# Virtual Power Plant Construction and Coordination Strategies and Methods for Optimal Market Operation

XIN LU

Supervisor: Prof. Jianguo Zhu

Associate Supervisor: Dr. Jing Qiu

*A thesis submitted in fulfilment of the requirements for the degree of*

*Doctor of Philosophy*

School of Electrical and Computer Engineering

Faculty of Engineering

The University of Sydney

Australia

*December 2023*

*To My Family*

# STATEMENT OF ORIGINALITY

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Name: Xin Lu

Signature: ____________ Date: ____________

# AUTHORSHIP ATTRIBUTION STATEMENT

Chapter 3 of this thesis is published as [X. Lu, J. Qiu, G. Lei and J. Zhu. Scenarios modelling for forecasting day-ahead electricity prices: Case studies in Australia. Applied Energy, 2022, 308: 118296] and [X. Lu, J. Qiu, G. Lei and J. Zhu. An Interval Prediction Method for Day-Ahead Electricity Price in Wholesale Market Considering Weather Factors. IEEE Transactions on Power Systems, 39(2), 2558-2569]. I designed the study, analyzed the data, and wrote the drafts.

Chapter 4 of this thesis is published as [X. Lu, J. Qiu, C. Zhang, G. Lei and J. Zhu. Assembly and Competition for Virtual Power Plants with Multiple ESPs Through a "Recruitment–Participation" Approach. IEEE Transactions on Power Systems, 39(2), 4382-4396]. I designed the study, analyzed the data, and wrote the drafts.

Chapter 5 of this thesis is published as [X. Lu, J. Qiu, C. Zhang, G. Lei and J. Zhu. Seizing Unconventional Arbitrage Opportunities in Virtual Power Plants: A Profitable and Flexible Recruitment Approach. Applied Energy, 358 (2024): 122628]. I designed the study, analyzed the data, and wrote the drafts.

Chapter 6 of this thesis is published as [X. Lu, J. Qiu, C. Zhang, G. Lei and J. Zhu. Deposit and Withdrawal Service Guided Aggregation of Shared Energy Storage with High Price-Tolerance Prosumers. IEEE Transactions on Energy Markets, Policy and Regulation (Under Review)]. I designed the study, analyzed the data, and wrote the drafts.

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Student Name: Xin Lu

Signature: ____________ Date: ____________

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Supervisor Name: Jianguo Zhu

Signature: ____________ Date: ____________

# ACKNOWLEDGEMENTS

This work could not have been accomplished if it was not for the help and contribution of several people. I would like to thank them all from the bottom of my heart.

First and foremost, I extend my deepest gratitude to my supervisor, Prof. Jianguo Zhu His mentorship has been instrumental in my pursuit of a Ph.D. degree. Prof. Zhu's guidance in shaping my research direction and his encouragement to delve deeper into scientific inquiry have been pivotal. His engaging discussions, extending beyond academics and enriched with his vast experiences, have profoundly influenced my journey. My sincere thanks also go to my associate supervisor, Dr. Jing Qiu. Dr. Qiu’s insights into the practical applications of our academic work have been tremendously beneficial. His unique perspectives have greatly enhanced the practicality and applicability of my research.

I want to thank Dr. Cuo Zhang and Dr. Gang Lei for their inspiration and valuable suggestions regarding my research. My heartfelt also appreciation goes out to my friends –Dr. Jun Lin, Ms. Yuan Meng, Mr. Xiao Liu, Ms. Chenxi Zhang, Mr. Yi Yang, Mr. Qiangqiang Xiao, and Mr. Jie Guo. Their collaboration and camaraderie have enriched my academic experience and added joy to my life.

A special note of thanks goes to my parents, whose selfless financial and emotional support has been my foundation. Their unwavering support has enabled me to pursue my academic dreams with dedication and without reservations.

I deeply thank to my wife, Ms Xinyan Yu, for her endless love, patience, and encouragement. She has been my confidante, pillar of strength, and guiding light throughout this journey. Her presence in my life has made all the difference, and I am forever grateful for her companionship.

# LIST OF PUBLICATIONS

The following papers have been completed in the course of my PhD studies, where I was listed as the first author.

[1] **Lu, X.**, Qiu, J., Lei, G., & Zhu, J. (2024). Seizing Unconventional Arbitrage Opportunities in Virtual Power Plants: A Profitable and Flexible Recruitment Approach. ***Applied Energy***, 358 (2024): 122628.

[2] **Lu, X.**, Qiu, J., Lei, G., & Zhu, J. (2024). An Interval Prediction Method for Day-Ahead Electricity Price in Wholesale Market Considering Weather Factors. ***IEEE Transactions on Power Systems***, 39(2), 2558-2569.

[3] **Lu, X.**, Qiu, J., Zhang, C., Lei, G., & Zhu, J. (2024). Assembly and Competition for Virtual Power Plants with Multiple ESPs Through a "Recruitment–Participation" Approach. ***IEEE Transactions on Power Systems***, 39(2), 4382-4396.

[4] **Lu, X.**, Qiu, J., Lei, G., & Zhu, J. (2023). State of Health Estimation of Lithium Iron Phosphate Batteries based on Degradation Knowledge Transfer Learning. *IEEE Transactions on Transportation Electrification*, 9(3), 4692-4703.

[5] **Lu, X.**, Qiu, J., Lei, G., & Zhu, J. (2022). Degradation Mode Knowledge Transfer Method for LFP Batteries. ***IEEE Transactions on Transportation Electrification***, 9(1), 1142-1152.

[6] **Lu, X.**, Qiu, J., Lei, G., & Zhu, J. (2022). Scenarios Modelling for Forecasting Day-Ahead Electricity Prices: Case Studies in Australia. ***Applied Energy***, 308, 118296.

The following manuscripts have been submitted and are currently under review:

[7] **Lu, X.**, Qiu, J., Lei, G., & Zhu, J. Deposit and Withdrawal Service Guided Aggregation of Shared Energy Storage with High Price-Tolerance Prosumers. *IEEE Transactions on Energy Markets, Policy and Regulation* (Under Review)

# ABSTRACT

The rapid growth of distributed energy resources (DERs), including rooftop photovoltaics (PVs) and batteries, has significantly transformed the landscape of renewable energy supply. With the emergence of virtual power plants (VPPs), energy service providers (ESPs) can now aggregate diverse DERs and empower them to participate in energy markets as prosumers. This creates an attractive opportunity for prosumers to generate profits in various electricity markets, such as the spot and ancillary services markets. Previous studies have primarily concentrated on the aggregation and profitability of VPPs in the energy market, employing optimized strategies. However, these studies often assume the existence of a VPP and overlook the crucial aspect of recruiting DER owners to participate in the VPP.

Nonetheless, previous studies often overly idealize prosumers' willingness to participate in VPPs, failing to account for many concerns of DER owners, such as low returns and limited plan options. Consequently, the existing studies lack the necessary incentives to encourage prosumers to participate in VPPs. Additionally, concerns persist among DER participants regarding the profits of ESPs, as these service providers determine the existing profit allocation rules within VPPs. Furthermore, the conventional fixed recruitment approach for VPPs lacks the flexibility required to accommodate temporary expansions in participant numbers, resulting in limited VPP profitability and hampering its potential for maximizing returns. Hence, it is imperative to reevaluate the current approach to VPP construction.

This thesis aims to present innovative frameworks and methods for VPP construction and coordination to address the research gaps highlighted earlier. Besides the three chapters of introduction, literature review, and conclusion, the main body of the thesis can be divided into four parts, each aligned with distinct VPP construction frameworks catering to various prosumer types and associated technologies.

The first part is Chapter 3, where we introduce the conditional time-series generative adversarial network (CTSGAN). This network is the foundation for our research by providing predictive models for wholesale electricity price.

The second part, represented by Chapter 4, proposes a recruitment-participation approach as an alternative to the conventional ESP-centered VPP construction model. ESPs offer diverse risk-return strategies in this approach and engage DERs in collaborative VPP creation. A fair and incentive-based payoff allocation method is also introduced.

Chapter 5 constitutes the third part, presenting a flexible and profitable VPP recruitment approach optimized to seize unconventional arbitrage opportunities (UAOs). This approach integrates regular and casual recruitment mechanisms to accommodate both conservative and ambitious DER participants.

The fourth part, found in Chapter 6, introduces a framework for aggregating shared energy storage (SES) among residential prosumers with a high price tolerance for price fluctuations. This framework includes a credit-based electricity deposit and withdrawal service (DWS) and incorporates a matching mechanism that streamlines prosumer participation without additional actions or ESP coordination.

Throughout Chapters 4, 5, and 6, we leverage data-driven methods to optimize incentives, coefficients, and energy trading strategies, effectively improving the overall profitability of VPPs.

This thesis delves into utilizing frameworks, models, and solution methods within the VPP construction process. A series of benchmark tests are employed to showcase the effectiveness of the proposed method. Experimental studies are conducted to compare the proposed method with existing approaches from the literature. The findings unequivocally demonstrate that our VPP construction and coordination methods introduces enhanced flexibility, thereby enabling ESPs and DER participants to attain superior profitability.

Our proposed VPP construction and coordination frameworks and methods can effectively harmonize various prosumer types with diverse preferences, thereby facilitating a more flexible and efficient operation of the VPP. They prove to be practical and conducive to the market-oriented operation of VPPs.

# CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# GLOSSARY

| | |
|---|---|
| AEMO | Australian Energy Market Operator |
| ANN | Artificial Neural Network |
| ARIMA | Auto Regressive Integrated Moving Average |
| BESS | Battery Energy Storage System |
| BOM | Bureau of Meteorology |
| BPNN | Back Propagation Neural Network |
| CEPR | Combined Experience Pool Replay |
| CNEPR | Combined Neighboring Experience Pool Replay |
| CNN | Convolutional Neural Network |
| CP | Coverage Probability |
| CPI | Consumer Price Index |
| CTSGAN | Conditional Time Series Generative Adversarial Network |
| DDPG | Deep Deterministic Policy Gradient |
| DER | Distributed Energy Resource |
| DRL | Deep Reinforcement Learning |
| DWS | Deposit and Withdrawal Service |
| ECM | Equivalent Circuit Model |
| EIS | Electrochemical Impedance Spectroscopy |
| ELM | Extreme Learning Machine |
| ESP | Energy Service Provider |
| ESS | Energy Storage System |
| EV | Electric Vehicle |
| FIT | Feed-in Tariff |
| GAN | Generative Adversarial Network |
| GDP | Gross Domestic Product |
| GRU | Gated Recurrent Unit |
| LASSO | Least Absolute Shrinkage and Selection Operator |
| LSTM | Long Short-Term Memory |

| | |
|---|---|
| LUBE | Lower Upper Bound Estimation |
| MADRL | Multi-Agent Deep Reinforcement Learning |
| MADDPG | Multi-Agent Deep Deterministic Policy Gradient |
| MAE | Mean Absolute Error |
| MAPE | Mean Absolute Percentage Error |
| MEPR | Multiple Experience Pool Replay |
| MLEPR | Multi-Level Experience Pool Replay |
| MDP | Markov Decision Process |
| ML | Machine Learning |
| NCP | Nominal Coverage Probability |
| NEM | National Electricity Market |
| NN | Neural Network |
| NSW | New South Wales |
| OC | Operation Cycle |
| PI | Prediction Interval |
| PSO | Particle Swarm Optimization |
| PV | Photovoltaics |
| QR | Quantile Regression |
| QRRF | Quantile Regression Random Forest |
| RMSE | Root Mean Square Error |
| RMSProp | Root Mean Square Propagation |
| RNN | Recurrent Neural Network |
| SA | South Australia |
| SES | Shared Energy Storage |
| SOH | State of Health |
| SV | Shapley Value |
| TD3 | Twin Delay Deep Deterministic Policy Gradient |
| TOU | Time-Of-Use |
| TSGAN | Time Series Generative Adversarial Network |
| TSM | Threshold Select Machine |

| | |
|---|---|
| t-SNE | t-Distributed Stochastic Neighbour Embedding |
| UAO | Unconventional Arbitrage Opportunity |
| UAT | Unconventional Arbitrage Threshold |
| VPP | Virtual Power Plant |
| WNN | Wavelet Neural Network |
| WS | Winkler Score |

# CHAPTER 1

# Introduction

*This chapter first introduces the background of VPPs, including the current state of VPPs worldwide, wholesale electricity prediction, and existing construction and coordination methods. It then presents the motivations, clarifying the research gaps. Following this, the chapter outlines the contributions of this thesis, categorized according to the prediction frameworks and methods, and the construction and coordination frameworks and methods. Finally, the chapter provides an outline of the thesis, describing the content of each chapter and presenting a graph illustrating the relationships between the chapters.*

## 1.1 Background

The rapid increase in the number of distributed energy resources (DERs), such as rooftop photovoltaics (PVs) and batteries, has brought about significant changes in the renewable energy supply [1, 2]. DERs, intended initially for self-consumption, were previously precluded from direct involvement in energy trading. Recently, energy service providers (ESPs), which are companies that manage the distribution and sale of electricity to consumers, have started to play a crucial role in integrating DERs into the electricity grid. ESPs facilitate the connection between individual prosumer and the electricity market, enabling these small-scale prosumers to become active participants in electricity trading [3]. This development has been significantly supported by the advent of virtual power plants (VPPs), which aggregate the capacities of various DERs to provide a reliable and flexible energy supply. Through VPPs, ESPs help to optimize energy usage, improve grid stability, and promote the use of renewable energy sources.

VPPs have changed the global energy landscape [4, 5]. In Europe, Next Kraftwerke stands out as one of the largest VPP operators, connecting more than 10,000 units including biomass, hydroelectric power, and solar energy, with a combined capacity of over 8 GW. Similarly, Enel X operates one of the world's most extensive VPPs, with over 6 GW of aggregated capacity across various countries, incorporating demand response services, battery storage, and renewable sources to promote grid flexibility on a global scale. In Germany, the SonnenCommunity links residential solar battery systems, enabling energy sharing among members and showcasing the potential of residential storage in energy markets. Meanwhile, in the United States, Virtual Peaker employs cloud-based software to connect residential DERs like smart thermostats and water heaters, creating a versatile VPP that assists utilities in managing demand, cutting costs, and facilitating the shift towards renewable energy. There are also several commercial VPP projects in Australia by the end of 2023, ranging from small-scale trials to larger deployments. AGL, one of Australia's largest energy retailers, initiated a VPP project that initially began as a trial involving 1,000 residential batteries, aiming to reach a combined storage capacity of 5 MW [6]. A VPP has been developed in support of the government of South Australia (SA), Tesla, energy retailers and energy locals [7]. This project aimed to eventually connect 50000 households, creating a 250 MW solar generation and 650 MWh storage VPP. These commercial VPP projects typically aim for profitability, with dispatch strategies for battery charging, discharging, and participation in electricity market transactions, being the main subjects of research.

Electricity price prediction technology in the wholesale market plays a crucial role in VPPs, which allows ESPs to make decisions in advance, such as the time scheduled to charge or discharge energy storage. This helps in buying electricity when it is low and selling or dispatching when the price is high, therefore increasing the economic viability of VPPs. In addition, accurate electricity price prediction can provide ESPs with insights into potential price spikes or drops, allowing them to manage risks associated with electricity trading.

Currently, VPPs are in their pilot phase or early stages of commercial operation, and there is a significant variance in their internal structures and operating models. For instance, the AGL VPP is initiated and aggregated by the electricity retailer AGL, while the SA VPP is organized and aggregated by the battery retailer Tesla. However, since these ESPs often provide a uniform VPP plan, which cannot meet the various requirements of DER owners with different risk tolerances and payoff expectations [8, 9], DER owners are currently not enthusiastic about participating in VPPs. Moreover, since VPPs often have a fixed number of DER participants, there is a lack of flexibility for the VPP to adapt to fluctuations in electricity prices. Finally, for ESPs, recruiting batteries is often more profitable than recruiting PVs. Therefore, the existing ESP-centred VPP construction approaches must be revisited, and VPP recruitment and aggregation modes will become more flexible and profitable.

## 1.2 Motivations

ESPs aggregate DERs as VPPs [3] to participate in electricity markets [4, 10] to gain extra profits. In the electricity market, buying low and selling high is the most commonly used strategy, and the most crucial issue within it is the accurate prediction of electricity prices. Price prediction is typically divided into deterministic point prediction and indeterministic interval prediction. Currently, there is a lack of a unified prediction method that can balance the accuracy of deterministic predictions and the diversity of indeterministic predictions.

Conventional research typically focuses on optimizing VPP scheduling problems while overlooking the more practical issues during the VPP formation. We identify the following research gaps: 1) When multiple ESPs offer different VPP plans within a specific region, how do ESPs compete and provide incentives to meet the DER participants with various risk-return preferences while striving to recruit more DER participants to maximize their own profits? 2) In VPP operations, there are many typical unconventional arbitrage opportunities (UAOs), and accurately predicting and capitalizing on these UAOs is crucial for VPP profitability. Current VPP models often use long-term fixed recruitment rather than short-term flexible recruitment, which limits their ability to benefit from these UAOs.

Existing research often assumes that DER participants are price-sensitive, but residential users typically have a higher price tolerance, making it challenging to motivate them with minor profit incentives. There is a lack of an effective mechanism to identify and select suitable high price-tolerance prosumers to participate in the shared energy storage (SES). As a result, these high price-tolerance prosumers are often excluded from the aggregation of SES. Bridging the above gaps needs an efficient framework.

## 1.3 Contributions

This thesis aims to develop the VPP construction and coordination approaches to fill the above research gaps. The contributions of this thesis are categorized according to the prediction frameworks and methods, construction and coordination frameworks and methods, as follows:

On Prediction Frameworks:

1) A novel conditional time series generative adversarial network (CTSGAN) deep learning framework is proposed to capture the distribution of each period and preserve temporal variations. In addition, Wasserstein Distance, weights limitation, and root mean square propagation (RMSProp) optimizer are used to ensure the stability of the CTSGAN training process.
2) A UAO is proposed to describe and assess future profit opportunities. Moreover, a CTSGAN-based UAO prediction method that accounts for weather conditions is proposed, which can help an ESP anticipate future UAO scenarios.

On Prediction Methods:

1) CTSGAN is introduced to generate highly realistic generated scenarios, therefore, the electricity point can be predicted. In addition, by combining predicted electricity price points, the prediction intervals (PIs) can be obtained. By adjusting the random noise input of CTSGAN, the point prediction model is eventually transformed into an interval prediction model.
2) A novel scenario-based PIs construction method is proposed by using a threshold select machine (TSM), which can adjust the truncated threshold of random noise input in CTSGAN to construct PIs that meet the nominal coverage probability (NCP). Weather factors are incorporated into the TSM, further improving the reliability and sharpness of the constructed PIs.

On Construction and Coordination Frameworks:

1) A recruitment-participation approach is proposed to replace the ESP-centered VPP construction model, where ESPs offer different risk-return strategies and recruit DERs to build a VPP together.
2) A flexible and profitable VPP recruitment approach that integrates regular and casual recruitment mechanisms, is proposed to maximize profits when UAO occurs, catering to conservative and ambitious DER participants.
3) A framework that includes construction and operation steps is proposed to aggregate an SES among high price-tolerance residential prosumers, This framework allows prosumers to participate without requiring additional actions or coordination with ESP.

On Construction and Coordination Methods:

1) A bet-on VPP mode is developed to eliminate the prerequisite for DERs and the ESP to reach a consensus before VPP formation. This allows for aggregating specific DER participants who have reservations about expected profits. If these expected profits are not realized, the concerned DER participants are eligible to receive compensation from the ESP.
2) A payoff allocation method based on fairness and incentives is introduced. A

fair profit allocation method is developed for individual DERs and ESPs based on the Shapley Value (SV) method. The incentive approach aims to encourage DERs to participate in VPPs through rebates.

3) A multi-ESP VPP model that considers DERs' behaviors and allows for ESP selection to meet diverse needs of DERs is proposed. Multi-agent deep reinforcement learning (MADRL) is employed to solve the rebate competition problem among ESPs, achieving a stable solution among ESPs.
4) Different incentive approaches are developed for different recruitment modes. A combined experience pool replay (CEPR)- twin delay deep deterministic policy gradient (TD3) algorithm is proposed to optimize the incentive coefficient, which can encourage DER participants and help the ESP achieve high profits.
5) A credit-based electricity deposit and withdrawal service (DWS) method is developed, which allows prosumers to have the freedom to deposit or withdraw electricity with dynamic coefficients and receive electricity gains. A matching mechanism is proposed to facilitate mutual selections between the ESP and prosumers, ensuring that prosumer electricity bills are reduced as expected. The ESP employs this mechanism to screen users to maximize overall profitability.
6) A combined neighbouring experience pool replay (CNEPR)-TD3 is proposed, which can search for similar training sets within a certain distance using multiple labels and reconstruct the experience pool. The modified TD3 can provide reasonable dynamic coefficients and energy trading strategies, effectively improving SES’s overall profit.

## 1.4 Thesis Outline

The rest of the thesis is organized as the following:

**Chapter 2** provides an overview of the existing research on VPP, including the conventional formation models and market models of VPP. It then reviews commonly used data-driven methods in VPP economic operation, such as deep learning and reinforcement learning. This chapter also reviews the wholesale electricity price prediction involved in VPP operations.

**Chapter 3** develops a novel generative model for electricity price prediction, which can capture the distribution of each period and preserve temporal variations of electricity prices. The proposed model can directly generate the deterministic electricity price point prediction. The deterministic model can be transformed into an interval prediction model by incorporating weather information and modifying random noise inputs. The generative model presented in this chapter is also employed for UAO prediction in Chapter 5. The electricity price prediction results in this chapter serve as the cornerstone for the research presented in Chapters 4 - 6.

**Chapter 4** proposes a recruitment-participation approach for multi-ESP VPP construction, where ESPs compete with each other and recruit DERs to build a VPP together by offering different risk-return strategies. A payoff allocation method based on fairness and incentives is developed in this chapter to encourage DER participation. This chapter first simulates the rebate competition between multi-ESPs with modified multi-agent deep deterministic policy gradient (MADDPG). The concept of 'recruitment-participation' that can transform the conventional ESP-centred VPP formation approach. In this recruitment-participation approach, both ESPs and DERs are equal VPP participants, which can change the current ESP-dominated profit allocation approach, further increasing participant profits and participation rates.

**Chapter 5** presents a flexible and profitable short-term recruitment approach for VPPs when ESPs forecast future UAOs. Combined with regular recruitment approach, this approach can cater to conservative and ambitious DER participants. This chapter introduce a bet-on VPP mode, which eliminates the prerequisite for DERs and the ESP to reach a consensus before VPP formation. This allows for aggregating specific

DER participants who have reservations about expected profits. A CEPR-TD3 algorithm is proposed to optimize the incentive coefficient, which can encourage DER participants and help the ESP achieve high profits.

**Chapter 6** proposes a framework to aggregate an SES among high price-tolerance residential prosumers with a credit-based electricity DWS. This framework allows prosumers to participate without requiring additional actions or coordination with ESP. A matching mechanism is proposed to facilitate mutual selections between the ESP and prosumers, ensuring that prosumer electricity bills are reduced as expected. The ESP employs this mechanism to screen users to maximize overall profitability. A combined neighbouring experience pool replay (CNEPR)-TD3, which can search for similar training sets within a certain distance using multiple labels and reconstruct the experience pool. The modified TD3 can provide reasonable dynamic coefficients and energy trading strategies, effectively improving SES's overall profit.

Finally, **Chapter 7** summarizes the main conclusions and suggests the directions of future works.

Fig. 1.1 illustrates the structure and relationship among the main technical contents in this thesis.

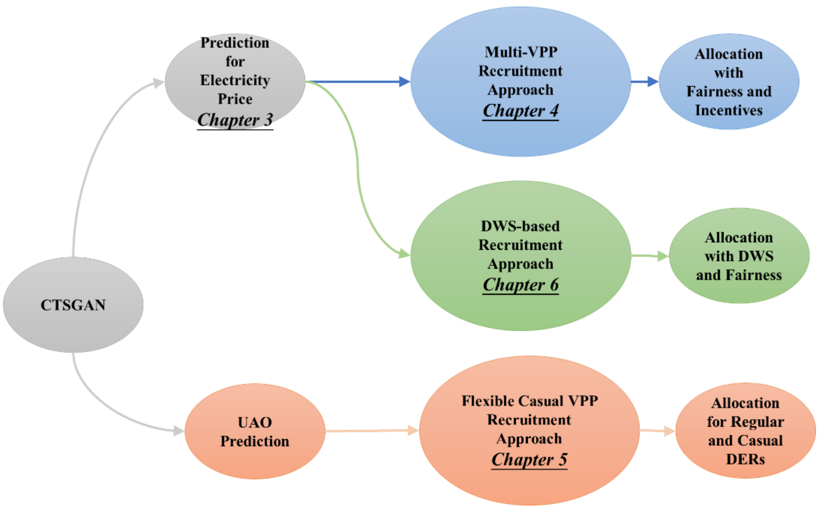



Fig. 1.1 Mind map for thesis outline

# CHAPTER 2

# Literature Review

*This chapter first reviews the conventional profitable VPPs with their DER components, then reviews wholesale electricity price prediction. The pricing mechanisms, such as dynamic incentives and rebate competition, are reviewed for DER participants in a VPP. Finally, this chapter reviews the VPP construction and coordination approaches with DER participants, focusing on VPP business models and recruitment approaches.*

## 2.1 Conventional Profitable VPPs

### 2.1.1 Conventional VPP and Components

The rapid increase in the number of DERs, such as rooftop PVs and batteries, has significant changed the renewable energy supply. PV systems, using solar panels, convert solar energy into electrical energy. They are an important part of renewable energy technology, helping to reduce reliance on fossil fuels and cut carbon emissions. PV systems can be installed in residential and commercial buildings, and in solar power stations. On the other hand, batteries are used to store electrical energy for later use. They play a key role in managing peak energy demand and improving the stability and efficiency of the energy system. These systems are especially useful in combination with renewable energy sources, allowing storing and releasing of excess energy as needed [1, 2]. However, DERs, intended initially for self-consumption, were precluded from direct involvement in energy trading. With the advent of VPPs, ESPs can now aggregate various DERs and enable them to participate in energy markets as prosumers [3]. This presents an opportunity for prosumers to earn profits in various electricity markets, such as the spot and ancillary markets [4, 5].

### 2.1.2 Profitable VPP Objectives

Liu et al. [11] proposed a risk-averse bidding strategy with a data-driven distributional robust model. A DRL method was proposed to model the bidding action and achieve a higher profit for VPPs [12]. A competitive framework was designed to model the participation of VPPs in the electricity market with multiple objectives, including profit maximization and cost minimization, through a multi-leader-follower game theory and Karush–Kuhn–Tucker conditions [13]. Cucchiella et al. [14] examined the residential PV and energy storage system (ESS) scale of VPP in an Italian electricity market. They found the self-consumption levels were crucial for the profitability of PV. However, they also concluded that household ESS did not yield profitability. Similarly, Joern et al. [15] investigated the PV-ESS VPP in Germany, considering eight electricity price scenarios from 2013 to 2022. They proposed a cost-optimal investment model and observed that the higher electricity retailer price and the lower wholesale price could enhance the profitability of ESS. Ren et al. [16] analyzed VPPs with different electricity price structures, PV and battery sizes in Australia. Another study conducted by Khalilpour et al. [17] revealed that the benefits of ESS became evident for VPPs when the feed-in tariff (FIT) was substantially lower than the retail electricity prices. While these studies emphasized profit maximization with an optimized operational strategy [18, 19] after successfully constructing a VPP, they overlooked the importance of incentivizing and recruiting DER participants.

## 2.2 Wholesale Electricity Price Prediction

Electricity price prediction has attracted considerable attention [20, 21] because an accurate forecast of electricity prices plays an essential role [22, 23] in supporting the behaviors of market participants. Electricity prices are usually determined by several dependencies, such as electricity demand, business activities, temperature, and holidays. In addition, the electricity prices may also be affected by various potential factors, such as the large-scale application of artificial intelligence algorithms in the electricity trading market [24] or even the personal bidding strategies of electricity traders. Uniejewski et al. [25] developed a seasonal component autoregressive forecasting model, consisting of a seasonal trend component and a stochastic component, for significant accuracy gains. By evaluating the merit order effect of wind power generation penetration on electricity prices in the national electricity market (NEM), Bell et al. [26] and Forrest et al. [27] concluded that the increasing wind generation penetration decreased wholesale spot prices in Australia. In [28], Nazifi et al. pointed out that carbon pricing would indeed be fully passed on to wholesale electricity spot prices resulting in higher electricity prices for consumers. Higgs et al. [29] concluded that the spot price exhibited stronger mean-reversion after a price spike than in the normal period, and the volatility was more than 14 times higher in spike periods than in normal periods. These factors are difficult to analyse quantitatively, making electricity price prediction significantly more challenging in recent years.

In terms of forecasting objectives, electricity price forecasting can be divided into point forecasting and probabilistic forecasting. Point forecasting is generally deterministic forecasting that directly gives the exact outcome. For example, the deterministic outcome of day-ahead electricity price forecasting is to provide each half-hourly price of the next 24 hours, a total of 48 electricity price points. Most of the previous studies [25, 30-32] are about point forecasting of electricity prices and they are concerned about the root mean square error (RMSE) and mean absolute error (MAE) between the forecasted and actual data [33, 34]. However, some uncertainties are inevitable, usually due to incomplete data or some unanticipated events. In recognition of the inherent limitations of traditional point prediction models, interval prediction of electricity prices is used to compensate for the low accuracy of point prediction and describe the uncertainties of electricity prices. Bootstrap method is the most common PIs method, considering historical prediction errors and parameter uncertainty [35]. A lower upper bound estimation (LUBE) method was proposed by Khosravi et al. [36] for PIs, directly giving the upper and lower bounds with an artificial neural network (ANN). A Pareto PIs construction method was proposed by Wan et al. [37] to balance the reliability and sharpness of PIs and achieve good prediction performance. However, the above LUBE methods were constructed directly by minimizing a cost function called the coverage width criterion. Selecting a reasonable penalty coefficient becomes difficult, and a huge penalty term may lead to PIs containing too many scenarios with low or zero probability, especially in price fluctuations, which will degrade the prediction performance. Marcjasz et al. [38] constructed the PIs from the point scenarios, which were obtained from five independent machine learning (ML) -based point prediction models, by quantile regression (QR) methods, and the scenario-based PIs construction method outputted accurate PIs with different quantiles. However, constructing point prediction and probabilistic prediction is often done through different methods, lacking a framework that can simultaneously provide both point prediction and probabilistic prediction.

## 2.3 VPP Profit Allocation Approach

### 2.3.1 Dynamic Pricing Mechanism

Several research works have been conducted on increasing the profits of DER participants in VPP. The dynamic pricing mechanism is the most common approach. Nguyen et al. [39] designed an optimal dynamic pricing model for demand response integration. The load-serving entity was considered the leader, and the load aggregators were followers in a Stackelberg game. A dynamic retail pricing scheme was studied in [40] with load demand and market price information based on a binary genetic algorithm. However, these VPPs are constructed by ESPs, commonly the profit allocation rule makers. Although ESPs provide DERs with dynamic prices, they occupy the interest space of DERs [41].

### 2.3.2 Profit Allocation with Fairness and Incentives

A commonly used approach to achieve fair profit allocation is to base it on the contribution of each VPP participant to the whole. The SV method is one of the most common solutions and was introduced in [42, 43] to allocate benefits among PVs, wind power plants, gas turbines, and batteries. However, the ESP's contribution was not evaluated, and no payoff was allocated to the ESP. In [44], Gougheri et al.

proposed an SV-based fair allocation method to divide profit among PVs, batteries, and other components. However, there are multiple PVs and batteries in a VPP, and the allocation method in [16] only evaluates the contributions of the PV and battery groups but cannot allocate precise profits to each battery or PV within the groups. Therefore, a fair method for profit allocation to every DER individual instead of DER groups is required. In addition to fair benefit allocation, providing DER participants with certain incentives is a commonly used method in VPP profit distribution. Customized prices or rebates are the most common ways. Chen et al. in [45] proposed a customized pricing mechanism to motivate VPP prosumers. The interaction between a retailer and prosumers is modelled as a Stackelberg game and solved by a Salp Swarm Algorithm. A customized rebate pricing mechanism was proposed in [46] to reward prosumers for supporting the power system operation, considering the heterogeneous prosumers' dynamic selection process based on an evolutionary game. Therefore, it is important to design a set of profit allocation measures that combine fairness and incentives, which are crucial for the aggregation and coordination of VPP.

## 2.4 VPP Construction Approach

### 2.4.1 Recruitment-Participation Approach

At the outset of a VPP formation, ESP typically advertises the participation benefits and payoffs in the VPP to potential DER owners to gain control authority over their DERs and eventually form the VPP [47, 48]. This process is commonly called recruitment-participation, with the former for ESPs and the latter for DER owners. However, ESPs often provide a uniform VPP plan to recruit DER participants, which is not enough to meet the requirements of DER owners with different risk tolerances and payoff expectations. In [49], Liu et al. proposed a conditional value-at-risk strategy for VPP operation, which enables the ESP to adjust the risk of the VPP as a whole but not for individual DER participants. Meanwhile, Chen et al. [4] proposed a bargaining-game-based VPP model that accounts for personal risk preferences but requires the ESP to negotiate with each participant for different payoffs, potentially leading to a complicated VPP formation process. A multi-ESP competition approach was proposed in [50] to model the incentivization and recruitment of DER participants with varying investment preferences. While the concept of flexible VPP recruitment was introduced, the focus lay on stable solutions in the multi-ESP game rather than profit optimization. Moreover, each VPP discussed in [50] only offered a single participation plan, failing to attract DER participants with diverse needs. Diverse VPP participation plans provided by multiple ESPs or a single ESP can more effectively meet the various needs of DER participants.

### 2.4.2 Participants' Willingness for VPP Construction

The studies mentioned above focused solely on the increase of individual DER profits, mainly from the perspective of ESPs, without modeling the behaviors of DER owners. Moreover, it remains unclear whether these studies provide incentives for the participation rate of DER participants. Currently, there is a lack of literature on modeling the willingness of DERs to participate in VPPs or to choose ESPs. However, some studies have proposed nonlinear models to represent the willingness to install solar panels [51], the behavior of EVs to accept dispatch [52], and the behavior of consumers in selecting electricity retailers [53]. Nevertheless, these models do not entirely address the issue, as consumers primarily aim to reduce their bills. DER owners consider more than just profit when selecting ESPs, as their demands are diverse. Some prioritize high short-term returns while others seek stable long-term returns. Therefore, differentiated VPP plans are necessary to accommodate the varying demands of different DER owners. Since introducing the multi-retailer model has brought dynamism and increased efficiency to the electricity market [54], introducing ESP competition in VPPs can ensure their economic operation with full competitiveness and provide differentiated services to attract participants. Xu et al. modelled competitive behaviors between two VPPs based on a non-cooperative pricing game in [3] with limited DERs. They partitioned the original optimization problem into two sub-problems. An interactive dispatch model of VPPs was proposed in [55]. The aforementioned studies focused on the competition between two or three VPPs, but more complex competition among a larger number of VPPs still needs to be investigated for practical implementation. The rebate competition problem among multiple ESP is difficult to solve using conventional mathematical methods. In recent years, DRL algorithms, combining conventional reinforcement learning and deep learning, have provided a new opportunity to address such problems. The deep deterministic policy gradient (DDPG) algorithms and TD3 algorithms [56] have been used to solve the profit game between EV charging stations and VPPs. Ma et al. designed a DDPG-based incentive determination agent to maximize the profits of both operators and prosumers simultaneously [57]. MADRL was proposed in [58] to realize coordinated control and improve the performance of VPPs in the frequency regulation market. MADDPG was deployed to obtain the optimal dispatching strategy of the main network and microgrids in a bi-level energy management framework. However, the aforementioned studies often assume that price is the sole factor influencing participants' choices of VPP plans without considering the impact of other factors. To simulate participants' choice preferences more realistically, it is necessary to consider other non-price factors.

### 2.4.3 Consensus VPP Construction

DER participants in a VPP seek returns on assets like PVs and batteries, while the ESP aims to grow the VPP and increase profits. During the VPP formation process, ESP is viewed as investee and DER participants as investors [59]. Both investment parties must reach a mutual understanding of risks and returns [60], commonly called achieving consensus. However, DER participants may have concerns about participating in VPPs due to the fear of actual returns falling short of expected returns. Such concern often leads DER participants to resist being controlled by the ESP, thereby refusing to join the VPP. To address this, Saeed et al. [61] suggested accounting for and addressing uncertainties to control the risk of profit variability. Despite this, it remains unclear if this approach can fully alleviate the doubts of DER owners and establish a consensus with the ESP. Many studies have explored the issue from a technical perspective, proposing various mechanisms to ensure that DERs reach a consensus and achieve economic dispatch. These include consensus control [62] and blockchain technology [63]. However, when consensus cannot be reached, how ESP recruits DER participants remains unresolved. A potential solution is to introduce a bet-on agreement, also known as a valuation-adjustment mechanism, commonly used in the venture capital industry and related to several factors, such as severe information asymmetry, high initial investment, and uncertainty [64]. The following issues are also present in VPP formation: (a) Operating strategies and market conditions are often confidential, and the ESP does not share this information with DER owners; (b) The initial investment in VPP is high, and the ESP cannot afford all PVs and batteries; (c) The electricity market is uncertain, and the expected payoffs are not always realizable. Therefore, introducing a bet-on agreement [65] in forming a VPP is viable. The agreement establishes a set of pre-determined payoff conditions, the fulfillment or non-fulfillment of which confers the DER owners a contractual right to get compensation from the ESP.

### 2.4.4 High Price-Tolerance Prosumers VPP Construction

Individual prosumers invest in the ESS to complement the surplus power generated by their PV. The primary motivation for these prosumers to install an ESS is to enhance the flexibility of accessing and utilizing the stored energy according to their demands. However, once the VPP comes into play as an aggregated system, the liberty of ESS becomes more of a luxury than a given due to the constraints enacted by the governing SES regulation. Additionally, the current VPP often requires participants to perform additional actions, like optimizing rent capacity [66], reporting electricity usage information [67], submitting demands and bidding prices through auctions [68], and participating in internal energy trading [69]. It is worth questioning the attractiveness of the conventional VPP for high price-tolerance residential prosumers who experience diminished control over ESS and face increased complexity of additional actions. For the high price-tolerance prosumers, directly involving them in constructing a VPP may not be suitable. Their unregulated charging and discharging behaviors could introduce uncertainty to the VPP.

Moreover, their reluctance to engage in additional actions may limit the potential profitability of the VPP. Therefore, it is essential to propose a new model for the high price-tolerance prosumers, which requires no further actions and prioritizes freedom storage as a prerequisite for VPP construction. Additionally, a matching mechanism is necessary to identify suitable prosumers and ensure profitability for all participants.

## 2.5 Chapter Summary

This chapter extensively reviews the literature concerning conventional profitable VPPs and their components, methods of predicting wholesale electricity prices, pricing mechanisms for DER participants, and strategies for VPP construction and coordination. While significant progress has been made in existing research, numerous challenges remain for further exploration. This chapter highlights the prevailing research gaps in these areas and lays the groundwork for the thesis in the subsequent chapters.

# CHAPTER 3

# Prediction Methods for Day-ahead Electricity Price in Wholesale Market

*Electricity prices in wholesale market are volatile and can be affected by various factors, such as generation and demand, system contingencies, local weather patterns, bidding strategies of market participants, and uncertain renewable energy outputs. Because of these factors, electricity price prediction is challenging. This chapter proposes a scenario modeling approach to improve forecasting accuracy, conditioning time series generative adversarial networks on external factors. After data pre-processing and condition selection, a CTSGAN is designed to forecast electricity prices. Wasserstein Distance, weights limitation, and RMSProp optimizer are used to ensure that the CTSGAN training process is stable. By changing the random noise input, the point prediction model can be transformed into an interval prediction model. For electricity price point prediction, the proposed CTSGAN model has better accuracy and has better generalization ability than the TSGAN and other deep learning methods. The proposed CTSGAN model can significantly improve the coverage probability (CP) and winkler score (WS) for interval prediction.*

## 3.1 Prediction Model for Electricity Price

A good generative model should be able to capture not only the distribution of features at each time but also the complex dynamics of these variables across time. However, the existing generative methods that bring GAN into the sequential setting cannot adequately address the temporal correlations unique to time-series data. A new framework for generating realistic time-series that integrates the versatility of unsupervised training with the control of supervised training is proposed to learn to encode features, generate representations, and iterate across time simultaneously. The embedding network provides the latent space, and the adversarial network operates within this space. The latent dynamics of both real and synthetic data are synchronized through a supervised loss. Since electricity price forecasting involves various exogenous conditions, CTSGAN is proposed by using relevant conditions to modify TSGAN. In addition, the Wasserstein Distance is introduced to accelerate the training process and keep it stable. Moreover, two methods are applied to ensure the CTSGAN stability by (a) limiting the weights of the discriminator network and (b) introducing the RMSProp optimizer to replace the Adam optimizer. This section introduces the proposed CTSGAN model and the steps to forecast the day-ahead price scenarios.

### 3.1.1 Conventional GAN Model

The GAN model is mainly composed of two neural networks (NNs), i.e., the generator $G(\cdot)$ and the discriminator $D(\cdot)$. The generator's responsibility is to generate a massive amount of realistic scenarios with different random noise inputs $z$, under a well-defined noise distribution, $Z \sim \mathbb{P}_z$ (Gaussian distribution or uniform distribution) until the discriminator cannot distinguish them from the real scenarios $x$. The responsibility of the discriminator, which is essentially a classifier, is to distinguish the generated scenarios and the real historical scenarios. Denote the generator neural network and discriminator neural network as $G(z;\theta^{(G)})$ and $D(z;\theta^{(D)})$, and the weights of neural networks in the generator and discriminator as $\theta^{(G)}$ and $\theta^{(D)}$, respectively. The ideal output of $G(z;\theta^{(G)})$ should follow the real historical data distribution, $X \sim \mathbb{P}_x$, after being well trained. The discriminator is trained to identify $\mathbb{P}_x$ from $\mathbb{P}_z$, and thus to maximize the difference between $D(z;\theta^{(D)})$ and $D(G(z;\theta^{(G)});\theta^{(D)})$. According to the objectives of the generator and the discriminator, loss functions $L_G$ and $L_D$ are formulated to train and optimize the weights of the neural networks of the generator and the discriminator, respectively. A small $L_G$ indicates that the generator can synthesize realistic scenarios which as if comes from the real historical scenarios. A small $L_D$ reflects that the discriminator cannot identify whether the generated scenarios are from the historical scenarios. Define the $L_G$ and $L_D$ as:

$$L_G = -\mathbb{E}_Z\left[log\left(1 - D\left(G\left(z;\theta^{(G)}\right);\theta^{(D)}\right)\right)\right] \tag{3.1}$$

$$L_D = -\mathbb{E}_X\left[log\left(D\left(x;\theta^{(D)}\right)\right)\right] - \mathbb{E}_Z\left[log\left(1 - D\left(G\left(z;\theta^{(G)}\right);\theta^{(D)}\right)\right)\right] \tag{3.2}$$

where $\mathbb{E}_X$ and $\mathbb{E}_Z$ denote the mathematical expectation over all real scenarios and the mathematical expectation over all random noise input to the generator, respectively. Thus, these two neural networks should be combined to construct the min-max optimization model:

$$\min_{\theta^{(G)}} \max_{\theta^{(D)}} \mathbb{E}_X\left[log\left(D\left(x;\theta^{(D)}\right)\right)\right] + \mathbb{E}_Z\left[log\left(1 - D\left(G\left(z;\theta^{(G)}\right);\theta^{(D)}\right)\right)\right] \tag{3.3}$$

During training, the two neural networks in the generator and discriminator are optimized simultaneously, while both neural networks achieve the Nash equilibrium. When receiving random noise inputs, the generator can generate rather realistic scenarios that the discriminator cannot distinguish from the real scenarios. Moreover, different scenarios can be generated by changing the random noise input.

### 3.1.2 Time-Series GAN Model

A good GAN model for time-series data should preserve the temporal dynamics, and the generated scenarios can respect the original distribution and relationship between variables across the whole period. Yoon [70] proposed a time series generative adversarial network (TSGAN) framework for synthesizing realistic scenarios, which integrates the versatility of unsupervised training with the control of supervised training. Unlike the basic GAN model, the TSGAN model consists of four neural network components: an embedding function, a recovery function, a generator, and a discriminator. The structure of the TSGAN model is shown in Fig. 3.1.

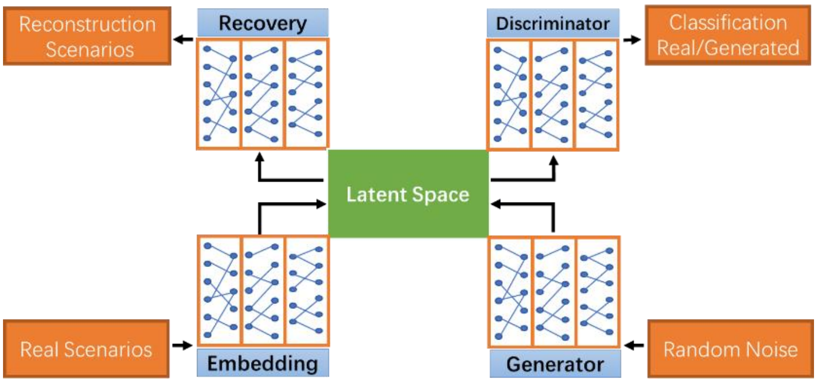



Fig. 3.1. Structure of the TSGAN model

Compared to the basic GAN, this model has two additional sections: the embedding and recovery sections. These two sections provide the mapping between the feature and latent space, helping GAN learn the underlying temporal dynamics of the training scenarios via low-dimensional representations. Denote the embedding neural network and the recovery neural network as $e(\cdot)$ and $R(\cdot)$, respectively. The reconstructed real date input, $\widetilde{x_t}$, and the output of the embedding network, $h_t$, at time $t$ can be obtained respectively as the following:

$$h_t = e\left(h_{t-1}, x_t; \theta^{(e)}\right) \tag{3.4}$$

$$\widetilde{x_t} = R\left(e\left(h_{t-1}, x_t; \theta^{(e)}\right); \theta^{(R)}\right) \tag{3.5}$$

where $x_t$ and $h_{t-1}$ are the inputs of the embedding neural network at time $t$, $h_t$ denotes the output of the embedding neural network at time $t$; $\theta^{(e)}$ and $\theta^{(R)}$ are the weights of the embedding and recovery neural networks, respectively. As the embedding and recovery neural networks should enable accurate reconstruction of the original data from latent space, the first objective is the reconstruction loss expressed as:

$$L_R = \mathbb{E}_X\left[\sum_t \left(\left\|R\left(e\left(h_{t-1}, x_t; \theta^{(e)}\right); \theta^{(R)}\right) - x_t\right\|_2\right)\right] \tag{3.6}$$

where $\|\cdot\|_2$ denotes the Euclidean norm, and $\theta^{(e)}$ and $\theta^{(R)}$ denote the weights of the two neural networks in embedding and recovery functions, respectively. In TSGAN, the training period can be divided into two modes: the open-loop and closed-loop modes. In the open-loop mode, the generator and the discriminator work in the same way as the basic GAN. The second objective is the unsupervised loss, which is essentially the same as $L_D$ in the basic GAN, as the following:

$$L_U = \mathbb{E}_X\left[\sum_t\left(log\left(D\left(e\left(h_{t-1}, x_t; \theta^{(e)}\right); \theta^{(D)}\right)\right)\right)\right] + \mathbb{E}_Z\left[\sum_t\left(log\left(1 - D\left(G\left(h_{t-1}, z_t; \theta^{(G)}; \theta^{(D)}\right)\right)\right)\right)\right] \quad (3.7)$$

To achieve the distribution of generated scenarios more efficiently, in the closed-loop mode, the generator receives the embedding network output to generate the next latent vector. The third objective is the supervised loss expressed as the following:

$$L_S = \mathbb{E}_{X,Z}\left[\sum_t\left(\left\|e\left(h_{t-1}, x_t; \theta^{(e)}\right) - G\left(h_{t-1}, z_t; \theta^{(G)}\right)\right\|_2\right]\right. \quad (3.8)$$

### 3.1.3 Proposed Conditional TSGAN Model

Conventional electricity price forecasting using only historical electricity price data is generally challenging as electricity prices are related to multiple factors. Therefore, modern electricity price forecasting methods consider a wide range of conditions. In this chapter, the conditions related to electricity prices are added to the inputs of the generator and embedding networks. Combining the TSGAN and conditions, we can obtain a new model known as CTSGAN. Fig. 3.2 illustrates the structure of CTSGAN.

