Deep Learning Based Forecasting Studies for Real VPP Applications

SENYAO WANG

Master of Philosophy

THE UNIVERSITY OF SYDNEY

Supervisor: Jin Ma

A thesis submitted in fulfilment of
the requirements for the degree of
Master of Philosophy

School of Electrical and Computer Engineering
Faculty of Engineering
The University of Sydney
Australia

15 February 2024

Statement of Originality

I hereby affirm that this thesis is exclusively my own work and a product of my own original research. To the best of my knowledge and belief, it does not contain any material previously published or written by another individual, except where appropriate acknowledgment has been made in the text.

The research and work presented in this thesis were carried out by myself, commencing from the initiation of my higher degree by research program. It does not encompass a significant portion of any work that has been submitted for the awarding of any other degree or diploma in any university or other tertiary institution.

I further certify that all information sources and assistance received during the research and writing of this thesis have been acknowledged.

Name: Senyao Wang

Date: 15/2/2024

Acknowledgements

First and foremost, I extend my deepest gratitude to my supervisor, Professor Jin Ma, for his unwavering support and mentorship throughout my MPhil journey at the University of Sydney. Our weekly meetings have been a cornerstone of my academic development, providing a platform for intellectual exchange and critical thinking. Professor Ma has a unique teaching philosophy that I have not encountered in my previous educational experiences. Instead of simply providing answers, he has always guided me to think critically, teaching me not just how to solve problems, but more importantly, how to approach them. This "teaching a man to fish" approach has been incredibly enriching and has fundamentally changed the way I engage with academic challenges. Professor Ma's rigorous academic approach has had a profound impact on me. His emphasis on scientific rigor and precision has not only shaped my research methodology but also deepened my understanding of the subject matter. His invaluable guidance extends beyond academic expertise to include life advice and emotional support, contributing significantly to both my professional development and personal growth.

I would also like to express heartfelt thanks to my co-supervisor, Professor Wei Li. His constructive feedback, expert advice, and continuous encouragement have been instrumental in shaping my research and overcoming challenges. His support has been a cornerstone in my academic journey, and I am truly honored to have had the opportunity to work under his guidance.

The growth I have experienced over the past year from knowing little to achieving so much is a testament to the incredible support I have received from both of my supervisors. Their faith in me has been a constant source of motivation and has empowered me to strive for excellence.

Certainly, I must extend special thanks to my parents for their unconditional love and support. Their financial backing, including covering my tuition fees, has been a pillar that allowed me to focus on my studies and personal growth. I am eternally grateful for their sacrifices and the opportunities they have provided me. I would like to give a special mention to my mother, who even took the time to visit me in Sydney. Her presence and the home-cooked meals she prepared were not just a comfort but a reminder of the strong support system I have back home. Her visit alleviated the loneliness and stress that often accompany academic pursuits abroad, and for that, I am incredibly grateful. Their unwavering faith in me has been a constant source of strength and motivation, and I dedicate this achievement to them.

I would also like to express my heartfelt gratitude to my senior and junior lab mates, as well as my fellow students in Room 329. The harmonious relationships we've built and the academic atmosphere we've fostered have been nothing short of inspiring. From early mornings to late nights, the collective diligence and commitment to learning have been a driving force behind my own academic pursuits. The camaraderie and intellectual stimulation provided by this group have significantly enriched my time at the University of Sydney, and for that, I am deeply thankful.

Lastly, I would like to thank my peers and everyone else who has been a part of this journey. Your camaraderie and support have made this experience both enjoyable and rewarding.

Published Articles and Conferences

- Wang, Senyao and Ma, Jin (2023). 'A Novel GBDT-BiLSTM Hybrid Model on Improving Day-ahead Photovoltaic Prediction'. In: Scientific Reports.
- Wang, Senyao and Ma, Jin (2023). 'A Novel Ensemble Model for Load Forecasting: Integrating Random Forest, XGBoost, and Seasonal Naive Methods'. Presented at: ACFPE 2023 Conference.

Authorship Attribution Statement

I hereby certify that the work included in this thesis contains scholarly contributions for which I am an author. The written statement is included as part of the thesis, endorsed by my supervisor, attesting to my contribution to the scholarly work.

Chapter 3 of this thesis is published as Wang, Senyao, and Jin Ma. "A novel GBDT-BiLSTM hybrid model on improving day-ahead photovoltaic prediction." Scientific Reports 13.1 (2023): 15113. For this publication, I designed the study, analyzed the data, and wrote the drafts.

Chapter 4 of this thesis is published as Wang, Senyao and Ma, Jin (2023). "A Novel Ensemble Model for Load Forecasting: Integrating Random Forest, XGBoost, and Seasonal Naive Methods." For this publication, 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.

Name: Senyao Wang

Date: 15/2/2024

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

Name: Jin Ma

Date: 15/2/2024

Abstract

As renewable energy gains prevalence and machine learning technology advances, the need for accurate forecasting in Virtual Power Plants (VPPs) has become increasingly urgent, especially given the enhanced computational capabilities that make high-accuracy deep learning algorithms feasible, underscoring its importance for grid stability and efficiency. Considering these developments, this dissertation focuses on three key forecasting areas essential for VPP operations: solar photovoltaic (PV) output, load demand, and electricity prices. This thesis employ various deep learning algorithms to make these forecasts, each tailored to the specific characteristics of the data and the operational needs. For both solar PV and load demand, this thesis provide forecasts at two different time intervals: 1-hour and 5-minute. The 1-hour solar PV forecasts aim to provide a 24-hour outlook for day-ahead markets. The 5-minute forecasts are further divided into two categories: immediate next 5 minutes and up to 4 hours into the future. The rationale for these different forecasting horizons is twofold: The 5-minute forecasts are designed for real-time grid management, while the 4-hour forecasts align with the typical charge and discharge cycles of battery storage systems, thereby facilitating more effective battery scheduling. Electricity price forecasts are also made at 5-minute intervals for both immediate and medium-term scenarios. This research has been applied in a production simulation for a retail electricity company, demonstrating its practical utility. The integrated forecasting and optimization model developed in this study offers tangible benefits, enabling consumers to save on electricity costs while increasing profitability for the company.

Contents

Statement of Originality ii

Acknowledgements iii

Published Articles and Conferences v

Authorship Attribution Statement vi

Authorship Attribution Statement vi

Abstract vii

Contents viii

List of Figures xii

Chapter 1 Introduction 1

1.1 Renewable Energy Development Background 1

1.2 Background and Challenges of VPPs 3

1.3 Thesis Outline 7

Chapter 2 Literature review 12

2.1 Virtual Power Plants 12

2.1.1 Structure and Functionalities of VPPs 12

2.1.2 Overview of VPPs 14

2.2 Forecasting Methods in VPPs 15

2.3 Application Scenarios in VPPs 19

2.4 Research Contribution and Technical Scope 21

Chapter 3 PV Forecast 24

3.1 Introduction to PV Forecasting in VPPs 24

3.2 Methods 25
3.2.1 GBDT 25
3.2.1.1 Gradient Boosting 25
3.2.1.2 Decision Tree 25
3.2.1.3 GBDT 26
3.2.2 LSTM 28
3.2.2.1 BiLSTM 29
3.2.2.2 Adam Algorithm in BiLSTM 30
3.2.3 GBDT-LSTM Model With Teacher Forcing 31
3.3 Performance Metrics 34
3.4 PV Forecast Forecast Result 35
3.4.1 One Hour Interval Forecast Result 36
3.4.1.1 Data Preprocessing 37
3.4.1.2 GBDT Modeling Results 38
3.4.1.3 GBDT-BiLSTM Modeling Results 40
3.4.1.4 Model Comparison 42
3.4.2 Five Minutes Interval Forecast Result 45
3.4.2.1 Data Preprocessing and LSTM Model Definition 46
3.4.2.2 5-Minute for Next 5 Minutes Result 47
3.4.2.3 5-Minute for Next 30 Minutes Result 49
3.5 Chapter Summary 50

Chapter 4 Load Forecast **52**
4.1 Introduction to Load Demand Forecasting in VPPs 52
4.2 Methods 53
4.2.1 Transformer Algorithm 53
4.2.1.1 Encoder 54
4.2.1.2 Input Embedding 55
4.2.1.3 Multi-Head Self-Attention Mechanism 55
4.2.1.4 Feed-Forward Neural Networks 56
4.2.1.5 Residual Connection (Add & Norm) 56

4.2.1.6 Softmax in Multi-Head Attention ... 56
4.2.1.7 Linear Transformation Layers (LinearFC and LinearFC2) ... 57
4.2.1.8 Positional Encoding ... 57
4.2.2 Random Forest Algorithm ... 58
4.2.3 XGBoost Algorithm ... 59
4.2.4 Seasonal Naive Algorithm ... 60
4.2.5 Performance Metrics ... 61
4.3 Load Forecast Result ... 61
4.3.1 One Hour Interval Forecast ... 61
4.3.1.1 Data Preprocessing ... 63
4.3.1.2 Hybrid Model Forecast Result ... 64
4.3.2 Five Minutes Interval Forecast ... 66
4.3.2.1 Data Preprocessing and Transformer Model Definition ... 67
4.3.2.2 5-Minute for Next 5 Minutes Result ... 68
4.3.2.3 5-Minute for Next 4 Hours Result ... 71
4.4 Chapter Summary ... 73

Chapter 5 Electricity Price Forecast **74**
5.1 Introduction to Electricity Price Forecasting in VPPs ... 74
5.2 Methods ... 75
5.2.1 Data Preprocessing ... 75
5.2.2 Performance Metrics ... 76
5.3 Price Forecast Forecast Result ... 76
5.3.1 5-Minute to Next 5 Minutes Result ... 77
5.3.2 5-Minute for Next 4 Hours Result ... 79
5.4 Model Comparison ... 80
5.5 Chapter Summary ... 82

Chapter 6 VPP Integration Program Development **84**
6.1 VPP Optimization Model ... 84
6.1.1 Motivation Developing the VPP Program ... 85

6.1.2 VPP Model Function 86
6.1.3 Model Constraints 87
6.1.4 Optimization Methods 88
6.2 Forecast Program Integration 89
6.2.1 VPP-Prediction Models Interface Design 92
6.2.2 Algorithmic Steps and Equations 95
6.3 Chapter Summary 97

Chapter 7 Conclusion **98**
7.1 Major Contributions 98
7.2 Model Limitation 101
7.3 Future Outlook 102

Bibliography **104**

List of Figures

1.1	VPP application scenarios	9
1.2	Research content road map	10
2.1	Technical road map	23
3.1	A principle schematic diagram of Gradient Boosting	25
3.2	LSTM logic diagram	28
3.3	GBDT-BiLSTM logic diagram	34
3.4	The comparison between power generation and the PV irradiation	37
3.5	Three-dimensional data map of the total radiant power generation from June 2020 to December 2021	38
3.6	Last 300 hours of the GBDT model test set	39
3.7	Last 300 hours of the BiLSTM model training set	40
3.8	MSE curves of the model training set and the test set	41
3.9	Last 300 hours of the BiLSTM model test set	42
3.10	GBDT-BiLSTM predicts power generation on December 30	43
3.11	Last 300-hour forecast curves of multiple model test sets	44
3.12	Multiple models predict power generation on December 30	45
3.13	LSTM train set MSE curve	47
3.14	PV generation of 300 hours training set	48
3.15	PV generation of 300 hours test set	48
3.16	PV generation of a single day	49
3.17	Composite PV generation of 6 step forecast	50
4.1	Transformer logic diagram	54
4.2	XGBoost logic figure	59
4.3	Technology roadmap of the hybrid model	62
4.4	The result of aggregate feature importance analysis	63

4.5 Overall comparison of predicted and actual load values 65
4.6 Comparing true and predicted values on a certain day 66
4.7 Transformer train set MSE curve with no positional encoding 68
4.8 Transformer model's 300-hour training set fitting curve with no positional encoding 69
4.9 Transformer train set MSE curve 69
4.10 Transformer model's 300-hour training set fitting curve 70
4.11 Original transformer model's 300-hour training set fitting curve 70
4.12 Original transformer model's 300-hour test set fitting curve 71
4.13 Original transformer model's single day test set fitting curve 72
4.14 Original transformer model's 4 hours test set fitting curve 72
5.1 LSTM price forecast train set MSE curve 77
5.2 Comparison of actual and predicted electricity prices of 300-step training set 78
5.3 Comparison of actual and predicted electricity prices of 300-step test set 78
5.4 Comparison of actual and predicted electricity prices of one day 79
5.5 Composite electricity price forecast of multiple steps within 4 hours 80
6.1 VPP double-layer three-prediction parallel logic figure 91
6.2 PV, load and Price training layer historical data simulate rolling design figure 93
6.3 PV, load and Price training layers simulate rolling training design figure 93
6.4 PV, Load and Price optimization layer interface design figure 94
6.5 Data anti-collision module 1 design logic figure 95
6.6 Data anti-collision module 2 design logic figure 95

CHAPTER 1

Introduction

1.1 Renewable Energy Development Background

In the context of global climate change and environmental degradation, renewable energy, particularly solar energy, has emerged as a key solution. According to the report "Net Zero by 2050: A Roadmap for the Global Energy Sector" from the International Energy Agency (IEA) [1], the global total electricity capacity was 7,795 GW in 2020 and is projected to reach 14,933 GW by 2030, and further increase to 33,415 GW by 2050. During this period, solar photovoltaic (PV) capacity will grow from 737 GW in 2020 to 4,956 GW by 2030, and then to 14,458 GW by 2050. These figures collectively indicate that renewable energy, particularly solar PV, will play a critical role in the global energy transition, with significant growth expected between 2020 and 2030. Detailed data is shown in Tables 1.1 and 1.2. Complementing this shift toward renewables, Virtual Power Plants (VPPs) have emerged as a cutting-edge solution for managing this new energy landscape. VPPs not only aggregate distributed energy resources, including solar PV, but also offer the flexibility and dispatchability needed to balance the fluctuations and uncertainties inherently associated with renewable sources [2].

Among the challenges posed by the rapid advancement of renewable energy, the “California Duck Curve [3]” is a phenomenon that warrants attention. This curve vividly illustrates the rapid intraday changes in net load in the electrical grid due to the large-scale integration of solar energy. Specifically, as the sun sets, there is a need for the rapid dispatch of traditional energy sources to meet demand, leading to significant price volatility and increased risks to grid stability [4]. Moreover, it’s not only the evening surge that presents challenges. During

TABLE 1.1: Global Electrical Capacity (GW) From 2019 2050.

	2019	2020	2030	2040	2050
Total generation	**7484**	**7795**	**14933**	**26384**	**33415**
Renewables	2070	2994	10293	20732	26568
Solar PV	603	717	4956	10980	14458
Hydro	1306	1327	1804	2282	2599
Nuclear	415	415	515	730	812
Wind	623	737	3101	6525	8265
Unabated fossil fuels	**4351**	**4368**	**320**	**1151**	**677**

TABLE 1.2: Global Electricity Generation (TWh) From 2019 2050.

	2019	2020	2030	2040	2050
Total generation	**26922**	**26778**	**37316**	**56553**	**71164**
Renewables	7153	7660	22817	47521	62333
Solar PV	665	821	6970	17031	23469
Hydro	4294	4418	5870	7445	8461
Nuclear	2792	2698	3777	4855	5497
Wind	1423	1592	8008	18787	24785
Unabated fossil fuels	**16941**	**16382**	**9358**	**632**	**259**

the day, the abundance of solar energy could potentially be stored in battery systems for later use. When the sun sets and the evening peak load arrives, that is when the real test for battery dispatch occurs. Efficiently managing these stored reserves can make a significant difference in balancing the grid and stabilizing electricity prices, especially during these high-demand periods [5].

Building on the challenges and opportunities presented by the rapid growth of renewable energy, it's clear that traditional grid management strategies are increasingly inadequate for handling the complexities introduced by renewables. The intermittent nature of solar and wind energy, for instance, requires more than just additional storage solutions; It calls for intelligent systems capable of making real-time decisions based on a multitude of factors. These systems must be able to adapt to fluctuating supply and demand conditions, integrate diverse energy sources, and execute rapid dispatch of stored energy, all while maintaining grid stability and economic efficiency. This is a tall order that traditional grid systems are ill-equipped to handle, thereby highlighting the urgent need for advanced solutions like Virtual Power Plants. These integrated systems are designed to meet these challenges head-on, offering a more flexible and responsive approach to grid management in the era of renewable energy.

In summary, the surge in renewable energy development marks a pivotal moment in the global transition toward a more sustainable and resilient energy landscape. As outlined by various international reports and projections, the capacity for renewable energy sources like solar PV is expected to grow exponentially in the coming decades. This growth is not merely a technological advancement but a necessity in the face of escalating climate challenges and environmental degradation. The rapid integration of renewables into the energy mix brings its own set of challenges, including grid stability and energy storage, which necessitates innovative solutions. It is within this context of burgeoning renewable energy development that Virtual Power Plants emerge as a critical component for the successful orchestration of diverse energy resources.

1.2 Background and Challenges of VPPs

Virtual Power Plants are a modern solution for the management of distributed energy resources within an electrical grid [6]. Unlike traditional power plants that generate electricity at a centralized location, VPPs aggregate a variety of distributed energy sources such as solar panels, wind turbines, and battery storage systems into a single, coordinated network. Through advanced software and communication technologies, VPPs can control these disparate resources as if they were a single, unified power plant. This enables VPPs to provide a range of services, from balancing supply and demand to offering ancillary grid support functions. The primary advantage of VPPs is their ability to enhance grid reliability and efficiency, especially as renewable energy sources become increasingly integrated into the energy landscape.

In the past, when renewable energy sources were less prevalent, the impact of solar PV forecasting, electricity price forecasting, and load forecasting on the system was relatively minor[7], [8]. However, as the share of renewable energy in the grid has increased, the unpredictability and volatility associated with these sources have become more pronounced.

With advancements in technology, we now have the capability to make more accurate forecasts to mitigate the impact of these uncertainties on the grid [9]. In this evolving context, VPPs significantly benefit from these advanced forecasting techniques. Accurate forecasting of solar PV output, load demand, and electricity prices enables VPPs to make real-time decisions, optimizing resource allocation and market participation. This further mitigates the challenges brought about by the increased share of renewables, as VPPs act as a flexible layer that can quickly adapt to changing conditions [10].

To effectively manage the increasing complexities and uncertainties in modern electrical grids, accurate forecasting of solar PV output, load demand, and electricity prices is of paramount importance. Accurate solar PV forecasting allows for more effective power grid scheduling, thereby reducing operational costs associated with uncertainty and volatility. Precise load forecasting enables more efficient resource allocation, which minimizes waste and enhances system reliability[11]. Electricity price forecasting stabilizes the market and provides more reasonable pricing for consumers, thereby encouraging the adoption of renewable energy [12]. These forecasts are not just standalone metrics; they are integral components that enable optimized scheduling of battery storage systems. With the aid of machine learning and deep learning technologies, these forecasts can be made with higher precision, thereby significantly enhancing the efficiency and reliability of battery storage systems. By integrating these accurate forecasts, battery storage systems can execute charging and discharging operations more precisely, effectively smoothing out the net load curve, reducing price volatility, and ultimately enhancing the overall stability and efficiency of the grid. This optimized scheduling not only saves money for consumers but also increases profitability for electricity retailers and improves the overall safety of the electrical grid [13].

In this context, the successful operation of VPPs is critically dependent on accurate forecasting of solar PV output, load demand, and electricity prices. This research aims to equip VPPs with more precise forecasting tools by leveraging machine learning and deep learning technologies. By integrating these high-accuracy forecasts, VPPs can more effectively schedule their internal and external resources, further smoothing out the net load curve, reducing price volatility, and enhancing the overall stability and efficiency of the grid [2]. This not only lowers the electricity bills for consumers but also opens up greater profit margins for electricity retailers, while improving the overall safety of the electrical system.

VPPs serve as a pivotal component in modern electrical systems, optimizing electricity generation and distribution by integrating Distributed Energy Resources (DERs)—such as wind, solar, energy storage systems, and demand response measures [14]. Within the context of a smart grid, VPPs have the capability to manage and dispatch multiple types of energy resources in real-time, thereby providing more reliable and economically efficient electricity services. According to the latest available data, there are hundreds of VPP projects at various stages of development and operation worldwide. The growing prevalence of Virtual Power Plants can be attributed in part to their effectiveness in stabilizing the electrical grid, a role that becomes increasingly crucial with the rising integration of renewable energy sources. Their relevance is further highlighted by an increasing number of academic research studies, pilot projects, and governmental initiatives aimed at improving grid stability and reducing carbon emissions.

Not only do VPPs play a crucial role in maintaining a balance between electricity supply and demand, effectively reducing the necessity for traditional peak-load power plants, but they also serve as a linchpin for the integration of a wide range of renewable energies like wind and solar through advanced real-time analytics and control systems. These analytics go beyond simple data collection; they utilize machine learning algorithms and predictive modeling to facilitate more intelligent decision-making processes. By optimizing electricity generation, distribution, and even storage, VPPs can significantly lower operational costs for utility companies, which in turn could lead to substantial savings for consumers [15]. This economic efficiency is not just a theoretical advantage; there are cases around the world where VPPs have proven to deliver cost benefits.

Moreover, VPPs are becoming increasingly vital in the larger context of climate goals and sustainability [16]. They contribute to reducing greenhouse gas emissions by optimizing the use of renewable energy sources, and in doing so, help countries meet their environmental objectives. As the energy landscape undergoes rapid transformation with a rising share of renewable sources, the role of VPPs in mitigating the challenges of volatility and uncertainties cannot be overstated. They serve as a flexible layer capable of rapid adaptation to fluctuating supply and demand conditions. This flexibility provides invaluable stability to electrical grids, which is becoming ever more essential as grids evolve to accommodate a diversified energy mix.

In the operational context of Virtual Power Plants, a series of challenges emerge that underscore the need for advanced forecasting techniques. The accuracy of forecasting is paramount, as VPPs rely on a variety of distributed energy resources, many of which are renewable and thus susceptible to environmental fluctuations. Real-time or near-real-time data is essential for the efficient operation of VPPs, demanding not only highly accurate forecasting algorithms but also complex data processing and real-time analytical capabilities. As machine learning and deep learning technologies find increasing application in forecasting models, the complexity of these models escalates, adding computational burden and potentially affecting the interpretability and reliability of the models. A yet-to-be-fully-resolved issue is how to effectively integrate these high-accuracy forecasts into VPP battery optimization scheduling and other resource management strategies for efficient and reliable operation of the electrical system. These challenges impact the operational efficiency and reliability of VPPs across multiple dimensions and call for multidisciplinary research and highly integrated solutions.

Given these challenges, the significance of employing accurate forecasting for optimized scheduling in modern electrical grids, particularly in the context of VPPs, becomes even more critical. VPPs serve as a vital layer that utilizes these forecasts to harmonize electricity supply and demand, thereby enhancing grid stability and efficiency. For consumers, accurate forecasting in solar PV output and electricity demand, facilitated by VPPs, leads to cost savings by allowing for more efficient resource allocation and reducing reliance on expensive peak-load energy. For electricity retailers, the focus of this research and its real-world applications, VPPs armed with accurate forecasts offer twofold benefits: They enable the stabilization of electricity prices and minimize the costs associated with energy imbalances and price volatility. By leveraging machine learning and deep learning technologies for precise solar PV and electricity price forecasting, this study aims to address the aforementioned challenges, providing VPPs with the tools needed for effective scheduling of internal and external resources, including battery storage systems. These systems can then execute charging and discharging operations more precisely, thereby smoothing out the net load curve, reducing price volatility, and enhancing grid stability. All these factors contribute to the overall safety, efficiency, and reliability of the electrical grid, marking a significant step towards a more sustainable and reliable energy future.

1.3 Thesis Outline

Given the multifaceted importance of accurate forecasting in VPPs, as outlined above, this dissertation is structured to provide a comprehensive exploration of the subject matter. It aims to bridge the existing gaps in the literature and offer innovative solutions to the challenges faced by VPPs in the modern energy landscape. The following chapters are designed to delve into each critical aspect of forecasting and optimization in VPPs, providing both theoretical insights and practical applications.

Chapter 2 will serve as a literature review, laying the groundwork for the rest of the study. It will outline the organizational structure and functionalities of VPPs, emphasizing the critical role of accurate forecasting in their operation. This chapter will also provide an overview of the current state of research in this area, pinpointing the existing gaps and limitations that this dissertation aims to address. Based on these identified needs, the chapter will conclude by presenting a technical roadmap for the research to be carried out in the subsequent chapters.

Following the literature review, Chapter 3 will delve into the intricacies of solar PV forecasting. The chapter will begin by elucidating the pivotal role that PV forecasting plays within the operational framework of VPPs. Subsequently, it will introduce the specific forecasting methodologies that will be employed in this research, culminating in an evaluation of the predictive performance of these methods. Chapter 4 will be dedicated to load demand forecasting. The chapter will initiate with a discussion on the importance of accurate load forecasting in the context of VPPs. This will be followed by an exposition of the load forecasting techniques that will be applied in this study, and will conclude with an assessment of their predictive accuracy. Chapter 5 will focus on electricity price forecasting. The chapter will open by underscoring the significance of price forecasting in the optimization of VPP operations. It will then proceed to outline the methodologies that will be used for this specific type of forecasting, wrapping up with an analysis of the expected results. Chapter 6 will concentrate on two main areas: Firstly, introducing the existing VPP battery optimization program, outlining its operational framework and key functionalities; and secondly, detailing the VPP-Prediction Models Interface design. This interface is instrumental in integrating the forecasting methods discussed in previous chapters into the VPP battery optimization model. The chapter will provide a thorough explanation of the steps, mathematical formulas, and operational workflows that enable this integration.

Finally, Chapter 7 will serve as the concluding chapter of this dissertation. It will offer a comprehensive summary of the research, encapsulating the key findings, contributions, and implications of the study. The chapter will also discuss the limitations of the current work and suggest avenues for future research, thereby providing a complete wrap-up of the research journey. The research content road map is shown in Figure 1.2.

As illustrated in Figure 1.1, this research encompasses four specific scenarios, each representing a unique aspect of the Virtual Power Plant (VPP) operational environment.The VPP application scenarios data scope is shown in Table1.3. These scenarios include residential communities with solar panels but without individual batteries (Scenario 1), large commercial or governmental buildings with solar installations (Scenario 2), residential areas with individual solar panels and home batteries (Scenario 3), and communities with shared solar panels and a communal battery system (Scenario 4). This diverse range of scenarios ensures that the forecasting models developed are versatile and applicable across various VPP configurations and scales within a suburban context. In the third chapter, the focus is on solar PV forecasting, utilizing data from Scenarios 1 and 2 to enhance the forecasting model. The fourth chapter delves into load forecasting, where the distinct electricity usage patterns in Scenarios 1 and 2 are analyzed. The fifth chapter addresses electricity price forecasting, which is crucial for effective VPP operation in all scenarios. Each chapter provides an in-depth exploration of the respective forecasting methodologies and their application in the context of these scenarios. Chapter 6 then explores the integration of these forecasts into the VPP battery optimization

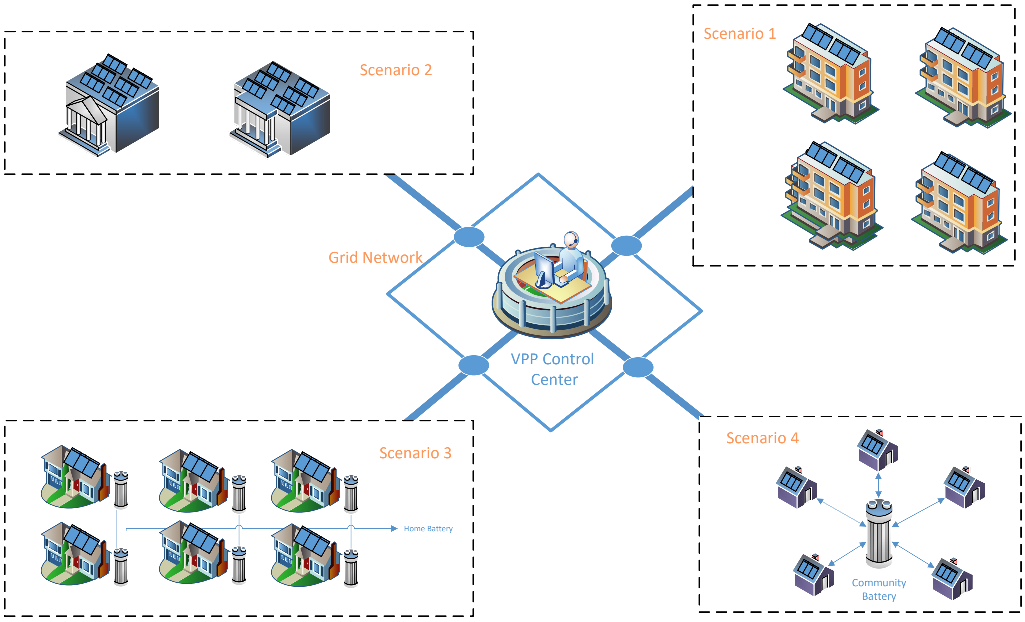

FIGURE 1.1: VPP application scenarios

scheduling. It discusses the relationship between solar PV, load, and electricity price forecasts and their impact on the VPP optimization process.

The outcomes of the PV and Load Forecasting from these scenarios are integrated into the VPP control center for simulation, calculation, and scheduling. This integration is crucial for effective energy management, allowing for real-time adjustments and optimization based on the predictive insights provided by the forecasting models. The adaptability of these models to different scales and complexities of energy systems highlights their potential in enhancing the operational efficiency and sustainability of modern electrical grids.

