A smart hierarchical transactive energy system in the presence of renewable energies, and demand-side management

Thesis by
Qiuyi Hong

A Thesis Submitted in Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in
Data Science

SCHOOL OF MATHEMATICS, STATISTICS AND ACTUARIAL SCIENCE
UNIVERSITY OF ESSEX

June 2024

Qiuyi Hong
ORCID: 0000-0001-5345-651X

ACKNOWLEDGEMENTS

I would like to extend my deepest gratitude to my PhD supervisors, Dr Fanlin Meng and Dr Felipe Maldonado, for their invaluable guidance and support throughout my PhD journey. Their expertise and insightful opinions have significantly shaped my research and played a crucial role in my academic development.

I am also immensely grateful to my parents, whose unwavering support and encouragement have been a cornerstone for me throughout this journey. Their belief in my abilities and their sacrifices have enabled me to pursue my academic goals and have been the source of my strength and motivation.

Thank you all for your enormous contributions to my PhD journey.

ABSTRACT

With the constant development of energy systems, multi-energy networks are becoming increasingly popular, which integrates renewable energies and demand-side management. This presents a significant challenge for developing smart energy management frameworks. Furthermore, unlike traditional energy systems, the transaction energy system is dynamic and complex, which is enriched by the interdependence among multiple energies, the uncertainty of renewable energies and the complexity of demand-side management. Therefore, advanced modelling techniques and solution methods are required to be developed to overcome the difficulties.

This thesis addresses the intricate challenges of developing a smart hierarchical transactive energy system that seamlessly marries multi-energy sources, renewable energy integration, and sophisticated demand-side management strategies. The research unfolds in three pivotal research topics: 1) The formulation and analysis of a game-theoretic decision-making model for energy retailers' strategic bidding and offering in both wholesale and local energy markets while considering customers' switching behaviour; 2) The introduction of a customised multi-energy pricing scheme, which is formulated as a bilevel optimisation model. The proposed model not only maximises the profit of energy retailers but also considers the multi-energy interdependencies and the diverse characteristics of microgrids; 3) The development of an innovative forecasting model named Patchformer, based on Transformer-based architectures and patch embedding method, for the prediction of long-term multi-energy loads. This model improves forecasting accuracy, which enables a more reliable and efficient energy system management by predicting energy demands with high precision.

This work presents a comprehensive approach to improving the effectiveness of transactive energy systems by merging advanced modelling techniques and machine/deep learning models. This thesis tackles the current challenges in the field of transaction energy systems while also providing information for future research that aims to unlock the full potential of smart energy management in smart grids.

PUBLISHED CONTENT AND CONTRIBUTIONS

The contributions made in this thesis are listed in the following peer-reviewed publications and in the preparation paper. Chapters 3-5 are associated either with published work or work in preparation for journal submission, as follows:

Chapter 3: Q. Hong, F. Meng, J. Liu, and R. Bo, "A bilevel game-theoretic decision-making framework for strategic retailers in both local and wholesale electricity markets," Applied Energy, vol. 330, p. 120311, 2023.

Chapter 4:

- Q. Hong, F. Meng, "Customized Multi-Energy Pricing in Smart Grids: A Bilevel and Evolutionary Computation Approach," The 21st UK Workshop on Computational Intelligence, Springer, 2022.
- Q. Hong, F. Meng, and J. Liu, "Customised multi-energy pricing: Model and solutions," Energies, vol. 16, no. 4, p. 2080, 2023.

Chapter 5: Q. Hong, F. Meng, and F. Maldonado, "Advancing Long-Term Multi-Energy Load Forecasting with Patchformer: A Patch and Transformer-Based Approach." *[In preparation for journal submission]*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

Abbreviations and Indices

DAW	Day-ahead Wholesale.
LPE	Local Power Exchange.
ISO	Independent System Operator.
ESS	Energy Storage System.
DR	Demand Response.
KKT	Karush-Kuhn-Tucher.
MPEC	Mathematical Programming with Equilibrium Constraints.
EPEC	Equilibrium Problems with Equilibrium Constraints.
MIQP	Mixed-Integer Quadratic Programming.
RES	Renewable energy sources.
PSO	Particle swarm optimisation.
GA	Genetic algorithm.
SA	Simulated annealing.
MILP	Mixed-integer linear program.
DERs	Distributed energy resources.
RTP	Real-time pricing.
TOU	Time-of-use.
EVs	Electric vehicles.
CVaR	Conditional value at risk.
IESP	Integrated energy service provider.
CHP	Combined heat and power.
ES,TS	Electrical storage, thermal storage.
LC,LS	Load curtailment, load shifting.
PV,WT	Photovoltaic, wind turbine.

LTTSF	Long-term time-series forecasting.
GHG	Greenhouse gas.
IMES	Integrated multi-energy systems.
ARIMA	Autoregressive integrated moving-average.
NLP	Natural language processing.
CV	Computer vision.
RNN	Recurrent neural networks.
LSTM	Long short-term memory.
GRU	Gated recurrent units.
FFNs	Feed forward networks.
MSE	Mean squared error.
MAE	Mean absolute error.
k	Index of the strategic retailer.
n	Index of all retailers.
g	Index of generators.
t	Index of time periods.
m	Index of the Microgrids.
a	Index of households participating in load shifting program.
c	Index of channels.
h	Index of attention heads.
l	Index of layers in the encoder and decoder.

Sets

$\mathcal{N}$	Set of retailers in the grid.
$\mathcal{G}$	Set of generators in the grid.
$\mathcal{T}$	Set of scheduling hours.
$\mathcal{M}$	Set of microgrids.
$\mathcal{A}$	Set of households participating in load shifting program.

Symbols

$\mathcal{X}, \mathcal{Y}$	Past and final prediction time sequence.
I, O	Past and prediction time sequence length.
C	Total number of channels.
N, M	Total number of layers in the encoder and decoder.
$\mathcal{X}^c$	Past time sequence for channel c.
Z	Total number of patches.
P	Patch length.
$\mathcal{P}^c_{patch}$	Two dimensional matrix for channel c with shape: $Z \times P$.
$W_{valEmbed}$	Learnable weight for value embedding.
$\mathcal{P}^c$	Patch embedded output.
$\mathcal{Q}^c_h, \mathcal{K}^c_h, \mathcal{V}^c_h$	Query, key and value for channel c, attention head h.
W^Q, W^K, W^V, W^O	Learnable weights for query, key, value and multi-head attention concatenation.
$\mathcal{P}^c_{attn}$	Multi-head attention score.
H	Total number of attention heads.
$\mathcal{P}^c_{norm}$	Value of layer normalisation.
μ^c, σ^c	Mean and standard deviation for channel c.
$\mathcal{P}^c_{ffn}$	Value of FFNs.
W^{ffn}_1, W^{ffn}_2	Learnable weights of the two FFNs.
b_1, b_2	Bias terms of the two FFNs.
$\mathcal{P}^{c,l}_{en,1}, \mathcal{P}^{c,l}_{en,2}$	Outputs of the first and the second layer normalisation blocks for channel c, l-th encoder layer in the encoder.
$\mathcal{X}^c_{de}$	Input of decoder for channel c.
$\mathcal{P}^{c,l}_{de,1}, \mathcal{P}^{c,l}_{de,2}, \mathcal{P}^{c,l}_{de,3}$	Outputs of the first, second and third layer normalisation blocks for channel c, l-th encoder layer in the decoder.
W_y	Learnable weights for linear transformation.
$\mathcal{Y}^c$	Final prediction time sequence for channel c.

Parameters

$\omega_{n,n}^{t}$	Self-elasticity of the retailer n at time t.
$\omega_{n,j}, \forall j \neq n$	Switching coefficient among retailers.
Δt	Duration of each time period.
ϵ_k	Self loss of the ESS of the retailer k.
c_k	Operation and maintenance cost of the retailer k.
c_g	The production cost of the generator g.
$\pi_i^{retail,t}$	The electricity price the retailer i sold to customers at time t.
$\pi_i^{bid,t}$	The electricity price the retailer i bought from the DAW market at time t.
$\pi_i^{LPE,t}$	The electricity price the retailer i bought from the LPE market at time t.
$\pi_k^{retail,min}$	Minimum retail price of the retailer k.
$\pi_k^{retail,max}$	Maximum retail price of the retailer k.
$\pi_k^{bid,min}$	Minimum bid price of the retailer k.
$\pi_k^{bid,max}$	Maximum bid price of the retailer k.
E_k^{min}, E_k^{max}	Minimum and maximum energy level of ESS of the retailer k.
$p_k^{c,min}, p_k^{c,max}$	Minimum and maximum charging power of ESS of the retailer k.
$p_k^{d,min}, p_k^{d,max}$	Minimum and maximum discharging power of ESS of the retailer k.
$q_{n,out}^{LPE,max,t}$	Maximum electricity volume the retailer n sold to the other retailers in LPE market.
$q_{n,in}^{LPE,max,t}$	Maximum electricity volume the retailer n bought from other retailers in LPE market.
q_g^{min}, q_g^{max}	Minimum and maximum electricity volume that the generator g sold to the DAW market.
$q_n^{bid,min,t}$	Minimum electricity volume that the retailer n bought from the DAW market.
$q_n^{bid,max,t}$	Maximum electricity volume that the retailer n bought from the DAW market.

α_m Proportion that the electricity price sold by the microgrid m back to the retailer against the retail price.

c_m^{CHP}, c_m^{pump} Operation and maintenance costs for CHP and heat pump in microgrid m.

$c_{CHP,m}^{st}, c_{CHP,m}^{sd}, c_{pump,m}^{st}, c_{pump,m}^{sd}$ Start-up and shut-down costs of CHP and heat pump in microgrid m.

c_m^{ES}, c_m^{TS} Electrical and thermal storage costs.

$c_{curtail,m}^{ele,t}, c_{curtail,m}^{gas,t}, c_{curtail,m}^{heat,t}$ Load curtailment cost of electricity, natural gas and heat in microgrid m.

$\eta_m^{CHP}, \eta_m^{e2h}$ Natural gas to electricity and electricity to heat conversion efficiency of the CHP in microgrid m.

$p_m^{CHP,min}, p_m^{CHP,max}$ Minimum and maximum electricity volume generated by the CHP in microgrid m.

$p_m^{CHP,init}, \delta_m^{init}$ Initial electricity volume and status of the CHP in microgrid m.

p_m^{RU}, p_m^{RD} Ramp-up and ramp-down limits of the CHP in microgrid m.

η_m^{pump} Electricity to heat conversion efficiency of the Heat pump in microgrid m.

$q_m^{pump,min}, q_m^{pump,max}$ Minimum and maximum heat volume generated by the heat pump in microgrid m.

$q_m^{pump,init}, \theta_m^{init}$ Initial heat volume and status of the heat pump in microgrid m.

q_m^{RU}, q_m^{RD} Ramp-up and ramp-down limits of the heat pump in microgrid m.

$E_m^{ES,init}$ Initial energy level of ES in microgrid m.

$\eta_m^{ES,c}, \eta_m^{ES,d}, \epsilon_m^{ES}$ ES charging, discharging and self-discharging rate in microgrid m.

$E_m^{ES,min}, E_m^{ES,max}$ Minimum and maximum of the ES energy level in microgrid m.

$p_m^{c,min}, p_m^{c,mac}, p_{d,min}, p_m^{d,max}$ Minimum and maximum of the ES charging and discharging volume in microgrid m.

$E_m^{TS,init}$ Initial energy level of TS in microgrid m.

$\eta_m^{TS,c}, \eta_m^{TS,d}, \epsilon_m^{TS}$ TS charging, discharging and self-discharging rate in microgrid m.

$E_m^{TS,min}, E_m^{TS,max}$ Minimum and maximum of the TS energy level in microgrid m.

$q_m^{c,min}, q_m^{c,mac}, q_{d,min}, q_m^{d,max}$ Minimum and maximum of the TS charging and discharging volume in microgrid m.

$\rho^{min}_{ele,m}, \rho^{max}_{ele,m}$ Minimum and maximum electricity curtailment rate in microgrid m.

$\rho^{min}_{gas,m}, \rho^{max}_{gas,m}$ Minimum and maximum natural gas curtailment rate in microgrid m.

$\rho^{min}_{heat,m}, \rho^{max}_{heat,m}$ Minimum and maximum heat curtailment rate in microgrid m.

D_{a_m} Shiftable load adjustable time window of the household a in microgrid m.

$T^{start}_{a_m}, T^{stop}_{a_m}$ Start and stop time of the load shifting program of the household a in microgrid m.

$d^{min}_{a_m}, d^{max}_{a_m}$ Minimum and maximum of the shiftable load of the household a in microgrid m.

E_{a_m} Total electricity consumption of the household a in microgrid m during the load shifting program.

$p^{PV,t,min}_m, p^{PV,t,max}_m$ Minimum and maximum of the PV-generated electricity volume in microgrid m at time t.

$p^{wind,t,min}_m, p^{wind,t,max}_m$ Minimum and maximum of the wind turbine-generated electricity volume in microgrid m at time t.

τ^{spin} Spinning reserve ratio.

$p^{min,t}_m, p^{max,t}_m$ Minimum and maximum of electricity volume that the microgrid m purchased from the retailer at time t.

$p^{min,t}_{export,m}, p^{max,t}_{export,m}$ Minimum and maximum of electricity volume that the microgrid m sold to the retailer at time t.

$p^{min}_{total}, p^{max}_{total}, g^{min}_{total}, g^{max}_{total}$ Total electricity and natural gas volume that the retailer purchased from the wholesale energy markets.

$\pi^{retail,min}_{ele,m}, \pi^{retail,max}_{ele,m}$ Minimum and maximum of retail electricity price for microgrid m.

$\pi^{retail,min}_{gas,m}, \pi^{retail,max}_{gas,m}$ Minimum and maximum of retail natural gas price for microgrid m.

AVG^{ele}, AVG^{gas} Average retail electricity and natural gas price over the scheduling hours.

η^c_k, η^d_k Charging and discharging efficiencies of the ESS of the retailer k.

Variables

$\gamma^{c,t}_k, \gamma^{d,t}_k$ Charging and discharging status of the ESS of the retailer k.

$\pi_k^{retail,t}$	Retail price of the retailer k at time t.
$\pi_k^{bid,t}$	Bid price of the retailer k at time t.
$q_n^{retail,t}$	Electricity volume the retailer n sold in the retail market at time t.
$q_n^{bid,t}$	Electricity volume that the retailer n bought from the DAW market at time t.
$p_k^{c,t}, p_k^{d,t}$	Charging and discharging power of the ESS of the retailer k at time t.
q_g^t	Electricity volume that the generator g sold to the DAW market at time t.
E_k^t	Energy level of the ESS of the retailer k at time t.
$q_n^{LPE,t}$	Electricity volume that retailer n bought from other retailers (if positive), or sold to other retailers (if negative) in the LPE market.
λ^t	DAW market-clearing price at time t.
$\lambda^{LPE,t}$	LPE market-clearing price at time t.
p_m^t, g_m^t	Electricity and natural gas volume that the microgrid m purchased from the retailer at time t.
$p_m^{export,t}$	Electricity volume that the microgrid m exports to the retailer at time t.
$\pi_{ele,m}^{retail,t}, \pi_{gas,m}^{retail,t}$	Retail electricity and natural gas price for the microgrid m at time t.
$p_m^{CHP,t}, q_m^{CHP,t}$	Electricity and heat volume generated by the CHP in microgrid m at time t.
$g_m^{CHP,t}$	Natural gas volume that consumed by the CHP in microgrid m at time t.
$\delta_m^t, \delta_m^{st,t}, \delta_m^{sd,t}$	CHP operational, start-up and shut-down status in microgrid m at time t.
$p_m^{pump,t}, q_m^{pump,t}$	Electricity consumed and heat generated by the heat pump in microgrid m at time t.
$\theta_m^t, \theta_m^{st,t}, \theta_m^{sd,t}$	Heat pump operational, start-up and shut-down status in microgrid m at time t.
$E_m^{ES,t}$	ES energy level in microgrid m at time t.
$p_m^{c,t}, p_m^{d,t}$	ES charging and discharging volume in microgrid m at time t.

$\gamma_m^{c,t}, \gamma_m^{d,t}$	ES charging and discharging status in microgrid m at time t.
$E_m^{ES,T}$	The final energy level of the ES in microgrid m.
$E_m^{TS,t}$	TS energy level in microgrid m at time t.
$q_m^{c,t}, q_m^{d,t}$	TS charging and discharging volume in microgrid m at time t.
$\zeta_m^{c,t}, \zeta_m^{d,t}$	TS charging and discharging status in microgrid m at time t.
$E_m^{TS,T}$	The final energy level of the TS in microgrid m.
$\rho_{ele,m}^t, \rho_{gas,m}^t, \rho_{heat,m}^t$	Electricity, natural gas and heat curtailment rate in microgrid m at time t.
$\mu_{a_m}^t$	Operational status of the household a in microgrid m at time t.
$d_{a_m}^t$	Shiftable load of the household a in microgrid m at time t.
$p_m^{PV,t}, p_m^{wind,t}$	Electricity generated by PV and wind turbine in microgrid m at time t.
$\psi_m^t, \psi_m^{export,t}$	Electricity importing and exporting status in microgrid m at time t.

Chapter 1

INTRODUCTION

1.1 Background and Problem Statement

With the increasing penetration of renewable energy sources (RES) and distributed energy resources (DERs) in electricity markets and systems, a transactive energy system has been developed to mitigate the pressure on electricity systems and achieve a more flexible and resilient electricity system and market operations. The formal definition of transactive energy is stated by GridWise Architecture Council as follows: *"A system of economic and control mechanisms that allows the dynamic balance of supply and demand across the entire electrical infrastructure using value as a key operational parameter"* [1]. The definition points out the importance of the economic perspective in the transactive energy system. Therefore, transactive energy markets have been proposed in the literature in order to provide a platform for prosumers (e.g., individual consumers with DERs, aggregators and microgrids (MGs)) to balance their power surplus and demand [2]. It can be understood as a solution to climate change and carbon emission reductions. [3] states many benefits of the local power market, such as mitigation of market power, providing flexibility and security to consumers and supporting medium-sized renewable energy power plants.

In a transactive energy system, there exist many research problems regarding the design and operation of the electricity and multi-energy market. For instance, strategic bidding and offering are two important research problems for both wholesale and local electricity markets where market participants attempt to maximise their own profits or minimise their costs by choosing optimal strategies. Many existing studies address the direction but mainly focus on the decision-making problem of electricity producers (e.g., generators). This is due to the fact that previously only electricity producers typically act as price-makers in the wholesale electricity markets [4]–[6]. However, with the development of smart grids and demand response (DR) management, the role of market players, such as energy retailers, has been changing. Traditionally, energy retailers act as price-takers in the wholesale market while offering fixed retail prices to their customers. With the increasing demand-side flexibility empowered by the penetration of DERs such as electric vehicles,

energy storage systems (ESS), photovoltaic, and DR programs [7], [8], energy retailers are now better positioned to make strategic bidding in the wholesale and local electricity markets and offer more flexible retail pricing decisions such as dynamic pricing to end customers [9], [10].

Moreover, since emerging smart grid technologies in the energy system have introduced new opportunities and challenges to both energy suppliers and customers [11], local market participants, such as microgrids and local energy communities, have been accelerating the pace of developing DERs, which has resulted in the increasing trend in developing integrated local energy systems, including electricity, natural gas and heat energy, and the expanding differences among the microgrids [12]. In this regard, the traditional retail pricing strategy of energy retailers, which offers uniform energy tariffs to customers regardless of their differences, cannot fully unlock the potential benefits of DERs and achieve the potential profit [13]. Therefore, energy retailers would need to consider the interdependence among different energy types from the demand side to make robust and reliable retail pricing decisions. Furthermore, energy retailers also need to take the differences among local customers into consideration to offer customised retail prices to each customer. This calls for a novel customised multi-energy pricing scheme capable of capturing the multi-energy interdependence and characteristics of differentiated customers to be developed for the retail energy markets.

In addition, improving the accuracy of energy demand predictions is crucial for energy suppliers to refine and optimise their strategic bidding, offering, and pricing decisions. This forecasting capability goes beyond operational utility and has become a strategic tool in the complex landscape of energy market operations. The importance of long-term energy load forecasting further impacts energy system operators' operating decisions, as they aim to improve the overall efficiency and resilience of the energy system and markets. This investigation into forecasting approaches not only addresses immediate challenges within the volatile energy market but also sets the stage for future innovations in predictive analytics and load forecasting. By doing so, it bridges operational hurdles with strategic advancements in energy management, highlighting the significant relationship between accurate forecasting and strategic energy market participation.

The aforementioned background and research problems show there are many challenges to developing a full transactive energy system considering multi-energy, the uncertainty of renewable energies and demand side management. The overall

general challenges in the transactive energy system can be summarised as follows.

1. Most of the existing studies about strategic bidding and offering problems focus on the energy producers, such as generators, which are considered price-makers. However, since the development of Smart Grid technologies and DR management, other energy participants, such as energy retailers and aggregators, can also contribute significantly to the energy pricing decisions of energy markets, such as wholesale and local/retail markets. There is a need to study the strategic bidding and offering behaviours of multiple strategic energy retailers as price-makers participating in energy markets.

2. With the development of Smart Grid and DR technologies, energy customers could gain the potential desire to change the energy carriers more frequently based on the energy price each energy carrier offers. There are only a few existing research addresses along the direction. Therefore, an advanced study of customers' switching behaviours is urgently needed.

3. Since accelerating the development of DERs, the trend of establishing integrated local energy systems, including electricity, natural gas and heat energies, is growing rapidly. In addition, the differences among energy customers, such as microgrids, also need to be taken into consideration as the DERs and DR technologies could differentiate energy customers significantly. Therefore, energy suppliers, such as energy retailers and aggregators, would need to develop a novel multi-energy pricing scheme to cope with the interdependence among different energies and characteristics of different energy customers.

4. With the recent development and popularisation of smart meters, the importance of developing advanced long-term energy load forecasting models for energy participants, such as energy retailers, has increased dramatically. In addition, forecasting models for long-term time-series data (e.g., energy load) need to be improved based on the latest development of machine/deep learning techniques, such as the Transformer-based models, adopted extensively in the field of natural language processing (NLP) and computer vision (CV).

1.2 Research Aims and Objectives

This research aims to develop a smart hierarchical transaction energy system that considers multi-energy, renewable energies and demand-side management. To achieve this aim, the following objectives are established below.

- To address Challenge 1, a novel optimisation model needs to be built to tackle the strategic bidding and offering problems where there are multiple energy retailers participating in energy markets, such as wholesale and retail markets. In addition, the retail competition among the energy retailers and the equilibrium energy price need to be simulated and calculated.

- To cope with Challenge 2, a model that can formulate the customers' behaviour of switching to different energy carriers needs to be developed based on the energy price that each energy carrier offers.

- To address Challenge 3, an innovative multi-energy pricing scheme for energy retailers is required to be developed to capture the interdependence among different energies, such as electricity, natural gas and heat energies, and the differentiated characteristics of different energy customers. Specifically, smart grid technologies, such as DERs, DR, and renewable energies, should be considered significant characteristics of energy customers.

- To address Challenge 4, a novel Transformer-based long-term time-series forecasting (LTTSF) model needs to be developed to outperform other LTTSF models at predicting multi-energy loads for energy suppliers, such as energy retailers and aggregators. The model should improve long-term prediction accuracy significantly by capturing the interrelationship among different energies and features, such as electricity, natural gas, heat and greenhouse gas (GHG), and long-range, past energy-related information.

1.3 Contributions

Based on the aforementioned challenges, research aims, and objectives, the thesis has made three major contributions to developing a smart hierarchical transactive energy system. The contributions are concluded as follows.

- A novel bilevel optimisation model for formulating strategic behaviours of multiple retailers as price-makers participating in both day-head wholesale and local markets is developed. The proposed bilevel model consists of multiple retailers, multiple electricity markets, and customers' ability to switch to different retailers. To the best of our knowledge, this is the first work from the bilevel game-theoretic perspective to investigate the problem for multiple retailers considering customers' switching behaviours and market share. In particular, the single retailer's bilevel problem is first transformed into

mathematical programming with equilibrium constraints (MPEC) problem via deriving Karush–Kuhn–Tucker (KKT) conditions from lower-level problems. The derived bilinear terms and complementarity slackness constraints, which result in the non-convexity of the MPEC, are solved by linearisation methods and lead to a tractable mixed-integer quadratic programming (MIQP) problem. In addition, the Bertrand competition model is used to formulate the retail competition among strategic retailers. This further forms the equilibrium problems with equilibrium constraints (EPEC) problem and is solved by the diagonalisation algorithm.

- An innovative customised multi-energy pricing scheme is proposed and formulated by a single-leader multi-follower bilevel optimisation model. In particular, the energy retailer's profit maximisation problem is modelled at the upper level, while the operational cost minimisation problem for each microgrid is formulated at the lower level. Furthermore, at the lower level, the energy management model for each microgrid equipped with energy converters (i.e., combined heat and power (CHP) and heat pump), electrical and thermal storage, RES (i.e., solar and wind) and DR programs (i.e., load curtailment and shifting) is formulated as a mixed-integer linear program (MILP) program. Three hybrid metaheuristic algorithms (i.e., particle swarm optimisation (PSO), genetic algorithm (GA) and simulated annealing (SA)) combined with the conventional MILP are developed to solve the proposed bilevel problem to conquer the non-convexity of lower-level problems.

- A novel Transformer-based model architecture, Patchformer, is proposed for long-term multi-energy load forecasting. It integrates the patch embedding block and encoder-decoder structure. In particular, the patch embedding block splits the multivariate time series into various distinct univariate inputs and segments each of them into subseries-level patches which can capture local semantic information within each univariate time series. The approach can also learn inter-channel relationships effectively since each channel shares the same embedding and Transformer weights. In addition, with the multi-head attention mechanism, the encoder-decoder structure facilitates capturing long-range dependencies and understanding the complexity of time-series data, which potentially improves forecasting accuracy.

1.4 Thesis Outline

The rest of this thesis is organised as follows. Chapter 2 reviews the relevant literature related to strategic bidding and offering, multi-energy pricing, and long-term multi-energy load forecasting problems, respectively. Chapter 3 proposes a bilevel game-theoretic decision-making framework that considers multiple energy suppliers' (i.e., retailers') strategic bidding and offering problems in day-ahead wholesale and local electricity markets while taking customers' switching behaviour into account. Chapter 4 develops a customised multi-energy retail pricing scheme considering the interdependence among different energies, such as electricity, natural gas and heat energies, and the differences among managed microgrids equipped with energy converters, storage, RES and DR programs. Chapter 5 proposes a long-term time series forecasting model named Patchformer for multi-energy load prediction, which is a significant shift towards adopting predictive analytics to better forecast future energy demands. Lastly, Chapter 6 concludes this thesis and discusses the directions of future research.

Chapter 2

LITERATURE REVIEW

This chapter presents background information and reviews relevant literature related to all research aims and objectives mentioned in the last chapter. Section 2.1 reviews the problem of strategic bidding and offering, the related bilevel modelling and its solutions. Section 2.2 introduces the recent development of retail pricing schemes for energy suppliers. The solution methods for a bilevel model with integer variables at the lower level are also discussed in detail. Lastly, the problem of long-term time-series forecasting for multi-energy datasets is reviewed in Section 2.3.

2.1 Strategic Bidding Among Energy Suppliers

The decision-making of participants in hierarchical systems (e.g., electricity markets) is often modelled as a bilevel optimisation problem or Stackelberg game [14][15]. In bilevel models for electricity markets, strategic participants (e.g., electricity generators and retailers) either maximise their profits or minimise their costs at the upper level. The lower level usually consists of a market-clearing problem solved by an independent system operator (ISO) or a customer-side energy management problem. The standard approach to solving the bilevel models is reformulating it as a single-level mixed-integer program by applying KKT conditions to the lower-level problem. There are numerous existing studies along this direction. For instance, in [16], a scenario-based bilevel model has been applied to a large consumer's profit maximisation problem where the wholesale market-clearing problem is considered at the lower level, and a heuristic method is introduced to solve one MPEC problem per scenario. [17] proposes a customised pricing framework for retailers for different residential users. The pricing framework is modelled as a bilevel program where retailers purchase electricity from wholesale markets and compete for market share. Although the bilevel models considering retailers, system operators, or generators are prevailing, there is increasing attention paid to other market participants, such as DR aggregators and microgrids. For instance, [18] introduces multi-energy players as aggregators to maximise their profits and mitigate their operational risks. The problem is modelled as a bilevel problem and interpreted as an MPEC problem. [19] focuses on the reserve management problem of the electric vehicles (EVs) aggregator. The upper level of the bilevel

model is formulated as the profit maximisation problem of the EVs aggregator. The lower level represents the optimal charging/discharging decisions of EVs owners. An exact and finite decomposition algorithm is proposed to solve the problem in an iterative manner. [20] proposes a bilevel program for EVs aggregators from a different perspective. Instead of maximising profit at the upper level, charging cost minimisation is formulated. The lower level represents the day-head wholesale (DAW) market-clearing problem. [21] develops a single-leader multi-follower game model where the market operator acts as the player at the upper level and smart grid entities at the lower level aim to optimally schedule their own renewable energy resources, energy storage, and DR resources. Likewise, [22] develops a bilevel model for microgrids to achieve optimal bidding strategy, in which the lower level is the distributed energy market's clearing problem and the upper level represents the optimal scheduling problem for a microgrid. [23] constructs a bilevel Stackelberg competition model to investigate the interaction between regulated and merchant storage investment. A merchant profit maximisation problem is modelled at the upper level, while an overall system cost minimisation problem is formulated at the lower level. [24] proposes a stochastic bilevel framework to model the interactions between a wind power producer at the upper level, and EVs and DR aggregators at the lower level. The wind power producer is also formulated to achieve optimal bidding decisions in the competitive wholesale markets.

From an economics point of view, existing studies on strategic bidding and offering problems can be classified based on whether the market participants are price-makers or price-takers [5]. If the market participants have relatively large-scale and flexible loads or supplies, they can be considered as price-makers. Along this direction, [9] develops a short-term planning model of a price-maker retailer with flexible power demand participating in the DAW electricity market. [25] develops a new scenario-based stochastic optimisation model for price-maker economic bidding in both day-ahead and real-time markets where a DR program with time-shiftable load is adopted to create load flexibility. [20] proposes an optimal bidding strategy for a large-scale plug-in electric vehicle (PEV) aggregator. The upper level represents the charging cost minimisation of the PEV aggregator, whereas the market-clearing problem is formulated at the lower level. In contrast, if the market participants are small-scale or have inelastic loads or supplies, they usually act as price-takers. Along this direction, [7] formulates a stochastic mixed-integer linear program to obtain an optimal bidding strategy for a DERs aggregator participating in the day-ahead market where the market-clearing prices are given by different scenarios. In [21], the lower

level of the bilevel program represents multiple smart grids' optimal scheduling problems, whereas the ISO clears the day-ahead market at the upper level. [26] takes both price-maker and price-taker positions into consideration. Specifically, the DR aggregator acts as a price-taker and a price-maker in the day-ahead and real-time market, respectively.

The decision-making of multiple retailers has also been studied in the literature either through a single-level model or a bilevel model. For the former, [27] addresses the portfolio optimisation model of retailers, which involves a risk-return optimisation method based on the Markowitz theory. [28] proposes a multistage stochastic optimisation approach to capture the uncertainties of electricity loads and prices for retailers' contract portfolios, which account for their risk preferences. For the latter, [29] proposes a bilevel multi-leader multi-follower game to investigate the benefit of aggregation of prosumers to revenue generation in wholesale and retail markets in which aggregated prosumers act as retailers (leaders) and end-users act as followers. [30] considers strategic firms as leaders in the upper-level problem, whereas electricity and natural gas market operators act as followers in the lower level. [31] presents a dynamic pricing framework for electricity and gas utility companies in the coupled retail electricity and natural gas markets by developing a two-leader multi-follower bilevel model. In particular, the electricity and gas utility companies acting as leaders serve energies to the integrated DR aggregators which are followers at the lower level. The competition among multi-energy retailers in the presence of integrated DR prosumers is formulated as a multi-leader-follower bilevel game in [32]. Lastly, [33] considers an EPEC framework to model the interaction among generation companies, microgrids, and load aggregators participating in the wholesale and distribution network electricity markets. In Chapter 3, we study multiple strategic retailers as price-makers participating in both wholesale and local/regional energy markets within the bilevel decision-making framework.

Existing studies can be further categorised based on whether market players participate in multiple levels of markets (e.g., wholesale vs. local/ retail) simultaneously. Most studies, however, are often based on a single electricity market, such as day-ahead market [4], [6], [7], [9], [16], [20], [21], [35] or retail market [5], [10], [34], [39], [42]. There are also a few studies that focus on analysing interactions among market participants in the wholesale (i.e., day-ahead and real-time) electricity markets [17], [25], [26]. Only a few studies in the literature consider multiple levels of markets simultaneously, such as wholesale and retail markets [18], [33], [40],

Table 2.1: Literature classification. ✓: Yes; ✗: No; – : Not applicable.

Literature	**Bilevel model**	**Price maker**	**Multi-market**	**Multi-leader**	**Customer behaviour**
[29], [31]–[33]	✓	✓	✓	✓	✗
[5], [6], [16], [19], [20], [22], [34]	✓	✓	✗	✗	✗
[4], [30], [35]	✓	✓	✗	✓	✗
[21]	✓	✗	✗	✗	✗
[7]	✗	✗	✗	–	✗
[9], [28], [36], [37]	✗	✓	✗	–	✗
[27]	✗	✓	✓	–	✗
[18], [38]	✓	✓	✓	✗	✗
[39]	✓	✓	✗	✗	✓
[10]	✗	✗	✗	–	✓
[17], [26], [40], [41]	✓	✓	✓	✗	✓
[25], [42]	✗	✓	✓	–	✓
[43]	✗	✓	✓	✓	✓

[43]. For instance, the aggregator in [18] participates in both the wholesale and local energy markets. [40] proposes a framework that can optimise the strategy of a distribution company owning DERs and ESS in the wholesale and retail energy markets. In Chapter 3, we also consider multiple levels of electricity markets (i.e., wholesale and local markets). Apart from the conventional retail market, we develop a novel local/regional energy exchange market named the local power exchange (LPE) market for retailers. In the literature, studying the local energy market typically focuses on modelling the operation of emerging market participants such as prosumers, DERs aggregators, and microgrids [36], [37]. For instance, in [36], a local power exchange centre is developed where a novel clustering algorithm is developed to cluster prosumers trading in the local energy market geographically. Another local energy exchange market design for energy trading among energy storage unit owners is studied in [37], where a novel local energy exchange market-clearing approach is proposed based on double auctions. However, modelling the established and traditional role of energy retailers in the local market is much less studied. In Chapter 3, we propose an LPE market for strategic retailers equipped with energy storage to manage their supply and demand deviation. Compared to the papers mentioned above, the uniqueness of our proposed LPE market lies in that

1) The participants in the LPE market are strategic retailers equipped with energy storage and arbitrage opportunities; 2) retailers in the LPE market can buy/ sell electricity from/ to other retailers; 3) the LPE market provides a platform for retailers to balance their supply and demand deviation in a local level market. This new local market for energy retailers will complement existing local energy markets to better facilitate the management of local and distribution energy systems.

In addition to the strategic decision-making problem of multiple retailers in multiple levels of electricity markets, customers' switching behaviours are also modelled in Chapter 3. There are only a few existing studies that address this direction. For instance, [43] considers customers' switching behaviours in the retail market where a single-level model is proposed to maximise the profit of strategic retailers. [41] presents a decision-making framework for an electricity retailer considering the rational response of consumers in a competitive environment. The retailer is considered as a price-taker in the day-ahead market, and the rival retailers' selling prices are assumed to be given. The switching behaviours of consumers are modelled as the switching cost for the hesitation of consumers to switch contracts between retailers. [39] adopts utility functions to model three categories of DR customers based on their sensitivity to retail prices from low, semi, to high flexibility. It should be noted that modelling customers' switching behaviours for the strategic offering of multiple retailers is particularly crucial to capturing the switching decisions of customers among different retailers, the implications and impacts on retailers' strategic decisions, and the market operations. To the best of our knowledge, there is no existing research that tackles this problem while considering the hierarchical nature of multiple competitive price-maker retailers and customers.

The above-reviewed literature is summarised in Table 2.1. To fill the research gap following the above analysis, we propose a bilevel game-theoretic framework in Chapter 3 to model the multiple retailers' (as price-makers) optimal decision-making problems when participating in both wholesale and local markets with customers' switching behaviours considered.

2.2 Retail Pricing Schemes of Energy Suppliers

Retail energy pricing is an important research problem and has been extensively studied in the literature. [44] systematically investigates and summarises the existence of retail pricing schemes, which demonstrates that the real-time pricing (RTP) strategy can well utilise the demand-side management flexibility over the static

pricing strategy, such as the time-of-use (TOU) scheme. [45] presents a dynamic, real-time energy pricing mechanism to accurately distribute power for the EVs charging process fairly when the microgrids are congested. A retail RTP scheme in the presence of hydrogen storage systems and EVs, compared with TOU and fixed pricing schemes, is proposed in [46]. The bi-objective problem is formulated, in which the average profit needs to be maximised, while the profit deviation should be minimised. The Pareto optimal solution is obtained by the epsilon constraint method and fuzzy satisfying approach. The numerical results show the privilege of the RTP scheme regarding the obtained profit of the retailer. [47] develops a conditional value at risk (CVaR)-based retail electricity pricing scheme to reduce the impact of risk caused by the uncertainties from RES generation and estimated wholesale electricity prices. An energy management and pricing method for the community energy retailer incorporating smart building consumers is investigated in [48]. The numerical results, which are solved by the bilevel chance-constrained programming approach, show that the proposed approach can benefit both retailers and customers. Lastly, [49] proposes a bilevel game-theoretic model for multiple strategic retailers' decision-making problems, which include the retail prices in the retail market, bid prices in the day-ahead wholesale market, and the bid/offer prices in the local power exchange market. The problem is formulated as an EPEC problem, which is solved by the diagonalisation algorithm.

The majority of the existing literature focuses on developing a retail pricing strategy in the electricity market, whereas only few studies analyse the retail pricing problem in a multi-energy context. In addition, since the development of a multi-energy pricing scheme consists of energy suppliers (e.g., retailers) and customers (e.g., microgrids and aggregators), a bilevel optimisation model, which can well present the intrinsic hierarchical structure of the energy system, has been widely adopted in the literature. For instance, [50] proposes a bilevel optimal retail pricing scheme for the retailer and multi-energy buildings to perform the price-based DR programs. A bilevel stochastic RTP model in the framework of the Markov Decision Process is formulated for the multi-energy system in [51]. A novel distributed online multi-agent reinforcement learning algorithm is developed to solve the proposed model. [52] proposes a bilevel multi-energy trading model between the multi-energy service provider and consumer by setting the optimal energy pricing scheme and energy economic dispatch at the upper level. Optimal consuming patterns of different energies are obtained for the multi-energy consumer in the lower-level problem. [53] develops an integrated energy service provider (IESP) as a retailer to effectively

set energy prices and energy management in the multi-energy market. The impact of DR and wholesale prices' uncertainties is considered in the proposed two-stage stochastic hierarchical framework. The day-ahead energy pricing and management method considering multi-energy DR programs for IESP in regional integrated energy systems is addressed in [54]. The bilevel Stackelberg game optimisation model is established and shows that the pricing scheme benefits both the energy supplier (i.e., IESP) and the consumer. The pricing behaviour of multi-energy players who can trade electricity, natural gas and heat energy to maximise their profits and reduce their operational risk is studied in [18]. The bilevel approach is applied to model the decision-making conflict of the multi-energy players with other energy players participating in the multi-energy system.

The bilevel optimisation model is typically solved by analytical mathematical methods, such as the KKT-based reformulations in the literature [18], [50], [52]–[54]. However, the premise of applying those methods may include the convexity of the lower-level problem. For the bilevel model, whose lower-level problem is proved as non-convex, such as including integer variables, traditional mathematical approaches cannot solve it effectively. Therefore, many metaheuristic algorithms, such as PSO, GA and SA, are introduced in the literature to overcome the non-convexity of the lower-level problem [55]. For instance, [56] develops a hierarchical market framework to apply real-time retail pricing between an energy supplier and multi-energy microgrids, where a hybrid solution method combining PSO and the branch and bound algorithm is proposed. In [57], a dynamic pricing profile is developed for utilising distributed energy storage and overcoming the intermittency from renewable generation. A novel non-cooperative Stackelberg game is proposed to formulate the pricing problem, in which the upper-level problem is solved by the PSO algorithm, and linear programming is applied at the lower level. In addition, [58] presents a bilevel price decision model for a load-serving entity to manage multiple multi-energy microgrids with DR. The microgrids' operation problems are formulated at the lower level and solved by a MILP program. A GA algorithm is applied to find the optimal price decisions for the load-serving entity at the upper level. A real-time pricing strategy is proposed in [59] to effectively adjust the power balance between the supply and demand and manage the microgrid's internal energy dispatch. The pricing strategy is formulated as a bilevel programming model, where the supplier's price decision-making at the upper level is solved by the GA algorithm. [60] develops an SA-based price control algorithm to solve the non-convex real-time pricing problem, which can reduce the peak-to-average load ratio and retailer's cost

through DR management in smart grid systems. [15] proposes a bilevel Stackelberg game to model hierarchical interactions between one profit-seeking energy retailer and multiple cost-minimising energy customers. A GA algorithm is developed to solve the bilevel model. Lastly, [10] proposes a bilevel model, where data-driven appliance-level customer behaviour learning models are developed at the lower level. The resulting hybrid optimisation–machine learning bilevel model is solved by GA.