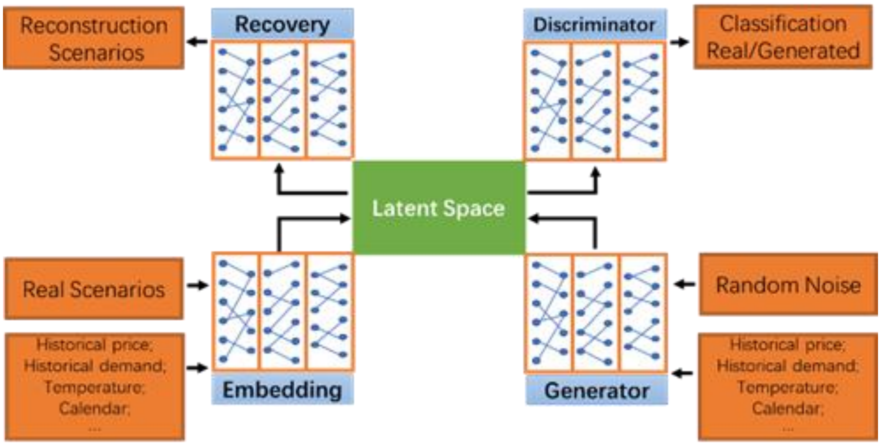



Fig. 3.2. Structure of the CTSGAN model

As shown, the inputs of generator networks contain not only random noise but also price-related conditions **C** . On the other hand, the embedding network maps the real scenarios and conditions to the latent space. Denoting the embedding and recovery neural networks as $e(\cdot)$ and $R(\cdot)$, respectively, one obtains:

$$h_t = e(h_{t-1}, x_t|\mathbf{C}; \theta^{(e)}) \tag{3.9}$$

$$\widetilde{x_t} = R\left(e\left(h_{t-1}, x_t\middle|\boldsymbol{C}; \theta^{(e)}\right); \theta^{(R)}\right) \tag{3.10}$$

The loss functions of reconstruction, including the embedding and recovery neural networks, can be presented as the following:

$$L_R = \mathbb{E}_X\left[\sum_t\left(\left\|R\left(e\left(h_{t-1}, x_t\middle|\mathbf{C}; \theta^{(e)}\right); \theta^{(R)}\right) - x_t|\mathbf{C}\right\|_2\right)\right] \tag{3.11}$$

The supervised and unsupervised losses in CTSGAN are:

$$L_S = \mathbb{E}_{X,Z}\left[\sum_t\left(\left\|e\left(h_{t-1}, x_t\middle|\mathbf{C}; \theta^{(e)}\right) - G\left(h_{t-1}, z_t\middle|\mathbf{C}; \theta^{(G)}\right)\right\|_2\right]\right. \tag{3.12}$$

$$L_U = \mathbb{E}_X\left[\sum_t\left(log\left(D\left(e\left(h_{t-1}, x_t\middle|\mathbf{C}; \theta^{(e)}\right); \theta^{(D)}\right)\right)\right)\right] + \mathbb{E}_Z\left[\sum_t\left(log\left(1 - D\left(G\left(h_{t-1}, z_t\middle|\mathbf{C}; \theta^{(G)}\right); \theta^{(D)}\right)\right)\right)\right] \tag{3.13}$$

Since the gradient of the logarithmic function in unsupervised loss (3.12) is very small in the early training stage, inspired by the Wasserstein GAN [71], the unsupervised loss function in the discriminator introduces the linear function instead of the logarithmic function to accelerate the training process and ensure training process stable. Thus, the unsupervised loss can be simplified as follows:

$$L_U = \mathbb{E}_X\left[\sum_t\left(D\left(e\left(h_{t-1}, x_t\middle|\mathbf{C}; \theta^{(e)}\right); \theta^{(D)}\right)\right)\right] - \mathbb{E}_Z\left[\sum_t\left(D\left(G\left(h_{t-1}, z_t\middle|\mathbf{C}; \theta^{(G)}\right); \theta^{(D)}\right)\right)\right] \tag{3.14}$$

In the optimization period, the reconstruction loss and the supervised loss are trained first, and then the generator and discriminator losses are trained. By combining the above three objectives, the final optimization target can be obtained as follows:

$$\min_{\theta^{(e)}, \theta^{(R)}}(\lambda L_S + L_R) \tag{3.15}$$

$$\min_{\theta^{(G)}} \left( \eta L_S + \max_{\theta^{(D)}} L_U \right) \tag{3.16}$$

where $\lambda \geq 0$ and $\eta \geq 0$ are hyperparameters that help balance both losses in each objective. Compared with the objectives of basic GAN, $L_S$ is a new optimization objective, indicating that the embedding function reduces the dimensions of the adversarial space and facilitates the generator in learning the underlying temporal dynamics.

Since GANs are difficult to keep stable during the training process, two methods are used to stabilize GAN. After each iteration of discriminator training, all the weights of the neural network $\theta^{(D)}$ in the discriminator are limited to a specific range $c$ as a clipping parameter. In addition, since Adam can learn the true distribution at intermediate iterations, but later on suffers from mode collapse and finally degenerates to noise. Therefore, in CTSGAN, we use the RMSProp optimizer to replace the Adam optimizer [72, 73].

Compared with TSGAN, the proposed CTSGAN model can learn the underlying relationship between electricity prices and price-related conditions, making the generated scenarios relatively realistic. The pseudocode of the training algorithm of CTSGAN for electricity price prediction is summarized in Algorithm 3.1.

**Algorithm 3.1** Scenarios Generation with CTSGAN

**Input:** Batch size $m$, number of iterations for each training period $n$, learning rate $\beta$, clipping parameter $c$

**Output:** $\theta^{(e)}$, $\theta^{(R)}$, $\theta^{(G)}$, $\theta^{(D)}$

*#Update parameter for Embedding and Recovery*

**for** $iteration = 1, 2, \ldots, n$ **do**

  Sample batch from historical data:

  $\{x^{(i)}\}_{i=1}^{m} \sim P_X$

  Update Embedding and Recovery nets using gradient descent:

  $g_{\theta^{(e)},\theta^{(R)}} \leftarrow \nabla_{\theta^{(e)},\theta^{(R)}} \left[\frac{1}{m}\sum_{t=1}^{m}\sqrt{\left(R(e(x^{(i)}|\mathbf{C})) - x^{(i)}|\mathbf{C}\right)^2}\right]$

  $\theta^{(e)} \leftarrow \theta^{(e)} - \beta \cdot RMSProp(\theta^{(e)}, g_{\theta^{(e)}})$

  $\theta^{(R)} \leftarrow \theta^{(R)} - \beta \cdot RMSProp(\theta^{(R)}, g_{\theta^{(R)}})$

**end for**

*#Update parameter for Generator only*

**for** $iteration = 1, 2, \ldots, n$ **do**

  Sample batch from historical data:

  $\{x^{(i)}\}_{i=1}^{m} \sim P_X$

  Sample batch from random noise ( Gaussian or Uniform):

  $\{z^{(i)}\}_{i=1}^{m} \sim P_Z$

  Update Generator net using gradient descent:

  $g_{\theta^{(G)}} \leftarrow \nabla_{\theta^{(G)}} \left[\frac{1}{m}\sum_{t=1}^{m}\sqrt{\left(e(x^{(i)}|\mathbf{C}) - G(e(z^{(i)}|\mathbf{C}))\right)^2}\right]$

  $\theta^{(G)} \leftarrow \theta^{(G)} - \beta \cdot RMSProp(\theta^{(G)}, g_{\theta^{(G)}})$

**end for**

*#Joint Training*

**for** $iteration = 1, 2, \ldots, n$ **do**

  *#Update parameter for Generator only (twice more than Discriminator)*

  **for** $iterationGen = 1, 2$ **do**

    Sample batch from historical data:

    $\{x^{(i)}\}_{i=1}^{m} \sim P_X$

    Sample batch from random noise ( Gaussian or Uniform):

    $\{z^{(i)}\}_{i=1}^{m} \sim P_Z$

    Update Generator net using gradient descent:

    $g_{\theta^{(G)}} \leftarrow \nabla_{\theta^{(G)}} \left[\frac{1}{m}\sum_{t=1}^{m}\left(\lambda\sqrt{\left(e(x^{(i)}|\mathbf{C}) - G(e(z^{(i)}|\mathbf{C}))\right)^2} + \left(D(e(x^{(i)}|\mathbf{C})) - D(G(e(z^{(i)}|\mathbf{C}))\right)\right)\right]$

    $\theta^{(G)} \leftarrow \theta^{(G)} - \beta \cdot RMSProp(\theta^{(G)}, g_{\theta^{(G)}})$

    Update Embedding net using gradient descent:

    $g_{\theta^{(e)}} \leftarrow \nabla_{\theta^{(e)}} \left[\frac{1}{m}\sum_{t=1}^{m}\left(\eta\sqrt{\left(e(x^{(i)}|\mathbf{C}) - G(e(z^{(i)}|\mathbf{C}))\right)^2} + \sqrt{\left(R(e(x^{(i)}|\mathbf{C})) - x^{(i)}|\mathbf{C}\right)^2}\right)\right]$

    $\theta^{(e)} \leftarrow \theta^{(e)} - \beta \cdot RMSProp(\theta^{(e)}, g_{\theta^{(e)}})$

  **end for**

  *#Update parameter for Discriminator only*

  Sample batch from historical data:

  $\{x^{(i)}\}_{i=1}^{m} \sim P_X$

  Sample batch from random noise ( Gaussian or Uniform):

  $\{z^{(i)}\}_{i=1}^{m} \sim P_Z$

  Update Discriminator nets using gradient descent:

  $g_{\theta^{(D)}} \leftarrow \nabla_{\theta^{(D)}} \left[\frac{1}{m}\sum_{t=1}^{m}\left(D(e(h_{t-1}, x_t|\mathbf{C})) - D(G(e(h_{t-1}, z_t|\mathbf{C})))\right)\right]$

  $\theta^{(D)} \leftarrow \theta^{(D)} - \beta \cdot RMSProp(\theta^{(D)}, g_{\theta^{(D)}})$

  $\theta^{(D)} \leftarrow clip(\theta^{(D)}, -c, c)$

**end for** 1

## 3.2 CTSGAN-Based Point Prediction Method

Fig. 3.3 illustrates a flowchart to predict electricity prices based on the proposed CTSGAN model. The prediction steps and implementation details are discussed as follows.

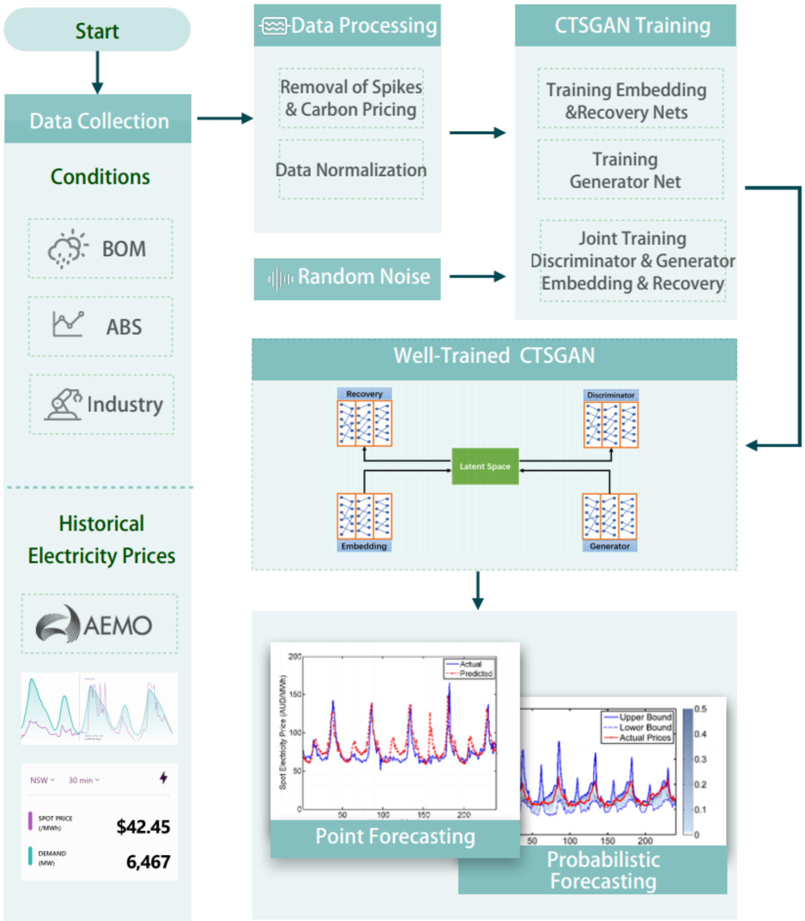



Fig. 3.3. Flowchart of price prediction based on CTSGAN model

**Step 1. Data Collection**

The dataset of half-hourly electricity spot prices in NSW from 15 September 2001 to 31 August 2020 is downloaded from the publicly available Australian Energy Market Operator (AEMO) website https://aemo.com.au. The electricity price data for each month of the four calendar years, 2012, 2014, 2016, and 2018, and five days in each of the four seasons of winter (28 July to 1 August), spring (27 October to 31 October), summer (28 January to 1 February) and autumn (28 April to 2 May) in two financial years of 2018-2019 and 2019-2020 are selected as the test sets. The remaining data is set as the training data.

In addition to electricity prices, the remaining electricity market-related conditions, like half-hourly electricity demand, are also obtained from the AEMO dataset. Electricity consumptions of different industrial components is from the annual report of AEMO. Besides, other conditions (such as weather, temperature, and economic index) are collected from the open-access websites of the BOM and ABS. The selection method of optimal conditions is described in the next section.

**Step 2. Data Pre-processing**

**Removal of positive and negative spikes:** The total historical electricity price dataset contains 332448 half-hourly electricity prices of 6926 days. Since power failures, transmission line maintenance, and extreme weather can result in electricity price spikes and affect the accuracy of forecasting models; it is necessary to limit the electricity prices to a certain range [27]. In this chapter, we set this range as [0,450] and the electricity price as A$450/MWh (symbol A$ represents the Australian Dollar) when it exceeds A$450/MWh and A$0/MWh when it is below A$0/MWh (negative electricity price is allowed in Australian Electricity Market). Counting through the dataset, we found 580 times of electricity prices over A$450/MWh and 61 times of electricity prices below A$0/MWh during this period, representing only 0.19% of the total data and thus having a limited impact on the forecasting model [74].

**Removal of carbon pricing:** Australia's carbon pricing scheme came into effect on 1 July 2012 and was repealed on 1 July 2014. Though this policy factor is not an input in the forecasting model, it impacts directly on the power generation from fossil fuels and renewable energy resources and thus significantly affects the electricity prices. To guarantee that the prediction model is not influenced by carbon pricing, the carbon tax should be removed from the dataset. The carbon pricing was A$23/t in the financial year of 2012-2013, and A$24/t in the financial year of 2013-2014. Considering that the carbon emissions factor in NSW is 0.9 t/MWh [75], one obtains that the carbon pricing included in the electricity prices in NSW is A$20.7/MWh in 2012-2013 and A$21.6/MWh in 2013-2014. This part should be removed during the training process and inserted back into the forecasts for the year of carbon price impact.

**Data normalization:** To better fit the dataset to the forecasting model, the data should be normalized to the range of $[0,1]$ with a min-max normalization method after removing the positive and negative spike prices and carbon pricing. The min-max normalization method is the linear transformation defined as the following:

$$p' = (p - p_{min})/(p_{max} - p_{min}) \tag{3.17}$$

where $p_{min}$ and $p_{max}$ are the maximum and minimum prices in the whole dataset of historical electricity prices, and $p$ and $p'$ the original and normalized prices, respectively.

**Step 3. Optimizing CTSGAN Weights**

Before using CTSGAN to forecast the electricity day-ahead price, the weights of the embedding and recovery networks in CTSGAN should be optimized. The inputs of the CTSGAN are the pre-processing data of actual electricity prices, random noise, and conditions. The dimensionality of the random noise varies according to the forecasting objectives, with a smaller value chosen for point forecasting and a large value for probabilistic forecasting. The training processes for the point and probabilistic predictions are independent of each other.

As shown in Algorithm 3.1, the training process can be divided into three stages. In the first stage, the optimization objectives are the weights of the embedding and recovery networks, which enable accurate reconstructions of the original electricity prices. In the second stage, the optimization objective is the weights of the generator. The generator receives the mapped data of the actual electricity price computed by the embedding network and generates the next latent vector. The final stage is the joint training stage, in which the optimization objectives are the weights of the generator and discriminator. The generator is trained twice as many times as the discriminator is trained [76] to ensure the convergence of CTSGAN. When the maximum number of iterations is reached, the weights of the embedding and recovery networks, generator and discriminator have been optimized.

## 3.3 CTSGAN-Based Interval Prediction Method

The CTSGAN model in the last section can generate a large number of scenarios with authenticity and diversity. By directly combining these scenarios, the PIs can be obtained. However, the CP of the PIs depends only on the generative model and cannot be adjusted by specifying the desired NCP. Although combining PIs can ensure stability, it cannot meet the requirements for sharpness. Therefore, there are difficulties in practical applications. This section proposes a high-quality TSM-based PI construction method with a given NCP, including three parts: 1) TSM and training process, and 2) PI construction with an optimal threshold.

### 3.3.1 Threshold Select Machine and Training

The diversity of CTSGAN-generated scenarios is closely related to the diversity of random noise input. Generally, the random noise inputs $z$ follow a standard Gaussian distribution as $z \sim \mathcal{N}(0,1)$, which results in rich diversity in the generated scenarios. Lee et al. [77] verified that the truncated threshold of random noise could be used to adjust the diversity, which directly affects the quality of interval in this chapter. Therefore, we use a truncated threshold $\tau$ (as shown in Fig. 3.4) to control the diversity of the input noise. Suppose $\pm 0.67$ is selected as $\tau$. A relatively concentrated 50% of the random noise $z$ can be used as input, such that the output scenarios are also relatively concentrated and narrow. When ±1.96 is selected as $\tau$, and this range contains 95% of the training $z$, which can be applied to predict more diverse scenarios, which can be combined wide intervals. However, as [77] did not provide a method for optimizing the truncated threshold $\tau$. To address this issue, we propose a TSM to truncate the random noise input and adjust the reliability and sharpness of PIs. In addition, we introduce the weather factors to adjust the quality of PIs.

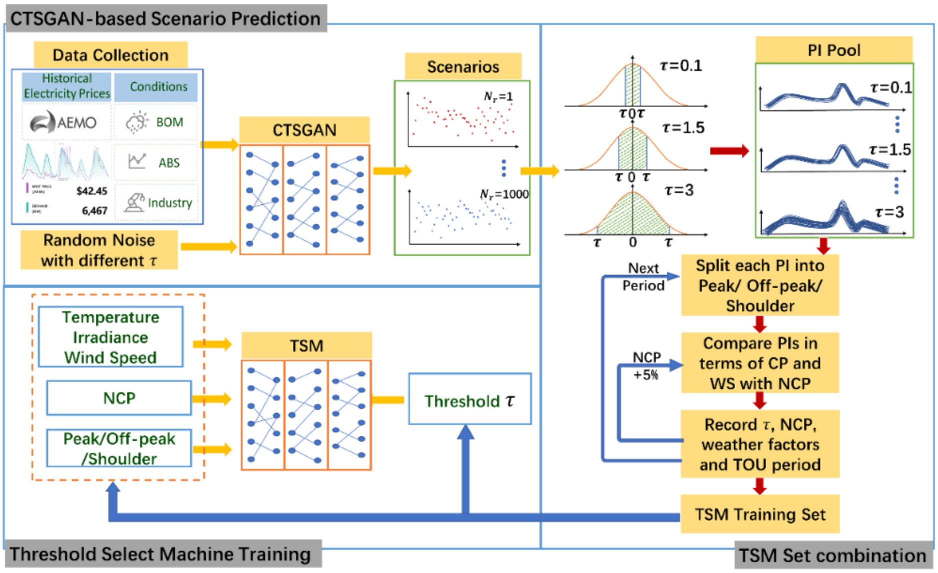



Fig. 3.4 Threshold select machine training steps

Fig. 3.4 illustrates the steps for training TSM, which consists of three steps. The first step is to generate scenarios with different diversity by modifying the threshold value $\tau$ using the well-trained CTSGAN model. Take the day-ahead interval prediction as an example. As shown in Fig. 3.4, CTSGAN generates the next 48 predictions of 30-mins as one scenario with $\tau = 0.1$, and then repeats the process $N_R$ times. We can obtain $N_R$ scenarios, which can be combined into one interval. After that, the TSM training set combination step is shown in Fig. 3.4, bottom right subfigure. By adjusting $\tau$ from 0.1 to 3.0, we can obtain 30 different PIs for each day with rolling predictions and split each PI into three parts in terms of time use period: peak, off-peak, and shoulder. For the PIs of each period, we introduce CP and the WS to evaluate the reliability and sharpness of PIs. An NCP should be set, typically from 70% to 95%, with an interval of 5%. Calculate the CP of all PIs; take out the PIs that satisfy the NCP; then compare the WS and select the best PI with the smallest WS. The optimal $\tau$ can also be obtained. The corresponding weather factors, NCP and use period should be recorded in the TSM training set. Adjusting the NCP and repeating the optimal $\tau$ selection process, we can obtain the optimal $\tau$ according to 6 different NCP from 70% to 95% with a step of 5%. Algorithm 3.2 presents the pseudocode of the training set combination.

**Algorithm 3.2**: TSM training set combination

**For** day $N_d = N_d^1, N_d^2, N_d^3, \cdots, N_d^n$
    *#PI Pools generation*
    **For** threshold $\tau$ **from** 0.1 **to** 3.0 **with step of** 0.1:
        **For** repeat time $N_r$ **from** 1 **to** $N_R$:
            Predict one day-ahead scenario by CTSGAN with $\tau$
        **End**
        Combine one interval with $N_R$ scenarios
    Split all PIs into 3 parts in terms of peak, off-peak and shoulder, and add them into 3 PI pools: $P_P$, $P_o$, and $P_s$
    **End**
*#TSM training set combination*
**For** PI pool $P = P_P, P_o, P_s$:
    **For** $NCP$ **from** 70% **to** 95% **with step of** 5%:
        Calculate CP and WS for each PI, select the best PI with min WS and CP>NCP
        Obtain the optimal $\tau$ of the best PI for NCP and $P$
        Obtain the values $V^1$ and variances $V^2$ of wind speed, irradiance, and temperature at the time corresponding to the best PI, as $V_W^1, V_W^2, V_I^1, V_I^2, V_T^1, V_T^2$.
        Save the $V_W^1, V_W^2, V_I^1, V_I^2, V_T^1, V_T^2, NCP$, $P$, and $\tau$ into the training set
    **End**
**End**
**End**

Finally, the TSM training process is given in Fig. 3.4, the left bottom subfigure. We can use the values and variances of the weather factors, NCP, and time-of-use (TOU) periods in the training set as inputs and the optimal $\tau$ as output to train the TSM. To simplify the system, an LSTM [78] is introduced as the TSM.

### 3.3.2 PIs Construction

Fig. 3.5 presents the PI construction framework. The first step is to select the truncated threshold $\tau$ of random noise by the TSM with weather factors, NCP, and time of use period as input. After that, using the random noise with $\tau$ and conditions as input of the prediction model, the CTSGAN model can generate a large number of scenarios for the next 48 30-minute points. In this chapter, 1,000 different scenarios are generated for each point and then combined into one interval for day-ahead prediction. Repeat the TSM and prediction step three times, each time modifying only the time of use period in the TSM input; three PIs for peak, off-peak, and shoulder time can be obtained. Finally, by combining the three PIs, the day-ahead price prediction with specified NCP can be constructed.

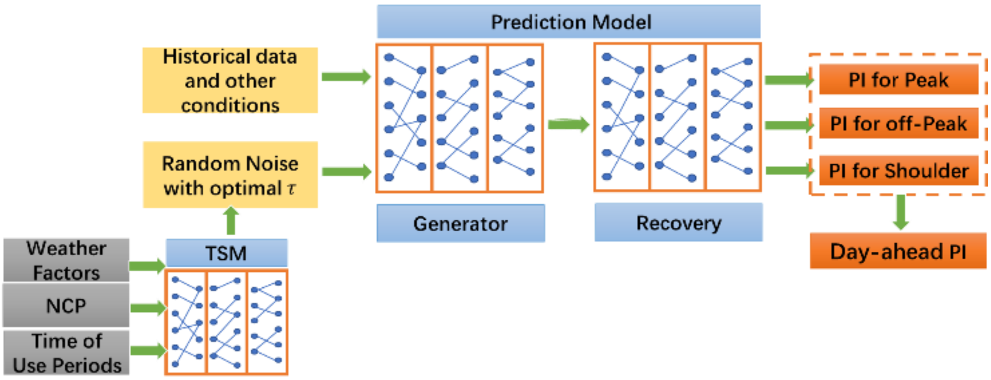



Fig. 3.5 PI construction framework

## 3.4 Implementation Details

### 3.4.1 Selection and Parameter Setting of Core NNs of CTSGAN

The choice of core NNs in GAN is also a significant issue. Any NN can be used for GAN, even like the back propagation neural network (BPNN). However, when handling large datasets, it is best to employ DL models to guarantee that the training process does not collapse. For sequence forecasting, such as the electricity price forecasting in this chapter, RNNs, including LSTM and GRU, are usually chosen. For model simplicity, LSTM is chosen as the core NNs of the embedding function, recovery function, generator, and discriminator in CTSGAN. The number of hidden layer units in LSTM is set to 100, and the number of training iterations is 10,000.

The parameters for the proposed CTSGAN are set as follows: the batch size is 7; the number for each training iteration is 10,000; the clipping parameter is 0.5; the learning rate is 0.02; the dimensionality of latent space is 100. The parameters for the TSGAN are set to be the same as those of the proposed CTSGAN.

### 3.4.2 Selection of Optimal Conditions

Another preparation task for forecasting day-ahead electricity prices is the selection of optimal conditions (features). Because electricity prices are related to many factors, by referring to the literature and relevant AEMO forecasting manual, data on various types of factors are downloaded from the public data websites of the Australian

Government, including the Bureau of Statistics, the Bureau of Meteorology (BOM), and the Department of Industry, Science, Energy, and Resources, to obtain a dataset of potential conditions, $\boldsymbol{C_p}$ , as the following:

$$\boldsymbol{C_p} = \begin{bmatrix} \boldsymbol{P}_{d-1}, \boldsymbol{P}_{d-2}, \boldsymbol{P}_{d-3}, \boldsymbol{P}_{d-4}, \boldsymbol{P}_{d-7}, \boldsymbol{P}_{d-14}, \boldsymbol{P}_{d-21}, \hat{P}_d^{coal}, \hat{P}_d^{gas}, \\ \hat{\boldsymbol{L}}_d, \boldsymbol{L}_{d-1}, \boldsymbol{L}_{d-2}, \boldsymbol{L}_{d-3}, \boldsymbol{L}_{d-4}, \boldsymbol{L}_{d-7}, \boldsymbol{L}_{d-14}, \boldsymbol{L}_{d-21}, \hat{\boldsymbol{T}}_d^{temp}, \hat{\boldsymbol{T}}_d^{cloud}, \\ \hat{\boldsymbol{T}}_d^{sunshine}, \hat{\boldsymbol{T}}_d^{shortwave}, \hat{\boldsymbol{T}}_d^{windspeed}, \hat{\boldsymbol{T}}_d^{winddirection}, \hat{T}_d^{hdd}, \hat{T}_d^{cdd}, \\ D_d^{week}, D_d^{month}, D_d^{holiday}, M_d^{popu}, M_d^{gdp}, M_d^{cpi}, M_d^{AEG}, M_d^{AREG}, \\ M_d^{NEG}, M_d^{NREG}, M_d^{NECC}, M_d^{NERC}, M_d^{NEIC1}, M_d^{NEIC2}, M_d^{NEIC3}, \\ M_d^{NEIC4}, M_d^{NEIC5}, M_d^{NEIC6} \end{bmatrix} \tag{3.18}$$

where $\boldsymbol{P}_{d-t}$ denotes the daily historical electricity price of $t$ days before the predicted day, and the length of daily electricity price is 48; $\hat{P}_d^{coal}$ and $\hat{P}_d^{gas}$ are the daily average coal spot price at the Newcastle port and natural gas price in AEMO, respectively; $\hat{\boldsymbol{L}}_d$ denotes the estimated load of the predicted day and $\boldsymbol{L}_{d-t}$ the daily historical load profiles of the $t$ day before the predicted day; $\hat{\boldsymbol{T}}_d^{temp}$ denotes the estimated temperatures of the predicted day, and the length of $\hat{\boldsymbol{T}}_d^{temp}$ is 24, with which the heating degree day $\hat{T}_d^{hdd}$ and cooling degree day $\hat{T}_d^{cdd}$ can be calculated [79]; $\hat{\boldsymbol{T}}_d^{cloud}, \hat{\boldsymbol{T}}_d^{sunshine}, \hat{\boldsymbol{T}}_d^{shortwave}, \hat{\boldsymbol{T}}_d^{windspeed}$ and $\hat{\boldsymbol{T}}_d^{winddirection}$ are the estimated weather conditions, including the duration of cloud, duration of sunshine, short wave radiation level, wind speed, and wind direction, respectively; $D_d^{week}$ and $D_d^{month}$ are the day types (day of the week and month of the year), and $D_d^{holiday}$ represents whether the predicted day is a holiday; $M_d^{popu}$, $M_d^{gdp}$ and $M_d^{cpi}$ denote the latest population, gross domestic product (GDP) and consumer price index (CPI) of the NSW, respectively; $M_d^{AEG}, M_d^{AREG}, M_d^{NEG}$ and $M_d^{NREG}$ represent the total electricity generation and the renewable electricity generation in Australia and the state of NSW; $M_d^{NECC}$, $M_d^{NERC}$ and $M_d^{NERCi}$ denote the commercial, residential, and industrial components of the electricity consumption, respectively, and $i$ in $M_d^{NERCi}$ represents different industrial sectors; $M_d^{NERC1}$ represents the electricity consumption of agriculture, forestry, and fishing, $M_d^{NERC2}$ the electricity consumption of mining, $M_d^{NERC3}$ the electricity consumption of manufacturing, $M_d^{NERC4}$ the electricity consumption of services of electricity, gas, water, and waste, $M_d^{NERC5}$ the electricity consumption of construction, and $M_d^{NERC6}$ the electricity consumption of transport, postal and warehousing. Although some data, such as the half-hourly electricity interregional transmission and electricity savings by policies, are unavailable, as their impacts on forecasting are limited, they are not discussed in this work.

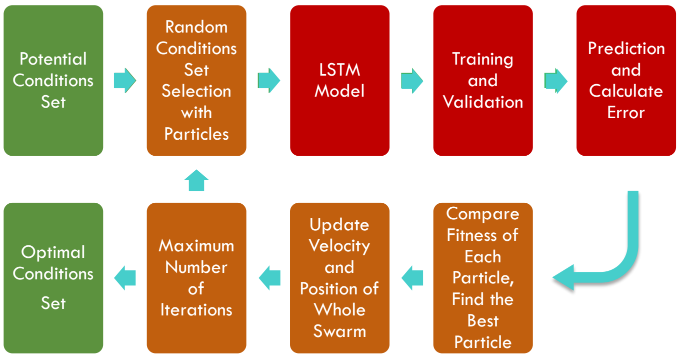



Fig. 3.6 Structure of proposed model for selecting optimal conditions

With the potential set $\boldsymbol{C_p}$ and inspired by [80], we propose a PSO-LSTM condition selection model, as shown in Fig. 3.6. The PSO algorithm is applied to choose a condition set randomly from the potential set $\boldsymbol{C_p}$ for training the LSTM model. The well-trained LSTM model can then be used to forecast the day-ahead electricity prices. The PSO algorithm can be used to optimize the random condition set gradually by minimizing the error between the forecasting results and the actual prices. After 10,000 iterations, the optimal condition set of electricity prices, $\boldsymbol{C}$, can be obtained as the following:

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{P}_{d-1}, \boldsymbol{P}_{d-2}, \boldsymbol{P}_{d-7}, \boldsymbol{P}_{d-14}, \hat{P}_d^{coal}, \hat{P}_d^{gas}, \\ \hat{\boldsymbol{L}}_d, \boldsymbol{L}_{d-1}, \hat{\boldsymbol{T}}_d^{temp}, \hat{T}_d^{hdd}, \hat{T}_d^{cdd}, D_d^{week}, \\ D_d^{holiday}, M_d^{cpi}, M_d^{NREG}, M_d^{NERC}, M_d^{NEIC3} \end{bmatrix} \tag{3.19}$$

From the final optimal condition set, it is evident that the day-ahead price is related to several factors, including the historical price data, price of coal and gas, predicted and historical loads, temperate and heating or cooling degree day, CPI, renewable electricity generation in the NSW, and electricity consumption of residential and manufacturing in the NSW. The total dimension of the input of the CTSGAN is

322=48*6+24+10.

### 3.4.3 Weather Input of TSM

The weather factor data is obtained from the Australian BOM website. Note that all the weather factors used in the training models are real historical data. In contrast, the weather factors used in the prediction models are synthetic data by adding an error term to the actual weather data. The mean absolute percentage error (MAPE) for temperature, wind speed, and irradiance are set at 10%, 15%, and 20%, respectively. The input of TSM consists of $W_{temp}^{ave}$, $W_{temp}^{var}$, $W_{wind}^{ave}$, $W_{wind}^{var}$, $W_{irra}^{ave}$, $W_{irra}^{var}$, $\Gamma_{NCP}$, $\Gamma_{TOU}$. The first six items represent the mean and variance of temperature, wind speed, and irradiance. Each is a 1*1-dimensional data and then is normalized over the training days. $\Gamma_{NCP}$ represents the NCP value, while $\Gamma_{TOU}$ represents the TOU, with 0, 0.5, and 1, respectively representing off-peak, shoulder, and peak periods. The TSM's input has a dimension of 8*1.

### 3.4.4 Evaluation Indicators

Different error measures are employed as the criteria to evaluate the performance of day-ahead electricity price point forecasting models. The most used indicators are MAE, the MAPE, and RMSE, defined as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - y_i| \tag{3.20}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}|(Y_i - y_i)/Y_i| \tag{3.21}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - y_i)^2} \tag{3.22}$$

where $Y_i$ and $y_i$ denote the actual and predicted electricity prices at the time $i$, respectively, and $n$ is the number of the prices in a day. In the Australian electricity market, the settlement process operates on a 30-minute basis, and therefore, $n$ is set to 48. Generally, a lower error measure indicates a better forecasting performance. As a measure of forecasting quality, the U2 proposed by Theil [81] is calculated by

$$U2 = \sqrt{\sum_{i=1}^{n}((Y_{i+1} - y_{i+1})/Y_i)^2} \Big/ \sqrt{\sum_{i=1}^{n}((Y_{i+1} - Y_{i+1})/Y_i)^2} \tag{3.23}$$

$U2$ is employed to compare the forecast quality between the proposed and naïve forecast results. The naïve forecast is a forecasting technique in which the actual electricity prices of the last moment are used as the forecast results of the current moment without adjusting them or attempting to establish causal factors. If $U2$=1, there is no difference between a naïve forecast and the proposed method; if $U2$<1, the proposed method is better than a naïve forecast; and if $U2$>1, the proposed method is worse than a naïve forecast.

Reliability is the most indispensable indicator of PIs [82]. Unconditional coverage is the simplest and the most common approach for assessing the reliability of PIs [35]. Averaging over the unconditional coverage, Khosravi et al. [36, 83] introduced the CP for PIs, and the definition of CP is given by:

$$CP = \frac{1}{T}\sum_{t=1}^{T} c_t, \quad \begin{cases} c_t = 1, if\ \theta_t \in [L_t, U_t] \\ c_t = 0, if\ \theta_t \notin [L_t, U_t] \end{cases} \tag{3.24}$$

where $T$ represents the sample size, $\theta_t$ the *t-th* sample; $U_t$ and $L_t$ denote the upper and lower bounds of forecasting intervals of electricity prices, respectively; $c_t$ is an indicator that is equal to 1 if $[L_t, U_t]$ covers $\theta_t$, and to 0 otherwise.

When faced with multiple PIs with similarly accurate CP, we aim to choose the narrowest PIs. WS is introduced to evaluate a central $(1-\alpha) \times 100\%$ interval by joint assessing reliability and interval width as the following:

$$WS(\theta_t) = \begin{cases} \delta_t + \frac{2}{\alpha}(L_t - \theta_t), \theta_t \le L_t \\ \quad \delta_t \quad , L_t < \theta_t < U_t \\ \delta_t + \frac{2}{\alpha}(\theta_t - U_t), \theta_t \ge U_t \end{cases} \tag{3.25}$$

where $\delta_t = U_t - L_t$ is the interval width at the *t-th* time. A lower WS infers a superior probabilistic performance regarding a certain interval.

### 3.4.5 Simulation Environment

The simulation for our work is implemented in Python 3.7 with the Pytorch packages on VSCode. All run in a PC with an AMD RYZEN 9 3950X CPU and an NVIDIA RTX 3080 GPU.

## 3.5 Case Studies

### 3.5.1 Different Setting Effects in Prediction Model

Properly selecting settings for the CTSGAN model is crucial for electricity price prediction. To accomplish this, we assess the model's performance with different settings This enables us to determine appropriate condition input, hyperparameters, clipping parameters and the dimensionality of the latent space in CTSGAN, and the networks in CTSGAN and TSM, respectively.

#### 3.5.1.1 Selection of Hyperparameters in Objectives of CTSGAN

Two hyperparameters $\lambda$ and $\eta$ can help balance both losses in objectives of CTSGAN as (3.15) and (3.16). In this section, we conducted experiments in terms of $\lambda$ and $\eta$, to compare the RMSE of predicted scenarios to determine the optimal hyperparameters $\lambda$ and $\eta$. Fig. 3.7 presents the results of RMSE under different hyperparameters.

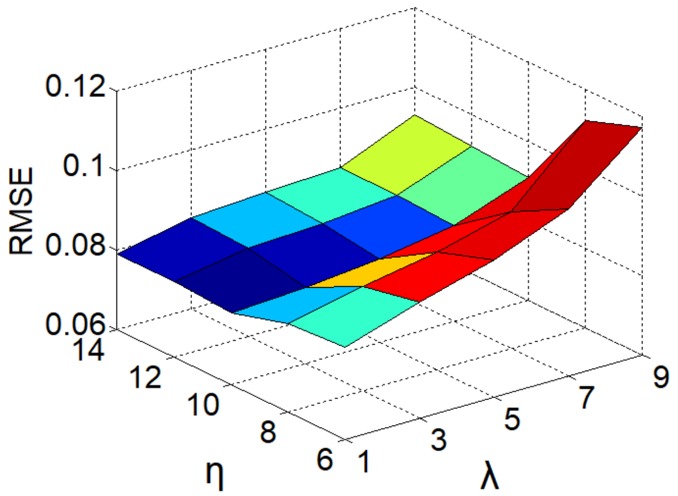



Fig. 3.7 RMSE of predicted scenarios in terms of different $\lambda$ and $\eta$

The results show the RMSE for the hyperparameters ranging from 1 to 9 for $\lambda$ and 6 to 14 for $\eta$. It can be observed that when $\lambda$ is between 1 to 3 and $\eta$ is between 10 to 14, the RMSE are nearly identical, at around 0.078. However, a notable increase in RMSE is observed when $\eta$ is set to 6. As proposed in [76], TSGAN is not sensitive to these two hyperparameters. Similarly, we find that CTSGAN is also not sensitive when $\lambda$ is between 1 to 3 and $\eta$ is between 10 to 14. Therefore, we select the hyperparameters with $\lambda$ set to 1 and $\eta$ set to 10 for our model.

#### 3.5.1.2 Selection of Clipping Parameter in CTSGAN

The choice of clipping parameter in CTGAN is an important decision that can affect the stability and performance of the model. One approach to selecting the clipping parameter value is to experiment with a range of values and observe the effects. Fig. 3.8 shows the error bars for RMSE in terms of different clipping parameters in the discriminator.

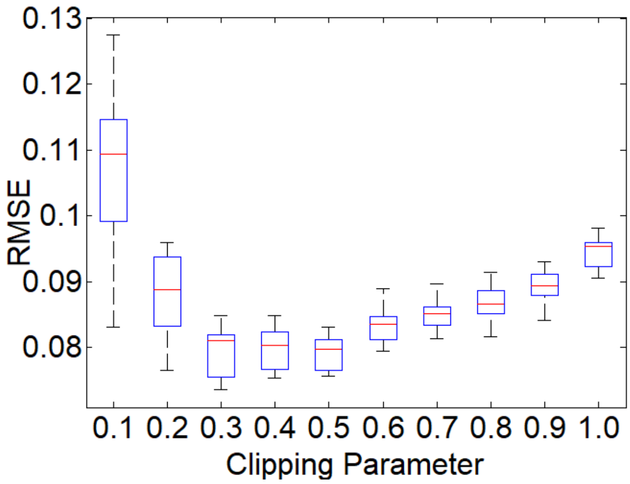



Fig. 3.8 Error bar of RMSE in terms of different clipping parameters

From Fig. 3.8, it can be found that as the clipping parameter exceeds 0.5, the average RMSE starts to increase, indicating that the convergence of CTSGAN begins to slow down, and there is a possibility of gradient vanishing; however, the range of errors does not significantly increase. On the other hand, as the clipping parameter decreases, especially when it is less than or equal to 0.2, the range of errors begins to increase significantly. When the clipping parameter is set to 0.1, the RMSE fails to converge to below 0.1 since the discriminator's weight can only be limited to a very small range. When the clipping parameter is set to 0.4 or 0.5, the range of errors is smaller, and the mean RMSE is also smaller. Although the minimum value at 0.4 is slightly smaller than at 0.5, considering the range of errors, a clipping parameter of 0.5 is chosen in our work.

#### 3.5.1.3 Selection of Dimensionality of Latent Space in CTSGAN

We perform a hyperparameter search, where a range of dimensionality in latent space is tested and the model with the best performance is selected. Fig. 3.9 presents the RMSE and training time.

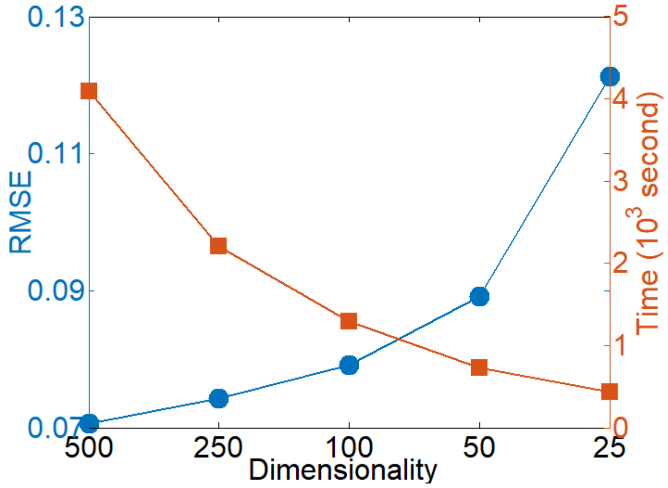



Fig. 3.9  RMSE and training time in terms of dimensionality of the latent space

Fig. 3.9 shows that when the dimensionality is greater than 100, the training time increases significantly from 1093 to 4128 seconds, while the RMSE only decreases from 0.079 to 0.071. On the other hand, when the dimensionality is reduced to 50, although the training time decreases, the prediction accuracy decreases significantly, with an RMSE of 0.089 when the dimension is 50 and 0.127 when it is 25. Therefore, to balance the training time and RMSE, we choose a latent space dimensionality of 100.

#### 3.5.1.4 Selection of Conditions as Input of CTSGAN

We propose a PSO-LSTM condition selection method to choose the optimal condition set as the input of CTSGAN because the core NNs of CTSGAN are LSTMs. In this section, the forecasts of electricity prices for the year 2019 are conducted with the historical electricity price set, all condition set, and the optimal condition set, respectively, to verify the effectiveness of the optimal condition selection method. Fig. 3.10 plots the bar charts of forecasting errors with three different condition sets, and Table 3.1 lists the results of four indicators and improvements.

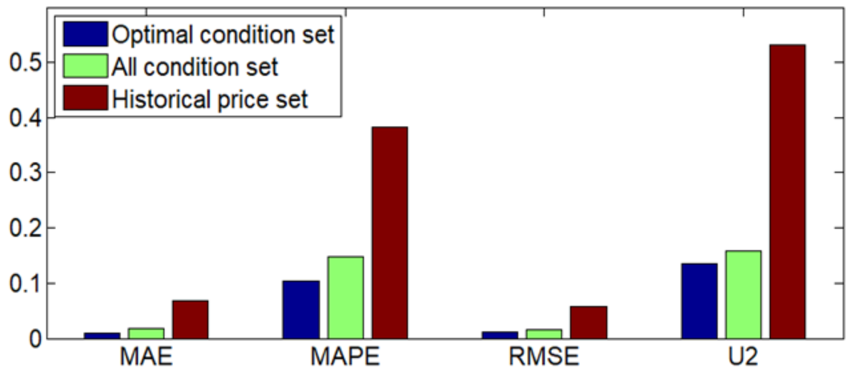



Fig. 3.10 Forecasting errors with optimal condition set and other two different condition sets

In Fig. 3.10, one can find that when the CTSGAN input condition set only contains historical prices, the forecasting errors almost double or triple the errors with all condition sets as input. However, all condition set contains too many irrelevant variables, and DL will learn the features, resulting in poorer forecasting results. The optimal condition set selected by PSO, which is applied as the input of the CTSGAN model, provides the best forecasting performance.

Table 3.1 Comparison of different condition sets as input

| Year | Indicators | Optimal condition set | All condition set | Historical price set |
|---|---|---|---|---|
| 2019 | MAE | **0.0093 (47.7%, 86.4%)** | 0.0178 | 0.0682 |
| | MAPE | **0.1034 (29.8%, 73.1%)** | 0.1473 | 0.3839 |
| | RMSE | **0.0108 (32.1%, 81.2%)** | 0.0159 | 0.0574 |
| | U2 | **0.1362 (13.9%, 74.4%)** | 0.1581 | 0.5312 |

The values in brackets in cells of the optimal condition set column represent the improvement of error by introducing the optimal condition set compared to all condition set and historical price set, elucidating that the PSO-LSTM condition selection method can reduce forecasting errors. Compared with all condition set input, MAE, MAPE, RMSE, and U2 are improved by 47.7%, 29.8%, 32.1%, and 13.9%, respectively. When the condition input is only the historical price set, the CTSGAN model provides the worst prediction error, with MAE of 0.0682, MAPE of 0.3839, RMSE of 0.0774, and U2 of 0.5312, all of which are several times larger than those with the optimal condition set.

