Demand Response for Efficient Power Generation in Smart Grids Using Hyperlocal Weather Predictions

By

Faraz

00000398859

Supervisor

Dr. Mehak Rafiq

A thesis submitted in conformity with the requirements for the degree of Master of Science in Computational Science and Engineering

Department of Sciences

School of Interdisciplinary Engineering & Sciences (SINES)

National University of Sciences & Technology (NUST)

Islamabad, Pakistan

January 2024

Annex A to NUST Letter No.
09/2/102/Exams/Thesis-Cert
dated 23 Dec 16.

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by Mr FARAZ Registration No. 00000398859 of SINES has been vetted by the undersigned, found complete in all aspects as per NUST Statutes/Regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfilment for award of MS/MPhil degree. It is further certified that necessary amendments, as pointed out by GEC members of the scholar, have also been incorporated in the said thesis.

Signature with stamp: DR. MEHAK RAFIQ Assistant Professor SINES, National University of Science & Technology H-12 Islamabad

Name of Supervisor: Dr Mehak Rafiq

Date: 27/12/23

Signature of HoD with stamp: Dr. Fouzia Malik HoD Sciences Professor SINES NUST, Sector H-12 Islamabad

Date: 27-12-2023

Countersign by

Signature (Dean/Principal): Dr. Hammad M. Cheema Principal & Dean SINES - NUST, Sector H-12 Islamabad

Date: 08 JAN 2024

Declaration

I, Faraz, declare that this thesis titled "Demand Response for Efficient Power Generation in Smart Grids Using Hyperlocal Weather Predictions" and the work presented in it are my own and have been generated by me as a result of my original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a Master of Science degree at NUST
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at NUST or any other institution, this has been clearly stated
3. Where I have consulted the published work of others, this is always clearly attributed
4. Where I have quoted from the work of others, the source is always given. Except for such quotations, this thesis is entirely my work
5. I have acknowledged all primary sources of help
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself

Faraz

00000398859

Copyright Notice

Dedicated to my beloved parents

Acknowledgements

Alhamdulillah – all glory to Allah, the Almighty, for allowing me to conduct this research in detail and with novel results.

I am grateful to my supervisor, Dr. Mehak Rafiq, for her guidance, dedication, and support. Her knowledge and feedback were essential to the success of this research. I am also grateful for her allowing me to choose my research topic and all the methodology from data collection to the end of the results.

I am also thankful to my co-supervisor, Dr. Hasan Ali Khattak, for his support, dedication, and guidance and for frequently providing valuable feedback throughout this journey.

I would also like to thank my evaluation committee members, Dr. Ammar Mushtaq and Dr. Muhammad Tariq Saeed, for their guidance and cooperation. I am grateful for their support throughout my coursework and research work.

Finally, I would like to thank everyone who indirectly contributed to this research, including my friends, classmates, and lab fellows at Data Analytics Lab, SINES, NUST. I appreciate their cooperation, which helped to make this research possible.

Abstract

Optimising power-generation efficiency has arisen as a critical challenge in the fast-changing environment of smart grids and urban ecosystems. Properly using energy resources has become essential in today's technologically advanced world. Smart Grids are at the forefront of this shift, giving a complete energy consumption and distribution approach. Given the delicate connection between climatic conditions and energy use patterns, including weather data in these networks is an improvement and a necessity now. Significant advances in the domain have used large amounts of weather data and sophisticated models. Future studies can improve on these findings by combining sophisticated time series models with meteorological data and optimising them for demand response techniques in smart grid power generation. This study presents a novel strategy combining smart grid electricity generation demand response mechanisms with hyperlocal weather forecasts. The project attempts to improve the accuracy and reliability of power generation estimates by using the capability of machine learning. Five distinct time series and machine learning models - SARIMAX, Prophet, Holt-Winters, XGBoost, and LSTM – have been integrated with hyperlocal meteorological data, encompassing precipitation, relative humidity, temperature, and cloud cover. SARIMAX has shone out among the individual models, with a MAPE of 4.92% and an MAE of 3.54, demonstrating its ability to capture subtle seasonal patterns and autocorrelations. The hybrid model, an ensemble of SARIMAX, Prophet, and Holt-Winters, outperformed the individual SARIMAX in predicted accuracy, boasting an impressive 0.06% MAPE and an MAE of 4.43. When paired with real-time data analytics, this demonstrates the transformational potential of machine learning algorithms. A new aspect of this research is introducing a user-centric dashboard, which provides a real-time display of anticipated data and model performance indicators. Its versatility is enhanced further by user-specific customisation options, which provide specialised forecasting insights over user-defined timeframes, increasing real-time decision-making processes. The combination of demand response tactics with powerful machine learning models, demonstrated by the hybrid model's excellent performance, offers a promising path toward increased flexibility and efficiency in smart grids and cities.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AI	Artificial Intelligence
ML	Machine Learning
DR	Demand Response
DER	Distributed Energy Resources
STLF	Short-Term Load Forecasting
PF	Power Factor
kPt	Active Power in Unit of Kilo
kQt	Reactive Power in Unit of Kilo
KSt	Apparent Power in a Unit of Kilo
ARIMA	Autoregressive Integrated Moving Average
ADF	Augmented Dicky-Fuller Test
LSTM	Long Short-Term Memory
SARIMAX	Seasonal Autoregressive Integrated Moving Average with Exogenous Variables
UTC	Universal Time Coordinates
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
MSE	Mean Squared Error
MAPE	Mean Absolute Percentage Error

Chapter 1 Introduction

This chapter presents the need for this research, the problem statement, aims and objectives, and the structure of this thesis.

1.1 Overview

Demand response comes into play when electric utility customers change their energy usage to match supply and demand more closely. Historically, utilities have had to alter power production rates or turn producing units on and off to match demand since electric energy could not be easily stored. However, the complexity of new energy integration and flexible demand response has increased the complexity of smart grid operating scenarios [1].

The need for energy is crucial because it determines how power systems are built and run. The overall infrastructure and capital expenses can be kept to a minimum by regulating and lowering peak demand through techniques such as demand response. Additionally, as peak demand is frequently used to scale power generating and transmission networks, reducing this demand can result in significant savings and more effective resource usage [2].

Smart Grids (SGs) make it possible to make more informed judgments about power use, which improves communication between electricity providers and customers. The demand response dynamism and flexibility make centralised load estimates more challenging and unpredictable. The conventional power grid, which has been in use for more than a century, was built to support only one-way communication and electricity transmission. However, a more dynamic and interactive system was required because of the rapid development of technology and the rising need for power, prompting the growth of the Smart Grid (SG). This modernised electrical grid employs information and communications technology to automatically gather and act on information, such as consumer and supplier behaviour, to increase the efficiency, dependability, and sustainability of electricity production and distribution [3].

The SG promises Numerous advantages, such as improved dependability, cost-effective production, and environmentally friendly energy generation [4]. The capacity of smart grids to incorporate a range of energy supplies, including renewable ones such as solar and wind, is one of their main advantages. This integration makes a more adaptable and robust power system capable of changing with demand and supply situations.