The retail pricing schemes mentioned in the above literature are classified into the uniform pricing scheme, where the decision maker optimises the pricing decisions without considering the different characteristics among underlying energy customers, such as the demand-side load profile and the specifications of the equipped DERs, such as energy converters and storage. Therefore, the elasticity on the demand-side management of each energy customer cannot be fully utilised under the uniform pricing scheme. To conquer the deficiency of the widely used uniform pricing scheme, a novel customised energy pricing scheme should be developed that takes the unique characteristics of each customer into consideration. It is designed to motivate the potential flexibility of demand-side management to benefit market participants, such as energy retailers. However, only few studies address the customised retail energy pricing problem in the literature. For instance, [61] develops a bilevel model for optimal differential pricing considering different customer groups characterised by different price sensitivities. [17] proposes a customised TOU electricity pricing scheme for different residential users depending on their load-consumption profiles, which is established based on the bilevel optimisation framework. The problem of customising TOU electricity retail prices based on load profile analysis by applying a clustering algorithm is addressed in [62]. Similarly, a realistic multiple dynamic pricing scheme for the segmented customers based on the different identification of load patterns is proposed in [13], which demonstrates the effectiveness of the clustering-based approach. Furthermore, the electricity retailer could achieve better profit gain under the proposed multiple pricing scheme. [63] presents a personalised RTP scheme using a bilevel model to improve the management of different electricity consumption, including both traditional and renewable energies. [64] proposes a bilateral energy-trading structure to coordinate the hierarchical personalised electricity pricing model between the energy trading agent and energy prosumers.

The above-reviewed literature is compared and summarised in Table 2.2. Although the above studies provide valuable insights regarding the customised retail pricing

Table 2.2: Literature comparison. ✓: Yes; ✗: No; –: Not applicable.

Literature	Uniform Pricing	Customised Pricing	Electricity Market	Multi-Energy Market	Bilevel Model	KKT-Based Approach	Metaheuristic-Based Approach
[45]–[47]	✓	✗	✓	✗	✗	–	–
[48], [49]	✓	✗	✓	✗	✓	✓	✗
[18], [50], [52]–[54]	✓	✗	✗	✓	✓	✓	✗
[51]	✓	✗	✗	✓	✓	✗	✗
[56]–[58]	✓	✗	✗	✓	✓	✗	✓
[10], [15], [59], [60]	✓	✗	✓	✗	✓	✗	✓
[61]	✗	✓	✓	✗	✓	✗	✓
[17]	✗	✓	✓	✗	✓	✓	✗
[62]–[64]	✗	✓	✓	✗	✓	✗	✗
[13]	✗	✓	✓	✗	✗	–	–

scheme, they are limited to electricity markets. Therefore, to the best of our knowledge, there is no existing research studying the *customised multi-energy pricing* problem. To fill the research gap following the above analysis, in Chapter 4, we propose a novel customised multi-energy pricing scheme for an energy retailer that manages multiple microgrids in the multi-energy market.

2.3 Long-Term Multi-Energy Load Forecasting for Energy Suppliers

The development of Integrated Multi-Energy Systems (IMES) marks a significant evolution in the energy sector, which reflects a shift towards more diversified, efficient, and sustainable energy management practices. IMES encompasses various energy sources and forms, including electricity, gas, heat, cooling and renewables, and it integrates into a comprehensive system. This integration facilitates a more holistic energy production, distribution, and consumption framework. The evolution of IMES reflects a growing recognition of the need for more resilient and adaptable energy systems, especially in the face of escalating global energy demands and the need for improving collective technical, economic, and environmental performance [65]. IMES has become an important strategic development direction in the energy field to deal with the global challenges in the fossil energy crisis, climate changes and environmental pollution. The role of energy forecasting is central to the effective operation of IMES. Precise prediction plays a crucial role in overseeing these complex systems, guaranteeing that energy generation and supply correspond with

demand trends. This alignment is essential for maximising the economic and environmental benefits of IMES as well as for their operational efficiency. Accurate load forecasting contributes to the economic and environmental sustainability of energy systems by optimising resource allocation and lowering operating costs. For instance, [66] shows that the annual economic loss could be up to 10 million pounds when every percentage point increases in the error of electricity load forecasting in the United Kingdom. In addition, [67] also illustrates that when the forecasting error is decreased by 1%, the total energy consumption of 58 million MW/h can be saved in one year in China. Moreover, effective energy load forecasting promotes the integration of renewable energy sources, assisting in the reduction of carbon/ greenhouse gas emissions and furthering the aims of sustainable energy from an environmental standpoint. The prediction's reliability and accuracy are critical for the development of future energy systems to be sustainable and able to satisfy energy demands economically.

The methods of time-series forecasting can be categorised into two groups: 1) traditional methods represented by time-series analysis and regression analysis. 2) artificial intelligence methods represented by machine learning and deep learning. Autoregressive integrated moving-average (ARIMA) [68], [69] is one of the most popular models in the former group. It predicts the data by obtaining the fitting equations for time series data and other variables. However, the traditional forecasting methods are mainly based on linear analysis and have limited capacity to deal with nonlinear problems [67]. The latter group of time-series forecasting methods includes various machine learning and deep learning techniques, such as Recurrent Neural Networks (RNN) [70], Long Short-Term Memory (LSTM) networks [71], Gated Recurrent Units (GRU) [72], [73], and Transformer-based methods [74]–[78], which have significantly enhanced forecasting capabilities. Among these, Transformer-based models stand out for their ability to handle large datasets and capture complex temporal relationships in multivariate data, offering substantial improvements over other techniques. This is due to their parallelisability and attention mechanism. In particular, the vanilla transformer model [79] uses a scaled dot product attention mechanism to calculate the point-wise correlation between two different data points. However, applying the full attention mechanism to LTTSF is computationally expensive due to its quadratic complexity in terms of the length of input sequence L, which makes it challenging to handle long-term sequences. To overcome this issue, many improved models for LTTSF have been developed. For instance, Reformer [74] has been designed to enhance the efficiency for training on

long sequences by introducing the Locality-Sensitive Hashing attention to reduce the computational complexity from $O(L^2)$ to $O(L \log L)$. Informer model [75], which also achieves $O(L \log L)$ complexity, introduces a ProbSparse self-attention mechanism, self-attention distilling, and a generative style decoder. These features collectively enhance the model's efficiency and prediction capacity, making it a robust solution for LTTSF. Although the computational complexity has been reduced, the above models still use the point-wise attention mechanism to understand the correlation between two data points. This may not be the preferred mechanism in time-series forecasting compared to NLP. It is because, unlike a word in a sequence sentence, a single time step in a time series does not have semantic meaning. Therefore, the point-wise correlation cannot capture the input sequence's local semantic information or pattern, which results in poor LTTSF performance. A few models have been developed to address the problem. For instance, Autoformer [76] incorporates a decomposition architecture with an Auto-Correlation mechanism calculated by Fast Fourier Transforms (FFT) based on series periodicity. It focuses on discovering dependencies and aggregating representations at the sub-series level. This mechanism is more efficient and accurate than traditional self-attention mechanisms, especially for long-term forecasting. Similarly, FEDformer [77] also introduces the seasonal-trend decomposition and Frequency Enhanced Attention block with Discrete Fourier Transform (DFT) to capture the sub-series level correlation in time-series sequences. Moreover, inspired by the Vision Transformer (ViT) [80] which truncates each image into 16×16 patches before feeding it into the vanilla transformer model, and the following influential work BEiT [81], PatchTST [78] segments time series into subseries-level patches as input tokens, which is designed to retain local semantic information, and employs channel-independence, where each channel contains a single univariate time series sharing the same embedding and Transformer weights which benefits for multivariate time-series forecasting. This design enhances long-term forecasting accuracy significantly compared to state-of-the-art Transformer models, reduces computation and memory usage, and allows the model to attend to a more extended history. Chapter 5 proposes a novel Transformer-based model which integrates the patch embedding mechanism and vanilla transformer's encode-decoder architecture to improve the LTTSF performance.

The methods for load forecasting also consist of both traditional and artificial intelligence methods. For instance, [82], [83] apply ARIMA and its variant seasonal autoregressive integrated moving-average (SARIMA) for load forecasting problems.

[84] proposes a novel pooling-based deep RNN for household load forecasting, which batches a group of customers' load profiles in a pool of inputs. [85] introduces an LSTM RNN-based framework to forecast the highly volatile and uncertain electric load of an individual energy customer. A novel short-term load forecasting method based on attention mechanism, rolling update and bi-directional long short-term memory (Bi-LSTM) neural network is proposed for short-term electricity load forecasting in [86]. One of the characteristics of multi-energy data is the interdependence among each energy. Forecasting models have been built to capture the interdependence to enhance the forecasting accuracy in the literature. For instance, [67] proposes an encoder-decoder model based on LSTM, considering the high-dimensional temporal dynamic characteristic. To capture the cross-coupling characteristic, a coupling feature matrix for multi-energy load is established. [87] presents a convolutional neural network (CNN)-Seq2Seq model with an attention mechanism based on a multi-tasking learning method for a short-term multi-energy load forecasting, which considering temperature, humidity, wind speed, and the coupling relationship of multi-energy. An improved multi-energy forecasting method, which uses a CNN-Attention-LSTM model based on federated learning to predict multi-energy load in the integrated energy microgrid is proposed in [88]. In time-series forecasting literature, the prediction length for LTTSF typically ranges from 96 to 720 time steps with hourly data [76]–[78], [89]. On the other hand, short-term energy load forecasting literature usually predicts for a few days or weeks [86], [87], [90], [91], while long-term energy load forecasting literature extends to months or years [92]–[94]. However, the daily or monthly data used in energy load prediction may have fewer time steps to predict than LTTSF literature. To predict multi-energy load with hourly data, this thesis follows the definition of long-term in LTTSF literature. Furthermore, long-term multi-energy load forecasting has witnessed a paradigm shift in recent years with the development of Transformer-based models. Many Transformer-based models are developed for multi-energy load forecasting in the literature. For instance, A one-encoder multi-decoder multi-task model is developed in [95] to capture the joint relationships among different energies. A similar idea is also adopted in [96] with the novel Bayesian multi-head attention mechanism. Despite the growing body of research on Transformer-based models in multi-energy forecasting, a research gap remains in the literature. To the best of our knowledge, there has been no exploration of Transformer-based models that incorporate patch embedding techniques in long-term multi-energy load forecasting, which has shown great promise in other domains, such as NLP and CV, for its ability to capture

local contextual information and reduce computational complexity. [91] proposes a PatchTCN-TST model, which applies a patching approach but only for short-term multi-load energy forecasting. The absence of patch embedding-based Transformer models in long-term multi-energy load forecasting is a significant neglect. Such models have the potential to enhance the model's ability to process and learn from multivariate time series data, capture local and global semantic information, and provide a deep understanding of energy consumption patterns. This approach could lead to more accurate and robust forecasting models, which are essential for effective energy management and planning in the face of increasing demand and the growing complexity of future energy systems.

2.4 Conclusion

This chapter reviews the related literature of three major research topics, which are: 1) strategic bidding among energy suppliers, 2) retail pricing schemes of energy suppliers, and 3) long-term multi-energy load forecasting for energy suppliers. In particular, due to the intrinsic hierarchical structure of the problem, the modelling of strategic bidding and offering for energy suppliers, such as retailers and aggregators, is discussed using a bilevel model approach. Furthermore, the KKT solution method is applied to solve the bilevel model. In addition, the role of the price-maker and price-taker is compared. It also reviews the problem of multiple energy retailers participating in multiple energy markets. For the second research topic, the retail energy pricing problem in electricity and multi-energy markets is presented in detail. The related solution methods, such as KKT and metaheuristic methods (e.g., PSO, GA and SA) are discussed. Finally, the customised multi-energy retail pricing scheme is proposed. In terms of the third research topic, the importance of multi-energy load forecasting is illustrated. Moreover, it reviews the development of general time-series forecasting and long-term multi-energy load forecasting in detail. The next three chapters, namely, Chapter 3-5, focus on each of the research topics and provide detailed contributions, model formulation, and numerical analysis.

Chapter 3

A BILEVEL GAME-THEORETIC DECISION-MAKING FRAMEWORK FOR STRATEGIC RETAILERS IN BOTH LOCAL AND WHOLESALE ELECTRICITY MARKETS

This chapter is reproduced with changes from:

- Q. Hong, F. Meng, J. Liu, and R. Bo, "A bilevel game-theoretic decision-making framework for strategic retailers in both local and wholesale electricity markets," Applied Energy, vol. 330, p. 120311, 2023.

In Chapter 2, the background information and literature review are illustrated for strategic bidding and offering problems. To tackle the problem, a multi-leader, multi-follower bilevel optimisation model for strategic retailers participating in both local and wholesale energy markets is proposed in this chapter.

3.1 Introduction

This chapter proposes a bilevel game-theoretic model for multiple strategic retailers participating in both wholesale and local electricity markets while considering customers' switching behaviours. At the upper level, each retailer maximises its own profit by making optimal offering decisions in the retail market and bidding decisions in the DAW and LPE markets. The interaction among multiple strategic retailers is formulated using the Bertrand competition model. For the lower level, there are three optimisation problems. First, the customers' welfare maximisation problem with their switching behaviours is formulated to capture the demand responses from customers. Second, a market-clearing problem is formulated for the ISO in the DAW market. Third, a novel LPE market is developed for retailers to facilitate their power balancing. In addition, the bilevel multi-leader multi-follower Stackelberg game forms an EPEC problem, which is solved by the diagonalisation algorithm. Numerical results demonstrate the feasibility and effectiveness of the EPEC model and the importance of modelling customers' switching behaviours. We corroborate that incentivizing customers' switching behaviours and increasing the number of retailers facilitates retail competition, which results in reducing strategic retailers' retail prices and profits. Moreover, the relationship between customers' switching

behaviours and welfare is reflected by a balance between the electricity purchasing cost (i.e., electricity price) and the electricity consumption level.

3.1.1 Contributions

The contributions of this chapter are summarised as follows:

• We propose a novel bilevel model to formulate strategic behaviours of multiple retailers as price-makers participating in both DAW and local markets. The proposed bilevel model consisting of multiple retailers, multiple electricity markets, and customers' abilities to switch to different retailers is particularly important to model practical scenarios. To the best of our knowledge, this is the first work from the bilevel game-theoretic perspective to investigate the problem for multiple retailers considering customers' switching behaviours and market share.

• The bilevel problem with a single retailer is firstly reformulated into an MPEC problem by deriving KKT conditions from lower-level problems. To overcome the non-convexity in the resulting MPEC problem introduced by the bilinear terms and complementarity slackness constraints, linearisation methods are conducted, which leads to a tractable MIQP problem. In addition, the Bertrand competition model is adopted to model the interaction among strategic retailers, which is formulated as an EPEC problem and solved by the diagonalisation algorithm.

• Comprehensive numerical results are provided to verify the feasibility and effectiveness of the proposed EPEC model and diagonalisation algorithm. In addition, the effects of customers' switching behaviours and the number of retailers in the markets on the strategic retailers' optimal decisions are extensively studied. Specifically, increasing customers' switching behaviours and the number of retailers promotes retail competition, which negatively correlates to strategic retailers' equilibrium retail prices and profits. The relationship between customers' switching behaviours and their welfare is also elaborated.

3.1.2 Chapter Organisation

The remainder of this chapter is organised as follows. The proposed bilevel model of a single retailer is developed in Section 3.2. Section 3.3 discusses the methodologies for reformulating the bilevel model into an MIQP model. Furthermore, the diagonalisation algorithm for solving the EPEC problem with multiple retailers is also proposed in this section. Numerical results are presented and discussed in detail in Section 3.4.

3.2 Bilevel Game-Theoretic Model

This section proposes a bilevel optimisation problem for a strategic retailer who maximises its profit. Specifically, the strategic retailer participates in DAW and local markets (i.e., retail and LPE markets). The detailed description of the proposed bilevel model is presented in Section 3.2.1. Furthermore, the upper and lower-level problems of the bilevel model are introduced and analysed in Section 3.2.2 and 3.2.3, respectively. Consequently, the complete bilevel model is formulated in Section 3.2.4.

3.2.1 Model Description

The proposed bilevel model with a single retailer can be interpreted as a single-leader multi-follower game where the strategic retailer acts as the leader, whereas customers, ISO, and the LPE market operator are followers. In particular, the strategic retailer optimises the ESS management and pricing decisions (i.e., retail prices in the retail market, bid prices in the DAW market, and bid/offer prices in the LPE market) at the upper level. Subsequently, customers react to the optimal load demand at the lower level based on their welfare. Market operators clear their corresponding markets (i.e., DAW and LPE markets) and send their cleared electricity volume back to the strategic retailer. The structure of the proposed bilevel model is shown in Figure 3.1. Specifically, the strategic retailer k maximises its profit at the upper level by setting its strategies when participating in all three electricity markets. These strategies include its retail prices $\pi_k^{retail,t}$ in the retail market, its bid prices in the DAW market $\pi_k^{bid,t}$, its bid/offer prices in the LPE market $\pi_k^{LPE,t}$ and its ESS charging/discharging volume $p_e^{c,t}/p_e^{d,t}$. Subsequently, there are three lower-level problems. The first lower-level problem describes customers' welfare maximisation problem. The welfare function is formulated as the difference between customers' utility and their cost of purchasing electricity [1]. The market share function of the retailer k, as opposed to other retailers participating in the retail market, is derived after reformulating the problem, which can be embedded directly into the upper-level problem as a constraint. The ISO's DAW market-clearing problem is constructed as the second lower-level problem. The ISO receives the bid prices and electricity load demand from retailers and offers prices and generation capacities from generators to clear the DAW market. As a result, generators receive the volume of electricity that needs to be produced in each time period, while retailers receive the volume of electricity allocated to each of them. The market-clearing price of the DAW market can also be obtained. The third lower-level problem represents

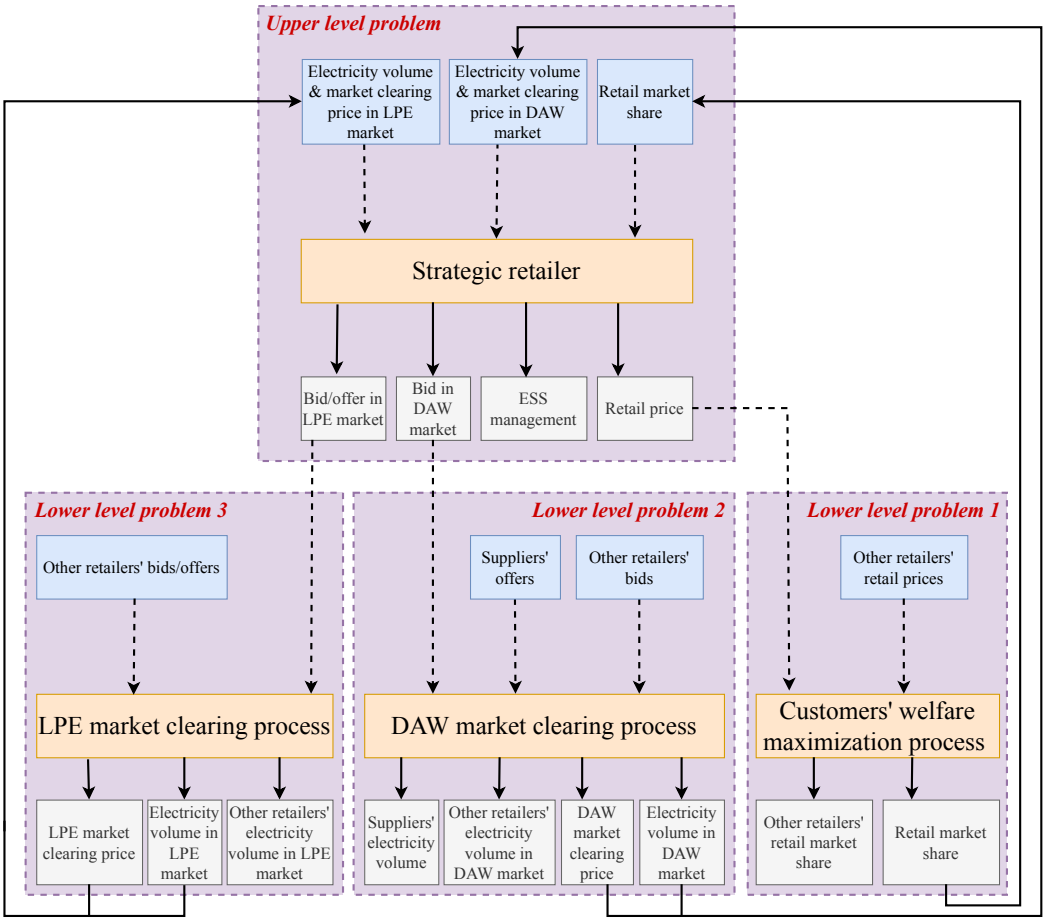


Figure 3.1: Bilevel model structure.

the LPE market-clearing problem, where the volume of electricity that each retailer needs to buy or sell is optimised. The market-clearing price of the LPE market can be derived simultaneously.

3.2.2 Upper Level Problem

The upper-level problem aims to maximise the profit of the strategic retailer k participating in the retail, DAW, and LPE markets. We assume that all three markets are operated on an hourly basis and scheduled on the same time horizon $\mathcal{T} = \{1, ..., T\}$ [25], [40]. It is also assumed that the retailer k owns the ESS, which aims to facilitate its energy operations. Mathematically, the upper-level problem is

modelled as follows:

$$\underset{\Xi_{upper}}{\text{Maximise}} \sum_{t\in\mathcal{T}} \left\{ \pi_k^{retail,t} q_k^{retail,t} - \lambda^t q_k^{bid,t} - c_k(p_k^{c,t} + p_k^{d,t})\Delta t - \lambda^{LPE,t} q_k^{LPE,t} \right\} \tag{3.1a}$$

Subject to:

$$\pi_k^{retail,min} \leq \pi_k^{retail,t} \leq \pi_k^{retail,max}, \forall t \in \mathcal{T} \tag{3.1b}$$

$$\pi_k^{bid,min} \leq \pi_k^{bid,t} \leq \pi_k^{bid,max}, \forall t \in \mathcal{T} \tag{3.1c}$$

$$\pi_k^{LPE,min} \leq \pi_k^{LPE,t} \leq \pi_k^{LPE,max}, \forall t \in \mathcal{T} \tag{3.1d}$$

$$E_k^{t+1} = E_k^t + \eta_k^c p_k^{c,t} \Delta t - \frac{1}{\eta_k^d} p_k^{d,t} \Delta t - \epsilon_k \Delta t, \forall t \in \mathcal{T} \tag{3.1e}$$

$$E_k^{min} \leq E_k^t \leq E_k^{max}, \forall t \in \mathcal{T} \tag{3.1f}$$

$$E_k^1 = E_k^{T+1} \tag{3.1g}$$

$$\gamma_k^{c,t} p_k^{c,min} \leq p_k^{c,t} \leq \gamma_k^{c,t} p_k^{c,max}, \forall t \in \mathcal{T} \tag{3.1h}$$

$$\gamma_k^{d,t} p_k^{d,min} \leq p_k^{d,t} \leq \gamma_k^{d,t} p_k^{d,max}, \forall t \in \mathcal{T} \tag{3.1i}$$

$$\gamma_k^{c,t} + \gamma_k^{d,t} \leq 1, \forall t \in \mathcal{T} \tag{3.1j}$$

$$\gamma_k^{c,t}, \gamma_k^{d,t} \in \{0, 1\}, \forall t \in \mathcal{T} \tag{3.1k}$$

$$q_k^{bid,t} + p_k^{d,t} \Delta t + q_k^{LPE,t} = q_k^{retail,t} + p_k^{c,t} \Delta t, \forall t \in \mathcal{T} \tag{3.1l}$$

The decision variables of the upper level problem are $\Xi_{upper} = \{\pi_k^{retail,t}, \pi_k^{bid,t}, \pi_k^{LPE,t}, p_k^{c,t}, p_k^{d,t}, E_k^t, \gamma_k^{c,t}, \gamma_k^{d,t}, \forall t \in \mathcal{T}\}$.

The upper-level objective function (3.1a) denotes the overall profit that the strategic retailer k can obtain. It consists of the revenue made in the retail market, the cost of purchasing electricity in the DAW market, the cost of operating the ESS, and the revenue or cost made in the LPE market. (3.1b)-(3.1d) constrain the pricing decisions of the retailer in the three markets, respectively. We define the operating constraints for the ESS following [97] [98]. In particular, (3.1e) represents the time-varying energy level of ESS. (3.1f), (3.1h), and (3.1i) ensure the energy level, charging and discharging power of the ESS at each time period follow the operational limitations. (3.1g) makes sure that by the end of the scheduling hours, the energy level of the retailer is equivalent to the initial energy level. (3.1j) and (3.1k) ensure the ESS can only be in either a charging or discharging state in a time period. (3.1l) represents the retailer's power balance constraint at each time period.

3.2.3 Lower Level Problems

The lower level of the proposed bilevel model consists of three different optimisation problems: customers' welfare maximisation problem and market-clearing problems of the DAW and LPE markets, respectively. It should be noted that we model aggregated customers' welfare and behaviour from the perspective of retailers to reflect customers' switching behaviours among different retailers. In addition, we follow [6], [18], [41], [99] in formulating the market-clearing problems by omitting the loss of direct current power flow and line congestion in transmission (i.e., DAW market) and distribution (i.e., LPE market) networks. Such a modelling choice will improve computational tractability and also allow us to focus on studying the strategic behaviours of retailers in different electricity markets.

3.2.3.1 Customers Welfare Maximisation

In the first lower-level problem, customers' satisfaction is considered and modelled as the utility function from microeconomics [3]. Following [43] [100], the utility function can be formulated as follows:

$$U(\boldsymbol{q}^{retail,t}) = \sum_{n\in\mathcal{N}} \alpha_n^t q_n^{retail,t} - \frac{1}{2}\left(\sum_{n\in\mathcal{N}} \beta_n^t q_n^{retail,t^2} + \sum_{n\in\mathcal{N}, i\in\mathcal{N}\setminus\{k\}} \beta_{n,i}^t q_n^{retail,t} q_i^{retail,t} \right) \tag{3.2a}$$

where $\mathcal{N} = \{1, ..., N\}$ represents a set of retailers in the markets. $\boldsymbol{q}^{retail,t} \in \mathcal{R}^N$ is a vector where each element denotes the electricity demand of customers from each retailer at time t. Moreover, customers' welfare is defined as the difference between the utility of all customers and the electricity purchase cost [1], which is formulated below:

$$\underset{\Xi_{lower1}}{\text{Maximise}} \sum_{t\in\mathcal{T}} \left\{ U(\boldsymbol{q}^{retail,t}) - \sum_{n\in\mathcal{N}} q_n^{retail,t} \pi_n^{retail,t} \right\} \tag{3.2b}$$

where the decision variables of the customer's welfare maximisation problem are $\Xi_{lower1} = \{q_n^{retail,t}, \forall n \in \mathcal{N}, \forall t \in \mathcal{T}\}$.

After deriving KKT optimality conditions from (3.2b), the market share function of each retailer is obtained below, which can be directly embedded in the upper-level optimisation problem of the retailer as a constraint.

$$q_n^{retail,t}(\boldsymbol{\pi}^{retail,t}) = \sum_{j\in\mathcal{N}} \omega_{n,j}^t \alpha_j^t - \omega_{n,n}^t \pi_n^{retail,t} - \sum_{j\in\mathcal{N}\setminus\{n\}} \omega_{n,j}^t \pi_j^{retail,t}, \quad \forall n \in \mathcal{N}, \forall t \in \mathcal{T} \tag{3.2c}$$

where $\boldsymbol{\pi}^{retail,t} \in \mathcal{R}^N$ is a vector that each element denotes the electricity retail price of each retailer at time t. The details of the derivation of (3.2c) can be found in the next section. In particular, elements along the main diagonal of $\boldsymbol{\Omega}^t$ (taking into account the negative sign) could be used to indicate the self-elasticity of the corresponding retailer's pricing decisions on its own customers. For instance, when the magnitude of $\omega^t_{n,n}$ becomes larger, it causes the load of customers served by the retailer n to reduce given that the unit retail price $\pi^{retail,t}_n$ increases. Furthermore, other off-diagonal elements of $\boldsymbol{\Omega}^t$ (taking into account the negative sign) could be used to indicate cross-impact effects among retail prices of different retailers, which can be interpreted as switching coefficients [43]. The switching coefficients indicate the impact on the retailer's market share when other retailers change their retail prices. A larger magnitude of the switching coefficient demonstrates a more significant impact on other retailers' retail price change to the retailer's market share. From the customers' perspective, (3.2c) implies that customers can switch among different retailers based on their offered retail prices. Specifically, customers prefer to switch to other retailers who offer lower retail prices when their subscribed retailer increases its retail price. Moreover, $\sum_{j\in\mathcal{N}} \omega^t_{n,j}\alpha^t_j$ indicates the market share potential of the retailer n, which is not affected by the price changes. It is also worth noting that (3.2c) indicates customers switch energy retailers at each time period t (e.g. on an hourly basis), which could be a viable business model in practice. This is because, with the development of information and communication technology and smart meter analytics, technical barriers to automatic and smart switching among retailers will be ultimately removed [101][102]. In addition, the proposed agile customer switching model could be modified and utilised to provide much-needed demand flexibility on short notice to help with demand and supply management (e.g. unexpected peak demand or excessive renewable generation in some hours due to forecast uncertainty).

Derivation of the market share function The combination of (3.2a) and (3.2b) can derive an unconstrained minimization problem as follows:

$$\underset{\Xi_{lower1}}{\text{Minimize}} \sum_{t\in\mathcal{T}} \left\{ \frac{1}{2}\left(\sum_{n\in\mathcal{N}} \beta^t_n q^{retail,t^2}_n + \sum_{n\in\mathcal{N}, i\in\mathcal{N}\setminus\{k\}} \beta^t_{n,i} q^{retail,t}_n q^{retail,t}_i \right) + \sum_{n\in\mathcal{N}} q^{retail,t}_n \pi^{retail,t}_n - \sum_{n\in\mathcal{N}} \alpha^t_n q^{retail,t}_n \right\} \tag{3.2.1a}$$

The first-order conditions of the objective function (3.2.1a) can be derived as:

$$\beta_n^t q_n^{retail,t} + \sum_{n\in\mathcal{N}, i\in\mathcal{N}\setminus\{n\}} +\beta_{n,i}^t q_i^{retail,t} + \pi_n^{retail,t} - \alpha_n^t = 0, \forall n \in \mathcal{N}, \forall t \in \mathcal{T} \quad (3.2.1b)$$

It can be reformulated to a compact form:

$$\boldsymbol{\pi}^{retail,t} = \boldsymbol{\alpha}^t - \boldsymbol{B}^t \boldsymbol{q}^{retail,t}, \forall t \in \mathcal{T} \quad (3.2.1c)$$

where $\boldsymbol{\alpha}^t \in \mathcal{R}^N$ is a vector that each element represents a parameter of each retailer. $\boldsymbol{B}^t \in \mathcal{R}^{N\times N}$ is a symmetric, strictly diagonally dominant matrix in which each element in a row/column represents the parameter of each retailer.

Let $\boldsymbol{\Omega}^t$ be the inverse matrix of $\boldsymbol{B}^t$, and (3.2.1c) can be reformulated as below:

$$\boldsymbol{q}^{retail,t} = \boldsymbol{\Omega}^t \boldsymbol{\alpha}^t - \boldsymbol{\Omega}^t \boldsymbol{\pi}^t, \forall t \in \mathcal{T} \quad (3.2.1d)$$

where $\boldsymbol{\Omega}^t = \begin{pmatrix} \omega_{1,1}^t & \dots & \omega_{1,N}^t \\ \dots & \dots & \dots \\ \omega_{N,1}^t & \dots & \omega_{N,N}^t \end{pmatrix}, \forall t \in \mathcal{T}$ are all symmetric matrices. Therefore, the market share function of each retailer can be derived as:

$$q_n^{retail,t} = \sum_{j\in\mathcal{N}} \omega_{n,j}^t \alpha_j^t - \omega_{n,n}^t \pi_n^{retail,t} - \sum_{j\in\mathcal{N}\setminus\{n\}} \omega_{n,j}^t \pi_j^{retail,t}, \quad \forall n \in \mathcal{N}, \forall t \in \mathcal{T} \quad (3.2.1e)$$

which is equivalent to (3.2c).

3.2.3.2 DAW Market-Clearing Problem

The ISO's DAW market-clearing problem is formulated to minimise the social cost among all generators and retailers participating in the DAW market [103]. Specifically, the bid prices $\pi_k^{bid,t}$ of the strategic retailer k are treated as known parameters in the lower-level problem. Furthermore, all generators are assumed to be non-strategic since we focus on the strategic behaviours of retailers in this chapter. The optimisation problem is therefore formulated below.

$$\underset{\Xi_{lower2}}{\text{Minimise}} \sum_{t\in\mathcal{T}} \left\{ \sum_{g\in\mathcal{G}} q_g^t c_g - \left(q_k^{bid,t} \pi_k^{bid,t} + \sum_{i\in\mathcal{N}\setminus\{k\}} q_i^{bid,t} \pi_i^{bid,t} \right) \right\} \quad (3.3a)$$

Subject to:

$$q_g^{min} \le q_g^t \le q_g^{max} : \underline{\mu_g^t}, \overline{\mu_g^t}, \forall g \in \mathcal{G}, \forall t \in \mathcal{T} \quad (3.3b)$$

$$q_k^{bid,min,t} \leq q_k^{bid,t} \leq q_k^{bid,max,t} : \underline{\zeta_k^t}, \overline{\zeta_k^t}, \forall t \in \mathcal{T} \tag{3.3c}$$

$$q_i^{bid,min,t} \leq q_i^{bid,t} \leq q_i^{bid,max,t} : \underline{\zeta_i^t}, \overline{\zeta_i^t}, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T} \tag{3.3d}$$

$$q_k^{bid,t} + \sum_{i \in \mathcal{N} \setminus \{k\}} q_i^{bid,t} - \sum_{g \in \mathcal{G}} q_g^t = 0, : \lambda^t, \forall t \in \mathcal{T} \tag{3.3e}$$

where $\Xi_{lower2} = \{q_g^t, q_k^{bid,t}, q_i^{bid,t}, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T}\}$ are the decision variables in this lower level problem. $\Xi_{lower2}^{dual} = \{\underline{\mu_g^t}, \overline{\mu_g^t}, \underline{\zeta_k^t}, \overline{\zeta_k^t}, \underline{\zeta_i^t}, \overline{\zeta_i^t}, \lambda^t, \forall g \in \mathcal{G}, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T}\}$ represents the set of dual variables of corresponding constraints.

The objective function (3.3a) minimizes the social cost of the DAW market. The production level of each generator is constrained in (3.3b). (3.3c) and (3.3d) constrain the demand level of strategic retailer k and other retailers, respectively. (3.3e) represents the electricity supply and demand balance. Furthermore, the dual variable λ^t in (3.3e) represents the market-clearing price of the DAW market.

3.2.3.3 LPE Market-Clearing Problem

The LPE market facilitates each retailer's electricity supply and demand balance. The LPE market operator acts as a non-profit entity (the same role as the ISO) and clears the LPE market as the social welfare maximisation problem. The mathematical formulation is shown as follows:

$$\underset{\Xi_{lower3}}{\text{Maximise}} \sum_{t \in \mathcal{T}} \left\{ \pi_k^{LPE,t} q_k^{LPE,t} + \sum_{i \in \mathcal{N} \setminus \{k\}} \pi_i^{LPE,t} q_i^{LPE,t} \right\} \tag{3.4a}$$

Subject to:

$$-q_{k,out}^{LPE,max,t} \leq q_k^{LPE,t} \leq q_{k,in}^{LPE,max,t} : \psi_{k,out}^t, \psi_{k,in}^t, \forall t \in \mathcal{T} \tag{3.4b}$$

$$-q_{i,out}^{LPE,max,t} \leq q_i^{LPE,t} \leq q_{i,in}^{LPE,max,t} : \sigma_{i,out}^t, \sigma_{i,in}^t, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T} \tag{3.4c}$$

$$\sum_{i \in \mathcal{N} \setminus \{k\}} q_i^{LPE,t} + q_k^{LPE,t} = 0 : \lambda^{LPE,t}, \forall t \in \mathcal{T} \tag{3.4d}$$

where the decision variables are $\Xi_{lower3} = \{q_k^{LPE,t}, q_i^{LPE,t}, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T}\}$. The dual variables of corresponding constraints are denoted as $\Xi_{lower3}^{dual} = \{\psi_{k,out}^t, \psi_{k,in}^t, \sigma_{i,out}^t, \sigma_{i,in}^t, \lambda^{LPE,t}, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T}\}$.

The objective function (3.4a) maximises the social welfare of the LPE market. (3.4b) and (3.4c) ensure the volume of electricity that each retailer buys or sells in the LPE market is bounded. Finally, (3.4d) represents the power balance constraint. The dual variable $\lambda^{LPE,t}$ represents the market-clearing price of the LPE market.

3.2.4 Bilevel Model

After formulating both the upper- and lower-level problems, the proposed bilevel model for the strategic retailer k can be summarised as follows.

$$\Xi_{upper} \in \arg\underset{\Xi_{upper}}{\text{maximise}} \quad \text{(3.1a)} \tag{3.5a}$$

Subject to:

$$\text{Constraints (3.1b)-(3.1l)} \tag{3.5b}$$

$$\Xi_{lower1} \in \arg\underset{\Xi_{lower1}}{\text{maximise}} \quad \text{(3.2b)} \tag{3.5c}$$

$$\Xi_{lower2}, \underline{\mu_g^t}, \overline{\mu_g^t}, \underline{\zeta_k^t}, \overline{\zeta_k^t}, \underline{\zeta_i^t}, \overline{\zeta_i^t}, \lambda^t \in \arg\underset{\Xi_{lower2}}{\text{minimise}} \Bigg\{ \text{(3.3a)}$$

$$\text{Subject to:} \tag{3.5d}$$

$$\text{Constraints(3.3b)-(3.3e)} \Bigg\}, \forall g \in \mathcal{G}, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T}$$

$$\Xi_{lower3}, \psi_{k,out}^t, \psi_{k,in}^t, \sigma_{i,out}^t, \sigma_{i,in}^t, \lambda^{LPE,t} \in \arg\underset{\Xi_{lower3}}{\text{maximise}} \Bigg\{ \text{(3.4a)}$$

$$\text{Subject to:} \tag{3.5e}$$

$$\text{Constraints(3.4b)-(3.4d)} \Bigg\}, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T}$$

(3.5a) and (3.5b) denote the strategies of the retailer k at the upper level. Furthermore, (3.5c)-(3.5e) represent the reactions from the three electricity markets given by the upper-level decisions, respectively. The bilevel model forms a single-leader-multiple-follower Stackelberg game, which can also be interpreted as an MPEC program [2]. The methods to solve the MPEC problem are discussed in detail in the next section.

3.3 Solution Methods

This section illustrates the solution methods for MPEC and EPEC problems. It first details the treatment of the MPEC problem, which is linearised and reformulated to an MIQP problem. Furthermore, the single-leader MPEC model is extended to the multi-leader EPEC model, which can be solved by the diagonalisation algorithm.

3.3.1 MPEC Problem

The bilevel model can be transformed into a single-level MPEC problem by deriving KKT optimality conditions for the lower-level problems into a system of equations and inequalities. The transformed MPEC problem is shown below:

$$\underset{\Xi_{MPEC}}{\text{Maximise}} \quad \text{(3.1a)} \tag{3.6a}$$

Subject to:

$$\text{Constraints (3.1b)-(3.1l), (3.2c)} \tag{3.6b}$$

$$c_g - \underline{\mu_g^t} + \overline{\mu_g^t} - \lambda^t = 0, \forall g \in \mathcal{G}, \forall t \in \mathcal{T} \tag{3.6c}$$

$$-\pi_k^{bid,t} - \underline{\zeta_k^t} + \overline{\zeta_k^t} + \lambda^t = 0, \forall t \in \mathcal{T} \tag{3.6d}$$

$$-\pi_i^{bid,t} - \underline{\zeta_i^t} + \overline{\zeta_i^t} + \lambda^t = 0, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T} \tag{3.6e}$$

$$q_k^{bid,t} + \sum_{i \in \mathcal{N} \setminus \{k\}} q_i^{bid,t} - \sum_{g \in \mathcal{G}} q_g^t = 0, \forall t \in \mathcal{T} \tag{3.6f}$$

$$0 \leq (q_g^t - q_g^{min}) \perp \underline{\mu_g^t} \geq 0, \forall g \in \mathcal{G}, \forall t \in \mathcal{T} \tag{3.6g}$$

$$0 \leq (q_g^{max} - q_g^t) \perp \overline{\mu_g^t} \geq 0, \forall g \in \mathcal{G}, \forall t \in \mathcal{T} \tag{3.6h}$$

$$0 \leq (q_n^{bid,t} - q_n^{bid,min,t}) \perp \underline{\zeta_n^t} \geq 0, \forall n \in \mathcal{N}, \forall t \in \mathcal{T} \tag{3.6i}$$

$$0 \leq (q_n^{bid,max,t} - q_n^{bid,t}) \perp \overline{\zeta_n^t} \geq 0, \forall n \in \mathcal{N}, \forall t \in \mathcal{T} \tag{3.6j}$$

$$-\pi_k^{LPE,t} - \psi_{k,out}^t + \psi_{k,in}^t + \lambda^{LPE,t} = 0, \forall t \in \mathcal{T} \tag{3.6k}$$

$$-\pi_i^{LPE,t} - \sigma_{i,out}^t + \sigma_{i,in}^t + \lambda^{LPE,t} = 0, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T} \tag{3.6l}$$

$$\sum_{i \in \mathcal{N} \setminus \{k\}} q_i^{LPE,t} + q_k^{LPE,t} = 0, \forall t \in \mathcal{T} \tag{3.6m}$$

$$0 \leq \psi_{k,out}^t \perp (q_k^{LPE,t} - q_{k,out}^{LPE,max,t}) \geq 0, \forall t \in \mathcal{T} \tag{3.6n}$$

$$0 \leq \psi_{k,in}^t \perp (q_{k,in}^{LPE,max,t} - q_k^{LPE,t}) \geq 0, \forall t \in \mathcal{T} \tag{3.6o}$$

$$0 \leq \sigma_{i,out}^t \perp (q_i^{LPE,t} - q_{i,out}^{LPE,max,t}) \geq 0, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T} \tag{3.6p}$$

$$0 \leq \sigma_{i,in}^t \perp (q_{i,in}^{LPE,max,t} - q_i^{LPE,t}) \geq 0, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T} \tag{3.6q}$$

where the decision variables of the MPEC problem are $\Xi_{MPEC} = \Big\{\pi_k^{retail,t}, \pi_k^{bid,t}, q_k^{retail,t}, q_i^{retail,t}, q_k^{bid,t}, q_i^{bid,t}, p_k^{c,t}, p_k^{d,t}, E_k^t, \pi_k^{LPE,t}, q_k^{LPE,t}, q_i^{LPE,t}, q_g^t, \gamma_k^{c,t}, \gamma_k^{d,t}, \underline{\mu_g^t}, \overline{\mu_g^t}, \underline{\zeta_j^t}, \overline{\zeta_j^t}, \lambda^t, \psi_{k,out}^t, \psi_{k,in}^t, \sigma_{i,out}^t, \sigma_{i,in}^t, \forall t \in \mathcal{T}, \forall i \in \mathcal{N} \setminus \{k\}, \forall g \in \mathcal{G}, \forall j \in \mathcal{N}\Big\}$.