#### 3.5.1.5 Selection of Networks in CTSGAN and TSM

The proposed TSM-CTSGAN model involves multiple neural networks, including four networks in CTSGAN and one in TSM. Although any networks can be used, deep networks such as LSTM or gated recurrent unit (GRU) are typically applied. Table 3.2 lists the prediction time and actual CP using LSTM and GRU in CTSGAN and TSM, respectively, when NCP=90%.

Table 3.2 Comparison of CP, WS and time for prediction with different networks in CTSGAN and TSM

| | TSM-CTSGAN | | | |
|---|---|---|---|---|
| | LSTM-LSTM | GRU-LSTM | LSTM-GRU | GRU-GRU |
| CP (%) | 90.14 | 90.16 | 90.28 | 90.20 |
| WS | 7.33 | 7.34 | 13.28 | 13.31 |
| Time (second) | 310.12 | 308.93 | 187.27 | 186.69 |

Table 3.2 demonstrates that utilizing LSTMs in CTSGAN significantly decreases the WS of the prediction results, from 13.28 (LSTM-GRU) to 7.33 (LSTM-LSTM), despite the increase in prediction time, which is predominantly determined by the scenarios generated by CTSGAN. Conversely, using LSTM and GRU in TSM shows a negligible difference of less than 2 seconds. Therefore, four LSTMs are employed in CTSGAN, while TSM adopts one LSTM. All LSTMs are set with the number of hidden layer units of 100, and the number of iterations of 10,000 for each training step.

### 3.5.2 Point Prediction Results and Analysis

#### 3.5.2.1 Comparison with Different Prediction Models

To verify the point forecasting performance of the proposed CTSGAN model, we introduce four other neural networks, TSGAN, LSTM, GRU, and BPNN models, and two linear regression model, auto regressive integrated moving average (ARIMA) and least absolute shrinkage and selection operator (LASSO) [84], for forecasting and comparison. Both the CTSGAN and TSGAN choose the LSTM as the core NN and the same parameter settings of LSTMs as described in Section 3.2.1. The GRU also has the same parameter as the LSTM. The numbers of layers and hidden units in

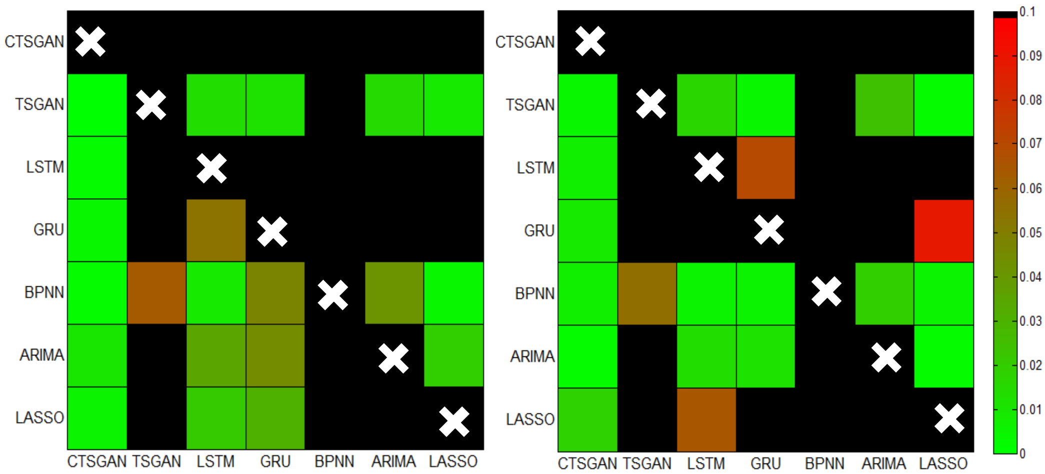
BPNN [85] are set to 3 and 200, respectively, and the number of iterations is set to 50,000. The forecasts of electricity prices for 2012, 2014, 2016, and 2018 are conducted, respectively. Table 3.3 lists the results of four indicators, MAPE and U2 in p.u., and MAE and RMSE in A$/MWh, which can reflect the real-world meaning well. Table 3.3 illustrates the average indicators from 2012 to 2018.

Table 3.3 Comparison of different point forecasting models for four years

| Year | Indicators | CTSGAN | TSGAN | LSTM | GRU | BPNN | ARIMA | LASSO |
|---|---|---|---|---|---|---|---|---|
| 2012 | MAE | **4.82** | 22.32 | 5.63 | 7.02 | 31.77 | 8.69 | 7.69 |
| | RMSE | **6.17** | 24.26 | 6.75 | 8.24 | 33.98 | 8.64 | 8.15 |
| | MAPE | **0.1342** | 0.7353 | 0.1656 | 0.1958 | 1.1626 | 0.2969 | 0.1961 |
| | U2 | **0.1617** | 0.8665 | 0.2067 | 0.2196 | 1.3718 | 0.4333 | 0.2234 |
| 2014 | MAE | **4.37** | 23.85 | 7.88 | 5.54 | 35.91 | 9.32 | 8.10 |
| | RMSE | **5.94** | 25.83 | 9.27 | 6.93 | 37.94 | 9.59 | 7.11 |
| | MAPE | **0.1765** | 0.8827 | 0.3238 | 0.2156 | 1.3038 | 0.5236 | 0.3375 |
| | U2 | **0.1555** | 0.9810 | 0.3108 | 0.1912 | 1.6995 | 0.3238 | 0.2186 |
| 2016 | MAE | **16.20** | 28.04 | 17.19 | 18.77 | 37.62 | 20.84 | 20.61 |
| | RMSE | **20.52** | 33.71 | 22.37 | 23.54 | 43.47 | 28.85 | 23.09 |
| | MAPE | 0.3069 | 0.4865 | **0.2959** | 0.3295 | 0.6681 | 0.4922 | 0.3081 |
| | U2 | 0.1668 | 0.2512 | **0.1555** | 0.1626 | 0.3790 | 0.1937 | 0.1684 |
| 2018 | MAE | **12.78** | 39.92 | 17.82 | 19.98 | 54.77 | 24.12 | 20.84 |
| | RMSE | **16.74** | 43.83 | 22.59 | 25.92 | 58.01 | 37.89 | 23.31 |
| | MAPE | **0.1570** | 0.4793 | 0.2959 | 0.2527 | 0.6597 | 0.3044 | 0.3012 |
| | U2 | **0.1253** | 0.3364 | 0.1960 | 0.2093 | 0.4405 | 0.3139 | 0.2330 |

In Table 3.3, the best or least value for each indicator in each year is marked in bold. As shown, overall, the proposed CTSGAN model gives the best performance in forecasting electricity prices, with a MAE of A$4.82/MWh and a RMSE of A$6.17/MWh, followed by the LSTM model, which also gives satisfactory forecasting performance in 2016 with a MAPE of 0.2959 and a U2 of 0.1555. The GRU model has a slightly inferior forecasting ability to the LSTM model. The predictive ability of the LASSO model is similar to that of the GRU model for the prediction of 2016, with an RMSE of A$23.09/MWh (LASSO) and A$23.54/MWh (GRU), then followed by the ARIMA model. The TSGAN model has poor predictive results due to not considering the conditions, with all indicators larger than those of the GRU, LSTM, LASSO, and ARIMA models. This further demonstrates that it is not accurate to rely on the historical electricity price data alone to forecast electricity prices, and it is necessary to consider the conditions, such as the weather, calendar, and demands. BPNN is the worst forecasting model, worse than the naïve forecasts for 2012 and 2014, with the U2 indicator greater than 1.

However, the values in Table 3.3 can only be used to provide a model ranking, Uniejewski et al. [86] pointed out that the DM [87] test can be introduced to verify whether the values in Table 3.3 are statistically significant. Fig. 3.11 depicts the results of the DM test for MAE and RMSE in 2012. The color bar on the right represents the range of p-values: the closer they are to zero (→greener), the more significant the difference between the forecasts of a set on the X-axis (better) and those on the Y-axis (worse); otherwise, the closer they are to 0.1 (→redder or black), the less statistically significant the difference between the forecasts of models on the X-axis and Y-axis. For example, the second row in both Fig. 3.11 is green except for one black square, which indicates that TSGAN gives poorer performance significantly than all other forecasting models except BPNN.




(a) (b)

Fig. 3.11. Results of DM test of forecast models in terms of (a) MAE and (b) RMSE

Fig. 3.11 clearly shows that the proposed CTSGAN is the best forecast model with a small p-value, and BPNN is the worst model. In addition, one can find that the prediction performance of the CTSGAN is statistically significant, with highly saturated green squares in the column labeled CTSGAN. In terms of both the DM test of RMSE, it is not outperformed significantly by the GRU, LASSO, and LSTM.

To compare the models graphically, Fig. 3.12 plots the rolling forecast (red) and actual (blue) electricity prices for five days in the winter, spring, summer, and autumn seasons for the 2019-2020 financial year by the BPNN, GRU, LSTM, TSGAN, CTSGAN, ARIMA, and LASSO models. As shown in Figs. 3.12 (a)-(d), the BPNN model cannot effectively capture the complex characteristics of electricity prices, and in the four seasons of the financial year 2019-2020, the forecasting results fluctuate between A$20/MWh and A$40/MWh. As shown in Figs. 3.12 (e)-(h), the forecast results of the GRU model are inaccurate for spring and autumn in 2019-2020. The actual autumn daily electricity price usually has two peaks (the blue curve in Fig. 3.12 (h)), while the forecasted result has only one peak (the red curve in Fig. 3.12 (h)). This phenomenon may be attributed to overfitting, where the GRU model learns features that do not belong to each season. The LSTM has also exhibited the same problem of overfitting, e.g., the winter forecasts in Fig. 3.12 (i). The TSGAN model also exhibits poor, unstable forecasting results, as shown in Figs. 3.12 (m)-(p), as it does not include the relevant conditions as the input and relies only on the historical data for forecasting. The forecasting prices given by TSGAN are good for only the days of 28 July 2019 (the first day in Fig. 3.12 (m)), 27 October 2019 (the last day in Fig. 3.12 (n)), 28-29 January 2020 (the first two days in Fig. 3.12 (o)), and 30 April 2020 (the third day in Fig. 3.12 (p)). The ARIMA model provides inaccurate forecasting prices, especially as shown in Fig. 3.12 (v), and the LASSO model forecasts better than the ARIMA model for the winter prices in Fig. 3.12 (y). It is very interesting to note that the GRU, LSTM, ARIMA, and LASSO models all forecast two price peaks incorrectly for the summer prices (Figs. 3.12 (g), (k), (w), (α)). In contrast, CTSGAN can accurately forecast only one electricity price peak during these five days. The proposed CTSGAN model captures temporal variation very well, with outstanding forecasts of electricity prices for all seasons from 2019-2020, with errors only occurring at the spikes, as shown in Figs. 3.12 (q)-(t).

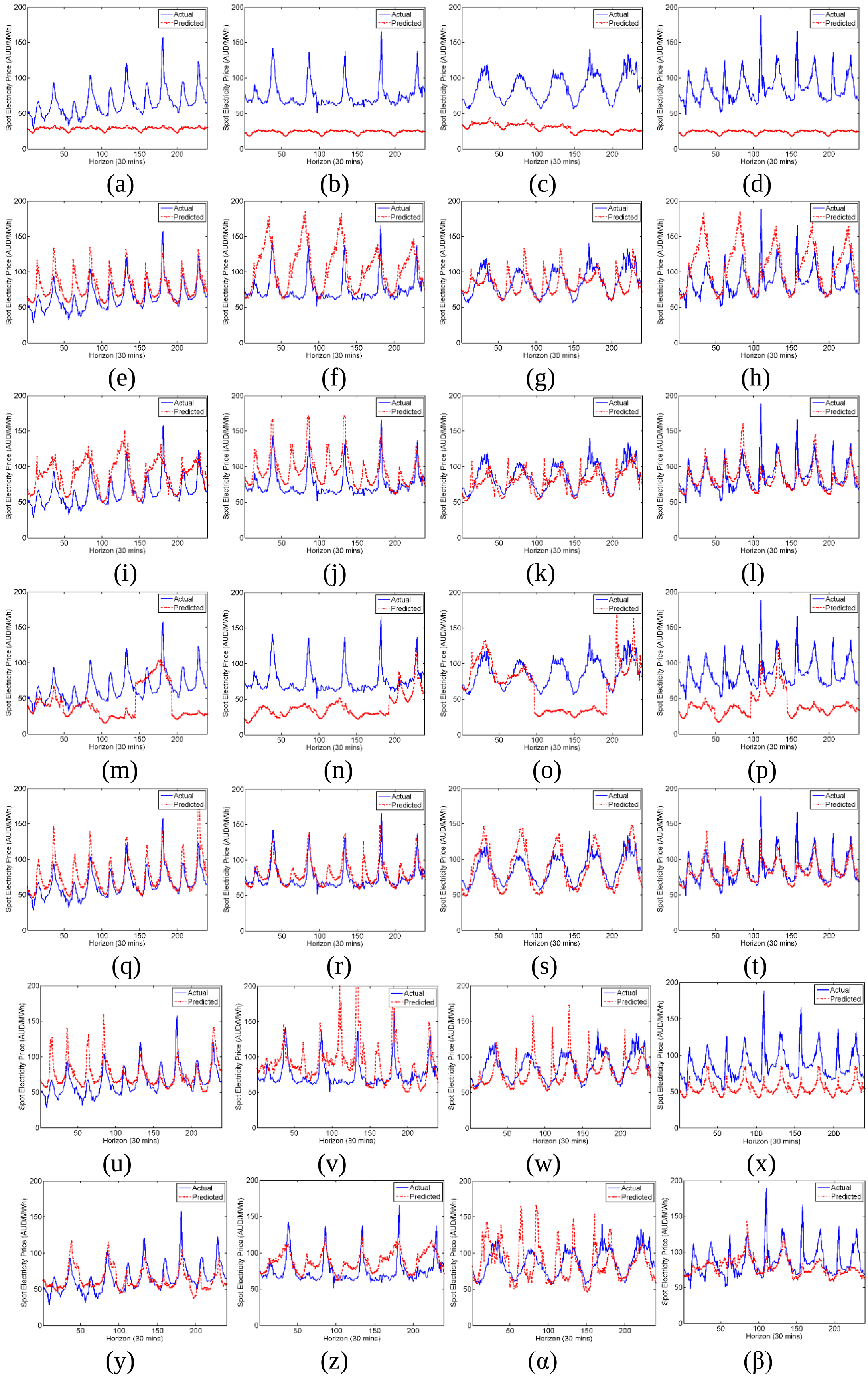



Fig. 3.12. Day-ahead forecasting for five days in four seasons of the financial year 2019-2020 by (a)-(d) BPNN, (e)-(h) GRU, (i)-(l) LSTM, (m)-(p) TSGAN, (q)-(t) CTSGAN models (u)-(x) ARIMA, and (y)-( β) LASSO

Table 3.4 Comparison of CTSGAN and LSTM point forecasting model in 2012

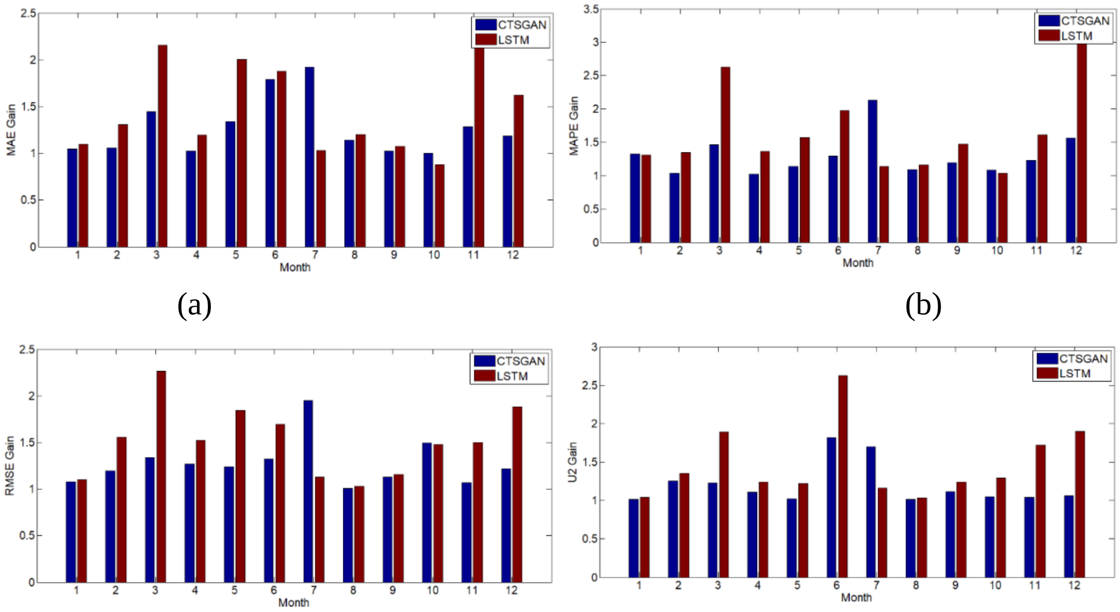
| Month | Indicator | CTSGAN | | | LSTM | | |
|---|---|---|---|---|---|---|---|
| | | Test | Training | Gain | Test | Training | Gain |
| Jan | MAE | **0.0048** | 0.0046 | 1.0445 | 0.0049 | 0.0045 | 1.096 |
| | MAPE | 0.0888 | 0.0671 | 1.3229 | **0.0875** | 0.0668 | 1.309 |
| | RMSE | **0.0057** | 0.0053 | 1.0772 | 0.0059 | 0.0054 | 1.099 |
| | U2 | **0.1214** | 0.1196 | 1.0149 | 0.1232 | 0.1183 | 1.041 |
| Feb | MAE | **0.0064** | 0.0061 | 1.0536 | 0.0078 | 0.0060 | 1.304 |
| | MAPE | **0.1199** | 0.1161 | 1.0328 | 0.1418 | 0.1055 | 1.343 |
| | RMSE | **0.0075** | 0.0063 | 1.1933 | 0.0092 | 0.0059 | 1.556 |
| | U2 | **0.1303** | 0.1037 | 1.2567 | 0.1442 | 0.1063 | 1.356 |
| Mar | MAE | **0.0055** | 0.0038 | 1.4444 | 0.0084 | 0.0039 | 2.155 |
| | MAPE | **0.0929** | 0.0637 | 1.4592 | 0.1390 | 0.0529 | 2.627 |
| | RMSE | **0.0064** | 0.0048 | 1.3391 | 0.0097 | 0.0043 | 2.261 |
| | U2 | **0.1223** | 0.0998 | 1.2255 | 0.1793 | 0.0947 | 1.893 |
| Apr | MAE | **0.0053** | 0.0052 | 1.0250 | 0.0061 | 0.0051 | 1.194 |
| | MAPE | **0.1022** | 0.1004 | 1.0174 | 0.1130 | 0.0831 | 1.360 |
| | RMSE | **0.0062** | 0.0049 | 1.2663 | 0.0072 | 0.0047 | 1.521 |
| | U2 | **0.1221** | 0.1104 | 1.1061 | 0.1395 | 0.1122 | 1.243 |
| May | MAE | **0.0111** | 0.0083 | 1.3407 | 0.0156 | 0.0078 | 2.002 |
| | MAPE | **0.1449** | 0.1271 | 1.1404 | 0.2016 | 0.1284 | 1.570 |
| | RMSE | **0.0128** | 0.0103 | 1.2386 | 0.0171 | 0.0093 | 1.842 |
| | U2 | **0.1534** | 0.1498 | 1.0239 | 0.1922 | 0.1577 | 1.218 |
| Jun | MAE | **0.0068** | 0.0038 | 1.7894 | 0.0082 | 0.0044 | 1.874 |
| | MAPE | **0.1085** | 0.0841 | 1.2900 | 0.1482 | 0.0749 | 1.978 |
| | RMSE | **0.0085** | 0.0064 | 1.3243 | 0.0096 | 0.0057 | 1.691 |
| | U2 | **0.1695** | 0.0931 | 1.8202 | 0.2342 | 0.0893 | 2.622 |
| Jul | MAE | 0.0219 | 0.0114 | 1.9232 | **0.0111** | 0.0108 | 1.027 |
| | MAPE | 0.2371 | 0.1113 | 2.1305 | **0.1288** | 0.1136 | 1.133 |
| | RMSE | 0.0261 | 0.0134 | 1.9492 | **0.0150** | 0.0133 | 1.130 |
| | U2 | 0.1660 | 0.0978 | 1.6974 | **0.1068** | 0.0922 | 1.158 |
| Aug | MAE | **0.0163** | 0.0143 | 1.1380 | 0.0173 | 0.0144 | 1.198 |
| | MAPE | **0.1413** | 0.1294 | 1.0920 | 0.1521 | 0.1309 | 1.161 |
| | RMSE | **0.0222** | 0.0221 | 1.0068 | 0.0229 | 0.0222 | 1.029 |
| | U2 | 0.1052 | 0.1039 | 1.0126 | **0.1063** | 0.1027 | 1.035 |
| Sep | MAE | **0.0105** | 0.0103 | 1.0210 | 0.0108 | 0.0101 | 1.073 |
| | MAPE | **0.1291** | 0.1082 | 1.1929 | 0.1353 | 0.0922 | 1.467 |
| | RMSE | **0.0131** | 0.0116 | 1.1267 | 0.0140 | 0.0121 | 1.154 |
| | U2 | **0.1847** | 0.1655 | 1.1161 | 0.2074 | 0.1672 | 1.240 |
| Oct | MAE | 0.0078 | 0.0078 | 0.9993 | **0.0073** | 0.0083 | 0.881 |
| | MAPE | **0.1104** | 0.1021 | 1.0814 | 0.1048 | 0.1013 | 1.034 |
| | RMSE | **0.0094** | 0.0063 | 1.4950 | 0.0095 | 0.0064 | 1.479 |
| | U2 | **0.1810** | 0.1733 | 1.0445 | 0.1987 | 0.1536 | 1.293 |
| Nov | MAE | **0.0145** | 0.0113 | 1.2836 | 0.0270 | 0.0125 | 2.162 |
| | MAPE | **0.1638** | 0.1333 | 1.2289 | 0.3221 | 0.2001 | 1.609 |
| | RMSE | **0.0222** | 0.0208 | 1.0655 | 0.0313 | 0.0209 | 1.497 |
| | U2 | **0.1749** | 0.1677 | 1.0429 | 0.2804 | 0.1631 | 1.719 |
| Dec | MAE | **0.0170** | 0.0143 | 1.1846 | 0.0249 | 0.0154 | 1.620 |
| | MAPE | **0.1713** | 0.1093 | 1.5670 | 0.3135 | 0.0978 | 3.205 |
| | RMSE | **0.0241** | 0.0198 | 1.2170 | 0.0286 | 0.0152 | 1.880 |
| | U2 | **0.3100** | 0.2931 | 1.0577 | 0.5681 | 0.2992 | 1.898 |
| **Average** | MAE | **0.0107 (14.4%)** | **0.0084** | **1.2706** | 0.0125 | 0.0086 | 1.466 |
| | MAPE | **0.1342 (19.0%)** | 0.1043 | **1.2963** | 0.1656 | **0.1040** | 1.650 |
| | RMSE | **0.0137 (8.7%)** | 0.0110 | **1.2749** | 0.0150 | **0.0104** | 1.512 |
| | U2 | **0.1617 (21.7%)** | 0.1398 | **1.2015** | 0.2067 | **0.1380** | 1.476 |

### 3.6.2.2 Verification of Generalization Ability with CTSGAN

To investigate why the proposed CTSGAN model gives a better prediction than the simple LSTM model, we compare the test and training errors of the CTSGAN and LSTM models for each month in the year 2012. The error gain between the test and training errors, which represents the generalization capability, is calculated.

Table 3.4 shows a monthly analysis of the forecasted electricity prices in 2012. The values in brackets in the bottom four cells of the CTSGAN/Test Column represent the improvement of error by using the CTSGAN model compared to the LSTM model, elucidating that the proposed CTSGAN model can give overall better forecasts for the test set. However, the forecasts in July are worse than those given by the LSTM model. The errors of forecasts given by the CTSGAN model are 0.0219 (MAE), 0.2371 (MAPE), 0.0261 (RMSE) and 0.1660 (U2), which are larger than those of the simple LSTM model by 0.0111 (MAE), 0.1288 (MAPE), 0.0150 (RMSE) and 0.1068 (U2), respectively. This may be attributed to the introduction of carbon pricing in July. Although carbon pricing has been removed from our model, it still profound affects prices and the amount of electricity generation and demand. In July, the first month of the introduction the carbon tax, the generator in the proposed CTSGAN does not fit well, and the discriminator requires more training to effectively distinguish between the real data and the generated data under the data mutation. Also, it can be found that the difference in fitting error between two models for the July training set is not significant, and only the generalization ability is affected.

After calculating the average of the training and test errors for each of the 12 months, it is evident that the proposed CTSGAN model can forecast better for the test set, with the average MAE of 0.0107, MAPE of 0.1342, RMSE of 0.0137, and U2 of 0.1617. Compared to the simple LSTM model, the 12-month average MAE, MAPE, RMSE, and U2 are improved by 14.4%, 19.0%, 8.7%, and 21.7%, respectively. However, when comparing the fitting errors of the training set, the fitting errors of the LSTM model are smaller, with the average MAPE of 0.1040, the average RMSE of 0.0104, and the average U2 of 0.1380. Overall, the LSTM model is an excellent fitting tool for the time series due to the selection of long and short network memory ability, resulting in a minimal training error, but at the same time may introduce the problem of overfitting. For the proposed CTSGAN model, due to the latent space, there are two other network errors, the embedding error and the recovery error. Compared to the simple LSTM model, the overall fitting error is inferior to that of the LSTM model.




(a) (b)

(c) (d)

Fig. 3.13 Monthly gains between CTSGAN and LSTM in terms of (a) MAE (b) MAPE (c) RMSE (d) U2

Furthermore, the CTSGAN/Gain Column and LSTM/Gain Column, obtained by dividing the test error by the training error, represent the generalization ability based one the CTSGAN and LSTM models. The smaller the value, the better the generalization ability. Fig. 3.13 illustrates the gains with CTSGAN and LSTM in terms of MAE, MAPE, RMSE, and U2 in 2012.

From Fig. 3.13 (b), compared to the gain between testing MAPE error and training error of LSTM, the gain of CTSGAN is significantly smaller, for example, which is only half of that of LSTM, with a MAPE of 1.45 and 1.57 in March and December, respectively. In terms of MAE, RMSE, and U2, the gains based on CTSGAN are much smaller in most months of the whole year, meaning that the CTSGAN model has a better generalization ability than the LSTM model. Although the training error of the CTSGAN model is not the same as that of the LSTM model, the gain of the CTSGAN model is significantly better than that of the LSTM model due to the generalization ability provided by the discriminator. Since the latent space retains the temporal dynamics, which improves the forecasting ability, point forecasting based on the CTSGAN model is effective.

#### 3.5.2.3 Verification of Contribution of Each Component in CTSGAN

To analyze the contribution of each component in CTSGAN, RMSE is used as a point prediction benchmark with the following modifications: (1) without conditions; (2) without the Wasserstein Distance mechanism; (3) without the supervised loss; (4) without the embedding networks; and (5) without joint training process. Table 3.5 lists the RMSE results of the ablation experiment. Overall, one can find that all five elements make important contributions in improving the quality of electricity prices forecasting. The condition set plays one of the most important roles in improving the accuracy of the prediction results. Taking the 2014 price prediction results as an example, one can find that the ranking according to the error range contribution is in the order of the joint training process (±0.0127), Wasserstein Distance (±0.0122), embedding networks (±0.0094), and supervised loss (±0.0068), which well reveals the reason for the stability of the proposed CTSGAN.

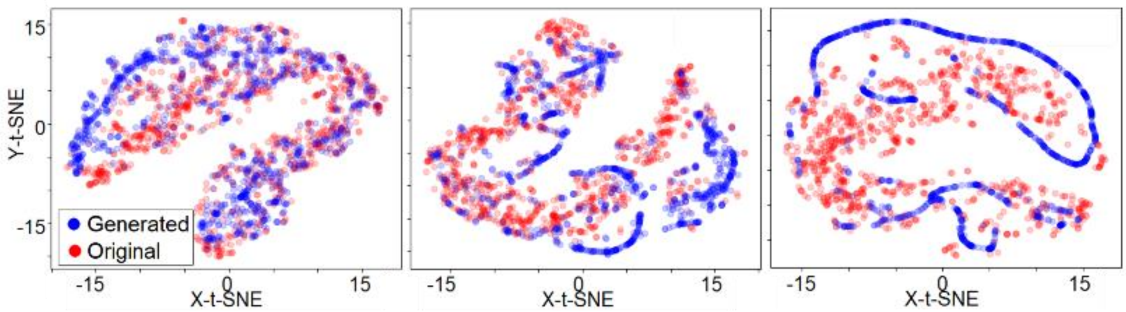
Table 3.5 Contribution of each component in CTSGAN

| | 2012 | 2014 | 2016 | 2018 |
|---|---|---|---|---|
| CTSGAN | 0.0137±0.0061 | 0.0132±0.0043 | 0.0456±0.0074 | 0.0372±0.0036 |
| w/o conditions | 0.0539±0.0073 | 0.0574±0.0047 | 0.0749±0.0094 | 0.0974±0.0055 |
| w/o Wasserstein | 0.0377±0.0143 | 0.0465±0.0122 | 0.0691±0.0101 | 0.0571±0.0098 |
| w/o supervised loss | 0.0471±0.0078 | 0.0493±0.0068 | 0.0731±0.0073 | 0.0837±0.0061 |
| w/o embedding | 0.0289±0.0086 | 0.0316±0.0094 | 0.0592±0.0082 | 0.0629±0.0063 |
| w/o joint training | 0.0363±0.0116 | 0.0485±0.0127 | 0.0605±0.0099 | 0.0774±0.0122 |

### 3.5.3 Interval Prediction Results and Analysis

#### 3.5.3.1 Verification of Authenticity of Generated Scenarios with CTSGAN

The scenario-based prediction we propose needs to address two key issues, one of which is authenticity. Compared to the conventional ML and DL methods, generative models can generate a larger set of scenarios. Therefore, the authenticity of the generated scenarios for different generative models determines whether this model is available. This section investigates whether the scenarios generated by CTSGAN, TSGAN, and GAN are realistic using t-distributed stochastic neighbour embedding (t-SNE). As t-SNE maps each of the 48-dimensional day-ahead electricity price scenarios to a two-dimensional point, the nearby points model similar scenarios and distant points different scenarios. Fig. 3.14 illustrates the t-SNE visualization on the generated scenarios (blue) and the original (red) electricity prices. Figs. 3.14 (b) and (c) show that the scenarios generated by the TSGAN and GAN are poorly distributed and mainly concentrated on a few curves. This means that some features are lost in the training process, resulting in a poor final overlap. Comparing Fig. 3.14 (a) with Fig. 3.14 (b), one can find that the scenarios generated by CTSGAN show better overlap with the original electricity prices than those obtained using other generative models. Therefore, the CTSGAN model can generate price scenarios that are relatively more realistic.




(a) (b) (c)

Fig. 3.14 Visualization with t-SNE for generated scenarios based on (a) CTSGAN (b) TSGAN, and (c) GAN

#### 3.5.3.2 Verification of Diversity of Generated Scenarios with CTSGAN

In addition to the above-mentioned authenticity, diversity is also an important issue. Only on the basis that the diversity of generated scenarios can be guaranteed the introduction of TSM can obtain PIs with a high CP. Fig. 3.15 provides a comparative experiment in terms of predictive time cost and CP with different numbers of generated scenarios to determine the optimal $N_R$.

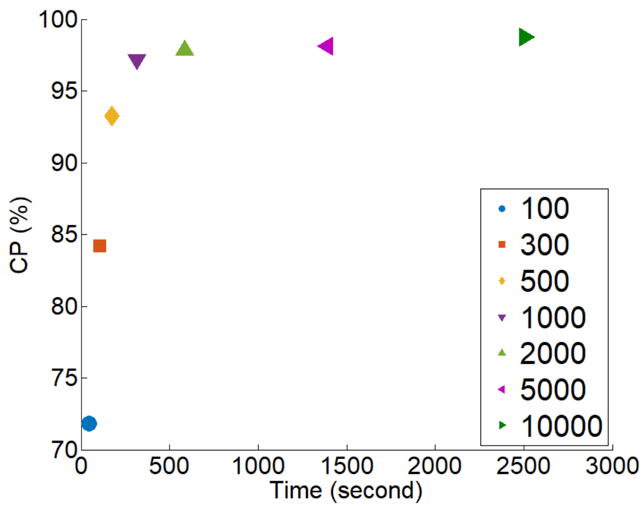



Fig. 3.15 Comparative experiment in terms of prediction time cost and CP with different $N_R$

It can be seen that when $N_R$ is less than or equal to 500, the CP of the generated scenarios is less than or equal to 93.23%, which means that it is impossible to achieve the CP requirement when NCP is set at 95%. When $N_R$ is 1000, the maximum CP can reach 97.16%, which generally satisfies NCP=95%. However, when $N_R$ continues to increase, even when it reaches 10000, the maximum CP cannot significantly increase and remains at 98.56%. On the other hand, the time cost of the prediction increases significantly, from 314.4 seconds ($N_R$=1000) to 2503.2 seconds ($N_R$=10000). Therefore, we choose $N_R$=1000, which can balance the prediction CP and time cost. In practical applications, if there is a demand for an NCP higher than 98%, the value of $N_R$ can be appropriately increased.

Then, we introduce the CTSGAN, TSGAN, and GAN methods to verify the diversity. Ten consecutive days for four seasons in 2020 are used as a tested dataset for PIs verifications. Day 1 to 10 in July are collected as winter test set, October as spring, January as summer, and April as autumn. Table 3.6 shows the CP for PIs for four seasons in 2020.

Table 3.6 Comparison of CP of PIs with different generative model

| Models | CTSGAN | TSGAN | GAN |
|---|---|---|---|
| Win. | 96.18 | 87.46 | 77.94 |
| Spr. | 96.89 | 92.59 | 76.73 |
| Sum. | 97.68 | 86.51 | 69.76 |
| Aut. | 97.49 | 93.10 | 82.13 |
| Avg. | 97.06 | 89.92 | 76.64 |

From Table 3.6, one can find that the three generative models: CTSGAN, TSGAN, and GAN, show different scenario-diversity generation abilities with different CPs. The basic GAN is an unsupervised model, which provides the worst quality PIs, with an average CP of 76.64%. The TSGAN introduces supervised loss to basic GAN and improves the CP from 76.64% to 89.92%. These two generative models are hard to be applied to construct PIs since they cannot provide high CPs. For example, the CP of PIs with TSGAN is 93.10% in autumn and 86.51% in summer, which means that even with the introduction of TSM, it is not possible to obtain PIs that satisfy an NCP of 95% for summer and 90% for winter. Similarly, GAN is not applicable. The CTSGAN can generate relatively diverse scenarios, and the combined PIs can cover 97.06% of real prices. With the scenarios generated by CTSGAN, the combined PIs can achieve a high CP, like 90% and 95%.

Although the PIs provided by the CTSGAN are diverse, they contain many low-probability scenarios, which directly affect the quality of the PIs with excessive sharpness. We introduce TSM to truncate the threshold to make the CP of PIs guarantee the NCP with minimum sharpness. Table 3.7 illustrates the CP of predicted PIs obtained by TSM-CTSGAN with different NCPs from 70% to 95%.

Since reliability and sharpness are contradictory, the increase in reliability is accompanied by an increase in sharpness. Therefore, the ideal PIs are those with CPs that just meet the NCP requirement, and the sharpness is minimized. Note that it is not that the larger the CP is, the better. From Table 3.7, one can find that the proposed TSM can effectively drive the CP to meet the NCP requirements, with an average CP of 95.24% (NCP=95%), 91.56% (NCP=90%), 85.53% (NCP=85%), 81.07% (NCP=80%), 76.72% (NCP=75%) and 72.28% (NCP=70%).

Table 3.7 Comparison of PIs generated by TSM-CTSGAN with different NCPs

| NCP | 95% | 90% | 85% | 80% | 75% | 70% | w/o |
|---|---|---|---|---|---|---|---|
| Win. | 95.23 | 90.83 | 85.71 | 81.54 | 76.76 | 71.94 | 96.18 |
| Spr. | 95.01 | 91.21 | 86.01 | 80.71 | 77.20 | 72.34 | 96.89 |
| Sum. | 95.68 | 92.27 | 85.29 | 81.92 | 76.98 | 72.45 | 97.68 |
| Aut. | 95.04 | 91.92 | 85.10 | 80.12 | 75.93 | 72.39 | 97.49 |
| Avg. | 95.24 | 91.56 | 85.53 | 81.07 | 76.72 | 72.28 | 97.06 |

#### 3.5.3.3 Comparative Simulation of PIs with TSM-CTSGAN Model Considering Weather Factors

To explore the effect of weather factors on threshold selection and PIs prediction, we retrain the TSM after removing the weather factor. The model without weather factors is called TSM2-CTSGAN. Table 3.8 shows CP and WS for PIs. The NCPs are set to 90%, 80%, and 70%, respectively. All PIs that do not satisfy the NCP are underlined, and meanwhile, these WS results yielded are marked in bold.

Table 3.8 Comparison of PIs with TSM-CTSGAN and TSM2-CTSGAN

| Models | | TSM-CTSGAN | | | TSM2-CTSGAN | | |
|---|---|---|---|---|---|---|---|
| NCP | | 70% | 80% | 90% | 70% | 80% | 90% |
| Win. | CP | 71.94 | 81.54 | 90.83 | 72.07 | 78.64 | 85.55 |
| | WS | 6.72 | 6.90 | 7.81 | 6.92 | **7.40** | **8.78** |
| Spr. | CP | 72.34 | 80.71 | 91.21 | 72.45 | 80.93 | 88.91 |
| | WS | 5.90 | 7.41 | 8.68 | 6.19 | 7.99 | **9.24** |
| Sum. | CP | 72.45 | 81.92 | 92.27 | 70.59 | 78.01 | 83.57 |
| | WS | 5.42 | 6.67 | 7.12 | 5.59 | **7.27** | **9.47** |
| Aut. | CP | 72.39 | 80.12 | 91.92 | 72.24 | 81.02 | 89.34 |
| | WS | 5.53 | 6.68 | 7.44 | 5.78 | 6.96 | **8.16** |
| Avg. | CP | 72.28 | 81.07 | 91.56 | 71.84 | 79.65 | 86.84 |
| | WS | 5.89 | 6.92 | 7.76 | 6.12 | 7.41 | 8.91 |

Table 3.8 also compares the quality of PIs with TSM-CTSGAN (with weather factors) and TSM2-CTSGAN (without weather factors). One can find that the CP of PIs with TSM-CTSGAN is able to reach the NCP, regardless of whether the NCP is 70% or 90%. However, the CP results of TSM2-CTSGAN are unstable; when the NCP is set to 70%, the CP can reach the NCP, with an average value of 71.84%; the CP cannot reach the NCP target when the NCP is set to 90%, with an average CP of 86.84%.

This means that for TSM2, the prediction performance decreases at high CP and suggests that the unpredictable prices at high CP are generally due to sudden weather changes, which significantly impact the prediction results after removing the weather factor with TSM2-CTSGAN. In summer, TSM can lead to a significantly improved reliability and sharpness with a high NCP (90%) compared to TSM2, with WS from 9.47 to 7.12 and CP from 83.57% to 92.27%. Because of the higher temperature in summer, the superimposed weather factors are more likely to lead to price fluctuations that are difficult to predict. The TSM-CTSGAN-based PIs method can solve well and ensure higher CP and lower WS in the PIs.

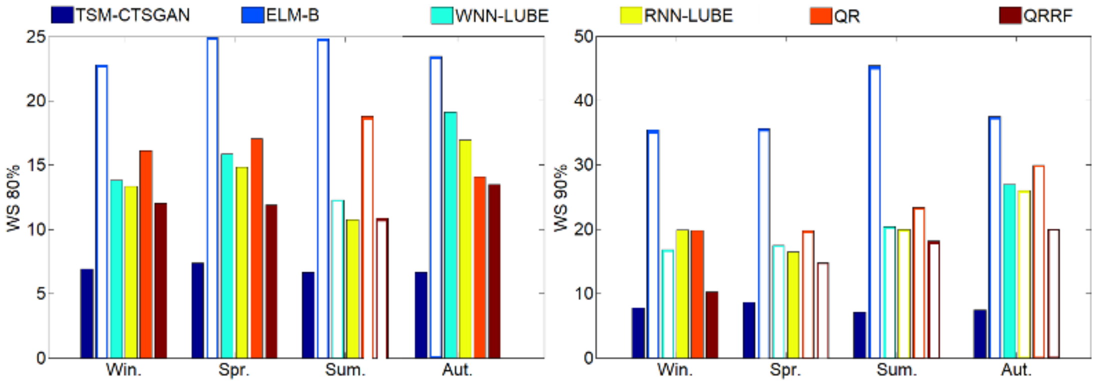
#### 3.5.3.4 Comparative Simulation of PIs with Different Models

This section introduces the conventional probabilistic forecasting models, such as Bootstrap, LUBE, and QR, to benchmark the proposed TSM-CTSGAN model. The parameters of extreme learning machine (ELM) combined with Bootstrap (ELM-B) are set the same as in [88], and two LUBE models are constructed by wavelet neural network (WNN) and recurrent neural network (RNN), known as WNN-LUBE and RNN-LUBE [85], respectively, with NCPs of 90% and 80%. The QR model and quantile regression random forest (QRRF) are set with the quantiles at 90% and 80%, respectively.

Table 3.9 Comparison of PIs with Different Prediction Models

| Seasons | | Win. | | Spr. | | Sum. | | Aut. | |
|---|---|---|---|---|---|---|---|---|---|
| NCP | | 80% | 90% | 80% | 90% | 80% | 90% | 80% | 90% |
| TSM-CTSGAN | CP | 81.54 | 90.83 | 80.71 | 91.21 | 81.92 | 92.27 | 80.12 | 91.92 |
| | WS | 6.90 | 7.81 | 7.41 | 8.68 | 6.67 | 7.12 | 6.68 | 7.44 |
| ELM-B | CP | <u>73.20</u> | <u>82.91</u> | <u>78.73</u> | <u>89.22</u> | <u>74.57</u> | <u>82.13</u> | <u>76.38</u> | <u>83.81</u> |
| | WS | **22.82** | **35.51** | **24.93** | **35.67** | **24.83** | **45.43** | **23.46** | **37.56** |
| WNN-LUBE | CP | 80.54 | <u>87.67</u> | 80.79 | <u>89.66</u> | <u>78.48</u> | <u>85.84</u> | 82.63 | 90.84 |
| | WS | 13.82 | **16.83** | 15.87 | **17.49** | **12.28** | **20.37** | 19.12 | 26.97 |
| RNN-LUBE | CP | 82.89 | 90.83 | 80.53 | 91.78 | 81.93 | <u>86.60</u> | 80.36 | <u>89.23</u> |
| | WS | 13.34 | 19.89 | 14.83 | 16.45 | 10.77 | **19.98** | 16.93 | **26.02** |
| QR | CP | 83.90 | 90.83 | 82.83 | <u>87.83</u> | <u>78.73</u> | <u>80.14</u> | 81.79 | <u>87.12</u> |
| | WS | 16.10 | 19.84 | 17.04 | **19.83** | **18.79** | **23.35** | 14.09 | **29.90** |
| QRRF | CP | 80.65 | 90.83 | 83.94 | <u>88.04</u> | <u>76.78</u> | <u>82.93</u> | 80.93 | <u>89.52</u> |
| | WS | 12.05 | 10.26 | 11.90 | **14.82** | **10.90** | **18.19** | 13.50 | **20.01** |

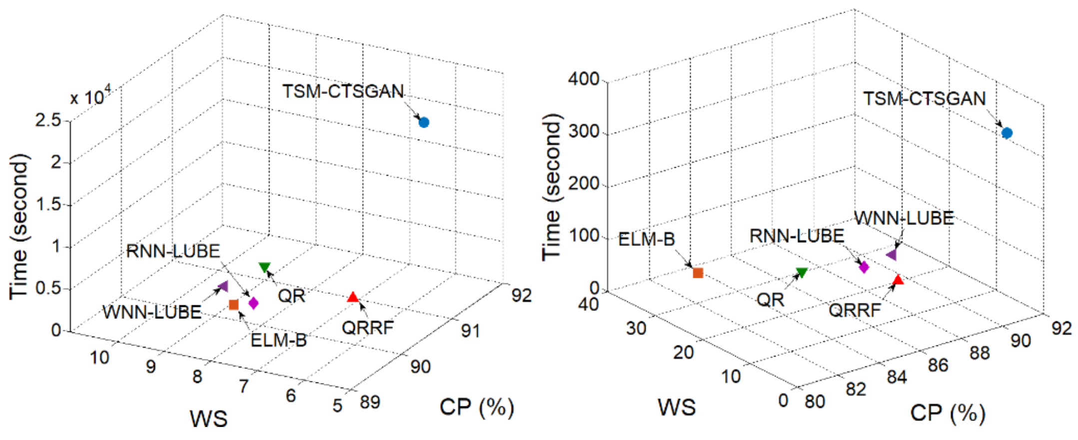
Table 3.9 illustrates the CP and WS for PIs with different prediction models with NCPs of 80% and 90%. All the PIs provided by the ELM-B model cannot achieve the NCP, which means that the reliability cannot be guaranteed. All PIs that do not satisfy the NCP are underlined in Table 3.9; meanwhile, the WS results yielded by these PIs are not meaningful and are marked in bold. For 90% NCP in summer, our proposed method can provide the PIs that meet the high NCP, while the Bootstrap, LUBE, and QR methods provide poor reliable PIs.




(a) (b)

Fig. 3.16 Winkler score comparison in terms of NCP of (a) 80% and (b) 90%

Figs. 3.16(a) and (b) illustrate the WS results in terms of NCP of 80% and 90%, respectively. As shown by hollow bars, some of the predicted results cannot reach the NCP, which means that WS is not meaningful. One can find that the TSM-CTSGAN provides the best PIs with the smallest WS for all seasons, while NCP is set as 80% and 90%. For example, in the autumn of 2020, all other methods provide higher CPs (82.63%, 80.36%, 81.79%, and 80.93%) than TSM-CTSGAN (80.12%). This is because these methods construct a very wide interval to achieve high coverage. Considering both reliability and sharpness, the WS provided by CTSGAN (6.68) is less than half of other methods, among which QRRF provides the smallest WS with 13.50. This indicates that our proposed method has a better predictive performance than other methods, always satisfying the NCP requirements and guaranteeing the smallest sharpness.




(a) (b)

Fig. 3.17 Comparisons for time-consumption, WS and CP for different PI models in process of (a) training and (b) test

The training and test time for day-ahead PI prediction is calculated with the NCP of 90%, plotted as Fig. 3.17. It can be found that during the training process, training time for the proposed TSM-CTSGAN is almost 2.2*104 second (6.1 hours), and that for the other predicted models is less than 2000 seconds (30 mins). The reason for this difference is that our model CTSGAN involves 4 core networks and ultimately 1000 scenarios are generated for each prediction point to ensure realism and diversity. Other methods usually construct the intervals directly, so the training time is much shorter. However, we found that the proposed model just needs to be updated once in about 3 months, and 6 hours for every time is acceptable. In the test process, it can be found that the prediction time is much smaller. Our proposed method costs 310 seconds, and other methods commonly take 10 seconds. This is for the same reason as described above. It takes longer on 1000 generated scenarios. For day-ahead problem, 310 seconds (5 mins) is acceptable. However, for the day-head PI prediction problem, the most important thing is not the prediction time, but whether the CP of the predicted results can meet the NCP and have a smaller WS. As can be seen from Fig. 3.17, although all algorithms can achieve NCP during the training process, only TSM-CTSGAN and WNN-LUBE can achieve it during the prediction process. In addition, the WS given by our model is only 0.28 of WNN-LUBE. Therefore, our TSM-CTSGAN has outstanding performance, and the time cost is acceptable.