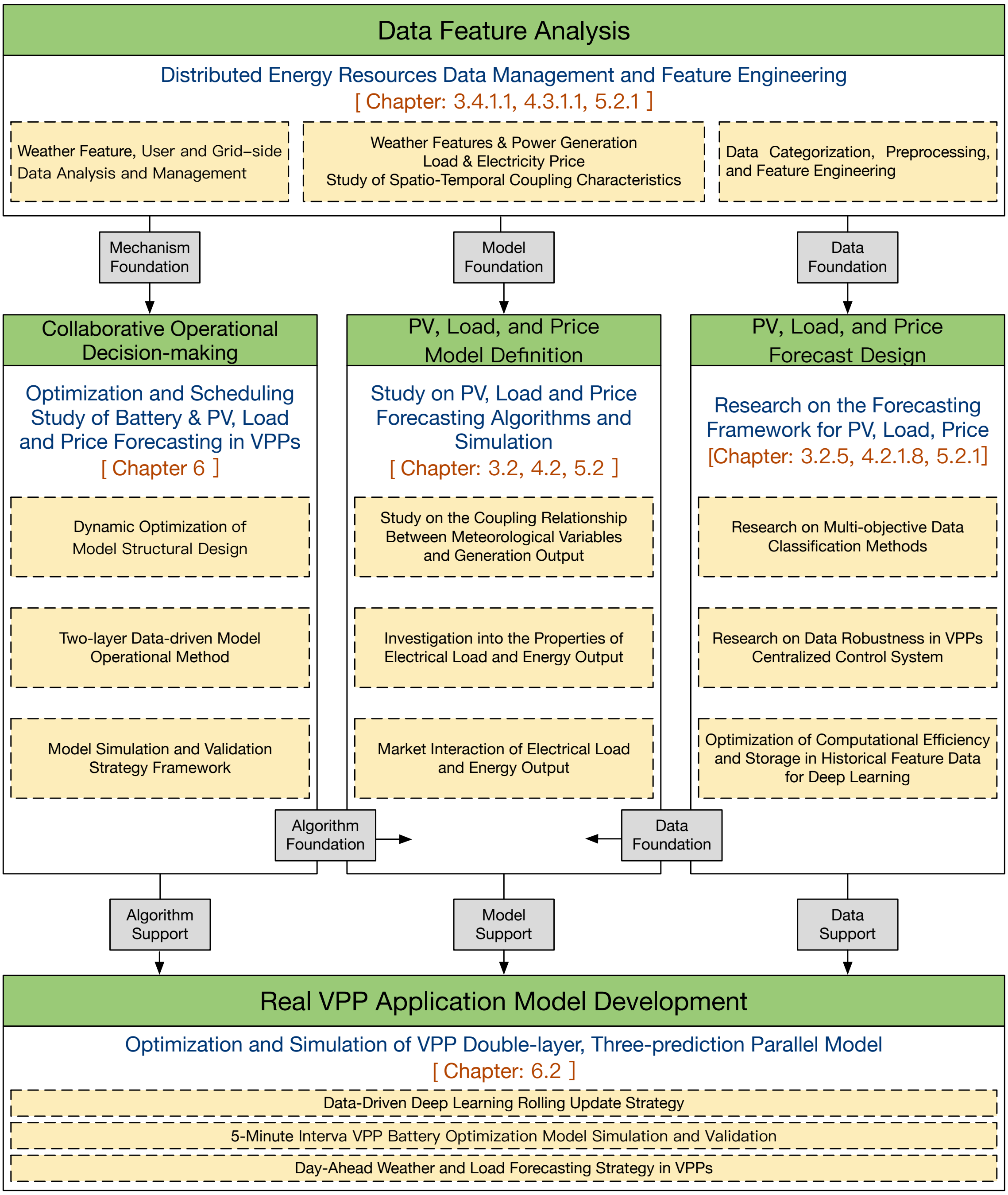

FIGURE 1.2: Research content road map

	Category	PV	Load	Battery	Price (AEMO)
Scenario 1	Residential Building	2022.6 - 2023.6		/	2022.1-2023.6
Scenario 2	Government & Commercial Building	2022.1 - 2022.12		/	2022.1-2023.6
Scenario 3	Community With Individual Battery	2022.1 - 2023.10			2022.1-2023.6
Scenario 4	Community With Shared	2022.1 - 2023.10			2022.1-2023.6

TABLE 1.3: VPP application scenarios data scope

CHAPTER 2

Literature review

Having established the critical role of accurate forecasting in modern electrical grids and its implications for both consumers and electricity retailers, it is essential to delve into the existing body of research that informs these practices. The following Literature Review section aims to provide a comprehensive overview of current studies and methodologies in the field of Virtual Power Plants and forecasting techniques. This will not only contextualize the significance of this research but also highlight the gaps and opportunities that this study aims to address.

2.1 Virtual Power Plants

Before delving into the sophisticated forecasting methods that enhance the operation of VPPs, it is essential to first understand the fundamental structure that underpins these systems. A clear comprehension of how a VPP is architecturally designed and functionally operates serves as a foundational basis for appreciating the intricacies of various forecasting algorithms employed. Let us now turn our attention to the key components that make up a VPP and the specific functions they serve.

2.1.1 Structure and Functionalities of VPPs

VPPs serve as an intricate, multi-faceted system that fundamentally redefines how electricity is generated, distributed, and consumed. Unlike traditional power plants, which hinge on a single large-scale generation source, VPPs bring together a myriad of smaller, distributed energy resources (DERs) into a cohesive and harmonious network. This aggregation offers a more versatile, scalable, and robust approach to energy management. By distributing the burden of electricity production across multiple points, VPPs enhance resilience against system failures and enable more efficient utilization of renewable resources. Through their decentralized yet coordinated nature, they harmonize the competing demands of electricity generation and consumption, bringing a new level of sophistication to grid management. This complex interplay of structure and functionality positions VPPs as a cornerstone of modern, smart grids, marking a significant departure from traditional energy systems towards a more adaptable, sustainable, and resilient infrastructure [17].

In terms of structure, a VPP can be conceptualized into three main tiers.

(1) Distributed Energy Resources (DERs). The foundational layer consists of DERs, which serve as the building blocks of the VPP ecosystem [18]–[20]. This tier comprises a variety of renewable energy sources such as wind turbines and solar panels, along with battery storage systems and even controllable loads like smart HVAC systems. In the context of this research, the primary focus is on the harmonious coordination and optimization between solar PV panels and battery storage systems. By concentrating on these specific types of DERs, the study aims to develop strategies that maximize the efficiency of solar energy production, storage, and consumption, thus furthering the objectives of sustainable and resilient energy management.

(2) Centralized Control System. Acting as the operational nucleus, this layer sits above the Distributed Energy Resources and plays a pivotal role in governing the entire VPP architecture. The centralized control system is equipped with advanced data analytics capabilities, machine learning algorithms, and various forecasting models tailored for both short-term and long-term predictions [21]. In this research, particular attention is given to employing state-of-the-art machine learning and deep learning algorithms for precise solar PV output and electricity price forecasting. Such a complex computational framework allows the system to harmonize the coordination among disparate DERs, like solar panels and battery storage, by dynamically adapting to real-time grid conditions and market variables. It not only schedules the charging and discharging operations of the battery storage systems but also determines the optimal amount of electricity to be generated or consumed based on predictive analytics. By doing so,

the centralized control system aims to achieve the overarching goals of maximizing energy efficiency, reducing operational costs, and bolstering the grid's reliability and resilience.

(3) Market Interaction. Functioning as the external interface of the VPP, this top layer is critical for integrating the system with external electricity markets and grid networks [22]. It serves as the trading hub where both real-time and future electricity bids are placed, facilitating the sale of excess generated power back to the grid or the purchase of additional electricity during times of peak demand. Advanced market algorithms are employed to scrutinize real-time market prices, demand trends, and even weather forecasts to make well-informed trading decisions. This layer can interact with multiple types of marketplaces, including wholesale electricity markets, ancillary service markets, and capacity markets. Assisted by the centralized control system, it formulates optimal bidding strategies and determines energy pricing. Through this intricate operation, the Market Interaction layer not only aims for maximum revenue but also focuses on cost-effectiveness and sustainability, adeptly adapting to market fluctuations while upholding grid stability.

Beyond these structural elements, VPPs serve multiple functionalities. They are designed to contribute to grid stabilization by rapidly adapting to supply and demand fluctuations. Moreover, by pooling various types of resources, VPPs can offer a range of ancillary services, from frequency regulation to voltage control, thereby enhancing grid reliability and resilience.

2.1.2 Overview of VPPs

VPPs have emerged as a pivotal innovation in the energy sector, particularly in the context of the growing integration of distributed energy resources (DERs) like solar panels, wind turbines, and battery storage systems [23]. By aggregating these diverse energy resources, VPPs function as a unified and flexible asset that can be controlled centrally, thereby mimicking the capabilities of a traditional power plant. This aggregation not only enhances grid stability but also facilitates the efficient utilization of renewable energy sources [6].

Research in the realm of VPPs has been extensive, covering various facets from operational optimization to control strategies. The primary focus has been on maximizing energy efficiency, minimizing operational costs, and ensuring grid stability. To achieve these objectives, researchers have employed a range of optimization algorithms. The paper [24] presents a Genetic Algorithm-based optimization approach for VPPs, integrating Demand Side Management and Distributed Energy Resources to reduce overall electricity market costs and alleviate peak usage burdens. The paper [25] uses machine learning to accurately estimate the longevity of aggregated response in a VPP, improving market operation flexibility. The paper [26] introduces a deep reinforcement learning (DRL) algorithm for optimal online economic dispatch in VPPs, reducing computational complexity and effectively handling the stochastic nature of distributed generation. The paper [27] presents a machine learning-based forecasting model using Long-Short-Term-Memory (LSTM) networks for VPPs, focusing on the prediction of wind, solar, and Combined Heat and Power generation outputs. Particle swarm optimization techniques have been applied to fine-tune the energy dispatch strategies, ensuring that the VPP can respond effectively to real-time market signals [28],[29],[30]. Mixed-integer linear programming has also been used for optimizing both the economic and environmental performance of VPPs, taking into account various constraints such as energy demand, grid stability, and emission levels [31], [32].

Moreover, machine learning and deep learning techniques are increasingly being integrated into VPP operations for predictive analytics. These advanced algorithms offer the potential for more accurate demand forecasting, price prediction, and even predictive maintenance, thereby further enhancing the operational efficiency and reliability of VPPs[33].

2.2 Forecasting Methods in VPPs

Accurate forecasting is crucial for the effective operation of VPPs. The literature on forecasting methods can be broadly categorized into traditional statistical methods and machine learning-based methods [34].

Traditional Statistical Methods: In the realm of VPPs, traditional statistical methods like AutoRegressive Integrated Moving Average (ARIMA) and Exponential Smoothing have long been the cornerstone for time-series forecasting. These methods are favored for their ease of implementation, low computational cost, and proven track record in various industries. ARIMA, for instance, is particularly adept at modeling univariate time-series data that exhibit trends or seasonality, making it a popular choice for load and price forecasting [35]. Exponential Smoothing, on the other hand, is known for its ability to capture the level, trend, and seasonality in historical data, offering a robust mechanism for short-term forecasting tasks.

However, these traditional methods come with their own set of limitations. Most notably, they often struggle to capture complex non-linear relationships and high-dimensional interactions in the data [36]. This is a significant drawback, especially in the context of modern VPPs, where multiple factors such as weather conditions, renewable energy generation, and market dynamics can introduce non-linearities and complexities. Additionally, these methods generally require the data to be stationary, a condition that is not always met in real-world scenarios where various dynamic factors can introduce variability. As a result, while traditional statistical methods continue to be useful for certain applications, they are increasingly being complemented or even replaced by more advanced machine learning and deep learning techniques to better capture the intricacies of modern energy systems.

Machine Learning-Based Methods: The advent of big data and increased computational power has significantly influenced the application of machine learning techniques in VPPs. Classical machine learning methods like Random Forests, Support Vector Machines (SVMs), and Genetic Algorithms have been increasingly applied to energy forecasting. The paper [37] proposes an online self-adaptive forecasting method based on random forest to address the challenges of electricity price forecasting in volatile markets, demonstrating higher accuracy than traditional methods. SVM [38] offer flexibility in capturing non-linear relationships, although they often require careful hyperparameter tuning to achieve optimal performance. The paper [39] presents a management system for residential VPPs that optimizes both economic and technical objectives using a genetic algorithm, demonstrating that such targets, usually reserved for larger VPPs, can be effectively managed in smaller setups.

These classical machine learning methods have their own set of advantages and limitations [40]. For instance, Random Forests are computationally intensive but provide high accuracy and feature importance scores, which are valuable for understanding the factors affecting energy forecasts. SVM are robust but may require a fine-tuned kernel for optimal performance. Genetic Algorithms are excellent for optimization but can be slow and computationally expensive. Despite these challenges, the ability of these machine learning methods to capture complex interactions makes them invaluable tools for improving the accuracy and efficiency of energy forecasting in VPPs.

Deep Learning Methods: The advent of deep learning technologies has significantly impacted the field of energy forecasting, particularly in VPPs. Various architectures have been explored to improve forecasting accuracy. The paper [41] employs a deep learning-based B-LSTM network for optimizing a VPP's bidding strategy in day-ahead markets, significantly enhancing accuracy and profitability. The paper [42] addresses challenges in time-series forecasting for VPPs using PV data, specifically data insufficiency and anomalies. Variational Pattern Networks (VPNNs) are another emerging technique, designed to capture complex non-linear relationships in the data. The paper [43] introduces a novel time series forecasting method using Transformer-based machine learning models, leveraging self-attention mechanisms to capture complex patterns. While these methods are computationally intensive and often require large datasets, their ability to capture complex relationships and temporal patterns makes them invaluable tools for high-accuracy forecasting.

Hybrid Methods: Building upon the advancements in both traditional statistical methods and machine learning, hybrid models have emerged as a compelling approach for energy forecasting in VPPs. These models aim to combine the best of both worlds: The robustness and ease of interpretation from statistical methods like ARIMA, and the high accuracy and pattern recognition capabilities from machine learning algorithms such as Random Forests or even deep learning methods like LSTMs. By integrating these diverse techniques, hybrid models are designed to offer superior forecast accuracy. They leverage statistical methods for capturing underlying trends and seasonality, while utilizing machine learning or deep learning algorithms to model complex non-linear relationships and temporal dependencies [44]. This results in a more accurate and computationally efficient forecasting model, making hybrid methods an increasingly popular choice for complex energy systems.

In summary, the literature presents a wide array of forecasting methods, each with its own set of advantages and limitations. The choice of method often depends on the specific requirements of the VPP, such as the forecasting horizon, data availability, and computational resources.

While the existing literature on forecasting methods in VPPs is extensive, several research gaps and limitations remain to be addressed:

1. Limited Accuracy of Simple Models: Traditional statistical methods like ARIMA and Exponential Smoothing, although computationally efficient, often fail to capture the complex non-linear relationships in the data. This results in suboptimal forecasting performance, particularly when dealing with the multifaceted nature of energy systems. 2. Computational Intensity of Complex Models: Advanced machine learning and deep learning methods offer high accuracy but are computationally intensive. This makes them less suitable for real-time battery optimization scheduling in VPPs, where quick decision-making is essential. 3. Lack of Comprehensive Systems and Multi-Faceted Nature of VPPs: The forecasting in VPPs involves simultaneous consideration of solar PV output, load demand, and electricity prices, each crucial for effective operation. However, there is a dearth of comprehensive systems that can be directly applied to VPPs, capable of handling the complexity and multi-dimensionality of modern VPPs, especially in residential settings where each household may have its own battery and solar PV system. 4. Scalability and Adaptability: Many existing models are not easily scalable or adaptable to different VPP configurations or market conditions. This limits their applicability in diverse settings, making it challenging to implement them in VPPs that have varying scales or are subject to different regulatory frameworks.

2.3 Application Scenarios in VPPs

From the perspective of the real-world operational requirements of VPPs, the role of accurate forecasting across different operational stages is paramount. These operational stages include Day-ahead, Real-time, and Frequency Regulation, among others.

In Day-ahead planning, solar panel output prediction is critical for optimizing the operations of VPPs. Accurate forecasts for PV energy production not only facilitate the effective scheduling of various distributed resources but also determine the optimal charge-discharge cycles for battery storage systems, enabling the most cost-efficient utilization of renewable assets like solar panels. The importance of this cannot be overstated, given the inherently intermittent nature of solar power, where accurate forecasting becomes pivotal for ensuring grid stability and reducing operational costs. Two methodologies that are particularly suited for day-ahead PV forecasting in VPPs are Gradient Boosting Decision Trees (GBDT) and Long Short-Term Memory (LSTM) networks. GBDT excels at capturing the non-linear relationships between weather conditions and solar output, offering robustness and model interpretability, while also being well-equipped to handle missing or imbalanced data commonly found in real-world energy scenarios. LSTM networks, conversely, are adept at capturing long-term temporal dependencies, making them ideal for handling time-series data influenced by a variety of time-dependent factors like cloud cover, temperature, and time of year. Both methodologies offer unique advantages, and when appropriately applied, they can significantly enhance the operational efficiency and reliability of VPPs.

Expanding on real-time market operations, VPPs face additional challenges that include not only the immediate generation and consumption of energy but also the integration and optimization of demand-side management and energy storage components. The real-time stage often involves intricate load-balancing tasks and frequency regulation that need to be executed swiftly. In this demanding environment, the agility of the Transformer model becomes a vital asset. Its ability to process sequence data in parallel eliminates the latency often encountered in traditional machine learning models that require sequence dependencies. Furthermore, Transformers are scalable and can be easily updated to incorporate new data streams or respond to changing grid conditions. This adaptability allows VPPs to nimbly navigate the challenges of real-time operations, be it sudden spikes in demand, fluctuating renewable energy outputs, or unforeseen grid disturbances, thereby upholding both grid reliability and operational efficiency.

In the context of Frequency Regulation, maintaining and stabilizing the grid's frequency is of utmost importance. In this highly dynamic setting, VPPs must perform precise second or minute-level forecasting for solar PV outputs and demand-side management while also continually adjusting other dispatchable resources in real-time. Crucial to this is the role of price-based optimization algorithms that govern the battery storage systems. These algorithms not only decide the optimal charging and discharging cycles of the battery storage units but also make real-time decisions on how much energy to either absorb from or inject into the grid based on the prevailing electricity prices. The capacity to do this accurately enables the VPP to rapidly counterbalance frequency deviations, thereby playing a pivotal role in maintaining grid stability. Such an approach adds another layer of complexity and computational demand, but the net effect is an optimized asset pool that can respond proactively to frequency imbalances, increasing both the financial viability and operational reliability of the VPP in real-time market scenarios.

In the current research, special attention is paid to the use of advanced machine learning and deep learning algorithms for the precise forecasting of solar PV panels and battery storage systems, aimed at achieving efficient operation across the aforementioned different segments of VPP functioning. Such multi-layered, multi-temporal scale prediction requirements do not only elevate the complexity of operating a VPP but also open up avenues for further optimization and enhancements in the overall performance of the electrical grid. By integrating sophisticated forecasting techniques into the VPP's operational framework, the research aims to create a more adaptable, resilient, and financially viable system. This effectively transforms the traditional grid infrastructure into a more intelligent and responsive entity, capable of better handling uncertainties, reducing energy wastage, and enhancing the grid's ability to integrate a diverse set of renewable energy resources.

2.4 Research Contribution and Technical Scope

This study introduces an integrated forecasting and optimization model for VPPs, distinctively addressing several limitations and challenges identified in existing research. The VPP in this study operates under a centralized control model, managed by the retailer. This model efficiently coordinates distributed energy resources and customer data, including battery status, load data, and solar PV generation. It optimizes operations from the retailer's perspective within existing regulatory frameworks, a contrast to decentralized models that may not fully harness the potential of centralized data processing and decision-making. Utilizing a variety of deep learning algorithms and ensemble learning methods, the model provides highly accurate forecasts across different time scales and data characteristics. In comparison to existing models that often rely on singular forecasting methods, this model employs a parallel three-way forecasting approach for solar PV, load, and electricity prices. This multifaceted approach not only enhances the system's robustness but also significantly reduces computational time, addressing the common challenge of computational intensity in complex models. Furthermore, the model generates independent weight files, eliminating data conflicts between the upper and lower layers, thereby enabling faster optimization scheduling. This innovative feature distinguishes it from other models that may struggle with data synchronization and real-time processing efficiency. The separation of layers allows for rapid execution of optimization scheduling in the lower layer, while the upper layer can be immediately updated upon the completion of model training, ensuring forecast accuracy. This aspect of the model addresses the scalability and adaptability challenges often encountered in VPP operations. This comprehensive model has been successfully applied in a production simulation for a retail electricity company, demonstrating its practical utility and effectiveness. The application of this model showcases its practical advantages over existing models, particularly in terms of operational efficiency and adaptability to real-world scenarios. Through this highly customized and multi-layered approach, the study offers not only an efficient and practical solution for battery optimization scheduling in VPPs but also enables consumers to save on electricity costs while increasing profitability for the company. This dual benefit of consumer savings and increased profitability for electricity retailers is a notable advancement over existing models, which may not effectively balance these two objectives.

The core focus of this study lies in the development and integration of three distinct forecasting models for solar PV output, load demand, and electricity prices, each employing advanced machine learning and deep learning techniques. This approach marks a significant departure from traditional methods, which often rely on singular or less sophisticated forecasting techniques. The high levels of accuracy and efficiency achieved by these models are a testament to the effectiveness of employing cutting-edge technologies in energy forecasting. While the VPP battery optimization program itself is not the primary subject of this research, the study makes a substantial contribution by seamlessly integrating these forecasting models into an existing VPP battery optimization framework. It addresses a critical gap in current VPP management strategies, which often lack the capability to effectively incorporate advanced forecasting results into operational decision-making.This integration of forecasting models into the VPP framework represents a significant advancement over existing models, which may not fully utilize the potential of advanced forecasting in operational scheduling. By doing so, the study addresses real-world challenges in energy management, particularly in terms of enhancing operational efficiency and decision-making accuracy. The practical implications of this research are far-reaching, offering a model that not only improves the operational efficiency of VPPs but also contributes to the broader goals of sustainable and reliable energy management.

The remainder of this dissertation is structured as follows: Chapter 3 delves into solar PV forecasting, employing GBDT-BiLSTM for day-ahead forecasts and LSTM algorithms for short-term forecasts ranging from 5 minutes to 1 hour. Chapter 4 focuses on load forecasting, utilizing an ensemble model of XGBoost, Random Forest, and Seasonal Naive for 1-hour forecasts, and a Transformer model with positional encoding for short-term forecasts from 5 minutes to 4 hours. Chapter 5 is dedicated to electricity price forecasting, providing insights into the methodologies and algorithms used. Chapter 6 discusses the optimization of VPP battery scheduling, elaborating on the implementation of a dual-layer, three-forecast parallel model for efficient operation. A technical roadmap of this thesis is provided in Figure 2.1.

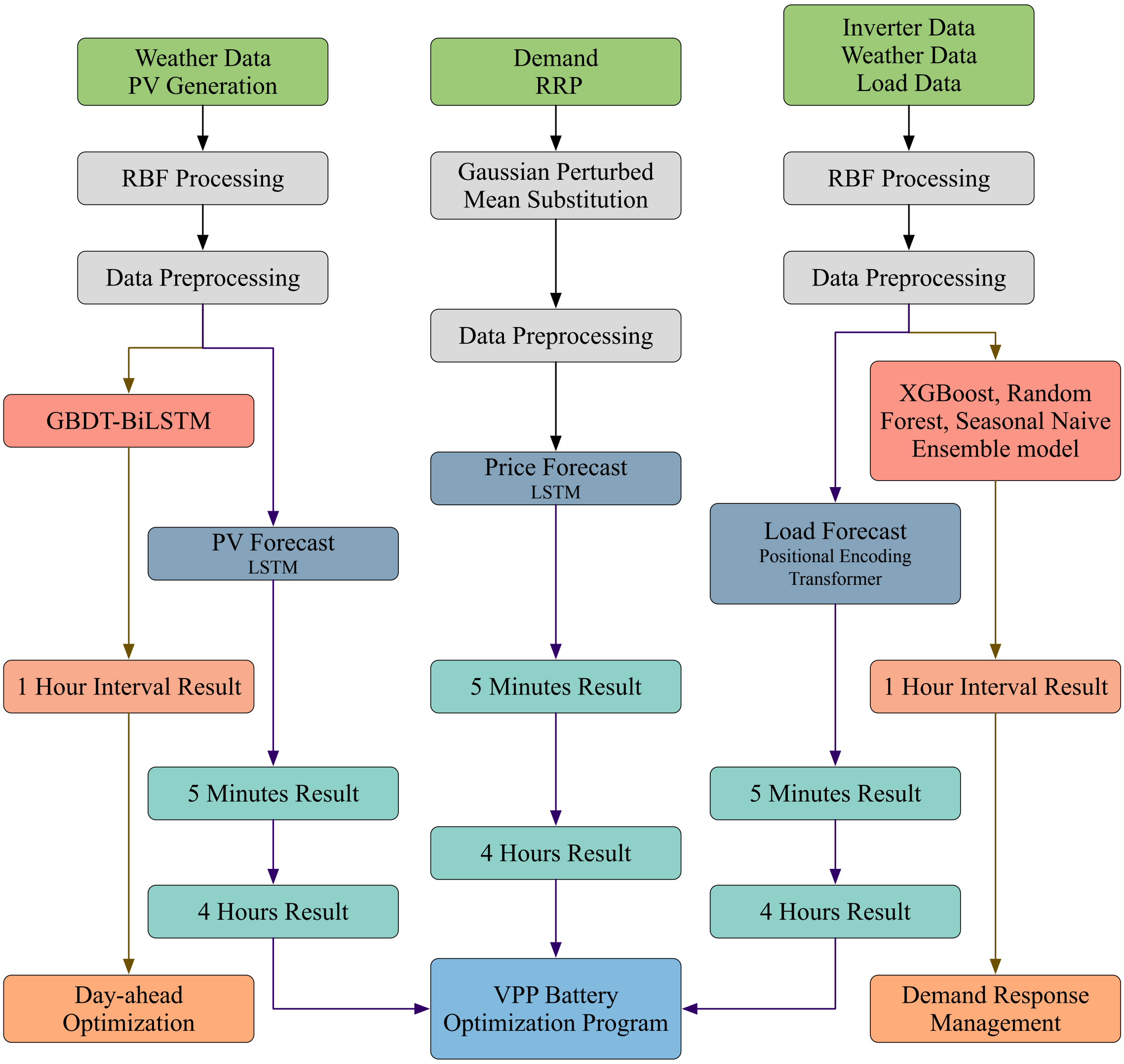

FIGURE 2.1: Technical road map

CHAPTER 3

PV Forecast

3.1 Introduction to PV Forecasting in VPPs

Solar PV forecasting plays a pivotal role in the efficient operation of VPPs. As VPPs aggregate various distributed energy resources, including solar PV installations, accurate forecasting of solar PV output becomes crucial for several reasons.

Firstly, the intermittent nature of solar energy introduces variability into the grid, posing challenges for maintaining grid stability. Secondly, inaccurate forecasts can lead to suboptimal resource allocation, resulting in increased operational costs and reduced energy efficiency. Lastly, precise solar PV forecasts are essential for effective participation in energy markets, where bidding strategies often rely on accurate predictions of available solar power.

Given these challenges, there is a pressing need for robust and accurate solar PV forecasting models tailored for VPP operations. This section delves into the mathematical models and technical approaches employed in this research to address these real-world engineering problems. This thesis focus on both short-term and long-term forecasts, employing a range of machine learning and deep learning techniques to optimize accuracy and computational efficiency.

The content in the section on PV Forecasting is an extension of the authors' previous work published in *Scientific Reports*, under the title "A Novel GBDT-BiLSTM Hybrid Model on Improving Day-ahead Photovoltaic Prediction".

3.2 Methods

3.2.1 GBDT

3.2.1.1 Gradient Boosting

The Gradient Boosting (GB) is often used to build regression models [45]. It approaches the minimum loss value in the gradient direction through repetition and iteration. As Figure 3.1 shows, the lost of original data is minimized in iterations, which produces a loss function for weak learners. After N times iterations, a more accurate and stronger learner can be trained. This algorithm can solve the optimization problem of spatial functions.

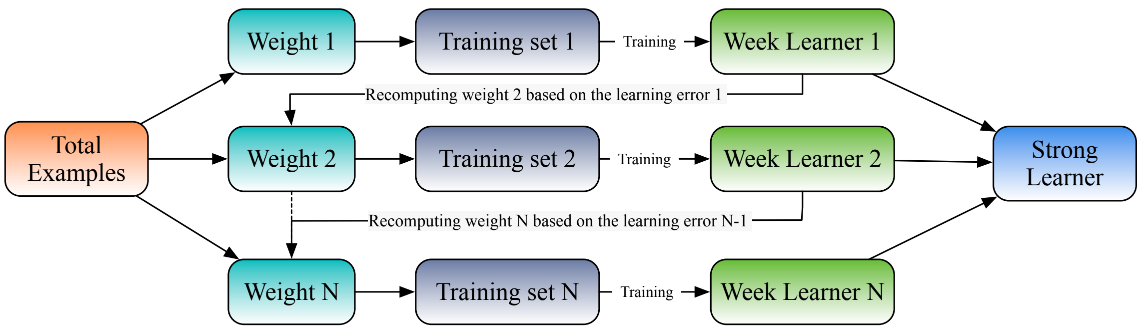

FIGURE 3.1: A principle schematic diagram of Gradient Boosting

3.2.1.2 Decision Tree

Decision tree is a supervised learning model based on the tree structure [46]. It utilizes the concept of entropy in informatics to solve basic classification and regression problems. The tree structure of the model can be adjusted in accordance with the characteristics of the data set[47], making decision tree models flexible in handling various types of data, including continuous and discrete values.

This thesis use the least square regression tree to express the model, which adapts to the impurity function of the regression tree. The standard formula of the decision tree can be shown as:

$$F\left(X,\left\{X_{1},X_{2}\right\}\right)=\sum_{x_{i}\in X_{1}}\left(y_{i}-\hat{y}_{1}\right)^{2}+\sum_{x_{i}\in X_{2}}\left(y_{i}-\hat{y}_{2}\right)^{2} \tag{3.1}$$

In this formula, X represents the set of all data points, while X_1 and X_2 are subsets of X that result from a split in the decision tree. Each x_i is an individual data point within these subsets. The terms y_i represent the actual values associated with each data point, and $\hat{y}_1$ and $\hat{y}_2$ are the predicted values for the subsets X_1 and X_2, respectively. The goal of the model is to find the optimal split that minimizes the sum of squared differences within each subset.

3.2.1.3 GBDT

Based on the above algorithm, the GBDT algorithm can be viewed as a regression model that combines all the weak learners of the model into a strong learner. The specific algorithm is as follows:

Step1. Initialize the weak learner, assuming a loss function L(y, f(x)). The mean value of c can be set as the mean value of sample y.

$$g_{0}(x)=\arg\min_{c}\sum_{i=1}^{n}L\left(y_{i},c\right) \tag{3.2}$$

Step2. Make M times iterations. First estimate the residual.

$$r_{im}=-\left[\frac{\partial L\left(y_{i},f\left(x_{i}\right)\right)}{\partial f\left(x_{i}\right)}\right]_{f(x)=f_{m-1}(x)} \qquad i=1,\ldots,n \tag{3.3}$$

Generate a regression tree from $\left\{\left(x_{i},r_{im}\right),i=1\ldots n\right\}$, and the node area of the number m tree is $R_{mj}(j=1,2,\ldots,J)$. Then find the optimal $\mathcal{C}_{mj}$ for j=1,2,...,n, and minimize the loss function. The formula can be shown as:

$$c_{mj}=\arg\min_{c}\sum x_{i}\in R_{mj}L\left(y_{i},f_{m-1}\left(x_{i}\right)+c\right) \tag{3.4}$$

$$h_m(x_i) = \sum_{j=1}^{|R_m|} c_{mj} I(x_i \in R_{mj}) \tag{3.5}$$

Step3. Update the model:

$$g_m(x) = g_{m-1}(x) + \sum_{j=1}^{J} c_{mj} I(x_i \in R_{mj}) \tag{3.6}$$

Step4. Until the completion of N iterations, we can use an additive model to represent the prediction function G(x) with higher accuracy:

$$G(x) = \sum_{i=1}^{M} \sum_{j=1}^{J} c_{mj} I(x_i \in R_{mj}) \tag{3.7}$$

This thesis still assume that the training set has n samples, the specific model is as follows:

Algorithm 1 The GBDT model

Ensure: The prediction function after obtaining m regression, r_{im}.