Additionally, the SG brings the idea of peer-to-peer energy trading, where customers may create energy in addition to consuming it and selling it back to the grid or to other consumers directly. This decentralised method of managing energy fosters competition in the energy market, reduces power outages, and increases the general effectiveness of power systems [5].

However, the switch to smart grids has its challenges. Integrating multiple energy resources and the growing interconnection raises new security and privacy risks. It is crucial to ensure the safety of the smart grid and the smart home, both essential components. While maintaining the entire system's security, the threats to the smart grid environment must be recognised and countered [4].

The electrification of the fleet of cars, such as plug-in hybrid electric vehicles (PHEVs), adds to the stress on the electrical grid. With their sophisticated regulatory systems, smart grids can reduce the energy used to charge PHEVs, maintaining a stable and adequate power system [6]. The development of smart grids is a meaningful change for the power sector, bringing many advantages and new difficulties. The future of energy management and consumption is anticipated to be significantly shaped by smart grids as research and development continue.

Energy demand has always been significantly influenced by weather, especially when it comes to renewable energy sources such as solar and wind. The production predictability of these renewable sources becomes increasingly important as our globe turns to them. Renewable energy sources are subject to the whims of nature, as opposed to conventional energy sources, which can be managed and changed in response to demand. There might be variations in the amount of energy produced if the sun does not shine and the wind does not blow.

This unpredictability highlights the importance of hyperlocal weather prediction. The operation of an SG depends on accurate forecasts of power demand and renewable energy production. However, the unpredictable nature of renewable energy sources and the wide range of user behaviours make the prediction process problematic [7]. For SG operators and solar energy suppliers, accurate forecasting of energy generation—especially solar power—is essential. Due to its considerable reliance on weather factors, renewable power needs accurate forecasts to guarantee power continuity and make suitable preparations for energy dispatch and storage [8].

Predicting energy demand and supply correctly is increasingly challenging as the SG becomes more innovative and interconnected. Without adequate planning, a rapid decline in renewable energy production might cause blackouts and unstable system conditions. Therefore, hyperlocal weather forecasts are essential for preserving the dependability and stability of the whole power system and maximising the use of renewable energy sources. Some machine learning algorithms, e.g., long short-term memory (LSTM) recurrent neural networks with weather features as input, have been proposed in models to improve the accuracy of power load forecasting for individual energy users because weather factors significantly impact residential load prediction [9].

1.2 Problem Statement

SGs are critical infrastructures in the constantly evolving environment of smart cities, guaranteeing efficient energy delivery and consumption. While sophisticated, these networks have the enormous problem of meeting changing energy needs tightly linked to local meteorological conditions. The idea of hyperlocal weather prediction, which provides detailed meteorological information at the micro-level, adds another degree of complication. While integrating hyperlocal meteorological data with smart grid systems has enormous potential, it poses many obstacles. The combination of these disparate variables necessitates the development of a forecasting system capable of deciphering the subtleties of hyperlocal weather patterns and correctly predicting energy consumption. The overriding goal is to create a system that can harmonise different statistics, guaranteeing that SGs are effective in the present while also being proficient at meeting the varied energy demands of future smart cities.

1.3 Aims and Objectives

1. To provide a mechanism that integrates hyperlocal weather forecasting with smart grid systems, providing accurate energy demand forecasts for current and future smart cities.
2. To Create a hybrid model that combines the characteristics of distinct machine learning models to achieve improved flexibility and precision in the face of hyperlocal weather data unpredictability.
3. To Create a real-time dashboard that visualises projected data, hyperlocal weather patterns, and model performance indicators, allowing for quick decision-making and increasing user involvement.
4. To Compare the proposed hybrid model to standard forecasting approaches, proving its usefulness in setting smart cities and measuring its resilience in dealing with the complexities of hyperlocal weather prediction.

1.4 Structure of the Thesis

This thesis is structured as follows: Chapter 1 gives the overview and an introduction of basic concepts and why they need to be addressed. Chapter 2 summarises related studies in this domain, and Chapter 3 gives an overview of data collection and pre-processing, feature engineering techniques, and experimental methodology to make this research possible. Chapter 4 discusses the results in detail, and Chapter 5 offers Future Direction and Discussion.

Chapter 2 Literature Review

Electrical systems have evolved during the last decade. Traditional power networks have undergone substantial changes to accommodate the dynamic nature of modern energy demands and the introduction of renewable energy sources. Initially, these networks were designed to provide a one-way power flow from centralised production sources to customers.

In the early 2010s, there was a surge of research on the grid integration of renewable energy sources. [10] stressed the importance of defining a roadmap for smart grid interoperability standards, citing the difficulty of combining multiple energy sources. Smart Grids (SGs) were planned to be sophisticated power grid systems that could employ digital communication technologies to monitor and respond to local variations in electricity usage [11].

The need for a two-way flow of information and electricity, allowing for better communication between suppliers and customers, was the primary impetus for the construction of SGs [12]. This adjustment proved critical in reducing blackouts and providing financial and environmental benefits to energy companies and end users.

As the decade progressed, the need for Demand Response Management (DRM) in SGs became clear. With the increasing integration of renewable energy sources, the output of which can be inconsistent and unpredictable, DRM systems have played an essential role in guaranteeing grid stability [13]. DRM programs have been enhanced further by expanding communication networks, particularly with the arrival of 5G Internet of Things (IoT) technologies [14]. These advancements made data transport quicker, more reliable, and more secure, and the ability to manage many connections conceivable.

Renewable energy sources and their difficulties, particularly their dependency on environmental conditions, have become a study focus in recent years. While smart meters enabled real-time demand prediction, it was found that exact models to estimate the electricity delivered by renewable sources were urgently needed [15]. These prediction models (PMs) were essential for grid stability, realistic scheduling, and energy management

[16]. For example, if a model forecasted that renewable energy would be depleted, the SG would need to smoothly transition to conventional energy sources to ensure that the generated electricity met the predicted demand.

The significance of hyperlocal weather forecasting in SGs has been emphasised. Temperature, wind speed, and humidity all impact renewable energy sources, causing fluctuations in the amount of electricity they generate [17]. Advanced machine learning and deep learning algorithms have been created to improve the accuracy of these predictions [18]. For example, the Pre-Attention Mechanism and Convolutional Neural Network-based Multivariate Load Prediction model were proposed to increase demand responsiveness in SGs [19].

The application of artificial intelligence (AI) in SGs has created new research prospects. To maximise demand-side management in SGs, AI-based renewable energy projection tools have been developed [20]. These algorithms provide efficient power production and provide actionable SG system performance data.

Combining demand response mechanisms with hyperlocal weather predictions is crucial for improving power-generating efficiency in smart networks. The usefulness of meteorological data and machine learning in improving power demand estimates was stressed by Inagata et al. 2023 [21]. Building on this subject, Kadlec et al. (2017) provided an algorithm that uses previous metering data and weather forecasts to arrange the activity intervals of water heaters in smart grids [22]. The technique is analogous to a perfectly tuned symphony, with each instrument (the water heaters) playing in sync with the unpredictable weather patterns and solar energy production cycles. This study emphasises the complex balance between energy demand and supply, which was accomplished using the predictive potential of meteorological data.