## 3.6 Chapter Summary

Accurate electricity price forecasting plays an essential role in the electricity market. This chapter proposed a new CTSGAN-based electricity price forecasting model, which can well preserve the temporal dynamics for the whole period. Firstly, for point prediction, the proposed CTSGAN model had better prediction accuracy than other conventional ML or DL models and had better generalization ability than the simple LSTM model due to the introduction of the discriminator. Secondly, we proposed a novel scenario-based PI construction method. A TSM considering weather factors was proposed to make the PIs controllable. The case studies demonstrated that 1) the CTSGAN training process can remain stable; 2) the CTSGAN model can generate realistic and diverse scenarios; 3) TSM can guide CTSGAN to achieve NCP; 4) introducing weather factors can improve the prediction performance; 5) TSM-CTSGAN can outperformed other PI methods significantly in terms of CP and WS.

# CHAPTER 4

# Assembly and Competition for Virtual Power Plants through A “Recruitment – Participation” Approach

*DER owners have many concerns when participating in VPPs, including low returns and a limited variety of plans. Therefore, it lacks incentives for them to participate in VPPs. This chapter proposes a flexible VPP assembly approach to recruit DERs by ESPs. The contributions of each ESP and DER to VPPs are evaluated separately as the basis for a fair payoff allocation. Rebates are adopted to incentivize DER owners to participate in the VPPs. A multi-ESP VPP assembly and competition method are proposed and simulated based on the above fundamentals, considering diverse preference behaviors and electricity price fluctuations. The existence of the solution to the proposed multi-ESP VPP rebate competition problem is discussed. The results indicate that DER owners are motivated to participate in VPPs with high payoffs, and the payoffs of ESPs increase by attracting more customers.*

## 4.1 A Recruitment – Participation Approach

In this chapter, VPPs are no longer ESP-centered. ESPs can recruit any DERs, and DERs can participate in any ESP at will. If the ESPs and DERs are considered modules, the VPP construction process, jointly carried out by ESPs and DERs, can be called VPP assembly. The following assumptions on the VPP assembly are made in this study.

***Assumption 1:*** DERs can choose ESPs according to their preferences. There is no cost for DERs to switch ESPs.

***Assumption 2:*** The operation cycle (OC) is defined as the smallest operation period for a VPP assembly. DERs cannot exit VPPs within each OC, and ESPs cannot change the published strategy.

***Assumption 3:*** An ESP recruits DERs through a prospectus. When the parties sign a contract, control of the DER is placed under the control of the ESP.

Fig. 4.1 shows the timeline of the entire 'Recruitment - Participation' approach. Initially, there are no VPP, only DER owners who are willing to participate in VPPs and potential ESPs as shown in Fig. 4.1(a). Then, the VPPs are assembled (Fig. 4.1(b)), followed by possible challenges of disassembly (Fig. 4.1(c)), reassembly (Fig. 4.1(d)), and competition. The profit-risk profiles are also changed in this process, as shown in Figs. 4.1(e)-(g). The details are introduced and explained in the next subsection.

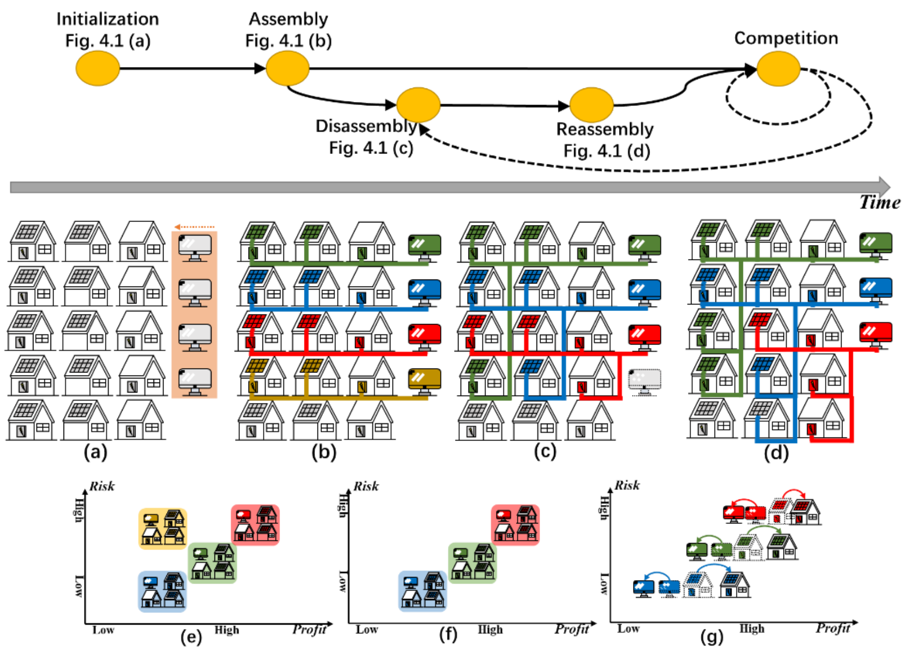



Fig. 4.1 Proposed 'Recruitment – Participation' approach for VPP. (a) Initialization, (b) Assembly, (c) Disassembly, (d) Reassembly, (e-g) Changes in risk-profit in VPP competition

### 4.1.1 Recruitment – Participation Approach

#### 4.1.1.1 VPP Assembly

Assume that there currently exists an area with several PV and battery owners. Some of whom are not satisfied with using PVs and batteries only for their own consumption and wish to obtain more profit. Therefore, they are highly willing to participate in VPP; however, based on their own preferences, their preferences for risk and return are diverse, which means that one single VPP plan cannot satisfy everyone's needs. Therefore, multiple VPP ESPs need to be introduced to provide differentiated plans with competition (Fig. 4.1 (a)).

**ESP:** Multiple ESPs join as participants, and through prediction, simulation and calculation, each ESP provides a prospectus to attract DERs before an OC.

**DER:** Partial DERs choose different ESPs to assemble VPPs according to the published prospectus and their own preferences, while the remaining DERs are in a wait-and-see situation, as shown in Fig. 4.1(b). During the initial stage of VPP assembly, DERs are not allowed to exit the VPP for some cycles, named closed OCs.

#### 4.1.1.2 VPP Disassembly and Reassembly

After the closed OCs, the actual profit and expected profit of VPP operation may differ. Therefore, DERs' confidence may change since the actual and expected payoffs differ. Some DERs will leave the current VPP and choose a new ESP to assemble the VPP.

**ESP:** With the loss of members, the high-risk and low-profit VPP ESPs (the yellow one in Fig. 4.1(e)) will eventually disassemble the VPP. Only the low-risk and high-profit VPPs (the green, blue and red ones in Figs. 4.1(e) and 4.1(f)) can be retained, and the ESPs will gradually become oligopolies by offering differentiated VPP plans.

**DER:** The DERs of the disassembled VPPs can choose other ESPs, as shown in Fig. 4.1(c). The participants of the other VPPs can also re-select different VPP plans with the existing ESPs according to their preferences.

#### 4.1.1.3 VPP Competition

After a period of operation, the number of DER participants in each VPP cannot be further increased. ESPs give DER owners rebates to attract non-participating DER owners, while ESPs compete to attract DERs from other VPPs.

**ESP:** ESPs provide rebates to DER participants to reward them [46]. In Fig. 4.1(g), the arrows represent the profit change of ESPs and other participants before and after the rebates. The dotted icons are profits before the rebates, and the solid ones are with the rebates.

**DER:** The rebates promote ESP reselection from PV and battery participants [89]. In addition, as shown in Fig. 4.1 (d), the high profits with rebates can attract wait-and-see DERs.

### 4.1.2 Operation Cycle

The OC is the smallest unit of time for VPP assembly. A short OC will lead to frequent changes in the number of DERs, which is a challenge for ESPs to participate in the energy market. However, if OC is very long, it is not conducive for DER participants to regularly re-select ESP based on actual benefits and risks. Note that OC can be set to any number of days. This chapter considers the weekly periodicity of the electricity spot price market, so OC is set to seven days. Fig. 4.2 illustrates the VPP assembly process for OCs.

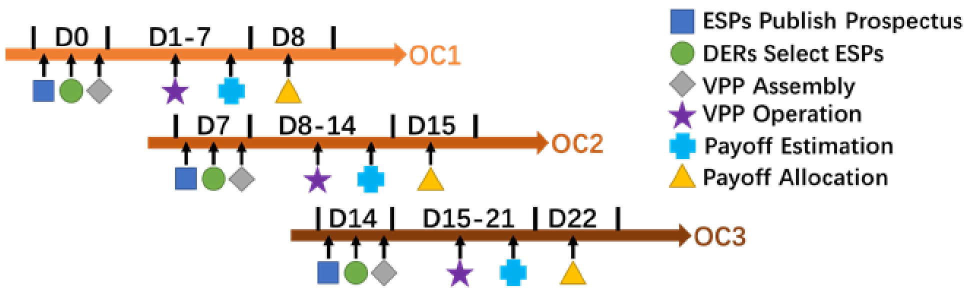



Fig. 4.2 VPP assembly flow for OCs

Before the VPP assembly, Day0 (D0) in Fig. 4.2, ESPs publish a prospectus. Potential DERs can select the appropriate ESPs based on their preferences and prospectus disclosures, and then they are assembled into VPPs. Starting from D1, the first day of OC1, ESPs perform VPP operations according to the established optimization strategy while giving the daily estimated payoffs for each DER until D7. Meanwhile, each ESP is required to publish a prospectus for OC2 on the last day of OC1, and DERs can re-select ESPs. On D8, the ESPs calculate the actual payoff of each DER based on the SV and make the allocation. D8 also serves as the first day of OC2, on which ESPs start a new VPP operation round and then repeat the above process.

### 4.1.3 Prospectus

The ESPs attract DERs by publishing a prospectus, which includes the following information:

- A summary of the ESP's background,
- Basic information on the VPP, including
  - A brief description of the strategy,
  - Predicted profits,
  - Nominal risks, and
  - Rebate coefficients;
- Historical information on VPP (if available), including
  - Actual VPP profits and fluctuations,
  - Actual long- and short-term returns for DERs, and
  - Actual size of DERs.

## 4.2 Behaviors in ESPs Competition and Proposed Models

Prospectus directly influences DERs' selection on ESPs. The choice of participants, in turn, promotes the ESPs' rebate competition. The participation, ESP selection, and rebate competition models are proposed to simulate the participants' behaviors in the VPP competition process.

### 4.2.1 Participation or Non-participation

Different DERs have various elasticities of willingness under different payoffs, and linear models are insufficient to portray the behavior of DERs participating in VPP. However, there is a lack of research on the relationship between payoffs and VPP participation rate for DERs. Based on the PV installation and EV dispatching models in [51, 52], a nonlinear model is assumed to describe the payoff increment and VPP participation rate as follows:

$$\mu = -\frac{1}{\left((\Theta^{w/}-\Theta^{w/o})/\Theta^{w/o}\right)*a+b} + 1 \tag{4.1}$$

where $\mu$ represents the overall participation rate for DERs; $\Theta^{w/}$ and $\Theta^{w/o}$ represent the average profits with and without participation in the VPPs with different ESPs; *a* and *b* are coefficients used to control for the effect of incremental marginal payoffs on the participation ratio, $a \geq 1$ and $b \geq 1$, when $\Theta^{w/} = \Theta^{w/o}$, and the initial participation rate $\mu = 1 - 1/b$. Therefore, the marginal incremental payoff diminishes the increase in the DER participation rate, meaning most DERs' willingness significantly increases with a slight payoff rise. In contrast, the willingness of a small number of DERs remains low even after a significant payoff increase.

### 4.2.2 ESP Selection by Participants

DER participants should carefully consider various factors when selecting ESPs to assemble VPPs. Since the multi-ESP competition model is proposed for the first time in this chapter, there is a lack of research on how DER participants choose ESPs. The model for electricity retailer selection [53] does not apply here, as customers have a single purpose, reducing their bills. However, the participation of DER owners in VPP by selecting ESPs can be considered as an investment, and it has many similarities to mutual fund investment. For example: 1) Fund investment requires investors to invest cash to obtain returns, while VPP participants need to provide control over PV and battery to obtain returns. 2) Both funds and VPPs offer diversification in risk and return and require professionals to manage risks. 3) Funds and VPPs can recruit investors and DER participants, respectively, through the disclosure of prospectus, including historical returns, risk, and scale. 4) Fund managers receive management fees, while VPP ESPs also receive payoff for bringing return growth to DER participants. Therefore, based on the mutual fund selection model [90-92], an ESP selection model is proposed as follows:

$$\Gamma = \xi_1 \Theta_{\mathrm{ReturnL}} + \xi_2 \Theta_{\mathrm{ReturnS}} + \xi_3 (1 - \Theta_{\mathrm{Risk}}) + \xi_4 (1 - \Theta_{\mathrm{Fluc}}) + \xi_5 \Theta_{\mathrm{Scale}} + \xi_6 \Theta_{\mathrm{Other}} \tag{4.2}$$

where $\Gamma$ denotes the potential for an ESP to be selected by DERs; coefficients $\xi_1, \xi_2, \cdots, \xi_6$ indicate the weights for investment preference of DER participants; $\Theta_{\mathrm{ReturnL}}$ and $\Theta_{\mathrm{ReturnS}}$ represent the historical average long- and short-term payoffs which can be obtained from the disclosed prospectus. Due to the lack of data on historical returns for the new assembly VPP, the expected returns can be used instead for the historical payoffs; $\Theta_{\mathrm{Risk}}$ represents the nominal risk rating; $\Theta_{\mathrm{Fluc}}$ represents the fluctuation of historical returns. of disclosed historical returns; and $\Theta_{\mathrm{Scale}}$ represents the average number of DER owners recruited by one ESP, respectively. The five factors can be obtained from the prospectus of VPP and should be normalized to 0 to 1. Apart from the deterministic factors mentioned above, investment decisions often involve some uncertain factors. In [93], Dimitrios et al. point out that individual investors may rely on newspapers, media and noise in the market when making their investment decisions. Reputation, brand effects, and other factors may also influence investors' choices [94]. However, these factors are often difficult to quantify, so we use a random number between 0 and 1 (denoted as $\Theta_{\mathrm{Other}}$) to describe the uncertainty of individual choices.

Like fund investors [90, 95], DER owners also have different preferences when choosing an ESP to assemble VPPs. We categorize DER owners into different groups based on their investment preferences, including professional, ambitious, moderate, conservative, and cautious [96]. Their risk profiles and profit expectations are shown in Table 4.1. Additionally, according to Table 4.1, we assumed weights $\xi_1, \xi_2, \cdots, \xi_6$ for different preference participants, which are detailed in the experiment setting of Section 4.5.

Table 4.1 Classification of participant groups

| Types | Risk Profile | Expectation |
|---|---|---|
| Professional | Takes Necessary Risks | Maximum Return |
| Ambitious | Highly Risk Taking | High Short-Term Return |
| Moderate | Comfortable Levels of Risk | Good, Steady Return |
| Conservative | Risk Averse | Regular Return |
| Cautious | Extremely Risk Averse | Minimum Return |

Algorithm 4.1 below simulates the ESP selection process for each DER. Firstly, the participant's investment type and the factors' coefficients are determined. The relevant variables $[\Theta_{\mathrm{ReturnL}}, \Theta_{\mathrm{ReturnS}}, \cdots, \Theta_{\mathrm{Scale}}]$ can be extracted from the prospectus published by each ESP. A random variable $\Theta_{\mathrm{Other}}$ is generated, and the participant's

probability of participation for each ESP, $\Gamma$, can be calculated. The $\Gamma$ of all ESPs are then collected and composed into multiple probability intervals, such as $[0, \Gamma_1]$ $[\Gamma_1, \Gamma_1 + \Gamma_2]$, $\cdots$, $[\sum_{j=1}^{J-1} \Gamma_j, \sum_{j=1}^{J} \Gamma_j]$. The participant generates a random value between $[0, \sum_{j=1}^{J} \Gamma_j]$ and determines in which ESP's probability interval the random value lies to simulate the participant's willingness. Repeat the above steps until all participants have made their decisions.

**Algorithm 4.1** Simulation of ESP Selection for DER Participants

1:**For** DER Participant $j = 1{:}J$:
2: Obtain related coefficients $\xi_1, \xi_2, \cdots, \xi_6$ for DER $j$.
3: **For** ESP $m = 1{:}M$:
4: Obtain related factors $\Theta_{\text{ReturnL}}, \Theta_{\text{ReturnS}}, \cdots, \Theta_{\text{Scale}}$ from prospectus published by ESP $m$.
5: Generate random factor $\Theta_{\text{Other}}$.
6: Calculate potential $\Gamma$ for ESP $m$ selected by DER $j$.
7: **End**
8: Collect all potentials $[\Gamma_1, \Gamma_2, \cdots, \Gamma_M]$ with different ESPs for DER $j$.
9: DER $j$ determines the selected ESP.
10: **End**

### 4.2.3 ESP Rebate Competition

ESPs that adopt homogeneous competition are highly susceptible to disassembly due to the loss of participants once they become less profitable than other ESPs offering the same risk plan. The ESPs offer differentiated plans to achieve differentiated competition to attract participants with different preferences. For further profit, ESPs give rebates to attract more participants. The ESP selection based on (4.2) is related to historical profits $\Theta_{\text{ReturnL}}$ and $\Theta_{\text{ReturnS}}$, profit fluctuations $\Theta_{\text{Fluc}}$, which are related to rebate coefficient $\lambda$. In addition, real-world ESPs tend to be bounded by rationality, meaning that their decisions do not immediately lead to optimal solutions.

Therefore, the rebate coefficient $\lambda_c^m$ for DER participant for each ESP is defined as follows:

$$\lambda_c^m = \begin{cases} [\lambda_{c,PV}^m, \lambda_{c,battery}^m], if\ c > 1 \\ \quad [0,0] \quad\quad , if\ c = 1 \end{cases} \tag{4.3}$$

where $\lambda_{c,PV}^{m}$ and $\lambda_{c,battery}^{m}$ denote the rebate coefficients of the *m-th* VPP for OC $c$ for PV and battery, respectively, and $\lambda_{c}^{m} \in [0,1]$. The rebate competition process involves several ESPs and time-varying variables, such as electricity price and solar power generation, leading to a more intense ESP rebate. Therefore, we use DRL to simulate the multi-ESP competition.

# 4.3 System Models for VPP Profit and Allocation

The ESPs recruit participants by publishing a complete prospectus and assembling the VPP with DER participants so that the VPP plan can be considered an expression of the will of all participating members. This section provides generalized system models for VPPs participating in the energy market.

## 4.3.1 Profit Seeking and Risk Control

The ESP makes profits by predicting electricity prices, solar power generation, and battery dispatching to buy low and sell high for electricity. Conventional VPP profit-seeking models are given by (4.4) to (4.10). The profit $\varphi_{t,\omega}$ of the VPP for the energy market at time $t$ with predicted scenario ω is calculated as follows:

$$\varphi_{t,\omega} = \rho_{t,\omega_{\rho}}^{RT} P_{t}^{VPP} \tag{4.4}$$

$$P_{t}^{VPP} = P_{t,\omega_{PV}}^{PV} - \mu_{t}^{ch} P_{t}^{ch} + \mu_{t}^{dch} P_{t}^{dch} \tag{4.5}$$

$$\mu_{t}^{ch} + \mu_{t}^{dch} \leq 1 \tag{4.6}$$

$$P_{t,\omega_{PV}}^{PV}, P_{t}^{ch}, P_{t}^{dch} \geq 0 \tag{4.7}$$

where $\rho_{t,\omega_{\rho}}^{RT}$ represents the probabilistic real-time electricity price scenario, and $P_{t}^{VPP}$ the output of VPP; $P_{t,\omega_{PV}}^{PV}$, $P_{t}^{ch}$, and $P_{t}^{dch}$ denote the PV generation power, battery charging, and discharging power, respectively; $P_{t}^{ch}$ and $P_{t}^{dch}$ are less than the rated power capacity $P^{m}$. These two battery components are non-negative values and are limited by binary variables $\mu_{t}^{ch}$ and $\mu_{t}^{dch}$. The battery state of charge (SoC) constraints are shown below:

$$SoC_t = SOC_{t-1} + \left(\mu_t^{ch} P_t^{ch} \eta^{ch} - \frac{\mu_t^{dch} P_t^{dch}}{\eta^{dch}}\right) \Delta t / E^m \tag{4.8}$$

$$SoC_{min} \le SoC_t \tag{4.9}$$

$$SoC_0 = SoC_T \tag{4.10}$$

where $\Delta t$ denotes the time duration of each time interval, which is 30 minutes in this chapter. $\eta^{ch}$ and $\eta^{dch}$ are the battery charging and discharging efficiencies, respectively. There is a lower limit regarding the battery $SoC_t$, which is set at 10%; the initial and final SoCs are assumed to be the same and equal to 20%, as in (4.10).

ESPs attract different DER owners by offering differentiated VPP strategies, including risk-neutral, risk-averse and risk-seeking strategies. ESPs seek to maximize the expected profit with different scenarios ω by scheduling the battery operations Ξ, which can be expressed as:

$$\underset{\Xi}{Max} \sum_{\omega} \sum_{t} \pi_{\omega} \, \varphi_{t,\omega} \tag{4.11}$$

where $\pi_\omega$ denotes the probability of scenario ω. The strategy in (4.11) can consider all possibilities and is thus called the risk-neutral strategy. However, conservative DER participants prefer stable profits and avoid risk. The risk-averse VPP plan, which can manage its risks in the worst-case scenarios, can be described as the follows:

$$\underset{\Xi}{Max} (1 - \beta_r) \sum_{\omega} \sum_{t} \pi_{\omega} \, \varphi_{t,\omega} + \beta_r \eta_r \tag{4.12}$$

where $\beta_r$ denotes the risk-averse degree parameter, and a larger $\beta_r$ indicates that the participant is more risk-averse. $\eta_r$ represents the conditional value at risk, a risk-averse measurement [11]. Aggressive participants are willing to take higher risks to obtain higher profits, which is called the risk-seeking VPP plan and can be described as follows:

$$\underset{\Xi}{Max} (1 - \beta_s) \sum_{\omega} \sum_{t} \pi_{\omega} \, \varphi_{t,\omega} + \beta_s \eta_s \tag{4.13}$$

where $\beta_s$ denotes the risk-seeking degree parameter, and $\eta_s$ represents the value at the best scenarios, a risk-seeking measurement [97].

$\eta_r$ and $\eta_s$ involve a confidence parameter, which is set to 0.90. It is worth noting that the above three risk control methods are just examples for the following simulations in this chapter, and the methods used for real-world VPP optimization scheduling are diverse.

### 4.3.2 Fair Profit Allocation

ESPs and DERs assemble VPPs in the recruitment–participation approach, where there is no profit allocation rule maker. ESPs and all DERs need to agree on an allocation method, and the best allocation method is based on their respective contributions [44]. Therefore, a fair profit distribution based on DER individual contribution instead of DER group is proposed, which contains two steps. The first step is profit allocation among ESPs and PV and battery groups with SV, and the second is the internal reallocation for individuals in PV and battery groups. Defining $X/\langle i\rangle$ to be the set of all parts after removing part $i$, the marginal contribution of part $i$ to a coalition $S$, $S \subseteq X/\langle i\rangle$, is

$$\epsilon_i(S) = v(S \cup \{i\}) - v(S) \tag{4.14}$$

where $v(S)$ represents the total payoffs that the members of $S$ can obtain by cooperation. The SV for each part in VPP can be calculated as follows:

$$\phi_i = \sum_{S \subseteq X/\langle i\rangle} \frac{|S|!(n-|S|-1)!}{n!} \epsilon_i(S) \tag{4.15}$$

where $n$ represents the total number of parts and is set to 3 since the VPP is divided into three parts: PVs, batteries and ESP. The existing studies generally allocate profits to PV and battery groups rather than to real individual owners. To motivate DER individuals to participate in VPP, a reasonable internal group profit allocation method for PV and battery owners, $\delta_{PV-j}$ and $\delta_{battery-k}$, based on actual contributions, is proposed as follows:

$$\delta_{PV-j} = \phi_{PV} \frac{\sum_{t=1}^{T} \rho_t^{RT} P_{t,j}^{PV}}{\sum_{j=1}^{J} \sum_{t=1}^{T} \rho_t^{RT} P_{t,j}^{PV}} \tag{4.16}$$

$$\delta_{battery-k} = \phi_{battery} \frac{E_k}{\sum_{k=1}^{K} E_k} \tag{4.17}$$

$$\delta_{ESP} = \phi_{ESP} \tag{4.18}$$

where $\rho_t^{RT}$ represents the wholesale real-time electricity price at time $t$; $P_{t,j}^{PV}$ is the available output of the *j-th* PV owner at time $t$, and $J$ represents the total number of PV participants; $E_k$ denotes the actual capacity of the *k-th* available for VPP, and $K$ the total number of battery participants. For PV individuals, the payoff is related to the amount of generation available for VPP and the corresponding real-time electricity price, as in (4.16). In contrast, for batteries, the payoff is related to the capacity available for VPP, as in (4.17).

In addition to the SV-based profit allocation method with contributions, other allocation methods, such as investment shares-based, may lead to a fixed aggregation of VPP, which can impede the flexibility of assembly. Furthermore, non-contribution-based profit allocation can lead to the emergence of profit distribution rule makers, which conflicts with the proposed non-ESP-centered VPP assembly.

### 4.3.3 DER Participants' Payoffs Including Rebates

The ESP adopts the rebates to DER owners to reward them. The rebate coefficients are defined as $\lambda_{c,PV}$ and $\lambda_{c,battery}$ for PVs and batteries in OC $c$, respectively. The total payoff for individual PV and battery can be calculated as follows:

$$\zeta_{PV-j}=\delta_{PV-j}+\lambda_{c,PV}\delta_{ESP}^{PV-j} \tag{4.19}$$

$$\zeta_{battery-k}=\delta_{battery-k}+\lambda_{c,battery}\delta_{ESP}^{battery-k} \tag{4.20}$$

where $\zeta_{PV-j}$ and $\zeta_{battery-k}$ denote the final PV and battery payoffs, respectively, including the allocated payoff and rebate; $\delta_{ESP}^{PV-j}$ and $\delta_{ESP}^{battery-k}$ denote the ESP payoffs exploited from the *j-th* PV and *k-th* battery, respectively, which can be calculated with SV.

### 4.3.4 ESP's Payoff Excluding Rebates

The proposed VPP assembly approach involves multi-ESPs. $\delta_{ESP-m}$ is proposed to represent the actual source of the total payoff of ESP before rebate competition, as the follows:

$$\delta_{ESP-m} = \sum_{j=1}^{J} \delta_{ESP-m}^{PV-j} + \sum_{k=1}^{K} \delta_{ESP-m}^{battery-k} \tag{4.21}$$

After the *m-th* ESP provides the rebate $\lambda_c^m$ in OC $c$, the number of DER participants changes, and the total payoff of ESP changes to $\zeta_{ESP-m}$ as follows:

$$\zeta_{ESP-m} = \underbrace{\left(1-\lambda_{c,PV}^m\right)\sum_{j=1}^{J^{Rebate}} \delta_{ESP-m}^{PV-j}}_{\zeta_{ESP-m}^{PV}} + \underbrace{\left(1-\lambda_{c,battery}^m\right)\sum_{k=1}^{K^{Rebate}} \delta_{ESP-m}^{battery-k}}_{\zeta_{ESP-m}^{battery}} \tag{4.22}$$

where $J^{Rebate}$ and $K^{Rebate}$ represent the numbers of PV and battery participants after the rebate competition, respectively. The first part $\zeta_{ESP-m}^{PV}$ represents the payoffs exploited from PVs, and the second part $\zeta_{ESP-m}^{battery}$ represents the payoffs from batteries. $J^{Rebate}$ is defined as the following:

$$J^{Rebate} = \mu \frac{\Gamma^m}{\sum_{m=1}^{M} \Gamma^m} N^{PV} = \frac{\Gamma^m}{\sum_{m=1}^{M} \Gamma^m} N^{PV} \left( -\frac{1}{\frac{\delta_{PV-j}+\lambda_{c,PV}^m \delta_{ESP}^{PV-j} - \Theta_{PV}^{w/o}}{\Theta_{PV}^{w/o}} * a^{PV} + b^{PV}} + 1 \right) \tag{4.23}$$

where $K^{Rebate}$ has the same structure as $J^{Rebate}$.

The best rebate response function of each ESP can be drawn by taking the first-order derivative of $\zeta_{ESP-m}$ with respect to $\lambda_{c,PV}^m$ and $\lambda_{c,battery}^m$. Since $\zeta_{ESP-m}^{battery}$ has the same structure with $\zeta_{ESP-m}^{PV}$, only the first-order derivative of $\zeta_{ESP-m}^{PV}$ is presented with respect to $\lambda_{c,PV}^m$ for brevity as follows:

$$\frac{\partial \zeta_{ESP-m}^{PV}}{\partial \lambda_{c,PV}^m} = \frac{\partial \left( \frac{\left(1-\lambda_{c,PV}^m\right)\Gamma^m N^{PV} \delta_{ESP-m}^{PV-j}}{\sum_{m=1}^{M} \Gamma^m} \left( -\frac{1}{\frac{\delta_{PV-j}+\lambda_{c,PV}^m \delta_{ESP}^{PV-j} - \Theta_{PV}^{w/o}}{\Theta_{PV}^{w/o}} * a^{PV} + b^{PV}} + 1 \right) \right)}{\partial \lambda_{c,PV}^m} \tag{4.24}$$

Let $A = \frac{\Gamma^m}{\sum_{m=1}^{M} \Gamma^m}$ , $B = N^{PV} \delta_{ESP-m}^{PV-j}$ , $C = a^{PV} \delta_{PV-j} + (b^{PV} - a^{PV}) \Theta_{PV}^{w/o}$ , $D = a^{PV} \delta_{ESP}^{PV-j}$ , and $E = a^{PV} \delta_{PV-j} + (b^{PV} - a^{PV} - 1) \Theta_{PV}^{w/o}$ . Equation (4.24) can be rewritten as follows:

$$\frac{\partial \zeta_{ESP-m}^{PV}}{\partial \lambda_{c,PV}^{m}} = AB\,\frac{(D\lambda_{c,PV}^{m})^2+2CD\lambda_{c,PV}^{m}+EC+ED-CD}{(D\lambda_{c,PV}^{m}+C)^2} = 0 \tag{4.25}$$

By setting the first-order derivative to zero, we can get the best rebate $\lambda_{c,PV}^{m}$ as follows:

$$(\lambda_{c,PV}^{m})^{*} = \frac{-\begin{pmatrix} a^{PV}\delta_{PV-j}+ \\ (b^{PV}-a^{PV})\Theta_{PV}^{w/o} \end{pmatrix} + \sqrt{\Theta_{PV}^{w/o}\begin{pmatrix} a^{PV}\left(\delta_{PV-j}+\delta_{ESP}^{PV-j}-\Theta_{PV}^{w/o}\right) \\ +b^{PV}\Theta_{PV}^{w/o} \end{pmatrix}}}{a^{PV}\delta_{ESP}^{PV-j}} \tag{4.26}$$

Since $\delta_{PV-j} + \delta_{ESP}^{PV-j}$ denote the all ESP payoff, which must be larger than $\Theta_{PV}^{w/o}$, $a^{PV}\left(\delta_{PV-j} + \delta_{ESP}^{PV-j} - \Theta_{PV}^{w/o}\right) > 0$ and $b^{PV}\Theta_{PV}^{w/o} > 0$ can be readily obtained, which means $(\lambda_{c,PV}^{m})^{*}$ must exist. Since $\zeta_{ESP-m}^{battery}$ has the same structure as $\zeta_{ESP-m}^{PV}$, $(\lambda_{c,battery}^{m})^{*}$ must exist too.

# 4.4 Rebate Competition with MADDPG Algorithm

## 4.4.1 Markov Game Model for Multi-ESP Competition

In a multi-agent environment, the rebate decision-making by each ESP is influenced by the joint actions of all the ESPs. A Markov decision process (MDP) model [58, 98, 99] for $I$ ESPs can be formalized as the follows:

$$\tau = < \mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{P} > \tag{4.27}$$

where $\mathcal{S}$, $\mathcal{A}$, $\mathcal{R}$ and $\mathcal{P}$ denote the state space, action space, reward space, and transition probability, respectively. At time $t$, ESP $m$ observes a state $s_m^t \in \mathcal{S}$ and choose an action $a_m^t \in \mathcal{A}$ based on a certain policy, which produces a new state $s_m^{t+1}$. The transition probability from $s_m^t, a_1^t, a_2^t, \cdots, a_I^t$ to $s_m^{t+1}$ can be represented as $\mathcal{P}(s_m^{t+1} | s_m^t, a_1^t, a_2^t, \cdots, a_I^t)$. In this work, $\mathcal{S}$ contains the numbers and returns of DER participants for each ESP, VPP profits for each ESP, historical rebate coefficients of each ESP, historical VPP size of each ESP, other public information can be obtained from prospectus. In addition, historical half-hourly electricity prices are included in $\mathcal{S}$.

$\mathcal{A}$ represents rebate coefficients $\lambda_{t,PV}^{m}$ and $\lambda_{t,battery}^{m}$ for each ESP at time $t$. For our work, the objective for each ESP is to maximize the expected accumulated payoffs by finding the optimal policy. The payoff model for each model in (4.21) can be re-formulated as:

$$Max \sum_{t\in T} \gamma^{T-t} \zeta_{ESP-m_t}(\lambda_{t,PV}^{m}, \lambda_{t,battery}^{m}) \tag{4.28}$$

where $\gamma$ represents the reward discount factor, and $T$ represents the total number of OC.

### 4.4.2 Multi-Agent Deep Deterministic Policy Gradient

In the MADDPG algorithm, multiple agents are involved, and the centralized training and distributed implementation mode [100] are used. For agents, ESPs in our work, the objective of reinforcement learning is to maximize the expected reward with the policy gradient algorithm by adjusting the parameter $\theta$ of policy $\boldsymbol{\pi}$ at the direction of $\nabla_{\theta} J(\theta)$, as:

$$\nabla_{\theta} J(\theta) = \mathbb{E}_{s\sim\rho^{\pi}, a\sim\pi_{\theta}}[\nabla_{\theta} log\pi_{\theta}(a|s) Q^{\pi}(s, a)] \tag{4.29}$$

where the policy $\pi_{\theta}$ is stochastic, which may cause large computational consumption. In [99], the authors proposed a deterministic policy gradient, which considers a deterministic policy $\boldsymbol{\mu}_{\theta}: \mathcal{S} \rightarrow \mathcal{A}$, to integrate over the state space. The deterministic policy gradient can be derived as:

$$\nabla_{\theta} J(\theta) = \mathbb{E}_{s\sim p^{\mu}}[\nabla_{\theta} \mu_{\theta}(s) \nabla_{a} Q^{\mu}(s, a)|_{a=\mu_{\theta}(s)}] \tag{4.30}$$

Compared with a policy gradient, the deterministic policy gradient can reduce the computational complexity and improve the learning efficiency. For a game with $I$ agents, $\boldsymbol{\mu}_{\theta} = \{\mu_{\theta_1}, \mu_{\theta_2}, \cdots, \mu_{\theta_I}\}$ is parameterized by $\boldsymbol{\theta} = \{\theta_1, \theta_2, \cdots, \theta_L\}$, which is a deterministic policy set for all agents. The gradient of the expected return for each agent can be obtained by:

$$\nabla_{\theta_m} J(\mu_{\theta_m}) = \mathbb{E}_{\boldsymbol{o}, \boldsymbol{a}\sim\mathcal{D}}\left[\nabla_{\theta_m} \mu_{\theta_m}(a_m|s_m) \nabla_{a_m} Q_m^{\boldsymbol{\mu}}(\boldsymbol{o}, a_1, a_2, \cdots, a_I)|_{a_m=\boldsymbol{\mu}_{\theta_m}(s_m)}\right] \tag{4.31}$$

where $\boldsymbol{o}$ denotes the observation information, which generally includes the states of all agents; $Q_m^{\boldsymbol{\mu}}(\boldsymbol{o}, a_1, a_2, \cdots, a_I)$ denotes a centralized action-value function; $\mathcal{D}$ denotes the experience replay buffer, which stores the experience of all agents, expressed as $(\boldsymbol{o}, \boldsymbol{o}', a_1, a_2, \cdots, a_I, r_1, r_2, \cdots, r_I)$. $\mathcal{D}$ can be transmitted to the critic network for centralized updating with the back-propagation approach by minimizing the loss function:

$$L(\theta_m) = \mathbb{E}_{\boldsymbol{o},\boldsymbol{o}',a,r}[(Q_m^{\boldsymbol{\mu}}(\boldsymbol{o}, a_1, a_2, \cdots, a_I) - y)^2] \tag{4.32}$$

$$y = r_m + \gamma Q_m^{\boldsymbol{\mu}'}(\boldsymbol{o}', a'_1, a'_2, \cdots, a'_I)|_{a'_m=\boldsymbol{\mu}'(\boldsymbol{o}'_m)} \tag{4.33}$$

The parameters $\theta_m^{\mu}$ and $\theta_m^{Q}$ are manipulated by the soft-update approach as below:

$$\theta_m^{\mu}{}' = \varrho\theta_m^{\mu} + (1-\varrho)\theta_m^{\mu}{}' \tag{4.34}$$

$$\theta_m^{Q}{}' = \varrho\theta_m^{Q} + (1-\varrho)\,\theta_m^{Q}{}' \tag{4.35}$$

### 4.4.3 MADDPG-based Multi-ESPs Rebate Competition

The MADDPG-based multi-ESPs rebate competition with the key models is given in Algorithm 4.2.

**Algorithm 4.2** Simulation of Multi-ESP Competition with MADDPG

**Input:** Structures of actor $\boldsymbol{\mu}$, critic $\boldsymbol{Q}$ networks, and initialized weights $\boldsymbol{\theta}^{\boldsymbol{\mu}}$ and $\boldsymbol{\theta}^{\boldsymbol{Q}}$.
**Output:** Actor networks' weights $\boldsymbol{\theta}_m^{\boldsymbol{\mu}}{}'$ for each ESP
1:**For** each episode:
2: **For** each training step:
3: **For** each OC:
4: **For** each **ESP agent $\boldsymbol{m}$**:
5: Actor of each ESP agent generates rebate action for current OC: $a_m^{ESP}$=$[\lambda_{PV}^m, \lambda_{battery}^m]$.
6: **End**
7: ESPs publish prospectus which contains all rebate actions $\boldsymbol{a}^{\boldsymbol{ESP}} = [a_1^{ESP}, a_2^{ESP}, \cdots, a_I^{ESP}]$.
8: DERs select ESPs using (4.1 and 4.2) and Algorithm 4.1. Each ESP obtain actual total numbers of PV and battery participants.
9: **For** each day in OC:
10: Obtain predicted PV generation and electricity price with the CTSGAN in [101].
11: VPPs operated by different ESPs with different risk control strategies using (4.11)-(4.13).
12: Payoffs allocation for each DERs and ESPs using (4.19)-(4.21) for one

day.

**End**

ESPs obtain the reward $\mathcal{R}$ for OC using (4.22) and new states $\boldsymbol{o}_{new}$.

Store transition sample $(\boldsymbol{o}, \boldsymbol{a}, \mathcal{R}, \boldsymbol{o}_{new})$ into $\mathcal{D}$.

Sample a random minibatch of samples $(\boldsymbol{o}^d, \boldsymbol{a}^d, \mathcal{R}^d, \boldsymbol{o}_{new}^d)$ from $\mathcal{D}$.

Set $y^d = \mathcal{R}^d + \gamma Q_m^{\boldsymbol{\mu}'}(\boldsymbol{o}_{new}^d, a'_1, a'_2, \cdots, a'_I)|_{a'_m=\boldsymbol{\mu}'(\boldsymbol{o}_m)}$.

Update critic by minimizing the loss function:

$L(\theta_m) = \frac{1}{N}\sum_{d=1}^{N}\left(y^d - Q_m^{\boldsymbol{\mu}}(\boldsymbol{o}_{new}^d, a_1, a_2, \cdots, a_I)^2\right)$

Update actor network using the sampled policy gradient:

$\nabla_{\theta_m} J = \frac{1}{N}\sum_{d=1}^{N}\left(\nabla_{\theta_m}\mu_{\theta_m}(\boldsymbol{o}_m^d)\nabla_{a_m} Q_m^{\boldsymbol{\mu}}(\boldsymbol{o}^d, a_1, a_2, \cdots, a_I)|_{a_m=\boldsymbol{\mu}_{\theta_m}(\boldsymbol{o}_m^d)}\right)$

Update target network parameters for each agent $m$:

$\theta_m^{\mu}{}' = \varrho\theta_m^{\mu} + (1-\varrho)\theta_m^{\mu}{}'$

$\theta_m^{Q}{}' = \varrho\theta_m^{Q} + (1-\varrho)\,\theta_m^{Q}{}'$

**End**

**End**

**End**

# 4.5 Case Studies

## 4.5.1 Experiment Settings

It is assumed that there are 5,600 households in an area, of which 4,200 are installed with PVs and 1,400 with batteries. In addition, 35 experienced ESPs are involved in this area and wish to assemble VPPs for participation in the energy market. ESPs publish prospectus for the relevant plans. The PV generation data are from Ausgrid (a major Australian power company) [102], and the energy market electricity prices for VPP participation are from the AEMO from July 2021 to June 2022 [103]. Table 4.2 lists the parameters for the 2,800 household batteries.

Table 4.2 Parameters of batteries

| $E^m$ | $\eta^{ch}$ | $\eta^{dch}$ | $P^m$ |
|---|---|---|---|
| 7kWh | 0.95 | 0.95 | 3.5kW |

In this section, several cases are studied to simulate the proposed approach for VPP assembly, disassembly, and competition as follows:

**VPPs Assembly.** It is assumed that only 20% of DER initially participate in assembling VPPs. Table 4.3 assumes seven common risk control strategies, including three risk-averse, three risk-seeking, and one risk-neutral, the corresponding risk degree parameters $\beta_r$ and $\beta_s$, and the number of ESPs, PVs and batteries. The first 10 OCs are set as closed periods, and the PV and battery participants cannot exit.

Table 4.3 Initial VPP assembly settings

| | Risk-Averse | | | Risk-Neutral | Risk-Seeking | | |
|---|---|---|---|---|---|---|---|
| Risk Degree | 0.3 | 0.2 | 0.1 | N.A. | 0.1 | 0.2 | 0.3 |
| ESP Number | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| PV Number | 120 | 120 | 120 | 120 | 120 | 120 | 120 |
| Battery Number | 40 | 40 | 40 | 40 | 40 | 40 | 40 |

**VPPs Competition.** ESP attracts more DER participants through rebates. The willingness of PV and battery owners to participate in VPP is based on (4.1). The involved coefficients $a^{PV}$, $b^{PV}$, $a^{battery}$, and $b^{battery}$ are assumed as 20, 1.25, 10, and 1.25, respectively [51]. The coefficients of the participants' preferences are assumed as listed in Table 4.4. In addition, we assume a standard normal distribution for all types of DERs, with the proportions of professional, ambitious, moderate, conservative, and cautious participants in order of 7%, 24%, 38%, 24%, and 7%.

Table 4.4 Coefficients of participant groups

| Types | $[\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6]$ |
|---|---|
| Professional | [0.30, 0.25, 0.15, 0.15, 0.10, 0.05] |
| Ambitious | [0.10, 0.50, 0.15, 0.10, 0.10, 0.05] |
| Moderate | [0.30, 0.15, 0.20, 0.20, 0.10, 0.05] |
| Conservative | [0.25, 0.10, 0.25, 0.25, 0.10, 0.05] |
| Cautious | [0.25, 0.05, 0.30, 0.25, 0.10, 0.05] |

It is noted that the coefficients are assumed based on the description in Table 4.1, which may not reflect the actual situation. Even for the same investment preference, the weights may be influenced by various factors, such as social development, education level, and local laws and regulations, leading to changes in the weights. Therefore, to ensure the effectiveness of the VPP competition model presented in this work, the relevant coefficients should be well-designed based on actual data before applying the model.

This work employs two base cases are employed for the VPP assembly stage (Case 1) and the VPP competition stage (Case 2).

**Case 1: VPP Assembly**

**Case 1.a:** ESPs initially assemble VPPs with recruited DER owners. VPPs with poor strategies will be disassembled.

**Case 1.b:** VPP profits are allocated to each DER participants.

**Case 1.c:** ESPs earn payoffs by helping DER participants increase their payoffs.

**Case 2: VPP Competition**

**Case 2.a:** ESPs provide rebates to attract DER participants, and the final state of competition can be stable with MADDPG.

**Case 2.b:** Changes in rebate competition with large electricity prices fluctuations.

**Case 2.c:** Changes in rebate competition with different participation related coefficients.

### 4.5.2 Simulation Results

#### 4.5.2.1 VPP Initial Assembly

**a) Comparison of VPP Profits with Multi-ESPs**

Thirty-five ESPs assemble 35 VPPs with 840 PVs and 280 batteries with multiple risk degree levels. The first 10 OC periods, from July 5 to September 12, 2021, are set as closed period. The multi-ESPs first use day-ahead prediction of electricity prices and PV generation and then apply an optimization algorithm to optimize the battery scheduling for profit maximization based on the chosen risk level. For simplicity, the prediction method used is CTSGAN [101], and the optimization method is the PSO algorithm. We set different hyperparameters for CTSGAN and PSO to simulate the performance of different ESPs. Fig. 4.3 illustrates the average profit, nominal risk degree level, and actual profit fluctuations for each ESP in 10 OCs. The sign A$ on the coordinate axis represents the Australian dollar. Dots of the same color represent the same risk degree level.

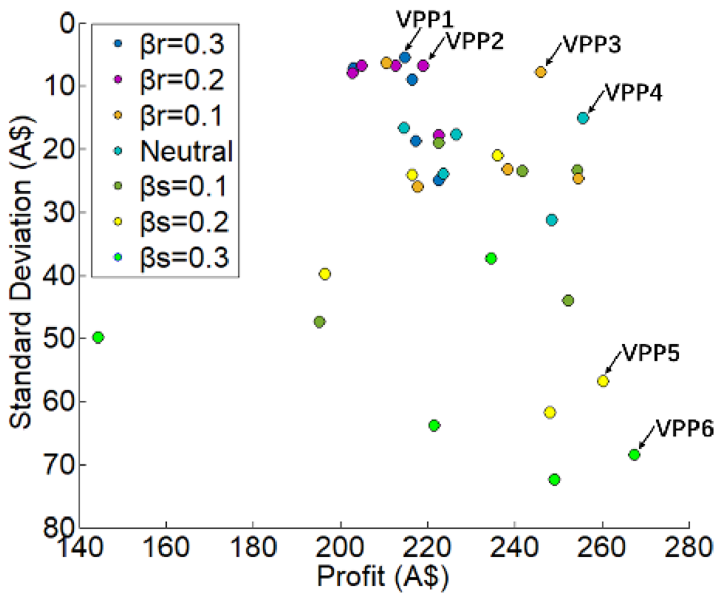



Fig. 4.3 Profit and standard deviation for 35 VPPs

One can find that all ESPs have different performances; only 6 VPPs lie on the Pareto frontier of standard deviation and total profit, including three risk-averse VPPs with $\beta_r$=0.3 (VPP1 in Fig. 4.3), $\beta_r$=0.2 (VPP2), and $\beta_r$=0.1 (VPP3), one risk-neutral VPP (VPP4), and two risk-seeking VPPs with $\beta_s$=0.2 (VPP5) and $\beta_r$=0.3 (VPP6), respectively. The average daily profits for risk-averse VPPs are approximately between A$215 and A$245 and are stable with a standard deviation of less than A$7. The profits for risk-seeking VPPs are slightly higher and more volatile, with a large standard deviation of more than A$55. The remaining VPPs have poor performance, which means that their DERs will choose other ESPs at the end of the closed period, which leads to the disassembly of the VPP. Note that the performance of the five risk-seeking VPPs with $\beta_s$=0.1 simulated in this chapter is poor, but this does not mean that $\beta_s$=0.1 is inappropriate for any VPPs.