1: **Initialize the weak learner:**
2: $F_0(x) = \arg\min_c \sum_{i=1}^{n} L(y_i, c)$
3: **Calculate learning rate:**
4: $r_{im} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)} \quad i = 1, \ldots, n$
5: **Use** (x_i, r_{im}) $(i = 1, \ldots, n)$ **to find the** m **times tree:**
6: $c_{mj} = \arg\min \sum x_i \in R_{mj} L(y_i, f_{m-1}(x_i) + c)$
7: $h_m(x_i) = \sum_{j=1}^{|R_m|} c_{mj} I(x_i \in R_{mj})$
8: **The prediction function after obtaining** m **regression is:**
9: $r_{im} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)} \quad i = 1, \ldots, n$

GBDT have been widely used for modeling various patterns and simulating the NWP model with weather information. However, GBDT lacks the ability to learn from sequential data, a critical factor in predicting time-dependent phenomena like weather patterns.

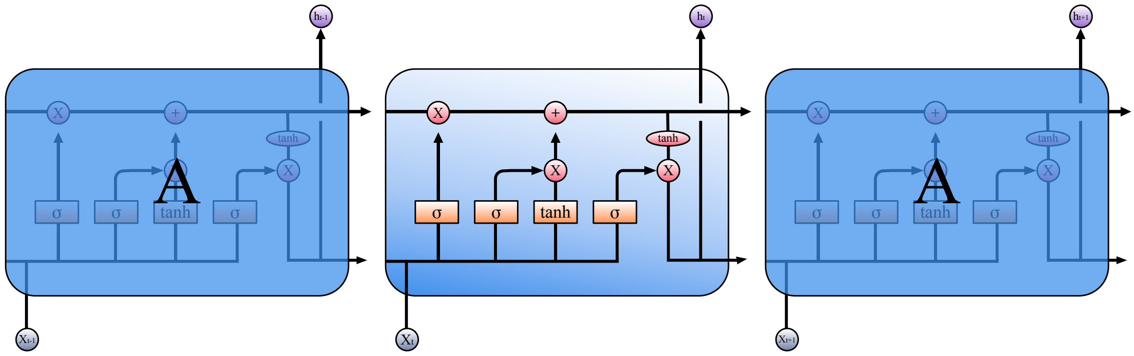

FIGURE 3.2: LSTM logic diagram

3.2.2 LSTM

To address this limitation, this thesis introduce the use of Long Short-Term Memory, which is well-suited for processing time series data. LSTM is a special Recurrent Neural Network (RNN) model [48] that utilizes the long-term dependence of valid data samples and has a long-term memory function at the same time. Compared with the ordinary RNN model, it has three more controllers: input gate, output gate and forget gate, which can effectively solve the problem of perception distance loss during conversion training in the RNN network [49]. The logic diagram is shown in Figure 3.2. Where ht means hidden state, xt means input, σ means sigmoid link layer, x means point wise multiplication, + means point wise addition. In addition, the special neuron mechanism of LSTM has the powerful information transmission and memory storage functions, so it is easier to capture the dependence of time series, reduces the information degradation rate, and increases the depth of calculation. Researchers nowadays take the advantage of the regularity of solar radiation to apply LSTM for short-term prediction, this algorithm is particularly effective when the data sample is limited. The following is a working logic of LSTM [50]:

The first step in LSTM is to decide what information we're going to throw away from the cell state. This decision is made by a sigmoid layer called the "forget gate layer". The forget gate can be expressed as :

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{3.8}$$

h and x respectively represent two vectors associated with each other. W_f is the weight coefficient matrix; σ is the sigmoid function; b_f is the bias parameter.

The second step is to decide what new information we're going to store in the two parts of the cell state. First, "input gate layer," a sigmoid layer, decides which values we'll update. Next, a tanh layer creates a vector of new candidate values, $\tilde{C}_t$, that could be added to the state.

$$\begin{aligned} i_t &= \sigma\left(W_i \cdot \left[h_{t-1}, x_t\right] + b_i\right) \\ \tilde{C}_t &= \tanh\left(W_c \cdot \left[h_{t-1}, x_t\right] + b_c\right) \end{aligned} \tag{3.9}$$

tanh is the tanh function; the weight coefficient matrix of the first part of W_i; W_c is the weight coefficient matrix of the second part; b_i and b_c are used as the bias item parameters of this input level.

The third step is to update the old cell state C_{t-1} into the new cell state C_t. Then we add $i_t * \tilde{C}_t$.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{3.10}$$

Following that the numerical unit coefficient matrix of the state at the moment is multiplied by the tanh function to get the final output result.

$$\begin{aligned} o_t &= \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right) \\ h_t &= o_t * \tanh\left(C_t\right) \end{aligned} \tag{3.11}$$

In the above formula, W_o represents the weight coefficient matrix of the output level; and b_o represents the bias coefficient of the output level.

3.2.2.1 BiLSTM

While LSTM offers significant advantages in handling sequential information, there exists an opportunity to further enhance the learning process. This leads us to explore Bi-directional Long Short-Term Memory (BiLSTM), a variant of LSTM.

BiLSTM is composed of forward LSTM and backward LSTM [51]. Just as Fig 3.3 shows, one string goes from the front to the back and the other string does the opposite. In Figure 3.3, we can see that the input of LSTM passes through Backward Layer and Forward Layer from T4 and T6 directions respectively. Finally, the output value is obtained after splicing T7 and T9. These two strings of LSTM results are calculated to obtain a composite result, which can better capture the time relationship for PV prediction [52].

3.2.2.2 Adam Algorithm in BiLSTM

The Adam (Adaptive Moment Estimation) algorithm is a gradient descent algorithm for optimizing neural network models. Combining the ideas of gradient descent and momentum optimization, it converges on the global optimal solution faster and is relatively robust to the selection of hyperparameters.

The Adam algorithm computes its mean and an exponentially weighted moving average of the mean of squared gradients, respectively. These moving averages are calculated by the following formulas:

$$\begin{aligned} m &= \beta_1 * m + (1-\beta_1) * g \\ v &= \beta_2 * v + (1-\beta_2) * g^2 \end{aligned} \tag{3.12}$$

Among them, m represents the average gradient, v represents the average square gradient, β_1 and β_2 are two hyperparameters, usually set to 0.9 and 0.999 respectively.

Calculate the bias correction coefficient: Since the initial values of the exponentially weighted moving averages are all 0, they need to be bias corrected to avoid the problem that the update step size is too large or too small in the early stage of training. These bias correction factors are calculated by the following formulas, respectively:

$$\begin{aligned} m_{\text{hat}} &= m / \left(1-\beta_1^t\right) \\ v_{\text{hat}} &= v / \left(1-\beta_2^t\right) \end{aligned} \tag{3.13}$$

Among them, t represents the number of current iterations. Since t will increase by 1 in each iteration, the power of β_1 and β_2 will continue to increase. Calculating the update step size: Based on estimates of the average gradient and the mean of the squared gradients, the update step size is calculated by the following formula:

$$\Delta_\theta = -\alpha * m_{\text{hat}} / \sqrt{\nu_{\text{hat}}} + \varepsilon \tag{3.14}$$

Among them, α is the learning rate, which controls the size of the update step, and epsilon is a small number to prevent the denominator from being 0. The sign of the update step is a negative sign, because the Adam algorithm is a gradient descent algorithm, and the parameters need to be updated along the opposite direction of the gradient.

3.2.3 GBDT-LSTM Model With Teacher Forcing

Having delineated the individual merits of GBDT and BiLSTM in previous sections, we now turn our attention to their fusion. The integration of GBDT and BiLSTM aims to leverage the unique strengths of both methods, but it poses a challenge in the context of day-ahead prediction of PV power generation.

As being detailed in the LSTM section, the RNN prediction ability in the early stage of its training iteration process is often weak, leading to ineffective generation results. This can cascade through the learning of subsequent units, causing a compounding effect. To tackle this, we introduce the Teacher Forcing mechanism, a rapid and efficient method to train the cyclic neural network model. Rather than using the output of the previous state as input for the next state, Teacher Forcing updates it with the newly available ground truth data.

However, in the domain of day-ahead PV prediction, power plants are required to submit their estimated 24-hour power generation data to the market trading organization in advance. The teacher forcing mechanism, when used to constrain the LSTM model, can only be updated hourly, making it unsuitable to predict the 24-hour power generation of the next day. This limitation underscores the need for an innovative approach to address the time scale problem.

To address this, our model utilizes the prediction results of GBDT to replace the conventional Ground Truth Predict data. This novel method not only improves the accuracy of the LSTM model but also effectively solves the aforementioned time scale problem.

The proposed GBDT-LSTM model operates in a sequence of steps to accurately predict PV power generation. The algorithm's computational flowchart can be referred to in Figure 3.3. Where ht means hidden state, Xn means LSTM input, x means GBDT weather data. This picture shows the overall logic of the GBDT-BiLSTM model. The output F(X) of GBDT is used as the input value of LSTM through the Teacher forcing mechanism. Here, we present the step-by-step process of the model:

(1) **Input features formation**: The weather data and historical PV data are combined to form the input features X, serving as the initial data for the model.

(2) **Tree splits and features transformation**: The input features X are processed through tree splits and features transformation. This step refines the data and makes it suitable for the GBDT model.

(3) **GBDT model prediction**: An updated learner $G(x)$ is derived from the transformed features, resulting in the GBDT model's prediction of PV power output. $G(x)$ represents the output of the GBDT model and contains valuable information about the underlying structural patterns of the data.

(4) **Integration with BiLSTM using TF mechanism**: Utilizing the TF mechanism, the results of $G(x)$ are fed into the BiLSTM's $X_{(n)}$. This step ensures that the time-dependent nature of the data is captured, and the information from the GBDT model is integrated with the sequential processing capability of the LSTM model.

(5) **BiLSTM neural units computation**: The data then passes through the T1 to T9 neural units of the BiLSTM model. During this stage, computations are performed in both the Backward Hidden Layer and Forward Hidden Layer to calculate the two layers of h_n. This two-layer structure enhances the model's ability to analyze the sequential information in the data.

(6) **Output layer fusion**: Finally, the fused model prediction result is obtained at the output layer. This result is a synthesis of the structural insights provided by the

GBDT model and the time series processing power of the LSTM model, leading to a more robust and accurate prediction of PV power generation.

The synergy between GBDT and BiLSTM, facilitated by the TF mechanism, creates a powerful model capable of handling both the structural and temporal aspects of PV prediction. The above steps collectively contribute to a comprehensive solution tailored for the specific challenges of PV power generation forecasting. And the mathematical output result of this model can be defined as:

$$\psi_t = O_t * \tanh\left(C_t\right) \tag{3.15}$$

Where O_t uses a sigmiod function to represent the output content, which can be defined as:

$$O_t = \sigma\left(W_0\left[h_{t-1}, G(x)\right] + b_0\right) \tag{3.16}$$

At the same time, we can redefine the update layer and forget layer functions in a LSTM cell as:

$$\begin{aligned} \tilde{c}_t &= \tanh\left(w_c \cdot \left[h_{t-1}, G(x)\right] + b_c\right) \\ i_t &= \sigma\left(w_i - \left[h_{t-1}, G(x)\right] + b_c\right) \end{aligned} \tag{3.17}$$

$$z_t = \sigma\left(W_f \cdot \left[h_{t-1}, G(x)\right] + b_f\right) \tag{3.18}$$

$G(x)$ represents the output result of the GBDT PV prediction portion, where x consists of the input features formed from historical weather data and historical PV series. This output captures the structural patterns within the data, specifically the PV prediction result generated by GBDT based on the structural relationship between weather conditions and historical PV series. ψ_t symbolizes the final output result of the model, synthesizing the structural insights from GBDT with the time-series processing capability of LSTM, to provide a more comprehensive and accurate forecast.

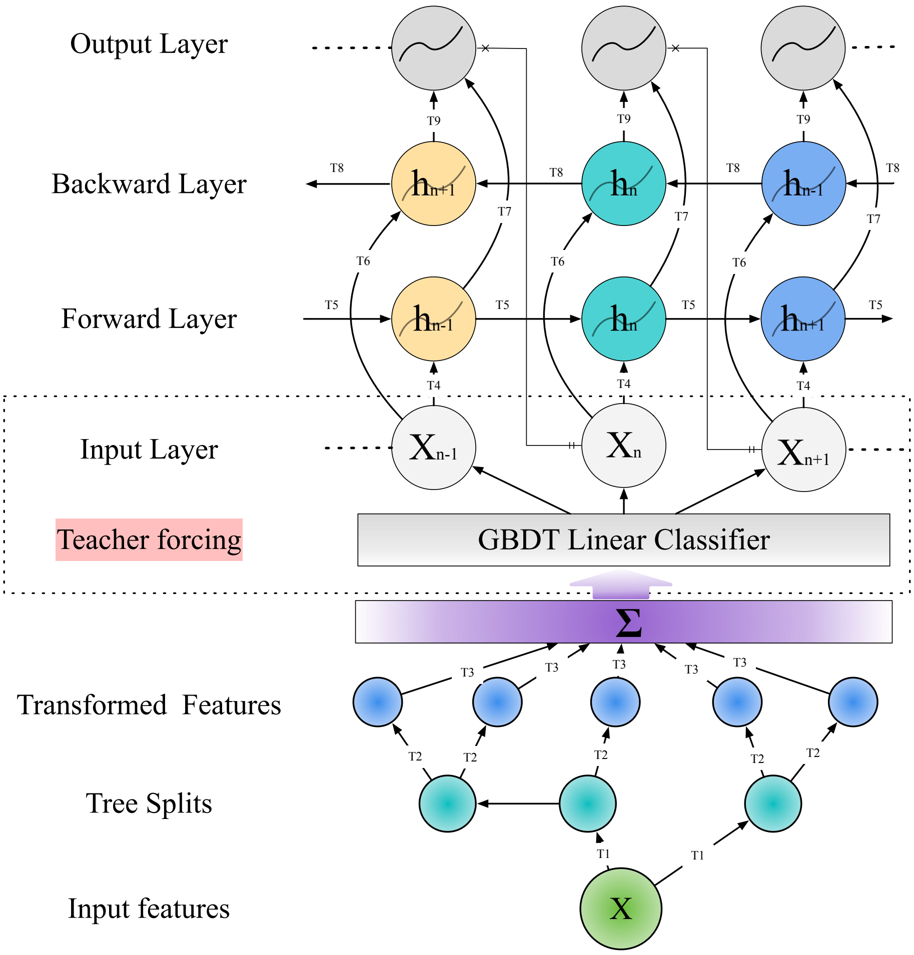

FIGURE 3.3: GBDT-BiLSTM logic diagram

3.3 Performance Metrics

Among various performance metrics, this thesis primarily focus on Mean Absolute Error (MAE), Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), R^2 (Coefficient of Determination), and Root Mean Square Error (RMSE) as the main evaluation criteria. The formulas for these metrics are given by:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{3.19}$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{3.20}$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{3.21}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3.22}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3.23}$$

Where n is the total number of samples, y_i is the actual value for the i^{th} sample, $\hat{y}_i$ is the predicted value for the i^{th} sample, and $\bar{y}$ is the mean of the actual values.

These metrics were chosen among various others because they offer an intuitive way to directly compare the performance of different models, highlighting their strengths and weaknesses.

3.4 PV Forecast Forecast Result

In this chapter, this thesis focus on the results of our predictive models specifically tailored for solar PV output forecasting. The models operate at two primary time intervals: 1-hour and 5-minute forecasts. For the 1-hour interval, the model outputs a one-time 24-hour forecast, which is intended to provide valuable insights for the day-ahead electricity market. Within the realm of 5-minute forecasts, this thesis further categorize them into short-term forecasts for the immediate next 5 minutes and medium-term forecasts extending up to four hours. The 5-minute short-term forecasts are crucial for real-time grid management, enabling quick adjustments to adapt to rapid changes in solar PV output. The medium-term forecasts, meanwhile, are designed to align with the typical charge and discharge cycles of battery storage systems, facilitating optimized battery scheduling. The forecasting approach for large-scale PV stations leverages the uniformity of environmental impacts due to their specific geographical locations. In these settings, similar weather zones are treated as a single forecasting unit, simplifying the prediction process. However, VPPs present a more complex scenario. Due to the varied locations of solar panels and diverse weather conditions, each unit within a VPP is smaller and more nuanced compared to a unit in a PV station. This granularity in VPPs necessitates individual adjustments for each unit, ensuring precise and reliable forecasting. This distinction in approach between VPPs and large-scale PV stations highlights the complexity and adaptability required in VPP forecasting models, crucial for their efficient and accurate operation.

3.4.1 One Hour Interval Forecast Result

To validate the efficacy of the proposed algorithm, various tests have been carried out based on the Pytorch operating environment using Azure Linux 23.02, 06 system with the virtual machine size being Standard-NC6 (6 cores, 56 GB RAM, 380 GB disk) and the processing unit is GPU - 1 x NVIDIA Tesla K80.

The experimental data used in this study were collected from a PV power station in Ganzi Tibetan Autonomous Prefecture, Sichuan Province, China. The data span a period of 18 months, from the second half of 2020 to 2021. This includes meteorological forecast data from the European Central Weather Center and power generation data from the PV power station. The meteorological data, which encompass temperature (2 meters above the ground), humidity (2 meters above the ground), air pressure, precipitation, ground surface wind speed (10 meters above the ground), wind direction (10 meters above the ground), and total radiation (from the PV panels of the power station), were collected at 1-hour intervals. However, the power generation data, derived from the historical records of the PV power station, were collected at daily intervals. In addition, before the data preprocessing phase, this thesis utilized the Radial Basis Function (RBF) interpolation to align both the hourly weather forecast data and the daily data from the PV power station, which includes temperature, irradiance, wind speed, and humidity. This ensures consistency and seamless integration for our predictive model.

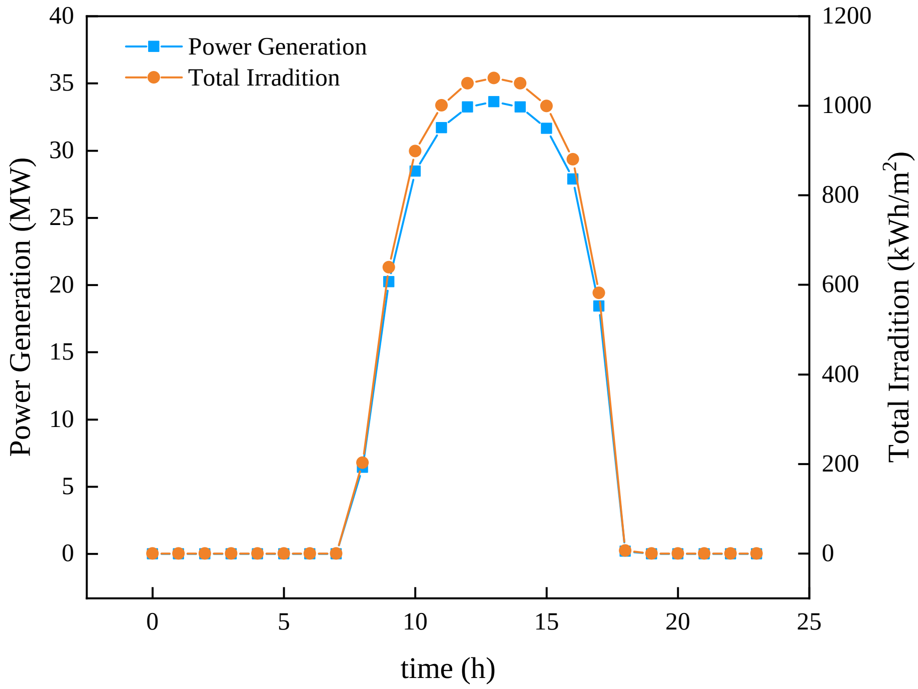

FIGURE 3.4: The comparison between power generation and the PV irradiation

3.4.1.1 Data Preprocessing

In this experiment, the last month data were used as the validation set, with an hourly sampling rate. The time window sampling used was a 24-hour period for both the GBDT model and the LSTM model. And the experiment focuses on the daily change in solar radiation. The total PV power generation for a day was discretized into hourly increments, with the discretized function representing the daily change in solar radiation. Figure 3.4 compares the 24-hour change in irradiation with the discrete change in power generation.

The outlier data are removed and their places as well as missing data are filled in by Difference Filling methods in order to reduce the impacts of the bad data on the LSTM algorithm. The PV power generation data for 18 months are displayed in Figure 3.5.

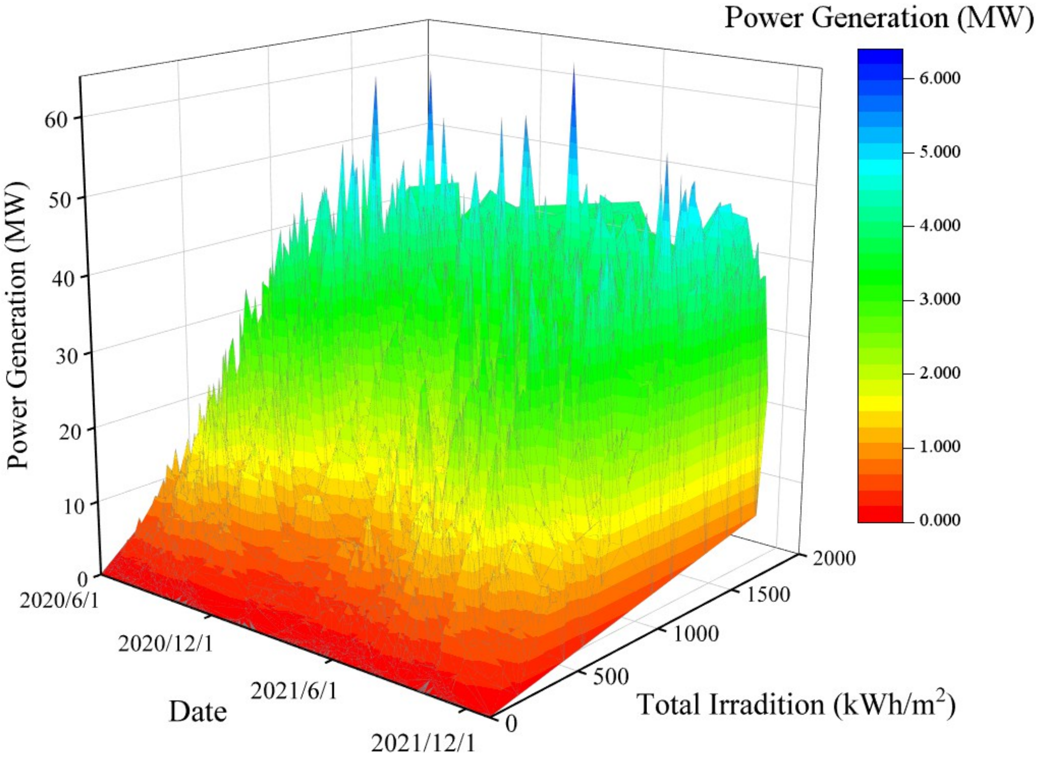

FIGURE 3.5: Three-dimensional data map of the total radiant power generation from June 2020 to December 2021

3.4.1.2 GBDT Modeling Results

Based on the NWP physical model, this experiment included the weather data, such as temperature, moderate humidity, precipitation, and wind speed. However, the NWP model requires high accuracy and data breadth in actual engineering applications. To address this, we adjusted the weight of missing values, and allowed prediction of future power generation even in the absence of certain weather features. To demonstrate this, we randomly set 30% of the data for each feature, excluding solar radiation, to missing values (-1). And then we refitted the training data and completed the prediction. The simulation fitting curve for the last 300 hours of the GBDT algorithm test set is shown in Figure 3.6.

The Figure 3.6 result shows that the prediction accuracy has declined after the operation in Missing value. However, our algorithm successfully capture the strong correlations between solar radiation characteristics and PV power generation, thus the absence of weather measurements to 30% has little impact on the prediction results and the prediction error is controlled

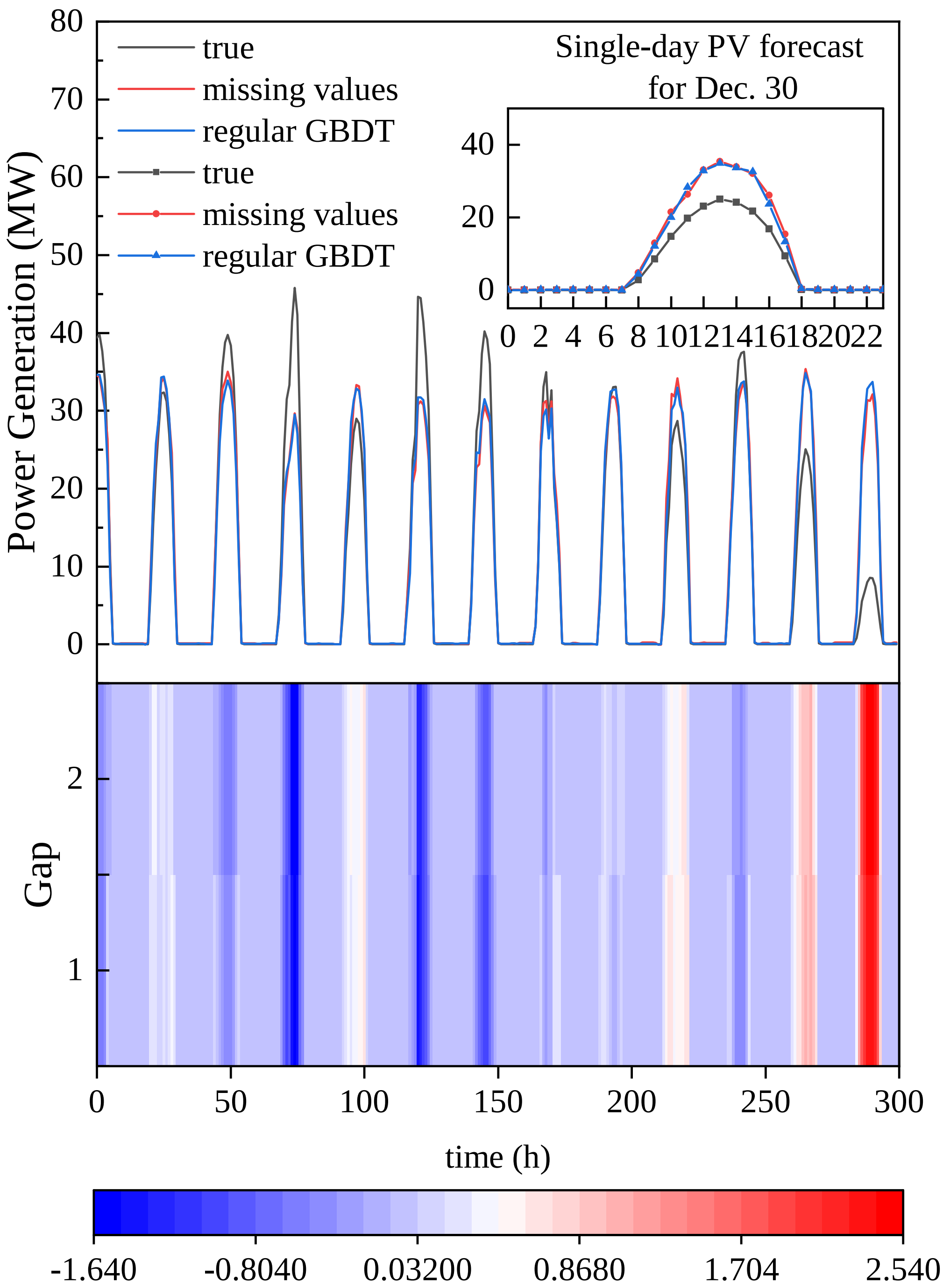

Figure 3.6: Last 300 hours of the GBDT model test set

within a reasonable range. Similarly, the 30% missing values set in the experiment have also been adjusted and modified. After the data verification, the root mean square error of

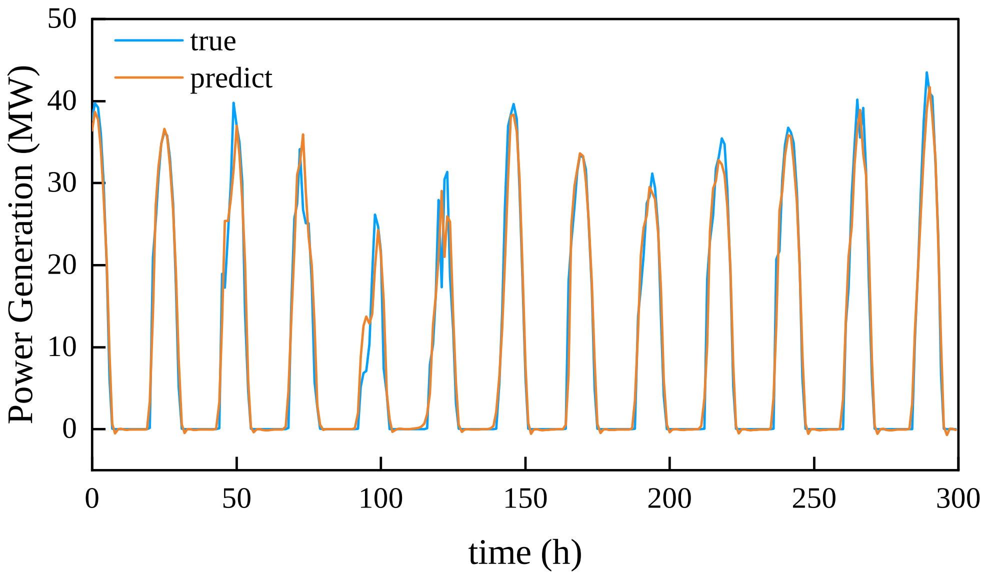

FIGURE 3.7: Last 300 hours of the BiLSTM model training set

this model was 2.0162 (when all features were intact) and 2.150 (when some features were missing), which verifies the robustness of the GBDT model in this part.