The work of Rao et al. (2021), who investigate IoT-based energy management systems [23], is at the forefront of this integration. Their research presents a picture of a future in which photovoltaic (PV) panels are active players in the energy system, conveying real-time meteorological data rather than passive energy harvesters. When this data is analysed in the cloud, it unlocks the possibility of optimising PV power output, thereby improving the grid's response to shifting energy demands.

Banerjee et al. (2021) broadened the canvas by investigating several demand response mechanisms in smart grids [24]. The study emphasised the vital interaction between renewable energy sources and weather patterns, demonstrating how combining smart grid communication technology with demand response programmes may result in more cost-effective and dependable power delivery, particularly during peak demand periods. This story emphasises the consumer's involvement in the energy ecosystem, where each user's consumption behaviour adds to the overall stability and efficiency of the grid.

Guo et al. (2022) investigated a day-ahead dispatch approach for high-penetration solar networks with virtual power plants (VPPs) [25]. The research demonstrated the potential of foresight, where weather predictions are used to prepare the energy infrastructure for the effects of the weather, illustrating how VPPs may adjust to the fluctuating levels and flow of solar energy by estimating photovoltaic power and implementing a price compensation scheme, providing a consistent and cost-effective energy supply.

Arumugham et al. 2023 [26] highlighted the need to integrate artificial intelligence and renewable energy forecasts in smart grid demand-side management. Singhal et al. 2022 [27] proposed a model that employs machine learning to estimate solar power generation from weather forecasts, emphasising smart grids' real-time adaptability. Islam et al. 2022 [28] proposed a data-driven technique for predicting event-driven data traffic in the cyber-physical layers of smart power grids, highlighting the necessity of such models in enhancing grid management. Thusitha et al. 2020 [29] emphasised the importance of robust demand response systems in smart grids, especially considering potential false data injection attacks.

Short-term load forecasting (STLF) has significantly benefited from machine learning methodologies. Bibi Ibrahim et al. examined machine learning in smart grids for short-term load forecasting [30]. Data collection, feature selection, pre-processing and transformation, model training, model assessment, and model selection were all part of their methodology. The study stressed the significance of climatic parameters in improving the accuracy of power forecasts. The study also investigated the importance of various parameters in forecasting power consumption, including the previous week's same-day same-hour load, the last day's same-hour load, and temperature.

Traditional statistical methods such as ARMA, ARIMA, exponential smoothing, linear regression, and the similar day approach are widely used. They typically, however, fail to represent the complex non-linear interactions between input and output variables [31]. As a result, their performance under STLF conditions may need to be improved.

Deep Learning techniques, on the other hand, have demonstrated promising success in overcoming the limits of statistical-based models. They can replicate complex non-linear mappings between inputs and outputs, find hidden patterns in vast datasets, and scale. Artificial Neural Networks (ANNs) are a great example of machine learning methodologies commonly used in STLF [32].

Integrating smart grids in urban areas is crucial for establishing smart cities, ensuring an efficient and reliable energy supply. The importance of smart grids in smart cities is highlighted in energy distribution, public safety, and other urban operations [33]. Because Renewable Energy Resources (RES) fluctuate and are influenced by meteorological factors, precise prediction models are necessary to ensure grid stability and effective energy management [26].

Several aspects of DR have recently been studied. Liu et al. (2023), for example, underlined the relevance of Virtual Power Plants (VPPs) in combining distributed energy resources (DERs) and real-time power production regulation to balance electricity supply and demand [35]. Arumugham et al. (2023) emphasised the need for renewable energy forecast models in smart grids for grid stability and effective energy management [26]. Bakare et al. (2023) comprehensively investigated the challenges of demand-side energy management in smart grids, including technological, economic, and regulatory constraints [37]. Salazar et al. (2023) created a reinforcement learning-based pricing and incentive system for smart grid demand response, emphasising the relevance of real-time pricing schemes in boosting demand displacement [38].

Table 1 provides a quick comparative overview of state-of-the-art work on the successful application of Artificial Intelligence in predictive analysis for SG energy generation and distribution.

Table 1 Comparison of the state-of-the-art work on the practical application of AI in predictive analysis for smart grid power generation and distribution.

Sr. No.	Title	Year	Contribution(s)	Limitation(s)
01	Incentive compatible demand response games for distributed load prediction in smart grids [39]	2014	With user participation, a game-theoretic demand response mechanism changes typical centralised load forecasting into a distributed system.	The problem of using theoretical models in real-world settings, where consumer behaviour and system dynamics are more intricate and unpredictable than in simulations.
02	Optimization Method with Prediction-Based Maintenance Strategy for Traction Power Supply Equipment Based on Risk Quantification [40]	2018	It creates a risk-quantified, prediction-based maintenance plan for traction power supply equipment and applies Bayesian classifiers based on historical data for fault prediction and maintenance optimisation.	Because the study is based on historical data, it cannot account for unexpected changes or irregularities in equipment behaviour.
03	Industrial load forecasting using machine learning in the context of smart grid [41]	2019	Investigates the use of machine learning in estimating power consumption for a meat processing business to enhance load forecasting in industrial smart grids.	Significant forecasting errors arise due to the unpredictable nature of industrial energy demand, indicating a need for existing machine-learning techniques for complex industrial settings.
04	Redills: Deep Learning-Based Secure Data Analytic Framework for Smart Grid Systems [42]	2020	A deep learning-based system that forecasts future load usage in SGs using LSTM models and a priority analyser system for improving time-of-use (ToU), assisting in energy use control and cost reductions.	Redills architecture has limitations in particularly volatile or unpredictable energy patterns when LSTM models fail to adequately capture rapid load consumption changes.
05	Short-Term Load Forecasting Based on Adabelief Optimized Temporal Convolutional Network and Gated	2021	Introduces a hybrid neural network with AdaBelief optimiser for enhanced short-term load forecasting accuracy by combining a Temporal	The study is limited regarding weather data detail and data feature dimensionality. The model's prediction efficacy must be