(3.6a) denotes the objective function of the MPEC model. In the following constraints, (3.6b) represents a collection of constraints from the upper-level problem and retailers' market share function. Equations (3.6c)-(3.6f) and (3.6k)-(3.6m) are stationary conditions of the KKT optimality conditions. Moreover, (3.6g)-(3.6j) and (3.6n)-(3.6q) represent the complementarity slackness.

3.3.2 Linearisation of the MPEC Problem

The MPEC model above is non-convex and difficult to solve due to the existence of bilinear terms in the objective function (3.6a) and complementarity slackness constraints (3.6g)-(3.6j) and (3.6n)-(3.6q). To overcome the difficulties, we first deal with the bilinear terms in the upper-level objective function (3.6a) through the strong duality theorem [104]. Therefore, the objective function of the MPEC model becomes:

$$\begin{aligned}\Phi = \sum_{t\in\mathcal{T}}\Bigg\{\sum_{g\in\mathcal{G}}\Big(q_g^t c_g - \underline{\mu_g^t} q_g^{min} + \overline{\mu_g^t} q_g^{max}\Big) - \sum_{j\in\mathcal{N}\setminus\{k\}}\Big(\pi_j^{bid,t} q_j^{bid,t} + \underline{\zeta_j^t} q_j^{bid,min} - \\ \overline{\zeta_j^t} q_j^{bid,max}\Big) + c_k\Big(p_k^{c,t} + p_k^{d,t}\Big)\Delta t - \pi_k^{retail,t}\sum_{j\in\mathcal{N}}\omega_{k,j}^t\alpha_j^t + \omega_k^t\pi_k^{retail,t^2} + \\ \pi_k^{retail,t}\sum_{j\in\mathcal{N}\setminus\{k\}}\omega_{k,j}^t\pi_j^{retail,t} + \sum_{i\in\mathcal{N}\setminus\{k\}}\Big(\sigma_{i,out}^t q_{i,out}^{LPE,max,t} + \sigma_{i,in}^t q_{i,in}^{LPE,max,t} - \\ \pi_i^{LPE,t} q_i^{LPE,t}\Big)\Bigg\}\end{aligned} \tag{3.7a}$$

The details of the derivation of objective function Φ are provided in the next section below.

Linearisation of the Objective Function of MPEC

Reformulation of bilinear terms The Lagrange function of the minimisation problem (3.3a)-(3.3e) is formulated as follows.

$$\mathcal{L}(\Xi_{lower2}|\Xi_{lower2}^{dual}) = \sum_{t\in\mathcal{T}}\Bigg\{\sum_{g\in\mathcal{G}} q_g^t c_g - \Bigg(q_k^{bid,t}\pi_k^{bid,t} + \sum_{i\in\mathcal{N}\setminus\{k\}} q_i^{bid,t}\pi_i^{bid,t}\Bigg)\Bigg\} \tag{3.7.1a}$$

$$+\sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{G}}\left(\underline{\mu_g^t}\left(q_g^{min}-q_g^t\right)+\overline{\mu_g^t}\left(q_g^t-q_g^{max}\right)\right)+\sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{N}}\left(\underline{\zeta_i^t}\left(q_i^{bid,min,t}-q_i^{bid,t}\right)+\overline{\zeta_i^t}\left(q_i^t-q_i^{bid,max,t}\right)\right)+\sum_{t\in\mathcal{T}}\left(\lambda^t\left(\sum_{i\in\mathcal{N}}q_i^{bid,t}-\sum_{g\in\mathcal{G}}q_g^t\right)\right)$$

Then, the dual program can be derived below:

$$\underset{\Xi_{lower2}^{dual}}{\text{Maximize}}\sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{G}}\left(\underline{\mu_g^t}q_g^{min}-\overline{\mu_g^t}q_g^{max}\right)+\sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{N}}\left(\underline{\zeta_i^t}q_i^{bid,min,t}-\overline{\zeta_i^t}q_i^{bid,max,t}\right) \tag{3.7.1b}$$

Subject to:

$$c_g-\underline{\mu_g^t}+\overline{\mu_g^t}-\lambda^t=0,\forall g\in\mathcal{G},\forall t\in\mathcal{T} \tag{3.7.1c}$$

$$-\pi_i^{bid,t}-\underline{\zeta_i^t}+\overline{\zeta_i^t}+\lambda^t=0,\forall i\in\mathcal{N},\forall t\in\mathcal{T} \tag{3.7.1d}$$

Since the primal program (3.3a)-(3.3e) is a linear program, the strong duality theorem holds. This indicates that the value of the primal objective function (3.3a) is the same as the value of the dual objective function (3.7.1b). Therefore, we can then obtain a system of equations:

$$\text{Objective function (3.3a) = Objective function (3.7.1b)} \tag{3.7.1e}$$

$$\text{Constraints (3.6d), (3.6e)} \tag{3.7.1f}$$

$$\underline{\zeta_i^t}(q_i^{bid,min,t}-q_i^{bid,t})=0,\forall i\in\mathcal{N},\forall t\in\mathcal{T} \tag{3.7.1g}$$

$$\overline{\zeta_i^t}(q_i^{bid,t}-q_i^{bid,max,t})=0,\forall i\in\mathcal{N},\forall t\in\mathcal{T} \tag{3.7.1h}$$

After solving the system of equations (3.7.1e)-(3.7.1h), we can derive the equality below.

$$\sum_{t\in\mathcal{T}}\lambda^t q_k^{bid,t}=\sum_{t\in\mathcal{T}}\sum_{g\in\mathcal{G}}\left\{q_g^t c_g-\underline{\mu_g^t}q_g^{min}+\overline{\mu_g^t}q_g^{max}\right\}-\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{N}\setminus\{k\}}\left\{\pi_j^{bid,t}q_j^{bid,t}+\underline{\zeta_j^t}q_j^{bid,min,t}-\overline{\zeta_j^t}q_j^{bid,max,t}\right\} \tag{3.7.1i}$$

Analogously, the Lagrange function of the problem (3.4a)-(3.4d) is formulated as follows.

$$\begin{aligned}\mathcal{L}(\Xi_{lower3}|\Xi^{dual}_{lower3}) = &\sum_{t\in\mathcal{T}}\left\{\pi^{LPEM,t}_k q^{LPEM,t}_k + \sum_{i\in\mathcal{N}\setminus\{k\}} \pi^{LPEM,t}_i q^{LPEM,t}_i\right\}\\ &+\sum_{t\in\mathcal{T}}\left\{\psi^t_{k,out}\left(q^{LPEM,t}_k + q^{LPEM,max,t}_k\right) + \psi^t_{k,in}\left(q^{LPEM,max,t}_{k,in} - q^{LPEM,t}_k\right)\right\}\\ &+\sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{N}\setminus\{k\}}\left\{\sigma^t_{i,out}\left(q^{LPEM,t}_i + q^{LPEM,max,t}_{i,out}\right) + \sigma^t_{i,in}\left(q^{LPEM,max,t}_{i,in}\right.\right.\\ &\left.\left.-q^{LPEM,t}_i\right)\right\} - \sum_{t\in\mathcal{T}}\left\{\lambda^{LPEM,t}\left(\sum_{i\in\mathcal{N}\setminus\{k\}} q^{LPEM,t}_i + q^{LPEM,t}_k\right)\right\}\end{aligned} \tag{3.7.1j}$$

The dual program of (3.4a)-(3.4d) is derived below.

$$\begin{aligned}\underset{\Xi^{dual}_{lower3}}{\text{Minimize}} &\sum_{t\in\mathcal{T}}\left\{q^{LPEM,max,t}_{k,out}\psi^t_{k,out} + q^{LPEM,max,t}_{k,in}\psi^t_{k,in}\right\}\\ &+\sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{N}\setminus\{k\}}\left\{\sigma^t_{i,out}q^{LPEM,max,t}_{i,out} + \sigma^t_{i,in}q^{LPEM,max,t}_{i,in}\right\}\end{aligned} \tag{3.7.1k}$$

Subject to:

$$\pi^{LPEM,t}_k + \psi^t_{k,out} - \psi^t_{k,in} - \lambda^{LPEM,t} = 0, \forall t \in \mathcal{T} \tag{3.7.1l}$$

$$\pi^{LPEM,t}_i + \sigma^t_{i,out} - \sigma^t_{i,in} - \lambda^{LPEM,t}, \forall i \in \mathcal{N}\setminus\{k\}, \forall t \in \mathcal{T} \tag{3.7.1m}$$

The primal program (3.4a)-(3.4d) is also a linear program. Therefore, the strong duality theorem holds. A system of equations can be derived as follows.

$$\text{Objective function (3.4a) = Objective function (3.7.1k)} \tag{3.7.1n}$$

Constraint (3.7.1l)

$$\psi^t_{k,out}(q^{LPEM,t}_k + q^{LPEM,max,t}_k) = 0, \forall t \in \mathcal{T} \tag{3.7.1o}$$

$$\psi^t_{k,in}(q^{LPEM,max,t}_{k,in} - q^{LPEM,t}_k) = 0, \forall t \in \mathcal{T} \tag{3.7.1p}$$

A solution of the system of equations (3.7.1n)-(3.7.1p), and (3.7.1l) is shown below.

$$\begin{aligned}\sum_{t\in\mathcal{T}}\lambda^{LPEM,t}q^{LPEM,t}_k = \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{N}\setminus\{k\}}&\left\{\sigma^t_{i,out}q^{LPEM,max,t}_{i,out} + \sigma^t_{i,in}q^{LPEM,max,t}_{i,in}\right.\\ &\left.-\pi^{LPEM,t}_i q^{LPEM,t}_i\right\}\end{aligned} \tag{3.7.1q}$$

Reformulation of the objective function of MPEC There are three bilinear terms in the objective function of the MPEC program, which are $\lambda^t q_k^{bid,t}$, $\lambda^{LPEM,t} q_k^{LPEM,t}$ and $\pi_k^{retail,t} q_k^{retail,t}$. The first two bilinear terms are linearised in (3.7.1i) and (3.7.1q), respectively. The last bilinear term can be linearised by substituting $\sum_{j\in\mathcal{N}} \omega_{k,j}^t \alpha_j^t - \omega_{k,k}^t \pi_k^{retail,t} - \sum_{j\in\mathcal{N}\setminus\{k\}} \omega_{k,j}^t \pi_j^{retail,t}$ for $q_k^{retail,t}$ based on (3.2c).

After linearising the bilinear terms, the final objective function of the MPEC program is derived as follows.

$$\begin{aligned}\Phi = \sum_{t\in\mathcal{T}} \Bigg\{ & \sum_{g\in\mathcal{G}} \left(q_g^t c_g - \underline{\mu_g^t} q_g^{min} + \overline{\mu_g^t} q_g^{max} \right) - \sum_{j\in\mathcal{N}\setminus\{k\}} \Big(\pi_j^{bid,t} q_j^{bid,t} + \underline{\zeta_j^t} q_j^{bid,min} \\ & - \overline{\zeta_j^t} q_j^{bid,max} \Big) + c_k \left(p_k^{c,t} + p_k^{d,t} \right) \Delta t - \pi_k^{retail,t} \sum_{j\in\mathcal{N}} \omega_{k,j}^t \alpha_j^t + \omega_k^t \pi_k^{retail,t^2} \\ & + \pi_k^{retail,t} \sum_{j\in\mathcal{N}\setminus\{k\}} \omega_{k,j}^t \pi_j^{retail,t} + \sum_{i\in\mathcal{N}\setminus\{k\}} \Big(\sigma_{i,out}^t q_{i,out}^{LPEM,max,t} + \sigma_{i,in}^t q_{i,in}^{LPEM,max,t} \\ & - \pi_i^{LPEM,t} q_i^{LPEM,t} \Big) \Bigg\} \end{aligned} \tag{3.7.2a}$$

Furthermore, the Fortuny-Amat transformation is used to linearise complementarity slackness by introducing additional binary variables and a relatively large integer constant M [105]. The resulting linearised constraints of (3.6g)-(3.6j) and (3.6n)-(3.6q) are shown in (3.7a)-(3.7j) and (3.7k)-(3.7t), respectively.

$$0 \le \underline{\mu_g^t} \le \underline{\iota_g^t} M, \forall g \in \mathcal{G}, \forall t \in \mathcal{T} \tag{3.7b}$$

$$0 \le q_g^t - q_g^{min} \le (1 - \underline{\iota_g^t}) M, \forall g \in \mathcal{G}, \forall t \in \mathcal{T} \tag{3.7c}$$

$$0 \le \overline{\mu_g^t} \le \overline{\iota_g^t} M, \forall g \in \mathcal{G}, \forall t \in \mathcal{T} \tag{3.7d}$$

$$0 \le q_g^{max} - q_g^t \le (1 - \overline{\iota_g^t}) M, \forall g \in \mathcal{G}, \forall t \in \mathcal{T} \tag{3.7e}$$

$$\underline{\iota_g^t}, \overline{\iota_g^t} \in \{0, 1\}, \forall g \in \mathcal{G}, \forall t \in \mathcal{T} \tag{3.7f}$$

$$0 \le \underline{\zeta_i^t} \le \underline{\xi_i^t} M, \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \tag{3.7g}$$

$$0 \le q_i^{bid,t} - q_i^{bid,min,t} \le (1 - \underline{\xi_i^t}) M, \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \tag{3.7h}$$

$$0 \le \overline{\zeta_i^t} \le \overline{\xi_i^t} M, \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \tag{3.7i}$$

$$0 \le q_i^{bid,max,t} - q_i^{bid,t} \le (1 - \overline{\xi_i^t}) M, \forall i \in \mathcal{N}, \forall t \in \mathcal{T} \tag{3.7j}$$

$$\underline{\xi_g^t}, \overline{\xi_g^t} \in \{0, 1\}, \forall g \in \mathcal{G}, \forall t \in \mathcal{T} \tag{3.7k}$$

$$0 \le \psi_{k,out}^t \le \rho_{k,out}^t M, \forall t \in \mathcal{T} \tag{3.7l}$$

$$0 \leq q_k^{LPE,t} + q_{k,out}^{LPE,max,t} \leq (1 - \rho_{k,out}^t)M, \forall t \in \mathcal{T} \tag{3.7m}$$

$$0 \leq \psi_{k,in}^t \leq \rho_{k,in}^t M, \forall t \in \mathcal{T} \tag{3.7n}$$

$$0 \leq q_{k,in}^{LPE,max,t} - q_k^{LPE,t} \leq (1 - \rho_{k,in}^t)M, \forall t \in \mathcal{T} \tag{3.7o}$$

$$\rho_{k,out}^t, \rho_{k,in}^t \in \{0,1\}, \forall t \in \mathcal{T} \tag{3.7p}$$

$$0 \leq \sigma_{i,out}^t \leq \delta_{i,out}^t M, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T} \tag{3.7q}$$

$$0 \leq q_i^{LPE,t} + q_{i,out}^{LPE,max,t} \leq (1 - \delta_{i,out}^t)M, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T} \tag{3.7r}$$

$$0 \leq \sigma_{i,in}^t \leq \delta_{i,in}^t M, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T} \tag{3.7s}$$

$$0 \leq q_{i,in}^{LPE,max,t} - q_i^{LPE,t} \leq (1 - \delta_{i,in}^t)M, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T} \tag{3.7t}$$

$$\delta_{i,out}^t, \delta_{i,in}^t \in \{0,1\}, \forall i \in \mathcal{N} \setminus \{k\}, \forall t \in \mathcal{T} \tag{3.7u}$$

3.3.3 MIQP Problem

After the linearisation, the MPEC model is reformulated into an MIQP problem and can be solved efficiently using off-the-shelf solvers. The complete MIQP model is formulated as follows.

$$\underset{\Xi_{MIQP}}{\text{Minimise}} \quad \Phi \tag{3.8a}$$

Subject to:

$$\text{Constraints (3.1b)-(3.1l), (3.2c), (3.6c)-(3.6f), (3.6k)-(3.6m), (3.7a)-(3.7t)} \tag{3.8b}$$

where $\Xi_{MIQP} = \Big\{\pi_k^{retail,t}, \pi_k^{bid,t}, q_k^{retail,t}, q_i^{retail,t}, q_k^{bid,t}, q_i^{bid,t}, p_k^{c,t}, p_k^{d,t}, E_k^t, \pi_k^{LPE,t}, q_k^{LPE,t}, q_i^{LPE,t}, q_g^t, \gamma_k^{c,t}, \gamma_k^{d,t}, \tau_{in}^t, \tau_{out}^t, \underline{\mu_g^t}, \overline{\mu_g^t}, \underline{\zeta_j^t}, \overline{\zeta_j^t}, \underline{\iota_g^t}, \overline{\iota_g^t}, \underline{\xi_j^t}, \overline{\xi_j^t}, \lambda^t, \psi_{k,out}^t, \psi_{k,in}^t, \sigma_{i,out}^t, \sigma_{i,in}^t, \rho_{k,out}^t, \rho_{k,in}^t, \delta_{i,out}^t, \delta_{i,in}^t, \forall t \in \mathcal{T}, \forall i \in \mathcal{N} \setminus \{k\}, \forall g \in \mathcal{G}, \forall j \in \mathcal{N}\Big\}$ represents the set of decision variables of the MIQP model.

The objective function (3.8a) shapes a quadratic form with respect to $\pi_k^{retail,t}$. Constraint (3.8b) consists of all the constraints in the upper-level problem, market share functions, KKT stationary conditions for the market-clearing problems of the DAW and LPE markets, and the linearised complementarity slackness constraints.

3.3.4 EPEC Problem

The Bertrand competition model is utilised to extend the MIQP model from a single strategic retailer to multiple strategic retailers. This results in a multi-leader, multi-follower Stackelberg game and can be reformulated as an EPEC problem [2], which is illustrated in Figure 3.2. Although the retailers share complete information among themselves in the theoretical setting of the EPEC problem, in practice,

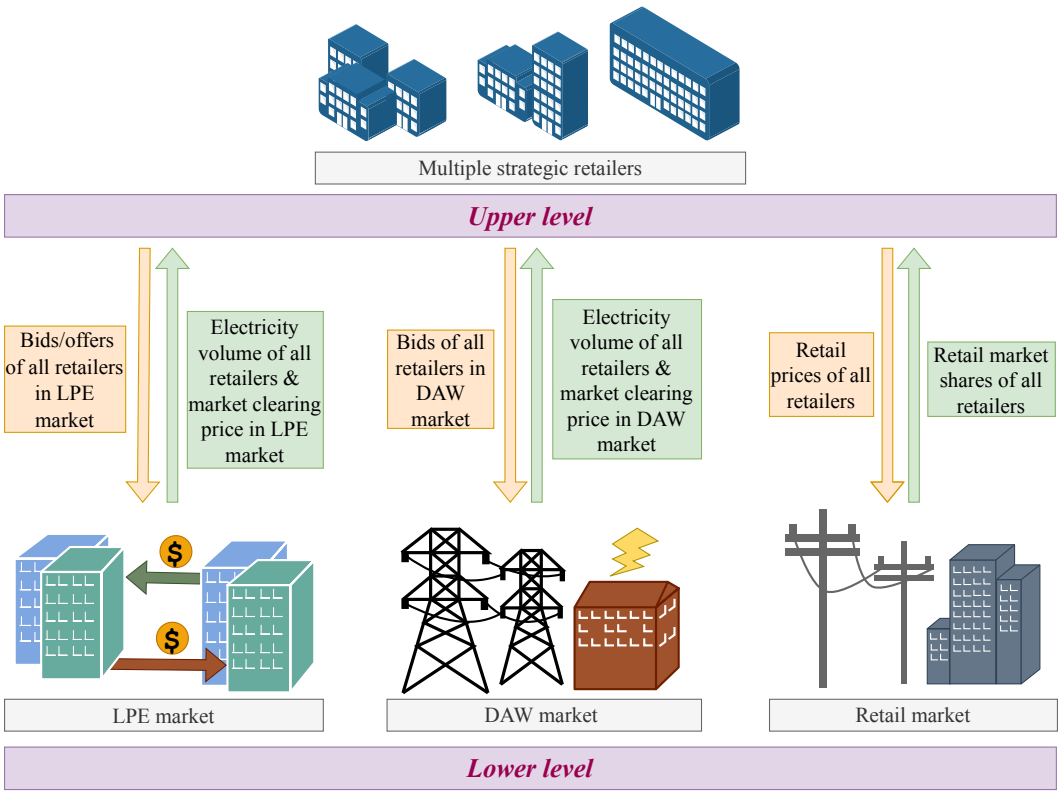


Figure 3.2: EPEC problem structure.

an independent market agent (e.g. ISO for wholesale markets) can play such a role in sharing required information among retailers. We adopt the diagonalisation algorithm in [106] to tackle our formulated EPEC problem where the converged strategies of strategic retailers represent a generalised Nash equilibrium. The diagonalisation algorithm considered for solving our EPEC problem is outlined in Algorithm 1. In Step 1, the strategy set is initialised as $\boldsymbol{S}^0$. The maximum iteration Y and convergence criterion ϵ are also predefined. The main iteration procedure of the diagonalisation algorithm is shown in Steps 2 – 13, which consists of an outer loop and an inner loop. In particular, the outer loop controls the iteration of the algorithm. For each iteration of the outer loop, Steps 3 – 6 define the inner loop and aim to solve the MIQP problem for each strategic retailer sequentially with the other retailers' strategies as given parameters. The convergence of the algorithm is checked in Steps 7 – 12 at each iteration of the outer loop. Specifically, in Steps 7 – 9, if the difference between the retailers' optimal decisions of two adjacent iterations is less than ϵ, the algorithm converges and terminates with retailers' optimal decisions. However, in Steps 10 – 12, if the algorithm reaches the maximum iteration Y without

Algorithm 1 Diagonalisation algorithm

Initialisation:

$$\boldsymbol{\mathcal{S}}^0 = \left\{\pi_n^{retail,t}, \pi_n^{bid,t}, \pi_n^{LPE,t}, p_n^{c,t}, p_n^{d,t}, E_n^t, \gamma_n^{c,t}, \gamma_n^{d,t}, \forall n \in \mathcal{N}, \forall t \in \mathcal{T}\right\};$$

maximum number of iterations Y; convergence criterion ϵ.

for $y = 1$ to Y **do**
  for $i = 1$ to N **do**
    Solve strategic retailer i's MIQP model assuming other retailers' strategies as given parameters.
    Update $\boldsymbol{\mathcal{S}}_i^y$;
  end for
  if $\|\boldsymbol{\mathcal{S}}_i^y - \boldsymbol{\mathcal{S}}_i^{y-1}\| \leq \epsilon, \forall i \in \mathcal{N}$ **then**
    The algorithm converges and terminates.
  end if
  if $y = Y$ **then**
    The algorithm fails to converge and terminates.
  end if
end for

convergence, it terminates and no optimal results are found.

3.4 Numerical Results

Numerical results are illustrated in this section to demonstrate the feasibility and effectiveness of the EPEC model and the diagonalisation algorithm. The effects of customers' switching behaviours and the number of strategic retailers on retail competition are discussed in detail. The proposed model is solved by Gurobi Optimiser (version 9.5.2) using the branch and bound algorithm under Pyomo [107] on Windows 10 Enterprise 64-bit with 4 cores CPU at 3.6GHz and 16GB of RAM.

3.4.1 Experimental Setup

Data used in this section comes mainly from the PJM datasets [108], such as the initial retail and DAW market bid prices for each retailer during the day. The initial LPE market bid/offer prices and the maximum of cleared electricity volume are based on PJM real-time market bid prices and cleared electricity for each retailer. We further calibrate the retailers' maximum cleared electricity volume in the LPE market to be 5% of the maximum bid load of retailers in the DAW market. The initial DAW market's maximum bid load of each retailer comes from the PJM DAW market bid load of different utility companies. In addition, the minimum and maximum retail, DAW market bid, and LPE market bid/offer prices are all set to be \$0/MWh

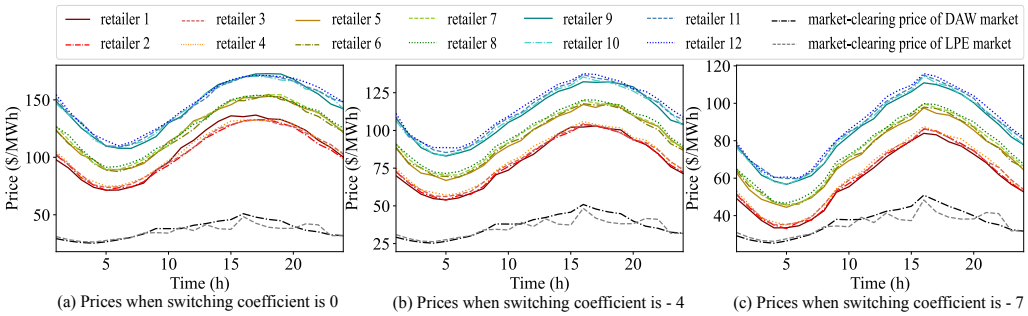


Figure 3.3: Time-varying retail prices and market-clearing prices of LPE and DAW markets with different switching coefficients.

and \$300/MWh, respectively. The minimum bid load for the retailers in the DAW market is considered to be 0.1 MW following the PJM day-ahead wholesale market [109]. The maximum iteration $Y = 30$ and the termination criteria $\epsilon = 1$ are chosen for the diagonalisation algorithm. The ESS-related parameters are modified based on [98]. In particular, the initial ESS energy level is set to be $80MWh$. The maximum and minimum charging and discharging rates are $60MW$ and $2MW$. The maximum and minimum ESS energy capacities are $200MWh$ and $30MWh$. The charging and discharging efficiencies are set to be 0.9. Lastly, the self-discharge rate $\epsilon_k = 0.002MW$ is considered.

In this chapter, we consider 24 time periods for the day starting from midnight. That is, each time period represents an hour. In this case, 12 strategic retailers are considered in the proposed EPEC model. Furthermore, the strategic retailers are classified into 3 groups based on their market share potential which the self-elasticity coefficient $\omega^t_{n,n}$ and parameter α^t_n are assumed to be time-varying. Specifically, retailers $1-4$ are classified into small market share groups (group 1). Retailers $5-8$ belong to the medium market share group (group 2). Lastly, retailers $9-12$ are in the large market share group (group 3). The input data of electricity prices and volume, self-elasticity coefficient $\omega^t_{n,n}$, and α^t_n for each retailer can be found in A.1.1. Additionally, we include 30 generators participating in the DAW market. The cost and maximum supply of each generator are shown in A.1.2.

3.4.2 Illustrative Examples

In this section, illustrative examples are given to discuss the results of the EPEC model when switching coefficients are set to be $0MWh/\$$, $-4MWh/\$$, and $-7MWh/\$$ respectively. The magnitude of the switching coefficient represents the ability of

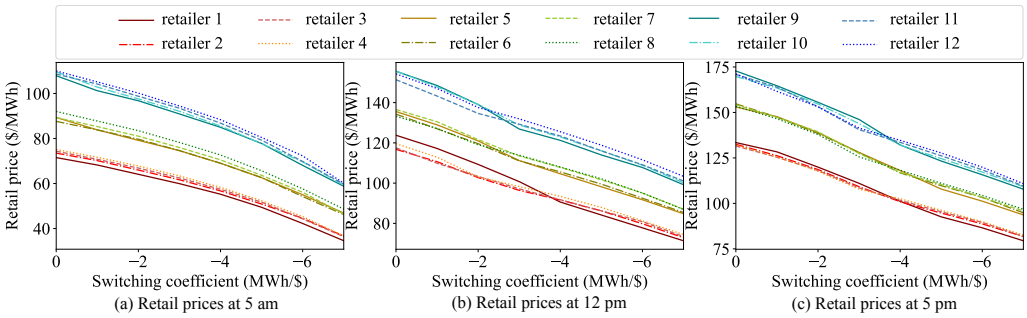


Figure 3.4: Retail prices of retailers with different switching coefficients at different times of the day.

customers to switch to other retailers and thus the competition level in the retail market. A larger magnitude of the switching coefficient indicates more competition in the retail market. Time-varying retail prices of each retailer and market-clearing prices of the LPE and DAW markets are shown in Figure 3.3. It can be found that both the retail and market-clearing prices decrease from 1 am to around 5 am, then increase until around 5 pm and drop down again afterwards, which follows customers' demand during the day.

Furthermore, the retail prices of all retailers are generally higher than the market-clearing prices of the LPE and DAW markets but become closer to the market-clearing prices with the increase of the magnitude of the switching coefficient. This can be explained that more competition in the retail market drives down the retail prices, and retailers' profit margins become lower. In addition, the retail prices are typically higher when the retailers have a larger market share (bigger retailers). This could be due to the fact that retailers with large market shares have more flexibility in their pricing decisions without worrying about losing customers.

It is also observed that the market-clearing prices of the LPE market are generally more volatile than the market-clearing prices of the DAW market. This could be explained by the fact that the market size (market-cleared electricity volume) of the LPE market is much smaller than the DAW market. Therefore, the unit difference in customers' demand has a more significant impact on the LPE market, which results in higher volatility of its market-clearing prices.

3.4.3 Retail Prices and Profits

Figure 3.4 presents the equilibrium retail prices among all strategic retailers given by different switching coefficients at 5 am, 12 pm, and 5 pm, respectively. It shows

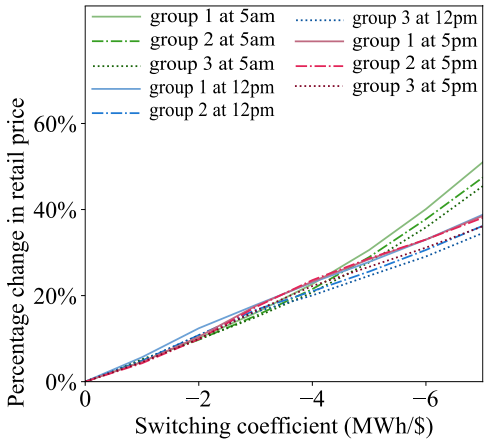


Figure 3.5: Percentage change in retail prices with different switching coefficients.

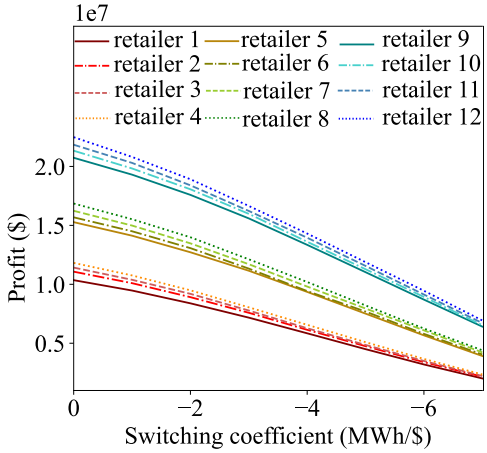


Figure 3.6: Profit of retailers with different switching coefficients.

that when the magnitude of the switching coefficient becomes larger, the retail prices among all retailers decrease dramatically. This is because retailers would like to reduce their retail prices to prevent customer losses as customers are more capable of switching their electricity retailers.

Moreover, the percentage changes in average retail prices of different retailer groups at 5 am, 12 pm, and 5 pm are shown against different switching coefficients in Figure 3.5. From the figure, we can find that with the increase of the magnitude of the switching coefficient, average retail prices of all retailer groups decrease consistently for different time periods. It should also be noted that when the magnitude is less than 4MWh/$, there is not much difference in price changes among the three retailer groups at different time periods. However, following the continuing increase of the magnitude, the price changes differ in different retailer groups and different time

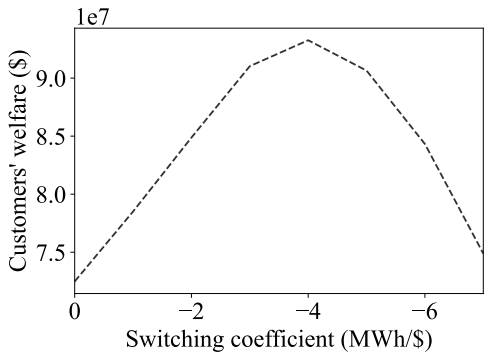


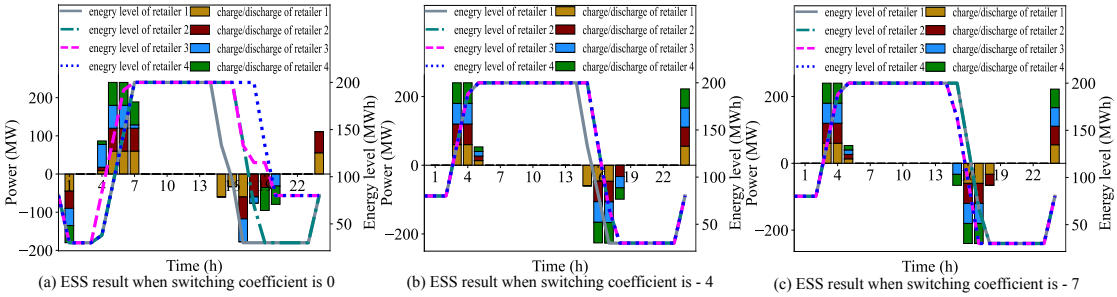
Figure 3.7: Customers' welfare with different switching coefficients.

periods. For instance, the price changes in all three retailer groups at 5 am are much higher than in other time periods. In addition, the price change of the small retailer group (e.g. group 1) is larger than that of the large retailer group (e.g. group 3). The above two phenomena enforce our findings that the switching coefficients have a larger impact on the prices of small retailers and low-demand time periods.

The impact of the switching coefficient on the profits of retailers is illustrated in Figure 3.6. Not surprisingly, the retailers' profits reduce significantly when increasing the magnitude of the switching coefficient. In addition, although the profits of bigger retailers are usually higher, the profit difference among retailers tends to decrease when the magnitude of the switching coefficient becomes larger (higher competition in the market). In other words, a market with higher competition provides a healthier environment for small players/ entrants.

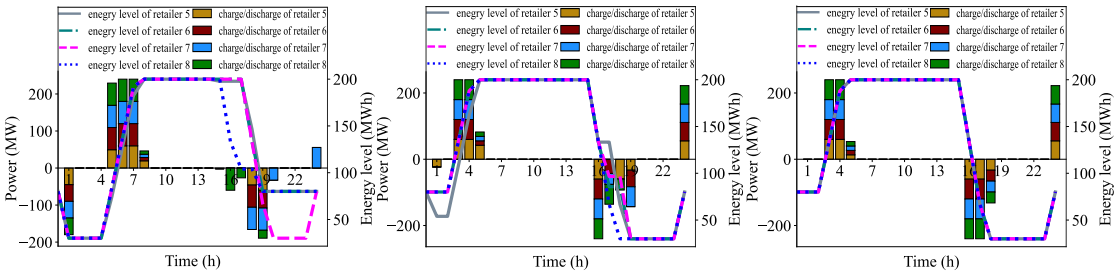
3.4.4 Customers' Welfare

The relationship between the switching coefficient and customers' welfare is displayed in Figure 3.7, which reflects the balance between the customers' utility (the amount of electricity consumed) and the electricity purchase cost. In particular, there is a positive correlation between the magnitude of the switching coefficient and customers' welfare until it reaches the peak with the switching coefficient around $-4MWh/\$$. Thereafter, the customers' welfare decreases drastically. Namely, compared to the situation of no switching behaviours being considered, increasing the magnitude of the switching coefficient at a certain level can increase customers' welfare since it can cause the reduction of the retail price whilst keeping the retailers' load supply at an acceptable level. However, when the magnitude of the switching coefficient becomes sufficiently large, it discourages the retailers from



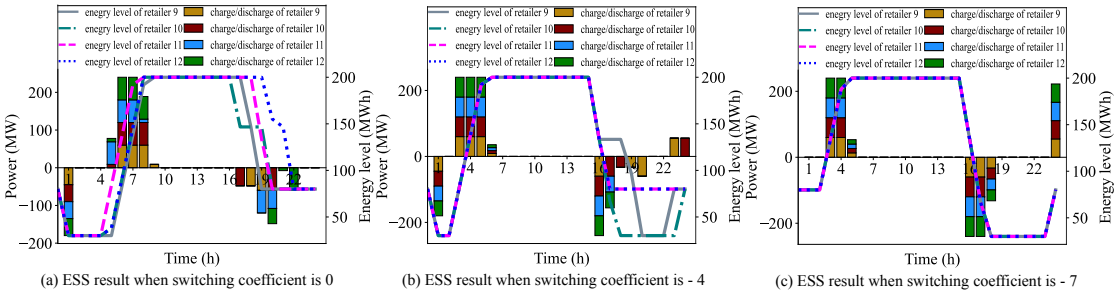
(a) ESS result when switching coefficient is 0 (b) ESS result when switching coefficient is - 4 (c) ESS result when switching coefficient is - 7

Figure 3.8: ESS energy level, charging and discharging power of retailers in group 1.



(a) ESS result when switching coefficient is 0 (b) ESS result when switching coefficient is - 4 (c) ESS result when switching coefficient is - 7

Figure 3.9: ESS energy level, charging and discharging power of retailers in group 2.



(a) ESS result when switching coefficient is 0 (b) ESS result when switching coefficient is - 4 (c) ESS result when switching coefficient is - 7

Figure 3.10: ESS energy level, charging and discharging power of retailers in group 3.

offering electricity supply with a smaller profit margin in return. In this regard, the customers are provided less electricity by the retailers, which results in the reduction of the customers' utility. Therefore, it leads to the customers' welfare losses.

Table 3.1: The effect of the number of retailers on the retail competition.

		Retailer	Average retail price ($/MWh)	Retail price by group ($/MWh)	Profit ($)	Profit by group ($)	Total profit ($)
Case 1	Group 1	1	106.26	105.64	1.04×10^7	1.12×10^7	1.95×10^8
		2	104.05		1.11×10^7		
		3	105.45		1.14×10^7		
		4	106.80		1.18×10^7		
	Group 2	5	124.23	124.73	1.53×10^7	1.60×10^7	
		6	123.35		1.57×10^7		
		7	125.29		1.62×10^7		
		8	126.04		1.69×10^7		
	Group 3	9	143.40	144.40	2.07×10^7	2.16×10^7	
		10	143.65		2.13×10^7		
		11	144.33		2.18×10^7		
		12	146.20		2.25×10^7		
Case 2	Group 1	1	141.38	143.77	3.73×10^7	3.83×10^7	3.20×10^8
		2	146.15		3.94×10^7		
	Group 2	3	168.23	171.16	5.16×10^7	5.28×10^7	
		4	174.09		5.41×10^7		
	Group 3	5	198.55	201.54	6.79×10^7	6.90×10^7	
		6	204.53		7.02×10^7		
Case 3	Group 1	1	223.20	223.20	1.16×10^8	1.16×10^8	4.71×10^8
	Group 2	2	268.61	268.61	1.56×10^8	1.56×10^8	
	Group 3	3	299.86	299.86	1.99×10^8	1.99×10^8	

3.4.5 ESS Result

This section discusses the ESS operation in the EPEC problem. Figure 3.8-3.10 show the ESS energy level, charging, and discharging power of each retailer in different retailers' market share groups, respectively. Particularly, Figure 3.8 (a) indicates the ESS result when there are no customers' switching behaviours. Figure 3.8 (b) and (c) show the ESS results when the switching coefficients are $-4MWh/\$$ and $-7MWh/\$$. Notice that the line plot in each figure denotes the ESS energy level, while the bar plot indicates the charging power (if positive) and discharging power (if negative) of the ESS. We conclude that the retailers typically charge their ESS when the DAW market-clearing price is low and discharge the ESS when the DAW market-clearing price is high regardless of the corresponding market share and the value of the switching coefficient.

Moreover, by comparing the ESS results under different switching coefficients, we can find that each retailer's charging/discharging strategy within each market share group becomes similar when the magnitude of the switching coefficient increases. The reason is that increasing the ability of customers' switching behaviours causes the convergence of the retailers' optimal strategies, including the ESS operating decisions.

3.4.6 The Number of Retailers on the Retail Competition

This section discusses the effect of the number of strategic retailers on the retail competition, where the results are shown in Table 3.1. We consider three different cases with different numbers of retailers. All cases have three retailer groups with different market share. To focus on the effect of the number of retailers, we do not consider switching behaviours in these three cases. The parameter setup for cases 2 and 3 can be found in A.1.3 and A.1.4, respectively. Compared to case 1, decreasing the number of retailers in cases 2 and 3 can significantly reduce the competition among retailers, resulting in much higher daily average retail prices in the larger market share group (e.g., group 3). Furthermore, the reduced retail competition surges the retail prices in each group consistently. For instance, the retailer's daily average retail price in group 3 of case 3 is \$299.86/MWh, which approaches the cap of the retail price (\$300/MWh). In addition, reducing retail competition causes the remarkable dilation of retailers' profit in each group and the total profit in each case. This is the result of the noticeably increased retail price and market power of the retailers in the absence of competition.

3.5 Discussion

This chapter proposes a bilevel game-theoretic framework for strategic retailers who aim to maximise their profits by participating in both DAW and local electricity markets. The proposed bilevel model consists of customers' switching behaviour and market-clearing problems for the DAW and local electricity markets, which is formulated as an MPEC problem and solved by the KKT conditions approach. The final multi-leader multi-follower bilevel model is reformulated as an EPEC problem as solved by the diagonalisation algorithm. The numerical analysis demonstrates the feasibility and effectiveness of the proposed bilevel decision-making framework and the effect of customers' switching behaviour on retailers' decision-making. The work can be further developed in the following directions. First, the modelling of customers' switching behaviours among different retailers could be considered in enhancing existing demand response programs such as load shifting and curtailment [15]. Second, the proposed bilevel strategic model could consider multi-energy scenarios involving electricity, natural gas, and heat energy. Moreover, the effect of network congestion and locational marginal prices on the main findings of this chapter is also worth investigating. In the next chapter, a customised multi-energy pricing problem is modelled as a single-leader-multiple-follower bilevel program and solved by three hybrid metaheuristic algorithms (i.e., PSO, GA and SA). Lastly, data-driven approaches can be employed to improve the modelling process, accuracy and performance. For instance, customers' switching behaviours, wholesale electricity prices and demand are all time-series data and can be learned from historical data through machine learning methods, such as RNN, GRU, LSTM and Transformer-based methods. To this end, chapter 5 proposes Patchformer, a Transformer-based model for long-term time-series forecasting. The model has demonstrated its superiority compared to many state-of-the-art time-series models in long-term forecasting.

