. O. Gupta. Particle swarm optimization based demand response for residential consumers. In *2015 IEEE Power Energy Society General Meeting*, pages 1–5, Jul. 2015.

[167] Remani T., Jasmin E.A., and Imthias Ahamed T.P. Load scheduling with maximum demand using binary particle swarm optimization. In *Proc. 2015 International Conference on Technological Advancements in Power and Energy (TAP Energy)*, pages 294–298, Kollam, India, Jun. 2015.

[168] S. A. Azad, A. M. T. Oo, and M. F. Islam. A low complexity residential demand response strategy using binary particle swarm optimization. In *Proc. 2012 22nd Australasian Universities Power Engineering Conference (AUPEC)*, pages 1–6, Australia, Sep. 2012.

[169] M. Ansari, A. T. Al-Awami, E. Sortomme, and M. A. Abido. Coordinated bidding of ancillary services for vehicle-to-grid using fuzzy optimization. *IEEE Transactions on Smart Grid*, 6(1):261–270, Jan. 2015.

[170] M. Yilmaz and P. T. Krein. Review of the impact of vehicle-to-grid technologies on distribution systems and utility interfaces. *IEEE Transactions on Power Electronics*, 28(12):5673–5689, Dec. 2013.

[171] S. Shao, M. Pipattanasomporn, and S. Rahman. Grid integration of electric vehicles and demand response with customer choice. *IEEE Transactions on Smart Grid*, 3(1):543–550, Mar. 2012.

[172] S. G. Yoon, Y. J. Choi, J. K. Park, and S. Bahk. Stackelberg-game-based demand response for at-home electric vehicle charging. *IEEE Transactions on Vehicular Technology*, 65(6):4172–4184, Jun. 2016.

[173] Cal Alumni Association. Supercharging more electric cars risks crashing the grid. Berkeley, USA, 2014.

[174] M. Shafie-khah, E. Heydarian-Forushani, G. J. Osorio, F. A. S. Gil, J. Aghaei, M. Barani, and J. P. S. Catalao. Optimal behavior of electric vehicle parking lots as demand response aggregation agents. *IEEE Transactions on Smart Grid*, 7(6):2654–2665, Nov. 2016.

[175] R. Yu, W. Zhong, S. Xie, C. Yuen, S. Gjessing, and Y. Zhang. Balancing power demand through ev mobility in vehicle-to-grid mobile energy networks. *IEEE Transactions on Industrial Informatics*, 12(1):79–90, Feb. 2016.

[176] J. C. Ferreira, V. Monteiro, and J. L. Afonso. Vehicle-to-anything application (v2anything app) for electric vehicles. *IEEE Transactions on Industrial Informatics*, 10(3):1927–1937, Aug. 2014.

[177] N. G. Paterakis, O. Erdinc, A. G. Bakirtzis, and J. P. S. Catalao. Optimal household appliances scheduling under day-ahead pricing and load-shaping demand response strategies. *IEEE Transactions on Industrial Informatics*, 11(6):1509–1519, Dec. 2015.

[178] Ronald L. Wasserstein and Nicole A. Lazar. The asa's statement on p-values: Context, process, and purpose. *The American Statistician*, 70(2):129–133, Mar. 2016.

[179] Regina Nuzzo. Scientific method: Statistical errors. *Nature*, 506:150–2, Feb. 2014.

[180] Y. Xu, W. Qu, Z. Li, G. Min, K. Li, and Z. Liu. Efficient k-means++ approximation with mapreduce. *IEEE Transactions on Parallel and Distributed Systems*, 25(12):3135–3144, Dec. 2014.

[181] Prihandoko, Bertalya, and M. I. Ramadhan. An analysis of natural disaster data by using k-means and k-medoids algorithm of data mining techniques. In *Proc. 2017 15th International Conference on Quality in Research (QiR): International Symposium on Electrical and Computer Engineering*, pages 221–225, Jul. 2017.

[182] J. S. Malik, A. Hemani, J. N. Malik, B. Silmane, and N. D. Gohar. Revisiting central limit theorem: Accurate gaussian random number generation in vlsi. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 23(5):842–855, May 2015.

[183] I. Cabria and I. Gondra. Potential- k- means for load balancing and cost minimization in mobile recycling network. *IEEE Systems Journal*, 11(1):242–249, Mar. 2017.

[184] T. S. Xu, H. D. Chiang, G. Y. Liu, and C. W. Tan. Hierarchical k-means method for clustering large-scale advanced metering infrastructure data. *IEEE Transactions on Power Delivery*, 32(2):609–616, Apr. 2017.

[185] C. Boutsidis and M. Magdon-Ismail. Deterministic feature selection for k-means clustering. *IEEE Transactions on Information Theory*, 59(9):6099–6110, Sep. 2013.

[186] [online] available: https://www.nationalgrid.com/uk.