	Recurrent Unit Hybrid Neural Network [43]		Convolutional Network with a Gated Recurrent Unit.	evaluated in various contexts to increase forecast accuracy, and novel load data preparation procedures must be examined.
06	A Stacked GRU-RNN-Based Approach for Predicting Renewable Energy and Electricity Load for Smart Grid Operation [44]	2021	An improved stacked Gated Recurrent Unit-Recurrent Neural Network (GRU-RNN) prediction technique for renewable energy generation and electrical load enhances accuracy in univariate and multivariate scenarios.	The inherent unpredictability and intermittent nature of renewable energy sources and the complexities of human behaviour in power use limit the model's efficiency.
07	Solar-Cast: Solar Power Generation Prediction from Weather Forecasts using Machine Learning [27]	2022	A prediction model for solar power generation based on weather forecasts using Linear Regression, Ridge Regression, and Lasso Regression to increase smart grid efficiency.	The intrinsic unpredictability of weather conditions limits the model's accuracy, reducing the dependability of 48-hour solar intensity forecasts.
08	A Machine Learning-Based Gradient Boosting Regression Approach for Wind Power Production Forecasting: A Step towards Smart Grid Environments [45]	2022	Five ML regression algorithms for wind energy forecasting are compared: Random Forest, Gradient Boosting Machine (GBM), K-Nearest Neighbour (kNN), Decision Tree, and Extra Tree Regression, with GBM outperforming the others in forecasting accuracy for Turkish Wind Farms.	The intermittent nature of wind and ever-changing meteorological conditions present problems that impair the forecasting accuracy of these models.
09	Optimal Adaptive Prediction Intervals for Electricity Load Forecasting in Distribution Systems via Reinforcement Learning [46]	2022	Introduces a reinforcement learning-based online, data-distribution-sensitive quantile estimation technique for building adaptive prediction intervals in electrical demand forecasting.	The variation in load patterns and the dependency on the data input quality and online learning effectiveness of the reinforcement learning model limits the performance.
10	AI-based forecasting for optimised solar energy management and smart grid efficiency [47]	2023	To improve the accuracy of solar power generation estimations and overall SG efficiency, a Deep Learning model based on LSTM approaches was developed.	The model's predicted accuracy falls with longer prediction horizons, revealing an issue sustaining forecast precision over extended durations.
11	A Comprehensive Analysis of Smart Grid Stability Prediction along with	2023	SG stability prediction based on explainable AI and data analytics,	As a result of renewable energy integration and bidirectional energy flow,

	Explainable Artificial Intelligence [48]		focusing on improving Decentralised SG Control using sophisticated feature engineering and a combination of classification and regression models.	there is inherent complexity in SG stability, emphasising the challenge of real-time data processing and analysis for successful management.

Chapter 3 Methodology

This chapter gives a detailed understanding of data acquisition methods, sources of data from where the datasets are sourced, understanding of variables present in the datasets, methods to pre-process and clean the data, description and usage of machine learning models used, results, and justification.

3.1 Data Acquisition

Data acquisition is the first method in the methodology. To address the problem, two datasets are sourced, for which the descriptions are given below separately. Both datasets are in comma-separated variables (CSV) format, making it easy to convert them to data frames and then merge them seamlessly using viable join strategies.

3.1.1 Smart Grid Batteries Data

This study used a broad collection of Smart Grid Battery Storage data gathered from IEEE Dataport [50]. This dataset is an information-rich repository, recording a wide range of electrical measurements from a smart grid storage system in Taiwan. It has 28 variables, including essential metrics such as power factor, frequency, current, and voltage measurements for each phase of a three-phase power system. It also offers statistics on apparent and real Power, among other factors, giving a detailed picture of the grid's functioning and energy flow.

Data collection, acquired throughout September 2022, consists of 76,966 distinct recordings, with data points recorded at a high-resolution frequency every 30 seconds. This high-frequency sampling provides a thorough temporal picture of the grid's dynamics, enabling exact study and modelling of energy storage characteristics.

The description for the first dataset is given in the table below.

Table 2 Description of variables from Smart Grid Battery Data acquired from IEEE Dataport

Variable	**Description**
Freq	Frequency, generally in Hz. Typically, around 50 or 60 Hz in power systems, depending on the country.
Ia, Ib, Ic	Currents in each of the three phases of a three-phase power system, usually in Amperes (A).
PF	Power Factor is a measure of how effectively the power system is using electricity. It varies between -1 and 1, with one being ideal.
UpdateTime	Timestamp for each reading or observation in the data.
Va, Vb, Vc	Voltages in each of the three phases of a three-phase power system, in Volts (V).
kPt, kQt, kSt	Represent the real power (kPt), reactive power (kQt), and apparent power (kSt) in kilo units.
kVARh_Q1, kVARh_Q2, kVARh_Q3, kVARh_Q4	Reactive energy, measured in kilovar-hours, for each of the four quadrants in power systems. Important in power factor correction.
Quadrants in Power Systems/Smart Grids	Quadrant I: Both active and reactive powers flow positively and are delivered to the consumer load (inductive influence, import condition). Quadrant II: Reactive power is positive, and active power flows negatively (export condition). Quadrant III: Both reactive and active power flow negatively (export condition). Quadrant IV: Reactive power flows negatively, active power flows positively (import condition).
kVAh+, kVAh-	Positive and negative apparent energy, measured in kilovolt-ampere hours.
kWh+, kWh-	Positive and negative real energy, measured in kilowatt-hours.
kWh_abs, kWh_net	The absolute and net real energy, in kilowatt-hours.

?IA, ?IB, ?IC	Error or uncertainty measurements for each of the three-phase currents.
?VA, ?VB, ?VC	Error or uncertainty measurements for each of the three-phase voltages.

3.1.2 Hyperlocal Weather Data

In parallel, hyperlocal meteorological data from Open-Meteo [51] was used to capture environmental elements that substantially impact energy use and generation. This information is customised to the smart grid's unique location, ensuring that the weather conditions are as precise and relevant for the research period as feasible. Temperature, Relative Humidity, Apparent Temperature, Precipitation, Cloud Cover, and Windspeed are among the significant meteorological variables measured hourly in the weather dataset. These factors are critical for comprehending and forecasting the effect of weather on energy storage and consumption patterns.

The descriptions for the variables given in the hyperlocal weather data are given in Table 3. It should be noted that the weather data is for the local area of the educational institute for which the smart grid data is available so that the research goal of integrating hyper-local weather data can be achieved.

Table 3 Weather variables and descriptions acquired from the open-meteo historical data repository.

Variable	**Description**
Time	The timestamp for the recorded data
Temperature	Temperature at 2 meters above the ground level, in degrees Celsius.
Apparent Temperature	Perceived temperature, in degrees Celsius.
Humidity	Relative humidity 2 meters above the ground level, in percentage.
Precipitation	Precipitation amount, in millimetres.
Cloud Cover	Percentage of the sky covered by clouds.
Wind Speed	Wind speed at 10 meters above the ground level, in kilometres per hour.

3.1.3 Merging Datasets

The combination of these two datasets gives a comprehensive perspective of the operation of the smart grid, contrasting technical electrical data with environmental context. This comprehensive approach is critical for constructing accurate forecasting models and comprehending the interaction between grid performance and weather conditions. The thorough and frequent data gathering highlights the possibility for nuanced insights into energy storage optimisation and the construction of more robust and efficient smart grid systems.

3.2 Data Preprocessing

Data pre-processing is an essential phase in data analysis, especially when working with time series data, which frequently contains complexities such as noise, missing values, and outliers. These concerns can reduce the accuracy of any forecasting models drastically applied to the data. As a result, pre-processing tries to clean and transform raw data into a format that can be efficiently used for future analysis and model training [52].

After importing both datasets, only useful variables were kept, while the rest were removed as they did not show much significance in the forecasting. Some of these variables have been removed as their values were not necessary in the analysis, and including them is not applicable. These variables include voltage and current in all three phases, reactive power from three quadrants (because of NaN values) and so on.