Chapter 4

CUSTOMISED MULTI-ENERGY PRICING: MODEL AND SOLUTIONS

This chapter is reproduced with changes from:

- Q. Hong, F. Meng, "Customized Multi-Energy Pricing in Smart Grids: A Bilevel and Evolutionary Computation Approach," The 21st UK Workshop on Computational Intelligence, Springer, 2022.
- Q. Hong, F. Meng, and J. Liu, "Customised multi-energy pricing: Model and solutions," Energies, vol. 16, no. 4, p. 2080, 2023.

The previous chapter proposes a multi-leader, multi-follower bilevel game-theoretic decision-making framework for multiple strategic retailers participating in both local and wholesale electricity markets. To extend the research from a single-energy market to a multi-energy market, in this chapter, a novel customised retail pricing scheme is proposed in the multi-energy context.

4.1 Introduction

With the increasing interdependence among energies (e.g., electricity, natural gas and heat) and the development of a decentralised energy system, a novel retail pricing scheme in the multi-energy market is demanded. Therefore, the problem of designing a customised multi-energy pricing scheme for energy retailers is investigated in this chapter. In particular, the proposed pricing scheme is formulated as a bilevel optimisation problem. At the upper level, the energy retailer (leader) aims to maximise its profit. Microgrids (followers) equipped with energy converters, storage, RES and DR programs are located at the lower level and minimise their operational costs. Three hybrid algorithms combining metaheuristic algorithms (i.e., PSO, GA and SA) with the MILP are developed to solve the proposed bilevel problem. Numerical results verify the feasibility and effectiveness of the proposed model and solution algorithms. We find that GA outperforms other solution algorithms to obtain a higher retailer's profit through comparison. In addition, the proposed customised pricing scheme could benefit the retailer's profitability and net profit margin compared to the widely adopted uniform pricing scheme due to the

reduction in the overall energy purchasing costs in the wholesale markets. Lastly, the negative correlations between the rated capacity and power of the energy storage and both retailer's profit and the microgrid's operational cost are illustrated.

4.1.1 Contributions

The main contributions of this chapter are summarised as follows:

- A bilevel optimisation model is developed to formulate the novel customised pricing scheme for an energy retailer that manages multiple microgrids in the multi-energy market. In particular, a retailer's profit maximisation problem is considered at the upper level. The energy management for each microgrid is detailed, and the operational cost minimisation is formulated at the lower level.
- The detailed energy management model for each microgrid equipped with energy converters (i.e., CHP and heat pump), electrical and thermal storage, RES (i.e., solar and wind) and DR programs (i.e., load curtailment and shifting) is formulated as a MILP program at the lower level. Specifically, load curtailment refers to the reduction in energy consumption, while in the load-shifting program, the electricity demand can be rescheduled and shifted to other scheduling hours.
- Three hybrid metaheuristic algorithms (i.e., PSO, GA and SA) combined with the conventional MILP program are developed to solve the proposed bilevel problem. The hybrid solution algorithms conquer the non-convexity of the lower level problems, which are proved difficult to solve with traditional mathematical methods, such as KKT-based solution methods. In numerical analyses, we test the performance of the three algorithms. The comparison between the customised and uniform pricing schemes is illustrated in detail. In addition, the effect of the rated capacity and power of electrical and thermal storage on the energy retailer's pricing decisions, profit, and microgrids' operational costs is thoroughly investigated.

4.1.2 Chapter Organisation

The remainder of this chapter is organised as follows. In Section 4.2, the proposed bilevel model is discussed in detail. Section 4.3 describes the three hybrid metaheuristic algorithms combined with the MILP program. The numerical results are presented in Section 4.4. Finally, the discussion is drawn in Section 4.5

4.2 Model Formulation

This section shows the formulation of the proposed bilevel optimisation model. In particular, Section 4.2.1 presents an overview of the bilevel MILP model. The customised multi-energy pricing problem is described in Section 4.2.2. Lastly, the lower and upper-level model formulations are discussed in Sections 4.2.3 and 4.2.4, respectively.

4.2.1 Bilevel MILP Model Overview

Bilevel optimisation refers to one of the categories of optimisation that cope with the problem with a hierarchical structure in nature, which includes two decision makers (i.e., leader and follower) located at the upper and lower levels, respectively. The bilevel model is formulated by an optimisation problem (lower level) embedded into another problem (upper level). The problem originates from the game theory in economics and was introduced by Heinrich Freiherr von Stackelberg in 1934 [110]. The decision variables of a bilevel model can be continuous and discrete. Since formulating the customised multi-energy pricing problem involves integer variables in the lower-level problem, in this section, an overview of the bilevel MILP model is introduced as follows.

The general form of a bilevel MILP model with a single leader and multiple independent followers is shown below:

$$\max_{x \in X} F(x, y_1, ...y_M)$$

Subject to:

$$G(x, y_1, ...y_M) \leq 0$$

$$y_i \in \arg\min_{y_i' \in Y_i} \left\{ f_i(x, y_i') : g_i(x, y_i') \leq 0, y_j' \in \mathbb{Z}, \forall j \in \mathcal{J}_i \right\}, \forall i \in \mathcal{M}$$

where $X \subset \mathbb{R}^n$ and $Y_i \subset \mathbb{R}^{m_i}$ are the feasible solution sets for both upper and lower level problems. The set of followers $\{1, ..., M\}$ is denoted as $\mathcal{M}$. n and m_i indicate the number of decision variables for the leader and the follower $i \in \mathcal{M}$. $F(x, y_1, ...y_M)$ and $f_i(x, y_i), \forall i \in \mathcal{M}$ represent the objective functions for the upper and lower level problems. On the other hand, the upper and lower level constraint functions are indicated as $G(x, y_1, ...y_M)$ and $g_i(x, y_i), \forall i \in \mathcal{M}$, respectively. Notice that $\mathcal{J}_i$ is the set of indices that the corresponding $\{y_j, \forall j \in \mathcal{J}_i\}$ are integer variables.

If $x \in X$ denotes the vector of the leader's decision variables, the feasible solution and rational reaction sets of each follower i can be represented as $Y_i(x) = \{y_i \in Y_i :$

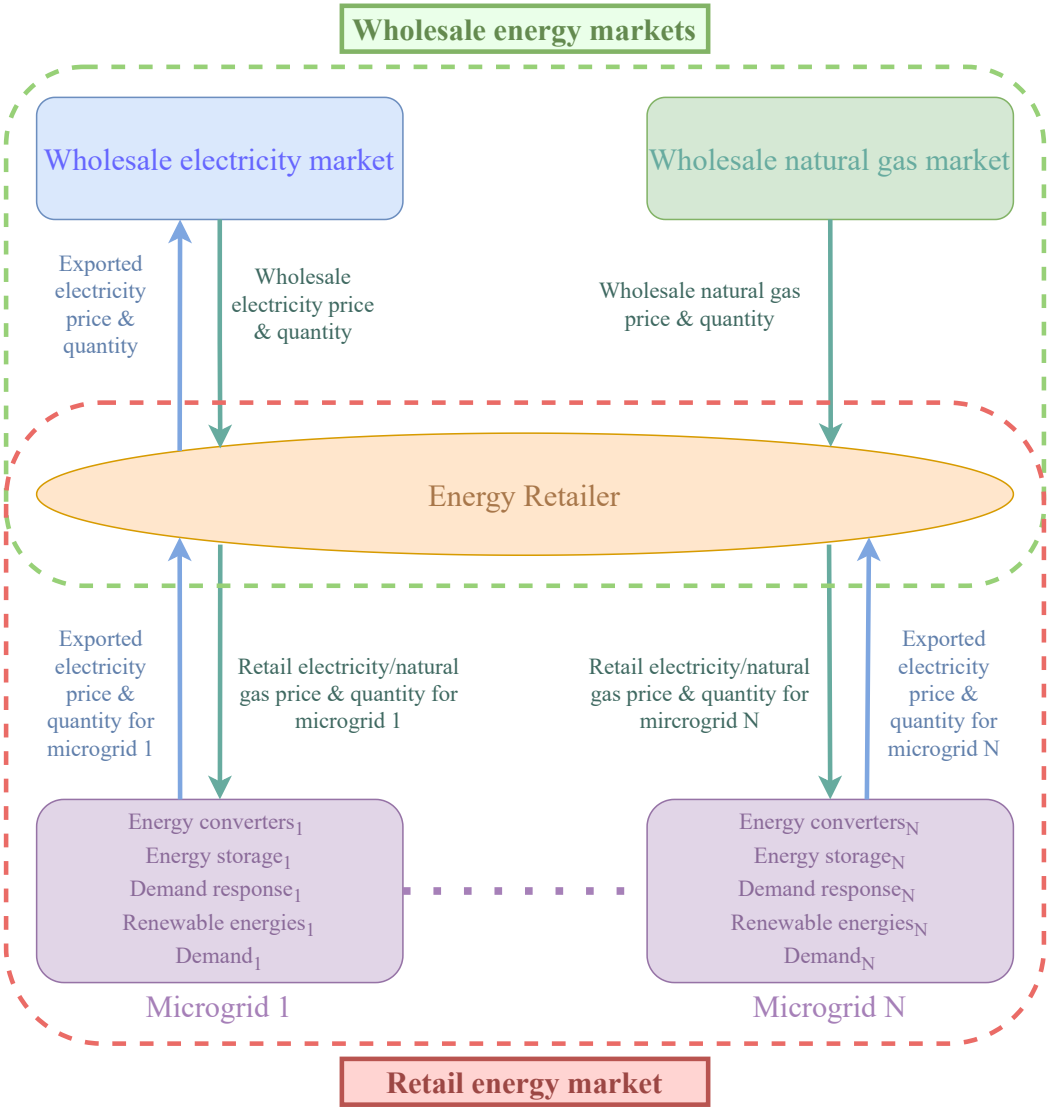


Figure 4.1: Structure of the multi-energy market.

$g_i(x, y_i) \leq 0, y'_j \in \mathbb{Z}, \forall j \in \mathcal{J}_i\}$ and $\Omega_i(x) = \left\{y_i \in Y_i : \arg\min_{y'_i \in Y_i(x)}\{f_i(x, y'_i)\}\right\}$. Finally, the bilevel MILP feasible set, which is also called the inducible region, is presented as $IR = \left\{(x, y_1, ..., y_M) : G(x, y_1, ...y_M) \leq 0, x \in X, y_i \in \Omega_i(x), \forall i \in \mathcal{M}\right\}$. The optimal solution of the bilevel model is denoted as $(x^*, y_1^*, ..., y_M^*) \in \arg\max\{F(x, y_1, ...y_M) : (x, y_1, ..., y_M) \in IR\}$.

4.2.2 Customised Multi-Energy Pricing Problem Description

In the proposed multi-energy market, which is shown in Figure 4.1, multiple microgrids are managed by a single energy retailer that purchases electricity and natural gas from the upstream wholesale markets. The microgrids are allocated with CHP,

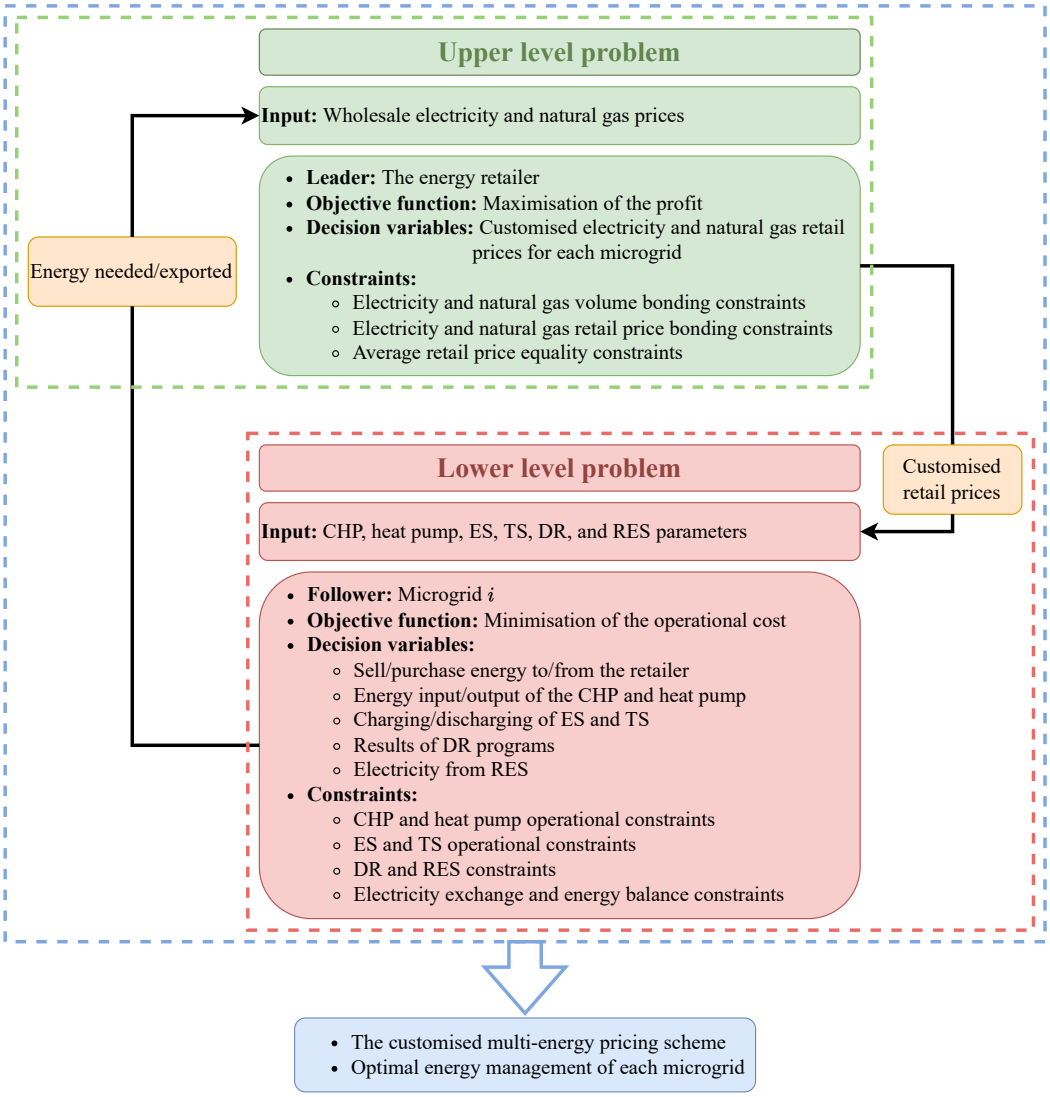


Figure 4.2: Framework diagram of the proposed bilevel model.

heat pump, electrical and thermal storage, RES and DR programs, which operate their energy management systems. Therefore, given the ability of microgrids to generate and transfer energy, the energy retailer can also purchase the electricity from the managed microgrids and sell it back to the wholesale market to make a profit. In addition, the detailed framework of the proposed bilevel model is presented in Figure 4.2. Particularly, at the upper level, to maximise the profit, the energy retailer optimises the retail pricing decisions based on the proposed customised multi-energy pricing scheme within the scheduling hours $\mathcal{T} = \{1, ..., T\}$ and announces them a day ahead to each microgrid m, respectively. After receiving the corresponding retail energy prices, each microgrid m reacts by minimising its

operational cost and reports the volume of energy to be exchanged (buy or sell) to the retailer. As a result, each microgrid's optimal customised retail pricing scheme and energy management are obtained. The detailed model formulations for both the lower and upper levels are shown in Sections 4.2.3 and 4.2.4 below.

4.2.3 Follower-Side/Lower-Level Problem

It is assumed that the microgrids in $\mathcal{M}$ managed by the energy retailer operate independently at the lower level. In this section, the detailed lower-level formulations for the microgrid $m \in \mathcal{M}$ are shown as follows.

4.2.3.1 Lower-Level Objective Function

The lower-level objective function (4.1) shows the total operational costs of the microgrid m. In particular, the first group of elements presents the energy exchange between the microgrid m and the retailer. p_m^t and g_m^t denote the amount of electricity and natural gas that the microgrid m purchases from the retailer. $p_m^{export,t}$ represents the amount of electricity that the microgrid m exports to the retailer. $\pi_{ele,m}^{retail,t}$ and $\pi_{gas,m}^{retail,t}$ denote the retail electricity and natural gas prices that the energy retailer announces to the microgrid m. Notice that the price of electricity sold by the microgrid m back to the retailer is proportional to the retail electricity price, denoted as α_m. The second and third groups of elements describe the CHP and heat pump costs, respectively, including operation and maintenance costs c_m^{CHP},c_m^{pump}, start-up cost $c_{CHP,m}^{st}$,$c_{pump,m}^{st}$ and shut-down cost $c_{CHP,m}^{sd}$,$c_{pump,m}^{sd}$. The fourth group of elements represents the electrical and thermal storage costs, which are denoted as c_m^{ES} and c_m^{TS}. The last group of elements shows the costs of the load curtailment program for all energies, which are denoted as $c_{curtail,m}^{ele,t}$, $c_{curtail,m}^{gas,t}$, and $c_{curtail,m}^{heat,t}$.

$$\begin{aligned}\min_{\Xi_{L_m}} f_m = \sum_{t\in\mathcal{T}} \Bigg\{ & \left[p_m^t \pi_{ele,m}^{retail,t} - p_m^{export,t} \alpha_m \pi_{ele,m}^{retail,t} + g_m^t \pi_{gas,m}^{retail,t}\right] + \left[c_m^{CHP} g_m^{CHP,t} + \right. \\ & c_{CHP,m}^{st} \delta_m^{st,t} + c_{CHP,m}^{sd} \delta_m^{sd,t}\Big] + \left[c_m^{pump} p_m^{pump,t} + c_{pump,t}^{st} \theta_m^{st,t} + c_{pump,m}^{sd} \theta_m^{sd,t}\right] + \\ & \left[c_m^{ES}(p_m^{c,t} + p_m^{d,t}) + c_m^{TS}(q_m^{c,t} + q_m^{d,t})\right] + \left[d_{curtail,m}^{ele,t} \rho_{ele,m}^t c_{curtail,m}^{ele} + \right. \\ & \left. \left. d_{curtail,m}^{gas,t} \rho_{gas,m}^t c_{curtail,m}^{gas} + d_{curtail,m}^{heat,t} \rho_{heat,m}^t c_{curtail,m}^{heat}\right]\right\} \end{aligned} \tag{4.1}$$

4.2.3.2 CHP Operational Constraints

The CHP unit, which converts natural gas into electricity and heat, is formulated in (4.2a)–(4.2m) inspired by Ref. [97]. The energy conversion constraints of the CHP are denoted in (4.2a) and (4.2b). $g_m^{CHP,t}$ represents the amount of natural gas consumed by the CHP. $p_m^{CHP,t}$ and $q_m^{CHP,t}$ denote the amount of electricity and heat generated by the CHP. The energy conversion efficiencies for different energies are denoted as η_m^{CHP} and η_m^{e2h}, respectively. The limitation of the CHP electricity output is shown in (4.2c), where δ_m^t denotes the CHP operational status. (4.2d)–(4.2g) represent the ramp-up and ramp-down power constraints of the CHP electricity output. The initial status of the CHP is denoted as δ_m^{init}. $p_m^{CHP,init}$ represents the last amount of CHP generated electricity during the last scheduling hours. p_m^{RU} and p_m^{RD} show the maximum amount of ramp-up and ramp-down electricity. In addition, the start-up and shut-down actions of the CHP are described in (4.2h)–(4.2l), where $\delta_m^{st,t}$ and $\delta_m^{sd,t}$ denote the CHP start-up and shut-down statuses. Lastly, (4.2m) presents the binary variables that appear in the CHP operation:

$$p_m^{CHP,t} = \eta_m^{CHP} g_m^{CHP,t}, \forall t \in \mathcal{T} \tag{4.2a}$$

$$q_m^{CHP,t} = \eta_m^{e2h} p_m^{CHP,t}, \forall t \in \mathcal{T} \tag{4.2b}$$

$$p_m^{CHP,min} \delta_m^t \leq p_m^{CHP,t} \leq p_m^{CHP,max} \delta_m^t, \forall t \in \mathcal{T} \tag{4.2c}$$

$$p_m^{CHP,t} - p_m^{CHP,init} \leq p_m^{RU}, t = 1 \tag{4.2d}$$

$$p_m^{CHP,t} - p_m^{CHP,t-1} \leq p_m^{RU}, \forall t \in \mathcal{T} \setminus \{1\} \tag{4.2e}$$

$$p_m^{CHP,init} - p_m^{CHP,t} \leq p_m^{RD}, t = 1 \tag{4.2f}$$

$$p_m^{CHP,t-1} - p_m^{CHP,t} \leq p_m^{RD}, \forall t \in \mathcal{T} \setminus \{1\} \tag{4.2g}$$

$$\delta_m^t - \delta_m^{init} \leq \delta_m^{st,t}, t = 1 \tag{4.2h}$$

$$\delta_m^t - \delta_m^{t-1} \leq \delta_m^{st,t}, \forall t \in \mathcal{T} \setminus \{1\} \tag{4.2i}$$

$$\delta_m^{init} - \delta_m^t \leq \delta_m^{sd,t}, t = 1 \tag{4.2j}$$

$$\delta_m^{t-1} - \delta_m^t \leq \delta_m^{sd,t}, \forall t \in \mathcal{T} \setminus \{1\} \tag{4.2k}$$

$$\delta_m^{st,t} + \delta_m^{sd,t} \leq 1, \forall t \in \mathcal{T} \tag{4.2l}$$

$$\delta_m^t, \delta_m^{st,t}, \delta_m^{sd,t} \in \{0, 1\}, \forall t \in \mathcal{T} \tag{4.2m}$$

4.2.3.3 Heat Pump Operational Constraints

The heat pump generates heat energy by consuming electricity. Equation (4.3a) is the energy conversion constraint. $p_m^{pump,t}$ and $q_m^{pump,t}$ denote the amount of electricity consumed and the amount of heat generated by operating the heat pump.

η_m^{pump} denotes the energy conversion efficiency. The amount of generated heat is bounded in (4.3b), where the heat pump operational status is presented as θ_m^t. The ramp-up and ramp-down constraints of the heat pump are depicted in (4.3c)–(4.3f), where $q_m^{pump,init}$ denotes the final amount of heat generated by the heat pump in the last scheduling hours. The maximum amount of ramp-up and ramp-down heat are represented as q_m^{RU} and q_m^{RD}, respectively. (4.3g)–(4.3k) define the heat pump start-up and shut-down actions, whose corresponding start-up and down statuses are denoted as $\theta_m^{st,t}$ and $\theta_m^{sd,t}$. The binary variables in this operation are shown in (4.3l):

$$q_m^{pump,t} = \eta_m^{pump} p_m^{pump,t}, \forall t \in \mathcal{T} \tag{4.3a}$$

$$q_m^{pump,min}\theta_m^t \leq q_m^{pump,t} \leq q_m^{pump,max}\theta_m^t, \forall t \in \mathcal{T} \tag{4.3b}$$

$$q_m^{pump,t} - q_m^{pump,init} \leq q_m^{RU}, t = 1 \tag{4.3c}$$

$$q_m^{pump,t} - q_m^{pump,t-1} \leq q_m^{RU}, \forall t \in \mathcal{T} \setminus \{1\} \tag{4.3d}$$

$$q_m^{pump,init} - q_m^{pump,t} \leq q_m^{RD}, t = 1 \tag{4.3e}$$

$$q_m^{pump,t-1} - q_m^{pump,t} \leq q_m^{RD}, \forall t \in \mathcal{T} \setminus \{1\} \tag{4.3f}$$

$$\theta_m^t - \theta_m^{init} \leq \theta_m^{st,t}, t = 1 \tag{4.3g}$$

$$\theta_m^t - \theta_m^{t-1} \leq \theta_m^{st,t}, \forall t \in \mathcal{T} \setminus \{1\} \tag{4.3h}$$

$$\theta_m^{init} - \theta_m^t \leq \theta_m^{sd,t}, t = 1 \tag{4.3i}$$

$$\theta_m^{t-1} - \theta_m^t \leq \theta_m^{sd,t}, \forall t \in \mathcal{T} \setminus \{1\} \tag{4.3j}$$

$$\theta_m^{st,t} + \theta_m^{sd,t} \leq 1, \forall t \in \mathcal{T} \tag{4.3k}$$

$$\theta_m^t, \theta_m^{st,t}, \theta_m^{sd,t} \in \{0,1\}, \forall t \in \mathcal{T} \tag{4.3l}$$

4.2.3.4 Electrical Storage (ES) Operational Constraints

Constraints (4.4a) and (4.4b) describe the change of the ES energy level $E_m^{ES,t}$ considering charging rate $\eta_m^{ES,c}$, discharging rate $\eta_m^{ES,d}$, and self-discharging rate ϵ_m^{ES}. $p_m^{c,t}$ and $p_m^{d,t}$ denote the amount of charged and discharged electricity. (4.4c) limits the energy level of the ES in each scheduling hour. In addition, for operational purposes, (4.4d) ensures that the energy level of the ES stays unchanged after the scheduling hours. $E_m^{ES,T}$ and $E_m^{ES,init}$ denote the final and initial energy level of the ES. The charging and discharging power are bounded in (4.4e)–(4.4h), where $\gamma_m^{c,t}$ and $\gamma_m^{d,t}$ represent the charging and discharging statuses:

$$E_m^{ES,t} = E_m^{ES,init} + \eta_m^{ES,c} p_m^{c,t} - \frac{1}{\eta_m^{ES,d}} p_m^{d,t} - \epsilon_m^{ES}, t = 1 \tag{4.4a}$$

$$E_m^{ES,t} = E_m^{ES,t-1} + \eta_m^{ES,c} p_m^{c,t} - \frac{1}{\eta_m^{ES,d}} p_m^{d,t} - \epsilon_m^{ES}, \forall t \in \mathcal{T} \setminus \{1\} \tag{4.4b}$$

$$E_m^{ES,min} \leq E_m^{ES,t} \leq E_m^{ES,max}, \forall t \in \mathcal{T} \tag{4.4c}$$

$$E_m^{ES,T} = E_m^{ES,init} \tag{4.4d}$$

$$\gamma_m^{c,t} p_m^{c,min} \leq p_m^{c,t} \leq \gamma_m^{c,t} p_m^{c,max}, \forall t \in \mathcal{T} \tag{4.4e}$$

$$\gamma_m^{d,t} p_m^{d,min} \leq p_m^{d,t} \leq \gamma_m^{d,t} p_m^{d,max}, \forall t \in \mathcal{T} \tag{4.4f}$$

$$\gamma_m^{c,t} + \gamma_m^{d,t} \leq 1, \forall t \in \mathcal{T} \tag{4.4g}$$

$$\gamma_m^{c,t}, \gamma_m^{d,t} \in \{0, 1\}, \forall t \in \mathcal{T} \tag{4.4h}$$

4.2.3.5 Thermal Storage (TS) Operational Constraints

The operational constraints of TS are similar to ES. Specifically, the TS energy level is represented in (4.5a) and (4.5b), where $E_m^{TS,t}$, $\eta_m^{TS,c}$, $\eta_m^{TS,d}$ and ϵ_m^{TS} denote the TS energy level, charging rate, discharging rate and self-discharging rate, respectively. $q_m^{c,t}$ and $q_m^{d,t}$ present the amount of charged and discharged heat. The TS energy level in each scheduling hour is bounded in (4.5c). The initial and final energy levels $E_m^{TS,T}$, $E_m^{TS,init}$ of the TS are imposed to be equal in (4.5d). Constraints (4.5e)–(4.5h) constrain the charging and discharging power of the TS, where $\zeta_m^{c,t}$ and $\zeta_m^{d,t}$ represent the TS charging and discharging statuses:

$$E_m^{TS,t} = E_m^{TS,init} + \eta_m^{TS,c} q_m^{c,t} - \frac{1}{\eta_m^{TS,d}} q_m^{d,t} - \epsilon_m^{TS}, t = 1 \tag{4.5a}$$

$$E_m^{TS,t} = E_m^{TS,t-1} + \eta_m^{TS,c} q_m^{c,t} - \frac{1}{\eta_m^{TS,d}} q_m^{d,t} - \epsilon_m^{TS}, \forall t \in \mathcal{T} \setminus \{1\} \tag{4.5b}$$

$$E_m^{TS,min} \leq E_m^{TS,t} \leq E_m^{TS,max}, \forall t \in \mathcal{T} \tag{4.5c}$$

$$E_m^{TS,T} = E_m^{TS,init} \tag{4.5d}$$

$$\zeta_m^{c,t} q_m^{c,min} \leq q_m^{c,t} \leq \zeta_m^{c,t} q_m^{c,max}, \forall t \in \mathcal{T} \tag{4.5e}$$

$$\zeta_m^{d,t} q_m^{d,min} \leq q_m^{d,t} \leq \zeta_m^{d,t} q_m^{d,max}, \forall t \in \mathcal{T} \tag{4.5f}$$

$$\zeta_m^{c,t} + \zeta_m^{d,t} \leq 1, \forall t \in \mathcal{T} \tag{4.5g}$$

$$\zeta_m^{c,t}, \zeta_m^{d,t} \in \{0, 1\}, \forall t \in \mathcal{T} \tag{4.5h}$$

4.2.3.6 DR Programs Constraints

Two types of DR programs are considered in microgrid m, which are load curtailment (LC) and load shifting (LS), which are formulated in (4.6a)–(4.6c) and (4.6d)–(4.6h), respectively. Both formulations of the DR programs are inspired by [97]. The detailed description and formulation are shown below.

Load curtailment:

It is assumed that the electricity, natural gas, and heat demand can all be curtailed during the scheduling hours, and the curtailment rates are denoted as $\rho^t_{ele,m}$, $\rho^t_{gas,m}$, $\rho^t_{heat,m}$, respectively. The bounding constraints for the three types of energies are presented in (4.6a)–(4.6c). Notice that the curtailable energy demand in each scheduling hour is predetermined by the energy retailer and represented as $d^{ele,t}_{curtail}$, $d^{gas,t}_{curtail}$ and $d^{heat,t}_{curtail}$:

$$\rho^{min}_{ele,m} \leq \rho^t_{ele,m} \leq \rho^{max}_{ele,m}, \forall t \in \mathcal{T} \tag{4.6a}$$

$$\rho^{min}_{gas,m} \leq \rho^t_{gas,m} \leq \rho^{max}_{gas,m}, \forall t \in \mathcal{T} \tag{4.6b}$$

$$\rho^{min}_{heat,m} \leq \rho^t_{heat,m} \leq \rho^{max}_{heat,m}, \forall t \in \mathcal{T} \tag{4.6c}$$

Load shifting:

We assume there are households in $\mathcal{A}_i$ that participate in the load-shifting program. Each household has the load-adjustable time window D_{a_i}, which is represented in (4.6d). $\mu^t_{a_i}$ denotes the household's operational status. The start and stop times of the load-shifting program for each household a_m are denoted as $T^{start}_{a_i}$ and $T^{stop}_{a_i}$. Constraint (4.6e) imposes that there is no shiftable load available in the scheduling hours outside of the adjustable time window. The shiftable load $d^t_{a_i}$ in each scheduling hour is flexible but bounded in (4.6f). Finally, (4.6g) makes sure the overall electricity consumption E_{a_i} is not affected by the load-shifting program:

$$\sum_{t=T^{start}_{a_i}}^{T^{stop}_{a_i}} \mu^t_{a_i} = D_{a_i}, \forall a_i \in \mathcal{A}_m \tag{4.6d}$$

$$\sum_{t=1}^{T^{start}_{a_i}-1} \mu^t_{a_i} + \sum_{t=T^{stop}_{a_i}+1}^{T} \mu^t_{a_i} = 0, \forall a_i \in \mathcal{A}_m \tag{4.6e}$$

$$d^{min}_{a_i}\mu^t_{a_i} \leq d^t_{a_i} \leq d^{max}_{a_i}\mu^t_{a_i}, \forall a_i \in \mathcal{A}_m, \forall t \in \mathcal{T} \tag{4.6f}$$

$$\sum_{t=T^{start}_{a_i}}^{T^{stop}_{a_i}} d^t_{a_i} = E_{a_i}, \forall a_i \in \mathcal{A}_m \tag{4.6g}$$

$$\mu^t_{a_i} \in \{0, 1\}, \forall a_i \in \mathcal{A}_m, \forall t \in \mathcal{T} \tag{4.6h}$$

4.2.3.7 RES Constraints

Two renewable energies, solar and wind power, which are generated by photovoltaics (PV) and wind turbines (WT) are considered in this chapter. To reduce the effect of

the uncertainties of the renewable energies in nature, the bounding constraints of the forecast of PV $p_m^{PV,t}$ and wind powers $p_m^{wind,t}$ in each scheduling hour are introduced in (4.7a) and (4.7b):

$$p_m^{PV,t,min} \leq p_m^{PV,t} \leq p_m^{PV,t,max}, \forall t \in \mathcal{T} \tag{4.7a}$$

$$p_m^{wind,t,min} \leq p_m^{wind,t} \leq p_m^{wind,t,max}, \forall t \in \mathcal{T} \tag{4.7b}$$

In addition, inspired by Ref. [111], the spinning reserve constraint (4.7c) is implemented to further maintain and secure the operation of microgrids and the power system. It presents that the maximum power supply of the microgrid m must be sufficient to provide at least $(1 + \tau^{spin})$ times load demand in each scheduling hour. τ^{spin} denotes the spinning reserve ratio:

$$\begin{aligned} &p_m^{CHP,max}\delta_m^t + p_m^{d,t} - p_m^{c,t} + p_m^t - p_m^{export,t} - p_m^{pump,t} + p_m^{PV,t} \\ &+ p_m^{wind,t} \geq (1 + \tau^{spin})\Big[d_m^{ele,t} + \sum_{a_i \in \mathcal{A}_m} d_{a_i}^t + d_{curtail,m}^{ele,t}(1 - \rho_{ele,m}^t)\Big], \forall t \in \mathcal{T} \end{aligned} \tag{4.7c}$$

4.2.3.8 Microgrid Electricity Exchange Constraints

The electricity imported to and exported from the microgrid m are constrained in (4.8a) and (4.8b), where the importing and exporting statuses are denoted as ψ_m^t and $\psi_m^{export,t}$, respectively. Additionally, (4.8c) imposes that electricity import and export cannot happen simultaneously:

$$\psi_m^t p_m^{min,t} \leq p_m^t \leq \psi_m^t p_m^{max,t}, \forall t \in \mathcal{T} \tag{4.8a}$$

$$\psi_m^{export,t} p_{export,m}^{min,t} \leq p_m^{export,t} \leq \psi_m^{export,t} p_{export,m}^{max,t}, \forall t \in \mathcal{T} \tag{4.8b}$$

$$\psi_m^t + \psi_m^{export,t} \leq 1, \forall t \in \mathcal{T} \tag{4.8c}$$

$$\psi_m^t, \psi_m^{export,t} \in \{0, 1\}, \forall t \in \mathcal{T} \tag{4.8d}$$

4.2.3.9 Energy Balance Constraints

The demand and supply balance constraint for each type of energy in the multi-energy system must be satisfied at every scheduling hour. Constraints (4.9a)–(4.9c) represent the energy balance constraints of electricity, natural gas, and heat, respectively. $d_m^{ele,t}$, $d_m^{gas,t}$, and $d_m^{heat,t}$ denote the critical/base demand for each type of energy:

$$p_m^{d,t} + p_m^t + p_m^{CHP,t} + p_m^{PV,t} + p_m^{wind,t} = d_m^{ele,t} + p_m^{export,t} + p_m^{pump,t} + p_m^{c,t}$$

$$+ \sum_{a_i \in \mathcal{A}_m} d^t_{a_i} + d^{ele,t}_{curtail,m}(1 - \rho^t_{ele,m}), \forall t \in \mathcal{T} \tag{4.9a}$$

$$g^t_m = d^{gas,t}_m + g^{CHP,t}_m + d^{gas,t}_{curtail,m}(1 - \rho^t_{gas,m}), \forall t \in \mathcal{T} \tag{4.9b}$$

$$q^{d,t}_m + q^{CHP,t}_m + q^{pump,t}_m = d^{heat,t}_m + q^{c,t}_m + d^{heat,t}_{curtail,m}(1 - \rho^t_{heat,m}), \forall t \in \mathcal{T} \tag{4.9c}$$

The decision variables of the lower-level problem for the microgrid m are $\Xi_{L_m} = \{p^t_m, g^t_m, p^{CHP,t}_m, p^{export,t}_m, p^{pump,t}_m, g^{CHP,t}_m, q^{CHP,t}_m, q^{pump,t}_m, \delta^t_m, \delta^{st,t}_m, \delta^{sd,t}_m, \theta^t_m, \theta^{st,t}_m, \theta^{sd,t}_m, \psi^t_m, \psi^{export,t}_m, E^{ES,t}_m, p^{c,t}_m, p^{d,t}_m, \gamma^{c,t}_m, \gamma^{d,t}_m, E^{TS,t}_m, q^{c,t}_m, q^{d,t}_m, \zeta^{c,t}_m, \zeta^{d,t}_m, \rho^t_{ele,m}, \rho^t_{gas,m}, \rho^t_{heat,m}, \mu^t_{a_i}, d^t_{a_i}, p^{PV,t}_m, p^{wind,t}_m\}$. Notice that the lower-level problem of each microgrid forms a MILP problem, which can be solved efficiently by off-the-shelf commercial solvers, such as CPLEX and GUROBI.

4.2.4 Leader-Side/Upper-Level Problem

We assume that the energy retailer manages multiple multi-energy microgrids by adopting the proposed customised multi-energy pricing scheme. The profit maximisation problem of the retailer is formulated at the upper level and shown as follows:

$$\max_{\Xi_U} F = \sum_{t \in \mathcal{T}} \left\{ \sum_{m \in \mathcal{M}} \left(\pi^{retail,t}_{ele,m} p^t_m - \alpha_m \pi^{retail,t}_{ele,m} p^{export,t}_m \right) - \pi^t_{ele} \sum_{m \in \mathcal{M}} \left(p^t_m - p^{export,t}_m \right) + \sum_{m \in \mathcal{M}} \left(\pi^{retail,t}_{gas,m} - \pi^t_{gas} \right) g^t_m \right\} \tag{4.10a}$$

subject to

$$p^{min}_{total} \leq \sum_{m \in \mathcal{M}} \left(p^t_m - p^{export,t}_m \right) \leq p^{max}_{total}, \forall t \in \mathcal{T} \tag{4.10b}$$

$$g^{min}_{total} \leq \sum_{m \in \mathcal{M}} g^t_m \leq g^{max}_{total}, \forall t \in \mathcal{T} \tag{4.10c}$$

$$\pi^{retail,min}_{ele,m} \leq \pi^{retail,t}_{ele,m} \leq \pi^{retail,max}_{ele,m}, \forall m \in \mathcal{M}, \forall t \in \mathcal{T} \tag{4.10d}$$

$$\pi^{retail,min}_{gas,m} \leq \pi^{retail,t}_{gas,m} \leq \pi^{retail,max}_{gas,m}, \forall m \in \mathcal{M}, \forall t \in \mathcal{T} \tag{4.10e}$$

$$\frac{\sum_{t \in \mathcal{T}} \pi^{retail,t}_{ele,m}}{T} = AVG^{ele}, \forall m \in \mathcal{M} \tag{4.10f}$$

$$\frac{\sum_{t \in \mathcal{T}} \pi^{retail,t}_{gas,m}}{T} = AVG^{gas}, \forall m \in \mathcal{M} \tag{4.10g}$$

The decision variables of the upper-level problem are $\Xi_U = \{\pi^{retail,t}_{ele,m}, \pi^{retail,t}_{gas,m}\}$, which denote the retail electricity and natural gas prices for each microgrid $m \in \mathcal{M}$

at each scheduling hour. The objective function (4.10a) depicts the overall profit of the retailer. π^t_{ele} and π^t_{gas} denote the wholesale electricity and natural gas prices. In particular, the first group of elements in (4.10a) represents the revenue obtained from all microgrids by exchanging electricity. The electricity purchasing cost from, or the selling revenue to, the wholesale electricity market is described in the second group of elements. The last group of elements denotes the profit of selling natural gas to the microgrids. Constraints (4.10b) and (4.10c) limit the amount of electricity and natural gas exchange between the energy retailer and wholesale markets. The retail electricity and natural gas prices customised for each microgrid are bounded in (4.10d) and (4.10e). Constraints (4.10f) and (4.10g), which are inspired by Ref. [13], impose the average retail electricity and natural gas prices within the scheduling hours to be equal to the predetermined constants AVG^{ele} and AVG^{gas}. Notice that the two equality constraints ensure a sufficient number of low-price scheduling hours. Without these constraints, in principle, the retailer could always choose the maximum retail prices to maximise its profit. As a result, the constraints provide a fair retail price image among microgrids, which are also crucial for the retailer's market share in the long term [112].

4.3 Solution Methods

The proposed bilevel model includes many binary decision variables in the lower-level problems, which make it extremely hard to solve by conventional analytical methods, such as KKT reformulation-related methods. To overcome the non-convexity of the model, we propose three categories of metaheuristic algorithms that simulate the behaviours of the energy retailer at the upper level to find an optimal solution efficiently, which are as follows: swarm-based, PSO; evolution-based, GA; and physics-based, SA. Therefore, the three metaheuristic algorithms combined with the lower-level MILP solver form different hybrid solution algorithms, which are illustrated in detail below.