3.4.1.3 GBDT-BiLSTM Modeling Results

In this phase of experiment, we used the deep learning algorithm of the BiLSTM model to simulate and predict historical power generation data first. Different from the conventional BiLSTM model, we used the prediction result of GBDT to replace the ground truth requirements in Teacher Forcing mechanism when updating the LSTM model. BiLSTM had the same time partition criterion with the GBDT model. The environment of BiLSTM's batch size was set to 1024. The Adam optimization algorithm is used here with the mean square error as the loss function and 200 times as data traversals. The training result is shown in the Figure 3.7. The mse and mae of the training set are 0.054821357 and 0.12106604, respectively. Compared with 0.17525618664000464 and 0.2018899794652092 of the GBDT model, it can be seen clearly that the fitting effect of the BiLSTM model is better than that of the GBDT model.

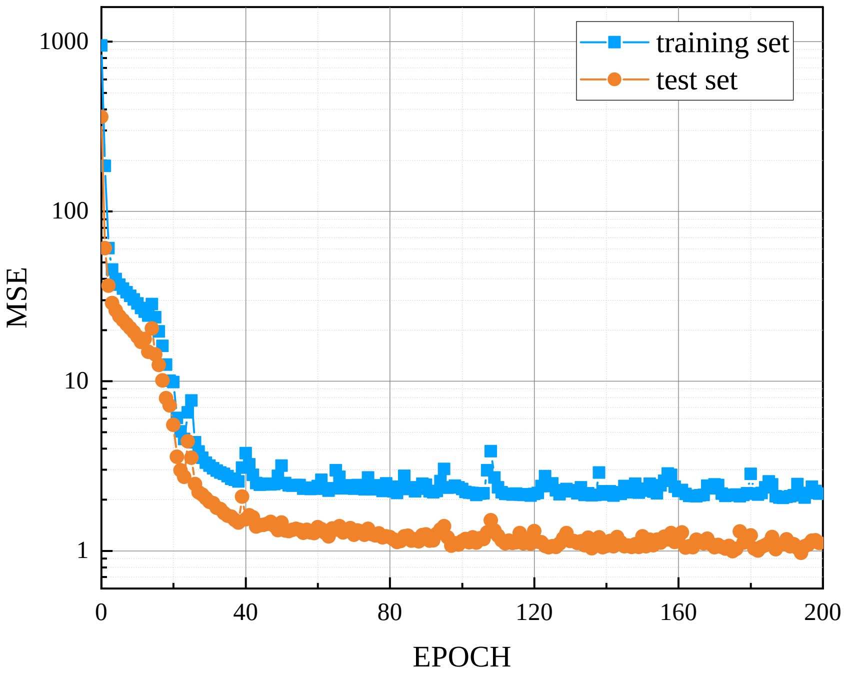

FIGURE 3.8: MSE curves of the model training set and the test set

Secondly, we used the test and data to perform the mse fitting test and save the mse change curve. After 200 iterations, the mse curves of the model training set and the test set tended to be consistent. The fitting results are shown in the Figure 3.8.

Then we output the last 300 hours training curve of the test set. Compared with Figure 3.9, it can be seen that the fitting effect of the BiLSTM model in the day-ahead PV prediction model is better.

Third, we imported the Teacher Forcing mechanism and used the prediction of GBDT as the ground truth data to update the prediction result of BiLSTM. Temperature, air humidity, air pressure, precipitation, surface wind speed and total radiation were imported as features and the time scale remained unchanged. The final prediction result is shown in the Figure 3.10. This picture includes five models: LSTM, GBDT, update predict LSTM, GBDT-LSTM and True.

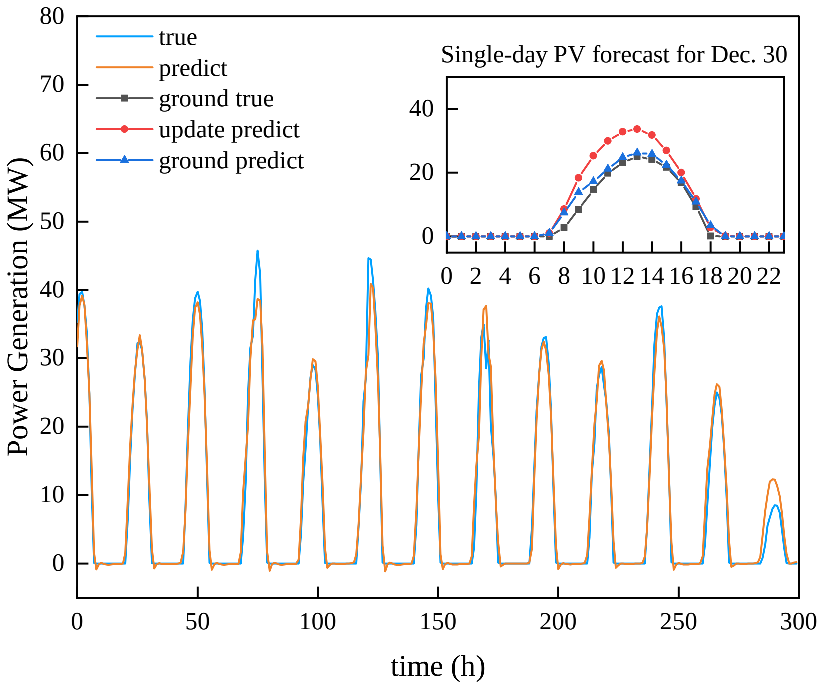

FIGURE 3.9: Last 300 hours of the BiLSTM model test set

Finally, as shown in Figure 3.10, there are five curves in this picture, among which the black one is the ground truth. The green one is the predicted value based on the Teacher Forcing mechanism provided by true power generation data. The red one is the GBDT-BiLSTM model, which is closer to the true value than the predictions of GBDT and the single BiLSTM models. The shaded region between the ground truth and the actual values represents the difference between them, signifying the ideal performance boundary of the model and indicating its potential upper limit.

3.4.1.4 Model Comparison

In order to verify the effectiveness of the new model, this paper compares the new model with SVM [53], Transformer, GBDT and LSTM models using the same database. The comparison results are shown in Figure 3.11. This picture includes five models: SVM, GBDT, update predict LSTM, GBDT-LSTM and update predict transformer. In the figure presented, it is

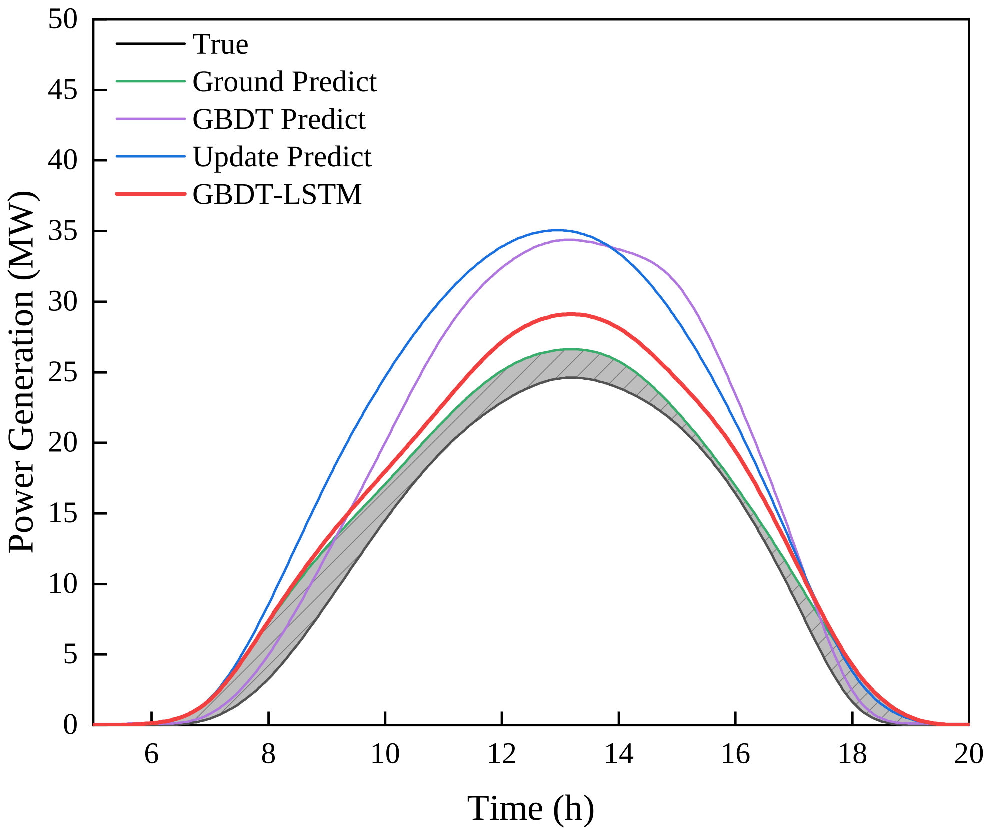

FIGURE 3.10: GBDT-BiLSTM predicts power generation on December 30

evident that the prediction fit of the newly proposed model in the last 300 hours of the test set is the closest to the actual value.

In Figure 3.12, December 30th serves as a case study to demonstrate the comparison between the new model and other models. Despite the time scale problem associated with the teacher forcing method in the domain of day-ahead PV prediction, the use of GBDT prediction results as a replacement for conventional Ground Truth Predict data has not only enhanced the accuracy but also solved this unique challenge. Furthermore, our results underscore that our model exhibits robust performance, even when confronted with missing or inaccurate weather data.

Focusing on the test set error values, the BiLSTM model with teacher forcing achieves the lowest values in MAE (0.08), MSE (0.151), and MASE (0.05), indicating its superior

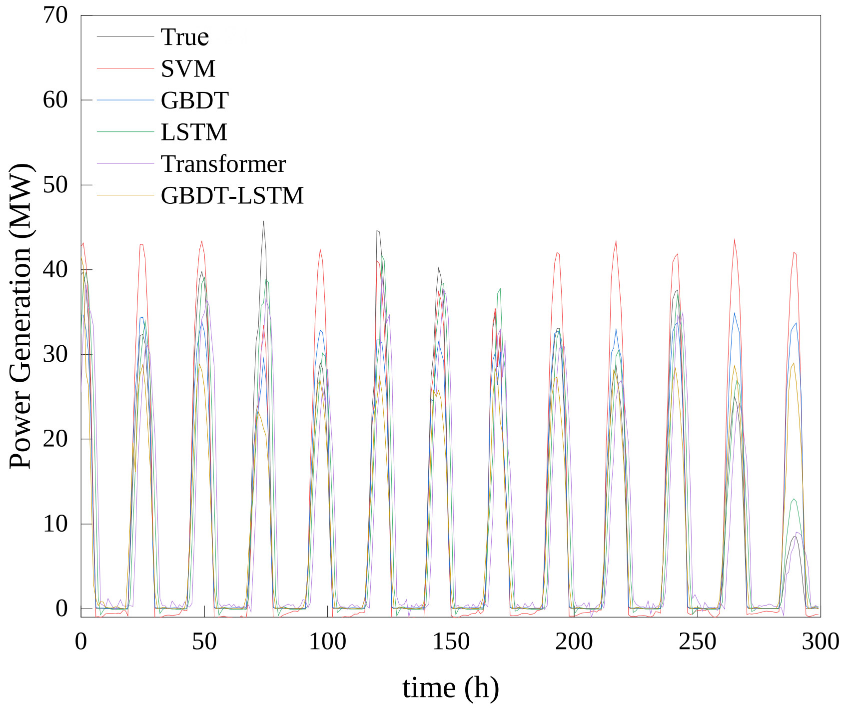

FIGURE 3.11: Last 300-hour forecast curves of multiple model test sets

prediction accuracy. The presence of LSTM with teacher forcing in the final model comparison results provides a pertinent benchmark, demonstrating the innovative approach of the hybrid GBDT-BiLSTM model. In addition, the GBDT-BiLSTM hybrid model also demonstrates good performance with an MAE of 0.112, MSE of 0.141, and MASE of 0.102, exhibiting an enhanced capability that approaches the ideal case of teacher forcing. On the other hand, other models such as SVM, GBDT, Improved GBDT, and Transformer exhibit relatively poorer performance across all metrics.

In the comparison of the Figure 3.12 and Table 3.1, our model stands out in terms of both prediction accuracy and robustness against inconsistent weather data. The 3.12 includes SVM, GBDT, update predict LSTM, GBDT-LSTM and ground true predict Transformer. The model is especially effective in situations where the weather data might be incomplete or not entirely accurate. This makes our model a reliable choice for real-world applications, ensuring consistent prediction even when faced with challenges in data quality.

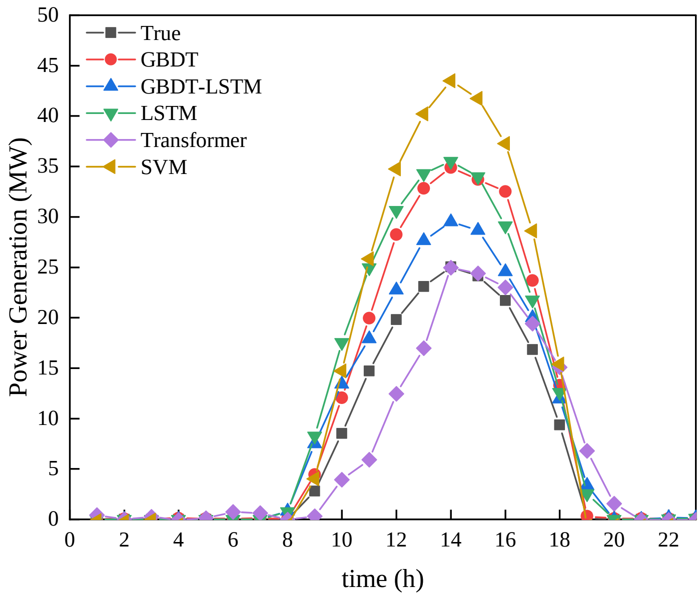

FIGURE 3.12: Multiple models predict power generation on December 30

TABLE 3.1: Comparison of prediction performance

Prediction Model	Training set		Test set	
	mae	mse	mase	mse
TF-LSTM (ideal)	0.08	0.151	0.05	0.124
SVM	0.253	0.216	0.219	0.187
GBDT	0.191	0.141	0.179	0.197
Improved GBDT	0.199	0.150	0.203	0.189
Transformer	0.339	0.312	0.345	0.297
GBDT-BiLSTM	0.112	0.141	0.102	0.128

3.4.2 Five Minutes Interval Forecast Result

To validate the efficacy of the proposed algorithm for 5-minute interval PV forecasting, various tests have been conducted based on the PyTorch operating environment running on a Windows 10 Professional system. The hardware specifications include a 13th Gen Intel(R) Core(TM) i7-13700F processor operating at 2.10 GHz, with 128 GB of available RAM. The system is a 64-bit operating system based on an x64 processor, and the graphics card is an NVIDIA 4070Ti.

The experimental data used for this study were collected from a user in Sydney, Australia, at 5-minute intervals. The data span a period of 10 months, from January 2022 to October 2022. The primary data sets include Meter Power (kW), Battery Power (KW), and Solar Power Output (KW), all recorded at these regular 5-minute intervals. Before proceeding to the data preprocessing phase, this thesis utilized Radial Basis Function (RBF) interpolation to align these three data sets and to impute any missing values. This ensures consistency and seamless integration for our predictive model operating at the same 5-minute intervals.

3.4.2.1 Data Preprocessing and LSTM Model Definition

In this study, specific methodologies for data preprocessing and model construction were employed to optimize performance. Initially, the feature "SOLAR POWER OUTPUT(KW)" was normalized to scale all the values within the range of [0, 1]. This step aids in accelerating the model's convergence and reducing the training time, as it ensures that all input features are on the same scale. Following normalization, features and labels were constructed based on the normalized power output data. Each feature sequence consists of 24 consecutive power output readings, aimed at capturing the daily pattern of power generation for more accurate forecasting.

For the division of training and testing sets, data from the last four months were designated as the test set, while the remaining data were used for training. This division allows for a more realistic simulation of the model's performance on unseen data.

In terms of model definition, a Bidirectional LSTM layer followed by a fully connected layer was employed. The LSTM layer accepts a single feature as input and has 32 hidden units. This layer is bidirectional and is stacked in two layers to increase the model's complexity and capture more intricate dependencies. During the forward propagation, the LSTM layer first processes the input sequence, and the output is then passed to the fully connected layer to generate the final prediction. Lastly, the model is trained using the Adam optimizer and Mean Squared Error (MSE) as the loss function. The Adam optimizer adaptively adjusts the learning rate, while MSE quantifies the difference between the predicted and actual values.

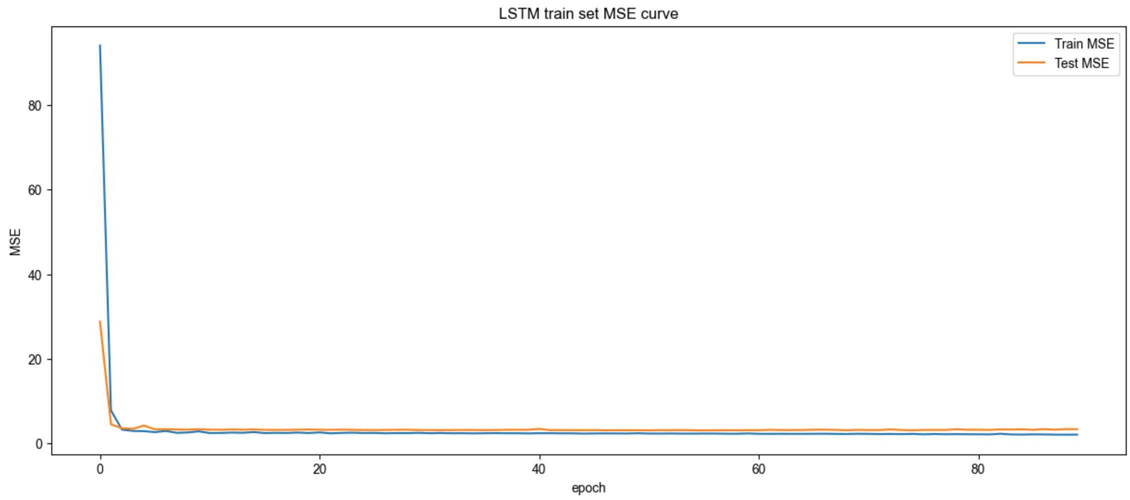

FIGURE 3.13: LSTM train set MSE curve

3.4.2.2 5-Minute for Next 5 Minutes Result

Building upon the methodologies and data preprocessing techniques outlined in the previous subsection, we now turn our attention to the results of the 5-minute interval forecasts for the immediate next 5 minutes. To provide a comprehensive understanding of the model's performance during the training process, Figure 3.13 illustrates the Mean Squared Error (MSE) curves for both the training and testing sets. The blue curve represents the MSE for the training set, while the orange curve depicts the MSE for the test set. This graphical representation serves as an initial insight into the model's ability to generalize from the training data to unseen data, as well as its convergence behavior.

To further evaluate the effectiveness of our training process, we present the following performance Figure 3.14, which showcases the model's predictions over a span of 300 hours on the training set. This extended timeframe allows us to scrutinize the model's capability to capture the underlying patterns and trends in the data, thereby providing a more detailed assessment of its predictive accuracy.

While the training set performance provides valuable insights into the model's learning capabilities, the true test of its generalizability and robustness lies in its performance on the testing set. The testing set serves as an unbiased indicator of the model's predictive power, as it comprises data points that the model has not encountered during the training phase.

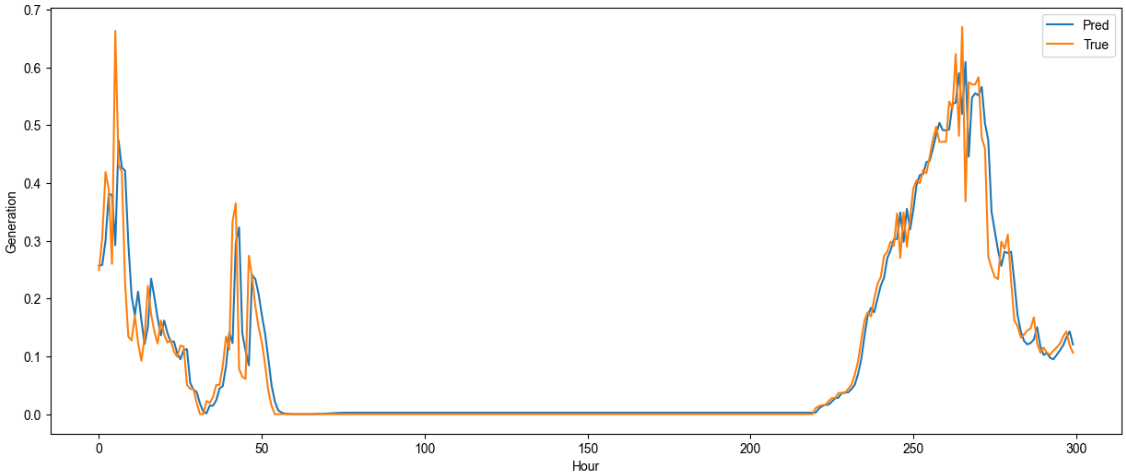

FIGURE 3.14: PV generation of 300 hours training set

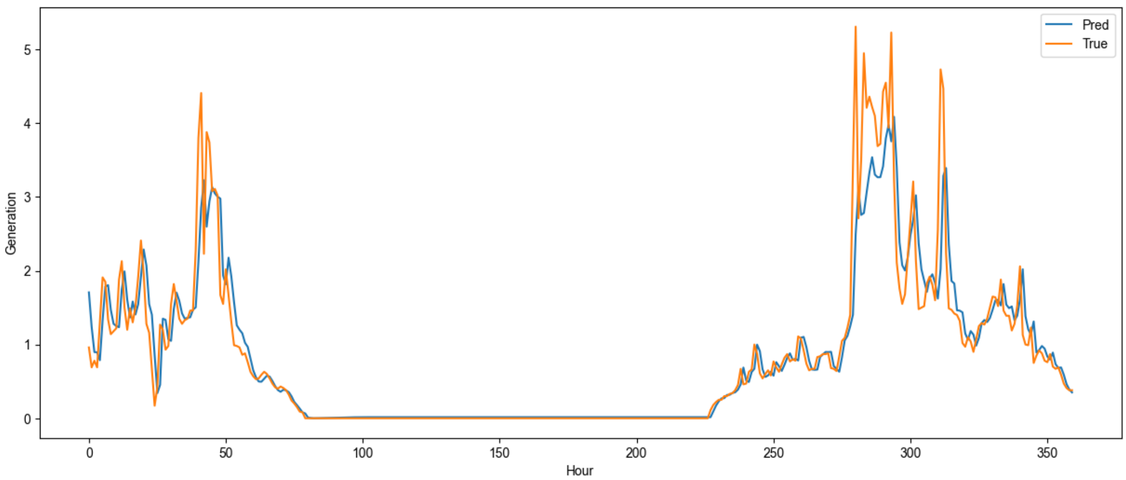

FIGURE 3.15: PV generation of 300 hours test set

Therefore, we present the Figure 3.15 to illustrate the model's predictions over a specific timeframe on the testing set.

After evaluating the model's performance on both training and testing sets, it is crucial to examine how well the model's predictions align with the actual solar power output over a 24-hour period. To make the results more interpretable, the predicted values have been inverse-transformed from their normalized state back to their original scale. Figure 3.16 presents a side-by-side comparison of the actual and predicted solar power output over a

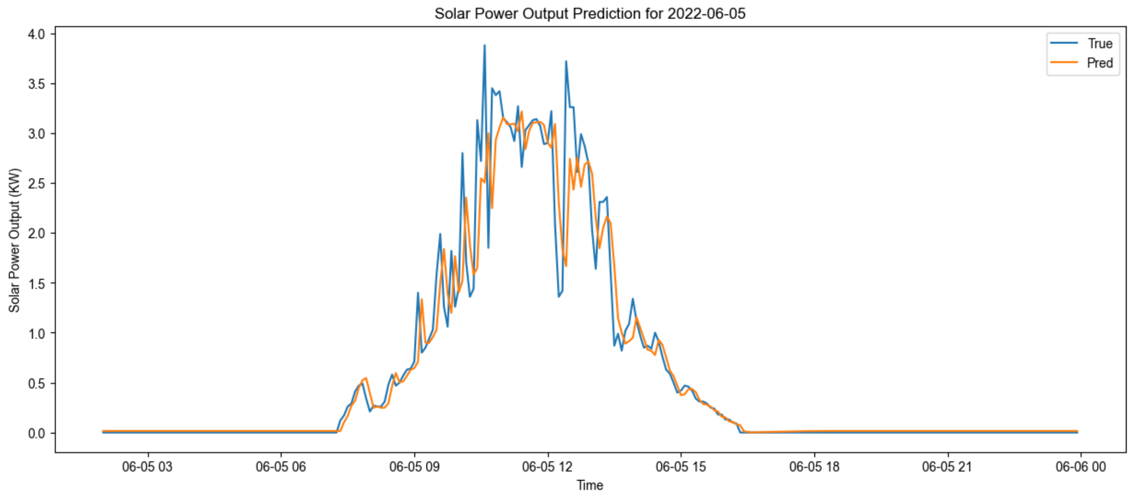

FIGURE 3.16: PV generation of a single day

single day, providing a detailed view of the model's forecasting accuracy in a real-world scenario.

3.4.2.3 5-Minute for Next 30 Minutes Result

Having examined the model's performance for immediate 5-minute forecasts, we now extend our focus to the next 30 minutes. This longer forecasting horizon is particularly relevant for scenarios that require more extended planning, such as optimizing battery charge and discharge cycles. The LSTM parameters and architecture remain consistent with the previous sub-section to maintain a uniform evaluation metric.

To offer a comprehensive view of the model's predictive capabilities over the next 30 minutes, Figure 3.17 presents a composite graph consisting of six sub-plots. The y-axis represents power in kilowatts (kW), and the x-axis indicates the steps over the observed period.. Each sub-plot represents the model's forecast at a different 5-minute interval within the 30-minute window. This multi-faceted representation allows us to assess the model's ability to maintain its predictive accuracy over a more extended period, thereby providing a robust evaluation of its performance for medium-term forecasting tasks.

As evidenced by the composite graph, the model demonstrates predictive accuracy over the 30-minute forecasting horizon. The close alignment between the predicted and actual values

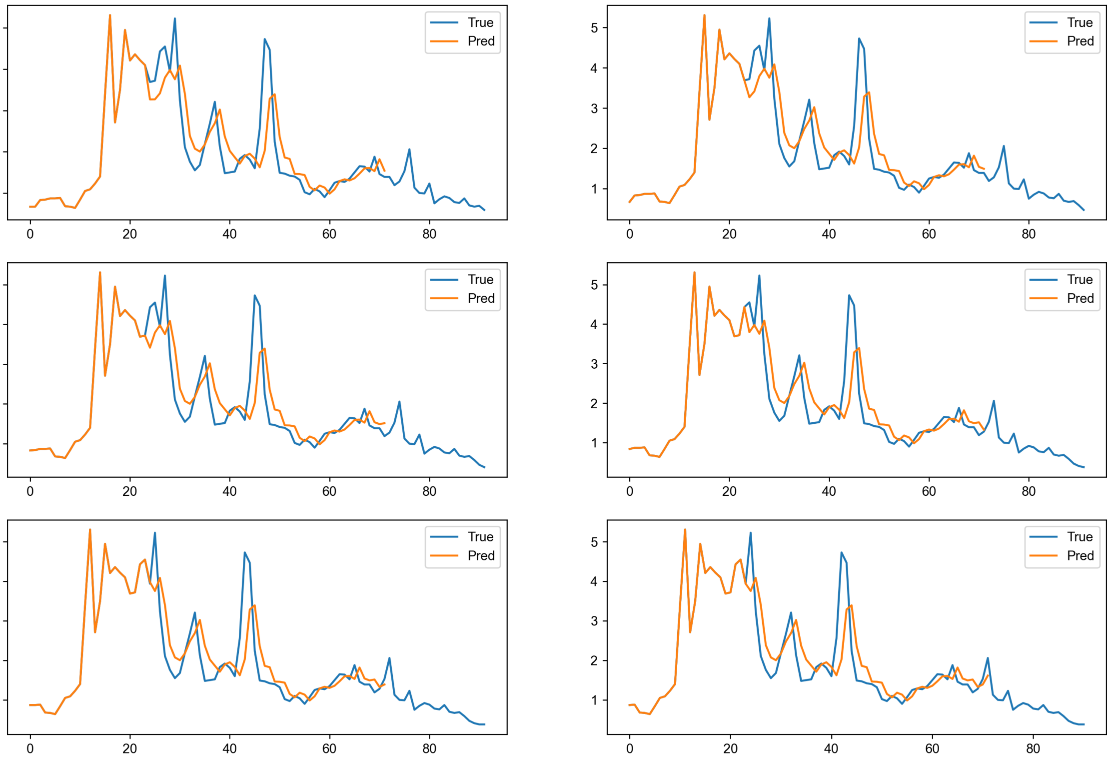

FIGURE 3.17: Composite PV generation of 6 step forecast

across the six sub-plots indicates that the model is capable of maintaining its performance even when extended to medium-term forecasting tasks. This robustness not only validates the model's utility for real-time grid management but also highlights its potential for more complex applications, such as optimizing battery storage systems over longer periods. Overall, the results affirm the efficacy of employing LSTM-based models for solar power forecasting at varying time intervals.

3.5 Chapter Summary

In summary, this chapter presents an array of forecasting models that are integral to the overall operation of solar panels and battery storage systems. Specifically, this thesis introduced a 24-hour forecasting model with a 1-hour interval, which incorporates weather variables for a comprehensive outlook and is crucial for day-ahead market planning. This thesis also detailed models with a 5-minute interval for immediate and up-to-30-minute forecasts, essential for real-time grid management. The high accuracy achieved across these models enables precise and efficient scheduling of battery charging and discharging cycles, thereby enhancing the stability and efficiency of grid operations. Importantly, these models are not just theoretical constructs but are part of an actual project for a retail company in Sydney, Australia. To further bolster the model's practical utility, this thesis have also developed an automated user classification program. By reducing uncertainties in solar power generation and offering actionable insights, these predictive models and additional features significantly contribute to the real-world applicability and effectiveness of integrated solar and battery systems.

CHAPTER 4

Load Forecast

4.1 Introduction to Load Demand Forecasting in VPPs

Load demand forecasting is another critical component in the effective management of VPPs. Accurate predictions of electricity consumption patterns are essential for a variety of operational aspects within a VPP.