The data was inconsistent as grid data was sampled with a frequency of 30 seconds while the weather data was sampled with 1 hour. It was necessary to resample the Grid data to join both datasets in an hourly sampled format. The steps that have been performed after importing the data are discussed in detail as follows:

3.2.1 Renaming Columns

To improve readability and accessibility, dataset columns are renamed, for example, temperature_2m (°C) to temperature, relativehumidity_2m (%) to Humidity, and so on. The reason for doing this is that when the columns are addressed in terms of using in code or interpreting while performing some analysis, it makes it easier to have variables with their real names or with names that can be easily understood. Similarly, while designing the dashboard and coding in general, it helps to know how humidity and cloud Cover are used in the analysis instead of seeing the column names and writing them in the code again.

3.2.2 Missing Values

After merging the two datasets by outer join, there was a chance of having missing values as it can be noticed that both datasets had a bit of inconsistency when it comes to frequency, i.e., the grid data is sampled by 30 seconds while the metrological data is sampled by one hour. The joined dataset is why it is sampled by hourly frequency. By doing so, the dataset now has no known inconsistency, and in this way, for each row in the dataset, there are variables for weather and smart grid battery data. Missing values are removed to avoid problems during machine learning implementation [54].

3.2.3 Moving Averages

Moving averages and other rolling window elements are intended to smooth out short-term swings and emphasise longer-term patterns in the data. These characteristics can assist models in better comprehending and forecasting the underlying trends in the series by finding the mean across a defined timeframe. Rolling window characteristics have been investigated in several forecasting systems and proven to improve prediction accuracy [55].

3.2.4 Stationarity

Establishing the dataset's stationarity is critical to assure the dependability of time series forecasting models. Stationarity means that the time series statistical features, such as mean, variance, and autocorrelation, remain constant. Non-stationary data might result in incorrect models due to patterns or seasonality [56].

The Augmented Dickey-Fuller (ADF) test is a famous statistical test that looks for unit roots in a time series, which indicates non-stationarity. The ADF test's null hypothesis states that the time series has a unit root and is non-stationary. The alternative hypothesis, on the other hand, indicates that the time series is stationary [57].

Non-stationarity in time series data needs procedures to stabilise the variance and mean. Among all those techniques, log transformation and lag feature development are two strategies used in this study.

3.2.5 Log Transformation

A log transformation is done to specific characteristics to normalise the distribution or lessen the skewness of the data [59]. Log transformation effectively reduces non-constant variance in a time series, a problem known as heteroscedasticity [60]. It may typically turn a non-linear connection into a linear one using a logarithmic scale on the data, which is required for many time series forecasting models [61]. This adjustment can help reduce the impact of any significant outliers, which can disproportionately impact the model's performance. The log transformation was performed to the kPt time series in our investigation to reduce variability and trend in the data and help meet the stationarity assumption. A simple numpy method implies that a log transformation with one addition is done to the target variable to ensure positive values.

3.2.6 Lag Features

Lag features are a type of feature engineering used in time series analysis. They add past time step values into the model as independent features [60], which might assist the model in capturing the temporal relationships within the data. For example, in this approach, the value of kPt an hour ago (lag1) and two hours ago (lag2) offers context to the model that is predictive of future values [61]. When dealing with autoregressive processes, where previous values systematically influence future values, including these lag aspects is critical.

These characteristics can assist in capturing the temporal interdependence inherent in time series data by moving the series by one or more periods. The necessity of picking relevant characteristics and calculating their effective window widths of lagged data was noticed. At the same time, it was felt that there were not enough variables in the data to provide significant and aimed results from the smart grid data; this can have a substantial impact on predicting success.

3.3 Evaluating Demand Response for Smart Grids

Understanding a grid's demand response is critical for balancing energy supply and demand, maintaining stability, and maximising resource allocation. While the information in question is restricted to a month's worth of data, it offers a unique chance to examine the grid's performance, particularly regarding energy storage and generation. The existence of a variable indicating positive energy generation is very advantageous. It enables daily data sampling to track the incremental energy stored in the batteries, which can then be used to calculate the average energy delivered to the grid. This conclusion is crucial since it aids in understanding not just consumption patterns but also the efficiency and adequacy of energy generation about grid demand.

Given that the hyperlocal environment is a Taiwanese educational facility, the energy consumption patterns should be unique to its operating dynamics. In previous studies, such as in [62], authors have measured the demand, providing a basic understanding of the institute's energy use in Mega Watt-hours (MWh). This historical data helps create a baseline against which to evaluate the present dataset's findings.

To better understand the grid's demand response, examine both the temporal changes in energy use and the possibility for scalability. Energy consumption trends at educational institutions are mainly predictable based on their academic calendar, with fluctuations over weekdays, weekends, and holidays. It may derive not just the average daily supply but also the peak demand periods and the grid's reaction to these changes by examining the daily rise in battery storage [65].

Furthermore, the scalability of the grid's energy supply is crucial. The grid must adjust as the institute grows or changes its energy usage habits. The data gives a snapshot that may be used to advise future grid expansions or changes to meet shifting demand. While the dataset's period is only one month, it provides insight into the grid's operating characteristics.

Equation 1 below gives the power demand calculated in Mega Watts-hour (MWh). In comparison, the equation is presented with the help of literature [62].

$$\text{Total Energy Consumption (MWh)} = \text{Energy Consumption per unit Area} * \left(\frac{\frac{\text{kWh}}{\text{m}^2}}{\text{month}}\right) * \text{Total Area}(\text{m}^2) * \frac{1}{1000}$$

Equation 1 Finding Power Demand (in MWh)

3.4 Machine Learning Models

The advent of Machine Learning has brought many advancements in research and development not only for computer sciences but also for interdisciplinary fields ranging from electrical engineering to archelogy. To employ machine learning, one must train data on some architecture known as machine learning models. The machine learning models are sophisticated enough to get the data and use the underlying mathematical concepts that are fed to the models for analysing the data and predict the values or results based on the underlying data on which these models are trained and the new data that is now being added to them for predicting the values.

There are many machine learning models, but the models dealt with here are Time Series models, i.e., Facebook Prophet, Holts-Winter Exponential Smoothing, and Seasonal Autoregression Integrated Moving Averages (SARIMAX). Deep learning models, i.e., Long Short-term Memory (LSTM) and the highly regularised, tuned, and boosted models, i.e., XGBoost.

Different time series and Conventional Machine Learning models have been deployed to predict and forecast Energy consumption patterns involving weather data. Absolute Power (kPt) has been chosen as the dependent variable as the primary goal of this research and analysis is to map or forecast the power generation from the smart grid while keeping metrological data in view.

The implemented models are also discussed in this chapter separately, along with their implementation.

3.4.1 Facebook Prophet Model (Trained on Smart Grid Data Only)

The Prophet model on Facebook is a model for forecasting time series data. It is based on an additive model incorporating non-linear trends and seasonal and holiday effects. Prophet is particularly effective for datasets with large seasonal patterns, and it is resistant to missing data and outliers, making it ideal for the Smart Grid Data discussed here.