4.3.1 PSO-Based Algorithm

The PSO algorithm is a swarm-based metaheuristic method that simulates the social behaviour of the movement of organisms in a bird flock or fish school [113]. Namely, when birds search randomly for food, for instance, all birds in the flock are able to share their knowledge and discovery to help the entire flock find the best location for hunting. Figure 4.3 shows the flowchart of the proposed PSO-based algorithm. Essentially, each particle contains two properties: position and velocity. During

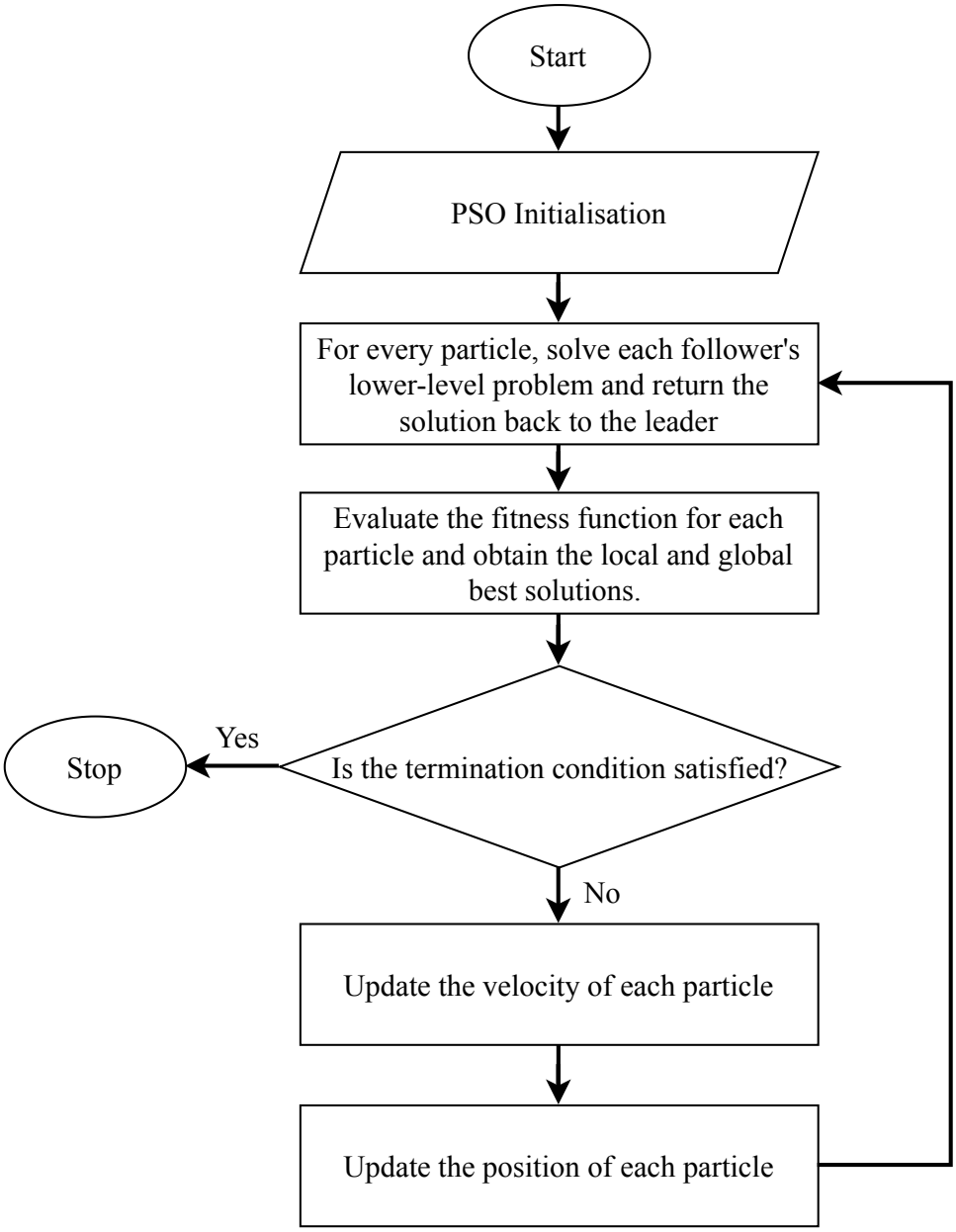


Figure 4.3: Flowchart of the PSO-based algorithm.

each iteration, the global best position of all particles and the best previously visited position of each particle are found by evaluating the fitness function of each particle. Each particle's position and velocity are updated via the equations as follows:

$$v_n^{y+1} = w v_n^y + c_1 r_1^y (x_n^{p,y} - x_n^y) + c_2 r_2^y (x^{g,y} - x_n^y) \tag{4.11a}$$

$$x_n^{y+1} = x_n^y + v_n^{y+1} \tag{4.11b}$$

where the superscript y indicates the number of iterations, w denotes the inertia

Algorithm 2 PSO-based algorithm.

1: Initialisation: Total number of particles N_{PSO}. Maximum iteration. Each particle's position x and velocity v.
2: Each particle's position denotes the strategy of the leader (i.e., energy retailer).
3: **for** $n_{PSO} = 1$ to N_{PSO} **do**
4: The energy retailer announces customised electricity and natural gas prices for each microgrid for the next 24 h.
5: After receiving the prices, each microgrid n_{PSO} reacts to the leader's strategy by solving the energy management problem and obtaining the optimal solutions, which are returned back to the energy retailer.
6: The energy retailer solves the profit maximisation problem and evaluates it using the fitness function after receiving the optimal solutions from all microgrids.
7: **end for**
8: **if** The termination condition is satisfied **then**
9: The algorithm terminates and returns the outputs.
10: **else**
11: Record the global best position of all particles as $x^{g,y}$.
12: Record the best previously visited position of the particle n_{PSO} as $x^{p,y}_{n_{PSO}}$.
13: Update each particle's velocity and position by using (4.11a) and (4.11b)
14: **end if**
15: Repeat steps 3–14 until the termination condition is reached.

weight, c_1 and c_2 represent cognitive constant and social constant, respectively, and r_1 and r_2 are uniformly distributed random numbers in $[0, 1]$. Notice that if any element in the updated particle position is out of the boundary set by (4.10d) and (4.10e), the nearest boundary is assigned to the element [4].

The detailed process of the PSO-based decision-making algorithm is shown in Algorithm 2. In particular, the maximum iteration, total number of particles N_{PSO} and each particle's position x and velocity v need to be initialised. Notice that each particle's position denotes the energy retailer's customised pricing decision for the next 24 h. Steps 3–7 show the interaction between the energy retailer and microgrids and are interpreted as follows. First, the energy retailer announces the customised electricity and natural gas prices for each microgrid for the next 24 h in step 4. Then, each microgrid solves its energy-management problem based on the received energy prices from the retailer and obtains the optimal solution, which is returned back to the energy retailer in step 5. In step 6, after receiving the optimal energy management solutions from all microgrids, the energy retailer solves its profit maximisation problem and evaluates the current pricing decisions using

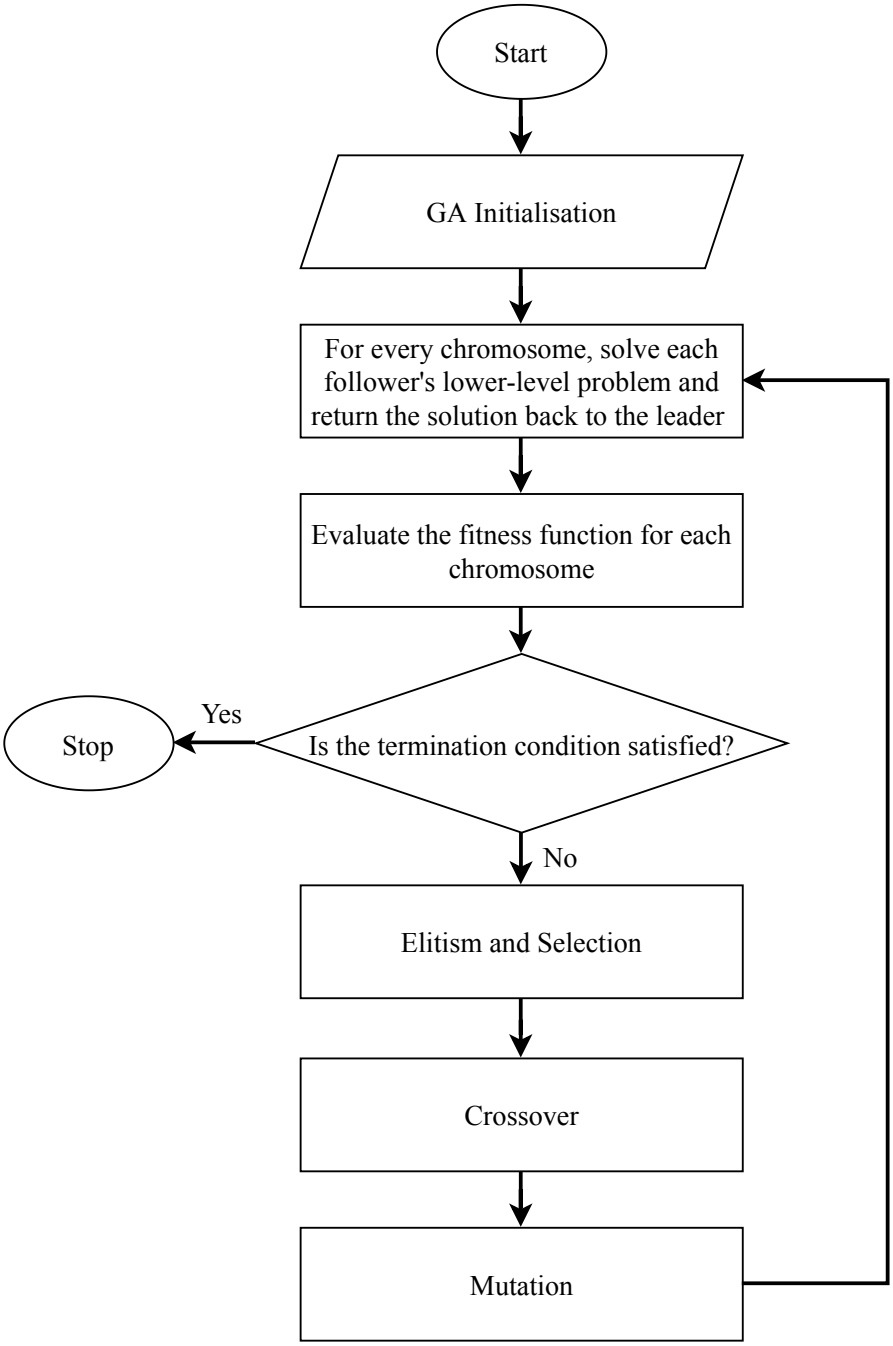


Figure 4.4: Flowchart of the GA-based algorithm.

the fitness function based on the penalisation method proposed by [114], [115]. Notice that the penalised fitness function formulation can not only be used for the PSO-based algorithm but also be applied to GA- and SA-based algorithms since the constraint-handling technique can be generalised in metaheuristic algorithms. When the predefined termination condition is not satisfied, the global best position

Algorithm 3 GA-based algorithm.

1: Initialisation: Maximum iteration. Population of N_{GA} chromosomes.
2: Each chromosome indicates the strategy of the leader (i.e., energy retailer).
3: **for** $n_{GA} = 1$ to N_{GA} **do**
4: The energy retailer announces customised electricity and natural gas prices for each microgrid for the next 24 h.
5: After receiving the prices, each microgrid n_{GA} reacts to the leader's strategy by solving the energy management problem and obtaining the optimal solutions, which are returned back to the energy retailer.
6: The energy retailer solves the profit maximisation problem and evaluates it using the fitness function after receiving the optimal solutions from all microgrids.
7: **end for**
8: **if** The termination condition is satisfied **then**
9: The algorithm terminates and returns the outputs.
10: **else**
11: The next generation of chromosomes is produced by selection, crossover and mutation.
12: **end if**
13: Repeat steps 3–12 until the termination condition is reached.

of all particles $x^{g,y}$ and the personal best previously visited position of each particle $x_n^{p,y}$ are recorded in steps 11 and 12. Each particle's position and velocity are then updated by applying (4.11a) and (4.11b) in step 13. The algorithm iterates until the termination condition is reached and outputs the optimal solutions for the energy retailer and each microgrid.

4.3.2 GA-Based Algorithm

GA algorithm is an evolution-based computational method inspired by genetics and natural selection [116]. The flowchart of the GA-based algorithm is shown in Figure 4.4. In particular, after evaluating the population of the current generation, the elite chromosome with the best fitness value is inherited by the next generation. Furthermore, the selection process, such as the roulette wheel, tournament and random selection, is applied to choose other chromosomes for the next generation. The chromosomes of the successive generation are finally generated by crossover and mutation processes [15]. Algorithm 3 explains the process of the solution algorithm. Firstly, the maximum iteration and population of chromosomes N_{GA} are initialised. Similar to the PSO-based algorithm, each chromosome indicates the energy retailer's pricing decisions, and steps 3–7 show the retailer and microgrids' interactions. Then, the next generation of chromosomes is created by applying selection, crossover and

mutation in step 11. Specifically, the stochastic uniform selection method is used to choose the next generation of chromosomes. The scattered crossover function is applied by creating a random binary vector and selecting the genes from the first parent chromosome when the entry is 1 and from the second parent chromosome when the entry is 0. Lastly, the Gaussian mutation is utilised to explore the search space, which adds a random number taken from the Gaussian distribution with a mean 0 to each gene of the chromosome. The algorithm is stopped when repeating steps 3–12 until the termination condition is satisfied.

4.3.3 SA-Based Algorithm

Simulated annealing is a physics-based probabilistic approach that emulates the standard annealing process used in metallurgy to improve the properties of solids. Essentially, after the solid is heated up by a significantly high temperature, the atoms gain the energy to explore their stable states. The optimal state of each atom can be then found following the annealing process. As a result, the solid is recrystallised and improves its ductility. Similarly, the simulated annealing algorithm is applied to find the optimal solution by cooling the heated atoms in the search space. Figure 4.5 shows the flowchart of the SA-based algorithm. In particular, it uses the Metropolis algorithm to generate the next trial atoms by randomly perturbing the current ones. If the fitness value of the trial atom is greater than the current fitness value, the trial atom replaces the current atom in the next iteration. Otherwise, the trial atom can still be accepted by the acceptance criterion, which is based on the Boltzmann distribution shown in (4.12) [117]:

$$Pr_{n_{SA}}^{y} = \exp\left(\Delta_{n_{SA}}/T^{y}\right) \tag{4.12}$$

where $\Delta_{n_{SA}}$ represents the difference of fitness values between the trial atom and current atom n_{SA}. The temperature in the current iteration is denoted as T^{y}. Notice that when the temperature is significantly high, the SA algorithm can accept the worse solution, which is particularly beneficial in search space exploration. As the temperature drops in every iteration, the acceptance criterion $Pr_{n_{SA}}^{y}$ also decreases, which leads to the lower possibility of accepting the worse solution. The SA algorithm can be considered to be the hill climbing algorithm when the temperature is significantly low (e.g., 1), which takes advantage of the efficient local search method.

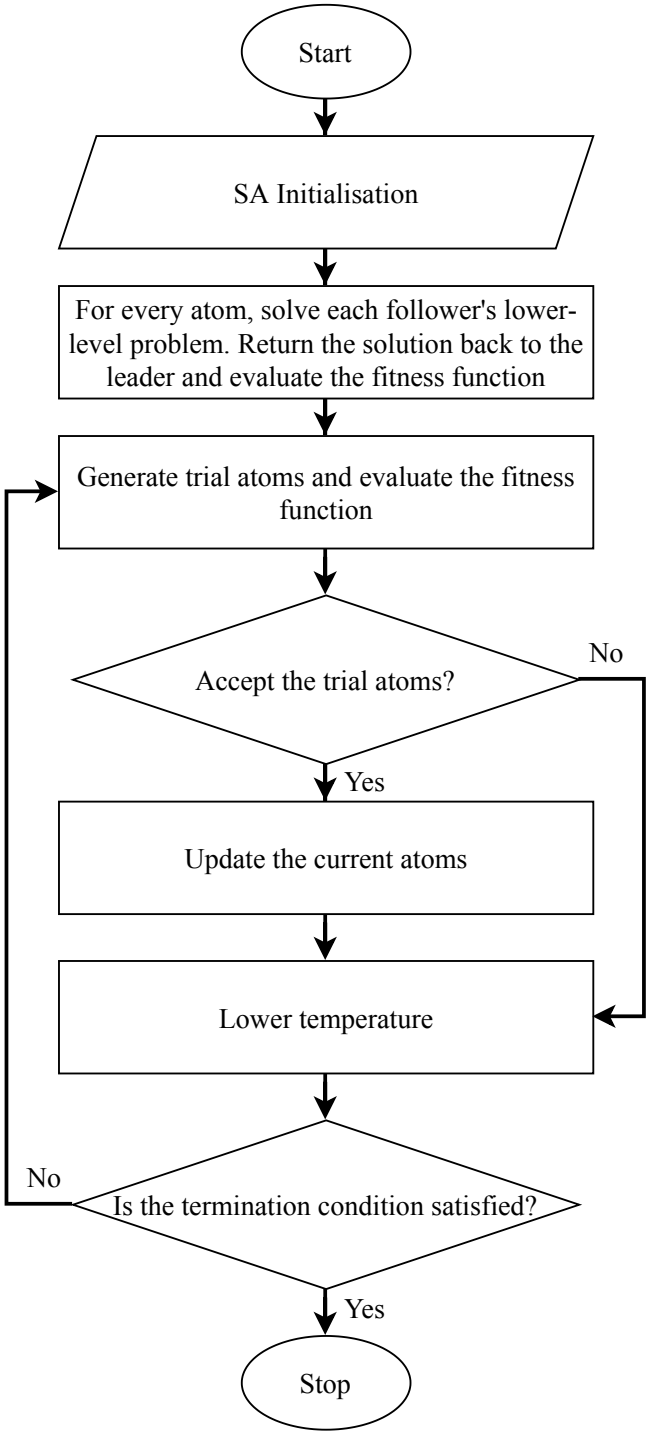


Figure 4.5: Flowchart of the SA-based algorithm.

Algorithm 4 shows the proposed SA-based algorithm to solve the bilevel model. The maximum iteration, total number of atoms, and initial and final temperature are initialised in step 1. Steps 2–7 are similar to Algorithm 2 and 3, which present

Algorithm 4 SA-based algorithm.

1: Initialisation: Maximum iteration. Population of atoms N_{SA}. Initial and final temperature.
2: Each atom indicates the strategy of the leader (i.e., energy retailer).
3: **for** $n_{SA} = 1$ to N_{SA} **do**
4: The energy retailer announces customised electricity and natural gas prices for each microgrid for the next 24 h.
5: After receiving the prices, each microgrid n_{SA} reacts to the leader's strategy by solving the energy management problem and obtaining the optimal solutions, which are returned back to the energy retailer.
6: The energy retailer solves the profit maximisation problem and evaluates it using the fitness function after receiving the optimal solutions from all microgrids.
7: **end for**
8: For each current energy prices strategy, the energy retailer generates trial energy prices and repeats steps 5 and 6.
9: The energy retailer decides to accept the trial energy prices against the current ones by comparing the fitness values and acceptance criterion (4.12).
10: **if** Trial energy prices are accepted **then**
11: The current energy prices are replaced by the trial prices in the next iteration.
12: **end if**
13: Decrease the temperature parameter.
14: **if** The termination condition is satisfied **then**
15: The algorithm terminates and returns the outputs.
16: **else**
17: Repeat steps 8–16 until the termination condition is reached.
18: **end if**

the interactions between the retailer and microgrids. In step 8, the energy retailer proposes a trial energy price decision for each current price strategy and collects the evaluation by repeating steps 5 and 6. If the fitness value from the trial prices is higher than the current ones, the retailer replaces the current prices with the trial prices in the next iteration. Otherwise, the retailer decides the accept or refuse the trial prices based on the acceptance criterion (4.12). This decision process is illustrated in steps 9–12. The temperature decreases in every iteration in step 13. The algorithm repeats steps 8–16 until the termination condition is fulfilled.

4.4 Numerical Results

The section discusses the results of numerical analyses to illustrate the feasibility and effectiveness of the proposed bilevel model and the solution algorithms. In particular, Section 4.4.1 describes the setup for the experiments. The three aforementioned

Table 4.1: ES parameters. – : Not applicable

Parameter	Microgrid 1	Microgrid 2	Microgrid 3
Charging rate	–	0.95	0.95
Discharging rate	–	0.95	0.95
Self-discharging rate (MWh)	–	0.002	0.002
Initial energy level (MWh)	–	250	500
Minimum energy level (MWh)	–	50	100
Rated Capacity (MWh)	–	500	1000
Minimum charging/discharging power (MW/h)	–	20	30
Rated charging/discharging power (MW/h)	–	200	300
Operation & maintenance cost ($/MWh)	–	3.5	3.5

Table 4.2: TS parameters. – : Not applicable

Parameter	Microgrid 1	Microgrid 2	Microgrid 3
Charging rate	–	0.95	0.95
Discharging rate	–	0.95	0.95
Self-discharging rate (MBtu)	–	0.004	0.004
Initial energy level (MBtu)	–	300	650
Minimum energy level (MBtu)	–	60	130
Rated Capacity (MBtu)	–	600	1300
Minimum charging/discharging power (MBtu/h)	–	20	43
Rated charging/discharging power (MBtu/h)	–	200	430
Operation & maintenance cost ($/MBtu)	–	3.5	3.5

metaheuristic algorithms for solving the proposed bilevel model are compared and analysed in Section 4.4.2. Moreover, the performance results of the proposed customised and uniform multi-energy pricing schemes are investigated in Section 4.4.3. Lastly, Section 4.4.4 presents the effect of the rated capacity and power of the ES and TS on the retailer's profit and the microgrids' operations.

4.4.1 Experimental Setup

In this section, we consider three different microgrids (i.e., microgrid 1, microgrid 2 and microgrid 3) that are managed by the energy retailer. Each microgrid's base demand and the configurations of its facilities (i.e., the capacity and rated power of ES and TS) are differentiated from others. The base electricity, natural gas, and heat demand for each microgrid come from the PJM dataset [108], United Kingdom Department of Education Gas dataset [118], and Open Power System Data [119], respectively, which are shown in Table B.1. Each microgrid's ES and TS

parameter setups are shown in Tables 4.1 and 4.2. Besides the ES and TS parameter setup, other facilities, such as CHP, heat pump, and load-shifting programs, share the same parameters among microgrids, shown in Tables B.2, B.3 and B.5. The minimum power of RES is zero, while the maximum power for each scheduling hour is shown in Table B.4, which originated from [120], [121]. The wholesale electricity and natural gas prices come from [122], which are presented in Table B.6. In addition, the curtailable load and costs of electricity, natural gas and heat are 300 MWh, 100 kcf and 350 MBtu, and \$70/MWh, \$20/kcf and \$60/MBtu, respectively. The minimum and maximum curtailment rates are 0 and 0.4. For each microgrid, the minimum and maximum electricity purchased from and exported to the energy retailer is 0 MWh and 5000 MWh, respectively. For the energy retailer, the minimum and maximum retail electricity prices are \$60/MWh and \$110/MWh, and \$15/kcf and \$60/kcf for the natural gas prices. The electricity price that each microgrid sells back to the energy retailer is set to be 90% of the current retail price. The average retail electricity and natural gas prices are predetermined as \$90/MWh and \$40/kcf.

All three hybrid algorithms are written in MATLAB R2022b and run on Windows 11 Pro 64-bit with 12 cores CPU @ 3.6 GHz and 32 GB of RAM. The coupled MILP problem is solved by Gurobi Optimiser (version 10.0.0) using the branch and bound algorithm. Each iteration for all three hybrid algorithms consists of upper-level operations and 600 (200 individuals $\times$ 3 microgrids) lower-level MILP problems, which take about 60 s to complete.

For the hybrid solution algorithms, we consider 100 iterations for a single run, which includes a population of 200 individuals (price signals). In particular, for the PSO-based algorithm, the inertia weight is 1.1. The cognitive and social constants are both 1.49. For the GA-based algorithm, the best 10 elite chromosomes survive to the next generation. In the SA-based algorithm, the initial and final temperatures are 100 and 1 degrees. Notice that all parameters are set after a mass of experiments, considering the balance between the quality of results and computation burden.

4.4.2 Solution Algorithms Comparison

The principal focus of this section is the numerical comparison among the aforementioned hybrid solution methods, i.e., PSO, GA and SA coupled with MILP algorithms. In this experiment, the three hybrid algorithms run 25 times independently for both proposed customised and uniform pricing models. The detailed

Table 4.3: Statistical results of three hybrid solution algorithms.

Scheme	Algorithm	Minimum ($)	Maximum ($)	Median ($)	Average ($)	Standard Deviation	IQR
Customised	PSO	1,747,893.23	3,237,259.44	2,637,117.20	2,661,933.25	305,700.17	304,077.64
	GA	2,793,836.04	3,621,174.98	3,226,859.38	3,235,457.87	188,685.73	262,036.86
	SA	2,157,779.58	3,253,958.04	2,652,633.36	2,794,698.37	313,398.40	534,933.18
Uniform	PSO	2,229,066.83	3,454,961.22	2,868,594.24	2,873,363.22	335,318.00	477,855.56
	GA	2,785,634.97	3,852,188.64	3,160,983.13	3,206,485.96	252,908.66	274,699.89
	SA	2,555,064.32	3,268,070.40	2,862,418.32	2,890,620.58	199,545.25	323,634.31

statistical analysis of the results of the energy retailer's profit is shown in Table 4.3. For the customised pricing scheme, the GA-based algorithm presents outstanding performance along with others, obtaining higher minimum, maximum, median and average values of the retailer's profit. Furthermore, the standard deviation and interquartile range (IQR) that measures the spread of the middle 50% of the data are the lowest using the GA-based algorithm, which indicates less data dispersion. On the other hand, for the uniform pricing scheme, although the standard deviation of SA-based results is the lowest among others, the GA algorithm still reveals significantly high values in all minimum, maximum, median and average statistical measurements. In summary, the GA-based algorithm is verified to provide the best performance for both customised and uniform pricing models, considering the value and stability of the results it can achieve. Therefore, the GA-based hybrid algorithm is applied to solve customised and uniform pricing models in the following experiments.

4.4.3 Customised and Uniform Multi-Energy Pricing Schemes

This section presents the difference between the performance of customised and uniform pricing schemes. Notice that the results come from the GA-based algorithm, which runs 25 times independently in the previous section. The cost and revenue of the retailer, net profit margin and each microgrid's cost are calculated based on the run, which generates the median value of the retailer's profit. The two pricing schemes are compared in Table 4.4. It reveals that the energy retailer can make more profit under the customised pricing scheme measured by both the average (+0.90%) and median (+2.08%) profit value. This is because the customised retail prices are tailored to each microgrid with different characteristics and load patterns, which makes the retailer's pricing decisions more flexible. In addition, Table 4.5 shows the microgrids' energy management results under the two pricing schemes.

Table 4.4: Results of retailer under customised and uniform pricing schemes.

Pricing Scheme	Average Profit of Retailer ($)	Median Profit of Retailer ($)	Cost of Retailer ($)	Revenue of Retailer ($)	Net Profit Margin
Customised	3,235,457.87	3,226,859.38	8,745,212.30	11,972,071.68	26.95%
Uniform	3,206,485.96	3,160,983.13	8,811,088.55	12,187,869.46	25.94%

Specifically, except for the operational cost, all other values in Table 4.5 are the sum of the particular result over the 24 scheduling hours. Notice that the amount of the microgrids' purchased energy is identical to the amount of energy the retailer bids from the wholesale markets. It turns out that the retailer purchases 7.93% less electricity and buys 3.84% more natural gas under the customised pricing scheme because natural gas is cheaper than electricity. Since the retailer purchases energy at a lower cost from the wholesale markets and gains the ability to customise the pricing strategy for each microgrid, the retailer's profit is increased compared to that under the uniform pricing scheme. Moreover, because the customised pricing scheme obtains higher profit with relatively less revenue, the retailer's net profit margin, which measures the amount of profit the retailer obtains per dollar of revenue gained, is 1.01% larger compared to the uniform pricing scheme. As a result, the proposed customised pricing scheme is superior to the uniform pricing scheme and beneficial for the retailer to acquire more profit and a higher net profit margin.

Each microgrid's energy demand under the customised pricing scheme is less than or equal to the demand under the uniform pricing scheme. This can show the effectiveness of the load curtailment program. Moreover, the CHP is heavily used as a cheaper alternative to generate electricity under the customised pricing scheme. On the contrary, the heat pump is less implemented since more heat demand is satisfied by the CHP. Additionally, because the rated capacity and power of the ES and TS in microgrid 3 are significantly larger than those in microgrid 2, the microgrid's ES and TS charging and discharging power is remarkably greater than those of microgrid 2 under both pricing schemes.

4.4.4 Effect of ES and TS Rated Capacity and Power

One of the objectives of this section is to identify the effect of the rated capacity and power of the ES and TS on the profit of the energy retailer. We consider three different scenarios to illustrate these effects. The identical base energy load demand

Table 4.5: Microgrids energy management results under customised and uniform pricing schemes. – : Not applicable.

	Customised			**Uniform**		
	Microgrid 1	**Microgrid 2**	**Microgrid 3**	**Microgrid 1**	**Microgrid 2**	**Microgrid 3**
Operational cost ($)	5,443,053.70	4,657,522.95	5,157,586.89	5,394,078.29	4,589,881.26	5,166,624.95
Purchased electricity (MWh)	24,118.78	23,275.57	20,984.30	28,006.46	22,121.82	24136.56
Purchased natural gas (kcf)	61,866.34	49,612.87	65,119.02	56,340.05	53,270.56	60460.13
Electricity demand (MWh)	27,949.10	23,022.80	25,180.26	28,189.10	23,072.72	25180.26
Natural gas demand (kcf)	11,844.79	8712.00	11,365.13	11,964.79	8832.00	11,365.13
Heat demand (MBtu)	26,293.81	24,378.08	27,594.91	26,453.64	24,529.00	27,591.80
CHP-generated electricity (MWh)	15,006.47	12,270.26	16,126.17	13,312.58	13,331.57	14728.50
Heat pump-generated heat (MBtu)	11,287.34	12,191.04	11,673.69	13,141.06	11,308.42	12988.86
ES charging power (MW/h)	–	1215.50	1682.92	–	1600.00	2294.51
ES discharging power (MW/h)	–	1096.94	1518.79	–	1443.95	2070.75
TS charging power (MBtu/h)	–	852.64	2101.07	–	1137.42	1286.82
TS discharging power (MBtu/h)	–	769.42	1896.13	–	1026.43	1161.26

Table 4.6: Profit of retailer with different scenarios.

	Scenario 1	**Scenario 2**	**Scenario 3**
Average of profit ($)	3,434,862.90	3,667,999.43	3,247,324.34
Median of profit ($)	3,433,100.16	3,656,147.99	3,248,191.95

(i.e., the base energy demand in microgrid 1) is applied for all three microgrids. In the first scenario (Scenario 1), each microgrid has differentiated ES and TS configurations (shown in Table 4.1 and 4.2). Notice that microgrid 1 does not equip either ES or TS. The rated capacity and power of ES and TS for microgrid 2 are 500 MWh and 200 MW/h, 600 MBtu and 200 MBtu/h. For microgrid 3, the rated capacity and power of ES and TS are 1000 MWh and 300 MW/h, and 1300 MBtu and 430 MBtu/h. In the second scenario (Scenario 2), none of the three microgrids are equipped with ES and TS. On the other hand, all three microgrids in the third scenario (Scenario 3) own the same configurations of ES and TS, which are the same as those of microgrid 3 in Scenario 1. Additionally, we analyse the retailer's profit and microgrids' operational costs over 15 independent runs for each scenario to obtain reliable results. The average and median values of the retailer's profit with each scenario are listed in Table 4.6. Scenario 1 represents a realistic context in which all three microgrids are equipped with different ES and TS configuration setups. On the other hand, Scenarios 2 and 3 can be considered as two extreme cases with the least and the greatest resources of ES and TS. It turns out that the retailer can make the highest profit in Scenario 2 since the underlying microgrids lack the ability to manage their energy by ES and TS. Therefore, the retailer can take advantage of this demand response deficiency. In contrast, microgrids in Scenario 3 own the highest rated capacity and power of ES and TS, which provide the most capability to manage their energy. As a result, it leads to the retailer making the least profit over the three scenarios.

To analyse the ES and TS energy management and the effect of their rated capacity and power on the operational cost of microgrids, we select the median value of the proposed bilevel problem solution in Scenario 1. Figures 4.6–4.8 display the operational results of load demand, CHP, heat pump, ES, TS, DR, renewable energies, and electricity exchange for each microgrid categorised by types of energy (i.e., electricity, natural gas and heat). The positive value of each energy indicates the energy supply for the load demand, and the negative value states the demand for

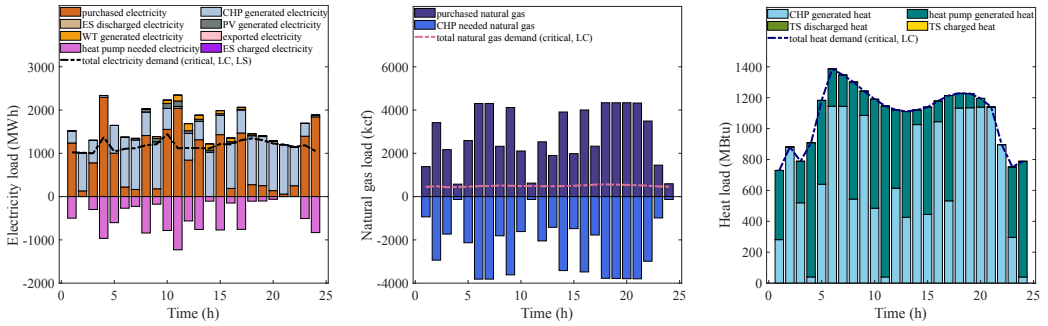


Figure 4.6: Operational results of microgrid 1.

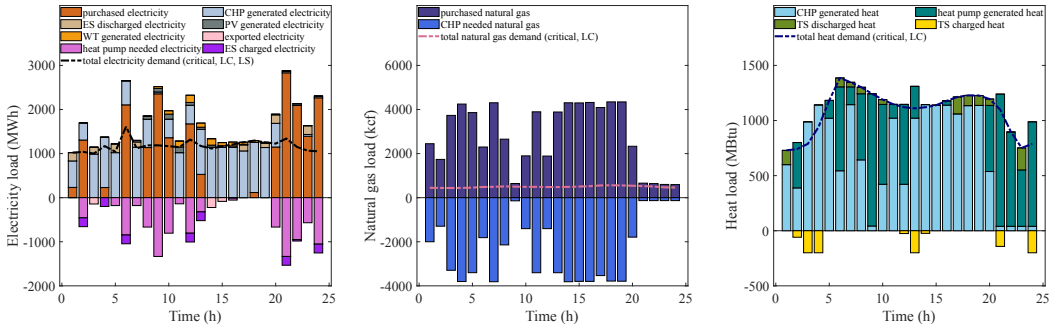


Figure 4.7: Operational results of microgrid 2.

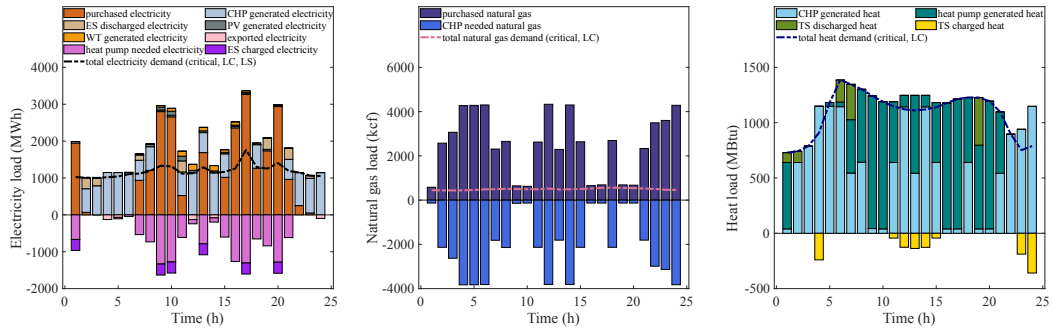


Figure 4.8: Operational results of microgrid 3.

each energy. In particular, the ES and TS operational results for microgrids 2 and 3 are shown in Tables 4.7 and 4.8. Microgrid 2 charges the ES with rated power in hours 2, 4, 6, 12, 13, 21 and 22 when the retail electricity prices are below the predetermined average price of $90/MWh. This can manage the energy usage for further use and reduce the potential cost. In contrast, ES is discharged with rated

Table 4.7: ES and TS results for microgrid 2. – : Not applicable

Time (h)	Retail Electricity Price ($/MWh)	ES Energy Level (MWh)	ES Charging Power (MW/h)	ES Discharging Power (MW/h)	Retail Natural Gas Price ($/kcf)	TS Energy Level (MBtu)	TS Charging Power (MBtu/h)	TS Discharging Power (MBtu/h)
1	100.01	50.00	0.00	190.00	41.30	163.89	0.00	129.30
2	73.72	240.00	200.00	0.00	33.06	220.01	59.08	0.00
3	106.93	76.02	0.00	155.78	33.51	410.00	200.00	0.00
4	68.81	266.02	200.00	0.00	20.26	600.00	200.00	0.00
5	105.15	55.49	0.00	200.00	44.33	600.00	0.00	0.00
6	65.89	245.49	200.00	0.00	41.39	514.83	0.00	80.91
7	99.23	120.01	0.00	119.20	31.70	471.08	0.00	41.56
8	91.80	120.01	0.00	0.00	42.80	408.90	0.00	59.07
9	80.04	120.01	0.00	0.00	55.36	408.89	0.00	0.00
10	85.11	120.01	0.00	0.00	44.86	363.23	0.00	43.38
11	89.07	120.00	0.00	0.00	27.63	363.22	0.00	0.00
12	66.12	310.00	200.00	0.00	40.06	387.41	25.46	0.00
13	74.45	500.00	200.00	0.00	28.80	577.33	199.93	0.00
14	106.84	500.00	0.00	0.00	21.27	600.00	23.87	0.00
15	98.33	500.00	0.00	0.00	38.25	600.00	0.00	0.00
16	98.75	499.99	0.00	0.00	34.21	599.99	0.00	0.00
17	104.40	357.63	0.00	135.24	47.71	438.23	0.00	153.67
18	94.22	357.63	0.00	0.00	39.77	340.75	0.00	92.61
19	109.85	254.05	0.00	98.40	26.59	246.44	0.00	89.58
20	103.01	50.00	0.00	193.85	51.12	185.41	0.00	57.97
21	66.43	240.00	200.00	0.00	53.62	320.54	142.24	0.00
22	86.17	270.53	32.14	0.00	54.54	320.53	0.00	0.00
23	109.31	60.00	0.00	200.00	56.92	110.00	0.00	200.00
24	76.35	250.00	200.00	0.00	50.99	300.00	200.00	0.00
Total:	**–**	**–**	**1432.14**	**1292.46**	**–**	**–**	**1050.57**	**948.05**

power in hours 1, 3, 5, 7, 17, 19, 20, and 23 when retail prices are above average to substitute the relatively expensive electricity source. On the other hand, since CHP and the heat pump are the primary heat source, the TS charging and discharging decisions depend on both retail electricity and natural gas prices. Analogous decisions are made by the microgrid 3.

In addition, Tables 4.7 and 4.8 indicate that increasing the rated capacity and power from microgrids 2 to 3 boosts ES and TS usage. Specifically, the total ES charging and discharging power in microgrid 2 are 1432.14 MW/h and 1292.46 MW/h, which are less than 1800.00 MW/h and 1624.45 MW/h in microgrid 3. A similar pattern happens to TS energy management (i.e., 1050.57 MBtu/h and 948.05 MBtu/h compared to 1264.09 MBtu/h and 1140.75 MBtu/h). Moreover, Table 4.9 shows the operational cost of three different microgrids. With ES and TS, microgrids 2 and 3 can reduce operational costs by 2.02% and 4.14% compared to microgrid 1,

Table 4.8: ES and TS results for microgrid 3. – : Not applicable

Time (h)	Retail Electricity Price ($/MWh)	ES Energy Level (MWh)	ES Charging Power (MW/h)	ES Discharging Power (MW/h)	Retail Natural Gas Price ($/kcf)	TS Energy Level (MBtu)	TS Charging Power (MBtu/h)	TS Discharging Power (MBtu/h)
1	72.75	785.00	300.00	0.00	45.98	556.00	0.00	89.30
2	105.28	469.21	0.00	300.00	49.69	449.34	0.00	101.32
3	102.46	245.64	0.00	212.39	50.22	449.34	0.00	0.00
4	105.44	245.63	0.00	0.00	35.79	677.53	240.20	0.00
5	109.89	245.63	0.00	0.00	31.66	677.52	0.00	0.00
6	102.56	245.63	0.00	0.00	29.20	467.25	0.00	199.75
7	102.04	100.00	0.00	138.34	54.17	130.01	0.00	320.38
8	83.42	100.00	0.00	0.00	27.65	130.01	0.00	0.00
9	62.57	385.00	300.00	0.00	42.99	130.00	0.00	0.00
10	67.83	670.00	300.00	0.00	54.99	130.00	0.00	0.00
11	106.24	354.20	0.00	300.00	33.50	171.21	43.38	0.00
12	94.18	354.20	0.00	0.00	18.17	291.18	126.29	0.00
13	64.51	639.20	300.00	0.00	57.00	421.18	136.84	0.00
14	106.82	639.20	0.00	0.00	31.65	541.81	126.99	0.00
15	89.62	639.20	0.00	0.00	29.78	582.66	43.00	0.00
16	81.47	639.19	0.00	0.00	55.27	582.66	0.00	0.00
17	61.14	924.19	300.00	0.00	49.24	582.65	0.00	0.00
18	89.72	924.19	0.00	0.00	25.13	582.65	0.00	0.00
19	105.77	608.40	0.00	300.00	53.40	130.01	0.00	430.00
20	62.25	893.40	300.00	0.00	44.58	130.01	0.00	0.00
21	100.28	577.61	0.00	300.00	42.99	130.00	0.00	0.00
22	94.81	577.60	0.00	0.00	40.82	130.00	0.00	0.00
23	96.90	500.00	0.00	73.72	35.15	309.00	188.42	0.00
24	92.04	500.00	0.00	0.00	20.99	650.00	358.95	0.00
Total:	**–**	**–**	**1800.00**	**1624.45**	**–**	**–**	**1264.09**	**1140.75**

Table 4.9: Operational costs of microgrids.