Firstly, understanding load demand is vital for grid stability. Mismatches between supply and demand can lead to frequency imbalances, affecting the overall reliability of the electrical grid. Secondly, accurate load forecasts enable more efficient resource allocation, minimizing the need for expensive last-minute adjustments or the activation of reserve generators. Thirdly, precise load forecasts are crucial for optimizing the charge and discharge cycles of battery storage systems, thereby enhancing the economic viability of the VPP.

Given these engineering challenges, there is a compelling need for advanced load forecasting models that are both accurate and computationally efficient, particularly in the context of VPPs. This section explores the mathematical models and technical methodologies employed in this research to tackle these real-world issues. We examine forecasts at various time intervals, utilizing a combination of machine learning and deep learning techniques to achieve optimal performance.

In this Results section, we present the outcomes of our predictive models across various forecasting horizons and time intervals, tailored to different operational scenarios. Specifically,

The content in the section on Load Forecasting is an extension of the authors' previous work presented at the ACFPE 2023 conference, under the title "A Novel Ensemble Model for Load Forecasting: Integrating Random Forest, XGBoost, and Seasonal Naive Methods".

we include solar PV forecasts with 1-hour and 5-minute intervals, load demand forecasts with 1-hour, 30-minute, and 5-minute intervals, and electricity price forecasts at 5-minute intervals. Within the 5-minute solar PV and load demand forecasts, we further distinguish between predicting the immediate next 5 minutes and forecasting up to four hours into the future. Similarly, the 5-minute electricity price forecasts are also categorized into short-term (next 5 minutes) and medium-term (up to four hours) forecasts.

The rationale for these different forecasting horizons is twofold. The 5-minute forecasts are particularly useful for real-time grid management, allowing for quick adjustments to meet immediate demand or to capitalize on sudden changes in electricity prices. The 1-hour and four-hour forecasts, on the other hand, are designed to align with different operational needs. The 1-hour load demand forecast aims to provide a near-future outlook, while the four-hour forecasts align with the typical charge and discharge cycles of battery storage systems, thereby facilitating more effective battery scheduling for optimized grid operations.

4.2 Methods

4.2.1 Transformer Algorithm

In this section, we propose a specialized Transformer model for time-series forecasting. The Transformer architecture, originally designed for natural language processing tasks, has been adapted to suit the nature of our time-series data. The logic of the specialized Transformer is shown in Fig 4.1. Unlike the standard Transformer model, which comprises both an Encoder and a Decoder, our model only employs the Encoder part for sequence modeling. The rationale behind omitting the Decoder is to leverage the Encoder's capability of capturing complex relationships in the input data and directly map it to a future value, thereby making the model computationally more efficient for our specific task.

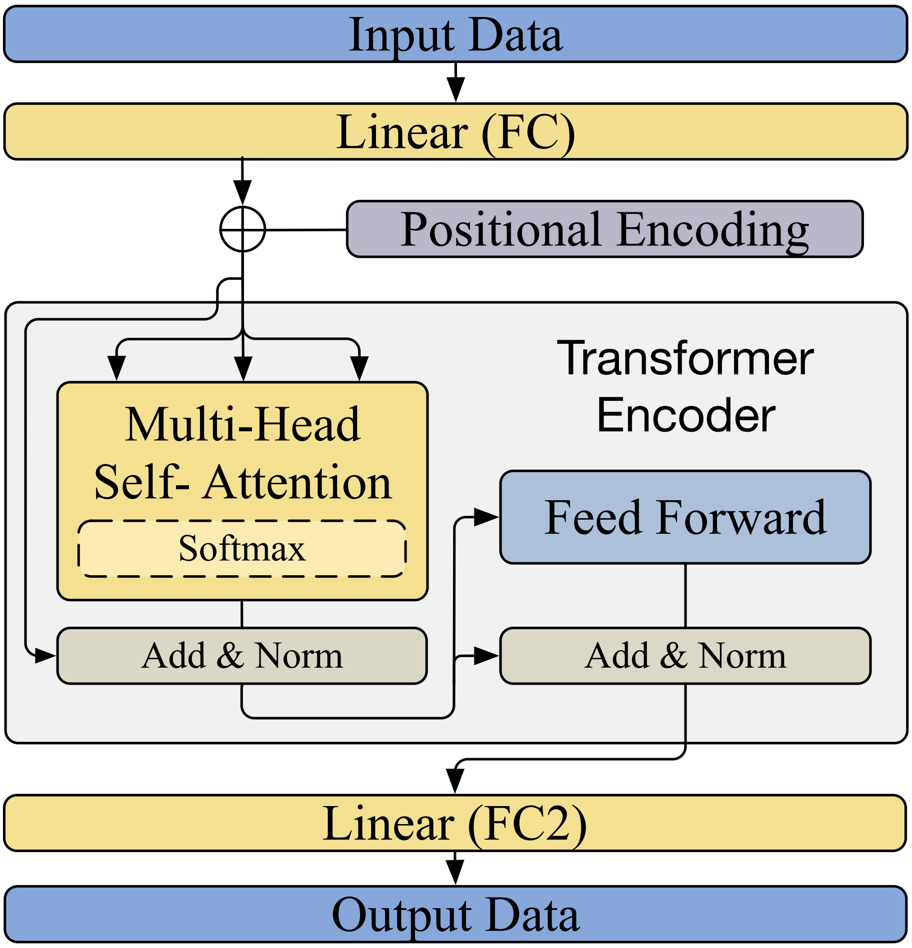

FIGURE 4.1: Transformer logic diagram

4.2.1.1 Encoder

The Encoder is the cornerstone of our model and is responsible for converting the input time-series data into a set of continuous latent variables. Each Encoder layer consists of two main components:

(1) Multi-Head Self-Attention Mechanism
(2) Feed-Forward Neural Networks

The architecture can be represented as follows:

Encoder Layer = Multi-Head Attention $\rightarrow$ Add & Norm $\rightarrow$ Feed-Forward $\rightarrow$ Add & Norm

4.2.1.2 Input Embedding

Before feeding into the Encoder, the input time-series data X of shape (T, D), where T is the sequence length and D is the feature dimension, is linearly transformed to a higher-dimensional space $\mathbb{R}^{T\times H}$, where H is the model's hidden size:

$$X' = XW + b$$

Additionally, positional encodings are added to the transformed input to provide the model with sequence order information.

4.2.1.3 Multi-Head Self-Attention Mechanism

The Multi-Head Self-Attention allows the model to focus on different parts of the input sequence for each output element. Given a query Q, a key K, and a value V, the attention mechanism can be described as:

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V$$

Where d is the dimension of the query and key.

In the multi-head context, this operation is performed multiple times in parallel, allowing the model to focus on different aspects of the input.

$$\text{Multi-Head Attention}(Q, K, V) = \text{Concat}(\text{head}_1, \text{head}_2, ..., \text{head}_n)W^O$$

Where each $\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$ and W^O is the output weight matrix.

4.2.1.4 Feed-Forward Neural Networks

Each attention output then passes through a feed-forward neural network, the same one for each position, followed by another Add & Norm operation.

$$\text{FFN}(x) = \text{ReLU}(xW_1 + b_1)W_2 + b_2$$

4.2.1.5 Residual Connection (Add & Norm)

One of the key features of our model is the incorporation of residual connections, also known as Add & Norm. After both the Multi-Head Attention and Feed-Forward Neural Network modules, the input sequence goes through a residual connection followed by layer normalization. Mathematically, the residual connection and layer normalization can be represented as:

$$\text{Add \& Norm}(x, \text{SubLayer}(x)) = \text{LayerNorm}(x + \text{SubLayer}(x))$$

Where $\text{SubLayer}(x)$ could be either the output from the Multi-Head Attention or the Feed-Forward Neural Network.

4.2.1.6 Softmax in Multi-Head Attention

The Softmax function in the Multi-Head Attention mechanism serves to assign varying degrees of importance to different parts of the input sequence. Given the scaled product of the query Q and key K, Softmax is applied along the last dimension to produce the attention weights:

$$\text{Softmax}\left(\frac{QK^T}{\sqrt{d}}\right)$$

These attention weights are then used to take the weighted sum of the value V, effectively allowing the model to focus more on parts of the sequence that are more relevant to the task at hand.

4.2.1.7 Linear Transformation Layers (LinearFC and LinearFC2)

In our model, two linear transformation layers (commonly known as fully connected layers) are employed. The first one, referred to as *LinearFC*, is used immediately after the input layer for feature transformation. This linear layer maps the original feature dimensions to the hidden dimensions as follows:

$$X' = XW_{\text{fc}} + b_{\text{fc}}$$

Where W_{fc} and b_{fc} are the weight matrix and bias term, respectively. This transformed feature set X' is then fed into the encoder layers.

The second linear layer, known as *LinearFC2*, is placed at the end of the encoder layers and serves as the output layer for the model. It maps the hidden dimensions back to the output dimensions as follows:

$$\hat{Y} = ZW_{\text{fc2}} + b_{\text{fc2}}$$

Here, Z is the output from the last encoder layer, and $\hat{Y}$ is the predicted output. W_{fc2} and b_{fc2} are the weight matrix and bias term for this layer, respectively.

4.2.1.8 Positional Encoding

To address the lack of sequence information in the standard Transformer architecture, a Positional Encoding is added to the inputs before they are fed into the Encoder. The Positional Encoding function PE is defined as:

$$PE_{(pos,2i)} = \sin\left(\frac{pos}{10000^{2i/d}}\right)$$
$$PE_{(pos,2i+1)} = \cos\left(\frac{pos}{10000^{2i/d}}\right)$$

Where pos is the position and i is the dimension. These Positional Encodings are added to the input embeddings at corresponding positions, thus providing the model with information about the order of the sequence.

4.2.2 Random Forest Algorithm

Random Forest is a collection of multiple Decision Trees (DTs) [54] and is known for its simple structure and strong anti-noise capabilities. It is highly resistant to overfitting, making it suitable for applications like power system load forecasting.

Initially, n samples are randomly selected with replacement from the original sample space to form a new training set. This process is repeated n_{tree} times, each time generating a new training set for a Classification and Regression Tree (CART).

The prediction Y made by a Random Forest model can be represented as the average of the predictions from each individual decision tree T_i:

$$Y = \frac{1}{n}\sum_{i=1}^{n} T_i(X) \tag{4.1}$$

Where, $T_i(X)$ is the prediction made by the i^{th} decision tree, and n is the total number of decision trees in the forest.

Each individual decision tree T_i in the Random Forest is constructed to minimize a certain impurity measure. In this paper, we use the Gini impurity measure method. The Gini impurity for a node t is calculated as:

$$\text{Gini}(t) = 1 - \sum_{i=1}^{c} p_i^2 \tag{4.2}$$

Where, P_i is the proportion of samples belonging to class i at node t, and c is the number of classes.

4.2.3 XGBoost Algorithm

While Random Forest excels in robustness, its capabilities can be further enhanced for complex data scenarios. To address this, we also employ XGBoost, known for its high predictive accuracy and efficiency. The XGboost logic diagram is shown in the Figure 4.2.

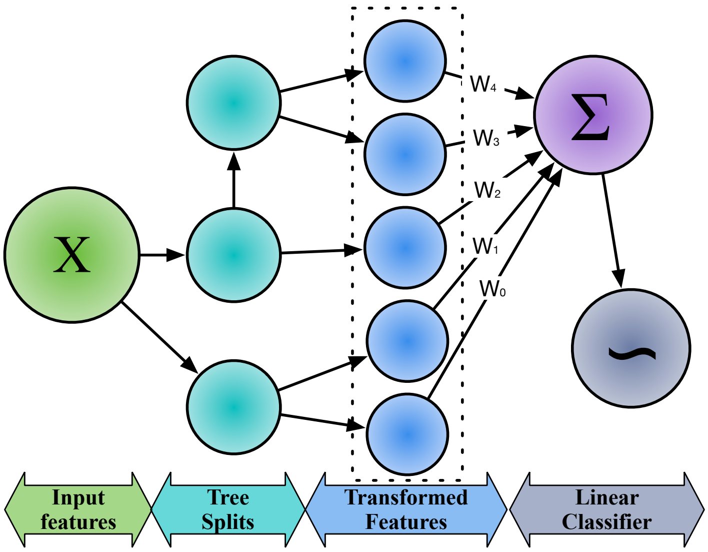

FIGURE 4.2: XGBoost logic figure

Thus, XGBoost serves as a complementary approach to Random Forest to improve the overall forecasting performance. The objective function $\mathcal{L}(\phi)$ in XGBoost can be represented as[55]:

$$\mathcal{L}(\phi) = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i\right) + \sum_{k=1}^{K} \Omega\left(f_k\right) \tag{4.3}$$

Where, $l\left(y_i, \hat{y}_i\right)$ is the loss function that measures the difference between the true value y_i, and the predicted value $\hat{y}_i$. $\Omega\left(f_k\right)$ is the regularization term for the k^{th} base learner. K is the number of base learners.

The regularization term $\Omega\left(f_k\right)$ in XGBoost is commonly defined as:

$$\Omega\left(f_k\right) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \tag{4.4}$$

Where, T is the number of leaf nodes in the tree f_k. w_j is the output value of the j^{th} leaf node. γ and λ are regularization parameters that control the complexity of the individual trees.

For optimizing the objective function, XGBoost uses both first and second-order derivatives (Gradient and Hessian):

$$g_i = \frac{\partial l\left(y_i, \hat{y}_i\right)}{\partial \hat{y}_i}, \quad h_i = \frac{\partial^2 l\left(y_i, \hat{y}_i\right)}{\partial \hat{y}_i^2} \tag{4.5}$$

4.2.4 Seasonal Naive Algorithm

For a given time series $y_1, y_2, \ldots, y_t$ with a seasonal period S, the Seasonal Naive forecast for the next seasonal point $\hat{y}_{t+S}$ is given by:

$$\hat{y}_{t+S} = y_t \tag{4.6}$$

Where, the forecast for the next seasonal point $\hat{y}_{t+S}$ is assumed to be the same as the last observed value y_t from the same season.

The advantage of this hybrid methodology lies in its ability to capture both linear and non-linear patterns, as well as seasonal trends, while also being mindful of computational time and resource requirements.

4.2.5 Performance Metrics

In assessing the performance of our load forecasting models, we employ the same set of metrics outlined in Section3.3. These metrics, including MAE, MSE, MAPE, R^2, and RMSE, provide a comprehensive evaluation of the model's accuracy and reliability, allowing for a direct comparison with other load forecasting models.

4.3 Load Forecast Result

This chapter delves into the results of our predictive models for load demand forecasting. The models operate at three different time intervals: 1-hour and 5-minute forecasts. The 1-hour forecasts aim to provide a near-future outlook for grid management. Similar to the solar PV forecasts, the 5-minute load demand forecasts are divided into short-term forecasts for the immediate next 5 minutes and medium-term forecasts for up to four hours. These varying time intervals serve different operational needs, from real-time adjustments to longer-term planning aligned with battery storage cycles.

4.3.1 One Hour Interval Forecast

This study introduces an ensemble forecasting model that incorporates multiple data preprocessing techniques and considers seasonal factors. Initially, the data undergoes two types of preprocessing: Sparse Normalize and Standardize. Sparse Normalize excels at capturing the sparsity of high-dimensional data, while Standardize eliminates the influence of different scales between features, thereby stabilizing the model. Under the Standardized data, the study employs three models weighted differently for training: Tuned XGBoost[56], Random Forest, and Auto XGBoost[57]. Additionally, XGBoost is applied once more to the Sparse Normalized data. Seasonal Naive methods are also integrated to account for the impact of seasonality and weather on load. Subsequently, we fuse the models by assigning different weights to each. Compared to existing research, this model enhances forecasting accuracy

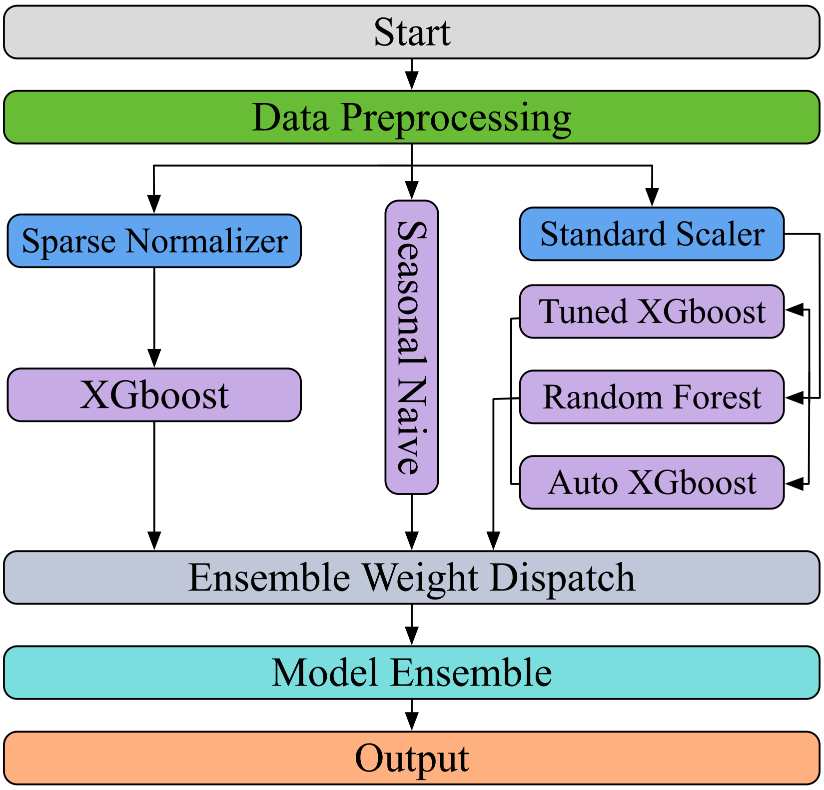

FIGURE 4.3: Technology roadmap of the hybrid model

and robustness through diverse data preprocessing and ensemble strategies. The technology roadmap is shown in Figure 4.3.

To validate the efficacy of the proposed algorithm, various tests have been carried out based on the Pytorch operating environment using Azure Linux V139, 06 system with a Standard-EDS-V4 virtual machine size (4 cores, 32 GB RAM, 150 GB disk) and the processing unit is GPU - 1 x NVIDIA Tesla K80 [58].

The experimental data is sourced from actual load measurements in a city in Northern China, spanning from January to November 2020, collected at hourly intervals. Before the data preprocessing process, missing values were imputed using the Radial Basis Function method. Additionally, an aggregate feature importance analysis was conducted. The specific results of this analysis are illustrated in the Figure 4.4. The features on the X-axis stand for the top-k important features that impact the predictions overall model predictions (a.k.a. global explanation).

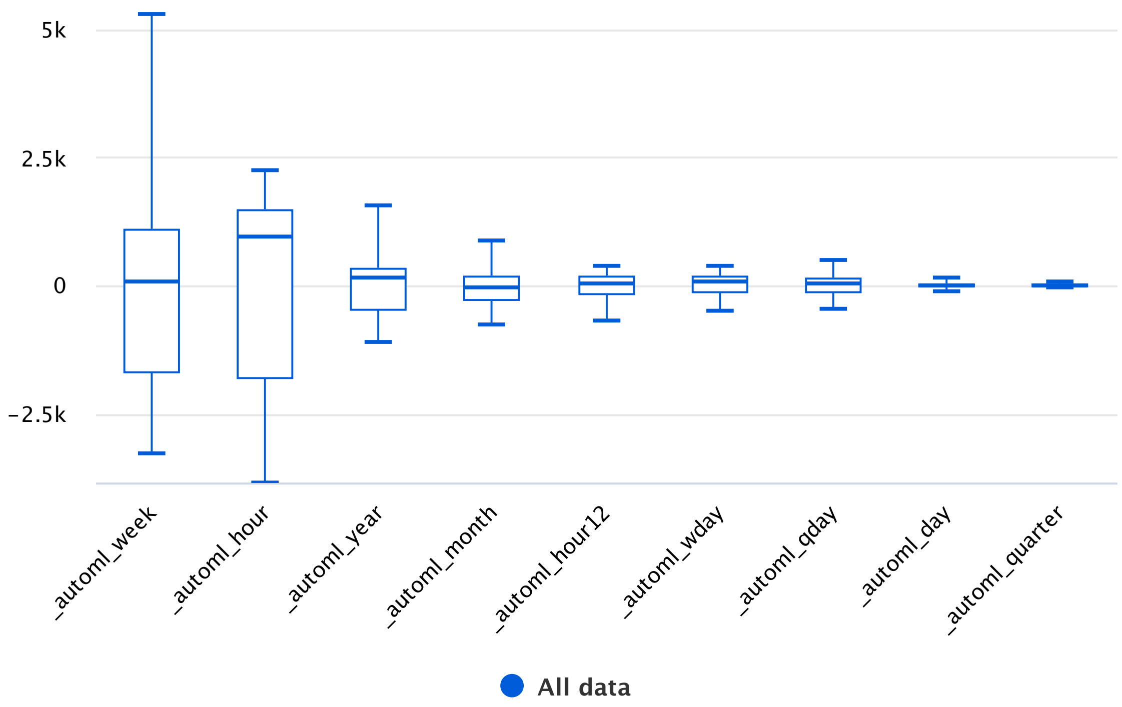

FIGURE 4.4: The result of aggregate feature importance analysis

4.3.1.1 Data Preprocessing

In the Data Preprocessing section, various techniques are employed to enhance the quality and interpretability of the dataset. To better capture the relationships within the data, the dataset is duplicated for multiple analyses. The load column is treated with an Imputation Marker, while the time column undergoes transformation through a Date Time Transformer. Subsequent steps include Short Series Handling, Frequency Detection, Data Size Validation for the TCNForecaster model, Non-stationarity Time Series Detection and Handling, and Missing Feature Values Imputation. For model validation, a K-fold cross-validation method with five folds is utilized.

After these preprocessing steps and data duplication, the dataset is subjected to three different treatments: Standardized Scaler, Sparse Normalized and Seasonal Naive model. These processed datasets are then used for training using Random Forest, XGBoost, and the Naive model, thereby ensuring a comprehensive and robust approach to load forecasting.

4.3.1.2 Hybrid Model Forecast Result

As depicted in Figure 4.3, the study employs XGBoost for training the data after Sparse Normalization. The gbtree is utilized as the booster, and multiple hyperparameters such as colsample-bylevel and eta are set. Regularization parameters are also incorporated into the model. The inclusion of regularization parameters serves to mitigate the risk of overfitting, thereby enhancing the model's generalization capabilities.

Under the Standardized Scaler, the model is divided into Random Forest, Auto XGBoost, and Tuned XGBoost algorithms. For the Random Forest model, the minimum number of samples required at a leaf node is set to 0.0284, the minimum number of samples required to split an internal node is 0.000899, and the number of trees (n-estimators) is set to 10 with bootstrap sampling enabled. Auto XGBoost is configured with multiple hyperparameters such as colsample-bylevel, colsample-bytree, eta, and grow-policy. These settings aim to enhance the model's robustness, mitigate overfitting, and optimize performance. Auto XGBoost serves as an additional reference for tuned XGBoost, and its results are weighted differently to improve the overall model accuracy. Additionally, the Seasonal Naive model is employed to account for the temporal and cyclical nature of the load.

In hybrid model weighting within the ensemble, the Tuned XGBoost under Standardized Scaler has a voting ensemble weight of 0.1333. The Random Forest model with Standardized Scaler is assigned a weight of 0.0667. The XGBoost Regressor with Sparse Normalizer has a weight of 0.3333. The Auto XGBoost with Standardized Scaler has a voting ensemble weight of 0.4, and the Seasonal Naive model carries a weight of 0.0667. The main performance metrics for this hybrid ensemble model configuration are presented in the Table 4.1. Other performance metrics are shown in the Table 4.2.

TABLE 4.1: Other hybrid model performance metrics

Explained variance	Spearman correlation
0.9693863	0.9806957
Median absolute error	normal root mean squared log error
247.5848	0.01950426

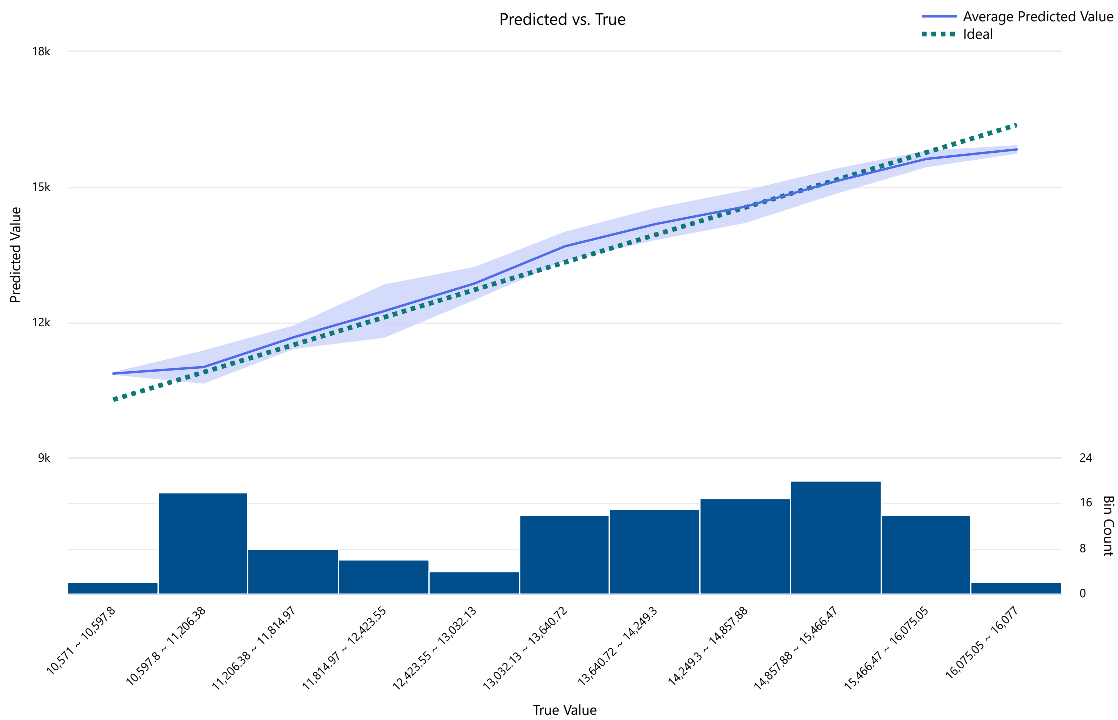

FIGURE 4.5: Overall comparison of predicted and actual load values

Following the presentation of key performance metrics in Table 4.1 and Table 4.2, there is another graphical representation to further substantiate the model's capabilities. The upcoming Figure 4.5 provides a comprehensive comparison between the predicted and actual load values. The x-axis represents the range of actual values, while the left y-axis indicates the predicted values and the right y-axis shows the Bin Count for a subset of actual values. The dashed line represents the ideal prediction, and the blue line illustrates the average predicted value within each range of actual values. This additional visualization serves as a prelude to the more specific comparison provided in Figure 4.6.

To visually complement the performance metrics presented in the Table 3.1, Figure 4.6 provides a side-by-side comparison of the actual and predicted load values for a specific day. The red area represents the predicted interval, the blue line represents the actual value, and the purple line represents the predicted value. This graphical representation further illustrates the

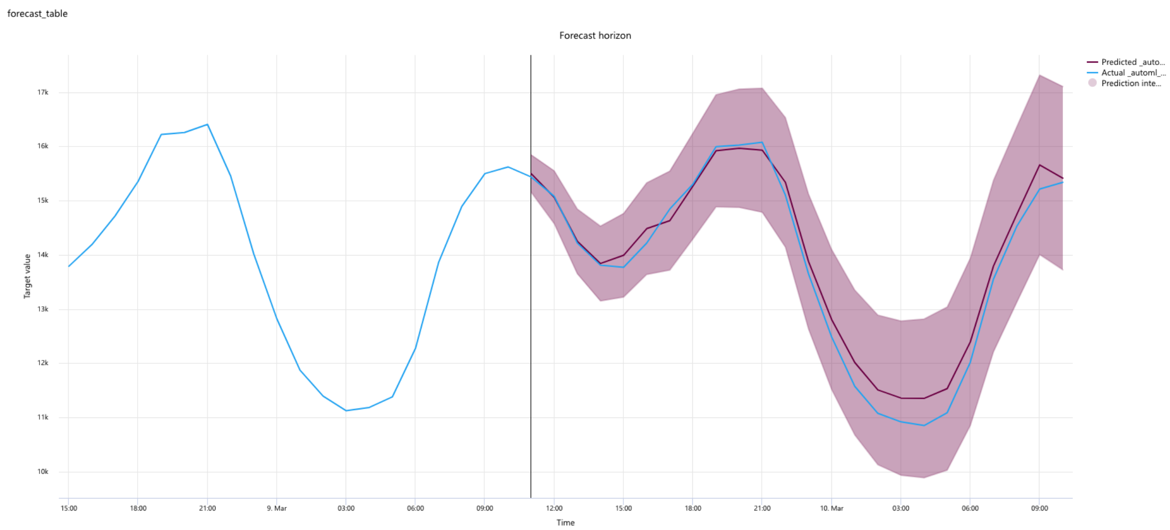

FIGURE 4.6: Comparing true and predicted values on a certain day

TABLE 4.2: Comparison of prediction performance

	MAE	MAPE	R-squared	RMSE	RMSLE
Hybrid model	281.47	2.1230	0.95941	332.70	0.02529
ElasticNet[59]	879.31	6.8055	0.58346	1100.758	0.082687
Random Forest	871.46	6.7542	0.614422	1037.0	0.077708
Seasonal Naive	402.57	2.9109	0.90832	481.40	0.034604
LightGBM[60]	639.07	5.2851	0.79660	767.43	0.059514

model's ability to closely match the real-world data, thereby validating the numerical scores obtained in the performance evaluation.

As the Table 4.2 presents a comparative analysis of the proposed hybrid model against other commonly used models, including LightGBM, standalone Random Forest, standalone XGBoost, and standalone Seasonal Naive. As evidenced by the performance metrics, our hybrid approach consistently outperforms these individual models across various evaluation criteria, further substantiating the efficacy of the ensemble technique employed in this study.

4.3.2 Five Minutes Interval Forecast

Building upon the 1-hour load forecasting models discussed earlier, this section delves into the intricacies of load forecasting at a finer granularity—specifically, at 5-minute intervals.

The computational environment remains consistent with the previous section 3.4.2. The experimental data for this study also originate from a user in Sydney, Australia, but extend over a longer period of 18 months, from January 2022 to June 2023. In addition to the previously mentioned variables—Meter Power (kW), Battery Power (kW), and Solar Power Output (kW)—this dataset also includes temperature and Load (kW) readings. Prior to data preprocessing, Radial Basis Function (RBF) interpolation was employed to align these variables and impute any missing values, ensuring a consistent and robust dataset for our predictive models.