First, a user-friendly FBProphet model asks the user how many hours of predicted energy use they desire in the future. This model has now been trained solely on Grid data. The reason for doing this forecast is to differentiate between the individual models with just smart grid data and the models merged with hyperlocal weather data.

Resampling is used to convert the data into an hourly time series, which is required for time series analysis and forecasting. The mean for each hour is calculated as shown in Equation (2).

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2}$$

Where, μ (mu): The mean or average of the dataset.

n: The total number of observations in the dataset.

∑ (Sigma): The summation symbol indicates that a sum is being taken.

i: An index used in the summation, typically representing each observation or element in the dataset.

x_i: The i-th observation or element in the dataset.

Training entails fitting the model to historical data to discover underlying patterns such as trends and seasonality. The Prophet model is represented in Equation (3).

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \qquad (3)$$

Where y(t) = The output or response variable.

G(t) = Trend component at time t.

S(t) = Seasonal component at time t.

H(t) = Holiday component at time t.

ϵ_t = Error component at time t.

3.4.2 Facebook Prophet Model Forecasting (On Grid and Weather Data)

Facebook's Prophet model is a powerful forecasting tool for time series data with significant seasonal trends and numerous seasons. It is especially well-suited for data with daily observations containing missing values and unique occurrences that may impact trends [13]. The time series decomposable model integrates trend, seasonality, and holiday impacts, providing flexibility and simplicity of understanding. The Prophet model was chosen for this approach along with SARIMAX because of its ability to manage the dataset's complicated seasonal cycles.

The dataset is reconstructed to meet the Prophet model's input requirements. The data frame has the date, target variable, and exogenous attributes labelled suitably to help the model identify and handle the data items. The Prophet model is configured to account for yearly and daily seasonality, ensuring that the model captures intrinsic patterns and swings at multiple time scales [65]. Exogenous variables are added as regressors to improve the model's prediction accuracy by accounting for external factors.

The implementation of Prophet combined with hyperlocal weather data is shown in the equations below.

The equation (3) is again used here first to train the basic Prophet model.

Equation (4) shows the trend component, Equation (5) shows the seasonality component and Equation (6) shows the holiday trend used in the model to catch the trend and seasonality from the Time Series Data. Equation (7) shows how exogenous variables are included in the model training.

$$\mathrm{g}(t) = \frac{C}{1 + e^{-k(t-m)}} \tag{4}$$

Where C: Carrying capacity.

k: Growth Rate

m: Offset Parameter

$$\mathrm{s}(t) = \sum_{n=1}^{N} \left(a_n cos\left(\frac{2\pi \mathrm{nt}}{P}\right) + b_n sin\left(\frac{2\pi \mathrm{nt}}{P}\right) \right) \tag{5}$$

The seasonal component $\mathrm{s}(t)$ at time t is modelled using the Fourier Series to provide flexibility.

Where P: Period (e.g., 30 for monthly seasonality)

a_n and b_n: Fourier coefficients

$$\mathrm{h}(t) = \sum_{i=1}^{H} I(t \in \mathrm{Holiday}_i) \cdot \delta_i \tag{6}$$

The holiday effect $\mathrm{h}(t)$ is modelled as an indicator function that adds a constant effect for holidays.

Where I: Indicator function

$\mathrm{Holiday}_i$: Represents each holiday.

δ_i: effect of the i-th holiday

$$\mathrm{y}(t) = \mathrm{g}(t) + \mathrm{s}(t) + \mathrm{h}(t) + \beta \mathrm{X}(t) + \varepsilon_t \tag{7}$$

Where $X(t)$: Represents the Exogenous variables

β: vector of coefficients for these variables

While the rest of the variables are the same as used in Equation (3).

3.4.3 Seasonal Autoregressive Integrated Moving Average with eXogenous Factors (SARIMAX) model forecasting

SARIMAX has been used as the data shows seasonality, and using just ARIMA did not justify results as simple ARIMA only supports some of the time series components that this data possesses. As the smart grid battery data has been integrated with exogenous factors, i.e., weather variables, the best time series approach in this situation can be SARIMAX [53]. Cyclic encoding for an hour, day of week, and month has been used in addition to exogenous factors to improve the previous methodologies and to have the best possible results.

Except for the past 24 hours, saving for the testing set, the training set contains all data points. This division guarantees that the model is assessed on previously unknown data, improving the performance indicators' validity. Specific parameters are set in the model to accommodate autoregressive terms, differencing, moving average terms, and seasonality [63]. Exogenous features are added to the model to improve its prediction performance.

The Equation below shows the notation of the SARIMAX model used in the model. Table 4 shows the parameters used in the model training.

$$\Delta^{d}\Delta_{s}^{D} y_t = \alpha + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{i=1}^{P} \Phi_i y_{t-i\times s} + \sum_{j=1}^{Q} \Theta_j \varepsilon_{t-j\times s} + \beta X_t + \varepsilon_t$$

(8)

Where, $\Delta^{d}\Delta_{s}^{D} y_t$ = Differenced and seasonally differenced dependent variable at time t.

α = Constant term.

ϕ_i = Co-efficient for the autoregressive term for the non-seasonal component.

y_{t-i} = The i-th lag of the dependent variable.

Θ_j = Coefficients for moving averages terms of the non-seasonal component.

ε_{t-j} = The j-th lag of the error term for the non-seasonal component.

Φ_i = Co-efficient for the autoregressive term for the seasonal component.

Θ_j = Co-efficient for moving average term for the seasonal component.

s = Seasonal period.

β = Co-efficient for Exogenous variable(s).

X_t = Exogenous variable(s) at time t.

ε_t = Error term at time t.

Table 4 Parameters used in the training process of the SARIMAX model

Parameter Type	**Autoregressive Order**	**Integration Order**	**Moving Average Order**	**Seasonal Period**
Order	4	1	0	-
Seasonal Order	4	1	1	24

3.4.4 Holt-Winters Exponential Smoothing Model Forecasting

Particularly for data with seasonality and trend, the Holt-Winters Exponential Smoothing model goes beyond conventional exponential smoothing by including trend and seasonality components, making it a powerful tool for forecasting complicated time series data. The Holt-Winters model is chosen for its ability to capture and model the dataset's complex patterns of seasonality and trend. The model's ability to generate accurate short-term forecasts and ease of implementation make it a popular choice for energy consumption and generation forecasting [64]. The parameters for the model are shown in Table 5. The model's additive nature and seasonality show a linear trend and seasonality, making it suited for datasets with these components changing consistently [65]. The 24 seasonal periods demonstrate the model's flexibility to datasets with diverse seasonal patterns, reflecting the hourly data's underlying daily changes.

Equation (9) shows the Level, Equation (10) shows the Trend, and Equation (11) shows the seasonal component used.

$$l_t = \alpha(y_t - s_{t-P}) + (1 - \alpha)(l_{t-1} + b_{t-1}) \tag{9}$$

Where, α: Smoothing parameter for the level, between 0 and 1.

y_t: Observed values at time t

s_{t-P}: Seasonal Component time $t - P$, P is Seasonal Period.

$$b_t = \beta^*(l_t - l_{t-1}) + (1 - \beta^*)b_{t-1} \tag{10}$$

Where,

β^*: Smoothing parameter for the trend (between 0 and 1).

$$s_t = \gamma(y_t - l_t) + (1 - \gamma)s_{t-P} \tag{11}$$

Where, γ: Smoothing parameter for the seasonality (between 0 and 1)

Table 6 shows the parameters used in the model training.