	Microgrid 1	Microgrid 2	Microgrid 3
Operational cost ($)	5,431,648.01	5,323,884.54	5,215,699.40

respectively. Furthermore, following the rise of ES and TS usage from microgrids 2 to 3, the operational costs of microgrids decrease notably. This is because microgrids with higher rated capacity and power of ES and TS are more capable of managing energy and reducing the potential cost.

4.5 Discussion

This chapter develops a customised multi-energy pricing scheme for an energy retailer that manages multiple microgrids equipped with energy converters, storage,

RES and DR programs. The proposed pricing problem is formulated as a single-leader-multiple-follower bilevel optimisation model. For future work, we would like to consider applying machine learning methods to predict the multi-energy prices, such as the maximum and minimum of the retail electricity and natural gas price for microgrids, also the multi-energy demand, such as maximum and minimum of the electricity volume the microgrids purchased from and sold to the retailer, and natural gas volume the microgrids purchased from the retailer. Chapter 5 proposes a Transformer-based model, Patchformer, to forecast long-term time-series data, such as energy prices and demand. Furthermore, we would like to develop a machine learning or data-driven model at the lower level to present the interaction between the upper-level energy retailer's price signals and the lower-level microgrids' energy management decisions. The method will benefit real-world applications due to the existence of imperfect information on lower-level agents, such as microgrids and aggregators (that is, their operation models may not be known by retailers).

Chapter 5

ADVANCING LONG-TERM MULTI-ENERGY LOAD FORECASTING WITH PATCHFORMER: A PATCH AND TRANSFORMER-BASED APPROACH

This chapter is reproduced with changes from:

- Q. Hong, F. Meng, and F. Maldonado, "Advancing Long-Term Multi-Energy Load Forecasting with Patchformer: A Patch and Transformer-Based Approach." *[In preparation for journal submission]*

In the previous chapter, a customised multi-energy pricing scheme is introduced along with its solution methods. Additionally, in Chapters 3 and 4, some predetermined data, such as electricity, natural gas and heat load demand, used to examine the model performance are simulated/ synthetic data, which may not reflect the real-world scenario. To address this issue, an effective load forecasting method needs to be developed to predict the future energy load based on the collected historical data. Therefore, this chapter proposes a Transformer-based time-series forecasting model for long-term multi-energy load prediction.

5.1 Introduction

In the context of increasing demands for long-term multi-energy load forecasting in real-world applications, this chapter introduces Patchformer, a novel model that integrates patch embedding with encoder-decoder Transformer-based architectures. To address the limitation in existing Transformer-based models, which struggle with intricate temporal patterns in long-term forecasting, Patchformer employs patch embedding, which predicts multivariate time-series data by separating it into multiple univariate data and segmenting each of them into multiple patches. This method effectively enhances the model's ability to capture local and global semantic dependencies. The numerical analysis shows that Patchformer obtains overall better prediction accuracy in both multivariate and univariate long-term forecasting on the novel Multi-Energy dataset and other benchmark datasets. In addition, the positive effect of the interdependence among energy-related products on the performance of long-term time-series forecasting across Patchformer and other compared models

is discovered, and the superiority of the Patchformer against other models is also demonstrated, which presents a significant advancement in handling the interdependence and complexities of long-term multi-energy forecasting. Lastly, Patchformer is illustrated as the only model that follows the positive correlation between model performance and the length of the past sequence, which states its ability to capture long-range past local semantic information.

5.1.1 Contributions

This chapter introduces a novel Transformer-based model that integrates patch embedding techniques for long-term multi-energy load forecasting. This model, which we have named the Patchformer, is designed to address the specific challenges of forecasting energy loads over extended periods. By processing the multivariate time series data into multiple univariate data and segmenting individual univariate data into patches, the Patchformer offers a unique approach to understanding and predicting energy consumption patterns. We believe this model represents an advancement in energy forecasting, filling a critical gap in the existing literature and improving the accuracy and efficiency across long-term forecasting models. The key contributions of this chapter are outlined as follows:

- Innovative Model Architecture: The Patchformer is designed to integrate the Patch Embedding block from PatchTST and the encoder-decoder structure from the vanilla transformer model. This Patch Embedding block treats each channel of the multivariate time series as a distinct univariate input and segments it into subseries-level patches. This approach captures local semantic information within each univariate time series and learns inter-channel relationships more effectively via a channel-independent approach, where each channel shares the same embedding and Transformer weights, enhancing efficiency in multivariate time-series forecasting. In addition, with its multi-head attention mechanisms, the encoder-decoder structure facilitates the importation of comprehensive information from the encoder to the decoder, potentially improving forecasting accuracy.

- First of its kind for Long-Term Multi-Energy Load Forecasting: To the best of our knowledge, the Patchformer is the first Transformer-based model employing a patch embedding method for long-term multi-energy load forecasting. This approach effectively addresses the complexities of predicting multi-energy load over extended periods and captures the interdependence

among different energies (e.g., electricity, gas and heat) and other energy-related products (e.g., GHG).

- Comprehensive Numerical Analysis: Experiments show the Patchformer model achieves better performance against other state-of-the-art Transformer-based models for multivariate long-term forecasting in the Multi-Energy dataset and six other benchmark datasets. In addition, the model also procures higher accuracy in univariate long-term forecasting when predicting the load of electricity and gas. Moreover, the numerical analysis also illustrates the positive effect of the interdependence among energies and energy-related products on the performance of the forecasting in the Multi-Energy dataset across Patchformer and other models via comparing the model accuracy between predicting the electricity, gas load and GHG emissions all at once and the average of the individual predictions. Lastly, the experiment demonstrates the distinct positive correlation between Patchformer's performance and the past sequence length, which shows its ability to capture long-range past local semantic information.

5.1.2 Chapter Organisation

The remainder of this chapter is organised as follows. Section 5.2 illustrates the proposed Patchformer model architecture in detail. In Section 5.3, numerical analysis has been developed and evaluated the performance of the Patchformer as well as other state-of-the-art Transformer-based models in different types of datasets, including one novel multi-energy dataset and six public benchmark datasets. In addition, the multi-energy analysis is also illustrated in the section. Lastly, Section 5.4 concludes the chapter and discusses future work.

5.2 Model Architecture

This section introduces the Patchformer model architecture, which is shown in Figure 5.1. Section 5.2.1 gives a high-level overview of the Patchformer model. Section 5.2.2 presents one of the core concepts of the model, which is the patching embedding approach. Sections 5.2.3, 5.2.4 and 5.2.5 depict each of the building blocks for the encoder and decoder. Finally, the structure of the encoder, decoder, and linear head are illustrated in Sections 5.2.6, 5.2.7 and 5.2.8, respectively.

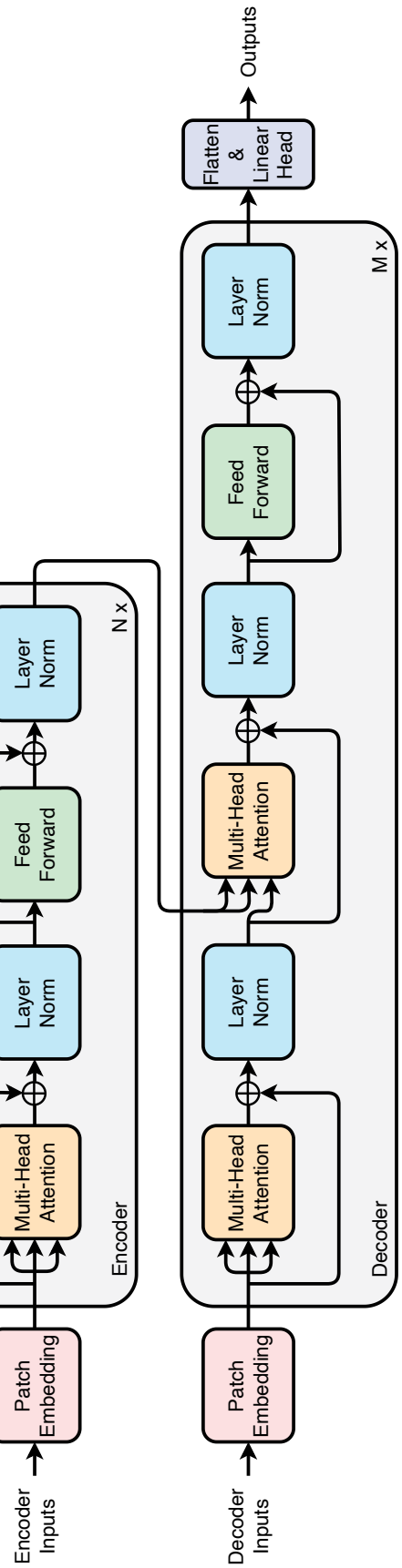


Figure 5.1: Patchformer Architecture

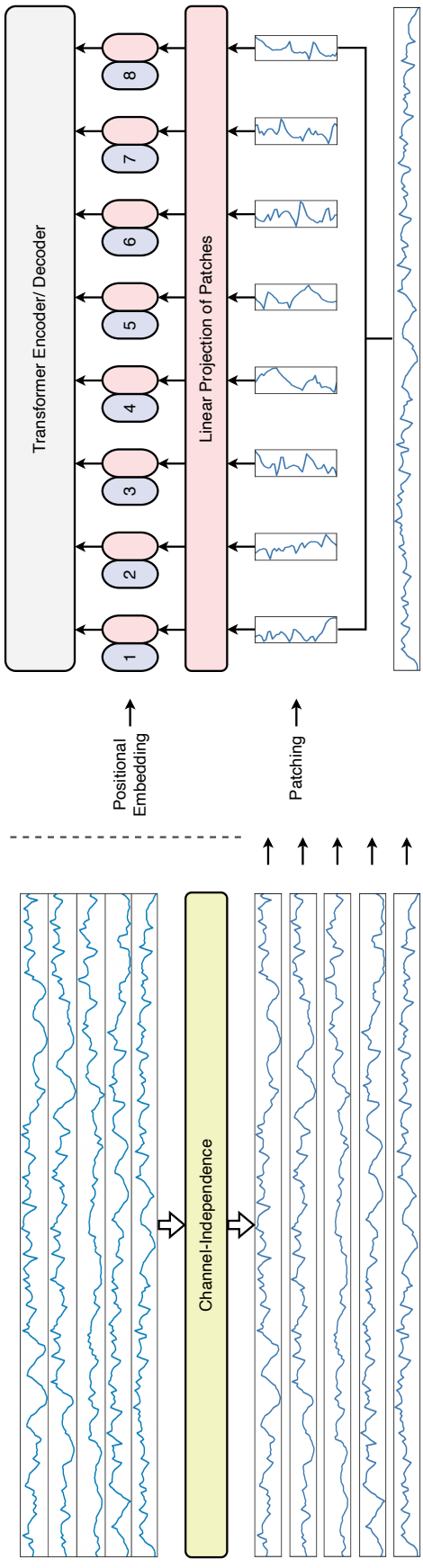


Figure 5.2: Patch Embedding

5.2.1 Model Overview

The model consists of an encoder and a decoder, while the maximum number of layers in the encoder and decoder are N and M, respectively. For the multivariate time-series forecasting, the past time sequence is denoted as $\mathcal{X} \in \mathbb{R}^{I \times C}$ with a total of length I and channels C. $\mathcal{X}^c_{t:I}$ represents a data point at time t and channel c. The future/prediction time sequence is represented as $\mathcal{Y} \in \mathbb{R}^{O \times C}$ with the total length O. Through the encoder, the information of its inputs is imported into the decoder to provide extra past information for the decoder to predict future sequences.

5.2.2 Patch Embedding

The past multivariate time sequence $\mathcal{X}$ is split as C univariate time sequences $\mathcal{X}^c \in \mathbb{R}^{1 \times I}$ before patching, which is the representation of the channel-independence. Each $\mathcal{X}^c$ is then segmented into multiple patches $\mathcal{X}^c_{z:Z} \in \mathbb{R}^{Z \times P}$ by the patch length P and stride S which is similar to the idea in CNN. Therefore, the total number of patches is calculated by $Z = \lfloor \frac{I-P}{S} \rfloor + 2$. Notice that the patching method always pads extra P time steps with the last value of the past time sequence to ensure all time-series data are in patches [78]. After patching, a one-dimensional univariate time series data $\mathcal{X}^c$ is converted to a two-dimensional matrix $\mathcal{P}^c_{patch}$ in which each row represents a patch. In addition, value embedding which projects $\mathcal{P}^c_{patch}$ from $\mathbb{R}^{Z \times P}$ dimensional space into $\mathbb{R}^{Z \times D}$ dimensional space and positional embedding is applied to optimise the patch representation and ordering. Figure 5.2 shows the procedure of the patch embedding approach. Lastly, the patch embedding block can be represented as $\mathcal{P}^c = \text{PatchEmbed}(\mathcal{X}^c)$ which are illustrated as follows:

$$\begin{aligned} \mathcal{P}^c_{patch} &= \text{Patching}(\text{Padding}(\mathcal{X}^c)) \\ \mathcal{P}^c &= \mathcal{P}^c_{patch} W_{valEmbed} + \text{PosEmbed}(\mathcal{P}^c_{patch}) \end{aligned} \tag{1}$$

where $W_{valEmbed} \in \mathbb{R}^{P \times D}$ is a learnable weight for value embedding and PosEmbed(.) denotes positional embedding. $\mathcal{P}^c \in \mathbb{R}^{Z \times D}$ is the patch embedded output.

5.2.3 Multi-Head Attention Block

The Patchformer employs the vanilla transformer's multi-head attention mechanism [79] to learn the complex local semantic information among patches. Figure 5.3 shows the scaled dot product attention and multi-head attention mechanisms. In particular, the attention layer is used to calculate the attention score by applying the softmax function to the dot product of the scaled similarity between Query $\mathcal{Q}^c_h$ and Key $\mathcal{K}^c_h$ by $\sqrt{d_k}$. The attention score is calculated by the product between the scaled dot product and Value $\mathcal{V}^c_h$. All the $\mathcal{Q}^c_h$, $\mathcal{K}^c_h$ and $\mathcal{V}^c_h$ are computed from

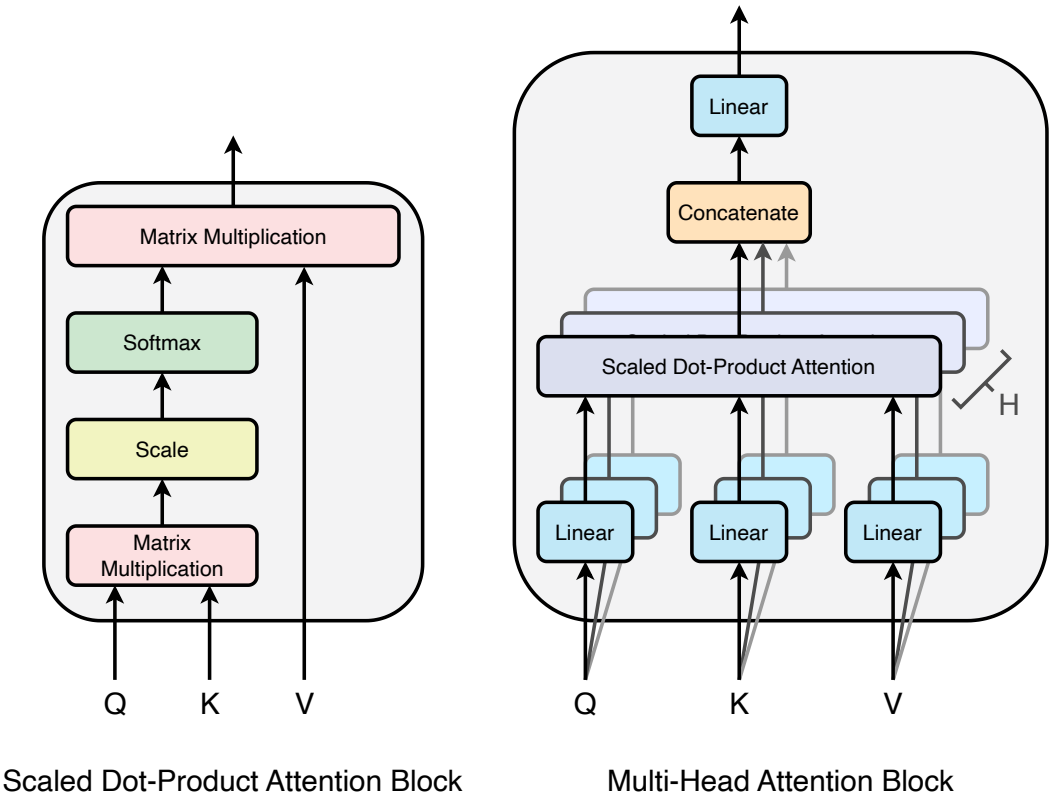


Figure 5.3: Scaled Dot-Product Attention Block (left) and Multi-Head Attention Block (right) including H attention layers (heads).

the dot product between input $\mathcal{P}^c$ and $W^{\mathcal{Q}}$, $W^{\mathcal{K}}$ and $W^{\mathcal{Q}}$, respectively. The multi-head attention block consists of H attention layers, namely, heads, which obtain the multi-head attention score by calculating the dot product between the concatenated attention scores and W^O. The multi-head attention block $\mathcal{P}^c_{attn} = \text{MultiHead}(\mathcal{P}^c)$ are formulated as follows:

$$
\begin{aligned}
\mathcal{Q}^c_h &= \mathcal{P}^c W^{\mathcal{Q}} \\
\mathcal{K}^c_h &= \mathcal{P}^c W^{\mathcal{K}} \\
\mathcal{V}^c_h &= \mathcal{P}^c W^{\mathcal{V}} \\
\mathcal{H}^c_h &= \text{Attention}(\mathcal{Q}^c_h, \mathcal{K}^c_h, \mathcal{V}^c_h) = \text{Softmax}(\frac{\mathcal{Q}^c_h \mathcal{K}^{c^T}_h}{\sqrt{d_k}})\mathcal{V}^c_h \\
\mathcal{P}^c_{attn} &= \text{Concat}(\mathcal{H}^c_1, ..., \mathcal{H}^c_H) W^O
\end{aligned}
\tag{2}
$$

where the trainable weights $W^{\mathcal{Q}}, W^{\mathcal{K}} \in \mathbb{R}^{D \times d_k}$ and $W^{\mathcal{V}} \in \mathbb{R}^{D \times D}$ result in $\mathcal{Q}^c_h, \mathcal{K}^c_h \in \mathbb{R}^{Z \times d_k}$ and $\mathcal{V}^c_h, \mathcal{H}^c_h \in \mathbb{R}^{Z \times D}$. Lastly, the multi-head attention output $\mathcal{P}^c_{attn} \in \mathbb{R}^{Z \times D}$

is obtained by projecting the row-wise concatenation among all heads $\mathcal{H}_h^c, \forall h \in \{1, ..., H\}$ with weight $W^o \in \mathbb{R}^{H*Z\times D}$.

5.2.4 Layer Normalisation Block

Layer normalisation is a technique designed to normalise the inputs across the features for each data sample in a mini-batch. Unlike batch normalisation, which normalises across the batch dimension, layer normalisation performs normalisation for each individual sample. For a given sub-layer output $\mathcal{P}_{attn}^c$, layer normalisation first computes the mean μ^c and standard deviation σ^c for each data sample independently. The output of the sub-layer is then normalised by subtracting the mean and dividing by the standard deviation. After normalisation, the process applies two learnable parameters, typically denoted as γ^c and β^c, which scale and shift the normalised value, respectively. These are trainable parameters which are learned during the training process. The layer normalisation is denoted as $\mathcal{P}_{norm}^c = \text{LayerNorm}(\mathcal{P}_{attn}^c)$ and the detailed process is shown below:

$$\begin{aligned}
\mu^c &= \frac{1}{Z\times D}\sum_{z=1}^{Z}\sum_{d=1}^{D}\mathcal{P}_{attn,z,d}^c \\
\sigma^c &= \sqrt{\frac{1}{Z\times D}\sum_{z=1}^{Z}\sum_{d=1}^{D}(\mathcal{P}_{attn,z,d}^c - \mu^c)} \\
\mathcal{P}_{norm}^c &= \gamma^c(\frac{\mathcal{P}_{attn}^c - \mu^c}{\sigma^c}) + \beta^c
\end{aligned} \tag{3}$$

Layer normalisation helps stabilise the training process and enables the training of deeper models by mitigating the vanishing or exploding gradient problems. It can lead to faster convergence in training, which is crucial for complex models like transformers that have a large number of parameters.

5.2.5 Feed Forward Block

The Feed Forward block includes two fully connected feed-forward networks (FFNs) with the ReLU activation function in between.

$$\mathcal{P}_{ffn}^c = \text{Max}(0, \mathcal{P}_{norm}^c W_1^{ffn} + b_1)W_2^{ffn} + b_2 \tag{4}$$

where the weights in the FFNs are $W_1^{ffn} \in \mathbb{R}^{D\times d_{ff}}$ and $W_2^{ffn} \in \mathbb{R}^{d_{ff}\times D}$, respectively. $b_1, b_2 \in \mathbb{R}^Z$ denote the bias terms of the FFNs. The feed forward block is summarised by $\mathcal{P}_{ffn}^c = \text{FeedForward}(\mathcal{P}_{norm}^c)$ with the output dimension (Z, D) unchanged compared to its input.

5.2.6 Encoder

With all the above blocks, the Patchformer encoder $\mathcal{P}_{en}^{c,l} = \text{Encoder}(\mathcal{P}_{en}^{c,l-1})$ can be summarised as below:

$$\begin{aligned} \mathcal{P}_{en,1}^{c,l} &= \text{LayerNorm}(\text{MultiHead}(\mathcal{P}_{en}^{c,l-1}) + \mathcal{P}_{en}^{c,l-1}) \\ \mathcal{P}_{en,2}^{c,l} &= \text{LayerNorm}(\text{FeedForward}(\mathcal{P}_{en,1}^{c,l}) + \mathcal{P}_{en,1}^{c,l}) \end{aligned} \tag{5}$$

where the $\mathcal{P}_{en,1}^{c,l}$ and $\mathcal{P}_{en,2}^{c,l}$ represent the outputs of the first and the second layer normalisation blocks, respectively. In addition, the output of the l-th encoder layer $\mathcal{P}_{en}^{c,l} = \mathcal{P}_{en,2}^{c,l}, \forall l \in \{1, ..., N\}$ and $\mathcal{P}_{en}^{c,0} = \mathcal{P}_{en}^{c}$ which is converted from the encoder inputs $\mathcal{X}_{en}^{c}$.

5.2.7 Decoder

Patchformer decoder's input $\mathcal{X}_{de}^{c} \in \mathbb{R}^{(\frac{I}{2}+O)\times 1}$ consists of two parts. The first part comes from the second half of the encoder's inputs $\mathcal{X}_{en}^{c}$, denoted as $\mathcal{X}_{en,\frac{I}{2}:I}^{c}$ to provide the most recent past information to the decoder. The second part of $\mathcal{X}_{de}^{c}$ are all zeros. The detailed formulation is shown below:

$$\mathcal{X}_{de}^{c} = \text{Concat}(\mathcal{X}_{en,\frac{I}{2}:I}^{c} + \mathcal{X}_{\text{zero}}) \tag{6}$$

where $\mathcal{X}_{\text{zero}} \in \mathbb{R}^{O\times 1}$ is used as a placeholder to form the decoder input.

After Patch Embedding, $\mathcal{P}_{de}^{c} = \text{PatchEmbed}(\mathcal{X}_{de}^{c})$, $\mathcal{P}_{de}^{c}$ is obtained as the input for the decoder. Notice that the inner and encoder-decoder multi-head attentions are designed to capture the local semantic information among input patches. By adopting the encoder's output $\mathcal{P}_{en}^{c,N}$, the decoder $\mathcal{P}_{de}^{c,l} = \text{Decoder}(\mathcal{P}_{de}^{c,l-1}, \mathcal{P}_{en}^{c,N})$ can be summarised as follows:

$$\begin{aligned} \mathcal{P}_{de,1}^{c,l} &= \text{LayerNorm}(\text{MultiHead}(\mathcal{P}_{de}^{c,l-1}) + \mathcal{P}_{de}^{c,l-1}) \\ \mathcal{P}_{de,2}^{c,l} &= \text{LayerNorm}(\text{MultiHead}(\mathcal{P}_{de,1}^{c,l}, \mathcal{P}_{en}^{c,N}) + \mathcal{P}_{de,1}^{c,l}) \\ \mathcal{P}_{de,3}^{c,l} &= \text{LayerNorm}(\text{FeedForward}(\mathcal{P}_{de,2}^{c,l}) + \mathcal{P}_{de,2}^{c,l}) \end{aligned} \tag{7}$$

where $\mathcal{P}_{de}^{c,l} = \mathcal{P}_{de,3}^{c,l}, \forall l \in \{1, ..., M\}$ represents the outputs for the l-th decoder layer. In addition, $\mathcal{P}_{de}^{c,0} = \mathcal{P}_{de}^{c}$. $\mathcal{P}_{de,i}^{c,l}, \forall i \in \{1, 2, 3\}$ denote the outputs of layer normalisation blocks in the l-th decoder layer, respectively.

5.2.8 Flatten and Linear Head

The Flatten and Linear Head block is designed first to flatten the output of the decoder $\mathcal{P}_{de}^{c,M}$ from dimension (Z_{de}, P) to $(1, Z_{de} * P)$. Second, the final prediction

sequence is obtained by converting the output dimension again to $(1, O)$. The detailed formulations are shown below:

$$\begin{aligned} \mathcal{Y}^c_{flattened} &= \text{Flatten}(\mathcal{P}^{c,M}_{de}) \\ \mathcal{Y}^c &= \mathcal{Y}^c_{flattened} W_y \end{aligned} \tag{8}$$

where $\mathcal{Y}^c_{flattened} \in \mathbb{R}^{1 \times Z_{de} * P}$, $\mathcal{P}^{c,M}_{de} \in \mathbb{R}^{Z_{de} \times P}$. The final prediction time sequence $\mathcal{Y}^c \in \mathbb{R}^{1 \times O}, \forall c \in \{1, ..., C\}$ is obtained by a linear transformation with weight $W_y \in \mathbb{R}^{Z_{de} * P \times O}$.

5.3 Numerical Analysis

In this section, the performance comparison and evaluation between the proposed Patchformer and other state-of-the-art models: Autoformer [76], Crossformer [89], and Transformer [79] and multi-energy analysis are discussed in detail. Firstly, the datasets which contain a novel Multi-Energy dataset, six public benchmark datasets and experimental setup are introduced in Section 5.3.1. The performance of multivariate forecasting for Patchformer and other models across different datasets are discussed in Section 5.3.2. Sections 5.3.3-5.3.5 analyse the Patchformer performance on the Multi-Energy dataset from various perspectives in detail. In particular, Section 5.3.3 studies the univariate forecasting performance among Patchformer and other models on the Multi-Energy dataset by predicting electricity, gas load and GHG emission. In addition, the effect of the interdependence among electricity, gas load and GHG emission on the performance of the LTTSF on the Multi-Energy dataset is illustrated in Section 5.3.4. Lastly, in Section 5.3.5, the forecasting performance is compared between Patchformer and other models with different past sequence lengths.

5.3.1 Datasets and Experimental Setup

The proposed Patchformer model is evaluated on seven datasets, including the novel and comprehensive Multi-Energy dataset [123] and other six datasets that are well known and have been utilised as benchmarks, publicly available on [76]. Here is the description of the seven datasets: (1) *Multi-Energy* dataset records energy-related data collected on the Temple campus at Arizona State University, which includes hourly electricity, gas, and heat load demand, renewable energy generation, and GHG emissions for each building from 24 July 2015 to 12 September 2020. (2) *Exchange* dataset collects the daily exchange rate of eight different countries from 1 January 1990 to 10 October 2010. (3) *Weather* dataset is a collection of 21 meteorological indicators (e.g., air temperature, humidity and precipitation) every

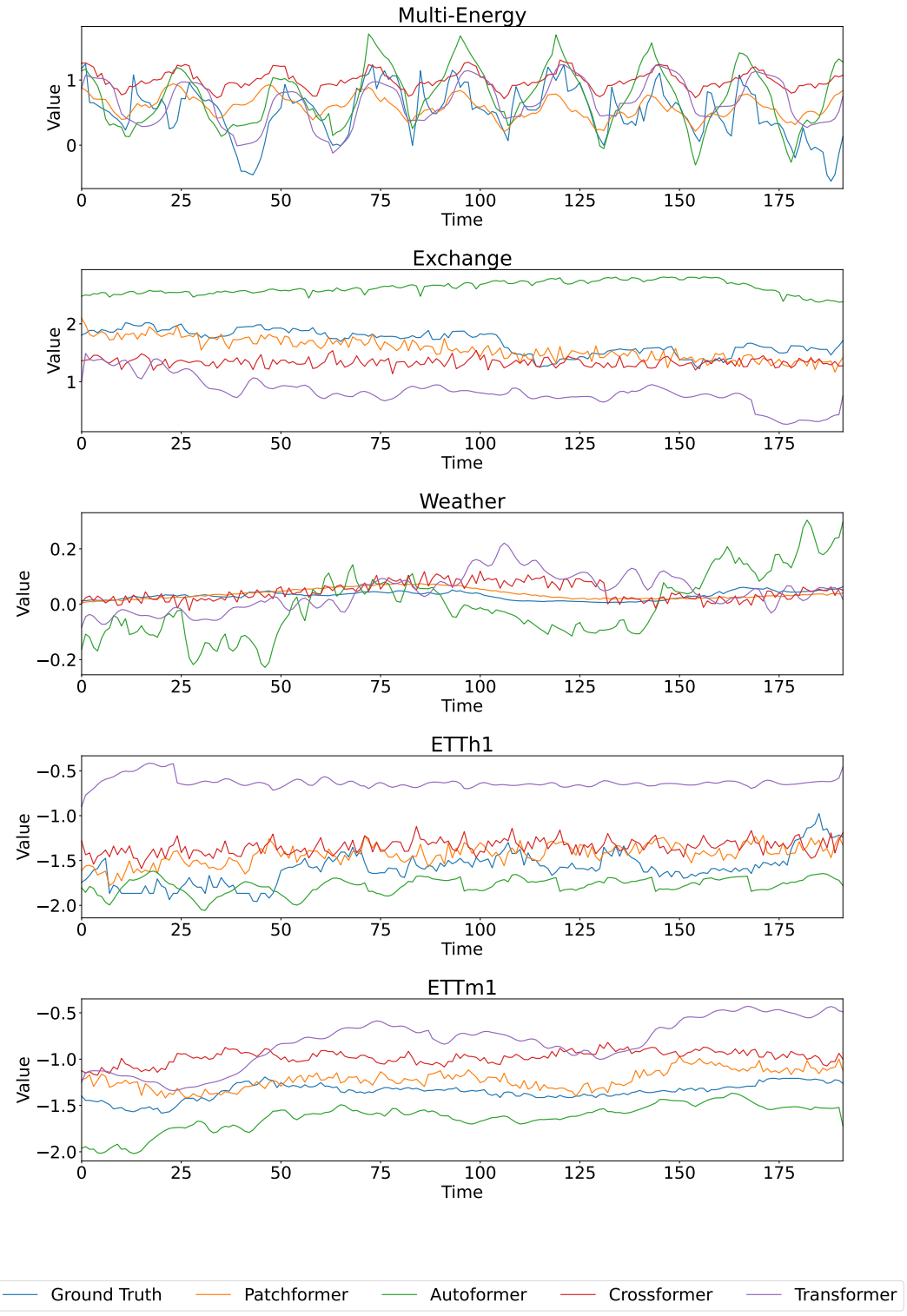


Figure 5.4: Visualisation of 192 step forecasting on Multi-Energy, Exchange, Weather, ETTh1 and ETTm1 datasets with the past sequence length = 96.

Table 5.1: Statistics of datasets.

Datasets	Multi-Energy	Exchange	Weather	ETTh1	ETTh2	ETTm1	ETTm2
Features	19	8	21	7	7	7	7
Length	49415	7588	52696	17420	17420	69680	69680

10 minutes in the entire year of 2020. (4) and (5) *ETTh1* and *ETTh2* datasets record the hourly data (e.g., load and oil temperature) from two different electricity transformers from 1 July 2016 to 26 June 2018. Similarly, (6) and (7) *ETTm1* and *ETTm2* datasets collect every 15 minutes data from the two electricity transformers in the same time period. The statistics of the seven datasets are shown in Table 5.1, which shows the total number of input features and length of the time-series observations. The Patchformer and other models are written in PyTorch and run on Ubuntu 22.04.3 LTS x86_64 with Intel Xeon (8) @ 2.000GHz and 52GB of RAM. The GPU uses NVIDIA Tesla V100 SXM2 16GB.

Moreover, in this section, the mean squared error (MSE) and mean absolute error (MAE) are applied as experimental evaluation indicators to reflect the forecasting accuracy of the proposed Patchformer and other comparison models. The definitions of the two evaluation indexes are formulated as follows:

$$\begin{aligned} \text{MSE} &= \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \\ \text{MAE} &= \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \end{aligned} \tag{9}$$

where n is the total number of the data. The actual and predicted values are denoted as y_i and $\hat{y}_i$ at the i-th time step of the dataset, respectively.

5.3.2 Multivariate Forecasting on Different Datasets

In this section, the performance of multivariate forecasting among different models on the above-mentioned seven datasets is compared. Multivariate forecasting considers the historical data of several variables to forecast one or more of these variables while taking into account the interdependence between multiple input variables. The same number of input and output variables is applied to analyse the performance of the multivariate forecasting in this section. The hyperparameters used in the section are shown in Table 5.2. The prediction and past sequence length are set to be $\mathcal{Y} \in \{96, 192, 336, 720\}$ and $\mathcal{X} = 96$, respectively. The evaluation results are shown in Table 5.3. Patchformer consistently outperformed competing

Table 5.2: Hyperparameters of different models for multivariate forecasting.

Models	Hyperparameters
Patchformer	patch length = 16, stride = 8, encoder number = 2, decoder number = 1, model dimension = 512, label length = sequence length/2, batch size = 32, head = 8, learning rate = 0.0001, dropout rate = 0.1, Optimiser = Adam, loss function = MSE, epoch = 10
Autoformer	topK = 5, encoder number = 2, decoder number = 1, model dimension = 512, label length = sequence length/2, batch size = 32, learning rate = 0.0001, dropout rate = 0.1, Optimiser = Adam, loss function = MSE, epoch = 10
Crossformer	topK = 5, encoder number = 2, decoder number = 1, model dimension = 512, label length = sequence length/2, batch size = 32, learning rate = 0.0001, dropout rate = 0.1, Optimiser = Adam, loss function = MSE, epoch = 10
Transformer	encoder number = 2, decoder number = 1, model dimension = 512, label length = sequence length/2, batch size = 32, head = 8, learning rate = 0.0001, dropout rate = 0.1, Optimiser = Adam, loss function = MSE, epoch = 10

models at prediction sequence lengths of 192, 336 and 720 on the Multi-Energy dataset introduced in this chapter. Notably, at the 720-step forecast, the Patchformer achieved an MSE of 0.121 and an MAE of 0.193, which is 15.70% and 11.40% less than the second-best scores, respectively. It indicates its capability to capture the complex interdependencies among the multiple energy vectors over long-term horizons. In addition, at the extended horizon of 96, 192 and 336 steps, the Patchformer's performance remains competitive, with all the best MSE and MAE, except for two second-best MSE at 96 and 336 steps, demonstrating the model's robustness in long-term forecasting. Its superiority in the multi-energy forecasting domain is critical for modern IMES systems.

For benchmark datasets in multivariate time series forecasting, the Patchformer exhibited varying degrees of efficacy. In the Exchange Rate dataset, the Patchformer performs the best across all prediction lengths and is optimal at shorter horizons but showed a significant decline as the prediction length increased, with the MSE rising sharply to 1.071 at the 720-step horizon. This suggests a potential vulnerability in the Patchformer's architecture when dealing with the non-stationary and volatile nature of financial time series over long-term periods.

In Weather forecasting, the Patchformer maintained competitiveness and the best performance across all horizons, with the best MSE and MAE for all prediction

Table 5.3: Multivariate time-series forecasting results with Patchformer. We use prediction lengths $\mathcal{Y} \in \{96, 192, 336, 720\}$ and past sequence length $\mathcal{X} = 96$. The best results are in **bold**, and the second best is in underlined.

Model		Patchformer		Autoformer		Crossformer		Transformer	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Multi-Energy	96	0.062	**0.135**	**0.061**	0.150	0.067	0.150	0.063	0.147
	192	**0.073**	**0.149**	0.075	0.161	0.086	0.172	0.079	0.161
	336	0.090	**0.167**	**0.083**	0.173	0.165	0.243	0.153	0.217
	720	**0.121**	**0.193**	0.140	0.215	0.221	0.274	0.257	0.273
Exchange	96	**0.089**	**0.217**	0.153	0.285	0.256	0.367	0.550	0.579
	192	**0.190**	**0.321**	0.277	0.383	0.468	0.508	0.934	0.734
	336	**0.387**	**0.471**	0.471	0.513	0.975	0.763	1.328	0.904
	720	**1.071**	**0.769**	1.107	0.818	1.620	1.029	2.565	1.336
Weather	96	**0.175**	**0.231**	0.342	0.385	0.177	0.242	0.353	0.412
	192	**0.213**	**0.274**	0.321	0.374	0.222	0.289	0.574	0.542
	336	**0.263**	**0.311**	0.347	0.384	0.276	0.338	0.631	0.584
	720	**0.339**	**0.369**	0.415	0.418	0.372	0.411	0.850	0.686
ETTh1	96	0.425	0.444	0.529	0.487	**0.419**	**0.439**	0.773	0.684
	192	**0.484**	**0.477**	0.509	0.486	0.539	0.517	0.886	0.744
	336	0.549	0.512	**0.508**	**0.494**	0.709	0.638	0.966	0.770
	720	0.603	0.566	**0.542**	**0.520**	0.721	0.622	1.016	0.800
ETTh2	96	**0.342**	**0.387**	0.375	0.410	0.790	0.612	2.633	1.291
	192	0.473	0.459	**0.443**	**0.449**	1.830	1.041	5.961	2.007
	336	**0.475**	**0.478**	0.501	0.496	1.863	1.088	5.811	1.948
	720	0.600	0.538	**0.496**	**0.499**	2.833	1.447	2.964	1.399
ETTm1	96	0.364	**0.393**	0.512	0.485	**0.362**	0.403	0.725	0.620
	192	0.411	**0.421**	0.539	0.494	**0.388**	0.422	0.870	0.703
	336	**0.437**	**0.447**	0.587	0.523	0.617	0.579	1.062	0.790
	720	**0.499**	**0.482**	0.650	0.535	0.931	0.722	1.063	0.789
ETTm2	96	**0.214**	**0.308**	0.230	0.314	0.250	0.333	0.469	0.500
	192	0.321	0.380	**0.281**	**0.340**	0.888	0.694	1.438	0.891
	336	0.373	0.415	**0.338**	**0.373**	1.451	0.850	1.154	0.818
	720	0.592	0.529	**0.459**	**0.441**	2.678	1.148	2.675	1.208

steps. This indicates the Patchformer's adeptness at modelling environmental time series data, which often have clearer temporal patterns and seasonality.

With the ETTh1, ETTh2, ETTm1, and ETTm2 datasets, which contain the data from electricity transformers, including load and oil temperature, the Patchformer displays either the best or second-best MSE and MAE for all prediction lengths, which indicates the model's robustness and reliability. In particular, Patchformer obtains the best MSE and MAE for all different prediction lengths on the ETTm1

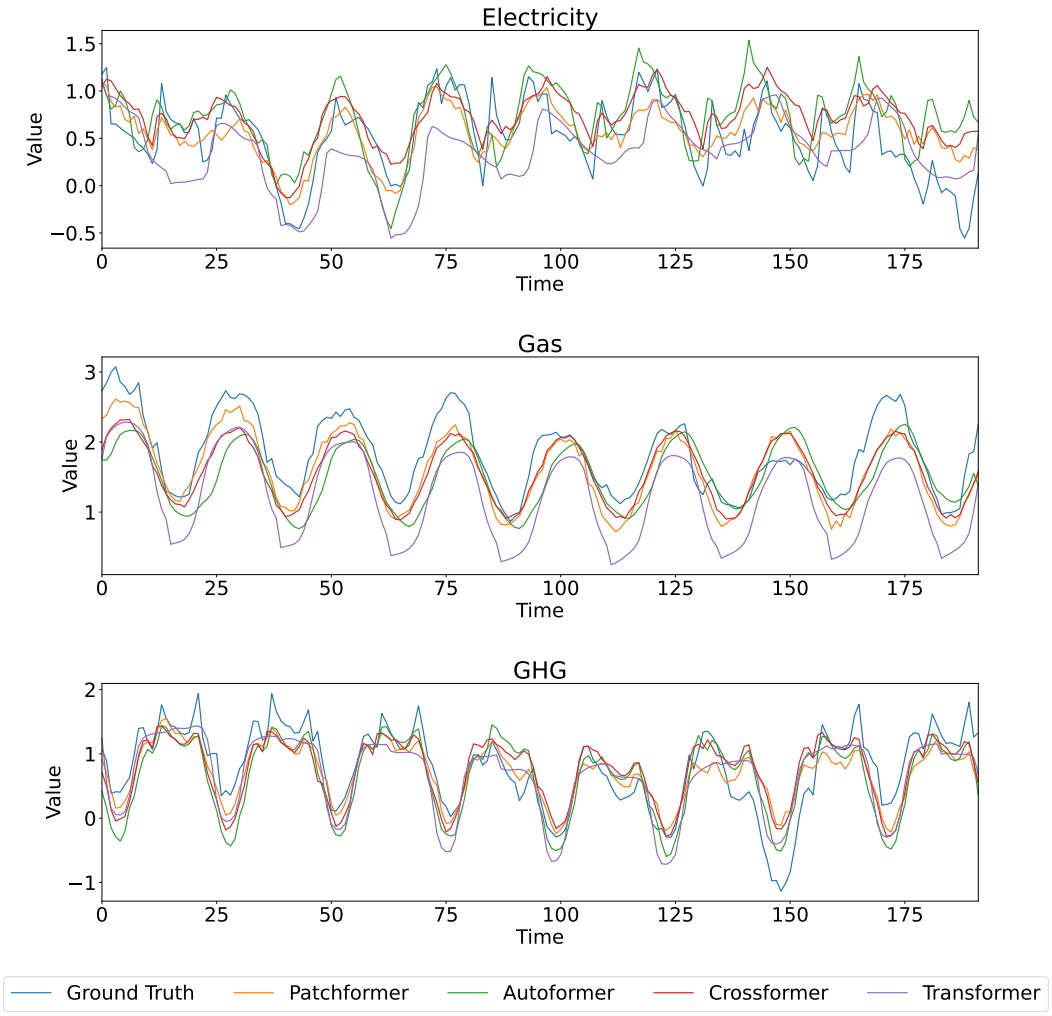


Figure 5.5: Visualisation of 192 step forecasting on Electricity, Gas and GHG in Multi-Energy dataset with the past sequence length = 336.

dataset, except for MSE at 96 and 192 steps, which are relatively close to the best results obtained by Crossformer.