4.3.2.1 Data Preprocessing and Transformer Model Definition

In this study, this thesis employ a Transformer-based deep learning model to predict electrical load. The model takes as input eight different dimensions related to electrical power, such as meter power, battery power, and solar power output. These features are normalized to the [0, 1] range using MinMaxScaler to ensure numerical stability during model training. A sequence length of 24 consecutive time points is chosen to form a data segment, capturing the day-to-day fluctuations and trends in electricity usage effectively.

The core of the model consists of a network structure composed of three layers of Transformer encoders. Unlike traditional time series models, the Transformer naturally handles time dependencies through its self-attention mechanism without relying on complex recurrent structures, significantly enhancing training efficiency. To capture the sequential information inherent in the time series, this thesis also incorporated Positional Encoding, making the model more sensitive to temporal dependencies.

For model optimization, this thesis use the Adam optimizer and Mean Squared Error (MSE) as the loss function. The Adam optimizer allows for adaptive learning rates for different parameters, accelerating model convergence. MSE is a commonly used loss function in regression problems, effectively quantifying the gap between the model predictions and the actual values. The model is trained and validated across 80 epochs, with model evaluation and prediction visualization conducted every five epochs. This cyclical evaluation and visualization aid in timely problem identification and resolution, enhancing the model's

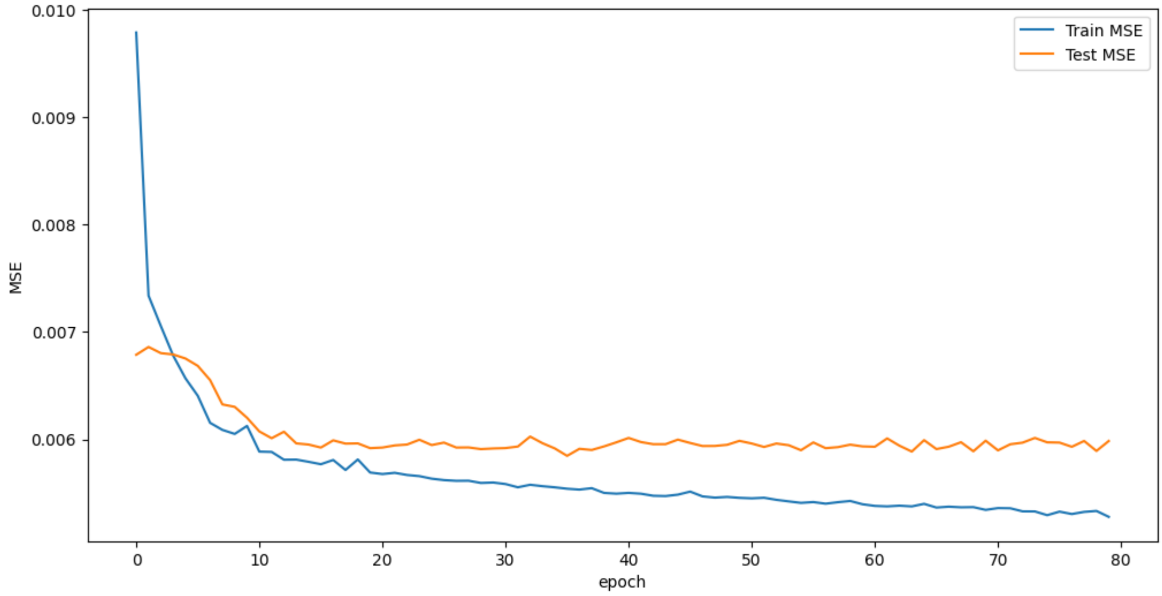

FIGURE 4.7: Transformer train set MSE curve with no positional encoding

predictive accuracy. Finally, the model's predictions are inverse-transformed back to the original range for practical application and performance evaluation.

4.3.2.2 5-Minute for Next 5 Minutes Result

Having elaborated on the data preprocessing steps and the architecture of our Transformer-based model, we now focus on evaluating the model's performance without Positional Encoding. This serves as a baseline to gauge the impact of Positional Encoding on predictive accuracy. In the following sections, figure 4.7 present the change in MSE over the training course of epoch times achieved in this configuration. This will provide valuable insights into how the Transformer algorithm adapts and improves during the training process, even without the added complexity of Positional Encoding.

After examining the MSE changes during the training epochs, we now delve into a more detailed evaluation of the model's performance. In the subsequent section, figure 4.8 present the model's fitting curve over a span of 300 hours on the training set. This will further illuminate the Transformer algorithm's capability to capture the intricate patterns in the electrical load data, serving as a comprehensive assessment of the model's predictive accuracy.

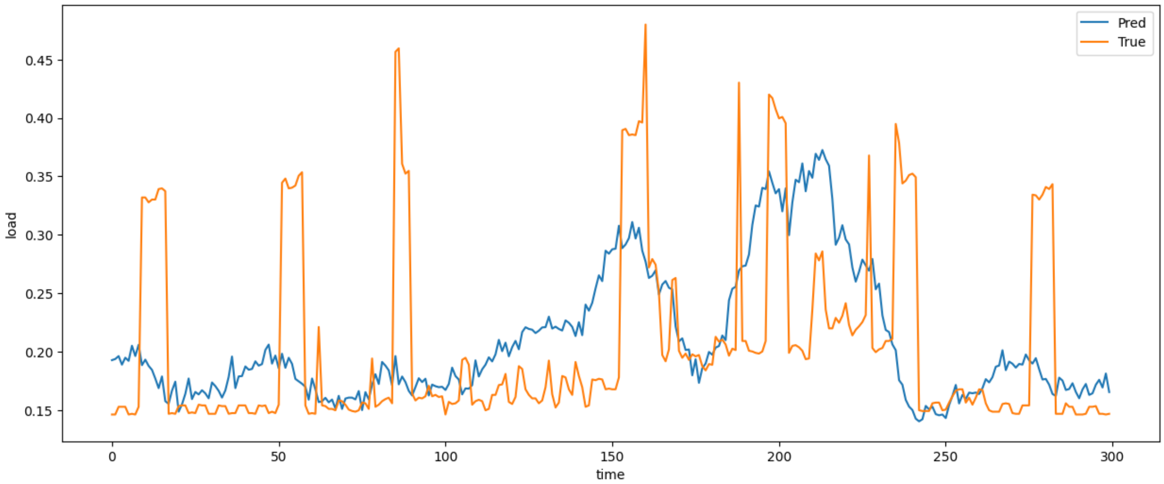

FIGURE 4.8: Transformer model's 300-hour training set fitting curve with no positional encoding

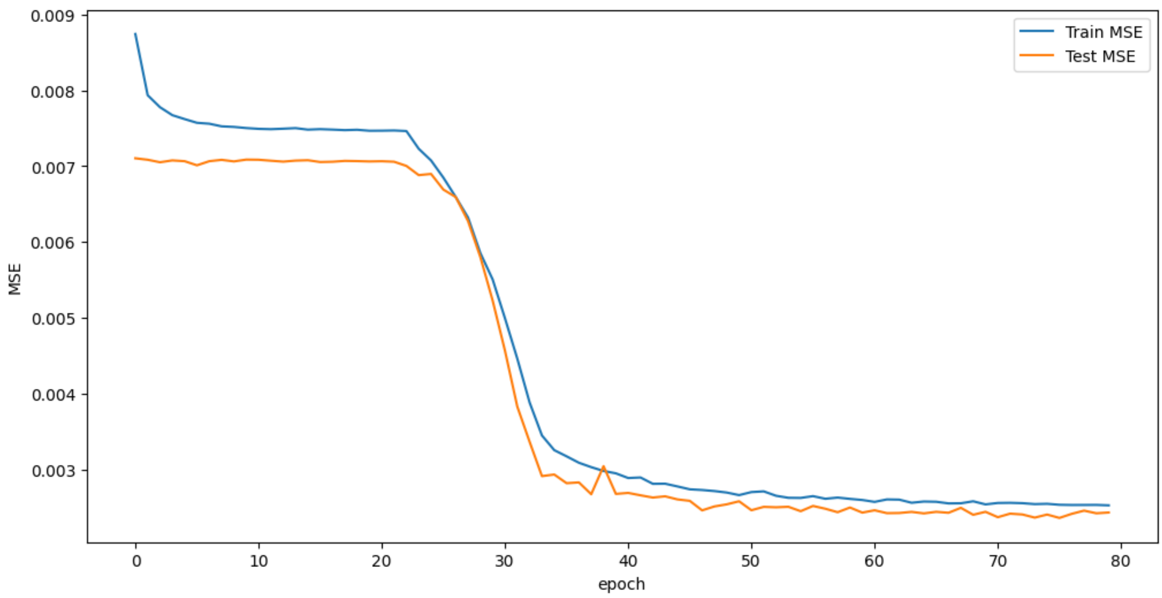

FIGURE 4.9: Transformer train set MSE curve

Having observed the model's performance without Positional Encoding, we now examine the impact of incorporating Positional Encoding within the same section. Next, figure 4.9 present the MSE curve with Positional Encoding across various training epochs. This will offer insights into how Positional Encoding influences the model's training convergence and overall predictive accuracy.

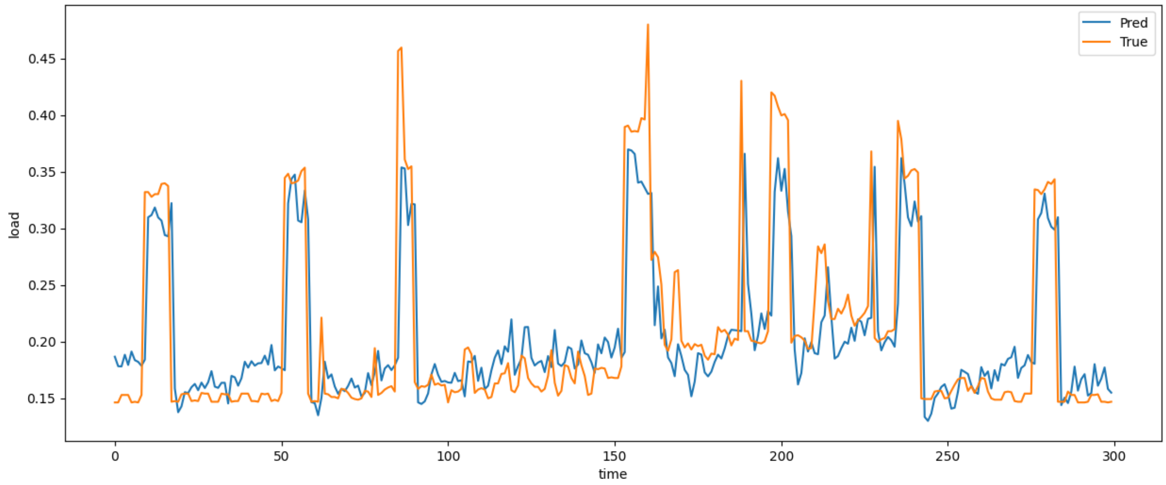

FIGURE 4.10: Transformer model's 300-hour training set fitting curve

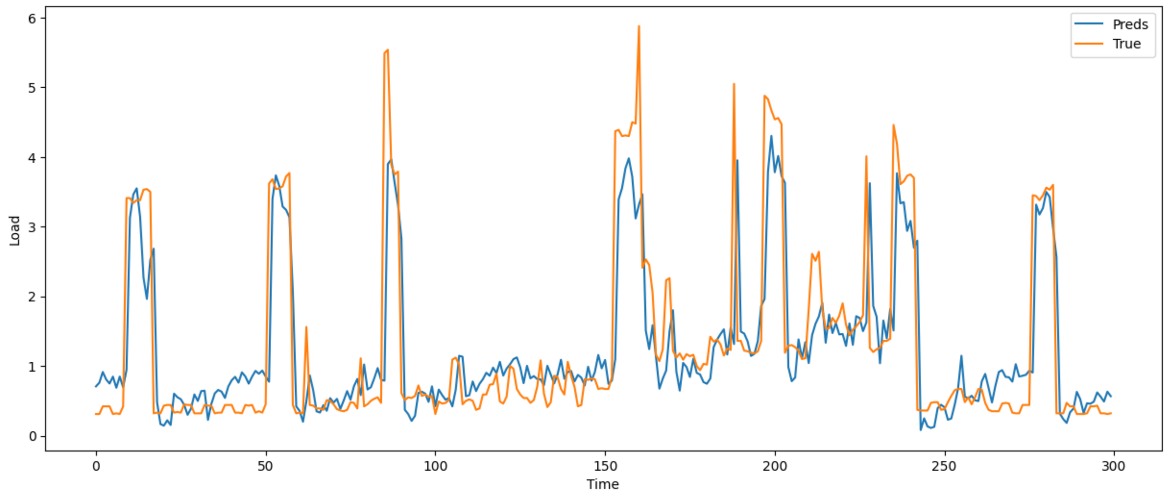

FIGURE 4.11: Original transformer model's 300-hour training set fitting curve

After discussing the MSE curves with and without Positional Encoding, we now delve into a more detailed evaluation of the model's performance. The figure 4.10 presents the model fitting curve for a span of 300 hours on the training set using the Transformer algorithm. This graphical representation will provide a comprehensive view of how well the model captures the underlying patterns in the electrical load data.

Following the normalized model fitting curve, figure 4.11 present the model's performance on the training set after inverse-transforming the predictions back to their original scale.

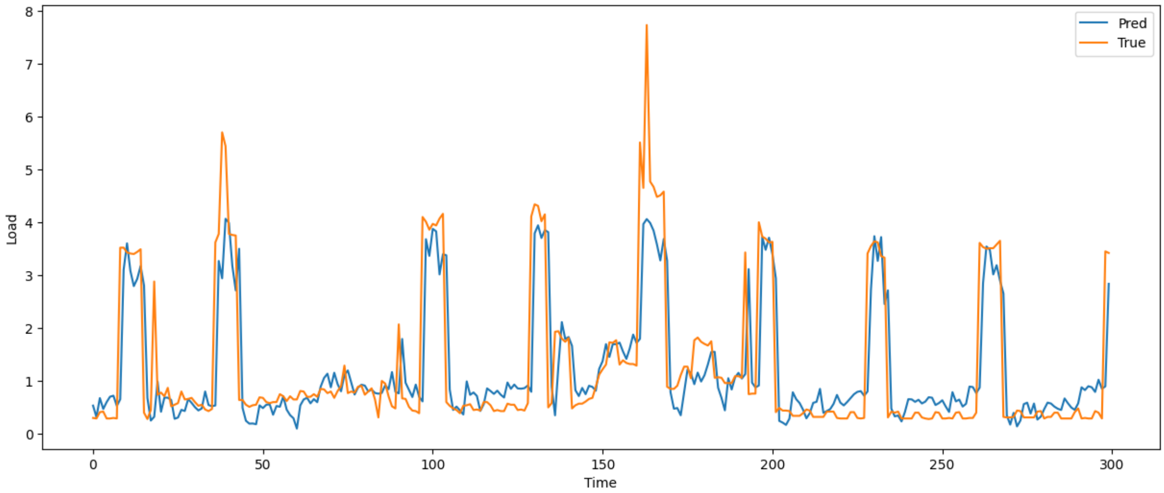

FIGURE 4.12: Original transformer model's 300-hour test set fitting curve

Having examined the model's performance on the training set, we now turn our attention to the testing set, which serves as a more critical evaluation of the model's generalization capabilities. Figure 4.12 display the model's predictive accuracy on the testing set, providing a more robust assessment of its effectiveness in capturing the underlying patterns in the electrical load data.

To further illustrate the model's predictive capabilities, we present a comparison between the actual and predicted electrical load values over a 24-hour period in the training set. The figure 4.13 aims to provide a detailed view of how well the model captures the daily fluctuations and patterns in electrical load.

4.3.2.3 5-Minute for Next 4 Hours Result

Having discussed the model's performance in predicting 5-minute intervals for immediate future load, we now turn our attention to a longer forecasting horizon—specifically, 4-hour forecasts at 5-minute intervals. The rationale for this extended forecasting window is twofold. First, a 4-hour forecast aligns well with the typical charge and discharge cycles of household battery storage systems. By knowing the expected load 4 hours in advance, the system can make more informed decisions about when to charge or discharge the battery, thereby optimizing its usage. Second, a longer forecast horizon provides retail energy providers

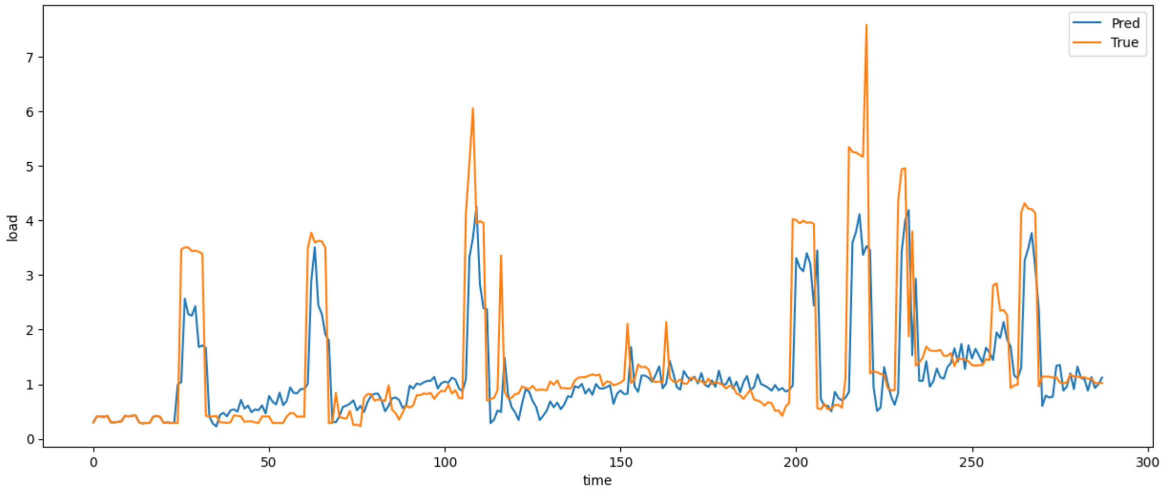

FIGURE 4.13: Original transformer model's single day test set fitting curve

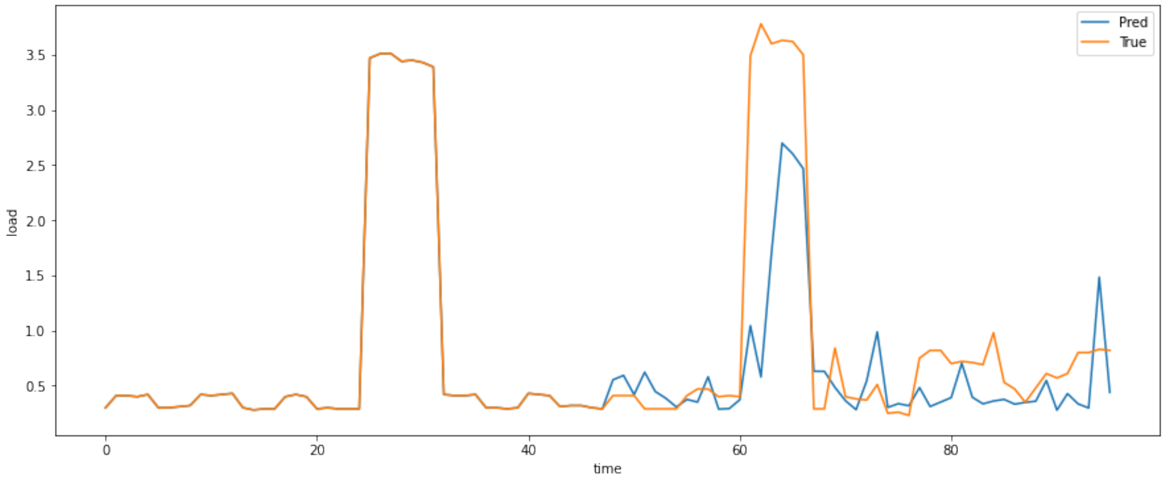

FIGURE 4.14: Original transformer model's 4 hours test set fitting curve

with a broader view of expected electricity consumption, allowing for more strategic grid management and operational planning.

Maintaining the same model parameters as in the previous sections, the following figure 4.14 illustrates a continuous prediction of 48 steps at 5-minute intervals, amounting to a 4-hour forecast. While it's evident from the graph that the predictive accuracy diminishes over the extended time horizon, the model still successfully captures the general trend of the user's electrical load. This provides valuable insights for both battery optimization and broader grid management strategies.

4.4 Chapter Summary

This chapter provides a comprehensive exploration of electrical load forecasting, presenting models tailored for different time intervals and operational scenarios. The first section of the chapter focuses on 1-hour interval load forecasting, offering a robust model that incorporates weather variables. This model serves as a valuable tool for short-term grid decision-making, allowing for more efficient resource allocation and operational stability.

The second section of the chapter delves into the nuances of 5-minute interval load forecasting. This part is further divided into two subsections: one for immediate next 5-minute forecasts and another for up-to-4-hour forecasts. The immediate 5-minute forecasts are essential for real-time grid management, enabling quick adjustments to meet immediate demand and to optimize battery storage systems. The 4-hour forecasts, on the other hand, align with the typical charge and discharge cycles of battery storage systems, providing a longer-term outlook that facilitates more strategic decision-making for battery scheduling. By employing these models, retailers gain a multi-faceted understanding of user load patterns, from short-term fluctuations to longer-term trends. This enables them to make more informed and timely decisions, particularly in optimizing battery charging and discharging cycles. The ability to predict user behavior with high accuracy enhances the overall stability and efficiency of the electrical grid, making these forecasting models indispensable tools for modern grid management.

CHAPTER 5

Electricity Price Forecast

5.1 Introduction to Electricity Price Forecasting in VPPs

Electricity price forecasting is a cornerstone in the efficient operation of VPPs. Accurate and timely price forecasts are indispensable for both short-term and medium-term operational planning within a VPP.

Firstly, accurate price forecasts are essential for real-time market participation. They enable the VPP to make informed decisions on when to buy or sell electricity, thereby maximizing profitability. Secondly, price forecasts are crucial for strategic planning, helping the VPP to negotiate contracts and manage risks effectively. Thirdly, precise price forecasts can significantly influence the charge and discharge schedules of battery storage systems, which in turn affects the overall economic performance of the VPP.

Given these engineering challenges, there is a pressing need for advanced electricity price forecasting models that are both accurate and computationally efficient. This section delves into the mathematical models and technical methodologies employed in this research to address these real-world challenges. We focus on 5-minute interval forecasts, employing a range of machine learning and deep learning techniques to achieve high accuracy and reliability.

5.2 Methods

In this chapter on Electricity Price Forecast, the methodological approach shares foundational similarities with the BiLSTM model 3.2.2.1 used for solar PV forecasting. However, it's crucial to note that electricity prices do not exhibit the same strong temporal correlation with time as solar PV output does. Therefore, the focus of this chapter leans heavily towards identifying the most accurate predictive models for electricity prices.

In pursuit of this goal, we have experimented with a diverse array of models, including but not limited to LSTM, CNN-LSTM, Transformer, LSTM with weather factors, and Transformer with weather factors. Each model was rigorously tested to gauge its predictive accuracy, given the volatile nature of electricity prices. Additionally, the data preprocessing techniques employed here diverge from those used in previous chapters. Unlike the radial basis function used for filling missing values in solar PV forecasting, we have adopted alternative methods to handle the potential spikes in electricity prices. This chapter will delve into these various models and preprocessing techniques, aiming to provide a comprehensive understanding of the most effective methods for electricity price forecasting.

5.2.1 Data Preprocessing

In this work, the data pre-processing for electricity price prediction comprises a unique combination of interval partitioning, mean substitution, and Gaussian perturbed mean substitution. The sequence of steps involved in the pre-processing are as follows:

(1) **Interval Partitioning**: The target column, in this case, the 'RRP' (Regional Reference Price), is divided into discrete intervals denoted by $[b_1, b_2, \ldots, b_n]$. Here, each b_j represents the boundary of an interval, with b_1 being the lower boundary of the first interval and b_n the upper boundary of the last interval.

(2) **Mean Substitution**: For each of these intervals, the mean value, μ_j, is calculated using the formula,

$$\mu_j = \frac{b_j + b_{j+1}}{2}$$

In this formula, μ_j represents the mean value of the j^{th} interval, calculated as the average of its lower and upper boundaries, b_j and b_{j+1}.

(3) **Gaussian Perturbed Mean Substitution**: An additional layer of complexity is added by generating a Gaussian random variable, ϵ, which perturbs the mean value of each interval.

$$r'_i = \mu_j + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \sigma)$$

Here, r'_i denotes the perturbed mean value for the i^{th} data point, and ϵ is a Gaussian random variable with mean 0 and standard deviation σ, introducing a stochastic element to the mean substitution.

(4) **Value Replacement**: The original values falling within each interval are then replaced by the perturbed mean value, r'_i. This step involves replacing the actual price values in each interval with their corresponding perturbed mean values, thus smoothing the data and reducing the impact of outliers.

This pre-processing strategy combines determinism with randomness through mean substitution and Gaussian perturbed mean substitution, respectively. This ensures effective data smoothing and outlier handling while preserving the original distributional characteristics of the data.

5.2.2 Performance Metrics

For the evaluation of our models in electricity price forecasting, we adhere to the metric standards set forth in Section 3.3. By employing MAE, MSE, MAPE, R^2, and RMSE, we maintain a uniform evaluation framework that allows for a meaningful comparison with the other price forecast model.

5.3 Price Forecast Forecast Result

In this chapter, we present the outcomes of our predictive models for electricity price forecasting. The models are designed to operate at 5-minute intervals, with a further distinction

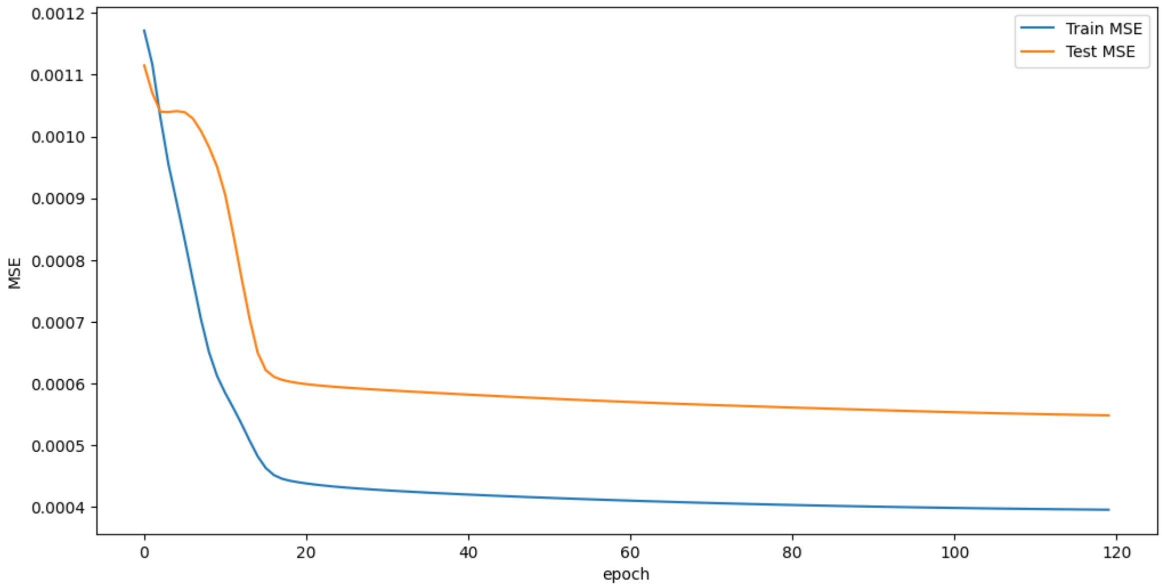

FIGURE 5.1: LSTM price forecast train set MSE curve

between short-term forecasts for the immediate next 5 minutes and medium-term forecasts for up to four hours. The short-term forecasts are crucial for real-time grid management, allowing stakeholders to capitalize on sudden changes in electricity prices. The medium-term forecasts, on the other hand, are designed to facilitate more effective battery scheduling, aligning with the typical four-hour charge and discharge cycles of battery storage systems. n conducting the electricity price forecasting, the data used as input for our models were sourced from the Australian Energy Market Operator (AEMO). Specifically, we utilized the Regional Reference Price (RRP) and total demand data, which are key indicators in the electricity market.

5.3.1 5-Minute to Next 5 Minutes Result

To provide a quantitative measure of their performance, Figure 5.1 present the Mean Squared Error (MSE) variation curves for both the training and test datasets. These curves will offer valuable insights into the models' robustness and predictive accuracy, setting the stage for a more in-depth discussion of the forecast results.

Following the discussion on the MSE variation curves, Figure 5.2 presents a plot comparing the actual and predicted electricity prices over a 300-step period in the training set. This

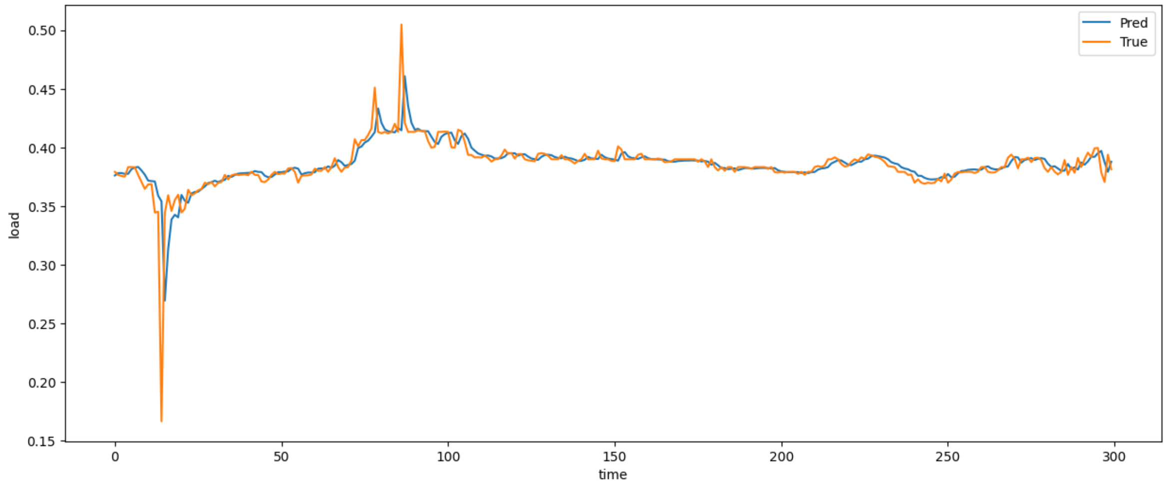

FIGURE 5.2: Comparison of actual and predicted electricity prices of 300-step training set

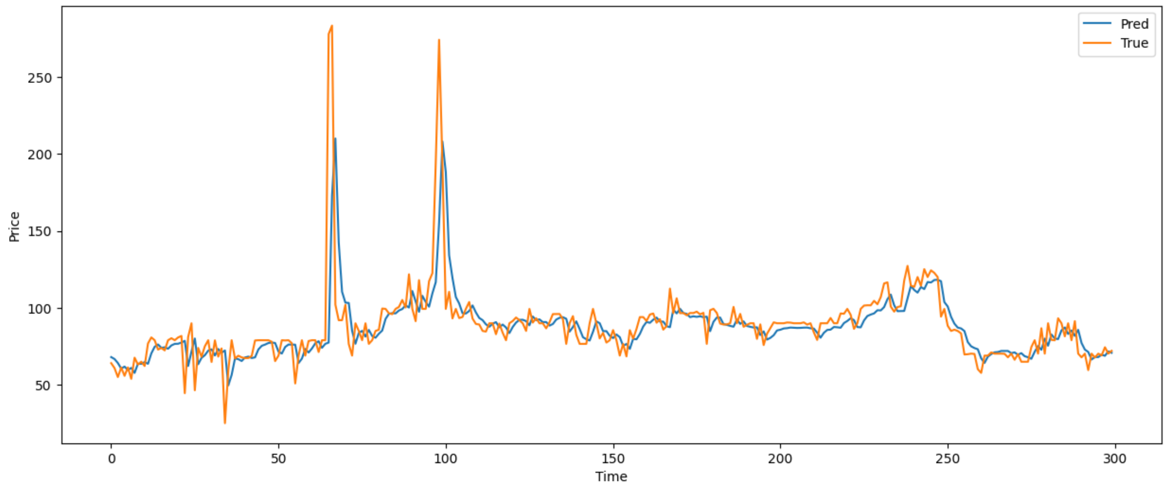

FIGURE 5.3: Comparison of actual and predicted electricity prices of 300-step test set

visual assessment aims to provide a nuanced understanding of the LSTM model's predictive accuracy, specifically its capability to capture the complexities and volatilities in electricity pricing.