Table 5 Parameters Used for Training the Holt-Winters Exponential Smoothing Model

Parameter	**Value**
Trend	Additive
Seasonality	Additive
Seasonal Periods	24

3.4.5 XGBoost Model Forecasting

XGBoost (Extreme Gradient Boosting) is a well-known machine learning method that is particularly well-liked for its performance and quickness. It is a fast and efficient implementation of gradient-boosted decision trees. Because of its capacity to handle multiple forms of structured data well, XGBoost can be especially useful in time series forecasting. It can take missing values and help with tree-building parallelisation [66,69].

The XGBoost model predicts power generation by first dividing it into training and testing sets. Based on the mean absolute error criteria, the model is subjected to hyperparameter tuning using GridSearchCV to determine the ideal parameters [70] that produce the most outstanding performance. The ideal parameters are then utilised on the training data to train the XGBoost model. Exogenous characteristics such as temperature, humidity, wind speed, cloud cover, and cyclical encoding of time variables such as hour, day of the week, and month are used to train the model [68]. The trained model is then used to forecast data for the following 24 hours.

XGBoost optimises a composite objective function [71] consisting of a loss function and a regularisation term. The objective at iteration t is represented as:

$$\mathrm{Obj}^{(t)} = \sum_{i=1}^{n}\left(y_i - \left(\widehat{y_i^{(t-1)}} + f_t(x_i)\right)\right)^2 + \lambda\sum_{j=1}^{T} w_j^2 + \alpha\sum_{j=1}^{T}|w_j| \qquad (12)$$

Where y_i: Actual Value

$y_i^{(t-1)}$: Predicted Values at Iteration $\mathrm{t} - 1$

$f_t(x_i)$: Decision Tree Output at iteration t

λ, α: Regularisation Parameters

w_j: Weight of j-th leaf in the tree.

T: Number of leaves in the tree.

Further, The XGBoost in this approach uses first and second-order derivatives (Gradient and Hessian) for optimisation. For a given loss function l, the gradient and Hessian are:

$$g_i = \partial_{\widehat{y^{(t-1)}}} \mathrm{l}\left(y_i, \widehat{y_i^{(t-1)}}\right) \tag{13}$$

Where, g_i = Gradient of the loss function l with respect to the predicted value.

$\widehat{y_i^{(t-1)}}$ at time t-1 for the i-th observation.

The gradient g_i is vital in optimisation methods because it indicates the direction and size of the sharpest rise in the loss function and is used to update model parameters during training.

$$h_i = \partial^2_{\widehat{y^{(t-1)}}} \mathrm{l}\left(y_i, \widehat{y_i^{(t-1)}}\right) \tag{14}$$

Where, h_i = Hessian or second order derivative of loss function l

$\widehat{y_i^{(t-1)}}$ at time t-1 for the i-th observation.

The Hessian offers information on the loss function's curvature, illustrating how the gradient varies when the model parameters are changed. This is especially essential in optimisation algorithms that consider the gradient's rate of change, such as second-order optimisation techniques.

Whereas the score of Tree Structure Q in this model is given by:

$$\text{Score}(Q) = -\frac{1}{2}\sum_{j=1}^{T} \frac{G_j^2}{H_j + \lambda} \tag{15}$$

Where G_j: Sum of Gradients in j-th leaf.

H_j: Sum of Hessians in the j-th leaf.

λ: L2 Regularization parameter.

Table 6 shows the parameters used in the model training.

Table 6 Parameters used in the Hyperparameter Tuning and Training of XGBoost Model

Parameter	**Value**
Max depth	3, 5, 7
Number of Estimators	50, 100, 200
Learning Rate	0.01, 0.1, 0.2
Cross-Validation	3-fold
Scoring Metric	Negative Mean Absolute Error

3.4.6 Long Short-Term Memory (LSTM) Model Forecasting

Deep learning models are used because of their capacity to predict complicated, non-linear connections in data, which is especially useful for time series forecasting. Deep learning, such as LSTM, can grasp subtle patterns and temporal relationships [68, 72] in the context of forecasting values, providing improved accuracy. These models extract essential characteristics automatically, reducing the need for human feature engineering, and can handle massive amounts of data, enhancing prediction precision and dependability.

The basic concept underlying LSTM is a memory cell that stores essential information throughout the sequence's processing and non-linear gating units that govern information flow in the cell. Long-term temporal associations can be recorded by memory, and the effects of short-term memory can be decreased, i.e., information from earlier time steps can find its way to later time steps. LSTMs comprise memory cells and gates that govern information flow, avoiding the disappearing and exploding gradient problems plaguing standard RNNs [72]. LSTMs are notable for their capacity to capture and represent complex patterns and temporal relationships in sequential data, providing better performance and accuracy in various applications. They are trained to utilise backpropagation across time [73].

In this case, LSTM estimates the batteries' power when exogenous factors, such as weather data, are introduced. The LSTM model may discover patterns in data that are too complicated or have too many sequential relationships for simpler models.

An LSTM cell has three gates, input, forget, and output, and two states: cell state and hidden state. These gates and states have the following equations:

The forget gate is shown in Equation (16), the Input Gate in Equation (17), the Candidate Layer in Equation (18), the Cell State Update in Equation (19), and the Output Gate in Equation (20).

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{16}$$

The activation vector for forget gate at time t. It decides the information to discard from the state of the cell using the previous hidden state h_{t-1}, weight W_f, input x_t, and bias b_f.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{17}$$

The activation vector for the input gate at time t. It's responsible for updating the cell state by controlling the flow of new information, using the previous hidden state h_{t-1}, weight W_f, input x_t, and bias b_f.

$$\widetilde{C_t} = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{18}$$

The candidate vector for cell state at time t. It creates a vector of new candidate values that can be added to the state, h_{t-1}, and the current input x_t.

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C_t} \tag{19}$$

The cell state at time t. This is an element-wise combination of the previous state and the new candidate state formulated by the forget gate f_t and input gate i_t.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{20}$$

The activation vector for the output layer at time t. It formulates the next hidden state and output of the LSTM unit by using the up-to-date cell state and previously hidden state h_{t-1} and current input x_t.

The parameters used for LSTM are shown in Table 7.

Table 7 Parameters Used in the Training of the LSTM Model

Parameter	**Value**
Input Shape	(1,30)
LSTM Units	100
Dense Units	4
Loss Function	Mean Squared Error
Optimizer	Adam
Batch Size	10
Epochs	100

3.5 Hybrid Time Series Model

The hybrid model, which combines the SARIMAX, Holt-Winters, and Prophet models, exemplifies the constructive interaction of many forecasting methodologies, each providing unique characteristics to improve overall predicted accuracy. SARIMAX excels in capturing complicated seasonal patterns and autocorrelations, which may be attributed to its robust parameterisation and incorporation of exogenous factors [74]. However, its effectiveness can be limited by non-linear patterns and numerous seasonalities, which can be effectively handled using the Prophet model.