[187] [online] available: http://www.aemo.org.au/. In *AEMO-Austrilian Energy Market Operator*.

[188] M. Kesaniemi and K. Virtanen. Direct least square fitting of hyperellipsoids. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(1):63–76, Jan. 2018.

[189] B. Renczes, I. Kollar, and T. Daboczi. Efficient implementation of least squares sine fitting algorithms. *IEEE Transactions on Instrumentation and Measurement*, 65(12):2717–2724, Dec. 2016.

[190] Tang Guilin and Qiu Yunming. Improved least square method apply in ship performance analysis. In *Proc. 2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*, volume 5, pages 594–596, Chengdu, China, Aug. 2010.

[191] S. H. Yang, W. J. Huang, J. F. Tsai, and Y. P. Chen. Symbiotic structure learning algorithm for feedforward neural-network-aided grey model and prediction applications. *IEEE Access*, 5:9378–9388, 2017.

[192] Y. Wang, Q. Liu, J. Tang, W. Cao, and X. Li. Optimization approach of background value and initial item for improving prediction precision of gm(1,1) model. *Journal of Systems Engineering and Electronics*, 25(1):77–82, Feb. 2014.

[193] J. S. Lee and Y. C. Lee. An application of grey prediction to transmission power control in mobile sensor networks. *IEEE Internet of Things Journal*, 5(3):2154–2162, Jun. 2018.

[194] Peng Hui, Wenqi Qu, Junjian Tang, and Jin Chen. Traffic indexes prediction based on grey prediction model. In *Proc. 2013 Sixth International Symposium on Computational Intelligence and Design (ISCID)*, volume 1, pages 244–247, Oct. 2013.

[195] Na Xu and Xin-Rui Zhang. Traffic volume prediction based on improved grey self-adaptable prediction formula. volume 2, pages 1027–1030, Jul. 2010.

[196] Mehdi Khashei and Mehdi Bijari. An artificial neural network (p,d,q) model for timeseries forecasting. *Expert Systems with Applications*, 37(1):479–489, 2010.

[197] L. D. Zhang, L. Jia, and W. X. Zhu. Overview of traffic flow hybrid ann forecasting algorithm study. In *Proc. 2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*, volume 1, pages V1–615–V1–619, Taiyuan, China, Oct. 2010.

[198] H. Zhao, L. Peng, T. Takahashi, T. Hayashi, K. Shimizu, and T. Yamamoto. Ann based data integration for multi-path ultrasonic flowmeter. *IEEE Sensors Journal*, 14(2):362–370, Feb. 2014.

[199] T. Lian, M. Xie, J. Xu, L. Chen, and H. Gao. Modified bp neural network model is used for oddeven discrimination of integer number. In *Proc. 2013 International Conference on Optoelectronics and Microelectronics (ICOM)*, pages 67–70, Harbin, China, Sep. 2013.

[200] D. Zhu, T. Zhang, and G. Mao. Back-propagation artificial neural networks for water supply pipeline model. *Tsinghua Science and Technology*, 7(5):527–531, Oct. 2002.

[201] K. Methaprayoon, C. Yingvivatanapong, W. J. Lee, and J. R. Liao. An integration of ann wind power estimation into unit commitment considering the forecasting uncertainty. *IEEE Transactions on Industry Applications*, 43(6):1441–1448, Nov. 2007.

[202] Xin Zhao, Shuangxin Wang, and Tao Li. Review of evaluation criteria and main methods of wind power forecasting. *Energy Procedia*, 12:761–769, 2011. Proc. International Conference on Smart Grid and Clean Energy Technologies.

[203] Nijat Mehdiyev, David Enke, Peter Fettke, and Peter Loos. Evaluating forecasting methods by considering different accuracy measures. *Procedia Computer Science*, 95:264–271, Nov. 2016.

[204] Mauno Vihinen. How to evaluate performance of prediction methods? measures and their interpretation in variation effect analysis. *BMC Genomics*, 13(4):S2, Jun. 2012.

[205] S. Shao, M. Pipattanasomporn, and S. Rahman. Demand response as a load shaping tool in an intelligent grid with electric vehicles. *IEEE Transactions on Smart Grid*, 2(4):624–631, Dec. 2011.

[206] [online] available: https://www.tesla.com.

[207] [online] available: http://www.bydauto.com.cn.

[208] K. L. Lo and Y. K. Wu. Analysis of relationships between hourly electricity price and load in deregulated real-time power markets. *IEE Proceedings-Generation, Transmission and Distribution*, 151(4):441–452, Jul. 2004.

[209] T. Blank, J. Badeda, J. Kowal, and D. U. Sauer. Deep discharge behavior of lead-acid batteries and modeling of stationary battery energy storage systems. In *Intelec 2012*, pages 1–4, Sep. 2012.

[210] [online] available: http://www.britishgas.co.uk.