Following the evaluation of the LSTM model's performance on the training set, we now turn our attention to its predictive accuracy on the test set. The Figure 5.3 presents a comparison between the actual and predicted electricity prices over a 300-step period in the test set, offering further insights into the model's generalization capabilities.

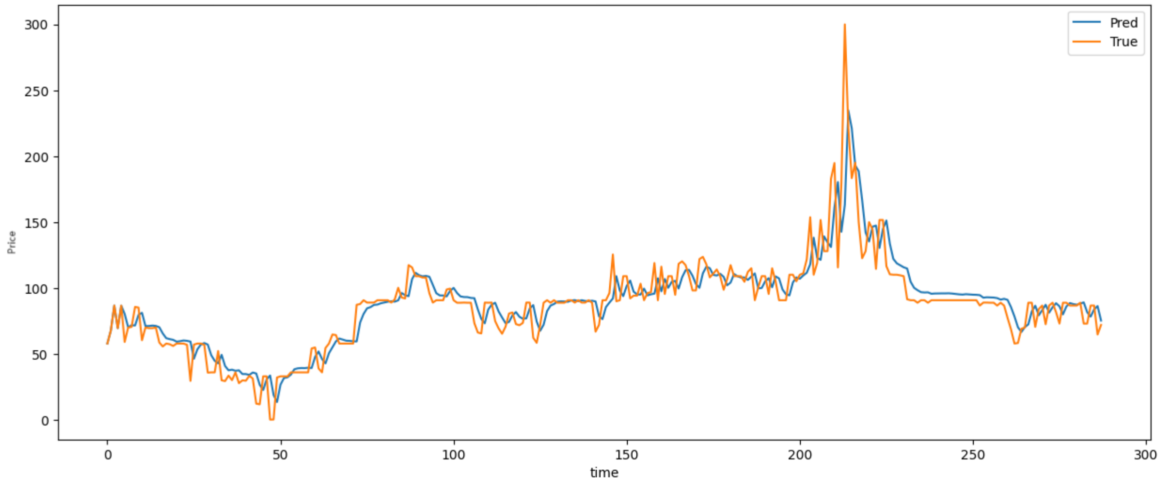

FIGURE 5.4: Comparison of actual and predicted electricity prices of one day

After examining the model's performance over extended periods in both the training and test sets, we narrow our focus to a more granular level. The next figure illustrates the model's forecasting accuracy over a single 24-hour period, providing a detailed view of how well the model captures the intricacies of daily fluctuations in electricity prices.

5.3.2 5-Minute for Next 4 Hours Result

After evaluating the model's performance for immediate 5-minute forecasts, we extend our analysis to a 4-hour forecasting horizon. This extended timeframe is especially relevant for tasks requiring longer-term planning, such as strategic energy trading and battery scheduling. The architecture and parameters of the LSTM model remain consistent with the previous sections to ensure a uniform evaluation metric.

To provide a comprehensive understanding of the model's predictive capabilities over the next 4 hours, Figure 5.5 displays a composite graph consisting of multiple sub-plots. Each sub-plot represents the model's forecast at different 5-minute intervals within the 4-hour window. This layered representation enables us to gauge the model's ability to sustain its predictive accuracy over an extended period, offering a robust evaluation of its suitability for longer-term forecasting tasks.

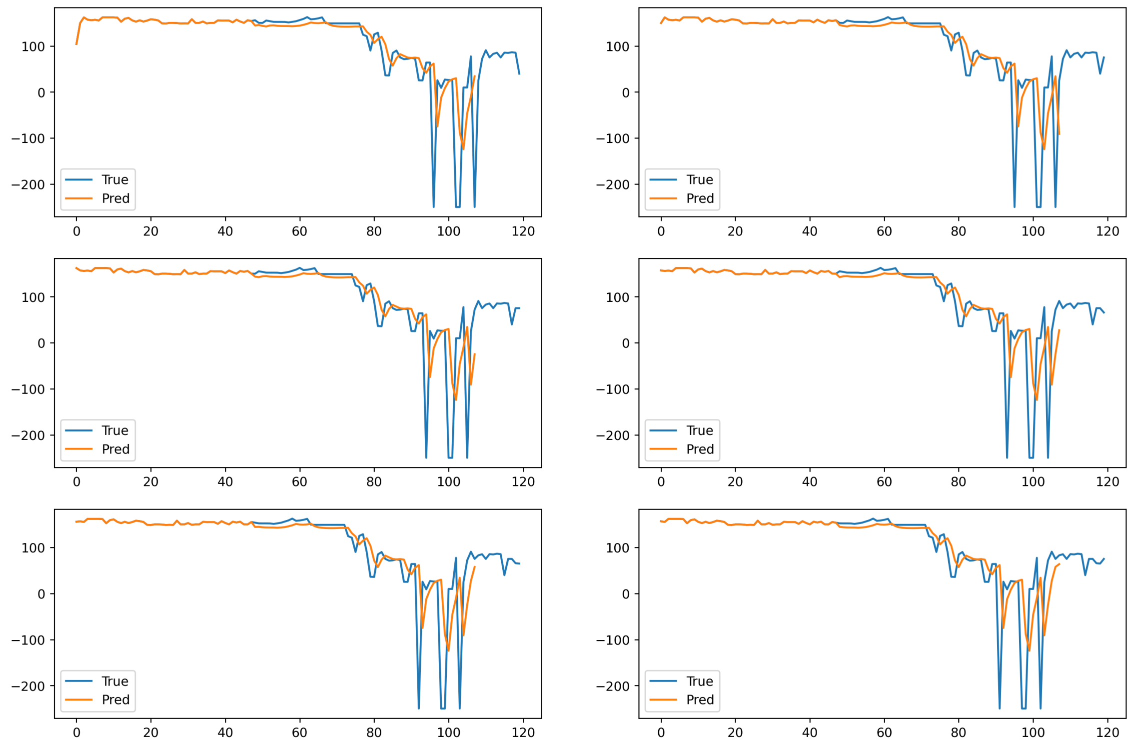

FIGURE 5.5: Composite electricity price forecast of multiple steps within 4 hours

As illustrated by the composite graph, the model maintains a high level of predictive accuracy across the 4-hour forecasting horizon. The tight correspondence between the predicted and actual values in the various sub-plots confirms the model's ability to sustain its performance for longer-term forecasting tasks. This consistency not only substantiates the model's applicability for real-time energy trading but also underscores its potential for more intricate applications, such as long-term battery scheduling and risk management. Overall, the results reinforce the effectiveness of utilizing LSTM-based models for electricity price forecasting at diverse time intervals.

5.4 Model Comparison

The evaluation metrics are crucial for assessing the effectiveness of the forecasting models. As part of our comprehensive analysis, we have included Table 5.1 that showcases the Mean

TABLE 5.1: Performance Metrics and Model Configurations for Various Forecasting Models

	5 Min Interval	MAE	MSE	Model Explaination
1	LSTM	6.7881446	292.10425	Original LSTM
2	Transformer	6.149606	80.72679	Original Transformer
3	LSTM_update	4.9369555	57.514942	Transformer with 18 months update data
4	Transformer_update	5.9143457	73.88758	Transformer with 18 months update data
5	LSTM_TF_5 to 2h	4.4190297	44.135574	LSTM with Teacher Force mechanism. 5 minutes forecast next 2 hours.
6	LSTM_TF_30 to 2h	29.985058	1669.3336	LSTM with Teacher Force mechanism. Merge 5 minutes interval to 30 minutes and then forecast next 2 hours.
7	CNN-LSTM	6.2516087	132.57918	CNN-LSTM model
8	LSTM_dif1	4.8573	58.828514	Put the first difference of the price to the input of the LSTM model
9	LSTM_dif2	4.404904	41.870243	Put the second difference of the price to the input of the LSTM model
10	LSTM_dif1+2	7.169315	285.12384	Put both first and second difference of the price to the input of the LSTM
11	LSTM + weather	5.4660077	55.756172	Put weather feature (temperature and humidity) to the input of LSTM
12	LSTM+ weather_update	5.309131	63.811275	Put 18 months update price and weather data to the LSTM model
13	Trans+weather_update	8.38679	120.45791	Put 18 months update price and weather data to the transformer model
14	PosTrans+weather_update	7.5778155	123.2493	Put 18 months update price and weather data to the positional encoding transformer

Absolute Error (MAE) and Mean Squared Error (MSE) for various model configurations, offering a detailed insight into their relative performance.

In analyzing the table, several key insights emerge that shed light on the effectiveness of different model configurations.

1. Comparing Models 1 and 2 with Models 3 and 4, it becomes evident that increasing the volume of training data significantly improves model accuracy. This aligns with the general principle that more data can lead to better generalization, thereby enhancing predictive performance.

2. The comparison between Models 5 and 6 reveals that aggregating data into larger time intervals (from 5 minutes to 30 minutes) adversely affects model accuracy. This could be due to the loss of finer temporal patterns that are crucial for short-term forecasting.

3. Model 7, which employs a CNN-LSTM architecture, surprisingly underperforms when compared to the original LSTM model (Model 1). This suggests that the convolutional layers may not be capturing the essential features in the data as effectively as the LSTM layers, or that the added complexity of the CNN-LSTM model doesn't offer a commensurate improvement in this specific application.

4. Models 8, 9, and 10 were designed to explore the potential nonlinear relationships in electricity prices by incorporating first and second-order differences. However, the results indicate that these additional features do not significantly improve model performance. This could imply that the electricity prices in the dataset may not have strong nonlinear characteristics that these differences could capture.

In summary, these findings underscore the importance of thoughtful data selection and feature engineering in model performance. While adding more data generally improves accuracy, the choice of model architecture and features requires careful consideration to ensure they are aligned with the underlying patterns in the data. The results also caution against unnecessary model complexity, suggesting that simpler models like LSTM can often yield comparable or even better performance.

5.5 Chapter Summary

In this study, this thesis have meticulously developed an electricity price forecasting model based on Long Short-Term Memory (LSTM) networks. The model underwent a comprehensive process of data preprocessing, feature engineering, and architecture design. It was trained over 120 epochs and evaluated using Mean Squared Error (MSE) as the performance metric.

After rigorous comparisons with various other models, including Transformer models, CNN-LSTM hybrids, and models incorporating weather factors, our LSTM-based model emerged as the most accurate for electricity price forecasting. Unlike other models that may excel in capturing certain aspects of the data, our chosen LSTM model demonstrated a balanced and superior performance in capturing the complex dynamics of electricity prices.

In summary, among the plethora of models tested, our LSTM-based model stands out for its superior predictive accuracy, making it an invaluable tool for electricity price forecasting. This succinct summary serves as a testament to the effectiveness and practical utility of our chosen model.

CHAPTER 6

VPP Integration Program Development

This section provides an in-depth exploration of the existing VPP Optimization Model, a cornerstone in the field of energy management. Understanding its framework, functionalities, and operational considerations is essential for setting the groundwork for the subsequent chapters. Section 6.1 delves into the VPP Optimization Model, elucidating its underlying motivations, functionalities, constraints, and methodologies. Section 6.2, the crux of this research, focuses on the primary contributions of this study. It details the design of the interface that facilitates the integration of high-accuracy forecasting models for solar PV output, load demand, and electricity prices. Furthermore, it elaborates on the two-layer, three-forecast parallel logic design, which serves as the linchpin for enhancing the existing VPP Optimization Model.

6.1 VPP Optimization Model

In today's complex and rapidly evolving electricity markets, battery storage systems have risen to critical importance as essential tools for efficient energy management. The increasing integration of renewable energy sources like wind and solar has brought about challenges, especially their intermittent nature, which leads to high volatility in electricity supply and demand. To manage these challenges, energy systems need to be more agile and responsive than ever. This is further complicated by the growing demands for electrification in various sectors such as transportation and industrial processes.

The model under discussion is not a newly minted academic prototype but a sophisticated system previously developed by a certain company for real-world applications. It serves as a comprehensive, highly configurable platform that goes beyond simple state-of-charge (SOC) monitoring. This model is an intelligent decision-making system that incorporates real-time data on electricity prices, power balance, and battery states to produce actionable insights. Algorithms within the model facilitate rapid and accurate decisions on when and how to charge or discharge batteries to maximize economic benefits. These decisions are also aligned with broader goals such as maintaining grid stability, integrating renewable energy, and reducing carbon emissions.

In the context of this work, the model will be seamlessly integrated with three specialized forecasting models developed in the first three chapters. This integration aims to reinforce the essential role of accurate forecasting in effective optimization. The combination of this advanced battery management system with state-of-the-art forecasting models will offer a comprehensive energy management solution capable of navigating the complexities and challenges of contemporary electricity markets.

6.1.1 Motivation Developing the VPP Program

The impetus for developing this comprehensive battery management model stems from the ever-evolving complexities of today's electricity markets and the intricate operational requirements of modern electrical grids. Traditional battery management systems are often limited in scope, designed for specific use-cases such as peak shaving or load leveling. In contrast, this model serves as a versatile platform, engineered to handle a multitude of scenarios and challenges. It does so by dynamically integrating real-time data on a variety of parameters, including electricity prices, power balances, and battery states. Through the utilization of sophisticated decision-making algorithms, the model is not merely reactive but proactively optimizes operations to maximize economic benefits for electricity retailers and Virtual Power Plant operators.

One of the model's most compelling features is its adaptability, making it a particularly valuable asset in our current volatile energy landscape, characterized by rapid fluctuations in electricity prices and ever-changing grid conditions. Economic viability is a central consideration in the model's architecture, intended to not just enhance profitability but also contribute to overall system reliability and efficiency.

Furthermore, the model is designed to synergize with predictive models developed in earlier chapters of this work, enabling a more cohesive and data-driven approach to energy management. This integration allows the system to leverage predictive analytics for better forecasting accuracy, thus informing more strategic and effective decision-making. The integration of this optimization model with predictive models from earlier chapters aims to offer a more nuanced approach to battery management. By combining these elements, the system seeks to improve forecast accuracy and make more informed decisions, contributing to ongoing efforts in the field to enhance the efficiency and reliability of energy systems.

6.1.2 VPP Model Function

The VPP model is an integrated platform that brings together a set of key functionalities aimed at optimizing battery management and market interactions. It encompasses the dynamic balance of energy supply and demand, economic optimization based on real-time electricity prices, meticulous State of Charge (SOC) management, real-time bidding capabilities, and a remarkable level of adaptability for handling various operational scenarios. Each of these functionalities serves a specific yet interconnected role, contributing to the model's overall effectiveness in a volatile energy landscape.

(1) Energy Balance: The model uses advanced algorithms to dynamically manage the charging and discharging cycles of the battery storage system. This ensures that the energy consumed is in equilibrium with the energy stored, a crucial feature for stabilizing the grid during periods of high volatility in supply or demand.

(2) Economic Optimization: One of the standout features of this model is its ability to interact with real-time electricity prices. Leveraging price forecasting models, the system can make strategic decisions about when to store energy during low-price periods and when to release it during high-price intervals. This real-time economic optimization not only maximizes profits but also contributes to grid stability

(3) State of Charge Management: The model continuously monitors the state of charge of the battery. This ensures that the battery operates within predetermined limits, enhancing its lifespan and efficiency. SOC management is particularly important for preventing scenarios like overcharging or undercharging that could impair battery performance.

(4) Real-Time Bidding: This feature allows for immediate interactions with electricity markets. The model can automatically submit bids for buying or selling electricity based on its current state and market conditions. This dynamic bidding capability makes the system highly adaptable to market fluctuations.

(5) Adaptability: What sets this model apart is its built-in flexibility to adapt to a multitude of operational scenarios such as peak shaving and load leveling. Depending on the current conditions, it can automatically adjust its operational constraints and objectives, making it a highly versatile tool for energy management.

In summary, this VPP model is a comprehensive solution for managing the complexities of modern electricity markets, capable of energy balancing, economic optimization, SOC management, real-time bidding, and adaptability.

6.1.3 Model Constraints

Building on the functionalities outlined earlier, the model also incorporates a set of carefully designed considerations or constraints that govern its operation. These are not just mere guidelines but pivotal elements encoded into the optimization algorithms to ensure the system's efficiency, safety, and economic viability. Here are the key considerations:

(1) SOC Constraints: The State of Charge (SOC) of the battery is a critical parameter for its efficient and safe operation. The model imposes both lower and upper SOC limits to prevent scenarios like deep discharge or overcharging, which could impair battery performance and lifespan. These limits are carefully integrated into the decision-making algorithms to ensure that the system operates within these bounds.

(2) Power Balance: This model adheres to the principle that energy input must equal energy output in real-time. By continuously monitoring the power flow, it intelligently adjusts the battery's charging and discharging states to maintain this balance. This is particularly significant for ensuring grid stability, especially during periods where there is high volatility in energy supply or demand.

(3) Economic Limits: While the model aims to make economically optimal decisions, it also operates within predefined economic boundaries. These could be set as certain price thresholds beyond which the system will not buy or sell electricity. By doing so, it ensures that the operations are not just economically beneficial but also sustainable in the long run.

(4) Inverter Limits: The model also incorporates constraints related to the inverter power limits. The inverter plays a critical role in converting DC power from the battery to AC power for the grid and vice versa. Hence, understanding its operational limits is crucial for the efficient functioning of the entire system. These constraints are integrated into the optimization algorithm to ensure that the inverter operates within its capacity limits, thereby preventing any potential system failures or inefficiencies.

6.1.4 Optimization Methods

Transitioning from the detailed functionalities and considerations, it's imperative to discuss the underlying methods and algorithms that empower the model to meet its objectives. These computational elements are pivotal in enabling the system to handle the complexities of real-time electricity markets and grid conditions. Specifically, the model leans on Generalized Disjunctive Programming (GDP) and a set of optimized decision-making algorithms.

The first foundational method utilized in the model is Generalized Disjunctive Programming. This mathematical programming approach is particularly effective for handling the complex optimization challenges that this model faces. GDP enables the model to manage a variety of operational scenarios, disjunctions, and logical constraints concurrently. It integrates a variety of factors such as real-time electricity prices, battery State of Charge, and power balance requirements into one unified optimization problem. This approach allows the model to adapt to rapidly changing market conditions and optimize its operational strategies for both economic benefit and system reliability.

The second pillar of the model's computational prowess is a set of optimized Decision-making Algorithms. These algorithms are tailored for real-time decision-making, considering multiple constraints and objectives. Factors such as economic thresholds, SOC limits, and power balance requirements are taken into account to make rapid yet accurate decisions. These algorithms empower the system to adapt to fast-changing market conditions and fluctuating grid states, thereby maximizing both economic gains and operational efficiency.

6.2 Forecast Program Integration

This study is predicated on the existing VPP battery optimization framework already implemented within the company. The principal aim is not to construct a new optimization model, but rather to augment the existing system through the integration of sophisticated forecasting algorithms for solar PV output, load demand, and electricity prices. This integration, or "merge", aims to significantly enhance the efficiency, stability, and economic performance of the VPP's battery optimization operations. By incorporating these accurate forecasts into the existing framework, this thsis aim to create a more robust and adaptive VPP battery optimization program capable of responding to real-time fluctuations and challenges in the energy sector.

The preceding chapters have laid the groundwork by individually addressing the forecasting algorithms for solar PV output, load demand, and electricity prices. While each of these components is critical in its own right, the true power of these forecasting models is realized when they are integrated into a cohesive system for VPP battery optimization.

After delving into the intricacies of load forecasting, solar PV forecasting, and electricity price forecasting, this chapter shifts its focus to the practical applications of these predictive models. Specifically, this thesis explore their critical role in optimizing battery scheduling within a VPP. The accurate and timely forecasts generated by these models serve as the backbone for efficient energy management, enabling the VPP to make informed decisions for real-time grid management and battery scheduling. This, in turn, contributes to grid stability, economic efficiency, and sustainable energy utilization. The models and strategies discussed in this chapter have been successfully applied in a production simulation for a retail electricity company, underscoring their practical relevance and effectiveness in a real-world setting. Additionally, this chapter will provide a detailed explanation of the operational logic behind the optimization and scheduling model.

this study introduces a two-layered "merge" model designed to integrate the forecasting results for load, solar PV, and electricity price with a VPP battery optimization scheduling program. Figure 6.1 is the VPP double-layer three-prediction parallel logic figure.

The first layer is the Model Training Layer, where new data for solar PV, load, and electricity price are fed into their respective forecasting models every five minutes for the Xth round of model training. It's important to note that due to the large volume of data and the complexity of the models, the training process may not complete within this 5-minute window. However, to avoid conflicts between data reading and writing, the training layer will not read new data until the Xth model training is complete, thereby ensuring data timeliness.

Upon the completion of each training cycle, a new set of training weights is generated and stored in a weight file. This weight file becomes a critical component for the second layer, known as the Optimization Layer. This layer simply reads the weight file to quickly generate forecast results without additional training. In essence, the same input data fed into the Optimization Layer yields rapid forecast outputs, which are then integrated into the VPP Battery Optimization Program for subsequent battery scheduling tasks.

A noteworthy feature is the Data Anti-collision Module, represented as the yellow module. This module is designed to prevent any interference or read-write conflicts between the Model Training Layer and the Optimization Layer. Its primary function is to ensure that the weight file from the upper training layer is not being written to during the 5-minute read cycle of the lower layer, thereby avoiding errors. Given that the Optimization Layer generates forecast results quickly, it can ensure a new forecast result every five minutes. Similarly, the VPP

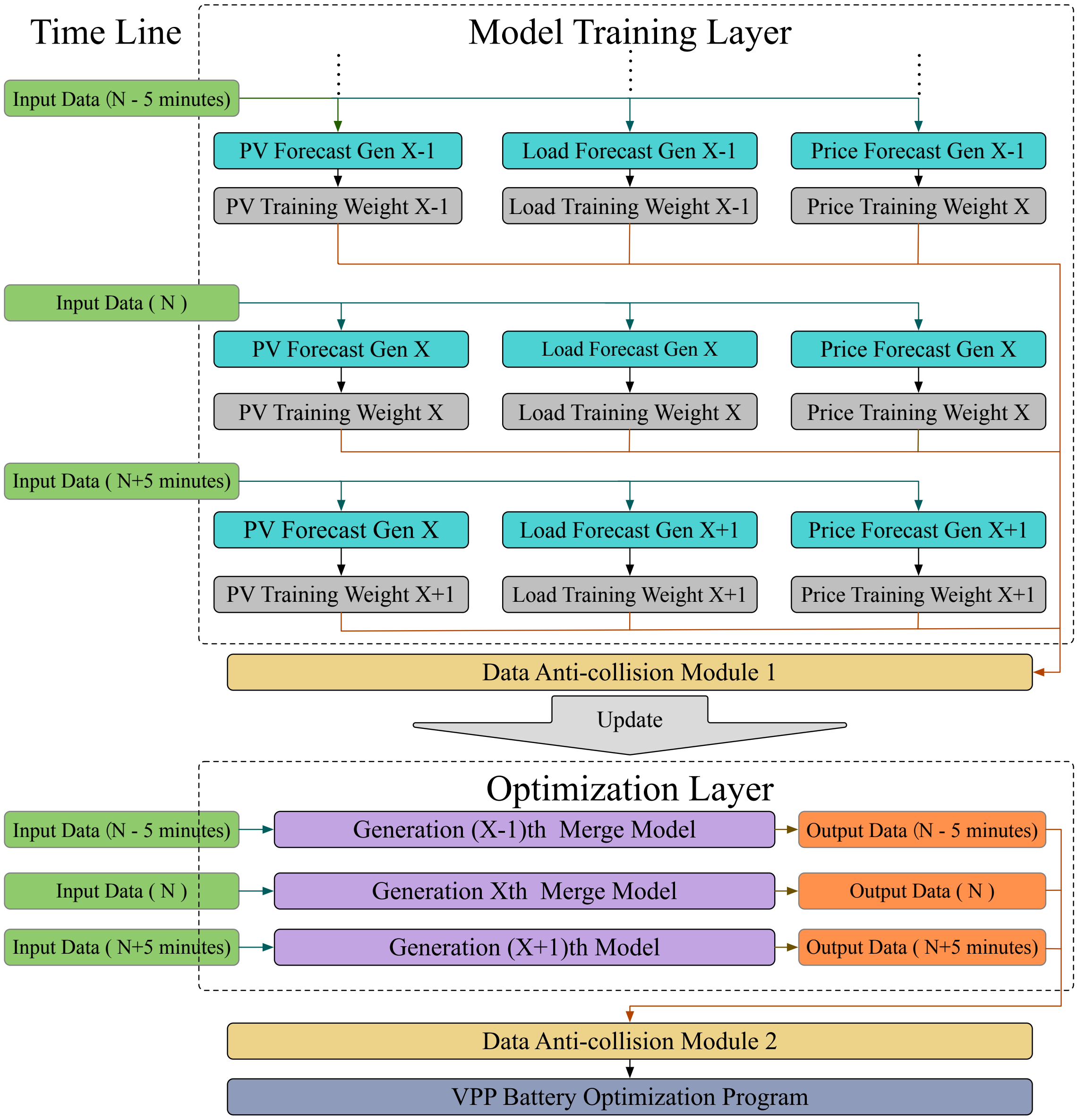

FIGURE 6.1: VPP double-layer three-prediction parallel logic figure

Battery Optimization Program operates at a high speed, allowing it to send a new battery scheduling command every five minutes. This anti-collision module serves to synchronize these operations, ensuring seamless and error-free functioning.

6.2.1 VPP-Prediction Models Interface Design

Building on the insights gained from the previous section and the existing VPP battery optimization program, this subsection transitions into a detailed exposition of the interface design between the VPP and the predictive models for solar PV output, load demand, and electricity prices. The primary objective is to detail how these models are seamlessly integrated into the existing VPP system. This integration aims to substantially improve the operational efficiency, stability, and economic performance of the VPP's battery optimization. Through the interface design, we aspire to create a VPP battery optimization program that is not only robust but also highly adaptive to the real-time challenges posed by modern electricity markets.

In this simulation approach, due to engineering constraints that make immediate data acquisition infeasible, the method illustrated in Figure 6.2 is employed. The length of available historical data for each user's solar PV generation and load demand varies. The starting point for the simulation is set at the most recent common timestamp where all three types of historical data—solar PV, load, and price—are available. The furthest point is based on the most recent year's worth of overlapping data for all three variables. For instance, if the data ranges from January 1, 2022, to January 1, 2023, the simulation starts from the latest point where all three types of data are available, even if more recent price data exists beyond January 1, 2023.

To mimic the real-world scenario of rolling updates, the one-year dataset is divided into two segments. The first 90% remains static and serves as the training set, marked as the "Simulation Start Point" in Figure 6.2. The remaining 10% is gradually introduced into the model at five-minute intervals until the "Simulation End Point" is reached. This approach effectively simulates the rolling update of data in real-time for the training layer and serves as a robust mechanism for evaluating how the system adapts to new incoming data.

In Figure 6.3, this thesis elaborate on the process of simulating the rolling data updates. Given that the Load and PV data are stored in one CSV file while the Price data is in another, the update process is handled in two separate parts. Initially, 90% of the complete historical data

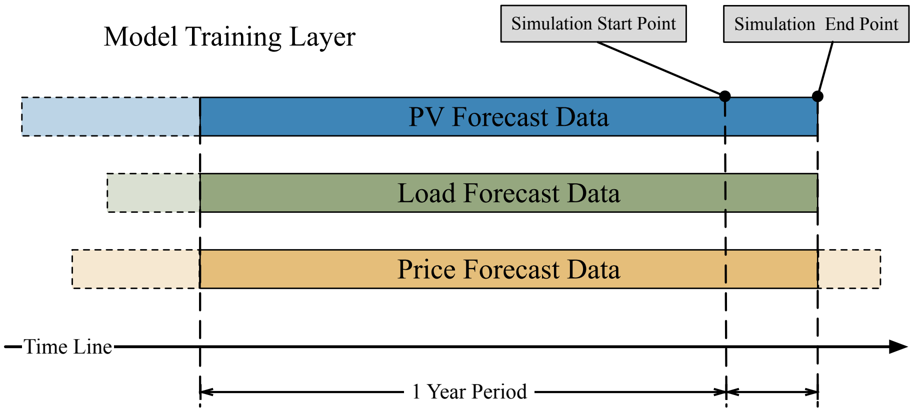

FIGURE 6.2: PV, load and Price training layer historical data simulate rolling design figure

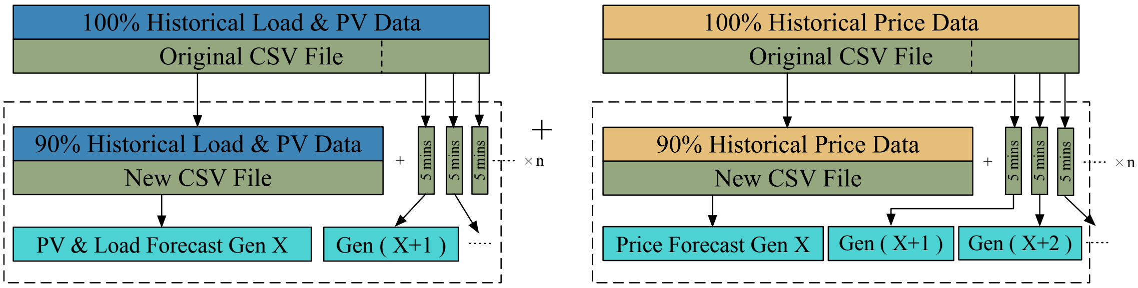

FIGURE 6.3: PV, load and Price training layers simulate rolling training design figure

set is used for the first round of training for the PV, Load, and Price models. Once this initial training is completed, an additional 5-minute data point is appended to the existing 90% data set. This updated set is then used for the second round of training for the PV, Load, and Price models.