**b) Comparison of VPP Profits Allocation for Each Participant**

Take VPP1-VPP6 in Fig. 4.3 as an example. The total profit is allocated to ESPs and PV and battery groups using SV based on the contribution. Fig. 4.4 illustrates the average payoff for ESP, PV and battery groups. The PV group's payoff is stable at around A$138, while the payoff of the battery group increases from A$61.09 to

A$89.93 as the risk increases. The ESP payoff is similar to that of the battery group, meaning that the payoff comes mainly from the arbitrage of battery charging and discharging.

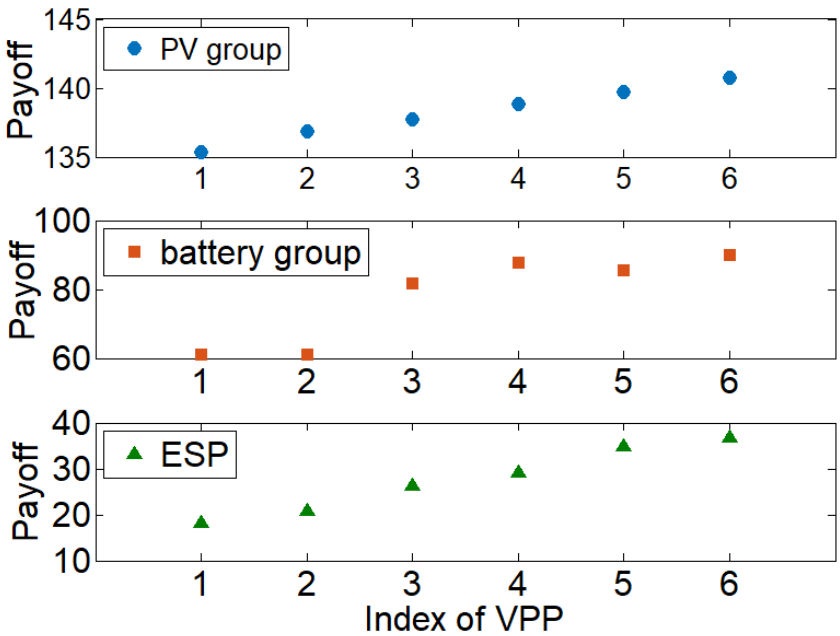



Fig. 4.4 Average payoff for ESP, PV and battery groups based on SV

Table 4.5 Average daily payoff for each participant (A$/Day)

| | VPP1 | VPP2 | VPP3 | VPP4 | VPP5 | VPP6 | Retailer-led |
|---|---|---|---|---|---|---|---|
| PV | 1.128 | 1.140 | 1.156 | 1.157 | 1.173 | 1.189 | 1.080 |
| Battery | 1.527 | 1.530 | 2.043 | 2.191 | 2.142 | 2.249 | 1.456 |
| ESP | 18.221 | 20.828 | 26.403 | 29.094 | 34.999 | 36.809 | NA. |

Based on the above DER group payoff, the payoff allocation of the individuals in PV and battery groups can be calculated based on (4.16)-(4.17). Table 4.5 provides the average daily payoffs for PV and battery participants in the first ten cycles with the proposed VPP assembly model and those of current Australian existing retailer-led VPPs [104]. It can be seen that the daily payoff of individual PV participants is stable between A$1.12 and A$1.19, while battery payoff varies from A$1.5 to A$2.2 for different VPPs. In addition, the payoff of individual PV is improved by at least 4.44% (VPP1) and at most 10.09% (VPP 6) compared to participation in the existing VPP. The battery participant's payoff increment is more obvious, increasing by 54.46% for VPP6. The payoff received by ESPs according to their contributions differs widely from A$18.221 to A$36.809. Fig. 4.5 presents the actual payoff of PV and battery participants for 70 consecutive days. It can be seen that the daily payoff for PV fluctuates very little, even for the risk-seeking VPP, such as VPP 6, with fluctuations ranging between A$1.18 and A$1.22 in the 7th OC. In contrast, battery payoff is significantly more volatile for the same VPP 6 and OC, ranging from A$2.11 to A$2.65.

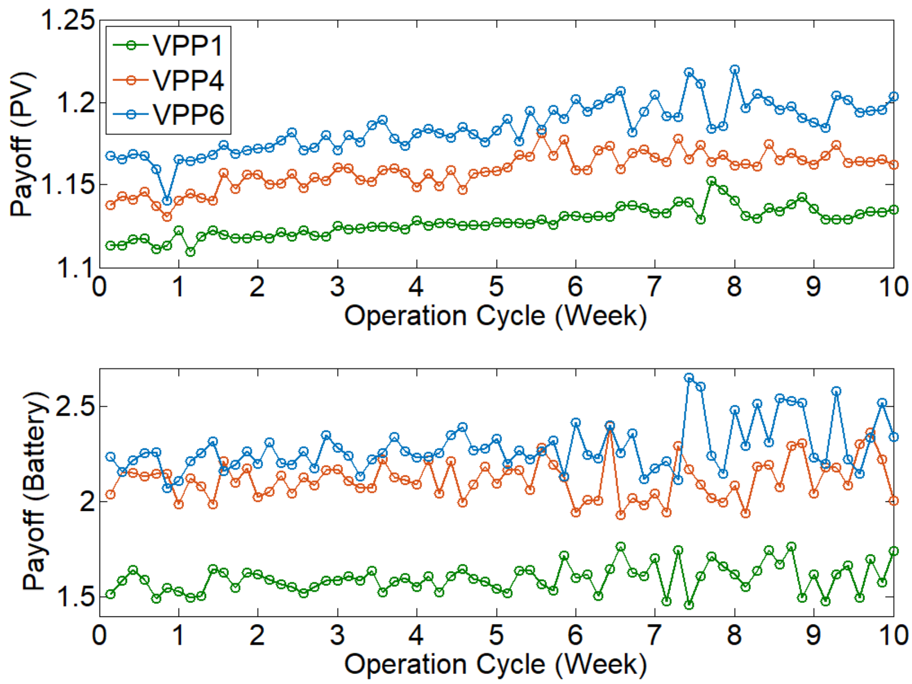



Fig. 4.5 Payoff for PV and battery individuals for VPP1, VPP4 and VPP 6

### c) Payoff Sources for VPP ESPs

Although ESPs do not make money directly by generating electricity or providing energy storage, ESPs get returns through dispatch, which are coming from each PV and battery participant. Table 4.6 lists the average daily payoff for ESP earned from each PV and battery individual for the first 10 OCs. One can find that the payoffs for ESPs provided by PVs are small and stable, with less than A$0.05 per PV per day. Payoffs from batteries account for a significant portion of ESPs and increase significantly with the increase in total VPP profits, from A$0.372 to A$0.833 per battery per day.

Table 4.6 ESPs' payoffs earned from each PV and battery (A$/Day,10E-2)

| | VPP1 | VPP2 | VPP3 | VPP4 | VPP5 | VPP6 |
|---|---|---|---|---|---|---|
| PV | 4.764 | 4.794 | 5.828 | 5.831 | 5.869 | 5.910 |
| Battery | 37.273 | 43.700 | 57.524 | 64.245 | 78.891 | 83.293 |

#### 4.5.2.2 VPP Rebate Competition

By the end of the closed period, all rational participants will leave those ESPs with high volatility and low profit, resulting in the majority of the total VPPs being disassembled and only 6 VPPs remaining. This section will explore the rebate competition process within the remaining VPPs as examples. Although new ESPs or eliminated ESPs can reassemble VPPs, we focus on simulating rebate competition and therefore do not consider the case of assembling new VPPs.

#### a) Simulation for VPP Rebate Competition

The MADDPG method is used to simulate the rebate competitive process for 20 OCs between September 13, 2021 and January 31, 2022. Fig. 4.6 presents the training process of rebate competition for each ESP.

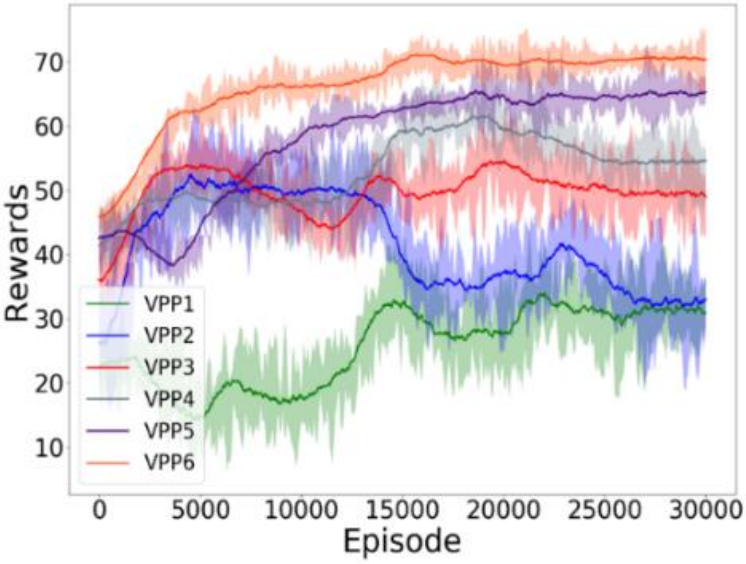



Fig. 4.6 Training process of the rebate competition

Table 4.7 ESPs rebate competition results (%)

| | VPP1 | VPP2 | VPP3 | VPP4 | VPP5 | VPP6 |
|---|---|---|---|---|---|---|
| $\lambda_{PV}$ | 26.96 | 25.90 | 23.28 | 18.14 | 17.68 | 13.42 |
| $\lambda_{battery}$ | 38.01 | 36.64 | 34.75 | 32.46 | 26.50 | 24.85 |

One can find that the ESP of VPP6 is the first to reach a stable payoff, followed by the ESPs for VPP5, VPP4, and VPP3. The competition between ESP1 and ESP2 is more intense; therefore, the reward does not stabilize until around episode 27,000. Table 4.7 tabulates the rebate competition results. One can find that ESPs give lower rebates to PVs than batteries, partly because PVs bring lower profits to ESPs, and partly because PV participants are more likely to be incentivized to participate in VPPs by low rebates. In addition, the high-risk-return VPP gives a low rebate coefficient, with values of 13.42% ($\lambda_{PV}^{6}$) and 24.85% ($\lambda_{battery}^{6}$) for VPP6, while those of the low-risk ones are high, with values of 36.64% and 38.01% for battery groups in VPP2 and VPP1. The reason is mainly the higher profitability of ESP for VPP6, where a small rebate coefficient can lead to a large actual rebate. Therefore, due to the increased risk and unstable profitability, providing a large rebate coefficient would result in extremely low profitability for ESP in VPP6 in specific cases.

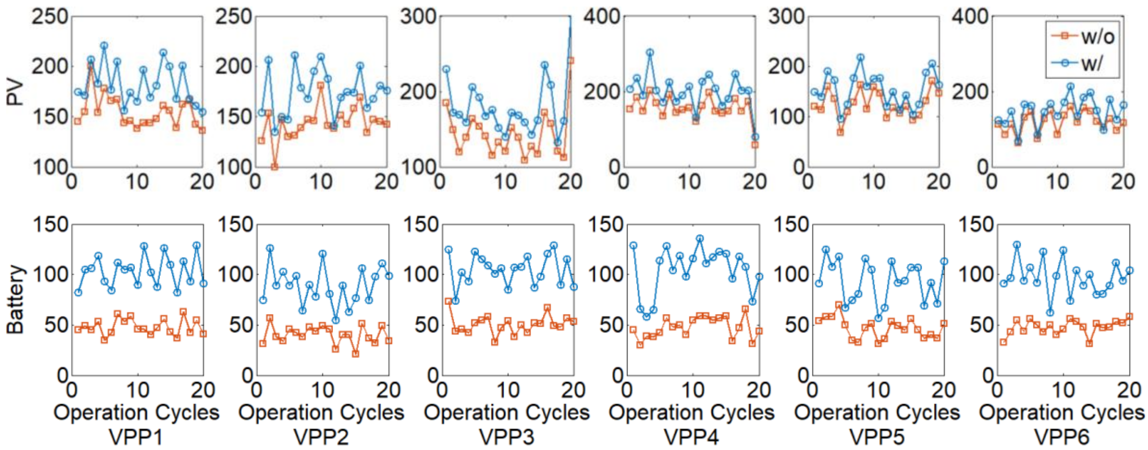


Fig. 4.7 Change in the number of PV and battery participants for 20 OCs

With the rebates provided by ESPs, DER participants can re-select the VPP operation before each cycle since the change of return, payoff stability, and VPP scale with (4.2). In addition, the total number of DERs increases due to payoff increment obtained from rebates. Fig. 4.7 shows the change in PV and battery participants for each VPP. The number of participants in each VPP increases. The average number of PV owners who participate in VPPs with rebates is 1.23 times higher than that without rebates; in contrast, the number of battery participants with rebates is 2.11 times higher than that without rebates, from 46.83 to 98.23.

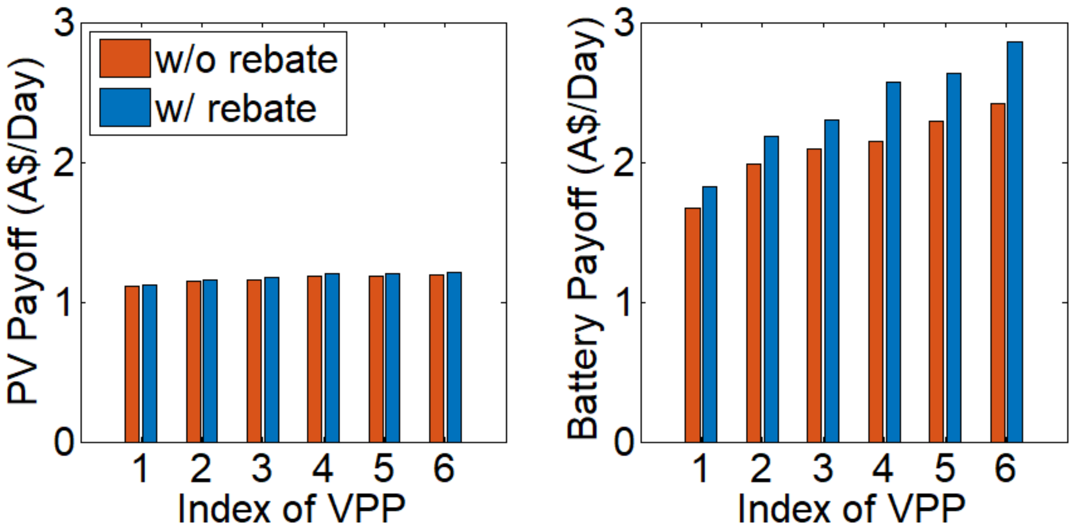


Fig. 4.8 Daily payoff for PV and battery participants

Table 4.8 Comparison of ESPs' daily payoff with and without rebate competition (A$/Day)

| | ESP1 | ESP2 | ESP3 | ESP4 | ESP5 | ESP6 |
|---|---|---|---|---|---|---|
| W/O Rebate | 26.861 | 26.619 | 36.377 | 40.196 | 45.989 | 49.629 |
| W/ Rebate | 31.451 | 32.774 | 46.737 | 53.870 | 63.283 | 71.491 |
| Improve | 0.171 | 0.231 | 0.285 | 0.340 | 0.376 | 0.440 |

Fig. 4.8 presents the daily payoffs for DERs for the 20 OCs with or without rebates. It can be seen that the payoffs of PV participants rise slightly with rebates compared to those without rebates, by approximately 0.62% to 1.79%. In contrast, the payoffs of battery participants increase more significantly, with a minimum of 9.28% in VPP1, from A$1.672 to A$1.826, and a maximum of 18.25% in VPP6, from A$2.428 to A$2.854. Table 4.8 lists the average payoff for ESPs. Although ESPs offer a portion of their payoffs through rebates to attract more participants, each ESP's daily payoff is increased significantly. The low-risk VPP1, with a small rebate, attracts only a limited number of participants, so the daily payoff increases from A$26.861 to A$31.451. In contrast, the high-risk VPP6 sees a more significant increment in daily payoff, with an improvement of 44.0% from A$49.629 to A$71.491.

**b) Impact of Electricity Price Fluctuations on Rebate Competition**

Price fluctuations have a direct relationship with VPP profitability. In this section, the prices between April 4 and June 12, 2022, are used to study the impact of electricity price fluctuations on the multi-ESP VPP competition. During this period, significant unplanned outages, challenges in the coal and natural gas supply chains, and the early onset of winter have combined to cause significant volatility in electricity prices [105], leading AEMO to temporarily suspend the Wholesale Market on June 15 [106]. Table 4.9 lists the daily profits of the participants before the rebate competition and the final rebate coefficients for each VPP. One can find that fluctuating electricity prices can bring higher payoffs to DER participants than those with normal electricity prices, where the payoff increment of batteries is significantly higher than that of PVs. For example, in VPP6, the average payoff for the battery individual increases from A$2.249 (normal) to A$6.392 (fluctuating), while that for PV individuals increases from A$1.189 (normal) to A$1.272 (fluctuating). In addition, there is no significant change in the battery rebate coefficients given by the high-risk-return ESPs compared to those with normal electricity prices ($\lambda_{PV}, \lambda_{battery}$ in Table 4.7), while that given by VPP1 increases significantly, from 38.01% to 52.16%. The ESP selection of participants with different preferences also changes with price fluctuation. Fig. 4.9 presents the changes in the number of participants with different preferences for normal and fluctuating electricity prices.

Table 4.9 Participant's average daily payoff (A$/Day) and ESP's rebate coefficients (%)

| | VPP1 | VPP2 | VPP3 | VPP4 | VPP5 | VPP6 |
|---|---|---|---|---|---|---|
| PV | 1.195 | 1.217 | 1.233 | 1.249 | 1.257 | 1.272 |
| Battery | 3.621 | 4.033 | 4.165 | 4.630 | 5.031 | 6.392 |
| $\lambda_{PV}$ | 27.93 | 24.91 | 23.04 | 19.70 | 16.92 | 13.53 |
| $\lambda_{battery}$ | 52.16 | 43.84 | 36.25 | 33.96 | 25.83 | 26.73 |

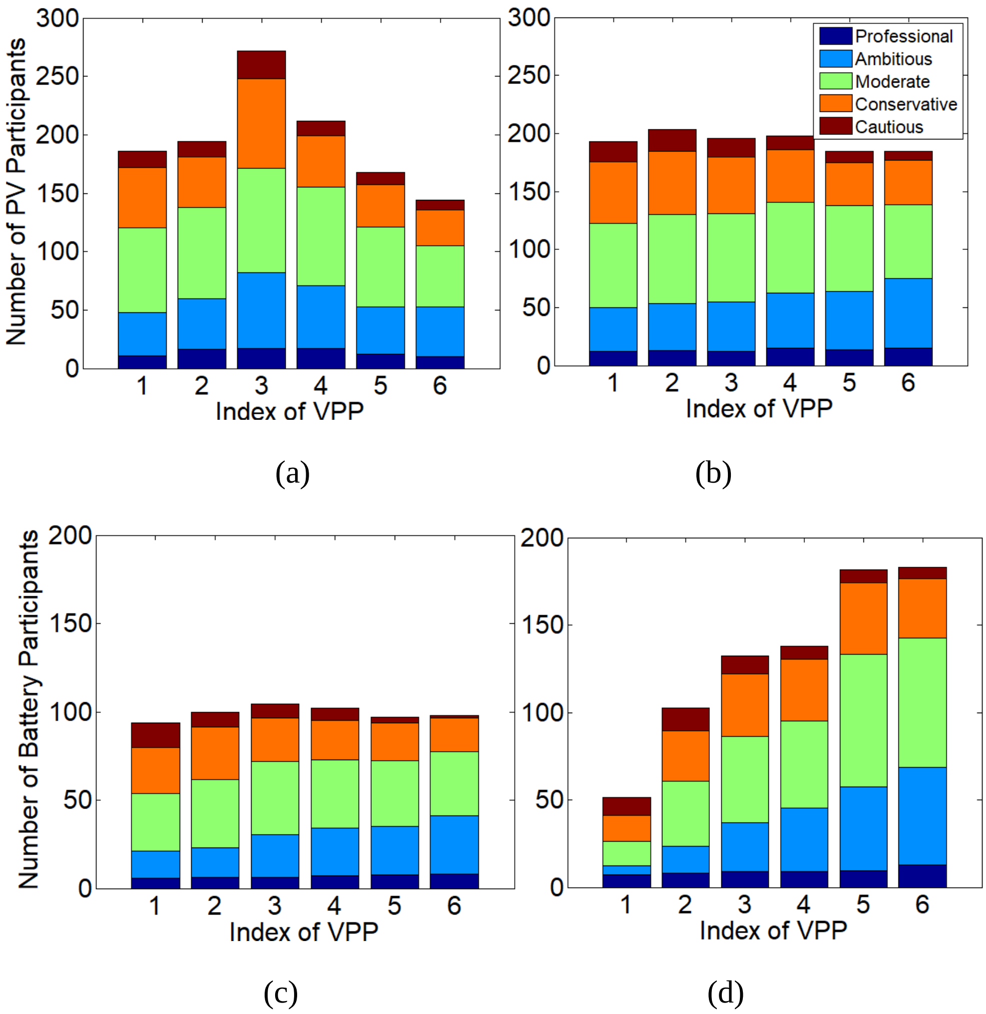



Fig. 4.9 Changes of operation selection for different participant groups for normal (a, c) and fluctuating (b, d) electricity price

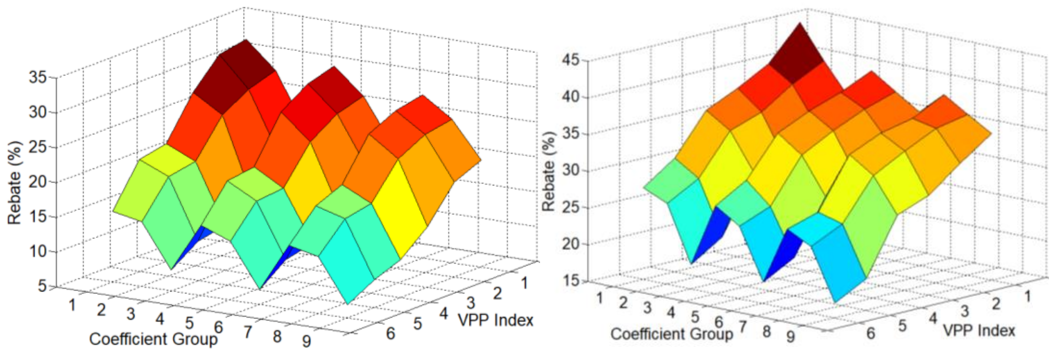
For PV participants, it can be seen that with normal prices, ESP3 recruits the highest number of participants and is about 100 more than ESP6. When the prices are volatile, ESP5 and ESP6 have significantly more participants, around 180. Ambitious participants prefer ESP3 with normal electricity prices, while ambitious ones are more inclined to ESP6 with high fluctuating prices. In addition, the change in the number of battery participants is more pronounced. The number of battery participants for each VPP is consistent with the normal electricity prices; however, that of VPP1 participants decreases significantly to 50 with the fluctuating prices, while the numbers of VPP5 and VPP6 participants are almost doubled. Conservative and cautious participants, driven by high returns, see a significant increase in participation in VPP5 and VPP6.

### c) Impact of VPP Participation Rate-Related Coefficients on Rebate Competition

In the above section, the VPP participation rate-related coefficients $a^{PV}$, $a^{battery}$, $b^{PV}$ and $b^{battery}$ are assumed to be 20, 10, 1.25, and 1.25 respectively. As there is currently a lack of literature describing the accurate relationship between VPP participation rates and payoff improvement, to explore further the response of different DER owners to payoff improvement, we assume the participation rates related coefficient group as shown in Table 4.10, which includes different $b^{PV}, b^{battery}$ pairs ranging from 1.11 to 1.43, $a^{PV}$ ranging from 10 to 40, and $a^{battery}$ ranging from 5 to 20.

Table 4.10 Assumed coefficients in the relationship between VPP participation rate and payoff increment

| Coefficients Group | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Initial Participation | | 10% | | | 20% | | | 30% | |
| $b^{PV}, b^{battery}$ | | 1.11 | | | 1.25 | | | 1.43 | |
| $a^{PV}$ | 10 | 20 | 40 | 10 | 20 | 40 | 10 | 20 | 40 |
| $a^{battery}$ | 5 | 10 | 20 | 5 | 10 | 20 | 5 | 10 | 20 |




(a) (b)

Fig. 4.10 Changes of rebate for PV (a) and battery (b) owners with assumed participation rate related coefficient group

Fig. 4.10 provides the changes of rebate coefficients for PV and battery with assumed coefficient groups in Table 4.10 using MADDPG. It can be observed that in (a), with the same initial participation rate, the rebate coefficients with larger $a^{PV}$ values significantly decrease compared to those with smaller $a^{PV}$ values. For instance, in VPP1, $\lambda_{PV}$ is 32.48% (coefficient group 1), 28.73% (coefficient group 2), and 21.21% (coefficient group 3). This phenomenon can be observed in different VPPs and in other coefficient groups, such as groups 4, 5, and 6. Moreover, it is found that for the same VPP, when the initial participation rate is high (groups 7, 8, and 9 with high $b^{PV}$), the difference in rebate coefficients is smaller than that when the $b^{PV}$ is low, with 28.45% (group 7) and 20.34% (group 9) compared to 32.48% (group 1) and 21.21% (group 3). Due to the low initial participation rate and the high remaining number of potential participants, ESP needs to increase the rebate coefficients to attract DER owners to participate, thereby increasing the number of participants and profits. These phenomena also exist in VPP competition for battery rebate. Compared with PV rebate competition, the fluctuation of battery rebates is more affected by VPP participation-related coefficients. For instance, in VPP1 with group 1, $\lambda_{battery}$ is 44.82%, while in VPP6, group 9, the $\lambda_{battery}$ is only 17.04%. Fig. 4.11 presents the average daily payoffs for ESPs with the above $\lambda_{PV}$ and $\lambda_{battery}$.

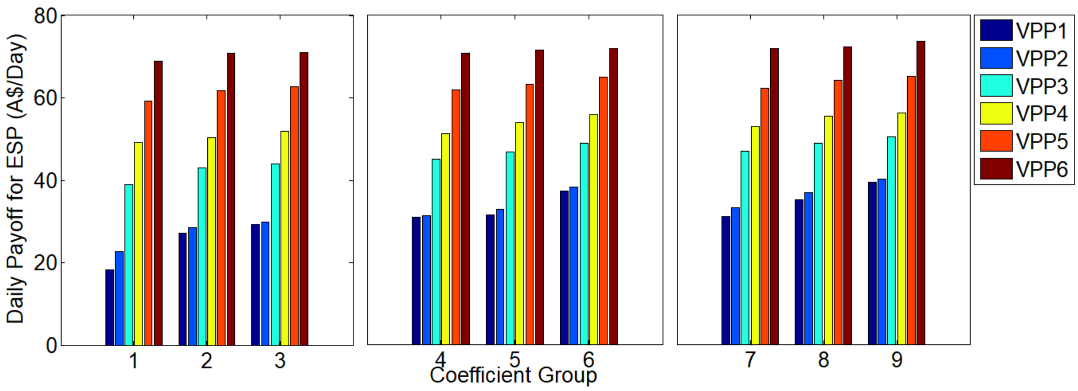



Fig. 4.11 Average daily payoffs for ESPs with assumed participation rate related coefficient group

Compared to the payoffs for 6 ESPs with group 6 (the payoffs in Table 4.8), when the ESPs face DER owners who are easily incentivized (group 7 with higher $a^{PV}$ and $a^{battery}$), the payoffs of each ESP increase, but the payoff for ESP1 increases the most, from A$31.45 to A$37.32, while that for ESP6 increases the least, just from A$71.49 to A$71.98. In addition, it can be seen that ESP6 is consistently able to maintain a higher and more consistent payoff when facing DER owners with coefficient group 1 to 9, ranging from A$68.93 to A$73.66. At the same time, while the return of ESP1 fluctuates greatly from A$18.21 to A$39.42. This means that ESP6 is always in a proactive position of concession in VPP competition, while ESP1 often needs to be forced to make concessions.

## 4.6 Chapter Summary

In this chapter, the ESPs were separated from the existing VPP models for the first time, and their payoffs can be allocated fairly based on their contributions to the whole VPP with SV. The recruitment–participation approach for the VPP assembly was proposed. The ESPs that meet DER preferences can remain, while the ESPs that DERs were not satisfied with will be eliminated. Moreover, this work introduced multiple ESPs into VPPs for the first time with reference to the existing multi-retailer competition model in the electricity market. Finally, the rebate competition of multi-ESPs was simulated for these DERs with diverse preferences and different electricity price fluctuations. The following conclusions are drawn: 1) High-risk-return ESPs can provide smaller rebate coefficients than low-risk-return ones; 2) ESPs can offer higher rebate coefficients to batteries compared to PVs; 3) both ESPs and DERs can improve payoffs after rebates; 4) DERs with different investment preferences has shown a larger difference under different scenarios of electricity price fluctuations.

# CHAPTER 5

# A Profitable and Flexible Distributed Energy Resource Recruitment Approach in Virtual Power Plant for Unconventional Arbitrage Opportunities

*A VPP is typically a collection of DERs aggregated by an ESP. However, recruiting DER owners to participate in a VPP is challenging. Therefore, we propose a profitable and flexible VPP recruitment-participation approach that incorporates both long-term regular recruitment and short-term casual recruitment. Casual recruitment caters to ambitious DER participants, consisting of fair and bet-on modes. The latter establishes a set of pre-determined payoff conditions, the fulfillment or non-fulfillment of which confers the participants a contractual right to get compensation from the ESP. To ensure the success of the proposed recruitment approach, we address two key problems. First, we introduce a new index, unconventional arbitrage opportunity (UAO), for evaluating future profits and propose a CTSGAN to predict UAO with weather conditions. Second, we introduce a payoff allocation method that combines fairness and incentives to motivate casual DER participants. The incentive coefficients are optimized using an improved DRL algorithm.*

## 5.1 Flexible Recruitment Framework

An ESP typically aggregates conventional VPP with a fixed number of DER participants for long-term involvement in the electricity market to gain profits. However, this model lacks flexibility and has limited profitability. For ESP, even if a high UAO is predicted for the future, ESP cannot immediately increase the number of DER participants to gain higher profits. As for DER participants, some wish to be involved in more flexible VPPs to achieve short-term profits rather than committing to the long-term.

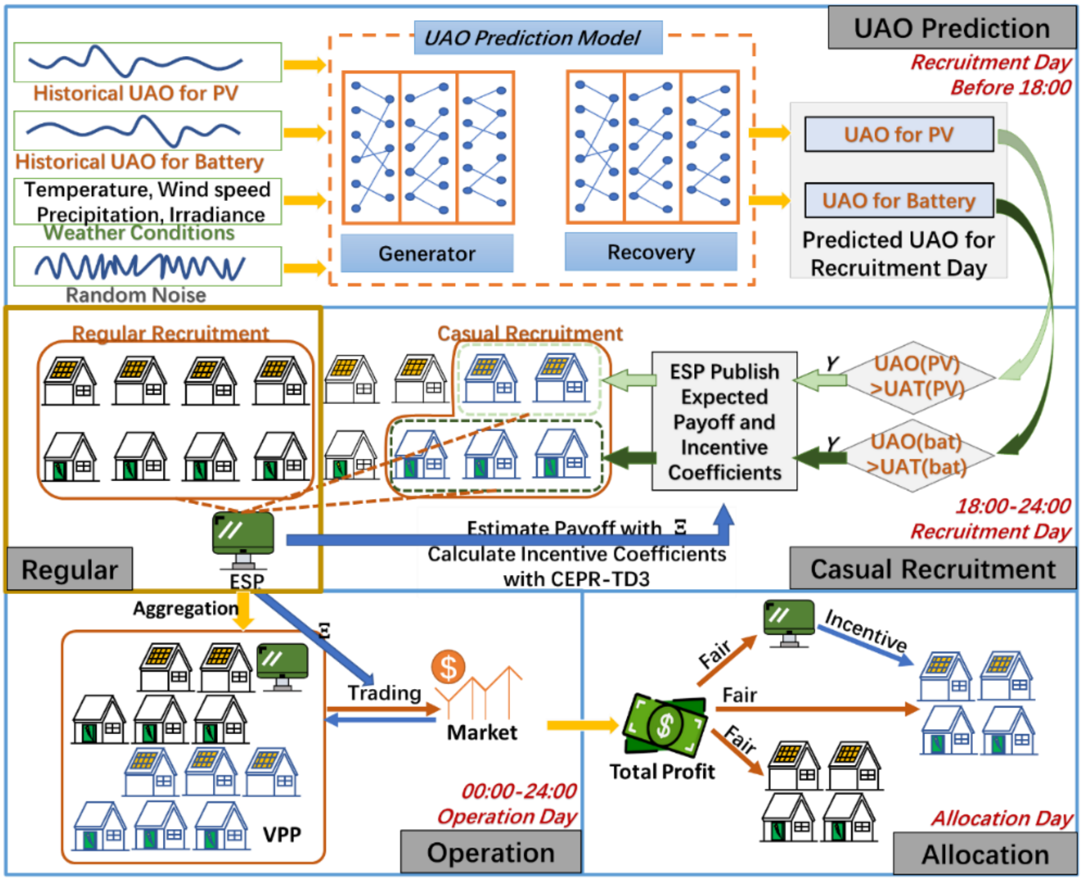



Fig. 5.1 Framework of the proposed DER participant recruitment approach

Therefore, a flexible and profitable DER participant recruitment approach is proposed in this chapter, as illustrated in Fig. 5.1. The entire process is divided into four steps as follows:

**Step 1 UAO Prediction**

The first step involves UAO prediction for PV and battery units on recruitment day. The prediction model is CTSGAN with a historical UAO dataset for PVs and batteries, weather conditions (temperature, wind speed, precipitation, and irradiance), and random noise as the input. The output is a UAO value for PVs and batteries for the recruitment day. ESP predicts the UAO to determine the need for casual recruitment. This step occurs before 18:00 on the recruitment day (the day before the operation day). If a significant UAO is forecasted, casual recruitment is conducted for DER owners who have not participated in VPP aggregation.

**Step 2 Casual Recruitment**

ESP compares the predicted UAO value with the UAO threshold (UAT), and if the UAO exceeds the UAT, the ESP will publish the incentive coefficients and expected payoffs to attract casual participants. Then, the casual DER participants are combined with the long-term regular recruitment ones to aggregate the VPP. Otherwise, there would only be regular and no casual recruited participants. It should be noted that since the UAO for PV and battery is predicted separately, the casual recruitment of PV and battery is also conducted independently. This step occurs from 18:00 to midnight on the recruitment day.

**Step 3 VPP Operation**

On the operational day (the day after recruitment day), the ESP manages all recruited PV and battery participants as a unified VPP to participate in the electricity market to increase profits.

**Step 4 Profit Allocation**

On the following day (allocation day), the profits are then allocated to the participants, which utilizes the SV method for regular recruitment participants and a fair and incentivized allocation method for casual recruitment participants. Afterward, the casual recruited VPP section will be disassembled, and only the long-term recruitment section will remain operational. The ESP continues with predictions, awaiting the next UAO, and casual recruitment will be conducted again.

The above DER recruitment process is based on the following three assumptions: (a) The DER owners decide whether to participate in a VPP solely based on the historical payoffs; (b) All necessary devices for the VPP, such as inverters and smart meters, are pre-installed. And the ESP does not incur additional expenses for casual participants; (c) The operational costs of the VPP under varying numbers of participants are all borne by the ESP.

## 5.2 DER Recruitment Approach and Payoff Allocation

A multiple-recruitment approach for VPP aggregation is proposed in this chapter, including regular and casual recruitments. Regular recruitment refers to aggregating existing VPPs, as discussed in [4, 5, 11]. Regular recruitment typically aims to attract conservative DER participants, and the cycle for regular recruitment is usually long, often measured in months or even years. Compared to regular recruitment, casual recruitment has a much shorter cycle, often measured in days. Participants in these short-term recruitments usually expect to receive higher returns in a short period of time. Fig. 5.2 illustrates the multiple-recruitment-mode-based approach for DER participant recruitment.

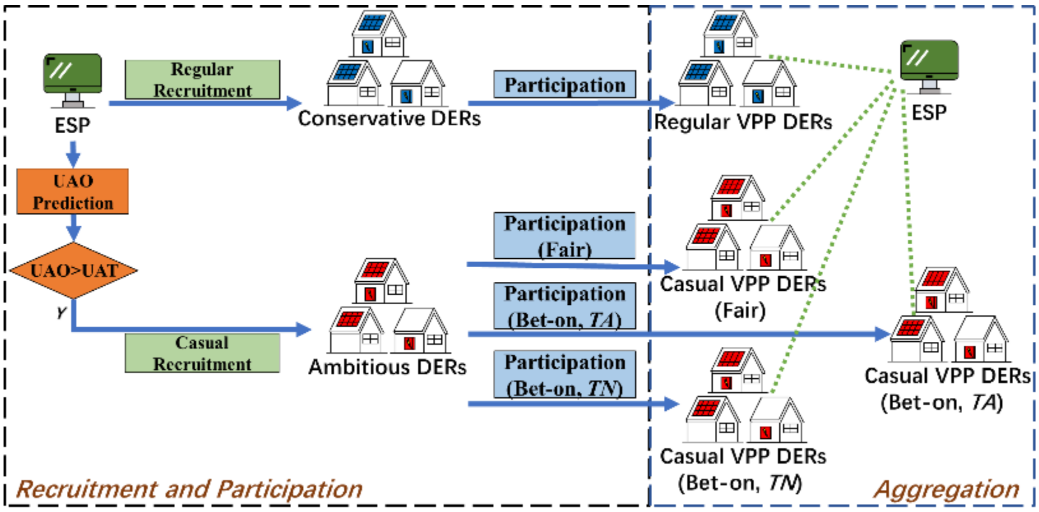



Fig. 5.2 Multiple-recruitment-mode-based approach for DER recruitment

### 5.2.1 Regular Recruitment and Conventional Payoff Allocation

Regular recruitment is currently the most commonly used method for aggregating VPPs and has been widely discussed in previous studies. Regular recruitment typically aims to attract conservative DER participants seeking a stable and consistent income over the long term and generally aversing to taking on risks. The ESP recruits conservative DER participants over the long term and distributes the profits between the ESP and each individual participant through the SV method [42]. Defining $X/\langle i \rangle$

to be the set of all parts after removing part$i$, the marginal contribution of part $i$ to a coalition $S$, $S \subseteq X/\langle i \rangle$, is

$$\epsilon_i(S) = v(S \cup \{i\}) - v(S) \quad (5.1)$$

where $v(S)$ represents the total payoffs that the members of $S$ can obtain by cooperation. The SV for each part in a VPP can be calculated as follows:

$$\phi_i = \sum_{S \subseteq X/\langle i \rangle} \frac{|S|!(n-|S|-1)!}{n!} \epsilon_i(S) \quad (5.2)$$

where $n$ represents the total number of participants, including PVs, batteries, and ESP. With the SV method, the payoffs for the $j^R$*-th* PV and the $k^R$*-th* battery can be obtained as $\phi_{PV-j^R}^{Regular}$ and $\phi_{bat-k^R}^{Regular}$, respectively.

### 5.2.2 Casual Recruitment and Incentive Payoff Allocation

Price-responsive models are often used in demand response scenarios [107], or for internal pricing mechanisms [45] within a VPP, but they have not yet been applied to the VPP formation stage. Therefore, price-responsive models are incorporated into the VPP formation process for the first time, and this approach is named "casual recruitment." Casual recruitment is targeted at ambitious participants who wish for high short-term returns and have risk tolerance. Unlike regular recruitment, casual recruitment participants are not inclined to relinquish control of their PVs or batteries to the ESP for an extended period. They only respond when they anticipate higher profits and engage with the VPP on a short-term basis.

The ESP publishes the incentive coefficients and expected payoffs to recruit ambitious PV or battery participants only when the ESP predicts a UAO for the recruitment day. Casual recruitment can be divided into two categories: fair mode and bet-on mode, the latter being riskier and allowing DER owners' participation with or without consensus.

It should be noted that conventional price-responsive models are usually risk-free. The incentive prices provided by ESPs are often fixed rather than expected. These types of price incentive triggering mechanisms are usually simple, such as triggering when the electricity wholesale price exceeds a certain threshold and stopping when it returns to normal. However, in VPPs, price incentives are subject to many uncertainties, such as wholesale prices, market behavior, weather conditions, and trading strategies. Additionally, due to the long time required for battery charging and discharging, it is not possible to have risk-free price incentives like demand response. Therefore, the proposed casual recruitment mode we propose allocates profits while also distributing risks, as detailed as follows:

#### 5.2.2.1 Fair Mode

In the fair mode, DER participants are not required to express their approval of the expected profit offered by the ESP, and they are exposed to risks consistent with those of regular recruitment. Additionally, casual participants are eligible for some incentives provided by the ESP. The total payoffs for the $j^F$*-th* PV and the $k^F$*-th* battery can be denoted as $\phi_{PV-j^F}^{fair}$ and $\phi_{bat-k^F}^{fair}$, defined as follows, respectively:

$$\phi_{PV-j^F}^{fair} = \phi_{PV-j^F}^{Casual} * (1 + \nu^{PV}) \tag{5.3}$$

$$\phi_{bat-k^F}^{fair} = \phi_{bat-k^F}^{Casual} * (1 + \nu^{bat}) \tag{5.4}$$

where $\phi_{PV-j^F}^{Casual}$ and $\phi_{bat-k^F}^{Casual}$ denote the payoffs for the $j^F$*-th* PV and the $k^F$*-th* battery with (5.1) and (5.2); $\nu^{PV}$ and $\nu^{bat}$ denote the incentive coefficients in the fair mode for casual recruitment, which are non-negative dependent parameters.

Under this model, apart from the incentives, payoffs are jointly distributed between the ESP and all DER participants, while the ESP and all DER participants collectively bear risks. In addition, the ESP further allocates a portion of its profits as incentive subsidies to casual participants, thereby increasing the total number of participants and ultimately boosting the ESP's overall profits. However, in fair mode, the incentive coefficients for casual participants are generally not large.

#### 5.2.2.2 Bet-on Mode

In fair mode, the incentive coefficients for casual participants are generally not large. Too large incentive coefficients would lead to a decline in the overall profits of the ESP. Additionally, residential users are often price-insensitive [108], so they don't respond strongly to minor price incentives in fair mode. To cater to the needs of various types of participants, we propose an alternative participation method for VPPs, which is more high-return but also high-risk compared to the fair mode, termed "bet-on" mode.

In the bet-on mode, two sub-modes are available for DER participants to choose from: consensus sub-mode and non-consensus sub-mode. In the consensus sub-mode, DER participants who believe that the announced profit target can be achieved are denoted as ***TA***; Otherwise, they are denoted as ***TN***. Upon completion of its daily operation, if the actual profit fulfills the expected target, incentives will be awarded to the participants who choose ***TA***; Otherwise, compensation will be awarded to the ones who choose ***TN***. In addition to the incentives, every participant is given a payoff based on the SV method to prevent the bet-on agreement from becoming a gamble.

More incentives can be gained if the correct choice is made. The total payoffs for the $j^B$-*th* PV and the $k^B$-*th* battery can be denoted as $\phi_{PV-j^B}^{bet-on}$ and $\phi_{bat-k^B}^{bet-on}$, defined as follows respectively:

$$\phi_{PV-j^B}^{bet-on} = \begin{cases} \phi_{PV-j^B}^{Casual} \times (1+\lambda_{PV}^{TA}), & if\ \boldsymbol{TA} \curvearrowright TA \\ \phi_{PV-j^B}^{Casual} \times (1+\lambda_{PV}^{TN}), & if\ \boldsymbol{TN} \curvearrowright TN \\ \phi_{PV-j^B}^{Casual}, if\ \boldsymbol{TA} \curvearrowright TN\ or\ \boldsymbol{TN} \curvearrowright TA \end{cases} \tag{5.5}$$

$$\phi_{bat-k^B}^{bet-on} = \begin{cases} \phi_{bat-k^B}^{Casual} \times (1+\lambda_{bat}^{TA}), & if\ \boldsymbol{TA} \curvearrowright TA \\ \phi_{bat-k^B}^{Casual} \times (1+\lambda_{bat}^{TN}), & if\ \boldsymbol{TN} \curvearrowright TN \\ \phi_{bat-k^B}^{Casual}, if\ \boldsymbol{TA} \curvearrowright TN\ or\ \boldsymbol{TN} \curvearrowright TA \end{cases} \tag{5.6}$$

where $\phi_{PV-j^B}^{Casual}$ and $\phi_{bat-k^B}^{Casual}$ denote the payoffs for the $j^F$-*th* PV and the $k^F$-*th* battery with SV, and $\lambda_{PV}^{TA}$, $\lambda_{PV}^{TN}$, $\lambda_{bat}^{TA}$ and $\lambda_{bat}^{TN}$ denote the incentive coefficients in the bet-on mode for casual recruitment, respectively. The bold italic $\boldsymbol{TA}$ and $\boldsymbol{TN}$ before the symbol $\curvearrowright$ in (5.5) and (5.6) represents the sub-mode participants select, and the italic $TA$ and $TN$ after $\curvearrowright$ represent the actual results; $\lambda_{PV}^{TA}, \lambda_{PV}^{TN}, \lambda_{bat}^{TA}$, and $\lambda_{bat}^{TN}$ are non-negative dependent parameters.

### 5.2.3 Designed Modes and Corresponding Risks and Returns

We design various recruitment modes to meet participants' different risks and return preferences. The following figure provides a schematic representation of the risk and return for each mode. It should be noted that this chart is solely for illustrating the relative relationships between risks and returns and does not represent the actual risks and returns.