Figure 5.4 visualises the time-series predictions of the four models on Multi-Energy, Exchange, Weather, ETTh1 and ETTm1 datasets. In particular, the predicted values of the electricity load on the Multi-Energy dataset are visualised to show the models' performance. As a result, the Patchformer exhibits strong performance across multiple datasets since it efficiently captures the trend and is closest to the ground truth. The experimental results indicate that the Patchformer can handle long-term predictions efficiently, especially in domains where the data has explicit seasonality patterns. With high volatility and irregular patterns, such as financial markets, the

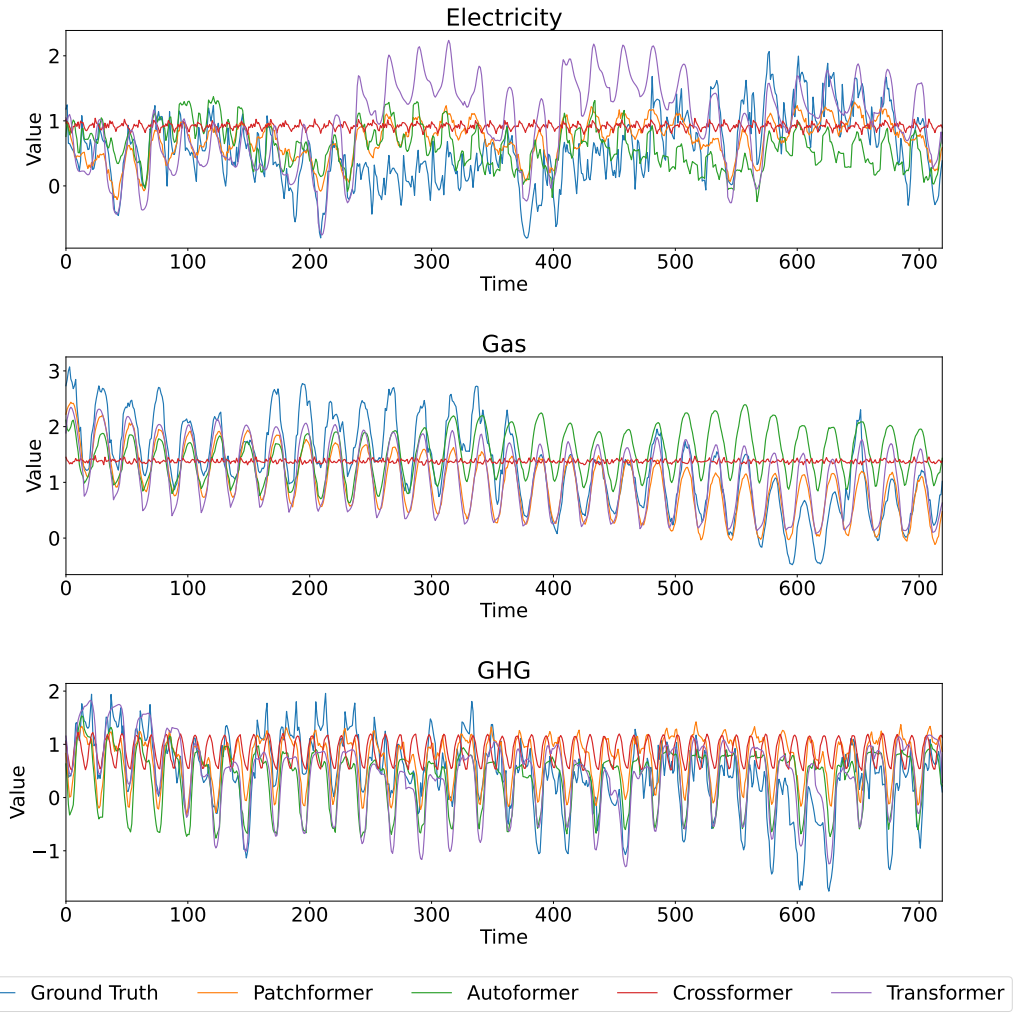


Figure 5.6: Visualisation of 720 step forecasting on Electricity, Gas and GHG in Multi-Energy dataset with the past sequence length = 336.

Patchformer's long-term forecasting performance can be improved in future work.

5.3.3 Univariate Forecasting

In addition to multivariate forecasting, this section presents univariate forecasting results of Patchformer and other compared models on the Multi-Energy dataset. Univariate forecasting focuses on forecasting the future values of that single variable based on its own historical data. The output is the prediction of future values of the same single variable without considering any other input variables. The features chosen to be analysed are electricity and gas load demand, as well as GHG emissions. The hyperparameters of Patchformer among different prediction lengths are shown in Table 5.4. The past sequence length for all models is fixed at

Table 5.4: Hyperparameters of Patchformer for univariate forecasting.

Prediction Length	Hyperparameters
96	patch length = 16, stride = 8, encoder number = 2, decoder number = 1, model dimension = 512, label length = sequence length/2, batch size = 32, head = 16, fully connected layer dimension = 2048, learning rate = 0.0001, dropout rate = 0.1, Optimiser = Adam, loss function = MSE, epoch = 10
192	patch length = 16, stride = 8, encoder number = 2, decoder number = 1, model dimension = 256, label length = sequence length/2, batch size = 32, head = 16, fully connected layer dimension = 1024, learning rate = 0.0001, dropout rate = 0.1, Optimiser = Adam, loss function = MSE, epoch = 10
336	patch length = 16, stride = 8, encoder number = 2, decoder number = 1, model dimension = 512, label length = sequence length/2, batch size = 32, head = 16, fully connected layer dimension = 2048, learning rate = 0.0001, dropout rate = 0.1, Optimiser = Adam, loss function = MSE, epoch = 10
720	patch length = 16, stride = 8, encoder number = 2, decoder number = 1, model dimension = 128, label length = sequence length/2, batch size = 32, head = 16, fully connected layer dimension = 1024, learning rate = 0.0001, dropout rate = 0.1, Optimiser = Adam, loss function = MSE, epoch = 10

336 time steps, and the forecasting horizons are 96, 192, 336, and 720 time steps. Table 5.5 shows univariate forecasting results. Patchformer outperforms all other models on electricity and gas forecasting as it receives the best MSE and MAE results across all prediction lengths. For the forecasting results on the GHG dataset, Patchformer has room for improvement, which implies there may not be a single model that universally outperforms others across all metrics, electricity and gas demand, GHG emissions, and prediction lengths. In addition, Figure 5.5 and 5.6 visualise 336 steps past sequence length forecasting with all models on Electricity, Gas and GHG in the Multi-Energy dataset when the prediction lengths are 192 and 720 steps, respectively. The time-series patterns among Electricity, Gas and GHG are shown to be different in both figures. Also, it is fairly obvious to observe that the predictions of Patchformer are the closest to the ground truth value when predicting electricity and gas in Figure 5.5 and 5.6, which indicates the excellent performance of Patchformer. Furthermore, from Figure 5.6, the prediction of Crossformer is stated as relatively consistent across time while other models fluctuate dramatically.

Table 5.5: Univariate multi-energy forecasting results with Patchformer. We use prediction lengths $\mathcal{Y} \in \{96, 192, 336, 720\}$ and past sequence length $\mathcal{X} = 336$. The best results are in **bold** and the second best are underlined.

Model		Patchformer		Autoformer		Crossformer		Transformer	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Electricity	96	**0.114**	**0.262**	0.176	0.326	0.116	0.264	0.154	0.304
	192	**0.148**	**0.300**	0.230	0.372	0.151	0.306	0.186	0.336
	336	**0.181**	**0.332**	0.291	0.427	0.227	0.375	0.226	0.369
	720	**0.272**	**0.412**	0.290	0.428	0.340	0.472	0.497	0.556
Gas	96	**0.071**	**0.195**	0.114	0.262	0.074	0.196	0.101	0.232
	192	**0.084**	**0.221**	0.121	0.269	0.095	0.230	0.120	0.262
	336	**0.106**	**0.244**	0.193	0.340	0.130	0.273	0.118	0.264
	720	**0.143**	**0.293**	0.335	0.439	0.270	0.405	0.169	0.307
GHG	96	0.141	0.290	0.242	0.368	**0.135**	**0.285**	0.172	0.323
	192	0.202	0.354	0.217	0.369	**0.159**	**0.312**	0.184	0.335
	336	0.280	0.413	0.226	**0.373**	**0.225**	0.377	0.266	0.399
	720	0.426	0.518	**0.311**	**0.442**	0.550	0.583	0.317	0.449

5.3.4 Multi-Energy Forecasting Comparison

In this section, the effect of the interdependence among energy-related products on the performance of the time-series forecasting in the Multi-Energy dataset is discussed. The features in the dataset chosen to be analysed are electricity and gas load demand, as well as GHG emissions since they are highly interrelated in nature. The prediction, past sequence length and hyperparameters of the Patchformer model are identical to Section 5.3.3. The forecasting results are shown in Table 5.6, in which All-at-Once means to predict electricity and gas load and GHG emission simultaneously, whereas Electricity, Gas and GHG are predicted individually. The feature Average is calculated by the mean of the features Electricity, Gas and GHG. In Table 5.6, Patchformer outperforms all other three models across all features selected from the Multi-Energy dataset as it achieves the most numbers of best MSE and MAE results than other models. In addition, the Patchformer forecasting results at 336 and 720 prediction lengths when predicting electricity, gas load, and GHG emission all at once are better than predicted individually. Furthermore, the percentage difference of Patchformer MSE and MAE between All-at-Once and Average at 96 and 192 prediction lengths are insignificant (MSE: 2.65% and 1.36%, MAE: 1.96% and 0.34%). Moreover, the results of All-at-Once are generally better than Average for all other models, especially Autoformer and Transformer. Overall, the pattern in which predicting multi-energy results all at once is better than predicting them individually demonstrates the interdependence among electricity,

Table 5.6: Multi-energy forecasting results with Patchformer. We use prediction lengths $\mathcal{Y} \in \{96, 192, 336, 720\}$ and past sequence length $\mathcal{X} = 336$. The best results are in **bold**, the second best is underlined, and the better results between All-at-Once and Average are shaded.

Model		Patchformer		Autoformer		Crossformer		Transformer	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
All-at-Once	96	**0.113**	**0.255**	0.139	0.286	0.130	0.274	0.146	0.282
	192	**0.147**	**0.292**	0.156	0.304	0.173	0.318	0.233	0.357
	336	**0.179**	**0.324**	0.187	0.336	0.368	0.467	0.318	0.408
	720	0.255	0.389	**0.234**	**0.378**	0.499	0.540	0.471	0.501
Electricity	96	**0.115**	**0.263**	0.154	0.306	0.134	0.284	0.186	0.321
	192	**0.145**	**0.297**	0.161	0.311	0.230	0.375	0.255	0.374
	336	**0.178**	**0.330**	0.209	0.355	0.246	0.391	0.491	0.518
	720	0.311	0.440	**0.289**	**0.427**	0.584	0.599	0.539	0.554
Gas	96	**0.073**	**0.198**	0.111	0.253	0.088	0.216	0.101	0.233
	192	**0.085**	**0.221**	0.131	0.281	0.114	0.253	0.159	0.294
	336	**0.101**	**0.241**	0.128	0.279	0.174	0.315	0.192	0.320
	720	**0.137**	**0.287**	0.196	0.330	0.382	0.467	0.210	0.338
GHG	96	**0.139**	**0.288**	0.171	0.323	0.154	0.305	0.188	0.333
	192	**0.206**	**0.355**	0.211	0.362	0.212	0.359	0.360	0.453
	336	0.312	0.438	**0.225**	**0.370**	0.311	0.439	0.432	0.492
	720	0.439	0.522	**0.273**	**0.411**	0.682	0.643	0.551	0.565
Average	96	**0.109**	**0.250**	0.145	0.294	0.125	0.268	0.159	0.295
	192	**0.145**	**0.291**	0.168	0.318	0.185	0.329	0.258	0.374
	336	0.197	0.336	**0.187**	**0.335**	0.244	0.382	0.371	0.443
	720	0.296	0.416	**0.252**	**0.389**	0.549	0.570	0.433	0.486

gas load and GHG emission can be captured by Patchformer and other models and improve the forecasting performance.

5.3.5 Different Length of Past Time Sequences

In this section, the forecasting performance between Patchformer and other models with different lengths of past time sequences has been compared and shown in Figure 5.7 and Table 5.7. Five different past sequence lengths are selected: $\{24, 48, 96, 192, 336\}$. Two prediction lengths are $\{96, 720\}$. Intuitively, the models' forecasting performance and the length of past sequences should be positively correlated. However, based on the argument in [124], this principle may not work for the majority of the Transformer-based models since they cannot capture the temporal local information efficiently. Figure 5.7 also proves the phenomenon, which shows that except for the proposed Patchformer, all other models do not follow the positive correlation between model performance and the past sequence length. This shows Patchformer's ability to capture long-range, past local semantic information.

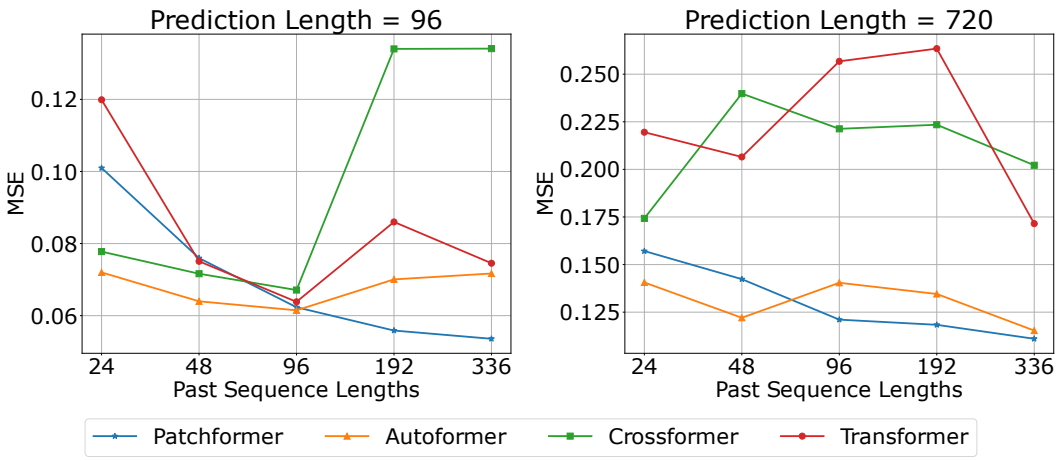


Figure 5.7: Performance analysis with different past time sequence lengths on Multi-Energy dataset.

Table 5.7: Multi-Energy load forecasting results with Patchformer on Multi-Energy dataset. We use prediction lengths $\mathcal{Y} \in \{96, 720\}$ and past sequence lengths $\mathcal{X} \in \{24, 48, 96, 192, 336\}$. The best results are in **bold** and the second best are underlined.

Model		Patchformer		Autoformer		Crossformer		Transformer	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Pred: 96	24	0.101	0.170	**0.072**	**0.155**	0.078	0.158	0.120	0.182
	48	0.076	0.150	**0.064**	**0.147**	0.072	0.155	0.075	0.151
	96	0.062	**0.135**	**0.061**	0.150	0.067	0.150	0.063	0.147
	192	**0.056**	**0.128**	0.070	0.167	0.134	0.212	0.086	0.165
	336	**0.054**	**0.126**	0.072	0.162	0.134	0.214	0.075	0.160
Pred: 720	24	0.158	0.224	**0.141**	**0.224**	0.174	0.250	0.220	0.256
	48	0.142	0.214	**0.122**	**0.210**	0.240	0.287	0.207	0.244
	96	**0.121**	**0.193**	0.140	0.215	0.221	0.274	0.257	0.273
	192	**0.118**	**0.191**	0.135	0.216	0.223	0.282	0.263	0.280
	336	**0.111**	**0.183**	0.115	0.223	0.202	0.289	0.172	0.226

In addition, Table 5.7 shows that, as the length of the past sequence increases, Patchformer's performance improves and surpasses all other models.

5.4 Discussion

The chapter presents Patchformer, a novel Transformer-based model for long-term multi-energy load forecasting. It addresses the challenges in IMES forecasting by predicting each feature in multivariate time-series data independently and segmenting it into patches, which can benefit the capture of interdependence among different

features and the reception of local semantic information. The Patchformer's architecture, including its patch embedding and encoder-decoder mechanism, is illustrated in detail. In numerical analysis, the Patchformer model demonstrates superior performance against other state-of-the-art models in multivariate long-term forecasting by the unique Multi-Energy dataset and six other benchmark datasets. Furthermore, the model's performance in univariate long-term forecasting when predicting electricity and gas load is superior to other models. The experiment also demonstrates the positive effect of the interdependence among energy-related products on the performance of the time-series forecasting in the Multi-Energy dataset across Patchformer and other models by comparing the forecasting results between predicting the electricity, gas load and GHG emissions all at once and the average of the individual predictions. In addition, the positive correlation between Patchformer's performance and the past sequence length is stated, which shows its capability to capture long-range past local semantic information. Future research directions can be addressed as follows: handling volatile data, like in financial markets, capturing long-term dependencies, and adapting to the non-stationarity of real-world datasets, which need to be improved in further study. In addition, for univariate forecasting, such as GHG data, the model performance needs to be enhanced accordingly. Lastly, the next chapter concludes the thesis and provides future research directions for Chapters 3-5.

Chapter 6

CONCLUSION AND FUTURE DIRECTIONS

The last three chapters illustrate the three research topics, respectively. In particular, Chapter 3 discusses the strategic bidding and offering problems among energy suppliers. Chapter 4 introduces the retail pricing schemes of energy suppliers. Lastly, the long-term multi-energy load forecasting problem for energy suppliers is studied in Chapter 5. In this chapter, the conclusion and future directions for each research topic are provided in Sections 6.1 and 6.2, respectively.

6.1 Conclusion

Chapter 3 proposes a bilevel game-theoretic framework for strategic retailers who aim to maximise their profits by participating in both DAW and local electricity markets. In terms of the proposed bilevel model, customers' welfare function and switching behaviours are considered to be the lower-level problems along with the market-clearing problems for the DAW and local electricity markets, respectively. Furthermore, the proposed model is formulated as an MPEC problem and then reformulated to a MIQP model. By extending the above bilevel model from a single leader (one retailer) to multiple leaders (multiple retailers), a Bertrand competition model is adopted to model the interactions among multiple leaders at the upper level. Finally, the resulting multi-leader, multi-follower Stackelberg game model is reformulated as an EPEC problem and solved by the diagonalisation algorithm. Extensive numerical results are present to demonstrate the feasibility and effectiveness of the proposed bilevel strategic decision-making framework, the effect of customers' switching behaviours on decision-making, and the benefits of different market players (e.g. retailers and customers). In particular, results show that incentivizing customers' switching behaviours can decrease strategic retailers' retail prices and profits. However, switching may not always benefit customers' welfare due to customers' need for balance between the electricity purchasing cost (i.e., electricity price) and the electricity consumption level. In addition, similar ESS charging/discharging decisions among strategic retailers are observed when customers' switching behaviours are enhanced.

In chapter 4, a customised multi-energy pricing scheme is proposed for an energy retailer that manages multiple microgrids equipped with energy converters, storage,

RES and DR programs. The proposed pricing problem is formulated as a bilevel optimisation model. The energy retailer is the leader at the upper level in maximising profit. Each multi-energy microgrid acting as a follower minimises the operational cost at the lower level. In addition, three hybrid metaheuristic algorithms (i.e., PSO, GA and SA) coordinated with the MILP program are developed to solve the model efficiently. Through numerical analyses, the GA-based hybrid solution algorithm has been proven to perform best against others. The customised pricing scheme presents superiority compared to the uniform pricing scheme. In addition, since increasing the rated capacity and power of the ES and TS can improve the microgrids' energy management capability, the retailer's profit and microgrids' operational costs will be reduced accordingly.

Chapter 5 introduces Patchformer, an innovative model designed for long-term multi-energy load forecasting. This model combines patch embedding with encoder-decoder Transformer architectures to tackle the inherent limitations in existing Transformer-based models, which often struggle with complex temporal patterns in long-term forecasting scenarios. By segmenting multivariate time series into multiple univariate patches, Patchformer significantly enhances its ability to recognise both local and global dependencies within the data. The chapter's numerical studies demonstrate Patchformer's superior performance over current state-of-the-art models in long-term multivariate and univariate time-series forecasting tasks, especially on the Multi-Energy dataset. Furthermore, the positive effect of the inter-dependence among energy-related products on the performance of the time-series forecasting in the Multi-Energy dataset is also illustrated in detail. In addition, Patchformer's positive correlation between its performance and the past sequence length is addressed to show its capacity to capture long-term past local semantic information. These advancements show a significant step forward in addressing the complexities of multi-energy load forecasting, presenting a promising new forecasting model for managing the intricacies of IMES. Patchformer's development is reflected in a broader shift towards more sophisticated, efficient, and accurate forecasting methods, essential for the sustainable and reliable operation of modern energy systems with growing demand and increasing system complexity.

6.2 Future Directions

For the study of strategic bidding and offering problems for energy suppliers, while considering customers' switching behaviour, the modelling of customers' behaviours could also consider existing demand response programs, such as load shifting and

curtailment, to better reflect the scenarios of future energy systems. Applying the network congestion and locational marginal prices to the proposed bilevel model and investigating the effects on the main findings of Chapter 3 is also a promising research direction. In addition, the strategic bidding and offering problems only focus on the electricity market in the thesis. Since the rapid growth of DER technologies, the IMES systems, including electricity, natural gas, and heat energies, tend to develop extensively. Therefore, there is a need to analyse the strategic bidding and offering problems in the multi-energy market. Lastly, data-driven approaches can be employed to improve the modelling process, accuracy and performance. For instance, customers' switching behaviours, wholesale electricity prices and demand are all time-series data and can be learned from historical data through machine learning and deep learning methods, such as RNN, GRU, LSTM and Transformer-based methods.

Future research for retail multi-energy pricing problems can align with the direction of developing data-driven models (e.g., machine learning and deep learning) at the lower-level problem to learn the pattern of the interaction between the retail prices set by energy retailers and the microgrids' energy management decisions. In addition, developing a novel approximation and numerical solution method for solving the bilevel model with the existence of binary variables in the lower-level problem can be another future research direction. The solution method can be compared with the proposed hybrid metaheuristic algorithms. Furthermore, since environmental factors, such as carbon cost/budget, are increasingly implemented by different organisations (e.g. microgrids or local energy communities) as required by retailers and governments, it could affect retailers' pricing decisions. Therefore, the objective functions and constraints of the bilevel model will consider such environmental factors in our future work. Lastly, we will investigate our modelling alternatives, such as cooperative game theory and bargaining mechanism, to model the interactions between the retailer and customers/microgrids.

There are some future research directions for long-term multi-energy forecasting problems. For instance, the performance of Patchformer for volatile data, such as financial data needs to be improved to adapt to the non-stationary real-world data. Furthermore, even though the Patchformer generally performs well for multivariate forecasting, the prediction accuracy for univariate forecasting tasks may sometimes be worse than other forecasting models and needs to be enhanced. Moreover, ablation/ sensitivity analysis, such as varying patch length, past sequence length, and

random seed, could be studied to examine the relevant effects on Patchformer forecasting performance. In addition, to further measure the Patchformer performance and allow users to understand the model's behaviour better, interpret and trust the results and output, Explainable AI (XAI) methods, such as SHapley Additive exPlanations (SHAP) [125] and local interpretable model-agnostic explanations (LIME) [126] could be implemented for future study.

BIBLIOGRAPHY

[1] P. Samadi, A.-H. Mohsenian-Rad, R. Schober, V. W. Wong, and J. Jatskevich, "Optimal real-time pricing algorithm based on utility maximization for smart grid," in *2010 First IEEE international conference on smart grid communications*, IEEE, 2010, pp. 415–420.

[2] D. Pozo, E. Sauma, and J. Contreras, "Basic theoretical foundations and insights on bilevel models and their applications to power systems," *Annals of Operations Research*, vol. 254, pp. 303–334, 2017.

[3] A. Mukherjee, "Price and quantity competition under free entry," *Research in Economics*, vol. 59, no. 4, pp. 335–344, 2005.

[4] G. Zhang, G. Zhang, Y. Gao, and J. Lu, "Competitive strategic bidding optimization in electricity markets using bilevel programming and swarm technique," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 6, pp. 2138–2146, 2010.

[5] M. Ghamkhari, A. Sadeghi-Mobarakeh, and H. Mohsenian-Rad, "Strategic bidding for producers in nodal electricity markets: A convex relaxation approach," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2324–2336, 2016.

[6] N. Mahmoudi, T. K. Saha, and M. Eghbal, "Demand response application by strategic wind power producers," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1227–1237, 2015.

[7] M. Di Somma, G. Graditi, and P. Siano, "Optimal bidding strategy for a der aggregator in the day-ahead market in the presence of demand flexibility," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 2, pp. 1509–1519, 2018.

[8] J. Liu, M. Ou, X. Sun, J. Chen, C. Mi, and R. Bo, "Implication of production tax credit on economic dispatch for electricity merchants with storage and wind farms," *Applied Energy*, vol. 308, p. 118 318, 2022.

[9] M. Song and M. Amelin, "Price-maker bidding in day-ahead electricity market for a retailer with flexible demands," *IEEE Transactions on power systems*, vol. 33, no. 2, pp. 1948–1958, 2017.

[10] F.-L. Meng and X.-J. Zeng, "A profit maximization approach to demand response management with customers behavior learning in smart grid," *IEEE Transactions on Smart Grid*, vol. 7, no. 3, pp. 1516–1529, 2015.

[11] P. Liu, T. Ding, Z. Zou, and Y. Yang, "Integrated demand response for a load serving entity in multi-energy market considering network constraints," *Applied Energy*, vol. 250, pp. 512–529, 2019.

[12] P. Mancarella, "Mes (multi-energy systems): An overview of concepts and evaluation models," *Energy*, vol. 65, pp. 1–17, 2014.

[13] F. Meng, Q. Ma, Z. Liu, and X.-J. Zeng, "Multiple dynamic pricing for demand response with adaptive clustering-based customer segmentation in smart grids," *Applied Energy*, vol. 333, p. 120 626, 2023.

[14] M. Carrión, J. M. Arroyo, and A. J. Conejo, "A bilevel stochastic programming approach for retailer futures market trading," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1446–1456, 2009.

[15] F.-L. Meng and X.-J. Zeng, "A stackelberg game-theoretic approach to optimal real-time pricing for the smart grid," *Soft Computing*, vol. 17, no. 12, pp. 2365–2380, 2013.

[16] S. J. Kazempour, A. J. Conejo, and C. Ruiz, "Strategic bidding for a large consumer," *IEEE Transactions on Power Systems*, vol. 30, no. 2, pp. 848–856, 2014.

[17] J. Yang, J. Zhao, F. Wen, and Z. Y. Dong, "A framework of customizing electricity retail prices," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2415–2428, 2017.

[18] M. Yazdani-Damavandi, N. Neyestani, M. Shafie-khah, J. Contreras, and J. P. Catalao, "Strategic behavior of multi-energy players in electricity markets as aggregators of demand side resources using a bi-level approach," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 397–411, 2017.

[19] W. Liu, S. Chen, Y. Hou, and Z. Yang, "Optimal reserve management of electric vehicle aggregator: Discrete bilevel optimization model and exact algorithm," *IEEE Transactions on Smart Grid*, vol. 12, no. 5, pp. 4003–4015, 2021.

[20] M. G. Vayá and G. Andersson, "Optimal bidding strategy of a plug-in electric vehicle aggregator in day-ahead electricity markets under uncertainty," *IEEE transactions on power systems*, vol. 30, no. 5, pp. 2375–2385, 2014.

[21] H. Haghighat, H. Karimianfard, and B. Zeng, "Integrating energy management of autonomous smart grids in electricity market operation," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4044–4055, 2020.

[22] Y. Wu, M. Barati, and G. J. Lim, "A pool strategy of microgrid in power distribution electricity market," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 3–12, 2019.

[23] Q. Huang, Y. Xu, and C. Courcoubetis, "Stackelberg competition between merchant and regulated storage investment in wholesale electricity markets," *Applied Energy*, vol. 264, p. 114 669, 2020.

[24] H. Rashidizadeh-Kermani, M. Vahedipour-Dahraie, M. Shafie-khah, and J. P. Catalão, "A bi-level risk-constrained offering strategy of a wind power producer considering demand side resources," *International Journal of Electrical Power & Energy Systems*, vol. 104, pp. 562–574, 2019.

[25] M. Kohansal and H. Mohsenian-Rad, "Price-maker economic bidding in two-settlement pool-based markets: The case of time-shiftable loads," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 695–705, 2015.

[26] R. Henríquez, G. Wenzel, D. E. Olivares, and M. Negrete-Pincetic, "Participation of demand response aggregators in electricity markets: Optimal portfolio management," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4861–4871, 2017.

[27] H. Algarvio and F. Lopes, "Agent-based retail competition and portfolio optimization in liberalized electricity markets: A study involving real-world consumers," *International Journal of Electrical Power & Energy Systems*, vol. 137, p. 107 687, 2022.

[28] J. Kettunen, A. Salo, and D. W. Bunn, "Optimization of electricity retailer's contract portfolio subject to risk preferences," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 117–128, 2009.

[29] Y. Xiao, X. Wang, P. Pinson, and X. Wang, "Transactive energy based aggregation of prosumers as a retailer," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3302–3312, 2020.

[30] S. Chen, A. J. Conejo, R. Sioshansi, and Z. Wei, "Equilibria in electricity and natural gas markets with strategic offers and bids," *IEEE Transactions on Power Systems*, vol. 35, no. 3, pp. 1956–1966, 2019.

[31] S. Chen, G. Sun, Z. Wei, and D. Wang, "Dynamic pricing in electricity and natural gas distribution networks: An epec model," *Energy*, vol. 207, p. 118 138, 2020.

[32] H. Aghamohammadloo, V. Talaeizadeh, K. Shahanaghi, *et al.*, "Integrated demand response programs and energy hubs retail energy market modelling," *Energy*, vol. 234, p. 121 239, 2021.

[33] S. D. Manshadi and M. E. Khodayar, "A hierarchical electricity market structure for the smart grid paradigm," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 1866–1875, 2015.

[34] X. Kong, D. Liu, C. Wang, F. Sun, and S. Li, "Optimal operation strategy for interconnected microgrids in market environment considering uncertainty," *Applied Energy*, vol. 275, p. 115 336, 2020.

[35] H. Guo, Q. Chen, Q. Xia, and C. Kang, "Electricity wholesale market equilibrium analysis integrating individual risk-averse features of generation companies," *Applied Energy*, vol. 252, p. 113 443, 2019.

[36] L. Park, S. Jeong, J. Kim, and S. Cho, "Joint geometric unsupervised learning and truthful auction for local energy market," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 2, pp. 1499–1508, 2018.

[37] Y. Wang, W. Saad, Z. Han, H. V. Poor, and T. Başar, "A game-theoretic approach to energy trading in the smart grid," *IEEE Transactions on Smart Grid*, vol. 5, no. 3, pp. 1439–1450, 2014.

[38] E. G. Tsimopoulos and M. C. Georgiadis, "Optimal strategic offerings for a conventional producer in jointly cleared energy and balancing markets under high penetration of wind power production," *Applied Energy*, vol. 244, pp. 16–35, 2019.

[39] R. Sharifi, A. Anvari-Moghaddam, S. H. Fathi, and V. Vahidinasab, "A bi-level model for strategic bidding of a price-maker retailer with flexible demands in day-ahead electricity market," *International Journal of Electrical Power & Energy Systems*, vol. 121, p. 106 065, 2020.

[40] F. H. Moghimi and T. Barforoushi, "A short-term decision-making model for a price-maker distribution company in wholesale and retail electricity markets considering demand response and real-time pricing," *International Journal of Electrical Power & Energy Systems*, vol. 117, p. 105 701, 2020.

[41] S. Sekizaki, I. Nishizaki, *et al.*, "Decision making of electricity retailer with multiple channels of purchase based on fractile criterion with rational responses of consumers," *International Journal of Electrical Power & Energy Systems*, vol. 105, pp. 877–893, 2019.

[42] Y. Zhou, W. Yu, S. Zhu, B. Yang, and J. He, "Distributionally robust chance-constrained energy management of an integrated retailer in the multi-energy market," *Applied Energy*, vol. 286, p. 116 516, 2021.

[43] C. Zhao, S. Zhang, X. Wang, X. Li, and L. Wu, "Game analysis of electricity retail market considering customers' switching behaviors and retailers' contract trading," *Ieee Access*, vol. 6, pp. 75 099–75 109, 2018.

[44] J. Yang, J. Zhao, F. Luo, F. Wen, and Z. Y. Dong, "Decision-making for electricity retailers: A brief survey," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4140–4153, 2017.

[45] T. M. Aljohani, A. F. Ebrahim, and O. A. Mohammed, "Dynamic real-time pricing mechanism for electric vehicles charging considering optimal microgrids energy management system," *IEEE Transactions on Industry Applications*, vol. 57, no. 5, pp. 5372–5381, 2021.

[46] S. Nojavan and K. Zare, "Optimal energy pricing for consumers by electricity retailer," *International Journal of Electrical Power & Energy Systems*, vol. 102, pp. 401–412, 2018.

[47] A. Ghasemi, H. J. Monfared, A. Loni, and M. Marzband, "Cvar-based retail electricity pricing in day-ahead scheduling of microgrids," *Energy*, vol. 227, p. 120 529, 2021.

[48] S. Su, Z. Li, X. Jin, K. Yamashita, M. Xia, and Q. Chen, "Bi-level energy management and pricing for community energy retailer incorporating smart buildings based on chance-constrained programming," *International Journal of Electrical Power & Energy Systems*, vol. 138, p. 107 894, 2022.

[49] Q. Hong, F. Meng, J. Liu, and R. Bo, "A bilevel game-theoretic decision-making framework for strategic retailers in both local and wholesale electricity markets," *Applied Energy*, vol. 330, p. 120 311, 2023.

[50] C. Wei, Q. Wu, J. Xu, Y. Wang, and Y. Sun, "Bi-level retail pricing scheme considering price-based demand response of multi-energy buildings," *International Journal of Electrical Power & Energy Systems*, vol. 139, p. 108 007, 2022.

[51] L. Zhang, Y. Gao, H. Zhu, and L. Tao, "Bi-level stochastic real-time pricing model in multi-energy generation system: A reinforcement learning approach," *Energy*, vol. 239, p. 121 926, 2022.

[52] F. Zeng, Z. Bie, S. Liu, C. Yan, and G. Li, "Trading model combining electricity, heating, and cooling under multi-energy demand response," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 1, pp. 133–141, 2019.

[53] H. Wang, C. Wang, W. Sun, and M. Q. Khan, "Energy pricing and management for the integrated energy service provider: A stochastic stackelberg game approach," *Energies*, vol. 15, no. 19, p. 7326, 2022.

[54] X. Zhu, Y. Sun, J. Yang, *et al.*, "Day-ahead energy pricing and management method for regional integrated energy systems considering multi-energy demand responses," *Energy*, vol. 251, p. 123 914, 2022.

[55] A. Sinha, P. Malo, and K. Deb, "A review on bilevel optimization: From classical to evolutionary approaches and applications," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 2, pp. 276–295, 2017.

[56] G. Yuan, Y. Gao, B. Ye, and R. Huang, "Real-time pricing for smart grid with multi-energy microgrids and uncertain loads: A bilevel programming method," *International Journal of Electrical Power & Energy Systems*, vol. 123, p. 106 206, 2020.

[57] M. Sheha, K. Mohammadi, and K. Powell, "Solving the duck curve in a smart grid environment using a non-cooperative game theory and dynamic pricing profiles," *Energy Conversion and Management*, vol. 220, p. 113 102, 2020.

[58] B. Li, R. Roche, D. Paire, and A. Miraoui, "A price decision approach for multiple multi-energy-supply microgrids considering demand response," *Energy*, vol. 167, pp. 117–135, 2019.

[59] G. Yuan, Y. Gao, B. Ye, and Z. Liu, "A bilevel programming approach for real-time pricing strategy of smart grid considering multi-microgrids connection," *International Journal of Energy Research*, vol. 45, no. 7, pp. 10 572–10 589, 2021.

[60] L. P. Qian, Y. J. A. Zhang, J. Huang, and Y. Wu, "Demand response management via real-time electricity price control in smart grids," *IEEE Journal on Selected areas in Communications*, vol. 31, no. 7, pp. 1268–1280, 2013.

[61] F. Meng, B. Kazemtabrizi, X.-J. Zeng, and C. Dent, "An optimal differential pricing in smart grid based on customer segmentation," in *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, IEEE, 2017, pp. 1–6.

[62] J. Yang, J. Zhao, F. Wen, and Z. Dong, "A model of customizing electricity retail prices based on load profile clustering analysis," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3374–3386, 2018.

[63] Y. Dai, X. Sun, Y. Qi, and M. Leng, "A real-time, personalized consumption-based pricing scheme for the consumptions of traditional and renewable energies," *Renewable Energy*, vol. 180, pp. 452–466, 2021.

[64] T. Huang, Y. Sun, M. Jiao, Z. Liu, and J. Hao, "Bilateral energy-trading model with hierarchical personalized pricing in a prosumer community," *International Journal of Electrical Power & Energy Systems*, vol. 141, p. 108 179, 2022.

[65] P. Mancarella, G. Andersson, J. Peças-Lopes, and K. R. Bell, "Modelling of integrated multi-energy systems: Drivers, requirements, and opportunities," in *2016 Power Systems Computation Conference (PSCC)*, IEEE, 2016, pp. 1–22.

[66] D. K. Ranaweera, G. G. Karady, and R. G. Farmer, "Economic impact analysis of load forecasting," *IEEE Transactions on Power Systems*, vol. 12, no. 3, pp. 1388–1392, 1997.

[67] S. Wang, S. Wang, H. Chen, and Q. Gu, "Multi-energy load forecasting for regional integrated energy systems considering temporal dynamic and coupling characteristics," *Energy*, vol. 195, p. 116 964, 2020.

[68] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[69] G. E. Box, G. M. Jenkins, and J. F. MacGregor, "Some recent advances in forecasting and control," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 23, no. 2, pp. 158–179, 1974.

[70] S. S. Rangapuram, M. W. Seeger, J. Gasthaus, L. Stella, Y. Wang, and T. Januschowski, "Deep state space models for time series forecasting," *Advances in neural information processing systems*, vol. 31, 2018.

[71] S. Elsworth and S. Güttel, "Time series forecasting using lstm networks: A symbolic approach," *arXiv preprint arXiv:2003.05672*, 2020.

[72] X. Zhang, F. Shen, J. Zhao, and G. Yang, "Time series forecasting using gru neural network with multi-lag after decomposition," in *Neural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, November 14–18, 2017, Proceedings, Part V 24*, Springer, 2017, pp. 523–532.

[73] H. J. Jo, W. J. Kim, H. K. Goh, and C.-H. Jun, "An improved time-series forecasting model using time series decomposition and gru architecture," in *International Conference on Neural Information Processing*, Springer, 2021, pp. 587–596.

[74] N. Kitaev, Ł. Kaiser, and A. Levskaya, "Reformer: The efficient transformer," *arXiv preprint arXiv:2001.04451*, 2020.

[75] H. Zhou, S. Zhang, J. Peng, *et al.*, "Informer: Beyond efficient transformer for long sequence time-series forecasting," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 35, 2021, pp. 11 106–11 115.

[76] H. Wu, J. Xu, J. Wang, and M. Long, "Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting," *Advances in Neural Information Processing Systems*, vol. 34, pp. 22 419–22 430, 2021.

[77] T. Zhou, Z. Ma, Q. Wen, X. Wang, L. Sun, and R. Jin, "Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting," in *International Conference on Machine Learning*, PMLR, 2022, pp. 27 268–27 286.

[78] Y. Nie, N. H. Nguyen, P. Sinthong, and J. Kalagnanam, "A time series is worth 64 words: Long-term forecasting with transformers," in *The Eleventh International Conference on Learning Representations*, 2023. [Online]. Available: https://openreview.net/forum?id=Jbdc0vTOcol.

[79] A. Vaswani, N. Shazeer, N. Parmar, *et al.*, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[80] A. Dosovitskiy, L. Beyer, A. Kolesnikov, *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," in *International Conference on Learning Representations*, 2021. [Online]. Available: https://openreview.net/forum?id=YicbFdNTTy.

[81] H. Bao, L. Dong, S. Piao, and F. Wei, "BEit: BERT pre-training of image transformers," in *International Conference on Learning Representations*, 2022. [Online]. Available: https://openreview.net/forum?id=p-BhZSz59o4.

[82] C.-M. Lee and C.-N. Ko, "Short-term load forecasting using lifting scheme and arima models," *Expert Systems with Applications*, vol. 38, no. 5, pp. 5902–5911, 2011.

[83] T. Fang and R. Lahdelma, "Evaluation of a multiple linear regression model and sarima model in forecasting heat demand for district heating system," *Applied energy*, vol. 179, pp. 544–552, 2016.