This approach effectively mimics the rolling updates of real-world data and allows us to observe the model's performance at various points in time. This, in turn, provides the opportunity for further optimizations and adjustments. The process ensures that the model remains adaptable to incoming data, making it more proficient in battery scheduling and grid management tasks.

In Figure 6.4, the focus is on the interface of the Optimization Layer. As indicated in Figure 6.1, this layer directly reads the model weights generated by the Training Layer. Because of this direct weight integration, the Optimization Layer is capable of swiftly generating

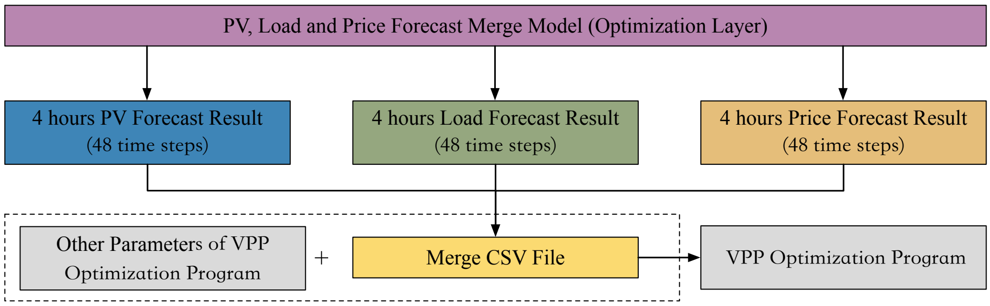

FIGURE 6.4: PV, Load and Price optimization layer interface design figure

forecasts for a 4-hour period, segmented into 48 data points at 5-minute intervals. These 48 forecast points are then consolidated into a new "Merge CSV File". This merged file not only contains these 48 points for PV, Load, and Price forecasts but also includes other parameters necessary for the VPP battery optimization program.

By arranging all the essential data into a single CSV file, the VPP optimization program can easily fetch all it needs for optimized operational commands by reading this one file. Subsequently, a comparison between these optimized commands and actual electricity prices enables the calculation of profit or loss, thereby providing a comprehensive view of the system's economic performance.

In Figures 6.5 and 6.6, the Data Anti-collision Modules 1 and 2 are elaborated, detailing their specific positions and functionalities within the system. In large-scale engineering applications with substantial data flow, simultaneous read and write operations on a single CSV file can result in multiple issues such as data corruption, read-write conflicts, and system inefficiencies. To mitigate these challenges, two Data Anti-collision Modules are introduced.

The working principle of Module 1 is tailored to accommodate the 5-minute data input intervals. The priority here is on writing new data to the "New CSV File". The system ensures that new data is first written into the file before any read operation is performed. This write-before-read operation is fundamental to maintain data integrity and timeliness.

Conversely, the focus of Module 2 is on the Optimization Layer, where the VPP battery optimization program must run every five minutes. Therefore, reading from the "Merge CSV

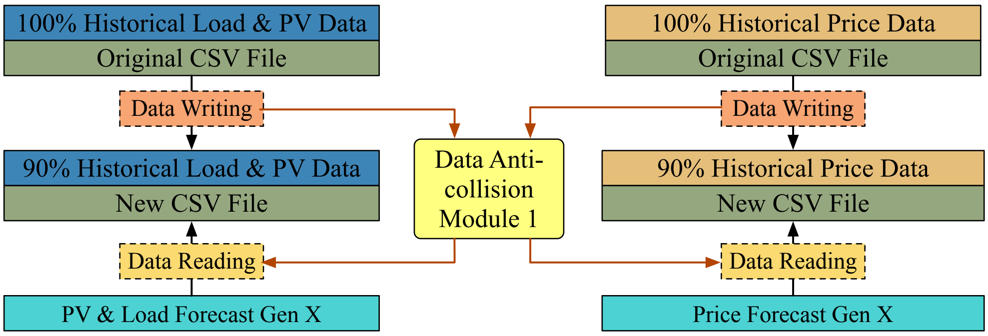

FIGURE 6.5: Data anti-collision module 1 design logic figure

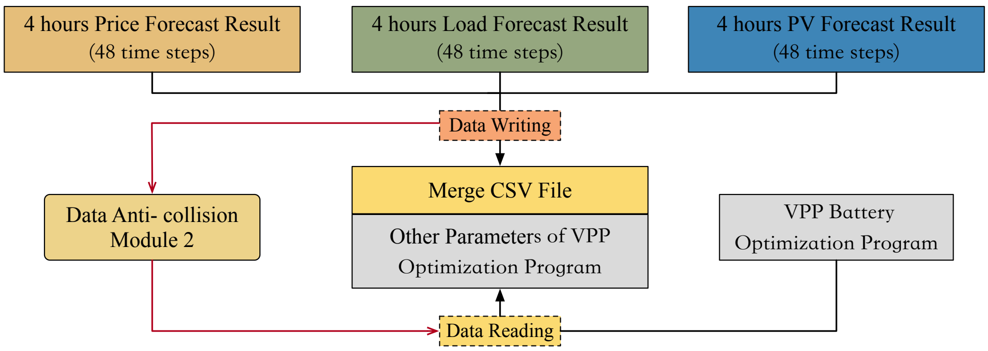

FIGURE 6.6: Data anti-collision module 2 design logic figure

File" takes precedence over writing. Here, the system ensures that the file is first read before any new data is written into it.

To synchronize these modules effectively, the read operation in Module 1 is slightly delayed compared to its write operation. Similarly, the write operation in Module 2 is set to occur a tad later than its read operation. This staggered timing approach ensures seamless, error-free data handling, thus enhancing the overall system reliability.

6.2.2 Algorithmic Steps and Equations

In this subsection, this thesis outline the algorithmic steps and corresponding equations for the two-layer model, which integrates solar PV, load, and electricity price forecasting with VPP battery optimization.

(1) **Initialize Model Training Layer and Optimization Layer:**

Initialize the models for solar PV, load, and electricity price forecasting.

$$\text{Initialize } M_{PV}, M_{Load}, M_{Price} \tag{6.1}$$

(2) **Data Ingestion:**

Ingest new data for solar PV, load, and electricity price into their respective models.

$$D_{PV}, D_{Load}, D_{Price} \leftarrow \text{New Data} \tag{6.2}$$

(3) **Model Training Layer:**

If the X-th training is complete, update the data for the next round of training.

$$\text{if } X\text{-th training complete, then } D_{PV}, D_{Load}, D_{Price} \leftarrow \text{Updated Data} \tag{6.3}$$

Train the models and generate new weight files.

$$W_{PV}, W_{Load}, W_{Price} \leftarrow \text{Train}(M_{PV}, M_{Load}, M_{Price}) \tag{6.4}$$

(4) **Data Anti-collision Module:**

Ensure that the weight files are not being written to while being read by the Optimization Layer.

$$\text{Ensure } W_{PV}, W_{Load}, W_{Price} \text{ not being written during read} \tag{6.5}$$

(5) **Optimization Layer:**

Use the weight files and original input data to quickly generate forecasts.

$$O \leftarrow \text{Forecast}(W_{PV}, W_{Load}, W_{Price}, D_{PV}, D_{Load}, D_{Price}) \tag{6.6}$$

(6) **VPP Battery Optimization Program:**

Use the forecasts to generate battery scheduling commands.

$$S \leftarrow \text{Schedule}(O) \tag{6.7}$$

6.3 Chapter Summary

In this chapter, this thesis initially introduced the existing VPP battery optimization program, detailing its functionalities and operational logic. Following that, the primary focus shifted to the integration of high-accuracy forecasting models for solar PV, load, and electricity price. A significant element in this integration is the VPP-Prediction Models Interface Design. This interface serves a crucial role in enhancing the efficiency, stability, and economic performance of the existing VPP battery optimization program. It is designed to accommodate the forecasted data and blend it seamlessly with other parameters essential for battery optimization within the VPP.

Due to engineering constraints that inhibit the use of real-time data, a simulated real-time data updating mechanism is employed. This simulation is instrumental in ensuring that the VPP program operates effectively and accurately, prerequisites for any further considerations of grid stability and economic benefits. Data Anti-collision Modules are also integrated within each operational layer to prevent read-write conflicts, thereby bolstering the system's reliability.

In summary, the chapter provides a comprehensive view of the complexities involved in integrating forecasting models into an existing VPP battery optimization program. The architecture, interface design, and anti-collision modules are specifically tailored to ensure the VPP program's effective and reliable operation, especially given the constraints of not having access to real-time data.

CHAPTER 7

Conclusion

Building on the insights gained from the preceding chapters, this thesis have a comprehensive understanding of the various aspects of deep learning-based forecasting in real-world VPP applications. Chapter 3 provided an in-depth look at photovoltaic (PV) forecasting, using advanced models like BiLSTM and GBDT-LSTM with Teacher Forcing. Chapter 4 focused on the complexities of load forecasting, offering solutions for short-term predictions. Chapter 5 delved into electricity price forecasting, emphasizing the role of data preprocessing and presenting various forecast results. Chapter 6 is devoted to introducing the VPP battery optimization program and detailing how the three forecasting models are integrated and applied within this program.

7.1 Major Contributions

In the case of solar PV output, the 5-minute forecasts using LSTM algorithms are designed for real-time grid management, allowing for immediate adjustments to solar energy input. The 1-hour forecasts using GBDT-BiLSTM are tailored for day-ahead markets, providing a 24-hour outlook that helps in strategic planning and bidding.

One of the standout innovations in this dissertation is the creation of a novel GBDT-BiLSTM hybrid model specifically designed for 1-hour solar PV output forecasting. This model ingeniously fuses the predictive power of Gradient Boosting Decision Trees (GBDT) with the sequence modeling capabilities of Bidirectional Long Short-Term Memory (BiLSTM) networks. The GBDT component excels at capturing the linear relationships and feature interactions in the data, making it highly effective for trend analysis and feature importance evaluation. On the other hand, the BiLSTM component is adept at modeling temporal sequences and capturing non-linear dependencies, which is crucial for understanding the volatile nature of solar PV output. By integrating these two algorithms, the hybrid model is able to capture a comprehensive range of data patterns, both linear and non-linear, thereby significantly enhancing the accuracy of 1-hour ahead solar PV forecasts. This is particularly beneficial for day-ahead market participation, where accurate 24-hour outlooks are essential for optimal bidding strategies.

For load demand forecasting, this dissertation employs a multi-faceted approach to address the varying needs of real-time and longer-term grid management. For the immediate, short-term 5-minute forecasts, a Transformer model enhanced with positional encoding is utilized. The inclusion of positional encoding allows the model to better understand the sequence of the data, making it highly effective for real-time grid adjustments. This is crucial for dynamically balancing supply and demand, thereby ensuring grid stability and minimizing the need for expensive peak-load energy sources.

On the other hand, for the 1-hour load demand forecasts, an ensemble model combining XGBoost, Random Forest, and Seasonal Naive methods is employed. Each of these algorithms brings its own set of advantages: XGBoost is known for its high predictive accuracy and ability to handle a variety of data types; Random Forest offers robustness and can capture complex interactions between features; and the Seasonal Naive method provides a baseline model that captures the seasonality in the load data. By integrating these methods into an ensemble model, the system gains the ability to make more accurate and robust longer-term forecasts. These 1-hour forecasts are particularly useful for daily operational planning, enabling more efficient resource allocation and reducing operational costs.

For electricity price forecasting, the dissertation adopts a nuanced approach tailored to the specific requirements of real-time and medium-term market participation. The 5-minute forecasts are generated using Long Short-Term Memory (LSTM) algorithms, a type of recurrent neural network that excels in capturing temporal dependencies in time-series data. These forecasts serve dual purposes and are instrumental in multiple aspects of VPP operations. Firstly, the immediate 5-minute forecasts are crucial for real-time market participation. They enable the VPP to make quick adjustments in electricity pricing, thereby optimizing its position in the highly volatile short-term energy markets. This real-time pricing information is invaluable for making instantaneous decisions on energy trading, allowing the VPP to capitalize on short-term price fluctuations and improve profitability. Secondly, the LSTM-generated forecasts also extend to medium-term scenarios, providing insights for periods up to 4 hours into the future. These medium-term forecasts are particularly useful for strategic decision-making, such as determining when to store energy in batteries or when to release it back into the grid. By having a more accurate picture of future electricity prices, the VPP can make more informed decisions, optimizing its market participation strategy and maximizing revenue while minimizing costs. Thus, the 5-minute electricity price forecasts generated by the LSTM algorithms are not only versatile but also pivotal in enhancing the operational efficiency and market competitiveness of the VPP.

The design of the VPP-Prediction Models Interface serves as a significant contribution to this study. It tackles the engineering constraints of not being able to use real-time data by employing a simulated real-time data updating mechanism. This interface not only ensures the seamless integration of the forecasting models into the existing VPP battery optimization program but also establishes a reliable system for the program to operate effectively. The introduction of Data Anti-collision Modules within the layers further strengthens the system's ability to function without read-write conflicts. Thus, the interface design ensures that the VPP battery optimization program can run both effectively and correctly, laying the groundwork for achieving grid stability and economic efficiency.

Each of these models is not just theoretically sound but has been applied in a production simulation for a retail electricity company, demonstrating their practical utility. The choice of these specific algorithms and time intervals is meticulously designed to meet the operational needs of VPPs, thereby facilitating both real-time grid management and effective battery scheduling.

7.2 Model Limitation

While the models developed in this dissertation offer significant advancements in forecasting accuracy and operational efficiency for VPPs, they are not without limitations.

Firstly, the GBDT-BiLSTM model for 1-hour solar PV output forecasting, while innovative, is computationally demanding, which could be a constraint in real-time applications requiring quick decisions. Additionally, the model's accuracy is highly dependent on the quality and completeness of the input data, making it sensitive to missing or erroneous information. These factors necessitate careful data preprocessing and model tuning, adding to the overall complexity of its deployment.

Secondly, the Transformer model with positional encoding employed for 5-minute load demand forecasting is effective but comes with its own set of challenges. Primarily, the model requires a substantial amount of historical data for training to achieve optimal performance. This could pose a limitation in scenarios where data is either sparse, inconsistent, or not well-maintained. Additionally, the model's complexity makes it computationally expensive, which might not be ideal for real-time applications where quick forecasting is crucial. Therefore, while the model is powerful, its deployment requires careful consideration of computational resources and data availability.

Thirdly, the LSTM algorithms used for electricity price forecasting are also computationally demanding, especially when scaling up to handle larger datasets or more complex market dynamics. Additionally, LSTMs are known for their sensitivity to hyperparameter settings, which means that considerable time and expertise are required for model tuning to achieve optimal performance.

Lastly, while the models are designed to be robust and adaptable, they have been tested in specific scenarios and may require further validation to generalize across different operational conditions or geographical locations.

7.3 Future Outlook

The first area of focus is on optimization and data sensitivity. As the landscape of renewable energy and Virtual Power Plants continues to evolve, the forecasting models presented in this dissertation offer a robust foundation for current operational needs and pave the way for future advancements. One immediate area for future work is the optimization of computational efficiency for the GBDT-BiLSTM and Transformer models. Given the real-time requirements of VPPs, reducing the computational burden of these models is crucial. Techniques such as model pruning, quantization, or leveraging specialized hardware accelerators could be explored to speed up both training and inference times. Furthermore, the models' sensitivity to data quality presents another avenue for improvement. Real-time data correction mechanisms or robust outlier detection algorithms could be integrated into the forecasting pipeline to mitigate the impact of missing or erroneous data, enhancing the models' reliability and robustness.

The second area delves into model enrichment and the development of a unified framework. Another promising avenue for future research is the integration of additional data sources to enrich the forecasting models. For instance, incorporating real-time weather forecasts, market trends, or even social factors could significantly enhance the accuracy and reliability of electricity price predictions. Advanced feature engineering techniques could be employed to effectively combine these diverse data sources. Moreover, the development of a unified framework that seamlessly integrates all three key forecasting models—solar PV output, load demand, and electricity prices—would represent a significant leap forward in the field. Such a comprehensive system could facilitate real-time VPP management by providing a one-stop solution for all forecasting needs, streamlining operational workflows and enabling more agile and informed decision-making.

Lastly, the focus shifts to practical applications and industry collaboration. The practical utility of these forecasting models could be further validated through extended real-world applications. Collaborative efforts with energy companies, grid operators, or even regulatory bodies would provide an invaluable platform to truly assess the models' impact on various key metrics such as grid stability, resource allocation, and market efficiency. These real-world trials could also serve as iterative feedback loops for model refinement, allowing for adjustments and optimizations based on actual operational challenges and performance outcomes. Moreover, such collaborations could facilitate the development of industry-specific benchmarks or best practices, thereby contributing to the standardization of forecasting methods in the context of VPPs. By engaging in these practical applications and partnerships, the models can be rigorously tested and fine-tuned, ensuring that they are not just theoretically sound but also practically impactful in enhancing the operations and sustainability of modern energy systems.

Bibliography

[1] IEA, *Net zero by 2050 – analysis*, License: CC BY 4.0, Location: Paris. [Online]. Available: https://www.iea.org/reports/net-zero-by-2050.

[2] B. Goia, T. Cioara and I. Anghel, 'Virtual power plant optimization in smart grids: A narrative review', *Future Internet*, vol. 14, no. 5, p. 128, 2022.

[3] P. Denholm, M. O'Connell, G. Brinkman and J. Jorgenson, 'Overgeneration from solar energy in california. a field guide to the duck chart', National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2015.

[4] S. Srinivasan, S. Kumarasamy, Z. E. Andreadakis and P. G. Lind, 'Artificial intelligence and mathematical models of power grids driven by renewable energy sources: A survey', *Energies*, vol. 16, no. 14, p. 5383, 2023.

[5] V. Bagalini, B. Zhao, R. Wang and U. Desideri, 'Solar pv-battery-electric grid-based energy system for residential applications: System configuration and viability', *Research*, vol. 2019, 2019.

[6] H. Saboori, M. Mohammadi and R. Taghe, 'Virtual power plant (vpp), definition, concept, components and types', in *2011 Asia-Pacific power and energy engineering conference*, IEEE, 2011, pp. 1–4.

[7] S. Souabi, A. Chakir and M. Tabaa, 'Data-driven prediction models of photovoltaic energy for smart grid applications', *Energy Reports*, vol. 9, pp. 90–105, 2023.

[8] 'Method and system for management charge and discharge of electric energy by prediction photovoltaic power generation and load', 2019.

[9] J. Darmstadter, 'The economic and policy setting of renewable energy: Where do things stand?', Tech. Rep., 2003.

[10] S. Ghavidel, L. Li, J. Aghaei, T. Yu and J. Zhu, 'A review on the virtual power plant: Components and operation systems', in *2016 IEEE international conference on power system technology (POWERCON)*, IEEE, 2016, pp. 1–6.

[11] M. Bai, X. Zhao, Z. Long, J. Liu and D. Yu, 'Short-term probabilistic photovoltaic power forecast based on deep convolutional long short-term memory network and kernel density estimation', *arXiv preprint arXiv:2107.01343*, 2021.

[12] N. Rezaei, R. Rajabi and A. Estebsari, 'Electricity price forecasting model based on gated recurrent units', in *2022 IEEE International Conference on Environment and Electrical Engineering and 2022 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, IEEE, 2022, pp. 1–5.

[13] M. Tan, Z. Dai, Y. Su, C. Chen, L. Wang and J. Chen, 'Bi-level optimization of charging scheduling of a battery swap station based on deep reinforcement learning', *Engineering Applications of Artificial Intelligence*, vol. 118, p. 105 557, 2023.

[14] M. S. Islam, B. Omidi, I. Alouani and K. N. Khasawneh, 'Vpp: Privacy preserving machine learning via undervolting', in *2023 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*, IEEE, 2023, pp. 315–325.

[15] A. K. Podder, S. Islam, N. M. Kumar *et al.*, 'Systematic categorization of optimization strategies for virtual power plants', *Energies*, vol. 13, no. 23, p. 6251, 2020.

[16] X. Kong, W. Lu, J. Wu *et al.*, 'Real-time pricing method for vpp demand response based on per-ddpg algorithm', *Energy*, vol. 271, p. 127 036, 2023.

[17] S. Riaz and P. Mancarella, 'On feasibility and flexibility operating regions of virtual power plants and tso/dso interfaces', in *2019 IEEE Milan PowerTech*, IEEE, 2019, pp. 1–6.

[18] S. Suo, X. Che, A. Xu *et al.*, 'Design scheme of distributed energy resources interconnection oriented the vpp system', in *Journal of Physics: Conference Series*, IOP Publishing, vol. 2527, 2023, p. 012 004.

[19] Y. Lee, 'Energy management system with power offering strategy for a microgrid integrated vpp', *be published, Computers, Materials and Continua*, 2022.

[20] R. S. Ravi, A. Jolfaei, D. Tripathy and M. Ali, 'Secured energy ecosystems under distributed energy resources penetration', *Internet of Things and Cyber-Physical Systems*, vol. 2, pp. 194–202, 2022.

[21] M. Pasetti, S. Rinaldi and D. Manerba, 'A virtual power plant architecture for the demand-side management of smart prosumers', *Applied Sciences*, vol. 8, no. 3, p. 432, 2018.

[22] E. Ghorbankhani and A. Badri, 'A bi-level stochastic framework for vpp decision making in a joint market considering a novel demand response scheme', *International Transactions on Electrical Energy Systems*, vol. 28, no. 1, e2473, 2018.

[23] D. Pudjianto, C. Ramsay and G. Strbac, 'Virtual power plant and system integration of distributed energy resources', *IET Renewable power generation*, vol. 1, no. 1, pp. 10–16, 2007.

[24] S. Mohanty, S. Panda, B. K. Sahu and P. K. Rout, 'A genetic algorithm-based demand side management program for implementation of virtual power plant integrating distributed energy resources', in *Innovation in Electrical Power Engineering, Communication, and Computing Technology: Proceedings of Second IEPCCT 2021*, Springer, 2022, pp. 469–481.

[25] P. MacDougall, A. M. Kosek, H. Bindner and G. Deconinck, 'Applying machine learning techniques for forecasting flexibility of virtual power plants', in *2016 IEEE electrical power and energy conference (EPEC)*, IEEE, 2016, pp. 1–6.

[26] L. Lin, X. Guan, Y. Peng, N. Wang, S. Maharjan and T. Ohtsuki, 'Deep reinforcement learning for economic dispatch of virtual power plant in internet of energy', *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6288–6301, 2020.

[27] T. V. Sarathkumar and A. K. Goswami, 'Renewable energy resources forecasting model for virtual power plant in the deregulated electricity market using machine learning', in *2022 IEEE International Conference on Power Electronics, Smart Grid, and Renewable Energy (PESGRE)*, IEEE, 2022, pp. 1–6.

[28] M. Hannan, M. G. Abdolrasol, M. Faisal, P. J. Ker, R. Begum and A. Hussain, 'Binary particle swarm optimization for scheduling mg integrated virtual power plant toward energy saving', *IEEE Access*, vol. 7, pp. 107 937–107 951, 2019.

[29] D. Yu, X. Zhao, Y. Wang, L. Jiang and H. Liu, 'Research on energy management of a virtual power plant based on the improved cooperative particle swarm optimization algorithm', *Frontiers in Energy Research*, vol. 10, p. 785 569, 2022.

[30] D. Hropko, J. Ivanecký and J. Turček, 'Optimal dispatch of renewable energy sources included in virtual power plant using accelerated particle swarm optimization', in *2012 ELEKTRO*, IEEE, 2012, pp. 196–200.

[31] M. Zdrilić, H. Pandžić and I. Kuzle, 'The mixed-integer linear optimization model of virtual power plant operation', in *2011 8th International Conference on the European Energy Market (EEM)*, IEEE, 2011, pp. 467–471.

[32] I. Kuzle, M. Zdrilić and H. Pandžić, 'Virtual power plant dispatch optimization using linear programming', in *2011 10th International Conference on Environment and Electrical Engineering*, IEEE, 2011, pp. 1–4.

[33] A. Dogan and D. Cidem Dogan, 'A review on machine learning models in forecasting of virtual power plant uncertainties', *Archives of Computational Methods in Engineering*, vol. 30, no. 3, pp. 2081–2103, 2023.

[34] S. Hanifi, X. Liu, Z. Lin and S. Lotfian, 'A critical review of wind power forecasting methods—past, present and future', *Energies*, vol. 13, no. 15, p. 3764, 2020.

[35] S. Stevenson, 'A comparison of the forecasting ability of arima models', *Journal of Property Investment & Finance*, vol. 25, no. 3, pp. 223–240, 2007.

[36] S. Siami-Namini, N. Tavakoli and A. S. Namin, 'A comparison of arima and lstm in forecasting time series', in *2018 17th IEEE international conference on machine learning and applications (ICMLA)*, IEEE, 2018, pp. 1394–1401.

[37] P. Wang, K. Xu, Z. Ding *et al.*, 'An online electricity market price forecasting method via random forest', *IEEE Transactions on Industry Applications*, vol. 58, no. 6, pp. 7013–7021, 2022.

[38] W. S. Noble, 'What is a support vector machine?', *Nature biotechnology*, vol. 24, no. 12, pp. 1565–1567, 2006.

[39] E. González-Romera, E. Romero-Cadaval, C. Roncero-Clemente, M.-I. Milanés-Montero, F. Barrero-González and A.-A. Alvi, 'A genetic algorithm for residential virtual power

plants with electric vehicle management providing ancillary services', *Electronics*, vol. 12, no. 17, p. 3717, 2023.

[40] G. Cho, J. Yim, Y. Choi, J. Ko and S.-H. Lee, 'Review of machine learning algorithms for diagnosing mental illness', *Psychiatry investigation*, vol. 16, no. 4, p. 262, 2019.

[41] S. Sadeghi, H. Jahangir, B. Vatandoust, M. A. Golkar, A. Ahmadian and A. Elkamel, 'Optimal bidding strategy of a virtual power plant in day-ahead energy and frequency regulation markets: A deep learning-based approach', *International Journal of Electrical Power & Energy Systems*, vol. 127, p. 106 646, 2021.

[42] T. Park, K. Song, J. Jeong and H. Kim, 'Convolutional autoencoder-based anomaly detection for photovoltaic power forecasting of virtual power plants', *Energies*, vol. 16, no. 14, p. 5293, 2023.

[43] N. Wu, B. Green, X. Ben and S. O'Banion, 'Deep transformer models for time series forecasting: The influenza prevalence case', *arXiv preprint arXiv:2001.08317*, 2020.

[44] A. Al Mamun, M. Sohel, N. Mohammad, M. S. H. Sunny, D. R. Dipta and E. Hossain, 'A comprehensive review of the load forecasting techniques using single and hybrid predictive models', *IEEE Access*, vol. 8, pp. 134 911–134 939, 2020.

[45] A. Natekin and A. Knoll, 'Gradient boosting machines, a tutorial', *Frontiers in neurorobotics*, vol. 7, p. 21, 2013.

[46] F. Farhangi, A. Sadeghi-Niaraki, S. V. Razavi-Termeh and S.-M. Choi, 'Evaluation of tree-based machine learning algorithms for accident risk mapping caused by driver lack of alertness at a national scale', *Sustainability*, vol. 13, no. 18, p. 10 239, 2021.

[47] L. Breiman, *Classification and regression trees*. Routledge, 2017.

[48] A. Graves, A.-r. Mohamed and G. Hinton, 'Speech recognition with deep recurrent neural networks', in *2013 IEEE international conference on acoustics, speech and signal processing*, Ieee, 2013, pp. 6645–6649.

[49] R. C. Staudemeyer and E. R. Morris, 'Understanding lstm–a tutorial into long short-term memory recurrent neural networks', *arXiv preprint arXiv:1909.09586*, 2019.

[50] S. Hochreiter and J. Schmidhuber, 'Long short-term memory', *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[51] A. Sherstinsky, 'Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network', *Physica D: Nonlinear Phenomena*, vol. 404, p. 132 306, 2020.

[52] S. Siami-Namini, N. Tavakoli and A. S. Namin, 'The performance of lstm and bilstm in forecasting time series', in *2019 IEEE International Conference on Big Data (Big Data)*, 2019, pp. 3285–3292. DOI: 10.1109/BigData47090.2019.9005997.

[53] Y. Bao, T. Xiong and Z. Hu, 'Multi-step-ahead time series prediction using multiple-output support vector regression', *Neurocomputing*, vol. 129, pp. 482–493, 2014.

[54] P. Rzepakowski and S. Jaroszewicz, 'Decision trees for uplift modeling with single and multiple treatments', *Knowledge and Information Systems*, vol. 32, pp. 303–327, 2012.

[55] M. Amjad, I. Ahmad, M. Ahmad, P. Wróblewski, P. Kamiński and U. Amjad, 'Prediction of pile bearing capacity using xgboost algorithm: Modeling and performance evaluation', *Applied Sciences*, vol. 12, no. 4, p. 2126, 2022.

[56] S. Dalal, B. Seth, M. Radulescu, C. Secara and C. Tolea, 'Predicting fraud in financial payment services through optimized hyper-parameter-tuned xgboost model', *Mathematics*, vol. 10, no. 24, p. 4679, 2022.

[57] Z. Pan, W. Lu, H. Wang and Y. Bai, 'Groundwater contaminant source identification based on an ensemble learning search framework associated with an auto xgboost surrogate', *Environmental Modelling & Software*, vol. 159, p. 105 588, 2023.

[58] G. Webber-Cross, *Learning Microsoft Azure*. Packt Publishing Ltd, 2014.

[59] H. Moayedi, D. T. Bui, A. Dounis, Z. Lyu and L. K. Foong, 'Predicting heating load in energy-efficient buildings through machine learning techniques', *Applied Sciences*, vol. 9, no. 20, p. 4338, 2019.

[60] G. Ke, Q. Meng, T. Finley *et al.*, 'Lightgbm: A highly efficient gradient boosting decision tree', *Advances in neural information processing systems*, vol. 30, 2017.