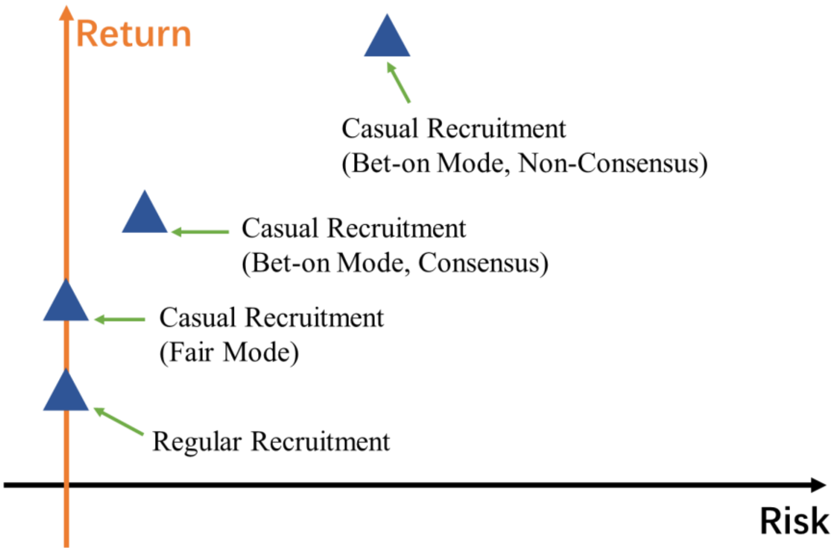



Fig. 5.3 Risk-return for regular and casual participants

Fig. 5.3 provides a schematic representation of the return and risk for participants. It can be found that regular recruitment is risk-free, but the corresponding returns are limited. Casual recruitment in fair mode is a price-responsive recruitment method that offers higher returns than conventional recruitment, with the same level of risk but fewer chance of casual recruitment. In contrast, the bet-on mode of casual recruitment increases risks compared to the two mentioned modes, but also offers increased returns. Under the consensus sub-mode, participants agree with ESP's strategies and anticipated profits. Since ESP has expertise in VPP operations, the risks are significantly lower than in the non-consensus sub-mode, although the returns are also less than in the non-consensus sub-mode. In the non-consensus mode, participants are essentially betting against ESP and those in the consensus mode, leading to significantly increased risks and returns.

The incentive coefficients, $\nu^{PV}$ and $\nu^{bat}$, in the fair mode and the incentive coefficients, $\lambda_{PV}^{TA}, \lambda_{PV}^{TN}, \lambda_{bat}^{TA}$ and $\lambda_{bat}^{TN}$, in the bet-on mode should be well optimized to encourage DER owners to participate in the VPP. Another issue that needs to be addressed is how to choose the appropriate time for casual recruitment; therefore, we propose a UAO scenario prediction method in the next section.

# 5.3 UAO Scenario Prediction by CTSGAN Model

The ESP can use the casual recruitment approach to enable ambitious participants to capture UAO with significantly higher profits than the conventional ones. The definition of UAO is proposed in this section. Moreover, the ESP should develop a predictive model for the day-ahead UAO.

## 5.3.1 Conventional Model for VPP Profit

Assume the VPP scheduling strategy is Ξ, and the total profit $\varphi_{t,\omega}$ of VPP for the energy market at time $t$ with predicted scenario ω is calculated as follows:

$$\varphi_{t,\omega}(\Xi) = \rho_{t,\omega_\rho}^{RT} P_t^{VPP} \tag{5.7}$$

$$P_t^{VPP} = P_{t,\omega_{PV}}^{PV} - \mu_t^{ch} P_t^{ch} + \mu_t^{dch} P_t^{dch} \tag{5.8}$$

$$\mu_t^{ch} + \mu_t^{dch} \leq 1 \tag{5.9}$$

$$P_{t,\omega_{PV}}^{PV}, P_t^{ch}, P_t^{dch} \geq 0 \tag{5.10}$$

where $\rho_{t,\omega_\rho}^{RT}$ represents the probabilistic real-time electricity price scenario, and $P_t^{VPP}$ represents the output of VPP; $P_{t,\omega_{PV}}^{PV}$, $P_t^{ch}$, and $P_t^{dch}$ denote PV generation power, battery charging, and discharging power, respectively; The two battery power components, $P_t^{ch}$ and $P_t^{dch}$, are non-negative values less than the rated power capacity, $P^m$, and are limited by binary variables, $\mu_t^{ch}$ and $\mu_t^{dch}$.

### 5.3.2 Unconventional Arbitrage Opportunity

Over an extended period of time, the average profit yielded by implementing the aforementioned conventional model with (5.7)-(5.10) may be referred to as the conventional profit. Nevertheless, numerous opportunities can bring profits that exceed conventional profits by taking advantage of the spot electricity price fluctuations. The unconventional arbitrage opportunity, or UAO, can be defined as follows:

$$UAO^{unit} = \frac{\sum_{t,d}^{T} \varphi_t^{unit}(\Xi)}{\frac{1}{D}\sum_d^D \sum_{t,d}^T \varphi_t^{unit}(\Xi))} \tag{5.11}$$

where $UAO^{unit}$ denotes UAO for each DER unit, which represents 1 kW PV panel ($UAO^{PV}$) and 1 kWh battery ($UAO^{bat}$); $\varphi_t^{unit}(\Xi)$ denotes the actual profit for a $unit$ at time $t$ in day $D$ with scheduling strategy Ξ; D is the number of days in the past. It should be noted that the occurrence of $UAO^{PV}$ and $UAO^{bat}$ may not coincide, as the former is influenced by the amount of electricity generated by PV panels and wholesale market prices. In contrast, the magnitude and duration of electricity price fluctuations determine the latter.

### 5.3.3 UAO Prediction with CTSGAN Model

The UAO prediction is more difficult than the electricity price prediction. Take $UAO^{bat}$ as an example. Even with the same RMSE as the predicted electricity price, $UAO^{bat}$ may be different because of the different error distributions. This chapter proposes a method to generate UAO scenarios directly instead of using the conventional method by predicting electricity prices and PV generation first and then calculating UAO. The training for UAO prediction is given in Fig. 5.4.

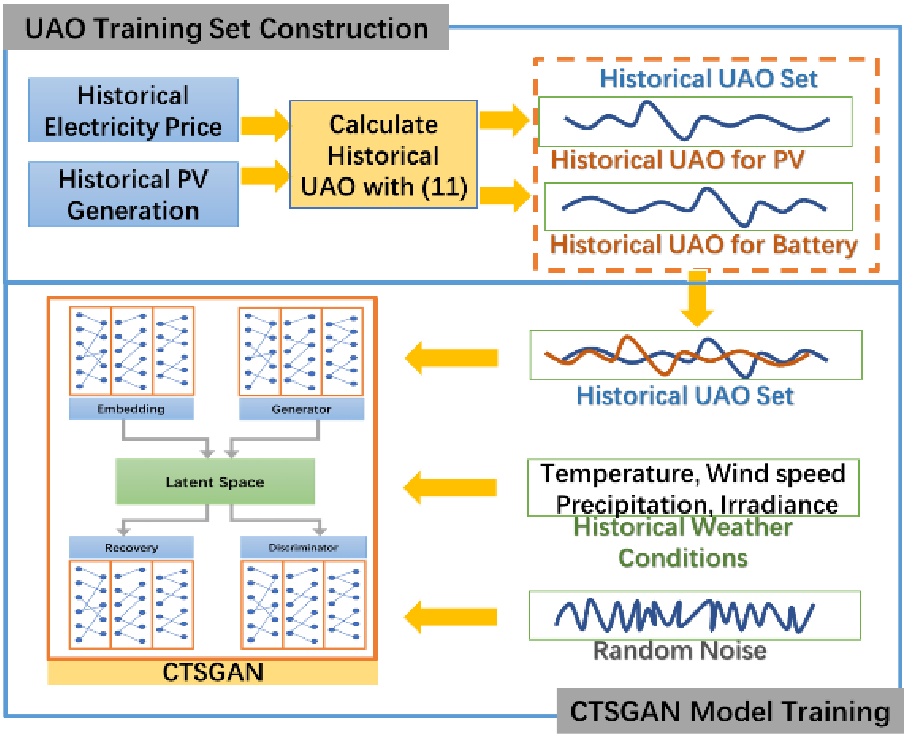



Fig. 5.4 UAO prediction model training

The first step is to prepare the UAO training sets for PVs and batteries. We collect historical data on electrical price and PV generation at interval of 30 minutes and then calculate the daily UAO with (5.11) for PVs and batteries, respectively. These daily UAO values serve as the UAO training set. The historical daily $UAO^{PV}$, $UAO^{bat}$, weather conditions (temperature, wind spend, precipitation and irradiance), and random noise are then applied to train the CTSGAN prediction models. Note that the CTSGAN models of $UAO^{PV}$ and $UAO^{bat}$ are trained separately. Once the CTSGAN models are successfully trained, they can be used to predict $UAO^{PV}$ and $UAO^{bat}$.

## 5.4 ESP's Profit Maximization and Dynamic Incentives Optimization with CEPR-TD3

The profit of ESP essentially comes from a commission earned by scheduling DERs to help them obtain high payoffs. This section first presents the profit target model of the ESP and the relationship model between DER owners' participation rate and payoffs. Finally, the modified DRL is used to optimize the incentive coefficients.

### 5.4.1 ESP' Profit Maximization

The payoff of an ESP depends on its contribution to the VPP's overall profit, which can be calculated using the SV method. The ESP's payoff can also be seen as a contribution-based draw after the ESP helps each DER to increase its individual payoff. To achieve the ESP's aim to attract more DER participants to the VPP, the ESP provides dynamic incentives to casual participants, whereby the ESP gives up a portion of their profits to increase the payoffs of DER owners. The ESP's aim to maximize its payoff can be expressed as:

$$MaxU^{ESP} = Max(\phi_{ESP} - \delta_{fair} - \delta_{bet-on} - \delta_{cost}) \tag{5.12}$$

$$\delta_{fair} = \sum_{j=1}^{J^F} \left(\phi_{PV-j^F}^{Casual} * \nu^{PV}\right) + \sum_{k=1}^{K^F} \left(\phi_{bat-k^F}^{Casual} * \nu^{bat}\right) \tag{5.13}$$

$$\delta_{bet-on} = \sum_{j=1}^{J_{PV}^{TA\sim TA}} \left(\phi_{PV-j^B}^{Casual} * \lambda_{PV}^{TA}\right) + \sum_{j=1}^{J_{PV}^{TN\sim TN}} \left(\phi_{PV-j^B}^{Casual} * \lambda_{PV}^{TN}\right) + \sum_{k=1}^{K_{bat}^{TA\sim TA}} \left(\phi_{bat-k^B}^{Casual} * \lambda_{bat}^{TA}\right) + \sum_{k=1}^{K_{bat}^{TN\sim TN}} \left(\phi_{bat-k^B}^{Casual} * \lambda_{bat}^{TN}\right) \tag{5.14}$$

$$\delta_{cost} = a^{PV} * J^2 + b^{PV} * J + c^{PV} + a^{bat} * K^2 + b^{bat} * K + c^{bat} \tag{5.15}$$

where $\phi_{ESP}$ denotes the SV-based ESP's payoff; $\delta_{fair}$ and $\delta_{bet-on}$ are the incentive costs for fair and bet-on mode participants, respectively. Additionally, the operational cost $\delta_{cost}$ for VPP with different numbers of participants is considered and incorporated into ESP's costs, and $a$, $b$, and $c$ are non-negative dependent parameters [46].

### 5.4.2 Willingness of DER Participants

DER owners exhibit varying levels of willingness to participate in the VPP based on different payoffs, and linear models cannot accurately represent their behavior. Diminishing marginal effect is a common phenomenon that can be observed in many systems in social sciences and economics systems, as it represents the gradual decrease in the effect of one variable on another variable as the degree of change increases. In addition, the diminishing marginal effect can be found in EV aggregation willingness and price incentive [52], PV installations and FIT [51], SES profit growth, and leasing market price [109]. In [50], Lu et al. presented a diminishing marginal effect model to describe the relationship between payoffs and VPP participation rates of DERs. We introduce this nonlinear model to describe the relationship between casual participation rate and historical payoffs as follows:

$$\mu = -\frac{1}{\varrho^{Part} \times a + b} + 1, \ where \ \ \varrho^{Part} = \varrho^{PV} \ or \ \varrho^{bat} \tag{5.16}$$

$$\varrho^{PV} = \frac{\Sigma_T\left(\left(\Sigma_{J^F}\phi_{PV-j^F}^{fair} + \Sigma_{J^B}\phi_{PV-j^B}^{bet-on}\right)/(J^F + J^B)\right)/T}{\Sigma_T\left(\left(\Sigma_{J^R}\phi_{PV-j^R}^{Regular}\right)/J^R\right)/T} - 1 \tag{5.17}$$

$$\varrho^{bat} = \frac{\Sigma_T\left(\left(\Sigma_{K^F}\phi_{bat-k^F}^{fair} + \Sigma_{K^B}\phi_{bat-k^B}^{bet-on}\right)/(K^F + K^B)\right)/T}{\Sigma_T\left(\left(\Sigma_{K^R}\phi_{bat-k^R}^{Regular}\right)/K^R\right)/T} - 1 \tag{5.18}$$

where $\mu$ represents the overall participation rate for DER owners; $\varrho^{PV}$ and $\varrho^{bat}$ denote the payoff gain rate for PV and battery participants, respectively, in the casual recruitment model compared to the regular model.

It is worth noting that the payoff of battery participants is more volatile than that of PV participants, and battery participants require larger incentives than PV participants due to battery degradation. In this study, $a$ is assumed as 0.5 and 2 for battery and PV, respectively, and $b$ is set as 1.

### 5.4.3 Dynamic Incentive Coefficients Optimization

The ESP needs to provide reasonable incentive coefficients to the casual DER participants. This process can be modeled as an MDP [58, 98], denoted as $< \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R} >$, where $\mathcal{S}$, $\mathcal{A}$ and $\mathcal{R}$ stand for the state, action, and reward, respectively. In this work, the observed environment state contains predicted $UAO^{PV}$ and $UAO^{bat}$, actual historical $UAO^{PV}$ and $UAO^{bat}$, expected payoffs for individual PV and battery, actual number of casual and regular participants, historical $\nu^{PV}$, $\nu^{bat}$, $\lambda_{PV}^{TA}$, $\lambda_{PV}^{TN}$, $\lambda_{bat}^{TA}$ and $\lambda_{bat}^{TN}$, and actual historical $U^{ESP}$. The action contains six incentive coefficients of $\nu^{PV}$, $\nu^{bat}$, $\lambda_{PV}^{TA}$, $\lambda_{PV}^{TN}$, $\lambda_{bat}^{TA}$, and $\lambda_{bat}^{TN}$. The reward represents $U^{ESP}/U^{init}$, where $U^{init}$ denotes the ESP's initial payoff with all incentive coefficients of 0. A modified TD3 with CEPR is proposed to provide optimized incentive coefficients for maximizing $U^{ESP}$.

#### 5.4.3.1 Twin Delayed Deep Deterministic Policy Gradient

TD3 is an upgraded version of DDPG, which is developed based on the architecture of the actor-critic algorithm. The state and action are represented by the critic and actor networks [110]. The critic network can be presented by a $Q$ function and be written as $y = r + \gamma Q_\theta(s', a')$, where $r$ denotes the reward and $\gamma$ the discount factor. The critic network can be updated by the following mean square error:

$$L = \frac{1}{N} \sum_t (Q_\theta(s, a) - y)^2 \tag{5.19}$$

where $N$ represents the size of mini-batches. The actor-network parameter $\phi$ can be updated by the deterministic policy gradient method as the following:

$$\nabla_\phi J \approx \frac{1}{N} \sum_t \left( \nabla_a Q(s, a)|_{s=(o_i), a=\pi_{\phi_i}(s)} \nabla_\phi \pi_\phi(s) \Big|_{s=o_i} \right) \tag{5.20}$$

With a double critic network, TD3 can overcome the overestimation of the $Q$ problem in DDPG. The target can be updated by the following two critic networks:

$$y_1 = r + \gamma Q_{\theta_1'}(s', a') \tag{5.21}$$

$$y_2 = r + \gamma Q_{\theta_2'}(s', a') \tag{5.22}$$

A minimum Q function is then selected from the above two critic networks to replace the conventional target in DDPG:

$$y = r + \gamma \min_{i=1,2} Q_{\theta_i'}(s', a') \tag{5.23}$$

#### 5.4.3.2 Combined Experience Pool Replay -TD3

The conventional TD3 algorithm incorporates an experience replay buffer pool to store transfer experience, from which the agent samples a mini-batch of transitions to train the actor and critic networks. This process, known as experience replay, can eliminate sample correlation and improve sample utilization rate. However, it is important to note that not all empirical samples have the same impact on the training process. Randomly selecting samples with equal probability may result in low-value and even valueless samples being included in the training process, potentially leading to suboptimal results in TD3 training. To address the problem, prior research proposed the multiple experience pool replay (MEPR) strategy [111] and the multi-level experience pool replay (MLEPR) strategy [112] to select prioritized samples for updating networks in TD3. To address the specific challenge of our proposed task, it is necessary to construct multiple individual experience replay pools for experiences under similar predicted UAO since different UAO predictions are decisive for participants' behavior. Therefore, a CEPR strategy is proposed to update the networks based on the priority of empirical samples, including similar and high-value ones.

The CEPR strategy (as illustrated in Fig. 5.5 incorporates multiple independent buffer pools, each labeled with a different UAO, or $\left\{UAO^1, \cdots, UAO^j, \cdots, UAO^J\right\}$. During the network initialization process, the average reward of the two pools with different UAO labels is set to 0. Whenever a sample with the same UAO label is added, its reward value is recalculated. This reward value is then compared to the average reward value for the same UAO label. If the reward value is less than the average, the sample is added to pool 1; Otherwise, it is added to pool 2. During the training, samples are randomly selected from pools 1 and 2 under each UAO label with probabilities $\xi$ and $1-\xi$, respectively, and are combined into a new experience buffer pool. Then, the real $UAO^j$ label is obtained corresponding to the current training data, and the samples from each combined pool are obtained to reconstruct the new pool with such a probability that $UAO^j$ accounts for 50%, $UAO^{j-1}$ and $UAO^{j+1}$ for 12.5%, $UAO^{j-2}$ and $UAO^{j+2}$ for 3.125%, respectively, and so on. Finally, the mini-batch samples are randomly selected from the reconstructed pool.

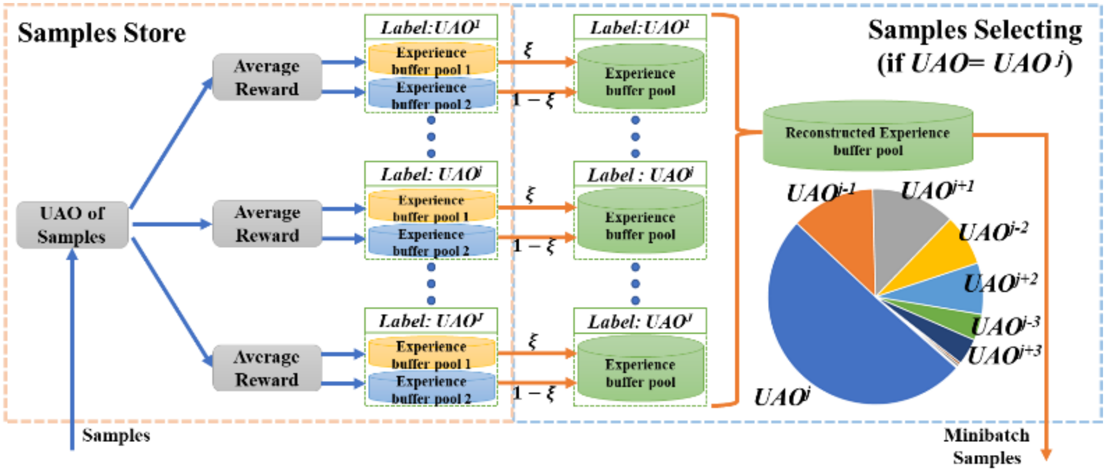



Fig. 5.5 Architecture of CEPR strategy

## 5.5 Case Studies

Accurate UAO prediction is the prerequisite for executing casual recruitment; therefore, the effectiveness of UAO prediction is first verified in **Case 1**. The optimal UAT applicable to ESP casual recruitment is given by experimentation. Subsequently, we assume casual recruitment with a fair mode under a no-incentive coefficient and verify its effectiveness through **Case 2**. After that, in **Case 3**, the performance of the proposed CEPR-TD3 algorithm is verified. The optimized incentive coefficients and payoffs are analyzed. We consider the proportion of participants who choose consensus and non-consensus within the bet-on mode and provide optimization results. Finally, in **Case 4** we explore the impact of considering weather factors on the profitability of DER participants and ESP. All simulations are implemented in Python 3.7 with the Pytorch packages on VSCode and run on a PC with an AMD RYZEN 9 3950X CPU and an NVIDIA RTX 3080 GPU.

### 5.5.1 Case 1: UAO Prediction Considering Weather Conditions

The core neural networks and hyperparameters for the proposed CTSGAN are selected the same as those in [101, 113]. The training set is from 1 January 2010 to June 2021, and the test set is from July 2021 to June 2022. The half-hourly electricity spot prices in New South Wales (NSW), Australia, are downloaded from the publicly available AEMO website https://aemo.com.au. The weather factor data is obtained from the Australian BOM website: http://www.bom.gov.au/climate/data/. Based on these historical data, we calculate the UAO historical data to serve as the training dataset, as described in Section 5.4.

#### 5.5.1.1 Comparative UAO Predictions with Different Models

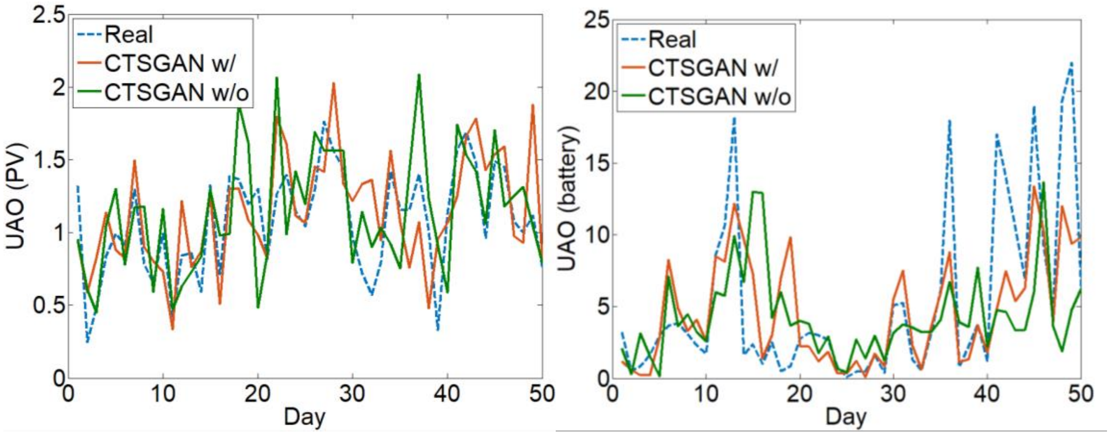
CTSGAN-based UAO prediction models with and without weather factors, denoted as CTSGAN w/ and CTSGAN w/o, respectively, are presented and benchmarked against the time-series GAN (TSGAN), long short-term memory (LSTM), and GRU models. Table 5.1 lists the RMSEs of the predicted UAO values for each model. The results indicate that both CTSGAN-based models achieve the best prediction performance with the smallest RMSEs for PV (0.364 and 0.415) and battery (4.878 and 5.554). It is worth noting that the UAO prediction errors of battery are significantly larger than those of PV. This can be attributed to the fact that the output curve of PV is relatively stable, and UAO mainly depends on the electricity price. In contrast, the profit of battery comes from buying low and selling high, and the electricity price and the timing of high and low points are crucial. These factors are also affected by the charging and discharging rates of the battery, making UAO prediction for the battery more challenging.

Table 5.1 RMSEs of predicted UAO with different models

| Models | | CTSGAN w/ | CTSGAN w/o | TSGAN | LSTM | GRU |
|---|---|---|---|---|---|---|
| RMSE | PV | 0.364 | 0.415 | 0.549 | 0.494 | 0.529 |
| | Battery | 4.878 | 5.554 | 6.427 | 6.132 | 6.892 |

#### 5.5.1.2 Impacts of UAO Prediction Considering Weather Factors

To investigate the impacts of weather factors on UAO prediction, we further analyze the UAO prediction results. Fig. 5.6 presents the predicted UAO for (a) PV and (b) battery with CTSGAN w/ and CTSGAN w/o models. It can be found that the CTSGAN w/ model can significantly enhance UAO prediction accuracy for PV. Specifically, the CTSGAN w/ model can successfully prevent the occurrence of incorrect spikes in the UAO prediction, which can be observed in the CTSGAN w/o model on days 18 and 37. Conversely, the UAO prediction for the battery system displays a considerably higher degree of fluctuation, ranging from 0 to 22, which presents a more significant challenge for accurate prediction. Notably, the CTSGAN w/o model fails to predict the UAO spikes between days 40 and 50, while the CTSGAN w/ model can effectively anticipate them. Despite some remaining discrepancies between the predicted and actual UAO values, our study suggests that the ability to anticipate the presence of UAO is of greater significance than predicting its precise value for VPP formation.




(a) (b)

Fig. 5.6 UAO prediction for (a) PV and (b) battery

Table 5.2 presents the percentage of accurately predicted UAO occurrences from July 2021 to June 2022, with a UAT of 1.0. As shown, including weather conditions can improve the overall correct prediction rates from 65.76% to 75.08% for the PV and from 71.23% to 76.45% for the battery.

Table 5.2 Accurate prediction rate of UAO occurrence

| | PV | | Battery | |
|---|---|---|---|---|
| Models | w/ | w/o | w/ | w/o |
| Correct | 75.08% | 65.76% | 76.45% | 71.23% |

Table 5.3 Comparison of predicted $UAO^{PV}$ with CTSGAN w/ and CTSGAN w/o models

| PV | | Real UAO > UAT | |
|---|---|---|---|
| | | True | False |
| Predicted UAO > UAT | True | 35.34% / 36.43% | 10.95% / 22.19% |
| | False | 13.97% / 12.05% | 39.74% / 29.33% |

Table 5.4 Comparison of predicted $UAO^{bat}$ with CTSGAN w/ and CTSGAN w/o models

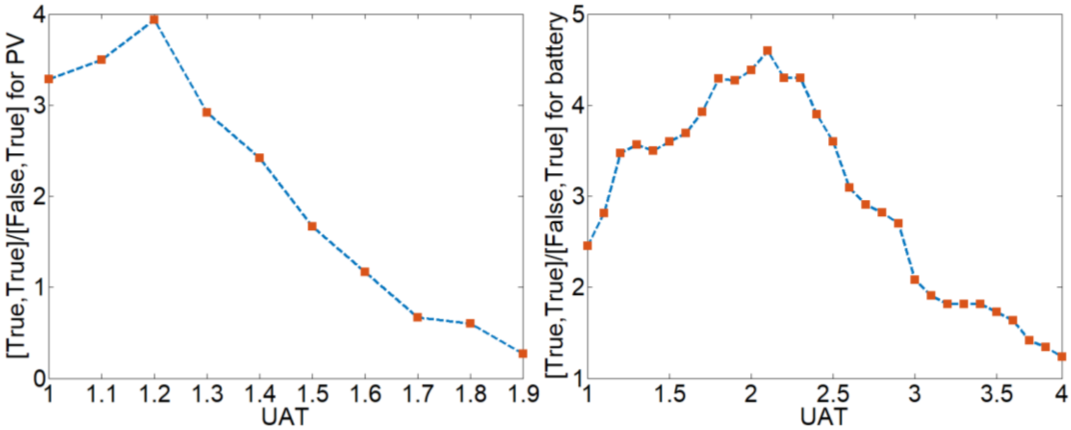
| Battery | | Real UAO > UAT | |
|---|---|---|---|
| | | True | False |
| Predicted UAO > UAT | True | 47.12% / 43.83% | 19.45% / 26.30% |
| | False | 4.10% / 2.47% | 29.33% / 27.40% |

However, simply reducing the prediction error is insufficient for the proposed casual recruitment problem. We propose a novel evaluation approach to assess the validity of UAO predictions with a binary confusion matrix [114]. There are four kinds of results in total, denoted as $\lfloor Bool, Bool \rfloor$. The first element signifies whether the actual UAO exceeds the UAT, while the second element indicates whether the predicted UAO is greater than the UAT. Tables 5.3-5.4 utilize four colors to represent the four possible results. Specifically, we denote $\lfloor True, True \rfloor$ by green, $\lfloor False, True \rfloor$ by orange, $\lfloor True, False \rfloor$ by blue and $\lfloor False, False \rfloor$ by grey. Accurate predictions of UAO, $\lfloor True, True \rfloor$, are crucial for ESP to generate excess profits, as incorrect predictions of UAO can lead to totally different outcomes. For instance, failing to predict the emergence of UAO, $\lfloor True, False \rfloor$, may prevent ESP from earning excess profits while predicting UAO that does not materialize, $\lfloor False, True \rfloor$, may result in losses for the ESP since high incentives for casual participants with bet-on $\boldsymbol{TN}$ mode. Therefore, it is more important to reduce the $\lfloor False, True \rfloor$ than $\lfloor True, False \rfloor$. Tables 5.3-5.4 illustrate the percentage of predicted UAO that meet both column and row headers, with values before the slash provided by CTSGAN w/ model and values after the slash by CTSGAN w/o model. It can be found that with the introduction of weather factors, the $\lfloor False, True \rfloor$ decreases from 22.19% to 10.95% for PV and $\lfloor False, True \rfloor$ from 26.30% to 19.45% for the battery. Although there is a small increase in $\lfloor True, False \rfloor$, it is within acceptable limits.

#### 5.5.1.3 Selection of Optimal UAT

The selection of UAT has a direct impact on both $\lfloor True, True \rfloor$ and $\lfloor False, True \rfloor$. The primary objective is to maximize the value of $\lfloor True, True \rfloor$ and minimizing the value of $\lfloor False, True \rfloor$. Fig. 5.7 illustrates the $\lfloor True, True \rfloor / \lfloor False, True \rfloor$ ratio for

(a) PV and (b) battery. The results show that the optimal UAT value for the PV is 1.2, with a corresponding $[True, True]/[False, True]$ ratio of 3.94. Similarly, the optimal UAT value for the battery system is found to be 2.1, with a corresponding $[True, True]/[False, True]$ ratio of over 4.63. Based on these findings, this work selects 1.2 and 2.1 as the optimal UAT values for the PV and battery. When the predicted UAO value exceeds these optimal UAT values, the ESP starts casual recruitment for DER owners.




(a) (b)

Fig. 5.7 $[True, True]/[False, True]$ ratios with different UATs for (a) PV and (b) battery

### 5.5.2 Case 2: Payoff Comparison with Casual Recruitment

To verify the effectiveness of casual recruitment, we assume a total of 3000 DER owners, comprising 2000 PV owners and 1000 battery owners. Half of the DER owners participate in regular recruitment, while the other half participate in casual recruitment in a fair mode with $v^{PV}$ and $v^{bat}$ values of 0. The average daily payoff for DER participants and ESP in casual recruitment and the average number of casual participants are calculated and compared with those obtained from regular recruitment for the differences. Tables 5.5 and 5.6 provide essential information on the PV, battery, and the operational cost parameters in (20). Table 5. 7 summarizes the increments in daily payoff and the number of participants for a period of four months.

Table 5.5 Parameters for PV and battery

| PV | Battery | | | |
|---|---|---|---|---|
| $P^m$ | $E^m$ | $\eta^{ch}$ | $\eta^{dch}$ | $P^m$ |
| 7kW | 10kWh | 0.95 | 0.95 | 4kW |

Table 5.6 Parameters of operation cost

| $a^{PV}$ | $b^{PV}$ | $c^{PV}$ | $a^{bat}$ | $b^{bat}$ | $c^{bat}$ |
|---|---|---|---|---|---|
| 0.01 | 0.01 | 0 | 0.01 | 0.01 | 0 |

Table 5.7 shows a significant increase in average payoff for PV and battery participants in the casual recruitment approach, with PV participants increasing by at least 17.6% (Nov. 2021) and up to 46.3% (May 2022), while battery participants increased by at least 72.3% (Jul. 2021) and up to 209.3% (May. 2022). The numbers of PV and battery participants in casual recruitment are also 23.9% to 49.6% and 34.2% to 59.4%, respectively, higher than in regular recruitment. However, it is important to note that the actual growth in payoff for PV and battery participants is not synchronized, as the increment in Nov. is 17.6% for PV (the lowest) and 183.9% for battery (the second highest). This discrepancy can be attributed to different expectations among PV and battery participants. The former expects higher electricity prices, while the latter expects more volatile electricity prices. The results also show a significant increase in payoffs for the ESP in the casual recruitment approach, with a 72.9% increase compared to regular recruitment in May 2022. Therefore, casual recruitment can effectively motivate and incentivize PV and battery participants while increasing ESP's profits simultaneously.

Table 5.7 Increments of payoffs and participant numbers for casual recruitment compared to regular recruitment

| | | Average Payoff | | | Number | |
|---|---|---|---|---|---|---|
| | | PV | battery | ESP | PV | battery |
| Increment | Jul. 2021 | 23.7% | 72.3% | 42.6% | 32.7% | 34.2% |
| | Nov. 2021 | 17.6% | 183.9% | 68.2% | 23.9% | 56.8% |
| | Feb. 2022 | 39.5% | 86.1% | 59.1% | 41.8% | 40.4% |
| | May. 2022 | 46.3% | 209.3% | 72.9% | 49.6% | 59.4% |

### 5.5.3 Case 3: Dynamic Coefficient Optimization and Effectiveness Verification

#### 5.5.3.1 Incentive Coefficients Optimization with Different DRL Algorithms

Based on Case 2, assuming there are 1000 additional PV participants and 500 potential battery participants for casual recruitment with bet-on mode. The ratio of participants choosing $\boldsymbol{TA}$ and $\boldsymbol{TN}$ remains 1:1, which is invisible to ESP. Table 5.8 presents the hyperparameters for the proposed CEPR-TD3.

Table 5.8 Hyperparameter settings for CEPR-TD3

| | |
|---|---|
| Critic learning rate | 0.001 |
| Actor learning rate | 0.001 |
| Discount factor | 0.9 |
| Selection probability for experience pool 1 | 0.8 |
| Capacity of experience pools 1 and 2 | 100000 |
| Update interval of the strategy network | 2 |
| UAO Interval for PV | 0.2 |
| UAO Interval for battery | 0.4 |
| Maximum number of episodes | 6000 |

Two modified experience-pool-based TD3 [111, 112], the conventional TD3 and DDPG are introduced as benchmarks. The comparison results are presented in Fig. 5.8. The vertical axis represents the ratio between the $U^{ESP}$ with optimized incentive coefficients and $U^{init}$ with all coefficients of 0. It can be observed that all modified TD3 algorithms exhibit superior convergence performance. In contrast, the conventional TD3 and DDPG algorithms exhibit poor convergence, characterized by large oscillations at the end of the learning process. The MEPR strategy, which identifies high- and low-value samples, enables TD3 to converge faster but does not provide further improvements in the later stages. On the other hand, the MLEPR strategy, which utilizes a multi-level experience pool, slightly improves performance compared to MEPR. The proposed CEPR strategy, which utilizes UAO labels to guide the storage of high- and low-value samples and maintains some diversity while focusing on training the neighboring UAOs through pool reconstruction, can further improve the payoff of the ESP, achieving a ratio of 1.32 compared to $U^{init}$, and also exhibiting better convergence speed. Table 5.9 illustrates the average optimized incentive coefficients with CEPR-TD3.

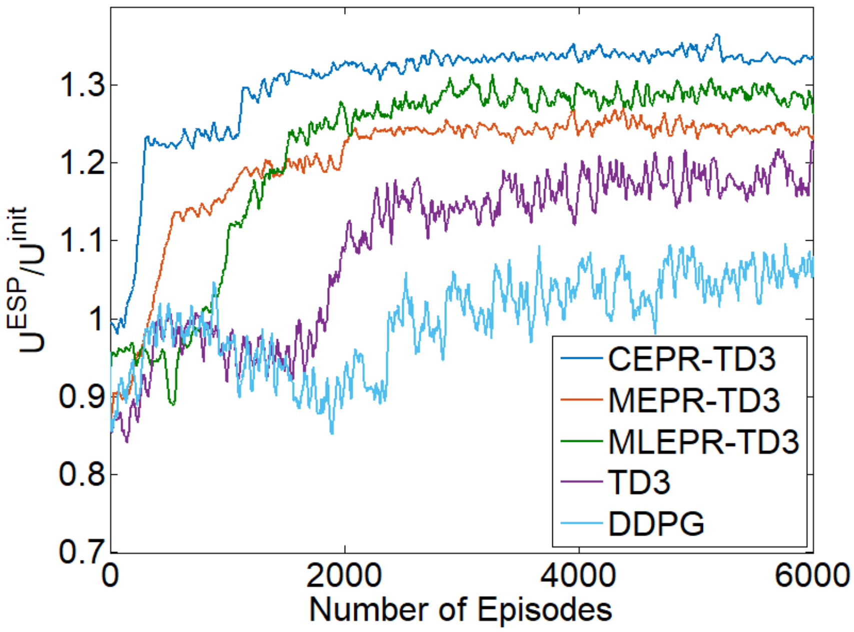
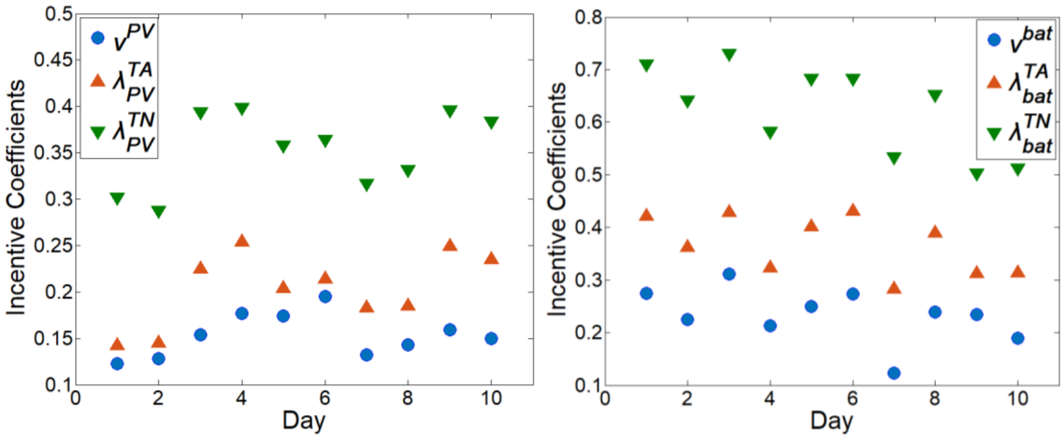



Fig. 5.8 Comparison of $U^{ESP}/U^{init}$ with different DRL algorithms

Table 5.9 Average optimal incentive coefficients for casual recruitment

| $\nu^{PV}$ | $\nu^{bat}$ | $\lambda_{PV}^{TA}$ | $\lambda_{PV}^{TN}$ | $\lambda_{bat}^{TA}$ | $\lambda_{bat}^{TN}$ |
|---|---|---|---|---|---|
| 0.148 | 0.227 | 0.201 | 0.356 | 0.373 | 0.610 |

Table 5.9 shows that the ESP can apparently offer higher incentive coefficients to battery participants than to PV participants in both fair and bet-on mode casual recruitment. It is worth noting that for bet-on mode, the incentive coefficients for ***TN*** choices are commonly higher than those for ***TA***, with $\lambda_{PV}^{TA}$ of 0.201, $\lambda_{bat}^{TA}$ of 0.373, $\lambda_{PV}^{TN}$ of 0.356 and $\lambda_{bat}^{TN}$ of 0.610. Fig. 5.9 presents the variation of all incentive coefficients for PV and battery participants across ten discontinuous days of casual recruitment. One can find that $\lambda_{PV}^{TA}$ and $\lambda_{bat}^{TA}$ are considerably closer to $\nu^{PV}$ and $\nu^{bat}$ compared to $\lambda_{PV}^{TN}$ and $\lambda_{bat}^{TN}$. Also, the fluctuation of $\lambda_{bat}^{TN}$ is markedly higher than that of $\lambda_{bat}^{TA}$ and similarly, the fluctuation of $\lambda_{PV}^{TN}$ is larger than that of $\lambda_{PV}^{TA}$. This underscores that the greater returns also entail a greater risk for those participants who opt for ***TN*** choices in bet-on mode.




(a) (b)

Fig. 5.9 Optimized incentive coefficients for (a) PV and (b) battery in ten casual recruitment days (discontinuous)

Table 5.10 Average daily payoff and participant number (casual)

| | | Average Payoff (A$/Day) | | | Participant Number | |
|---|---|---|---|---|---|---|
| | | PV | battery | ESP | PV | battery |
| Jul. 2021 | CEPR-TD3 | 1.312 | 3.184 | 610.49 | 1173 | 793 |
| | MEPR-TD3 | 1.283 | 3.083 | 549.65 | 1129 | 752 |
| | MLEPR-TD3 | 1.271 | 3.081 | 552.92 | 1098 | 733 |
| | TD3 | 1.239 | 2.837 | 521.28 | 973 | 684 |
| | DDPG | 1.225 | 2.782 | 471.86 | 956 | 677 |
| May. 2022 | CEPR-TD3 | 1.438 | 6.931 | 1141.82 | 1276 | 913 |
| | MEPR-TD3 | 1.399 | 6.782 | 1094.92 | 1254 | 896 |
| | MLEPR-TD3 | 1.382 | 6.831 | 1103.23 | 1246 | 903 |
| | TD3 | 1.304 | 5.470 | 907.16 | 1135 | 847 |
| | DDPG | 1.301 | 4.463 | 769.23 | 1124 | 834 |

Table 5.10 presents the average daily payoffs for participants engaged in casual recruitment, encompassing both the fair and bet-on modes, along with the total number of PV and battery participants for July 2021 (normal) and May 2022 (high and volatile), the sign A$ represents the Australian dollar. Table 5.10 reveals that the proposed CEPR-TD3 method yields the highest payoff for the ESP, with an average value of A$610.49 for July 2021, followed by MEPR-TD3 and MLEPR-TD3. Additionally, the proposed CEPR-TD3 method attracts the largest number of PV and battery participants, with an average daily payoff of A$1.312 and A$3.184, respectively, and 1173 PV and 793 casual battery participants. In May 2022, the proposed method significantly enhances the daily payoff for the battery compared to other methods, with values of A$6.931 (CEPR-TD3), A$5.470 (TD3), and A$4.463

(DDPG). Moreover, the daily payoff for the battery in May 2022 is substantially improved compared to that in July 2021 due to higher and more volatile electricity prices resulting in large UAO [115]. However, the daily payoffs for PV participants do not significantly increase from July 2021 to May 2022.

#### 5.6.3.2 Incentive Coefficients Optimization with Different Consensus Ratios for Casual Recruitment in Bet-on Mode

In the previous section, it is assumed that the number of participants in bet-on mode who chose ***TA*** and ***TN*** are in a 1:1 ratio. However, it is essential to note that this ratio is not necessarily stable, as it can be influenced by various factors. Therefore, we propose that the proportion of participants who choose ***TA*** to be denoted by $\varsigma$, and the proportion who choose ***TN*** to be denoted by 1-$\varsigma$. We assume that the value of $\varsigma$ follows a multiple normal distribution $\mathcal{N}(\zeta,\sigma)$, where $\zeta$ ranges from 0.1 to 0.9 and $\sigma$ is 0.1. If $\varsigma$ takes a value outside of the [0,1] range, it should be re-selected. MEPR-TD3 algorithm is introduced to optimize the incentive coefficients under different values of $\zeta$ in this section. The results are presented in Table 5.11.

Table 5.11 Average incentive coefficients with different consensus (***TA***) ratios for bet-on mode

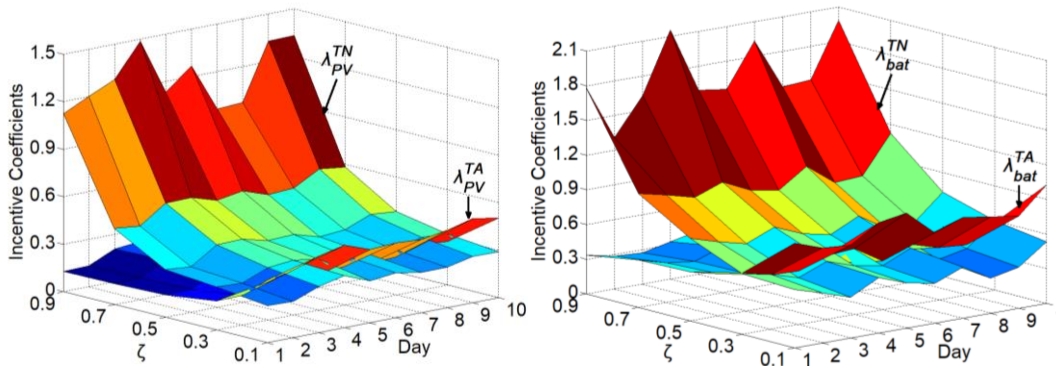
| | $\lambda_{PV}^{TA}$ | $\lambda_{PV}^{TN}$ | $\lambda_{bat}^{TA}$ | $\lambda_{bat}^{TN}$ |
|---|---|---|---|---|
| $\varsigma \sim \mathcal{N}(0.1, 0.1^2)$ | 0.476 | 0.285 | 0.868 | 0.395 |
| $\varsigma \sim \mathcal{N}(0.3, 0.1^2)$ | 0.252 | 0.311 | 0.520 | 0.432 |
| $\varsigma \sim \mathcal{N}(0.5, 0.1^2)$ | 0.199 | 0.352 | 0.357 | 0.603 |
| $\varsigma \sim \mathcal{N}(0.7, 0.1^2)$ | 0.192 | 0.526 | 0.351 | 0.920 |
| $\varsigma \sim \mathcal{N}(0.9, 0.1^2)$ | 0.184 | 1.311 | 0.302 | 1.661 |

One can find significant differences in the optimal incentive coefficients when $\varsigma$ satisfies different distributions. For instance, when $\varsigma \sim \mathcal{N}(0.5, 0.1^2)$, the incentive coefficients are close to those obtained in Case 3.1. In addition, the values of $\lambda_{PV}^{TA}$ and $\lambda_{bat}^{TA}$ increase as the proportion of participants selecting ***TA*** decreases, ranging from 0.184 to 0.476 and from 0.302 to 0.868, respectively. Furthermore, the incentive coefficients for battery participants tend to fluctuate more than those for PV participants. For instance, when $\varsigma \sim \mathcal{N}(0.9, 0.1^2)$, the value of $\lambda_{PV}^{TA}$ is 0.184, and the value of $\lambda_{bat}^{TA}$ is 0.302, while when $\varsigma \sim \mathcal{N}(0.1, 0.1^2)$, the value of $\lambda_{bat}^{TA}$ increases to

0.868 while the value of $\lambda_{PV}^{TA}$ only increases to 0.476. This implies that the ESP needs to adjust the incentive coefficients more accurately in response to changes in the selection preferences of battery participants than for PV participants. Finally, it is worth noting that when 90% of bet-on mode participants choose $\boldsymbol{TA}$, the values of $\lambda_{PV}^{TA}$ and $\lambda_{bat}^{TA}$ become very low, while the values of $\lambda_{PV}^{TN}$ and $\lambda_{bat}^{TN}$ increase significantly to greater than 1, with values of 1.311 for the PV and 1.661 for the battery, respectively.

Fig. 5.10 depicts the optimized incentive coefficients for 10 casual recruitment days for PV and battery participants with different consensus ratios. There are notable discrepancies in the incentive coefficients for each recruitment day. For instance, on Day 4, when $\zeta$=0.9, the coefficient $\lambda_{PV}^{TA}$ is the highest (1.48) among all 10 days, whereas on Day 8, it drops to 0.970. Additionally, the trends of $\lambda_{PV}^{TA}$ and $\lambda_{PV}^{TN}$ with respect to $\zeta$ are consistent with those presented in Table 5.11. Notably, when $\zeta$=0.9, the coefficient $\lambda_{bat}^{TN}$ is nearly twice the value of $\lambda_{PV}^{TN}$. Furthermore, when $\zeta$ lies between 1.5 and 2, the values of $\lambda_{PV}^{TA}$ and $\lambda_{PV}^{TN}$ are essentially the same, which implies that PV participants would receive similar payoffs regardless of their participation in either side of the bet-on mode. Similarly, when $\zeta$ falls between 1.75 and 2.5, the payoffs are the same for both sides (***TA*** and ***TN***) for battery participants.