[84] H. Shi, M. Xu, and R. Li, "Deep learning for household load forecasting—a novel pooling deep rnn," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5271–5280, 2017.

[85] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-term residential load forecasting based on lstm recurrent neural network," *IEEE transactions on smart grid*, vol. 10, no. 1, pp. 841–851, 2017.

[86] S. Wang, X. Wang, S. Wang, and D. Wang, "Bi-directional long short-term memory method based on attention mechanism and rolling update for short-term load forecasting," *International Journal of Electrical Power & Energy Systems*, vol. 109, pp. 470–479, 2019.

[87] G. Zhang, X. Bai, and Y. Wang, "Short-time multi-energy load forecasting method based on cnn-seq2seq model with attention mechanism," *Machine Learning with Applications*, vol. 5, p. 100 064, 2021.

[88] G. Zhang, S. Zhu, and X. Bai, "Federated learning-based multi-energy load forecasting method using cnn-attention-lstm model," *Sustainability*, vol. 14, no. 19, p. 12 843, 2022.

[89] Y. Zhang and J. Yan, "Crossformer: Transformer utilizing cross-dimension dependency for multivariate time series forecasting," in *The Eleventh International Conference on Learning Representations*, 2022.

[90] P.-H. Kuo and C.-J. Huang, "A high precision artificial neural networks model for short-term energy load forecasting," *Energies*, vol. 11, no. 1, p. 213, 2018.

[91] S. Cen and C. G. Lim, "Multi-task learning of the patchtcn-tst model for short-term multi-load energy forecasting considering indoor environments in a smart building," *IEEE Access*, 2024.

[92] N. A. Mohammed and A. Al-Bazi, "An adaptive backpropagation algorithm for long-term electricity load forecasting," *Neural Computing and Applications*, vol. 34, no. 1, pp. 477–491, 2022.

[93] S. R. Khuntia, J. L. Rueda, and M. A. Van der Meijden, "Long-term electricity load forecasting considering volatility using multiplicative error model," *energies*, vol. 11, no. 12, p. 3308, 2018.

[94] K. Lindberg, P. Seljom, H. Madsen, D. Fischer, and M. Korpås, "Long-term electricity load forecasting: Current and future trends," *Utilities Policy*, vol. 58, pp. 102–119, 2019.

[95] C. Wang, Y. Wang, Z. Ding, T. Zheng, J. Hu, and K. Zhang, "A transformer-based method of multienergy load forecasting in integrated energy system," *IEEE Transactions on Smart Grid*, vol. 13, no. 4, pp. 2703–2714, 2022.

[96] C. Wang, Y. Wang, Z. Ding, and K. Zhang, "Probabilistic multi-energy load forecasting for integrated energy system based on bayesian transformer network," *IEEE Transactions on Smart Grid*, 2023.

[97] Y. Zhang, F. Meng, R. Wang, B. Kazemtabrizi, and J. Shi, "Uncertainty-resistant stochastic mpc approach for optimal operation of chp microgrid," *Energy*, vol. 179, pp. 1265–1278, 2019.

[98] Y. Zhang, F. Meng, R. Wang, W. Zhu, and X.-J. Zeng, "A stochastic mpc based approach to integrated energy management in microgrids," *Sustainable cities and society*, vol. 41, pp. 349–362, 2018.

[99] X. Xiao, J. Wang, R. Lin, D. J. Hill, and C. Kang, "Large-scale aggregation of prosumers toward strategic bidding in joint energy and regulation markets," *Applied Energy*, vol. 271, p. 115 159, 2020.

[100] N. Singh and X. Vives, "Price and quantity competition in a differentiated duopoly," *The Rand journal of economics*, pp. 546–554, 1984.

[101] P. Van Aubel and E. Poll, "Smart metering in the netherlands: What, how, and why," *International Journal of Electrical Power & Energy Systems*, vol. 109, pp. 719–725, 2019.

[102] J. Hardy, "How could we buy energy in the smart future," *Imperial College London*, 2017.

[103] Y. Zhang and G. B. Giannakis, "Distributed stochastic market clearing with high-penetration wind power," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 895–906, 2015.

[104] C. Ruiz and A. J. Conejo, "Pool strategy of a producer with endogenous formation of locational marginal prices," *IEEE Transactions on Power Systems*, vol. 24, no. 4, pp. 1855–1866, 2009.

[105] J. Fortuny-Amat and B. McCarl, "A representation and economic interpretation of a two-level programming problem," *Journal of the operational Research Society*, vol. 32, pp. 783–792, 1981.

[106] S. A. Gabriel, A. J. Conejo, J. D. Fuller, B. F. Hobbs, and C. Ruiz, *Complementarity modeling in energy markets*. Springer Science & Business Media, 2012, vol. 180.

[107] W. E. Hart, C. D. Laird, J.-P. Watson, *et al.*, *Pyomo-optimization modeling in python*. Springer, 2017, vol. 67.

[108] *Pjm data directory*, `https://dataminer2.pjm.com/list`, Accessed: 2022.

[109] *Hourly day-ahead demand bids*, https://dataminer2.pjm.com/feed/hrl_da_demand_bids/definition, Accessed: 2022-10-25.

[110] S. Dempe and A. Zemkoho, "Bilevel optimization," in *Springer optimization and its applications. Vol. 161*, Springer, 2020.

[111] Y. Zhang, F. Meng, and R. Wang, "A comprehensive mpc based energy management framework for isolated microgrids," in *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, IEEE, 2017, pp. 1–6.

[112] M. Zugno, J. M. Morales, P. Pinson, and H. Madsen, "A bilevel model for electricity retailers' participation in a demand response market environment," *Energy Economics*, vol. 36, pp. 182–197, 2013.

[113] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95-international conference on neural networks*, IEEE, vol. 4, 1995, pp. 1942–1948.

[114] K. E. Parsopoulos, M. N. Vrahatis, *et al.*, "Particle swarm optimization method for constrained optimization problems," *Intelligent technologies–theory and application: New trends in intelligent technologies*, vol. 76, no. 1, pp. 214–220, 2002.

[115] M. S. Innocente and J. Sienz, "Constraint-handling techniques for particle swarm optimization algorithms," *arXiv preprint arXiv:2101.10933*, 2021.

[116] D. Whitley, "A genetic algorithm tutorial," *Statistics and computing*, vol. 4, no. 2, pp. 65–85, 1994.

[117] P. J. Van Laarhoven and E. H. Aarts, "Simulated annealing," in *Simulated annealing: Theory and applications*, Springer, 1987, pp. 7–15.

[118] *Department for education gas and electricity half hourly data*, https://www.data.gov.uk/dataset/fee711fd-b405-4939-8945-5f9189839ad0/department-for-education-gas-and-electricity-half-hourly-data, Accessed: 2022-11-10.

[119] *Open power system data: When2heat heating profiles*, https://data.open-power-system-data.org/when2heat/2019-08-06, Accessed: 2022-11-10.

[120] *Photovoltaic geographical information system*, https://re.jrc.ec.europa.eu/pvg_tools/en/, Accessed: 2022-11-12.

[121] *Smart meters in london*, https://www.kaggle.com/datasets/jeanmidev/smart-meters-in-london?select=weather_hourly_darksky.csv, Accessed: 2022-11-12.

[122] S. Qi, X. Wang, X. Li, T. Qian, and Q. Zhang, "Enhancing integrated energy distribution system resilience through a hierarchical management strategy in district multi-energy systems," *Sustainability*, vol. 11, no. 15, p. 4048, 2019.

[123] *Campus metabolism*, `https://cm.asu.edu/`, Accessed: 2024-03-19.

[124] A. Zeng, M. Chen, L. Zhang, and Q. Xu, "Are transformers effective for time series forecasting?" In *Proceedings of the AAAI conference on artificial intelligence*, vol. 37, 2023, pp. 11 121–11 128.

[125] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017.

[126] M. T. Ribeiro, S. Singh, and C. Guestrin, ""why should i trust you?" explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135–1144.

Appendix A

APPENDIX FOR CHAPTER 3

A.1 Input data for Chapter 3

A.1.1 Data in case 1

Table A.1: Initial retail prices of retailers in case 1 ($/MWh).

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	37.88	34.55	35.72	32.23	34.07	37.75	37.52	41.66	49.56	52.75	59.06	68.86	74.26	79.44	85.44	100.36	94.39	81.25	67.66	64.47	54.95	48.48	41.64	40.20
2	39.28	36.18	33.82	32.62	34.60	36.83	40.55	43.33	49.94	49.03	47.61	51.33	52.44	54.83	53.01	66.12	59.83	52.98	54.56	52.73	49.23	47.31	43.99	42.34
3	37.95	35.44	35.36	33.41	33.46	36.28	38.77	44.49	50.54	57.28	64.90	74.22	77.24	89.10	93.23	105.97	100.47	81.00	71.94	65.26	53.89	50.07	41.15	43.06
4	40.34	35.27	32.80	31.79	35.71	36.99	39.38	43.80	52.16	49.82	47.19	50.28	54.84	58.89	58.84	80.03	79.59	69.16	59.39	49.13	47.31	44.95	42.45	42.91
5	37.64	34.81	34.12	33.35	35.22	36.88	37.71	44.26	49.91	49.91	46.58	53.00	50.81	53.83	53.31	62.38	59.93	52.60	50.74	49.08	47.92	47.43	40.79	41.54
6	38.81	35.28	32.68	30.90	32.10	35.41	38.36	40.51	46.65	49.63	58.64	66.94	74.77	83.71	95.71	107.56	95.76	84.04	73.90	53.36	48.54	45.91	42.73	39.78
7	36.06	35.83	33.23	32.51	33.25	35.09	40.49	43.65	47.98	48.75	46.35	49.65	52.61	51.43	49.51	60.94	56.61	51.34	51.50	49.07	45.58	44.98	41.49	41.57
8	37.66	34.18	34.26	32.33	36.04	37.19	39.43	42.90	49.94	51.49	63.97	70.27	75.38	84.64	93.90	103.83	93.60	76.59	69.49	70.11	59.02	49.98	43.10	42.34
9	37.89	34.66	31.78	32.94	33.58	35.56	39.29	43.00	48.29	46.30	50.58	54.73	60.84	65.78	71.82	80.41	72.09	66.97	59.88	48.60	45.58	44.35	40.55	39.56
10	36.31	36.21	32.53	31.39	33.89	35.50	38.82	44.11	48.78	49.94	53.22	59.43	63.45	66.53	66.23	67.76	64.79	64.10	61.80	56.28	48.58	46.77	40.15	39.28
11	37.69	35.34	33.70	32.10	34.12	36.12	37.70	43.40	48.85	49.48	47.45	51.62	52.89	54.02	54.43	63.38	58.91	54.34	52.27	50.38	48.35	44.98	43.00	40.81
12	37.46	35.78	34.30	34.01	34.34	36.96	40.19	44.15	48.42	49.12	50.82	51.85	55.66	55.90	58.14	65.42	62.37	54.03	55.40	53.17	46.99	46.84	41.50	41.42

Table A.2: Initial DAW market bid prices of retailers in case 1 ($/MWh).

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	28.80	26.83	25.61	24.99	26.12	27.82	29.56	32.50	37.98	40.83	46.35	53.13	57.40	60.64	67.00	77.74	71.72	61.07	53.21	49.10	43.02	37.95	32.70	31.25
2	29.45	27.28	25.97	25.37	26.61	28.41	30.58	33.82	38.42	37.79	36.66	39.29	40.19	41.15	40.84	49.66	46.26	41.12	41.89	40.29	37.66	35.40	32.96	32.19
3	29.45	27.34	26.11	25.53	26.75	28.44	30.27	33.34	38.96	42.32	49.42	57.00	59.75	67.79	71.90	82.15	76.85	62.30	54.42	50.26	43.65	37.91	33.10	31.94
4	29.63	27.46	26.13	25.53	26.80	28.61	30.75	33.88	38.52	38.06	35.73	38.08	41.40	44.48	45.81	61.73	61.81	53.75	44.81	38.18	36.45	35.11	33.11	32.40
5	29.30	27.15	25.87	25.27	26.48	28.24	30.36	33.56	38.01	37.61	36.48	38.83	40.03	40.85	40.97	48.36	45.47	40.68	40.83	38.65	36.51	34.76	32.52	31.97
6	28.74	26.67	25.43	24.86	26.06	27.81	29.79	32.89	37.19	37.94	45.42	52.07	57.83	65.02	72.91	81.48	72.07	64.10	56.27	39.92	36.00	34.19	31.94	31.09
7	28.97	26.89	25.63	25.05	26.26	27.99	30.07	33.22	37.73	37.20	35.42	37.79	38.75	39.59	39.07	46.85	44.04	39.07	39.45	37.48	35.50	34.28	32.31	31.58
8	29.20	27.15	25.97	25.41	26.61	28.27	29.97	32.92	38.46	41.44	48.27	54.68	57.25	65.60	72.34	80.04	72.80	59.05	53.09	54.25	46.37	38.89	32.96	31.69
9	28.50	26.44	25.21	24.64	25.81	27.54	29.52	32.60	36.76	36.50	39.38	43.41	46.65	50.58	54.53	61.69	55.74	49.89	46.85	38.41	35.53	33.71	31.46	30.77
10	28.79	26.72	25.50	24.98	26.16	27.93	30.04	32.90	37.87	38.45	41.22	45.86	48.18	50.77	51.21	52.84	49.94	48.67	47.55	43.39	38.76	35.21	32.03	31.20
11	29.28	27.16	25.89	25.28	26.49	28.25	30.39	33.58	38.11	37.74	36.61	38.97	40.15	40.86	40.94	48.22	45.44	40.76	40.92	38.78	36.53	34.82	32.53	32.01
12	29.23	27.10	25.84	25.24	26.47	28.20	30.43	33.60	38.25	38.22	37.98	40.84	42.33	43.33	43.95	50.92	47.94	43.12	42.65	40.07	37.30	35.12	32.47	31.90

Table A.3: Initial LPE market bid/offer prices of retailers in case 1 ($/MWh).

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	30.63	28.06	26.38	25.95	27.23	28.75	29.94	33.99	34.99	35.55	39.75	40.37	47.26	47.82	50.22	117.52	70.51	56.30	44.02	40.82	46.75	43.10	32.50	31.92
2	31.24	28.55	26.84	26.43	27.74	29.17	30.36	34.31	34.80	33.85	39.53	36.07	41.90	37.68	36.92	47.14	41.88	36.28	38.11	38.17	41.81	41.49	32.16	31.94
3	31.19	28.53	26.81	26.39	27.73	29.26	30.50	34.57	35.51	36.22	40.21	41.62	47.42	47.69	43.67	104.86	59.44	45.74	42.10	41.54	48.01	43.81	32.44	32.21
4	31.52	28.77	27.03	26.61	27.96	29.47	30.62	34.48	34.98	34.06	39.87	36.27	42.39	37.77	37.31	46.94	42.61	271.33	38.61	38.57	42.11	42.03	32.46	32.24
5	30.97	28.35	26.69	26.33	27.64	29.03	30.18	34.03	34.40	33.65	39.11	35.91	41.39	37.81	36.48	47.78	41.08	37.77	37.61	37.87	41.63	40.99	31.78	31.69
6	30.26	27.71	26.09	25.72	27.01	28.42	29.57	33.44	33.98	33.28	38.57	35.56	40.99	37.43	35.68	46.85	39.83	38.59	36.73	37.16	40.87	39.92	31.14	31.01
7	30.67	28.07	26.42	26.02	27.30	28.71	29.86	33.72	34.18	33.28	38.77	35.32	41.02	36.88	35.92	44.67	40.18	41.09	37.08	37.27	40.84	40.71	31.60	31.43
8	31.01	28.38	26.67	26.26	27.60	29.14	30.35	34.38	35.41	36.75	39.97	43.51	49.21	51.43	46.08	129.88	66.77	46.94	43.35	42.59	50.28	43.68	32.27	32.03
9	30.06	27.57	25.95	25.58	26.84	28.23	29.35	33.12	33.63	32.96	38.17	35.31	40.45	37.22	35.12	46.26	39.02	37.68	36.33	36.88	40.63	39.61	30.88	30.83
10	30.24	27.74	26.13	25.79	27.06	28.33	29.63	33.57	34.21	34.17	38.83	37.92	39.01	41.69	38.18	71.84	47.25	40.06	38.48	38.54	43.64	41.24	31.53	31.42
11	31.00	28.39	26.75	26.37	27.68	29.08	30.26	34.09	34.44	33.69	39.11	35.98	41.39	37.88	36.54	48.37	41.21	37.94	37.57	37.83	41.66	41.07	31.87	31.78
12	31.00	28.40	26.76	26.38	27.70	29.09	30.33	34.22	34.64	34.05	39.35	36.61	41.75	39.01	37.50	54.56	43.35	39.15	38.13	38.21	42.36	41.37	31.98	31.87

Table A.4: Maximum LPE market bid/offer electricity volume of retailers in case 1 (MWh).

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	250	240	231	232	232	234	240	250	266	282	296	310	322	329	331	333	331	324	315	306	299	286	267	251
2	252	243	236	236	237	240	246	262	284	304	318	329	342	353	361	367	369	362	353	337	327	306	282	264
3	253	243	238	238	243	242	249	263	284	305	323	341	350	357	362	368	370	366	354	339	327	307	282	266
4	253	246	241	241	244	253	265	277	290	306	323	342	357	368	372	377	376	367	354	339	330	308	285	266
5	261	250	245	244	248	260	271	284	294	313	325	344	360	368	375	382	379	369	354	339	331	311	286	266
6	269	257	249	245	253	262	272	285	302	319	334	349	360	374	381	383	382	376	362	349	340	321	301	282
7	269	258	251	251	260	270	283	294	305	322	342	361	373	377	383	391	395	391	379	363	355	329	305	284
8	271	259	254	254	260	272	284	300	319	335	349	362	376	385	391	398	397	392	381	368	357	335	310	290
9	277	266	260	260	264	273	287	304	321	343	360	380	399	408	404	406	402	403	391	378	362	338	311	290
10	277	269	261	261	268	280	291	306	325	344	365	385	400	409	411	408	407	405	396	378	366	341	320	294
11	283	272	266	265	274	281	293	313	327	347	371	391	405	412	413	409	409	411	403	385	375	348	320	298
12	288	277	269	268	277	289	301	313	335	361	381	400	407	413	418	422	419	418	407	391	380	356	329	306

Table A.5: Maximum DAW market bid load of retailers in case 1 (MWh).

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	5184	4928	4739	4646	4645	4678	4866	5111	5338	5521	5667	5845	6056	6299	6602	6877	7039	6995	6744	6559	6352	5994	5607	5236
2	5191	4983	4827	4758	4757	4805	5082	5504	5956	6370	6760	7145	7440	7653	7822	7907	7923	7743	7482	7288	7070	6542	5982	5529
3	5220	4985	4871	4871	5037	5186	5371	5703	6111	6550	6882	7177	7485	7743	7912	8056	8099	7957	7639	7348	7076	6558	6029	5556
4	5443	5200	5047	5043	5161	5294	5651	5989	6267	6573	7007	7292	7542	7760	7996	8195	8290	8123	7767	7385	7107	6661	6090	5572
5	5611	5397	5239	5149	5187	5466	5852	6175	6441	6740	7062	7512	7936	8270	8422	8423	8333	8161	7968	7591	7203	6751	6327	5813
6	5871	5653	5517	5524	5692	5994	6452	6853	7180	7504	7837	8108	8320	8414	8563	8661	8646	8498	8048	7802	7500	6927	6334	5921
7	5897	5657	5525	5554	5753	6035	6473	6907	7326	7660	7958	8187	8409	8545	8656	8750	8745	8518	8122	7815	7576	7126	6630	6194
8	5927	5685	5537	5569	5765	6077	6492	6934	7339	7834	8204	8536	8846	8967	8810	8851	8999	8983	8772	8419	8048	7419	6796	6260
9	6109	5826	5652	5623	5785	6086	6569	7002	7434	7863	8378	8739	8889	9043	9210	9327	9362	9168	8776	8539	8270	7759	7096	6517
10	6178	5895	5709	5683	5848	6112	6570	7120	7569	8010	8423	8847	9225	9510	9700	9855	9889	9681	9320	8927	8476	7780	7184	6603
11	6310	6037	5862	5838	5983	6287	6710	7154	7659	8263	8799	9203	9521	9710	9863	10036	10077	9882	9398	8986	8572	7871	7203	6700
12	6335	6072	5903	5901	6086	6385	6865	7358	7854	8357	8835	9290	9654	9879	10087	10230	10294	10111	9611	9104	8602	7920	7264	6708

Table A.6: Alpha values of retailers in case 1.

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	144	133	124	119	115	112	117	123	132	143	153	163	171	178	183	187	189	186	183	177	168	163	153	142
2	148	134	126	119	117	117	120	127	134	144	155	165	175	182	186	189	189	188	184	183	173	165	154	146
3	147	138	128	123	118	118	122	130	138	149	158	168	176	183	188	191	194	192	187	184	176	168	159	149
4	153	140	133	125	121	121	125	132	140	152	162	169	181	189	193	195	195	196	192	188	180	172	162	152
5	175	167	156	150	147	144	148	157	164	177	185	194	205	212	218	220	219	221	215	211	204	194	186	176
6	179	167	160	154	148	149	152	161	169	179	188	199	206	214	218	223	224	222	219	213	205	198	189	179
7	184	173	162	157	151	152	156	164	171	181	192	201	211	216	221	225	226	228	223	217	209	202	191	180
8	186	175	167	160	157	155	158	167	174	185	195	205	213	219	225	227	231	228	225	221	213	206	194	186
9	211	199	192	185	182	182	187	193	200	211	222	231	240	244	251	254	256	255	252	246	240	230	220	210
10	215	205	198	189	184	185	188	196	203	215	225	234	242	250	254	258	259	259	254	251	244	232	225	215
11	218	208	199	192	189	187	194	200	207	217	229	237	247	253	260	263	263	261	257	254	245	238	228	218
12	220	212	204	196	192	193	197	200	214	223	232	241	251	257	261	268	267	267	262	259	251	242	231	224

Table A.7: Self-elasticity values of retailers in case 1.

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	130	122	117	113	113	110	116	117	123	128	133	133	138	140	144	143	144	143	143	141	137	136	133	129
2	127	120	116	111	110	111	111	115	120	125	129	133	135	139	142	142	142	141	140	140	136	132	130	128
3	124	119	114	110	107	108	110	114	119	125	128	132	133	138	139	139	140	139	139	138	136	129	126	123
4	122	117	112	106	106	107	108	112	116	122	126	129	132	134	136	138	138	139	137	134	131	128	126	122
5	114	109	103	101	98	99	101	104	107	114	118	120	124	127	127	130	131	129	129	125	124	119	119	114
6	112	106	103	100	94	97	100	102	105	113	115	118	121	125	127	127	129	126	127	125	122	119	116	113
7	111	105	100	97	94	95	96	97	106	112	114	117	120	122	124	124	125	125	125	122	120	117	115	110
8	110	101	97	95	95	92	94	97	104	108	112	115	117	121	123	125	124	123	123	121	118	115	113	108
9	103	96	91	85	86	85	88	91	95	102	104	110	109	111	113	115	116	118	115	112	112	107	103	102
10	99	93	90	85	85	84	87	89	92	101	102	106	108	111	113	115	115	116	113	110	108	105	101	100
11	96	91	87	82	80	82	84	86	93	96	100	104	104	111	112	111	115	113	112	109	107	102	101	97
12	95	91	85	81	80	79	82	86	90	96	98	101	104	109	109	112	110	110	110	108	105	102	99	95

A.1.2 Information of generators in DAW market

Table A.8: Information of generators in DAW market.

Information	Generator														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Cost ($/MWh)	10	12	15	17	20	23	25	27	30	34	36	38	40	45	46
Maximum supply (MWh)	5000	4350	3940	3460	5070	2810	5300	4250	4650	3910	3250	3500	4750	3000	5750

Information	Generator														
	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Cost ($/MWh)	48	51	53	56	60	65	68	70	74	76	78	80	84	88	90
Maximum supply (MWh)	2250	3460	3940	2290	1990	2600	3800	3000	2500	2000	1050	3860	4800	3900	3000

A.1.3 Data in case 2

Table A.9: Initial retail prices of retailers in case 2 ($/MWh).

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	38.77	38.38	33.33	34.95	33.91	32.60	32.75	32.67	33.96	33.96	36.60	37.59	37.46	40.12	42.48	43.63	48.24	49.18	54.56	47.28	58.28	44.30	69.34	51.51
2	39.41	39.63	35.56	36.69	33.49	34.37	32.71	32.48	34.73	34.12	38.24	36.95	40.15	38.62	45.22	42.06	50.61	49.65	48.19	53.38	46.85	62.54	50.74	71.40
3	38.85	36.90	35.52	34.66	33.40	33.48	34.03	32.24	37.17	34.47	36.86	35.44	39.12	37.87	44.47	41.68	50.57	47.26	56.06	46.03	63.88	51.07	74.42	56.89
4	38.30	38.42	36.01	37.20	32.46	33.77	32.14	31.68	36.83	35.75	35.34	36.09	39.91	36.92	44.00	43.16	50.29	47.24	50.59	50.23	46.39	52.83	50.22	60.54
5	37.61	38.15	36.12	33.35	33.75	31.75	31.93	33.33	34.28	34.04	36.42	37.57	40.17	38.72	43.92	43.18	51.82	48.80	48.58	48.71	46.98	47.73	50.75	49.04
6	36.79	36.68	35.91	37.70	34.44	33.94	33.34	32.98	35.53	35.08	35.92	35.53	38.09	39.87	41.53	43.46	49.69	49.75	51.31	51.13	58.76	48.01	67.64	52.61

Table A.10: Initial DAW market bid prices of retailers in case 2 ($/MWh).

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	28.80	28.97	26.83	26.89	25.61	25.63	24.99	25.05	26.12	26.26	27.82	27.99	29.56	30.07	32.50	33.22	37.98	37.73	40.83	37.20	46.35	35.42	53.13	37.79
2	29.45	29.20	27.28	27.15	25.97	25.97	25.37	25.41	26.61	26.61	28.41	28.27	30.58	29.97	33.82	32.92	38.42	38.46	37.79	41.44	36.66	48.27	39.29	54.68
3	29.45	28.50	27.34	26.44	26.11	25.21	25.53	24.64	26.75	25.81	28.44	27.54	30.27	29.52	33.34	32.60	38.96	36.76	42.32	36.50	49.42	39.38	57.00	43.41
4	29.63	28.79	27.46	26.72	26.13	25.50	25.53	24.98	26.80	26.16	28.61	27.93	30.75	30.04	33.88	32.90	38.52	37.87	38.06	38.45	35.73	41.22	38.08	45.86
5	29.30	29.28	27.15	27.16	25.87	25.89	25.27	25.28	26.48	26.49	28.24	28.25	30.36	30.39	33.56	33.58	38.01	38.11	37.61	37.74	36.48	36.61	38.83	38.97
6	28.74	29.23	26.67	27.10	25.43	25.84	24.86	25.24	26.06	26.47	27.81	28.20	29.79	30.43	32.89	33.60	37.19	38.25	37.94	38.22	45.42	37.98	52.07	40.84

Table A.11: Initial LPE market bid/offer prices of retailers in case 2 ($/MWh).

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	30.63	30.67	28.06	28.07	26.38	26.42	25.95	26.02	27.23	27.30	28.75	28.71	29.94	29.86	33.99	33.72	34.99	34.18	35.55	33.28	39.75	38.77	40.37	35.32
2	31.24	31.01	28.55	28.38	26.84	26.67	26.43	26.26	27.74	27.60	29.17	29.14	30.36	30.35	34.31	34.38	34.80	35.41	33.85	36.75	39.53	39.97	36.07	43.51
3	31.19	30.06	28.53	27.57	26.81	25.95	26.39	25.58	27.73	26.84	29.26	28.23	30.50	29.35	34.57	33.12	35.51	33.63	36.22	32.96	40.21	38.17	41.62	35.31
4	31.52	30.24	28.77	27.74	27.03	26.13	26.61	25.79	27.96	27.06	29.47	28.33	30.62	29.63	34.48	33.57	34.98	34.21	34.06	34.17	39.87	38.83	36.27	37.92
5	30.97	31.00	28.35	28.39	26.69	26.75	26.33	26.37	27.64	27.68	29.03	29.08	30.18	30.26	34.03	34.09	34.40	34.44	33.65	33.69	39.11	39.11	35.91	35.98
6	30.26	31.00	27.71	28.40	26.09	26.76	25.72	26.38	27.01	27.70	28.42	29.09	29.57	30.33	33.44	34.22	33.98	34.64	33.28	34.05	38.57	39.35	35.56	36.61

Table A.12: Maximum LPE market bid/offer electricity volume of retailers in case 2 (MWh).

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	1002	962	924	929	929	937	959	1000	1062	1128	1183	1241	1286	1316	1325	1332	1325	1296	1260	1226	1196	1145	1068	1005
2	1007	970	945	944	949	959	982	1049	1136	1214	1274	1314	1368	1412	1443	1468	1476	1450	1414	1347	1307	1226	1128	1057
3	1010	973	951	953	971	967	994	1051	1136	1218	1291	1365	1399	1430	1449	1471	1478	1463	1416	1357	1308	1227	1128	1063
4	1013	983	962	963	975	1012	1059	1109	1161	1222	1292	1367	1428	1470	1488	1507	1504	1467	1416	1357	1319	1232	1139	1064
5	1046	1001	978	978	992	1040	1084	1138	1178	1252	1300	1376	1439	1472	1502	1529	1517	1474	1418	1358	1322	1246	1145	1065
6	1074	1028	996	981	1012	1049	1087	1138	1207	1277	1336	1396	1441	1496	1523	1530	1529	1506	1449	1395	1361	1284	1204	1129

Table A.13: Maximum DAW market bid load of retailers in case 2 (MWh).

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	20735	19713	18956	18586	18582	18714	19464	20445	21350	22084	22668	23378	24225	25197	26409	27508	28154	27978	26977	26237	25408	23975	22428	20945
2	20764	19930	19306	19030	19027	19221	20328	22015	23826	25478	27038	28578	29760	30612	31288	31627	31690	30970	29928	29154	28280	26167	23929	22117
3	20881	19941	19483	19485	20149	20744	21486	22812	24443	26201	27528	28706	29939	30972	31646	32224	32397	31826	30555	29390	28304	26232	24117	22224
4	21773	20799	20186	20174	20645	21178	22604	23957	25069	26294	28027	29167	30166	31041	31986	32779	33161	32492	31067	29541	28428	26642	24360	22290
5	22442	21586	20957	20597	20750	21864	23409	24698	25765	26959	28247	30049	31744	33082	33690	33693	33333	32643	31872	30364	28811	27002	25310	23251
6	23484	22613	22068	22098	22769	23978	25807	27412	28722	30014	31350	32432	33278	33655	34252	34642	34585	33991	32191	31208	30000	27710	25336	23684

Table A.14: Alpha values of retailers in case 2.

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	242	230	222	216	212	210	215	221	230	240	250	261	269	275	281	284	287	284	281	274	266	260	251	239
2	252	238	231	224	222	222	224	231	239	248	260	270	279	286	290	294	294	292	289	287	278	269	259	251
3	293	284	274	269	264	264	269	276	284	295	304	314	322	329	334	337	340	338	333	330	322	314	305	295
4	307	294	287	279	275	276	279	287	294	307	316	324	335	343	347	349	350	351	346	342	334	326	316	306
5	354	346	335	329	327	323	327	336	343	356	364	373	384	391	397	399	398	400	395	390	383	374	365	355
6	367	355	348	342	337	337	340	349	357	367	376	387	394	402	407	411	412	410	408	401	393	387	377	367

Table A.15: Self-elasticity values of retailers in case 2.

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	128	120	115	111	111	108	114	115	121	126	131	131	136	138	142	140	142	141	140	139	135	134	130	127
2	123	116	111	107	106	107	107	111	116	121	125	129	131	135	138	138	138	137	135	136	132	128	126	124
3	112	107	102	98	95	96	98	102	107	113	116	120	121	126	127	127	128	128	128	126	124	118	114	111
4	108	103	98	92	92	93	94	98	102	108	112	115	118	121	122	124	124	125	124	121	118	114	112	108
5	98	93	88	86	83	84	85	89	92	99	103	105	109	111	112	115	116	114	114	110	109	104	103	99
6	95	89	86	83	77	80	83	85	88	96	98	101	104	108	110	110	112	109	110	108	105	102	99	96

A.1.4 Data in case 3

Table A.16: Initial retail prices of retailers in case 3 ($/MWh).

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	37.88	38.19	40.09	37.17	34.99	37.28	34.05	34.14	33.48	33.64	32.13	32.95	32.13	33.79	30.90	31.77	35.11	36.70	32.63	34.65	35.19	36.34	35.52	35.89
2	39.28	38.80	38.02	37.70	35.47	35.19	36.09	34.68	33.75	33.53	33.71	33.91	33.18	34.19	32.95	34.43	34.29	33.95	34.69	34.79	37.20	38.00	37.89	37.22
3	37.95	37.26	38.46	38.22	34.23	33.98	33.79	36.38	33.84	35.32	33.43	33.72	32.75	33.29	31.79	31.99	35.34	33.89	34.74	34.34	34.11	36.94	33.92	38.20

Table A.17: Initial DAW market bid prices of retailers in case 3 ($/MWh).

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	28.80	29.63	28.97	28.79	26.83	27.46	26.89	26.72	25.61	26.13	25.63	25.50	24.99	25.53	25.05	24.98	26.12	26.80	26.26	26.16	27.82	28.61	27.99	27.93
2	29.45	29.30	29.20	29.28	27.28	27.15	27.15	27.16	25.97	25.87	25.97	25.89	25.37	25.27	25.41	25.28	26.61	26.48	26.61	26.49	28.41	28.24	28.27	28.25
3	29.45	28.74	28.50	29.23	27.34	26.67	26.44	27.10	26.11	25.43	25.21	25.84	25.53	24.86	24.64	25.24	26.75	26.06	25.81	26.47	28.44	27.81	27.54	28.20

Table A.18: Initial LPE market bid/offer prices of retailers in case 3 ($/MWh).

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	30.63	31.52	30.67	30.24	28.06	28.77	28.07	27.74	26.38	27.03	26.42	26.13	25.95	26.61	26.02	25.79	27.23	27.96	27.30	27.06	28.75	29.47	28.71	28.33
2	31.24	30.97	31.01	31.00	28.55	28.35	28.38	28.39	26.84	26.69	26.67	26.75	26.43	26.33	26.26	26.37	27.74	27.64	27.60	27.68	29.17	29.03	29.14	29.08
3	31.19	30.26	30.06	31.00	28.53	27.71	27.57	28.40	26.81	26.09	25.95	26.76	26.39	25.72	25.58	26.38	27.73	27.01	26.84	27.70	29.26	28.42	28.23	29.09

Table A.19: Maximum LPE market bid/offer electricity volume of retailers in case 3 (MWh).

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	6260	6012	5777	5806	5806	5855	5995	6248	6639	7052	7392	7758	8038	8222	8282	8324	8282	8097	7876	7661	7472	7156	6674	6284
2	6296	6064	5905	5902	5928	5994	6140	6558	7098	7590	7960	8214	8547	8827	9018	9173	9228	9060	8837	8418	8170	7662	7050	6606
3	6314	6080	5946	5959	6072	6042	6215	6568	7102	7614	8070	8534	8744	8937	9056	9193	9239	9141	8848	8480	8174	7668	7052	6642

Table A.20: Maximum DAW market bid load of retailers in case 3 (MWh).

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	129595	123205	118475	116160	116135	116960	121650	127780	133440	138025	141675	146115	151405	157480	165055	171925	175965	174865	168605	163980	158800	149845	140175	130905
2	129775	124565	120665	118940	118920	120130	127050	137595	148910	159240	168990	178615	186000	191325	195550	197670	198065	193565	187050	182210	176750	163545	149555	138230
3	130505	124630	121770	121780	125930	129650	134285	142575	152770	163755	172050	179415	187120	193575	197790	201400	202480	198915	190970	183690	176900	163950	150730	138900

Table A.21: Alpha values of retailers in case 3.

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	411	399	391	385	381	379	384	390	399	409	419	430	438	444	450	453	456	453	450	443	435	429	420	408
2	503	489	481	474	472	472	474	481	489	498	510	520	530	537	540	544	544	543	539	537	528	519	509	501
3	599	590	580	575	570	569	574	581	590	601	610	619	628	635	640	642	645	644	639	636	627	620	611	600

Table A.22: Self-elasticity values of retailers in case 3.

Retailer	Time																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	123	116	111	107	107	104	110	111	117	122	127	127	131	134	137	136	138	137	136	135	131	130	126	122
2	109	102	98	93	92	93	93	97	102	107	111	115	117	121	124	124	124	123	122	122	118	114	112	110
3	95	90	85	81	78	79	81	85	90	96	99	103	104	109	110	110	111	111	111	109	107	101	97	94

Appendix B

APPENDIX FOR CHAPTER 4

B.1 Input data for Chapter 4

Table B.1: Base energy demand for microgrid 1-3.

	Microgrid 1			Microgrid 2			Microgrid 3		
Time (h)	**Electricity (MWh)**	**Natural gas (kcf)**	**Heat (MBtu)**	**Electricity (MWh)**	**Natural gas (kcf)**	**Heat (MBtu)**	**Electricity (MWh)**	**Natural gas (kcf)**	**Heat (MBtu)**
1	842.92	387.74	519.30	504.00	210.00	554.40	622.31	311.16	715.66
2	828.44	381.08	531.32	499.20	208.00	549.12	596.93	298.46	686.46
3	820.95	377.64	578.56	504.00	210.00	554.40	580.13	290.06	667.14
4	831.32	382.41	698.73	499.20	208.00	549.12	580.06	290.03	667.07
5	861.44	396.26	972.46	508.80	212.00	559.68	597.68	298.84	687.33
6	924.58	425.31	1176.48	518.40	216.00	570.24	629.41	314.71	723.82
7	939.25	432.06	1137.13	528.00	220.00	580.80	677.44	338.72	779.05
8	981.48	451.48	1091.99	696.00	290.00	765.60	725.28	362.64	834.07
9	956.98	440.21	1032.92	835.20	348.00	918.72	770.57	385.29	886.16
10	937.68	431.33	979.94	936.00	390.00	1029.60	825.62	412.81	949.46
11	935.63	430.39	936.56	998.40	416.00	1098.24	879.69	439.85	1011.64
12	919.00	422.74	911.10	1008.00	420.00	1108.80	928.89	464.45	1068.23
13	916.53	421.60	900.55	1003.20	418.00	1103.52	968.60	484.30	1113.89
14	934.69	429.96	910.40	998.40	416.00	1098.24	998.55	499.28	1148.33
15	961.06	442.09	929.05	1008.00	420.00	1108.80	1018.46	509.23	1171.23
16	1002.37	461.09	968.73	1017.60	424.00	1119.36	1034.82	517.41	1190.04
17	1075.71	494.83	1004.19	1070.40	446.00	1177.44	1038.35	519.17	1194.10
18	1090.57	501.66	1017.83	1027.20	428.00	1129.92	1016.55	508.27	1169.03
19	1074.07	494.07	1016.88	619.20	258.00	681.12	978.60	489.30	1125.39
20	1046.70	481.48	985.27	614.40	256.00	675.84	943.57	471.79	1085.11
21	1011.72	465.39	887.76	590.40	246.00	649.44	900.10	450.05	1035.12
22	968.94	445.71	686.22	489.60	204.00	538.56	831.58	415.79	956.32
23	885.73	407.44	541.79	470.40	196.00	517.44	762.76	381.38	877.18
24	871.34	400.82	578.48	508.80	212.00	559.68	704.32	352.16	809.97

Table B.2: CHP parameters.

Parameter	**Value**
Gas-to-power conversion rate	0.3
Power-to-heat conversion rate	1
Minimum power output (MW/h)	40
Maximum power output (MW/h)	1200
Ramp-up rate (MW/h)	600
Ramp-down rate (MW/h)	600
Operation & maintenance cost ($/kcf)	15
Start-up & shut-down cost ($)	3.48

Table B.3: Heat pump parameters.

Parameter	**Value**
Power-to-heat conversion rate	0.9
Minimum heat output (MBtu/h)	20
Maximum heat output (MBtu/h)	1200
Ramp-up rate (MBtu/h)	600
Ramp-down rate (MBtu/h)	600
Operation & maintenance cost ($/MBtu)	2
Start-up & shut-down cost ($)	3

Table B.4: Maximum power of RES.

Time (h)	PV power(MW/h)	Wind power (MW/h)
1	0.00	0.00
2	0.00	0.00
3	0.00	0.00
4	0.00	0.00
5	0.00	0.00
6	6.43	0.00
7	35.50	0.00
8	67.44	5.98
9	80.50	41.53
10	114.05	77.96
11	127.20	135.86
12	63.41	165.05
13	47.74	94.61
14	48.03	145.59
15	39.42	71.58
16	34.44	88.22
17	13.96	45.99
18	1.92	41.53
19	0.00	18.38
20	0.00	13.05
21	0.00	2.55
22	0.00	1.55
23	0.00	0.56
24	0.00	0.00

Table B.5: Load shifting program parameters.

Shiftable load	Total energy (MWh)	Min. power (MW/h)	Max. power (MW/h)	Time window (h)	Duration (h)
Task 1	250	25	150	2-18	5
Task 2	110	5	50	2-20	8
Task 3	180	20	80	5-22	6
Task 4	150	10	60	3-21	12
Task 5	200	15	100	8-22	10

Tasks 1-5 can be assigned to many appliances, such as dishwashers, electric vehicles and water heaters.

Table B.6: Wholesale electricity and natural gas prices.

Time (h)	**Electricity price ($/MWh)**	**Natural gas price ($/kcf)**
1	67.09	16.93
2	66.89	16.60
3	67.64	15.75
4	68.64	16.85
5	69.54	20.54
6	70.16	21.30
7	71.34	22.15
8	71.04	22.60
9	71.41	23.77
10	71.61	23.73
11	71.23	23.93
12	71.17	23.75
13	70.88	20.25
14	71.10	20.25
15	71.52	20.48
16	72.27	24.50
17	72.82	24.65
18	73.15	25.00
19	72.56	24.60
20	71.19	24.43
21	70.42	20.53
22	70.16	17.98
23	69.36	17.88
24	66.92	17.18