It should be noted that although this section discusses the different proportions of participants choosing $\boldsymbol{TA}$, the purpose is to validate that the proposed dynamic incentive coefficient optimization method is still applicable under different proportions. In fact, situations where $\varsigma \sim \mathcal{N}(0.1, 0.1^2)$ rarely occur. This is because most people are still risk-averse and are more inclined to choose the consensus mode that is consistent with ESP expectations.

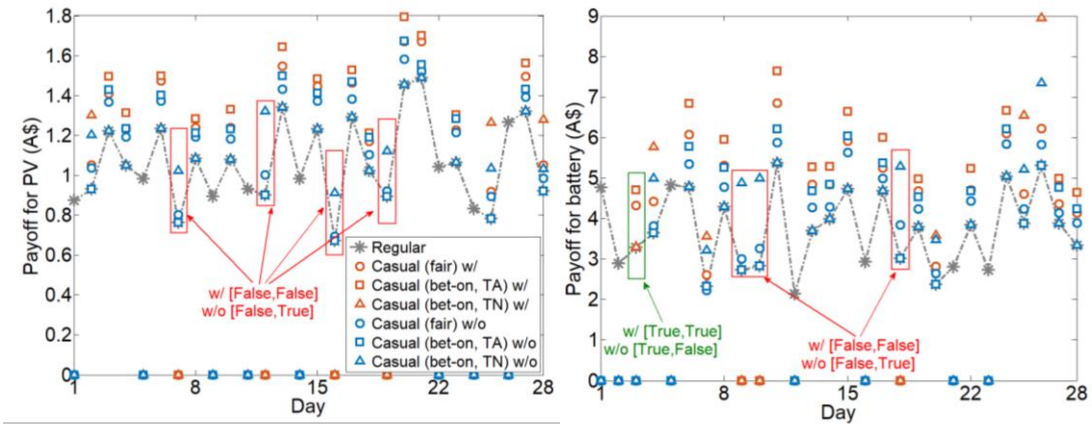



(a) (b)

Fig. 5.10 Optimized incentive coefficients for (a) PV and (b) battery with different consensus ratio

### 5.5.4 Case 4: Impact of Weather Conditions on Daily Payoffs

This section examines the impact of the introduction of weather conditions on DER recruitment. Fig. 5.11 depicts the daily payoffs for PV and battery participants over four consecutive weeks, with UAO predicted by two models: CTSGAN w/ and CTSGAN w/o, assuming $\varsigma \sim \mathcal{N}(0.5, 0.1^2)$. The red and green boxes in Fig. 5.11 highlight the differences in VPP formation. In Fig. 5.11 (a), ESP accurately recruits participants on 16 of the 28 days with CTSGAN w/ model, with $[True, True]$ of 13 days and with $[False, True]$ of 3 days. However, the CTSGAN w/o model may mistakenly recruit PV participants on Days 7, 12, 16, and 19, as indicated by the red boxes. Mis-recruitment can result in increased compensation for participants who choose ***TN***, thereby reducing profits for the opposite participants. Consequently, payoffs provided by the CTSGAN w/o model are lower than those of the CTSGAN w/ model. The daily payoffs for battery participants are generally higher than for PV participants. For example, on Day 26, the CTSGAN w/ model provides ***TN*** participants with a payoff of almost A$9, while the CTSGAN w/o model offers only A$7.3. In Fig. 5.11 (b), the first green box indicates that ESP with CTSGAN w/o model does not recruit battery participants on the day when a UAO occurs. However, as highlighted by the red boxes, the CTSGAN w/o model incorrectly recruits battery participants on Days 9, 10, and 18. The lower incentive coefficients provided by the CTSGAN w/o model led to lower payoffs for participants, resulting in fewer participants and reduced total payoffs for the VPP aggregator. Table 5.12 summarizes the daily payoffs for PV and battery participants and the ESP with the CTSGAN w/ and CTSGAN w/o models.




(a) (b)

Fig. 5.11 Comparison of average daily payoffs for (a) PV and (b) battery participants with CTSGAN w/ and CTSGAN w/o models

Table 5.12 Daily payoffs for DER owners and ESP with CTSGAN w/ and CTSGAN w/o (A$/Day)

| | | PV | | battery | | ESP | |
|---|---|---|---|---|---|---|---|
| | | w/ | w/o | w/ | w/o | w/ | w/o |
| Jul. 2021 | Regular | 1.055 | | 2.972 | | 1025.23 | 882.66 |
| | Casual (fair) | 1.289 | 1.002 | 3.283 | 3.103 | | |
| | Casual (bet-on, ***TA***) | 1.319 | 0.932 | 3.293 | 3.203 | | |
| | Casual (bet-on, ***TN***) | 1.320 | 0.947 | 3.297 | 3.037 | | |
| May. 2022 | Regular | 1.327 | | 4.537 | | 1693.42 | 1003.72 |
| | Casual (fair) | 1.440 | 1.091 | 6.776 | 3.973 | | |
| | Casual (bet-on, ***TA***) | 1.441 | 1.183 | 6.973 | 4.841 | | |
| | Casual (bet-on, ***TN***) | 1.442 | 1.130 | 6.983 | 4.502 | | |

In July 2021, the average daily payoffs for regular PV and battery participants are A$1.055 and A$2.972, respectively. With the application of CTSGAN w/, PV participants can receive payoffs of A$1.289 (fair mode), A$1.319 (bet-on, ***TA***) and A$1.320 (bet-on, ***TN***), while the payoffs for PV participants decrease significantly with CTSGAN w/o model. The benefits of using CTSGAN w/ are more pronounced for battery participants compared to PV participants. The payoffs for ESP also increase significantly from A$882.66 (CTSGAN w/o) to A$1025.23 (CTSGAN w/). The gain in ESP's payoff is more significant in May 2022 at 1.68 times than in July

2021 at 1.16 times, suggesting that a UAO prediction model incorporating weather factors can significantly improve the average ESP payoff when electricity prices are more volatile. The payoffs for battery participants who choose casual recruitment are very similar, with values of A$3.283, A$3.293, and A$3.297 in July 2021 and A$6.776, A$6.973, and A$6.983 in May 2022. However, with the CTSGAN w/o model, the differences in payoffs become significantly larger. Including weather conditions, the accurate UAO prediction allows the DRL method to learn optimal incentive coefficients.

## 5.6 Chapter Summary

The existing fixed recruitment approach for VPPs lacks the flexibility to expand the number of DER participants temporarily when facing UAO. This results in limited profitability for the VPP, preventing it from further maximizing its returns. In addition, DER participants have difficulty being satisfied with the uniform plan offered by the ESP, which leads to a low participation rate. To address this issue, a VPP formation approach combining regular and casual recruitment approaches was proposed to cater to DER participants with different risk-payoff profiles. A bet-on agreement was introduced in DER participant recruitment, where the fulfillment or non-fulfillment of expected payoffs will enable DER owners to receive compensation from the ESP. Moreover, two key issues were addressed: 1) The CTSGAN-based UAO prediction model incorporating weather conditions was proposed with superior performance; 2) The CEPR-TD3 algorithm was proposed to optimize the incentive coefficient with participation rate improvement and high payoff achievement.

# CHAPTER 6

# Deposit and Withdrawal Service Guided Aggregation of Shared Energy Storage with High Price-Tolerance Prosumers

*High price-tolerance residential prosumers lack the capability and motivation to participate in SES actively. To address the issue, an SES aggregation framework is proposed, which does not require prosumers to take additional actions and allows them to maintain existing energy storage patterns. This framework includes two steps, short-term construction and long-term operation. During the construction, an ESP aggregates all willing prosumers and gains control over their ESSs. Subsequently, the ESP offers electricity DWS with dynamic coefficients, enabling prosumers to withdraw more electricity than they deposit without additional actions. Additionally, a matching mechanism is proposed to align prosumers' electricity consumption behaviors with ESP's optimization strategies. Finally, an improved DRL algorithm optimizes the dynamic coefficients in DWS and trading strategies. Case studies are conducted to verify the effectiveness of the proposed SES aggregation framework with DWS and the matching mechanism.*

## 6.1 SES Aggregation Framework-Construction and Operation

Residential prosumers have a high tolerance level for electricity bills. This implies that incentivizing them to modify their electricity consumption behaviors with low rewards may be impractical. Therefore, when faced with high price-tolerance prosumers, their electricity consumption behaviors can be seen as an inherent attribute that is not easily changed. We propose an SES aggregation framework, as shown in Fig. 6.1, for the prosumers with high price tolerance. The framework consists of multiple cycles, each including a short-term construction step and a long-term operation step. When external environmental changes occur, such as seasonal transitions or alterations in prosumers' electricity consumption behaviors leading to reduced profits, it becomes crucial to initiate a new construction step before continuing with the long-term operation.

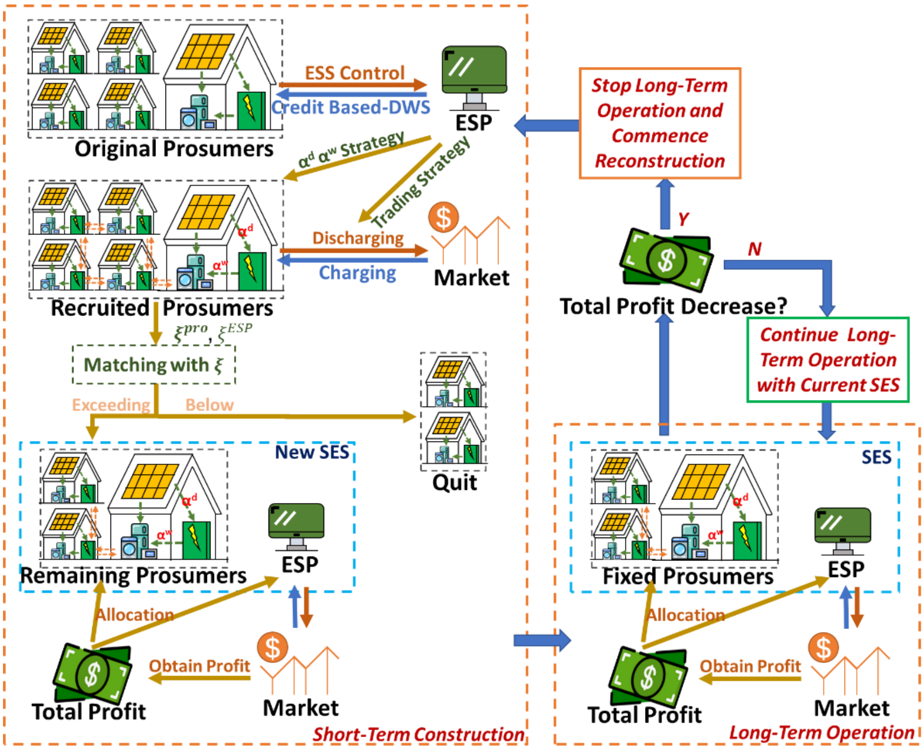



Fig. 6.1 Framework for SES construction and operation among price-tolerance prosumers

**SES Construction:** The ESP aggregates all PV-ESS prosumers willing to participate in the SES for a short-term construction step. Although the ESP gains control over the ESS to construct the SES, it is essential for the SES to prioritize meeting households' electricity deposit and withdrawal needs through DWS (to be detailed in 6.2). Therefore, the ESP needs to optimize the dynamic coefficients of DWS and optimize the SES's charging and discharging strategies to participate in the wholesale electricity market. In addition, the prosumers typically expect to reduce their electricity bills after participating in the SES. However, to seek maximum profitability, the ESP often screens the prosumers, proposing a matching mechanism (to be detailed in 6.3) between the ESP and the prosumers for mutual selections.

**SES Operation:** After a short period of operation, some prosumers are unable to achieve the expected profit targets, so they may choose to quit SES. The remaining prosumers and the ESP continue the long-term operation of the SES. The overall SES can generate additional profits through participating in energy trading. This portion of the profits will be allocated among the ESP and all remaining prosumers.

Regarding the proposed framework, several key issues need to be addressed: the reasonable dynamic coefficients for DWS, the matching mechanism between the ESP and the prosumers, and the optimization of the trading strategy and dynamic coefficient strategy. It is important to note that the framework is not only limited to the prosumers with high price tolerance but also applies to the easily incentivized prosumers. Through different DWS coefficients, easily incentivized prosumers will change their electricity consumption behaviors, making achieving the predetermined profit goals easier and becoming SES participants.

# 6.2 Credit-based Deposit and Withdrawal Service

## 6.2.1 Credit-based DWS Model

Credit is the concept introduced in the shared energy model to record prosumers' changes in charging and discharging [66]. However, the existing credit-based model cannot clearly measure whether prosumers' behaviors impact the SES trading and scheduling. Therefore, with the concept of credit-based energy sharing, a credit-based DWS model is proposed as follows:

$$\Gamma_{j,t+1} = \Gamma_{j,t} + \left(\alpha_t^{dps} * P_{j,t}^{dps} * \eta^{adm} - \alpha_t^{wtd} * P_{j,t}^{wtd}\right) * \Delta\mathrm{t} \tag{6.1}$$

where $\Gamma_{j,t+1}$ represents the credit points of prosumer j at time t; $P_{j,t}^{dps}$ represents the electric power deposited by prosumer j, indicating the charging process for SES; $P_{j,t}^{wtd}$ represents the electric power withdrawn by prosumer j, indicating the discharging process for SES; $\alpha_t^{dps}$ represents the dynamic deposit coefficient, $\alpha_t^{wtd}$ represents the dynamic withdrawal coefficient; from the prosumer side, $\eta^{adm}$ can be considered as the administration cost of SES during the deposit and withdrawal process for prosumers, but in reality, it's merely the loss associated with charging and discharging, denoted as $\eta^{adm} = \eta^{ch} / \eta^{dis}$, where $\eta^{ch}$ and $\eta^{dis}$ are the charging and discharging efficiency, respectively. With the proposed DWS, the ESP aims to ensure that the electricity loss during the deposit and withdrawal process is smaller than that when prosumers control the ESS themselves, which means that $\alpha_t^{dps} \geq 1$ and $\alpha_t^{wtd} \leq 1$. Indeed, the introduction of (6.1) is not primarily aimed at incentivizing prosumers to change their electricity consumption behaviors, especially for those with high price tolerance who are less likely to modify their consumption patterns. Instead, its purpose is to provide rewards to prosumers whose consumption behaviors align with the dispatching and trading needs of the SES.

### 6.2.2 Factors for Dynamic Coefficients

The deposit and withdrawal coefficients $\alpha_t^{dps}$ and $\alpha_t^{wtd}$ have three main factors: the supply-demand ratio of the SES, the storage state of the SES and the potential for future profitability. The supply-demand ratio $F_t^{SDR*}$ is proposed to describe the supply-demand status of all prosumers in the SES, which can be expressed as:

$$F_t^{SDR*} = \sum_{j=1}^{J} \frac{P_{j,t}^{dps} - P_{j,t}^{wtd}}{P_{MAX}} \tag{6.2}$$

where $P_{MAX}$ represents the max charging and discharging power for the SES. The second factor $F_t^{SOC*}$is related to the current storage state of the SES, as follows:

$$F_t^{SOC*} = \frac{E_t}{E_{MAX}} \tag{6.3}$$

where $E_t$ and $E_{MAX}$ represent the storage state at time t and max capacity of SES, respectively. The third factor, named arbitrage opportunity $F_t^{AO*}$, indicates the potential for future profitability, can be expressed as:

$$F_t^{AO*} = \lambda_{t+1}^{spot,Pred} - \lambda_t^{spot,Real} \tag{6.4}$$

where $\lambda_{t+1}^{spot,Pred}$ and $\lambda_t^{spot,Real}$ represent the predicted spot price at time t+1 and the actual spot price at time t in the electricity market, respectively. $F^{SDR*}$, $F_t^{SOC*}$ and $F_t^{AO*}$ need to be normalized to the range [0, 1], denoting as $F_t^{SDR}$, $F_t^{SOC}$ and $F_t^{AO}$, respectively.

### 6.2.3 Balance for Deposit and Withdrawal

We propose DWS with dynamic coefficients that allow prosumers to deposit and withdraw electricity through the SES similarly to their own ESS, without requiring additional actions. There are often cases of excessive deposits or withdrawals by the prosumers, which makes it challenging to ensure zero credit within the same day. To address this issue, we provide the following settlement method:

$$R^{CR} = \sum_{t=1}^{T} \sum_{j=1}^{J} (-\Gamma_{j,t} * \lambda^{wtd}) \,, if \ \Gamma_{j,t} < 0 \quad (6.5)$$

$$C^{CR} = \sum_{t=1}^{T} \sum_{j=1}^{J} (\Gamma_{j,t} * \lambda^{dps}) \,, if \ \Gamma_{j,t} > 0 \quad (6.6)$$

where $R^{CR}$ and $C^{CR}$ represent the unified settlement of unsettled revenue and cost at the end of the trading day. Positive credit indicates that the prosumer has excess electricity that is not withdrawn. This surplus is purchased from the prosumer at the price $\lambda^{dps}$, which is considered as the cost for the SES. On the other hand, negative credit indicates that the prosumer withdraws more electricity than the deposit. This surplus electricity is sold to the prosumer at a price represented by $\lambda^{wtd}$, which is considered as the profit for the SES. For the prosumer, these two components result from improper planning and impose an additional burden on the SES. Therefore, we set $\lambda^{dps}$ and $\lambda^{wtd}$ to be $\beta^{FIT} * \lambda^{FIT}$ and $\beta^{TOU} * \lambda^{TOU}$, with $\beta^{FIT} < 1$ and $\beta^{TOU} > 1$.

### 6.2.4 Virtual Storage and Extraction

From the prosumer's perspective, any $P_{j,t}^{dps}$ is considered as charging the SES. However, due to storage constraints or pricing considerations, there are some situations where the electricity deposited by prosumers cannot be charged to the SES. Instead, it is directly sold on the wholesale market. This situation is referred to as virtual storage, and it encompasses three scenarios: 1) when the net power exceeds a certain range, $F_t^{SDR} > \beta^{SDR,UP}$, 2) when the battery SOH exceeds a certain range, $F_t^{SOC} > \beta^{SOC,UP}$, and 3) future arbitrage opportunity is at a large level, $F_t^{AO} > \beta^{AO,UP}$.

The revenue for virtual storage can be expressed as:

$$R^{VS} = \sum_{t=1}^{T} \sum_{j=1}^{J} P_{j,t}^{dps,V} * \lambda_t^{SPO} * \Delta t \tag{6.7}$$

where $P_{j,t}^{dps,V}$ denotes the virtual storage power by prosumer $j$. On the contrary, when 1) the net power falls below a certain range, $F_t^{SDR} < \beta^{SDR,LOW}$, 2) the battery SOH falls below a certain range, $F_t^{SOC} < \beta^{SOC,LOW}$, and 3) electricity prices are at future arbitrage opportunity is at a low level, $F_t^{AO} < \beta^{AO,LOW}$ , virtual extraction is implemented, and the cost can be expressed as:

$$C^{VS} = \sum_{t=1}^{T} \sum_{j=1}^{J} -P_{j,t}^{wtd,V} * \lambda_t^{SPO} * \Delta t \tag{6.8}$$

where $P_{j,t}^{wtd,V}$ denotes the virtual extraction power by prosumer j.

## 6.3 Matching Mechanism Between ESP and Prosumers

Prosumers can participate in constructing the SES and obtain some benefits without changing their electricity consumption and storage patterns. They still expect to achieve the desired bill reduction threshold through DWS. As the ESP aims to maximize profits, it cannot aggregate all prosumers and should select participants who match their scheduling and trading strategy. To address this, based on the credit-based DWS, we propose a matching mechanism for mutual selections between the ESP and the prosumers.

### 6.3.1 Matching Mechanism-based Contract

For prosumers, DWS is an electricity gain service that can effectively reduce their bills. Prosumers want to maximize their benefits from the SES with DWS, and there is a lower threshold, assumed as $\xi^{pro}$, below which they do not participate, and above which they decide to participate in the SES. It should be noted that the $\xi^{pro}$ is an attribute of the prosumer group, which is determined by many factors of the prosumer group, including age, income, electricity bill level, education level, and even community groups' attitudes towards renewable energy [116].

On the other hand, for ESP, the DWS helps identify prosumers whose electricity consumption behaviors match the SES scheduling strategy, either based on their existing behaviors or the behaviors influenced by DWS incentives. When DWS can effectively help a prosumer reduce the bill, it means that the prosumer has a high level of matching with ESP's scheduling strategy. Aggregating these prosumers helps to increase the overall profit of the SES. Therefore, the lower threshold for the matching level of aggregated prosumers is denoted as $\xi^{ESP}$, and the final threshold $\xi$ needs to be selected as the maximum between $\xi^{pro}$ and $\xi^{ESP}$, expressed as:

$$\xi = \max \langle \xi^{pro}, \xi^{ESP} \rangle \tag{6.9}$$

With $\xi$ as the contract threshold, the matching mechanism-based contract between ESP and prosumers can be expressed as follows:

$$\mathcal{R}_{bill,j}/\mathcal{B}_j^{ORI} \geq \xi \tag{6.10}$$

$$\mathcal{R}_{bill,j} = \mathcal{B}_j^{ORI} - \mathcal{B}_j \tag{6.10.a}$$

$$\mathcal{B}_j^{ORI} = \sum_{t^b} \lambda_{t^b}^{TOU} * (D_{t^b}^{i} - G_{t^b}^{i}) * \Delta\text{t} \tag{6.10.b}$$

$$t^b \in \left\{ t \middle| G_{j,t} \leq D_{j,t} E_t^{ESS,j} \leq 0 \right\} \tag{6.10.c}$$

where $\mathcal{R}_{bill,j}$ represents the household electricity bill reduction for prosumer *j*, which can be expressed as (6.10.a); $\mathcal{B}_j^{ORI}$ represents the household electricity bill for prosumer *j* before participating in the SES, can be expressed as (6.10.b), and $\mathcal{B}_j$ represents that after participating in the SES, as detailed in Section 6.4; In (6.10.b), before aggregating the SES, prosumer *j* needs to pay the electricity bill for time $t^b$, as (6.10.c), when their demand is greater than PV generation, and their household ESS does not have electricity available for usage; $\xi$ represents the contract condition, only when this condition is satisfied, prosumer *j* continues to participate in the SES; otherwise, prosumer *j* should quit the SES.

### 6.3.2 Conventional Profit Allocation

The bill reduction $\mathcal{R}_{bill,j}$ that prosumer $j$ can effectively achieve can be considered as a reward for their highly matching electricity consumption behaviors. In addition to bill reduction, ESP can further generate profits by controlling the SES participants in the electricity wholesale market, which can be denoted as $\mathcal{R}_{SES}$ (to be detailed in 6.4). Since the SES units are invested by the prosumers and operated by the ESP, the final profit $\mathcal{R}_{SES}$ needs to be further allocated among two parts: ESP and prosumers. We introduce the Sharply Value for profit allocation between ESP, denoted by $\mathcal{R}_{ESP}^{SV}$, and all prosumers, denoted by $\mathcal{R}_{pro}^{SV}$. Defining $X/\langle i\rangle$ to be the set of all parts after removing part $i$, the marginal contribution of part $i$ to a coalition $S$, $S \subseteq X/\langle i\rangle$, is

$$\epsilon_i(S) = v(S \cup \langle i\rangle) - v(S) \tag{6.11}$$

where $v(S)$ represents the total payoffs that the members of $S$ can obtain by cooperation. The SV for each part in our model can be calculated as follows:

$$\phi_i = \sum_{S\subseteq X/\langle i\rangle} \frac{|S|!(n-|S|-1)!}{n!}\epsilon_i(S) \tag{6.12}$$

where $n$ represents the total number of parts. Since $\mathcal{R}_{SES}$ needs to be further allocated among two parts: ESP and prosumers, therefore, $n$ is set to 2 in our work.

With the introduction of Sharply Value, $\mathcal{R}_{ESP}^{SV}$ and $\mathcal{R}_{pro}^{SV}$ can be readily obtained. Then, we further allocate $\mathcal{R}_{pro}^{SV}$ to each prosumer by their capacities as follows:

$$\mathcal{R}_{pro,j}^{SV} = \frac{E_{MAX}^{pro,j}}{\sum_{j=1}^{J} E_{MAX}^{pro,j}} * \mathcal{R}_{pro}^{SV} \tag{6.13}$$

where $E_{MAX}^{pro,j}$ represents the max capacity for the *j-th* prosumer's household ESS.

## 6.4 Profit Maximization with Dynamic DWS Coefficients Optimization with CNEPR-TD3

### 6.4.1 Revenue of SES

The ESP maximizes profit by aggregating prosumers to participate in the electricity market through DWS. The objective function of the whole SES can be expressed as follows:

$$\mathcal{R}_{SES} = R^{TRD} + R^{CR} - C^{CR} + R^{VS} - C^{VS} - C^{SES} \tag{6.14}$$

where $R^{CR}$and $C^{CR}$ represent the unified settlement of unsettled credits at the end of the trading day; $R^{VS}$ and $C^{VS}$ represent the revenue and cost for virtual storage and extraction; $R^{TRD}$ represents the trading revenue in the electricity market, expressed as:

$$R^{TRD} = \sum_{t=1}^{T}\left(\lambda_t^{SPO} * \left(P_t^{SPO,ch} * \eta^{ch} - P_t^{SPO,dis}/\eta^{dis}\right) * \Delta\mathrm{t}\right) \tag{6.15}$$

where $\lambda_t^{SPO}$ represents the real electricity price in the wholesale market at time t; $P_t^{SPO,ch}$ and $P_t^{SPO,dis}$ represents the electric power purchased from the wholesale market that is charged to the SES, and the electric power discharged from the SES that is sold to the wholesale electricity market. $C^{SES}$ represents the degradation cost of all SES units, expressed as:

$$C^{SES} = \sum_{t=1}^{T}\frac{k_e}{100}\sum_{j=1}^{J}\gamma_j^{SES}\begin{pmatrix} P_t^{SPO,dis} \\ +\sum_{j=1}^{J}\left(P_{j,t}^{wtd} - P_{j,t}^{wtd,V}\right)\end{pmatrix} * \Delta\mathrm{t} \tag{6.16}$$

where $k_e$ represents the slope of the linear approximation of the battery life as a function of the cycles and $\gamma_j^{SES}$ the unit capital investment on the batteries for prosumer j [117, 118].

The energy storage state of SES at time t+1 $E_{t+1}^{SES}$, its constraint and the charging and discharging power constraints are respectively given as:

$$E_{t+1}^{SES} - E_t^{SES} = \sum_{j=1}^{J}\left(\left(P_{j,t}^{dps} - P_{j,t}^{dps,V}\right) * \eta^{ch} - \left(P_{j,t}^{wtd} - P_{j,t}^{wtd,V}\right)/\eta^{dis}\right) * \Delta\mathrm{t} +$$

$$\left(P_t^{SPO,ch} * \eta^{ch} - P_t^{SPO,dis}/\eta^{dis}\right) * \Delta\mathrm{t} \tag{6.17}$$

$$0 \le E_t^{ESS} \le E_{MAX}^{SES} \tag{6.18}$$

$$0 \le \sum_{j=1}^{J}\left(P_{j,t}^{dps} - P_{j,t}^{dps,V}\right) + P_t^{SPO,ch} \le P_{MAX}^{SES,ch} \tag{6.19}$$

$$0 \le \sum_{j=1}^{J}\left(P_{j,t}^{wtd} - P_{j,t}^{wtd,V}\right) + P_t^{SPO,dis} \le P_{MAX}^{SES,dis} \tag{6.20}$$

where $E_{MAX}^{SES}$ represent the maximum capacity of the SES; $P_{MAX}^{SES,ch}$ and $P_{MAX}^{SES,dis}$ represent the maximum charging and discharging power of the SES.

### 6.4.2 Bill of Prosumer

With the DWS provided by ESP, prosumer j can deposit excess electricity in SES and withdraw it when there is a demand. The electric power for storing and exacting can be described as follows:

$$P_{j,t}^{stg} = G_{j,t} - D_{j,t}, if\ G_{j,t} > D_{j,t} \tag{6.21}$$

$$P_{j,t}^{ext} = D_{j,t} - G_{j,t}, if\ G_{j,t} < D_{j,t}\ and\ \Gamma_{j,t} > 0 \tag{6.22}$$

where $D_{j,t}$ and $G_{j,t}$ represent the electricity demand and PV generation for prosumer j at time t.

Despite prosumers transferring control of ESS to ESP, they can still maintain their original ESS storage patterns through DWS services. Consequently, they can automatically deposit excess PV-generated electricity when it exceeds their household demand and withdraw it when their household demand exceeds the PV generation and their credit $\Gamma_{j,t}$ is positive. Since their credit $\Gamma_{j,t}$ is updated every hour, there is a possibility that the credit may become negative.

According to the DWS with dynamic coefficients $\alpha_t^{dps}$ and $\alpha_t^{wtd}$, the electricity bill for prosumer j can be represented as follows:

$$\mathcal{B}_j = C_j^{NET} - R_j^{CR} + C_j^{CR} \tag{6.23}$$

where $C_j^{NET}$ represents the cost of the net load of prosumer j, expressed as:

$$C_j^{NET} = \sum_{t=1}^{T} \left(\lambda_t^{TOU} * P_{j,t}^{net} * \Delta \mathrm{t}\right) \tag{6.24}$$

$R_j^{CR}$ and $C_j^{CR}$ represent the revenue and cost from settling credits at the end of each day, expressed as:

$$R_j^{CR} = \sum_{t=1}^{T} (\Gamma_{j,t} * \lambda^{dps}) \tag{6.25}$$

$$C_j^{CR} = \sum_{t=1}^{T} (\Gamma_{j,t} * \lambda^{wtd}) \tag{6.26}$$

### 6.4.3 Optimization Objective

The final optimization objective function includes two parts: the overall profit of the SES and the reduction in the prosumer's bills, which can be expressed as follows:

$$max\ Obj = \mathcal{R}_{SES} + \sum_{j=1}^{J} \mathcal{R}_{bill,j} \tag{6.27}$$

For our proposed DWS-based SES construction and operation, providing a larger $\alpha_t^{dps}$ or a smaller $\alpha_t^{wtd}$ to prosumers can effectively enhance their profit, thereby attracting a more significant number of prosumers. However, this undoubtedly affects the total profit obtained from the participation of SES in the electricity market. Conversely, prioritizing the profits from the SES's participation in the electricity market too much will impact the reduction in the prosumers' bills, thereby affecting their enthusiasm for prosumers and potentially leading to a decrease in the number of participants. Therefore, in order to consider the profits of both parties, we set the weight of both sides $\mathcal{R}_{SES}$ and $\sum_{j=1}^{J} \mathcal{R}_{bill,j}$ in the objective function to 1:1. We use a DRL algorithm to optimize the deposit-withdrawal coefficient strategy and the trading strategy simultaneously, aiming to maximize the final objective function.

### 6.4.4 Optimization of Dynamic DWS Coefficient and Trading Strategy

The ESP needs to provide reasonable dynamic coefficients to construct the SES with DWS for prosumers and then schedule the SES to participate in energy trading in the wholesale market. This process can be modeled as an MDP [98], which can be characterized by a tuple $< \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R} >$, where $\mathcal{S}$, $\mathcal{A}$, $\mathcal{P}$ and $\mathcal{R}$ stand for the state, action, transition function and reward, respectively. The DWS coefficient and trading problem can be reformulated as the following MDP, and the agent (ESP) aims to optimize its policy to get the maximum return in the whole trajectory. In our work, the state is the feedback from the environment based on the agent's action and previous state, which includes the supply-demand ratio $F_t^{SDR}$, storage state of the SES $F_t^{SOC}$, and potential arbitrage opportunity $F_t^{AO}$. The action contains DWS coefficients $\alpha_t^{dps}$ and $\alpha_t^{wtd}$, and exchanged power $P_t^{SPO,ch}$ and $P_t^{SPO,dis}$. The reward for the agent ESP is (6.27). A CNEPR modified TD3 is proposed to provide optimized dynamic and exchanged power in the electricity market.

#### 6.4.4.1 Twin Delayed Deep Deterministic Policy Gradient

TD3 is an upgraded version of DDPG, which is developed based on the architecture of the actor-critic algorithm. The state and action are represented by the critic and actor networks [110]. The critic network can be presented by a Q function and written as $y = r + \gamma Q_\theta(s', a')$, where $r$ denotes the reward and $\gamma$ the discount factor. The critic network can be updated by the following mean square error:

$$L = \frac{1}{N}\sum_t (Q_\theta(s,a) - y)^2 \tag{6.28}$$

where $N$ represents the size of mini-batches. The actor-network parameter $\phi$ can be updated by the deterministic policy gradient method as the following:

$$\nabla_\phi J \approx \frac{1}{N}\sum_t \left(\nabla_a Q(s,a)|_{s=(o_i),a=\pi_{\phi_i}(s)} \nabla_\phi \pi_\phi(s)\Big|_{s=o_i}\right) \tag{6.29}$$

With a double critic network, TD3 can overcome the overestimation of the Q problem in DDPG. The target can be updated by the following two critic networks:

$$y_1 = r + \gamma Q_{\theta_1'}(s', a') \tag{6.30}$$

$$y_2 = r + \gamma Q_{\theta_2'}(s', a') \tag{6.31}$$

A minimum Q function is selected from (6.30) and (6.31) to replace the conventional target in DDPG.

$$y = r + \gamma \min_{i=1,2} Q_{\theta_i'}(s', a') \tag{6.32}$$

#### 6.4.4.2 Combined Neighboring Experience Pool Replay - TD3

The conventional TD3 algorithm incorporates experience replay to reduce sample correlation and enhance sample utilization efficiency. However, it is essential to recognize that not all empirical samples have an equal impact on the training process. Randomly selecting samples with uniform probability may lead to including low-value or even insignificant samples in the training process. This could potentially result in suboptimal outcomes during TD3 training. Therefore, the MEPR strategies [111] and MLEPR strategies [112] are proposed to select prioritized samples for updating networks in TD3.

To address the specific challenge of our proposed task, it is crucial to create multiple individual experience replay pools based on neighboring labels, which encompass the supply-demand ratio $F_t^{SDR}$, storage state of the SES $F_t^{SOC}$, and potential arbitrage opportunity $F_t^{AO}$. Therefore, a CNEPR technique is proposed to update the networks based on the priority of empirical samples, giving higher importance to neighboring and high-value samples. Fig. 6.2 illustrates the CNEPR technique.

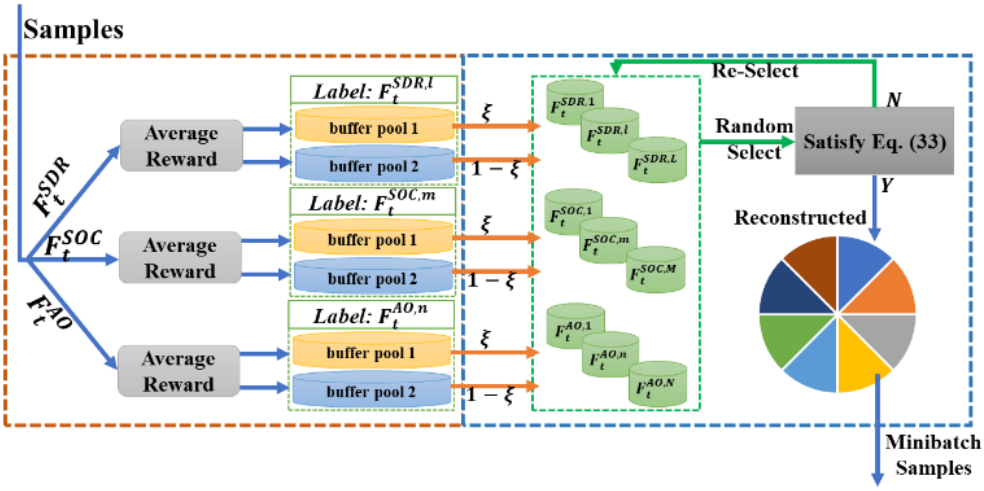



Fig. 6.2 Combined neighboring experience pool replay technique

The steps for constructing and utilizing multiple independent buffer pools through the CNEPR technique are as follows:

1) During the network initialization process, $F_t^{SDR}$, $F_t^{SOC}$ and $F_t^{AO}$ are divided equally into L, M and N sub-labels based on their ranges, respectively. Two experience pools are established separately for each sub-label to store high- and low-value

samples. There are a total of (L+M+N)*2 pools, with labels denoted as:

$< F_1^{SDR,1}, F_2^{SDR,1}, \cdots F_1^{SDR,l}, F_2^{SDR,l}, \cdots F_1^{SDR,L}, F_2^{SDR,L} >$, and

$< F_1^{SOC,1}, F_2^{SOC,1}, \cdots F_1^{SOC,m}, F_2^{SOC,m}, \cdots F_1^{SOC,M}, F_2^{SOC,M} >$, and

$< F_1^{AO,1}, F_2^{AO,1}, \cdots F_1^{AO,n}, F_2^{AO,n}, \cdots F_1^{AO,N}, F_2^{AO,N} >$.

where subscript 1 is used to store high-value samples, and subscript 2 to store low-value samples. The average reward of coupled pools with the same superscript, such as $F_1^{SDR,l}$ and $F_2^{SDR,l}$, is set to 0.

2) When a new sample is added with labels $F_k^{SDR,l}$, $F_k^{SOC,m}$ or $F_k^{AO,n}$, ($k$=*1*,*2*), first, obtain the rewards corresponding to the superscripts that match the three labels. Then calculate the reward of the sample itself. If the reward is greater than the corresponding reward, store it in the corresponding pool with subscript 1. Otherwise, store it in pool 2. Repeat this process three times corresponding to three rewards, respectively, and then renew the reward.
3) During the training, samples are randomly selected from pools 1 and 2 under the same label with probabilities $\xi$ and $1 - \xi$, respectively, and are combined into a new experience buffer pool. Therefore, there are a total number of (L+M+N) pools, as green pools in Fig. 6.2.
4) Then, the real label $F_k^{SDR,l\prime}$, $F_k^{SOC,m\prime}$ and $F_k^{AO,n\prime}$, (k=1,2), is obtained corresponding to the current training data, and the randomly selected pools with labels $F_k^{SDR,l}$, $F_k^{SOC,m}$ and $F_k^{AO,n}$, (k=1,2), from the new combined pools. If the randomly selected pools satisfy:

$$|l' - l| + |m' - m| + |n' - n| \leq \tau \quad (6.33)$$

where $\tau$ represents a distance, they can be used to reconstruct the new training pool; Otherwise, reselect and repeat step 4).
5) Finally, the mini-batch samples are randomly selected from the reconstructed pool.

## 6.5 Case Studies

### 6.5.1 Experiment Setting

In case studies, the DWS service operates 24 hours a day, with the DW coefficients updated every hour. The SES construction has a trial period of 7 days, after which it transitions to long-term operation once the construction is complete. During operation, the average profit of the SES is monitored. If the average profit remains below 80% of the profit during the setup period for a continuous period of 7 days, a reconstruction is initiated. Participants are re-matched and reselected for the next operating cycle. Table 6.1 lists the parameter settings for intraday balance and virtual storage extraction.

Table 6.1 Parameter setting for SES balance and virtual operation

| $\beta^{FIT}$ | $\beta^{TOU}$ | $\beta^{SDR,UP}$ | $\beta^{SDR,LOW}$ | $\beta^{SOC,UP}$ | $\beta^{SOC,LOW}$ | $\beta^{AO,UP}$ | $\beta^{AO,LOW}$ |
|---|---|---|---|---|---|---|---|
| 0.8 | 1.2 | 0.9 | 0.1 | 0.9 | 0.1 | 0.85 | 0.15 |

This study is assumed to contain 2000 high price-tolerance prosumers who have installed household PV and ESS and are potential SES participants. Table 6.2 lists the parameters for PV and battery. The PV home electricity data are obtained from the Ausgrid data [102], and the energy market electricity prices are from AEMO [103] from July 2021 to June 2022.

Table 6.2 Parameter setting for household PV-ESS

| PV | Battery | | | |
|---|---|---|---|---|
| $P^m$ | $E^m$ | $\eta^{ch}$ | $\eta^{dch}$ | $P^m$ |
| 10kW | 15kWh | 0.95 | 1.05 | 7.5kW |

Before providing dynamic DW coefficients, ESP should calculate the potential arbitrage opportunity $F_t^{AO}$, which involves the price prediction method. We choose the CTSGAN as the electricity price prediction method, which is the same as [101]. The hyperparameter settings of CNEPR-TD3 are given as follows: the critic learning rate and actor learning rate are set to 0.001, respectively; the discount factor is set to 0.9; the numbers of sub-labels for three labels are set to 5; the selection probability for experience pool 1 is set to 0.8; the distance $\tau$ is set to 0.6; the capacities of two experience pools are set to 100000; the update interval of the strategy network is set to 2; the maximum number of episodes is set to 10000. All simulations are implemented in Python 3.7 with the packages of Pytorch on VSCode and run on a PC with an AMD RYZEN 9 3950X CPU and an NVIDIA RTX 3080 GPU.

In subsequent experiments, Case 1 is used to verify the proposed CNEPR-TD3 algorithm, contract threshold $\xi$ and virtual storage and extraction trigger condition selection.

**Case 1:** The model uses a matching mechanism and credit-based DWS (w/ dynamic coefficients, w/ virtual operation). Dynamic coefficients and the scheduling strategy of the SES need to be optimized.

Afterwards, we employ Cases 2, 3, and 4 for comparison with Case 1, aiming to identify the contributions of our proposed DWS with dynamic coefficients and matching mechanism to the proposed model through ablation experiments.

**Case 2:** The model uses a matching mechanism and credit-based DWS (w/ dynamic coefficients, w/o virtual operation).

**Case 3:** The model uses a matching mechanism and credit-based DWS (w/ dynamic coefficients, w/ virtual operation). However, the re-matching process is discarded, which means that the SES will continue to run in its initial state indefinitely.

**Case 4:** The model uses credit-based DWS (w/ fixed coefficients, w/ virtual operation). The credit is solely used for recording the deposit and withdrawal of electricity [66], and no matching between ESP and prosumers is conducted. Only the scheduling strategy of the SES needs to be optimized.

Finally, with **Case 1**, the relationship between dynamic DW coefficients and three related factors is analyzed in Section 6.6.

### 6.5.2 Performance of CNEPR-TD3

The DW coefficients $\alpha_t^{stg}$ and $\alpha_t^{ext}$ and trading strategies $P_t^{SPO,ch}$ and $P_t^{SPO,dis}$ simultaneously in our proposed model (Case 1) to test the effectiveness of the proposed CNEPR-TD3 algorithm. Two modified experience-pool-based TD3 [111, 112], the conventional TD3 and DDPG are introduced as benchmarks. The comparison results are presented in Fig. 6.3.

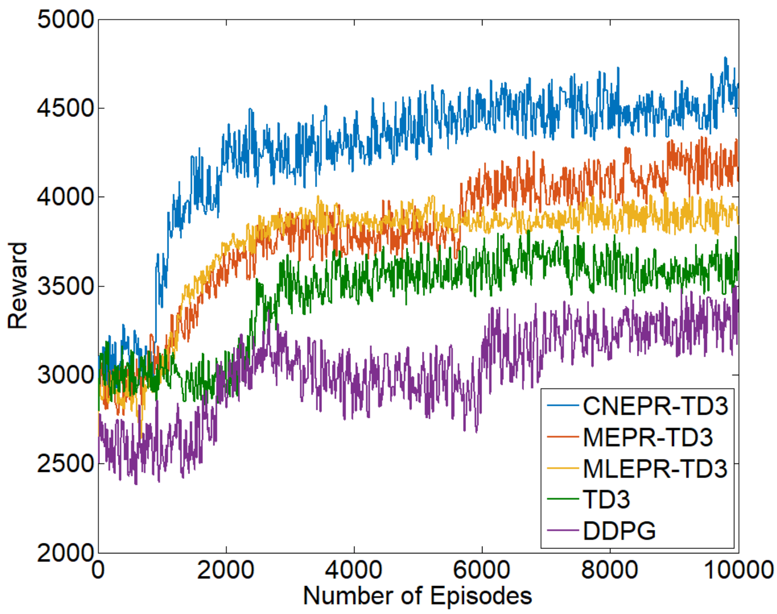



Fig. 6.3 Comparison of reward with different DRL algorithms

The vertical axis represents the reward. It can be observed that all TD3 algorithms exhibit superior convergence performance, while the conventional DDPG algorithm exhibits poor convergence, characterized by large oscillations at the end of the learning process. The MLEPR strategy, which utilizes a multi-level experience pool, slightly improves performance compared to basic TD3; however, it enables TD3 to converge faster but does not provide further improvements in the later stages. The convergence effect of MEPR and MLEPR is similar before 6000 iterations, but because MEPR can recognize high- and low-value samples, its performance surpasses MLEPR's after 6000 iterations. Based on the mechanism of recognizing high- and low-value samples in MEPR, the proposed CNEPR-TD3 can utilizes $F_t^{SDR}$, $F_t^{SOC}$ and $F_t^{AO}$ labels to find the neighboring training set through pool reconstruction, demonstrating relatively superior performance in the early stages of training.

Compared to MEPR-TD3 and MLEPR-TD3, CNEPR-TD3 can significantly improve the reward by more than A$500 per day.

Table 6.3 Profit comparison with different DRL methods for Jul. 2021 and Jan. 2022 (A$/Day)

| | | Bill Reduction | Participants' Profit | ESP's Profit |
|---|---|---|---|---|
| Jul. 2021 | CNEPR-TD3 | 1.56 | 0.72 | 1086.48 |
| | MEPR-TD3 | 1.57 | 0.70 | 1046.09 |
| | MLEPR-TD3 | 1.52 | 0.64 | 953.57 |
| | TD3 | 1.73 | 0.56 | 735.28 |
| | DDPG | 1.84 | 0.49 | 642.39 |
| Jan. 2022 | CNEPR-TD3 | 1.67 | 0.85 | 1435.60 |
| | MEPR-TD3 | 1.64 | 0.80 | 1257.65 |
| | MLEPR-TD3 | 1.68 | 0.75 | 1167.42 |
| | TD3 | 1.79 | 0.70 | 949.94 |
| | DDPG | 1.91 | 0.59 | 785.07 |

Table 6.3 lists the average bill reduction, participants' allocated profit, ESP's profit with the proposed CNEPR-TD3 and other benchmarking algorithms for July 2021 and January 2022. One can find that all algorithms can provide dynamic coefficients that meet participants' expectations, effectively reducing the bill. However, when comparing the three improved TD3, the dynamic coefficients provided by DDPG and basic TD3 result in excessive bill reductions for participants, with allocated profit obtained by participants of A$1.73/day and A$1.84/day in July 2021, respectively. For the SES, excessively high bill reductions indicate unreasonable dynamic coefficients. Moreover, it also leads to suboptimal optimization of market trading strategies, thereby affecting the overall profit. As a result, both individual profit allocation and ESP profit allocation are significantly reduced, with ESP's profit of A$642.39 (DDPG) and A$735.28 (TD3). All modified TD3 can effectively provide reasonable coefficients. Therefore, for participants, the daily bill reduction amounts are similar. However, our proposed CNEPR-TD3, which utilizes a pool reconstruction method based on neighboring experience replay, can provide better trading strategies. This results in daily ESP profit of A$1086.48 (July 2021) and A$1435.60 (January 2022), while MEPR-TD3 only achieves profits of A$1046.09 (July 2021) and A$1257.65

(January 2022), respectively.

Based on the above comparative analysis, it can be observed that our proposed CNEPR-TD3 algorithm exhibits superiority for trading strategy and dynamic coefficient optimization of DWS. Subsequent optimizations related to dynamic coefficients and trading strategies are conducted with CNEPR-TD3.

### 6.5.3 Comparison of Different $\xi^{pro}$ for Threshold $\xi$ Selection

Comparison of profit for ESP and prosumers with different $\xi^{pro}$ are simulated and analyzed to select the best contract threshold $\xi$ with Case 1. ESP needs to conduct a detailed investigation and set an appropriate value for $\xi^{pro}$ to ensure that prosumers are willing to accept DWS and participate in the SES. It should be noted that the $\xi^{pro}$ is an attribute of the prosumer group, which is determined by many factors of the prosumers [116]. For ESP, selecting the appropriate value for $\xi^{ESP}$ can help maximize the profit of the SES. According to (6.9), integrating $\xi^{pro}$ and $\xi^{ESP}$ can obtain the optimal value of $\xi$ for the current SES.

We test the performance of our proposed framework at different values of $\xi^{pro}$, ranging from 0% to 100% with intervals of 10%, to simulate the willingness of different prosumers. The results include the number of participants, average profits for prosumers, and corresponding profits for ESP. In addition, we propose a new index to indicate the deposit-withdrawal gain as: $G_{dw} = \frac{1}{T}\sum_{t=1}^{T} \alpha_t^{dps} / \alpha_t^{wtd}$. Since $\alpha_t^{dps} \geq 1$ and $\alpha_t^{wtd} \leq 1$, $G_{dw}$ must be larger than 1. Therefore, the SES has negligible loss. If $G_{dw} > 1/\eta^{adm}$, it indicates negative loss during the deposit-withdrawal process, which means that the withdrawal electricity exceeds the deposit amount.

Table 6.4 Comparison of remaining prosumer number, reduction bill and ESP's profit with different $\xi^{pro}$

| $\xi^{pro}$ | $Obj$ | Number | Bill Reduction | Participants' Profit | ESP's Profit | $G_{dw}$ |
|---|---|---|---|---|---|---|
| 0% | 1855.47 | 2000 | 0.13 | 0.40 | 793.33 | 1.03 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 10% | 2889.94 | 1823 | 0.56 | 0.51 | 935.81 | 1.13 |
| 20% | 4087.73 | 1734 | 1.04 | 0.66 | 1144.44 | 1.25 |
| 30% | 4695.30 | 1509 | 1.56 | 0.78 | 1171.99 | 1.37 |
| 40% | 4487.43 | 1193 | 2.08 | 0.84 | 1090.43 | 1.49 |
| 50% | 4131.86 | 937 | 2.58 | 0.91 | 855.79 | 1.62 |
| 60% | 2373.45 | 473 | 3.09 | 0.96 | 455.66 | 1.74 |
| 70% | 1587.18 | 278 | 3.58 | 1.06 | 295.61 | 1.85 |
| 80% | 413.64 | 65 | 4.14 | 1.11 | 72.37 | 1.99 |
| 90% | 117.83 | 17 | 4.66 | 1.14 | 19.32 | 2.11 |

From Table 6.4, one can find that there is a significant difference in the prosumers' bill reduction, allocated profit and ESP's profit under different assumed thresholds $\xi^{pro}$. When $\xi^{pro}$ is assumed to 10%, i.e., compared to the original electricity bill $\mathcal{B}_j^{ORI}$, using the DWS can further effectively reduce the bill by 10% (A$0.56/day) for 1823 out of 2000 prosumers. Through the profit allocation process, all participants can still receive A$0.51/day, and the ESP can receive A$935.81/day. The total deposit-withdrawal gain is 1.133. While $\xi^{pro}$ is set to 90%, although the daily bill for each participant can be reduced at most by A$4.66; only 17 out of 2000 individuals meet the threshold. Consequently, the ESP can only establish a small-scale SES to participate in the wholesale market, resulting in a small daily profit of only A$19.32. Additionally, we can observe that when $\xi^{pro} \geq 10\%$, all the deposit-withdrawal gains $G_{dw}$ are greater than 1.11. This implies that even considering $\eta^{adm} = 0.9$ in the SES, the final withdrawal amount of electricity is guaranteed to be greater than the deposit amount.

We can also observe that, without considering $\xi^{pro}$, the objective function reaches its maximum daily value of A$4695.30 when $\xi^{pro}$ is set to 30%. This implies that the optimal value for $\xi^{ESP}$ is 30%. At the same time, the ESP also achieves its maximum profit, which amounts to A$1171.99/day. It should be noted that the lower limit of $\xi$ depends on both $\xi^{ESP}$ and $\xi^{pro}$ with (6.9). If the prosumers are willing to accept $\xi^{pro} \leq 30\%$, the final threshold $\xi$ is 30%. However, if the prosumers just accept $\xi^{pro} > 30\%$, the ESP can only choose actual $\xi^{pro}$ as the threshold. In the subsequent case studies, assuming that prosumers can accept $\xi^{pro} \leq 30\%$, then the final $\xi$ is set to 30%.

### 6.5.4 Comparison of Different Trigger Conditions

When $F_t^{SDR}$, $F_t^{SOC}$ and $F_t^{AO}$ meet the trigger conditions, ESP triggers the virtual storage and virtual extraction; therefore, ESP directly sells or purchases the electricity that participants store or extract in real-time on the electricity market rather than engaging in physical-level storage or extraction. In consideration of the rapid degradation caused by overcharging and over-discharging of the SES, we set $\beta^{SOC,LOW}$ to 0.1 and $\beta^{SOC,UP}$ to 0.9. Additionally, for simplicity, we set $\beta^{SDR,LOW} = 1 - \beta^{SOC,UP}$, and $\beta^{AO,LOW} = 1 - \beta^{AO,UP}$. Fig. 6.4 presents the total profit of SES under different $\beta^{SDR,LOW}$ and $\beta^{AO,LOW}$ with an interval of 0.05 from 0 to 0.30 in July 2021.

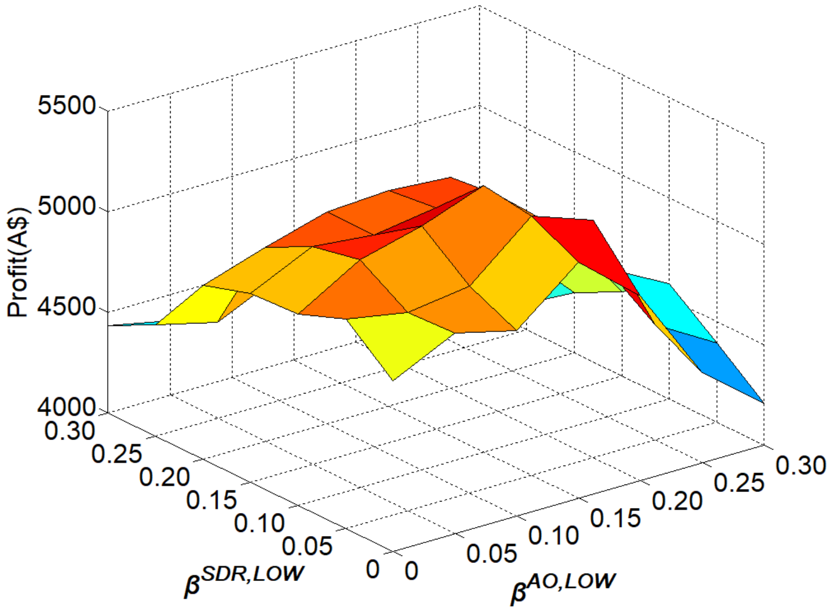



Fig. 6.4 Total profit of SES with different triggering conditions for virtual operation

From Fig. 6.4, one can find that different triggering conditions have a significant impact on the overall profit for the SES. When $\beta^{SDR,LOW} = 0$ and $\beta^{AO,LOW} = 0$, meaning that the virtual operation is only triggered when $F_t^{SOC} > 0.9$ or $< 0.1$, the daily total profit of the SES is approximately A$4571. When $\beta^{SDR,LOW} = 0.3$ and $\beta^{AO,LOW} = 0$, the overall profit decreases to about A$4462/day. However, when $\beta^{SDR,LOW} = 0$ and $\beta^{AO,LOW} = 0.3$, the overall profit further decreases and is about

A$4284/day. This indicates that compared to $\beta^{SDR,LOW}$, an unreasonable setting of $\beta^{AO,LOW}$ to an excessively large value will result in an even more significant decrease in overall profit. We also observe that when $\beta^{SDR,LOW}$ is set to 0.1 and $\beta^{AO,LOW}$ is set to 0.15, the overall profit is maximized at approximately A$4753. To further analyze the reasons, we compiled the trigger rates under different triggering conditions, as shown in Table 6.5.

Table 6.5 Profit, gain and trigger rates with different conditions

| Triggering Condition | | | Profit(A$) | Gain | Trigger Rate |
|---|---|---|---|---|---|
| $\beta^{SOC,LOW}$ | $\beta^{SDR,LOW}$ | $\beta^{AO,LOW}$ | | | |
| 0.1 | - | - | 4324.31 | - | 2.24% |
| | 0.1 | - | 4421.90 | 2.26% | 3.73% |
| | - | 0.15 | 4655.26 | 7.65% | 3.21% |
| | 0.1 | 0.15 | 4753.21 | 9.92% | 7.87% |
| | 0.3 | 0.3 | 4046.63 | -6.42% | 24.75% |

Table 6.5 presents the total profit, profit gain, and trigger rate under different triggering conditions. The first row of the table represents the baseline value, which corresponds to the profit when only $F_t^{SOC}$ triggers are considered. When $F_t^{SDR}$ is introduced with $\beta^{SDR,LOW}$ of 0.1 for virtual operations, the trigger rate increases from 2.24% to 3.73%, and the profit increases by 2.26%. However, if $F_t^{AO}$ is introduced instead of $F_t^{SDR}$, the trigger rate does not change significantly with 3.21%, but the profit significantly rises by 7.65%. Taking all three conditions into account, the profit increases to A$4753.21/day, and the corresponding trigger rate is 7.87%. However, if both $\beta^{SDR,LOW}$ and $\beta^{AO,LOW}$ are set to 0.3, the trigger rate increases to 24.75%, but the profit decreases by 6.42% compared to the baseline. This demonstrates that introducing reasonable virtual operations can contribute to profit improvement, but increasing the trigger rate through unreasonable coefficient values may lead to a decrease in profit.

### 6.5.5 Effectiveness Verification through Ablation Experiments

The above case studies validate the effectiveness of the proposed CNEPR-TD3 algorithm, threshold $\xi$ selection, and triggering conditions for virtual operations. In this section, we conduct ablation experiments and calculate the profits for four cases to verify the effectiveness of the proposed model. Table 6.6 illustrates the ESP's average daily profit, and prosumers' average daily profit, including bill reduction and allocated profit, and prosumer number from July 2021 to June 2022.

Table 6.6 Ablation experiments for profit (A$/day) and prosumer number

| | Case | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| Component | Dyn. Coef. | w/ | w/ | w/ | w/o |
| | Mat. Mech. | w/ | w/ | w/ | w/o |
| | Virt. Oper. | w/ | w/o | w/ | w/ |
| | Re-Const. | w/ | w/ | w/o | w/o |
| Average Profit (A$/day) | Prosumer | 2.43 | 2.19 | 1.63 | 0.71 |
| | ESP | 1327.92 | 1060.69 | 886.82 | 1426.68 |
| | Total | 4994.79 | 4242.76 | 3379.09 | 2853.36 |
| | Loss | - | 15.06% | 32.35% | 42.87% |
| Prosumer Number | | 1509 | 1453 | 1529 | 2000 |

Table 6.6 shows that our proposed model (Case 1) with dynamic coefficients and the matching mechanism exhibits the best total profits, with an average value of A$4994.79/day. In comparison to Case 1, Case 2 lacks the virtual operation component. It can be observed that both prosumers and ESPs experienced a decline in profits. The number of participants decreased by 53 compared to Case 1. Case 3 shows no re-matching and reconstruction after the profit decrease. As a result, the number of participants remains constant at 1529, the same as the initial combination. However, without reconstruction, the overall profits drop significantly. ESPs' profits are only A$886.82/day, and the prosumers' profits are A$1.63/day. In Case 4, the DWS coefficients are fixed, meaning that participants cannot obtain additional electricity gains from the SES. However, they can still receive profits from the allocation of the SES, which amounts to only A$0.71/day.

Additionally, due to the absence of a matching mechanism, it is assumed that all participants are involved. Even under this assumption, the overall profits are significantly lower than Case 1, totalling only A$2853.36/day. Based on the above analysis, it can be seen that dynamic coefficients and matching mechanisms are the most crucial components of our proposed model. Their removal would lead to a loss of 42.87%. Subsequently, re-matching and reconstruction are also important, with a loss of 32.35% when removed. Finally, removing the virtual operations results in a loss of 15.06%.

### 6.5.6 Dynamic Coefficients $\alpha_t^{dps}$ ,$\alpha_t^{wtd}$ with Related Factors

Fig. 6.5 illustrates the variations in the coefficients of $\alpha_t^{dps}$ and $\alpha_t^{wtd}$ and their relationship with $F_t^{SDR}$, $F_t^{SOC}$ and $F_t^{AO}$ throughout a 24-hour period on July 30, 2022. As shown, $\alpha_t^{dps}$ reaches its highest value of 1.53 at 18:00, indicating that the depositing electricity will be credited with a coefficient of 1.53. At this time, the corresponding $\alpha_t^{wtd}$ is 1, meaning that there is no discount on electricity withdrawal during this time period. By analyzing the relationship between the coefficients $\alpha_t^{dps}$ and $\alpha_t^{wtd}$ and related factors $F_t^{SDR}$, $F_t^{SOC}$, and $F_t^{AO}$, it can be determined that the high $\alpha_t^{dps}$ value at 18:00 is due to the following reasons: 1) A low $F_t^{SDR}$ implies that a smaller portion of the PV-generated electricity can be deposited while the withdrawal of electricity for consumption increases; 2) $F_t^{SOC}$ for the SES being at a low level contributes to the higher $\alpha_t^{dps}$ value; 3) High future $F_t^{AO}$ further contributes to the high value of $\alpha_t^{d}$. Additionally, one can find that $F_t^{SOC}$ drops below 0.1 at 20:00, triggering virtual extraction. This means that the SES still needs to purchase electricity directly from the spot market to meet the prosumers' demand. As a result, the $\alpha_t^{dps}$ and $\alpha_t^{wtd}$ remain high.

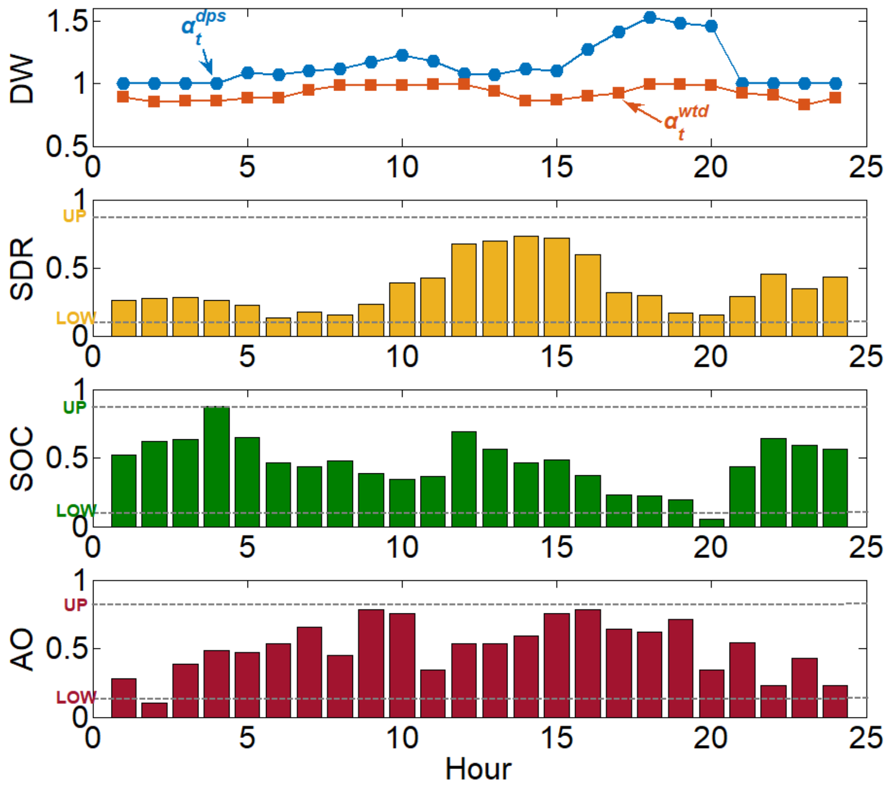



Fig. 6.5 Relationship between DW coefficients $\alpha_t^d$ ,$\alpha_t^w$ and related factors

## 6.6 Chapter Summary

In this chapter, we proposed a novel SES aggregation framework with DWS and a matching mechanism, considering high price-tolerance prosumers unwilling to cooperate with ESP to change their electricity consumption behaviors or carry out additional actions. A credit-based DWS method was designed, allowing the prosumers to deposit and withdraw electricity and gain benefits. Moreover, an ESP and prosumer matching mechanism was devised to help the ESP identify prosumers whose electricity consumption behaviors align with the ESP's trading strategy, facilitating the construction of the SES. The CNEPR-TD3 was developed to optimize the dynamic coefficients in DWS and the trading strategies for participating in the wholesale electricity market, effectively improving the overall profit of the SES.

The effectiveness of the proposed CNEPR-TD3 algorithm, threshold ξ selection, and triggering conditions for virtual operation were validated by simulation results. Four cases were used for ablation experiments, and the results showed that the design of dynamic DWS and the matching mechanism had the most significant impact on the overall SES profit, with an increase of 42.87%.

# CHAPTER 7

# Conclusions and Future Work

*This chapter summarizes the conclusion of this thesis and discusses its limitations and directions for future research. These include the prediction models, the selection preferences of VPP participants, and the potential impacts of more frequent settlement in electricity market transactions.*

## 7.1 Conclusions

This thesis exploits several recruitment-participation frameworks for VPP construction and coordination to achieve optimal market operations. Different DER participants often have different investment preferences, with some being more conservative and others more aggressive. This thesis aims to assist ESPs in obtaining high profits by constructing different VPPs with various business models, attracting more DER participants. For VPPs, profits typically come from buying low and selling high electricity prices, making price prediction technology the cornerstone of VPP profitable optimization.

A novel CTSGAN was proposed to capture the distribution of each period and preserve temporal variations of wholesale electricity prices. Then, with the proposed CTSGAN point prediction model, electricity prices can be generated. By adjusting the random noise input of CTSGAN, the point prediction model was eventually transformed into an interval prediction model, which can directly generate PIs with high reliability and low sharpness.

In order to meet the risk-profit requirements of different DER participants, this thesis presented a non-ESP-centric VPP construction approach where multiple ESPs offered different VPP plans and recruited DERs. A payoff allocation method based on fairness and incentives was proposed, and MADRL was employed to solve the rebate competition problem among ESPs, achieving a stable solution among ESPs.

Additionally, this thesis proposed a flexible and profitable VPP recruitment approach to maximize profits when UAO occurs, that integrated regular and casual recruitment mechanisms, catering to conservative and ambitious DER participants. A CEPR-TD3 algorithm was proposed to optimize the incentive coefficient, which can encourage DER participants and help the ESP achieve high profits. Finally, for high price-tolerance residential prosumers, this thesis proposed a framework to aggregate VPP, which allowed prosumers to participate without requiring additional actions or coordination with ESP. A credit-based electricity DWS method was proposed, allowing prosumers have the freedom to deposit or withdraw electricity with dynamic coefficients and receive electricity gains. A CNEPR-TD3 was proposed to optimize reasonable dynamic coefficients and energy trading strategies, effectively improving the VPP’s overall profit.

## 7.2 Future Work

Although this thesis covered multiple VPP construction and coordination approaches and novel business models for different risk-return DER participants, there are undoubtedly many directions for future work on the main topics addressed. Some are summarized below.

1) The CTSGAN prediction model presented in this thesis, which incorporates weather and other factors, can enhance the accuracy of electricity price prediction. Nevertheless, electricity price prediction remains challenging as prices are determined by the behaviors of multiple participants in the market, which various factors, including market psychology and policy changes, can influence. In the future, electricity price forecasting will involve more than simple regression and fitting; generative models considering multiple factors will require further research. Moreover, with the recent rise of large language models, it's possible to simulate the trading tendencies of electricity market participants using these models, which could further enhance the accuracy of predictions.

2) This thesis assumes that DER participant investment preferences are idealized. It is noted that the coefficients applied in investment preference may be influenced by factors such as social development, education level, and local regulations, leading to changes in the weights. Current research in relevant fields generally assumes that participants can be incentivized through an increase in profits, without conducting empirical investigations into actual participants' attitudes and consumption habits, limiting the transferability of these models to real-world applications. Future works can adopt a human-centered approach by utilizing advanced research methods to understand participants' perspectives and preferences better and identify the factors beyond pricing that motivate individuals to participate in VPPs.

3) The research discussed in this thesis commenced in July 2020, when the AEMO transactions operated on a 30-minutes settlement. Therefore, all forecasting, scheduling, and optimization aspects discussed in this thesis were based on 30-minute intervals. However, starting in October 2021, settlements shifted to a five-minute interval. Due to the shorter settlement period, market dispatch and operations require faster responses, potentially posing challenges for market participants and dispatch algorithms. The more frequent market trading and quicker decision-making behaviors may impact the market, warranting further exploration. The increased frequency of settlements may also influence risk management strategies. Market participants may need to reevaluate their risk exposure and adjust their risk management strategies to accommodate the faster market settlements. Thus, improved models and methods are necessary based on the current works in this thesis.

# BIBLIOGRAPHY


[1] Y. Wang, J. Qiu, Y. Tao, X. Zhang, and G. Wang, “Low-carbon oriented optimal energy dispatch in coupled natural gas and electricity systems,” *Applied Energy*, vol. 280, pp. 115948, 2020.

[2] X. Zhang, K. W. Chan, H. Wang, B. Zhou, G. Wang, and J. Qiu, “Multiple group search optimization based on decomposition for multi-objective dispatch with electric vehicle and wind power uncertainties,” *Applied Energy*, vol. 262, pp. 114507, 2020.

[3] Z. Xu, Y. Guo, and H. Sun, “Competitive Pricing Game of Virtual Power Plants: Models, Strategies, and Equilibria,” *IEEE Transactions on Smart Grid*, 2022.

[4] W. Chen, J. Qiu, J. Zhao, Q. Chai, and Z. Y. Dong, “Bargaining game-based profit allocation of virtual power plant in frequency regulation market considering battery cycle life,” *IEEE Transactions on Smart Grid*, vol. 12, no. 4, pp. 2913-2928, 2021.

[5] H. Saberi, C. Zhang, and Z. Y. Dong, “Capacity of virtual energy storage system for frequency regulation services via a data-driven distributionally robust optimization method,” *IEEE Transactions on Power Systems*, 2022.

[6] ARENA. "AGL Virtual Power Plant," https://arena.gov.au/projects/agl-virtual-power-plant/.

[7] Tesla. "South Australia's Virtual Power Plant: Frequently Asked Questions," https://www.tesla.com/en_au/support/energy/savpp-faqs.

[8] F.-f. Yang, and X.-g. Zhao, “Policies and economic efficiency of China's distributed photovoltaic and energy storage industry,” *Energy*, vol. 154, pp. 221-230, 2018.

[9] C. A. Hill, M. C. Such, D. Chen, J. Gonzalez, and W. M. Grady, “Battery energy storage for enabling integration of distributed solar power generation,” *IEEE Transactions on smart grid*, vol. 3, no. 2, pp. 850-857, 2012.

[10] K. N. Singh, N. Das, M. Maurya, A. K. Goswami, and N. B. D. Chudhury, "An Intelligent Bidding Strategy based on Social Welfare of Virtual Power Plant considering Carbon Trading." pp. 1-5.

[11] H. Liu, J. Qiu, and J. Zhao, “A data-driven scheduling model of virtual power plant using Wasserstein distributionally robust optimization,” *International Journal of Electrical Power & Energy Systems*, vol. 137, pp. 107801, 2022.

[12] Y. Ye, D. Qiu, M. Sun, D. Papadaskalopoulos, and G. Strbac, “Deep reinforcement learning for strategic bidding in electricity markets,” *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1343-1355, 2019.

[13] A. Zangeneh, A. Shayegan-Rad, and F. Nazari, “Multi-leader–follower game theory for modelling interaction between virtual power plants and distribution company,” *IET generation, transmission & distribution*, vol. 12, no. 21, pp. 5747-5752, 2018.

[14] F. Cucchiella, I. D'adamo, and M. Gastaldi, “Photovoltaic energy systems with battery storage for residential areas: an economic analysis,” *Journal of Cleaner Production*, vol. 131, pp. 460-474, 2016.

[15] J. Hoppmann, J. Volland, T. S. Schmidt, and V. H. Hoffmann, “The economic viability of battery storage for residential solar photovoltaic systems–A review and a simulation model,” *Renewable and Sustainable Energy Reviews*, vol. 39, pp. 1101-1118, 2014.

[16] Z. Ren, G. Grozev, and A. Higgins, “Modelling impact of PV battery systems on energy consumption and bill savings of Australian houses under alternative tariff structures,”

*Renewable Energy*, vol. 89, pp. 317-330, 2016.

[17] K. R. Khalilpour, and A. Vassallo, "Technoeconomic parametric analysis of PV-battery systems," *Renewable Energy*, vol. 97, pp. 757-768, 2016.

[18] L. Ju, Z. Yin, Q. Zhou, Q. Li, P. Wang, W. Tian, P. Li, and Z. Tan, "Nearly-zero carbon optimal operation model and benefit allocation strategy for a novel virtual power plant using carbon capture, power-to-gas, and waste incineration power in rural areas," *Applied Energy*, vol. 310, pp. 118618, 2022.

[19] N. Holjevac, T. Capuder, N. Zhang, I. Kuzle, and C. Kang, "Corrective receding horizon scheduling of flexible distributed multi-energy microgrids," *Applied energy*, vol. 207, pp. 176-194, 2017.

[20] H. Yang, and K. R. Schell, "GHTnet: Tri-Branch deep learning network for real-time electricity price forecasting," *Energy*, vol. 238, pp. 122052, 2022.

[21] C. Xiao, D. Sutanto, K. M. Muttaqi, M. Zhang, K. Meng, and Z. Y. Dong, "Online sequential extreme learning machine algorithm for better predispatch electricity price forecasting grids," *IEEE Transactions on Industry Applications*, vol. 57, no. 2, pp. 1860-1871, 2021.

[22] W. Li, and D. M. Becker, "Day-ahead electricity price prediction applying hybrid models of LSTM-based deep learning methods and feature selection algorithms under consideration of market coupling," *Energy*, vol. 237, pp. 121543, 2021.

[23] J. Lago, G. Marcjasz, B. De Schutter, and R. Weron, "Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark," *Applied Energy*, vol. 293, pp. 116983, 2021.

[24] T. Kempitiya, S. Sierla, D. De Silva, M. Yli-Ojanperä, D. Alahakoon, and V. Vyatkin, "An Artificial Intelligence framework for bidding optimization with uncertainty in multiple frequency reserve markets," *Applied energy*, vol. 280, 2020.

[25] B. Uniejewski, G. Marcjasz, and R. Weron, "On the importance of the long-term seasonal component in day-ahead electricity price forecasting," *Energy economics*, vol. 79, pp. 171-182, 2019.

[26] W. P. Bell, P. Wild, J. Foster, and M. Hewson, "Revitalising the wind power induced merit order effect to reduce wholesale and retail electricity prices in Australia," *Energy economics*, vol. 67, pp. 224-241, 2017.

[27] S. Forrest, and I. MacGill, "Assessing the impact of wind generation on wholesale prices and generator dispatch in the Australian National Electricity Market," *Energy policy*, vol. 59, pp. 120-132, 2013.

[28] F. Nazifi, and F. Nazifi, "The pass-through rates of carbon costs on to electricity prices within the Australian National Electricity Market," *Environmental economics and policy studies*, vol. 18, no. 1, pp. 41-62, 2016.

[29] H. Higgs, and A. Worthington, "Stochastic price modeling of high volatility, mean-reverting, spike-prone commodities: The Australian wholesale spot electricity market," *Energy economics*, vol. 30, no. 6, pp. 3172-3185, 2008.

[30] Y. Dong, J. Wang, H. Jiang, and J. Wu, "Short-term electricity price forecast based on the improved hybrid model," *Energy conversion and management*, vol. 52, no. 8, pp. 2987-2995, 2011.

[31] L. Peng, S. Liu, R. Liu, and L. Wang, "Effective long short-term memory with differential

evolution algorithm for electricity price prediction," *Energy (Oxford)*, vol. 162, pp. 1301-1314, 2018.

[32] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms," *Applied energy*, vol. 221, pp. 386-405, 2018.

[33] T. Jonsson, P. Pinson, H. A. Nielsen, H. Madsen, and T. S. Nielsen, "Forecasting Electricity Spot Prices Accounting for Wind Power Predictions," *IEEE transactions on sustainable energy*, vol. 4, no. 1, pp. 210-218, 2013.

[34] Z. Yang, L. Ce, and L. Lian, "Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods," *Applied energy*, vol. 190, pp. 291-305, 2017.

[35] J. Nowotarski, and R. Weron, "Recent advances in electricity price forecasting: A review of probabilistic forecasting," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1548-1568, 2018.

[36] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Lower Upper Bound Estimation Method for Construction of Neural Network-Based Prediction Intervals," *IEEE transactions on neural networks*, vol. 22, no. 3, pp. 337-346, 2011.

[37] C. Wan, M. Niu, Y. Song, and Z. Xu, "Pareto optimal prediction intervals of electricity price," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 817-819, 2016.

[38] G. Marcjasz, B. Uniejewski, and R. Weron, "Probabilistic electricity price forecasting with NARX networks: Combine point or probabilistic forecasts?," *International Journal of Forecasting*, vol. 36, no. 2, pp. 466-479, 2020.

[39] D. T. Nguyen, H. T. Nguyen, and L. B. Le, "Dynamic pricing design for demand response integration in power distribution networks," *IEEE Transactions on power systems*, vol. 31, no. 5, pp. 3457-3472, 2016.

[40] M. B. Rasheed, M. A. Qureshi, N. Javaid, and T. Alquthami, "Dynamic pricing mechanism with the integration of renewable energy source in smart grid," *IEEE Access*, vol. 8, pp. 16876-16892, 2020.

[41] S. Yin, Q. Ai, J. Li, D. Li, and Q. Guo, "Trading mode design for a virtual power plant based on main-side consortium blockchains," *Applied Energy*, vol. 325, pp. 119932, 2022.

[42] F. Fang, S. Yu, and M. Liu, "An improved Shapley value-based profit allocation method for CHP-VPP," *Energy*, vol. 213, pp. 118805, 2020.

[43] S. Tian, L. Ying, G. Zhu, and X. Xie, "Study on the Profit Distribution Under the Participation of Virtual Power Plants in Market Bidding." pp. 198-202.

[44] S. S. Gougheri, M. Dehghani, A. Nikoofard, H. Jahangir, and M. A. Golkar, "Economic assessment of multi-operator virtual power plants in electricity market: A game theory-based approach," *Sustainable Energy Technologies and Assessments*, vol. 53, pp. 102733, 2022.

[45] W. Chen, J. Qiu, and Q. Chai, "Customized Critical Peak Rebate Pricing Mechanism for Virtual Power Plants," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 4, pp. 2169-2183, 2021.

[46] W. Chen, J. Qiu, J. Zhao, Q. Chai, and Z. Y. Dong, "Customized Rebate Pricing Mechanism for Virtual Power Plants Using a Hierarchical Game and Reinforcement Learning Approach," *IEEE Transactions on Smart Grid*, 2022.

[47] H. Liang, and J. Ma, “Data-driven resource planning for virtual power plant integrating demand response customer selection and storage,” *IEEE Transactions on Industrial Informatics*, vol. 18, no. 3, pp. 1833-1844, 2021.

[48] A. Oshnoei, M. Kheradmandi, F. Blaabjerg, N. D. Hatziargyriou, S. Muyeen, and A. Anvari-Moghaddam, “Coordinated control scheme for provision of frequency regulation service by virtual power plants,” *Applied Energy*, vol. 325, pp. 119734, 2022.

[49] H. Liu, J. Qiu, J. Zhao, Y. Tao, and Z. Y. Dong, “A Customer-centric Distributed Data-driven Stochastic Coordination Method for Residential PV and BESS,” *IEEE Transactions on Power Systems*, 2022.

[50] X. Lu, J. Qiu, C. Zhang, G. Lei, and J. Zhu, “Assembly and Competition for Virtual Power Plants With Multiple ESPs Through a “Recruitment–Participation” Approach,” *IEEE Transactions on Power Systems*, 2023.

[51] C.-W. Hsu, “Using a system dynamics model to assess the effects of capital subsidies and feed-in tariffs on solar PV installations,” *Applied energy*, vol. 100, pp. 205-217, 2012.

[52] J. Hu, W. Ma, Y. Xue, L. Yao, and D. Xie, “Quantification of Reserve Capacity Provided by Electric Vehicle Aggregator Based on Framework of Cyber-Physical-Social System in Energy,” *Automation of Electric Power System*, vol. 46, no. 18, pp. 46-54, 2022.

[53] E. F. Bompard, G. Abrate, R. Napoli, and B. Wan, “Multi-agent models for consumer choice and retailer strategies in the competitive electricity market,” *International Journal of Emerging Electric Power Systems*, vol. 8, no. 2, 2007.

[54] P. Cramton, “Electricity market design,” *Oxford Review of Economic Policy*, vol. 33, no. 4, 2017.

[55] Y. Wang, X. Ai, Z. Tan, L. Yan, and S. Liu, “Interactive dispatch modes and bidding strategy of multiple virtual power plants based on demand response and game theory,” *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 510-519, 2015.

[56] J. Wang, C. Guo, C. Yu, and Y. Liang, “Virtual power plant containing electric vehicles scheduling strategies based on deep reinforcement learning,” *Electric Power Systems Research*, vol. 205, pp. 107714, 2022.

[57] R. Ma, Z. Yi, Y. Xiang, D. Shi, C. Xu, and H. Wu, “A Blockchain-Enabled Demand Management and Control Framework Driven by Deep Reinforcement Learning,” *IEEE Transactions on Industrial Electronics*, 2022.

[58] J. Li, T. Yu, and X. Zhang, “Coordinated load frequency control of multi-area integrated energy system using multi-agent deep reinforcement learning,” *Applied Energy*, vol. 306, pp. 117900, 2022.

[59] X. Yang, L. He, Y. Xia, and Y. Chen, “Effect of government subsidies on renewable energy investments: The threshold effect,” *Energy Policy*, vol. 132, pp. 156-166, 2019.

[60] L. E. Susskind, S. McKearnen, and J. Thomas-Lamar, The consensus building handbook: A comprehensive guide to reaching agreement: Sage publications, 1999.

[61] S. R. Dabbagh, and M. K. Sheikh-El-Eslami, “Risk assessment of virtual power plants offering in energy and reserve markets,” *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3572-3582, 2015.

[62] L. Dong, S. Fan, Z. Wang, J. Xiao, H. Zhou, Z. Li, and G. He, “An adaptive decentralized economic dispatch method for virtual power plant,” *Applied Energy*, vol. 300, pp. 117347, 2021.

[63] M. Foti, C. Mavromatis, and M. Vavalis, “Decentralized blockchain-based consensus for Optimal Power Flow solutions,” *Applied Energy*, vol. 283, pp. 116100, 2021.

[64] X. Li, and Q. Li, “Does corporate social responsibility affect the achievement of performance commitment in valuation adjustment mechanism agreement? Evidence from Chinese listed company,” *Managerial and Decision Economics*, vol. 43, no. 2, pp. 533-543, 2022.

[65] LehmanLaw. "Validity of Valuation Adjustment Mechanism ( ”VAM”) Agreement," https://www.lehmanlaw.com/resource-centre/faqs/financial-services-and-regulation/validity-of-valuation-adjustment-mechanism-vam-agreement.html.

[66] S. Lai, J. Qiu, and Y. Tao, “Credit-based pricing and planning strategies for hydrogen and electricity energy storage sharing,” *IEEE Transactions on Sustainable Energy*, vol. 13, no. 1, pp. 67-80, 2021.

[67] C. P. Mediwaththe, M. Shaw, S. Halgamuge, D. B. Smith, and P. Scott, “An incentive-compatible energy trading framework for neighborhood area networks with shared energy storage,” *IEEE Transactions on Sustainable Energy*, vol. 11, no. 1, pp. 467-476, 2019.

[68] W. Zhong, K. Xie, Y. Liu, C. Yang, and S. Xie, “Multi-resource allocation of shared energy storage: A distributed combinatorial auction approach,” *IEEE transactions on smart grid*, vol. 11, no. 5, pp. 4105-4115, 2020.

[69] Y. Tao, J. Qiu, S. Lai, and J. Zhao, “Integrated electricity and hydrogen energy sharing in coupled energy systems,” *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1149-1162, 2020.

[70] J. Yoon, "End-to-End Machine Learning Frameworks for Medicine: Data Imputation, Model Interpretation and Synthetic Data Generation," ProQuest Dissertations Publishing, 2020.

[71] L. Bottou, M. Arjovsky, D. Lopez-Paz, and M. Oquab, "Geometrical Insights for Implicit Generative Modeling," Lecture Notes in Computer Science, Springer International Publishing, 2018, pp. 229-268.

[72] X. Mao, Q. Li, H. Xie, R. Y. Lau, Z. Wang, and S. Paul Smolley, "Least squares generative adversarial networks." pp. 2794-2802.

[73] Y.-P. Hsieh, C. Liu, and V. Cevher, "Finding mixed nash equilibria of generative adversarial networks." pp. 2810-2819.

[74] N. Amjady, and F. Keynia, “A new prediction strategy for price spike forecasting of day-ahead electricity markets,” *Applied soft computing*, vol. 11, no. 6, pp. 4246-4256, 2011.

[75] T. Nelson, S. Kelley, and F. Orton, “A literature review of economic studies on carbon pricing and Australian wholesale electricity markets,” *Energy policy*, vol. 49, pp. 217-224, 2012.

[76] J. Yoon, D. Jarrett, and M. van der Schaar, “Time-series generative adversarial networks,” 2019.

[77] J. Lee, and J.-S. Lee, “TREND: Truncated Generalized Normal Density Estimation of Inception Embeddings for Accurate GAN Evaluation,” arXiv preprint arXiv:2104.14767, 2021.

[78] J. Lin, J. Ma, J. Zhu, and Y. Cui, “Short-term load forecasting based on LSTM networks considering attention mechanism,” *International Journal of Electrical Power & Energy Systems*, vol. 137, pp. 107818, 2022.

[79] J. Moral-Carcedo, and J. Pérez-García, “Temperature effects on firms’ electricity demand: An analysis of sectorial differences in Spain,” *Applied energy*, vol. 142, pp. 407-425, 2015.

[80] H. Shao, X. Deng, and Y. Jiang, “A novel deep learning approach for short-term wind power forecasting based on infinite feature selection and recurrent neural network,” *Journal of renewable and sustainable energy*, vol. 10, no. 4, pp. 43303, 2018.

[81] H. Theil, “The analysis of disturbances in regression analysis,” *Journal of the American Statistical Association*, vol. 60, no. 312, pp. 1067-1079, 1965.

[82] W. Can, X. Zhao, W. Yelei, D. Zhao Yang, and W. Kit Po, “A Hybrid Approach for Probabilistic Forecasting of Electricity Price,” *IEEE transactions on smart grid*, vol. 5, no. 1, pp. 463-470, 2014.

[83] H. D. Kabir, A. Khosravi, S. Nahavandi, and A. Kavousi-Fard, “Partial adversarial training for neural network-based uncertainty quantification,” *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 5, no. 4, pp. 595-606, 2019.

[84] C. Peng, H. Zhou, H. Gu, X. Yi, and Y. Chi, "A Power Market Nodal Price Forecasting Based on Historical Electricity Price and Unified Dispatch Load." pp. 1-8.

[85] S. Yanxia, W. Xu, and C. Jie, “Wind Power Forecasting Using Multi-Objective Evolutionary Algorithms for Wavelet Neural Network-Optimized Prediction Intervals,” *Applied sciences*, vol. 8, no. 2, pp. 185, 2018.

[86] B. Uniejewski, G. Marcjasz, and R. Weron, “Understanding intraday electricity markets: Variable selection and very short-term price forecasting using LASSO,” *International Journal of Forecasting*, vol. 35, no. 4, pp. 1533-1547, 2019.

[87] F. X. Diebold, and R. S. Mariano, “Com paring predictive accu racy,” *Journal of Business and Economic Statistics*, vol. 13, no. 3, pp. 253-263, 1995.

[88] X. Chen, Z. Y. Dong, K. Meng, Y. Xu, K. P. Wong, and H. Ngan, “Electricity price forecasting with extreme learning machine and bootstrapping,” *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2055-2062, 2012.

[89] T. J. Gerpott, and I. Mahmudova, “Determinants of price mark-up tolerance for green electricity–lessons for environmental marketing strategies from a study of residential electricity customers in Germany,” *Business strategy and the environment*, vol. 19, no. 5, pp. 304-318, 2010.

[90] L. Zheng, “Is money smart? A study of mutual fund investors' fund selection ability,” the *Journal of Finance*, vol. 54, no. 3, pp. 901-933, 1999.

[91] V. S. Mutswenje, “A survey of the factors influencing investment decisions: the case of individual investors at the NSE,” University of Nairobi, 2009.

[92] S. Dhar, S. K. Salema, and A. Saha, “Factors affecting individual investor behavior: empirical evidence from mutual fund investors in Dhaka city,” *Management*, vol. 31, no. 3&4, 2017.

[93] D. I. Maditinos, Ž. Šević, and N. G. Theriou, “Investors' behaviour in the Athens Stock Exchange (ASE),” *Studies in economics and Finance*, vol. 24, no. 1, pp. 32-50, 2007.

[94] T. Lal, “Investigating the Factors Affecting Investment Behaviour of Mutual Fund Investors.”

[95] M. Tamiz, R. A. Azmi, and D. F. Jones, “On selecting portfolio of international mutual funds using goal programming with extended factors,” *European Journal of Operational Research*, vol. 226, no. 3, pp. 560-576, 2013.

[96] H. Amiril, and A. M. Gil Lafuente, "Studying of the factors affecting on the mutual fund by individual investor in Iran, Malaysia, Turkey and US," *Modern Applied Science*, 2016, vol. 10, num. 9, p. 192-204, 2016.

[97] D. Xiao, H. Chen, C. Wei, and X. Bai, "Statistical Measure for Risk-Seeking Stochastic Wind Power Offering Strategies in Electricity Markets," *Journal of Modern Power Systems and Clean Energy*, 2021.

[98] Z. Zhu, K. W. Chan, S. Xia, and S. Bu, "Optimal Bi-Level Bidding and Dispatching Strategy between Active Distribution Network and Virtual Alliances Using Distributed Robust Multi-Agent Deep Reinforcement Learning," *IEEE Transactions on Smart Grid*, 2022.

[99] Y. Zhang, Z. Mou, F. Gao, J. Jiang, R. Ding, and Z. Han, "UAV-enabled secure communications by multi-agent deep reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11599-11611, 2020.

[100] H. Liu, C. Zhang, Q. Chai, K. Meng, Q. Guo, and Z. Y. Dong, "Robust regional coordination of inverter-based Volt/Var control via multi-agent deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 12, no. 6, pp. 5420-5433, 2021.

[101] X. Lu, J. Qiu, G. Lei, and J. Zhu, "Scenarios modelling for forecasting day-ahead electricity prices: Case studies in Australia," *Applied Energy*, vol. 308, pp. 118296, 2022.

[102] AusGrid. "Solar home electricity data," https://www.ausgrid.com.au/Industry/Our-Research /Data-to-share/Solar-home-electricity-data.

[103] AEMO. "Aggregated price and demand data," https://aemo.com.au/en/energy-systems/electricity/national-electricity-market-nem/data-nem/aggregated-data.

[104] "VPP Offers Available, 2020, Australian Energy Market Commission (AEMC) ".

[105] AEMO. "AEMO suspends NEM Wholesale Market," https://aemo.com.au/newsroom/media-release/aemo-suspends-nem-wholesale-market.

[106] AEMO. "Guide to Market Suspension in the NEM," https://aemo.com.au/en/energy-systems/electricity/emergency-management/guide-to-market-suspension-in-the-nem.

[107] R. Tang, S. Wang, and H. Li, "Game theory based interactive demand side management responding to dynamic pricing in price-based demand response of smart grids," *Applied Energy*, vol. 250, pp. 118-130, 2019.

[108] P. K. Adom, "The long-run price sensitivity dynamics of industrial and residential electricity demand: the impact of deregulating electricity prices," *Energy Economics*, vol. 62, pp. 43-60, 2017.

[109] T. Zhang, W. Qiu, Z. Zhang, Z. Lin, Y. Ding, Y. Wang, L. Wang, and L. Yang, "Optimal bidding strategy and profit allocation method for shared energy storage-assisted VPP in joint energy and regulation markets," *Applied Energy*, vol. 329, pp. 120158, 2023.

[110] J. Li, J. Geng, and T. Yu, "Grid-area coordinated load frequency control strategy using large-scale multi-agent deep reinforcement learning," *Energy Reports*, vol. 8, pp. 255-274, 2022.

[111] J. Li, T. Yu, X. Zhang, F. Li, D. Lin, and H. Zhu, "Efficient experience replay based deep deterministic policy gradient for AGC dispatch in integrated energy system," *Applied Energy*, vol. 285, pp. 116386, 2021.

[112] Y. Tan, Y. Shen, X. Yu, and X. Lu, "Low-carbon economic dispatch of the combined heat and power-virtual power plants: A improved deep reinforcement learning-based approach,"

*IET Renewable Power Generation*, 2022.

[113] X. Lu, J. Qiu, G. Lei, and J. Zhu, “An Interval Prediction Method for Day-Ahead Electricity Price in Wholesale Market Considering Weather Factors,” *IEEE Transactions on Power Systems*, 2023.

[114] T. Zhong, Z. Zhang, M. Chen, K. Zhang, Z. Zhou, R. Zhu, Y. Wang, G. Lü, and J. Yan, “A city-scale estimation of rooftop solar photovoltaic potential based on deep learning,” *Applied Energy*, vol. 298, pp. 117132, 2021.

[115] ACCC, Inquiry into the National Electricity Market, https://www.accc.gov.au/system/files/May%202022%20report%20-%20Addendum.pdf, 2022.

[116] K. K. Zander, G. Simpson, S. Mathew, R. Nepal, and S. T. Garnett, “Preferences for and potential impacts of financial incentives to install residential rooftop solar photovoltaic systems in Australia,” *Journal of Cleaner Production*, vol. 230, pp. 328-338, 2019.

[117] M. A. Ortega-Vazquez, “Optimal scheduling of electric vehicle charging and vehicle-to-grid services at household level including battery degradation and price uncertainty,” *IET Generation, Transmission & Distribution*, vol. 8, no. 6, pp. 1007-1016, 2014.

[118] X. Li, C. Li, X. Liu, G. Chen, and Z. Y. Dong, “Two-stage community energy trading under end-edge-cloud orchestration,” *IEEE Internet of Things Journal*, 2022.