Prophet, a forecasting tool created by Facebook, is well-known for its versatility and flexibility. It is designed to support various seasonalities, holiday impacts, and special events, providing a comprehensive solution for a wide range of time series data [75]. Despite its adaptability, Prophet can be computationally demanding and may only sometimes yield the best match, particularly in instances with complex data patterns and anomalies. The Holt-Winters model complements the hybrid ensemble's capacity to capture level, trend, and seasonality. While it excels at detecting systematic patterns in data, its performance might suffer when dealing with several seasonal patterns or non-linear trends [76].

Combining these models in the hybrid method is a complex integration that uses machine learning and statistical insights to improve forecast accuracy. The output of each model is weighted depending on its predictive ability, ensuring that the limits of any single model do not influence the overall prediction. The hybrid model includes a method for adapting to changing data patterns. It has a self-learning algorithm that recalibrates model parameters in real-time, guaranteeing that projections are current with the latest data trends [77]. This flexibility is critical given the volatile landscape of power demand, which is characterised by swings caused by behavioural, meteorological, and economic factors.

The Equation below shows the working of the Hybrid Time Series Model:

$$y_{\widehat{t,\mathrm{Hybrid}}} = \frac{y_{\widehat{t,\mathrm{Prophet}}} + y_{t,\widehat{\mathrm{SARIM}}\mathrm{AX}} + \widehat{y_{t,\mathrm{HW}}}}{3} \tag{21}$$

Table 8 consolidates the parameters of the SARIMAX, Holt-winters, and Prophet models combined to create the hybrid forecasting model.

Table 8 Parameters Used for the Training of Hybrid Model

Model	Parameter	Description	Value
Prophet	Daily Seasonality	Whether to include Daily Seasonality	True
SARIMAX	Order	The order of Autoregression, Integration, and Moving Average components	(4,1,0)
	Seasonal Order	Order of seasonal components	(4,1,1,24)
	Disp	Whether to display Convergence Statistics	False
Holt-Winters	Trend		Additive
	Seasonal		Additive
	Seasonal Periods	Number of periods in a complete seasonal cycle	Depending on the number of hours of forecast desired, i.e., 1,2,3,12,24,36

3.6 Error Evaluation Metrics

The error metrics are critical for analysing machine learning models to understand better where a particular model excels and falls short compared to other models. The models are evaluated using the four metrics listed below.

Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y_i}| \tag{22}$$

Mean Squared Error (MSE)

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2 \tag{23}$$

Root Mean Squared Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2} \tag{24}$$

Mean Absolute Percentage Error (MAPE)

$$\text{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \widehat{y_i}}{y_i}\right| \tag{25}$$

Chapter 4 Results

This chapter explains the results of the methodology involved in this research. The results include the Exploratory Data Analysis (EDA) for the two datasets to learn about the underlying pattern in the data and how to make them helpful for this research. Further, different machine learning models ranging from SARIMAX to LSTM are evaluated and briefly discussed.

4.1 EDA Monthly Analysis

Figure 4-1 below depicts the power output of smart grid batteries in kilowatts (kW). Power production varies significantly between September 2022, ranging from roughly 25 kW to a maximum above 175 kW. Power-generating spikes may correspond to increased energy demand or availability periods, whereas troughs imply lesser power production.

Peaks in power output may occur due to increasing sunshine if the smart grid is supplemented by solar power (as the source of smart grid energy storage is unclear from the dataset's documentation), resulting in higher energy production during the day. Troughs, particularly at night, might be caused by a lack of solar input, suggesting a dependence on stored energy or other generating sources.

The temperature graph in Figure 4-2 depicts daily temperature variations in degrees Celsius, ranging from the lower 20s to well beyond 30°C. The results show a continuous cyclical pattern, which most likely reflects the daily temperature fluctuation from day to night. Warmer days are followed by colder nights, with occasional intervals of higher or lower average temperatures.

Higher daytime temperatures boost the efficiency of solar panels if they are connected to the smart grid, perhaps corresponding with higher kPt values. Cooler night-time temperatures can reduce power consumption for cooling systems, resulting in lower kPt values at night.

Figure 4-3 displays the daily humidity levels for the same one-month period, given as a percentage. Humidity levels fluctuate between 40% and 90%, reflecting various atmospheric moisture conditions. Regular rhythmic patterns indicate a predictable variation in humidity, which daily weather variations or local climatic circumstances may influence.

Higher humidity, particularly in the early morning or late evening, can be associated with lower temperatures, implying a reduced efficiency in power generation or a higher power demand due to the discomfort induced by high humidity.

Figure 4-4 depicts the variation in wind speed measured in kilometres per hour over a month, from September 1 to October 1, 2022. It depicts a varying pattern of wind speed with periodic peaks indicating gusts or windy situations. The lowest wind speeds are near zero kilometres per hour, indicating calm times, while the most significant peaks exceed 40 kilometres per hour, indicating the presence of severe winds. This information is crucial for understanding weather patterns and might be used to improve the performance of wind-dependent systems or activities.

Higher windspeeds imply that if wind turbines are integrated, more electricity may be generated during windy circumstances, corresponding with higher kPt values. In contrast, calm times may diminish wind-generated power, resulting in lower kPt values.

Specific days with considerable differences can be found in the graphs by searching for significant deviations from the typical trends. For instance, wind energy contributes to the system, so a day with exceptionally high windspeeds coincides with increased electricity generation. A colder day results in less energy for cooling, resulting in reduced power output, as shown in Figure 4-1.

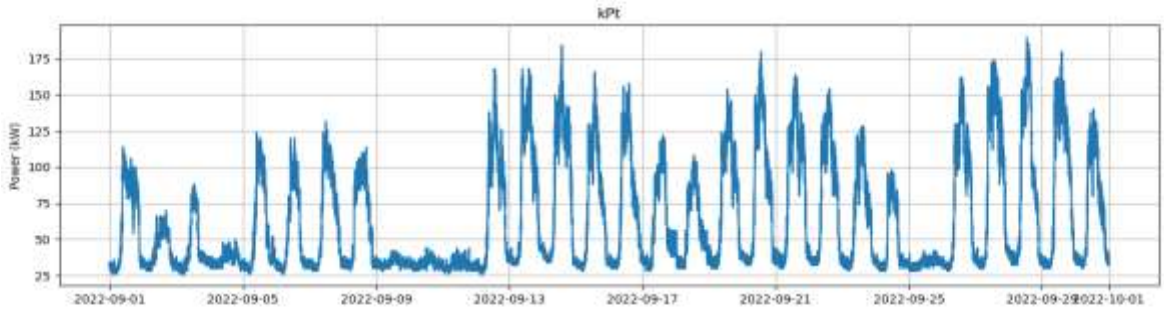

Figure 4-1 Power generated from smart grid batteries (in unit kilo) for the month of September 2022

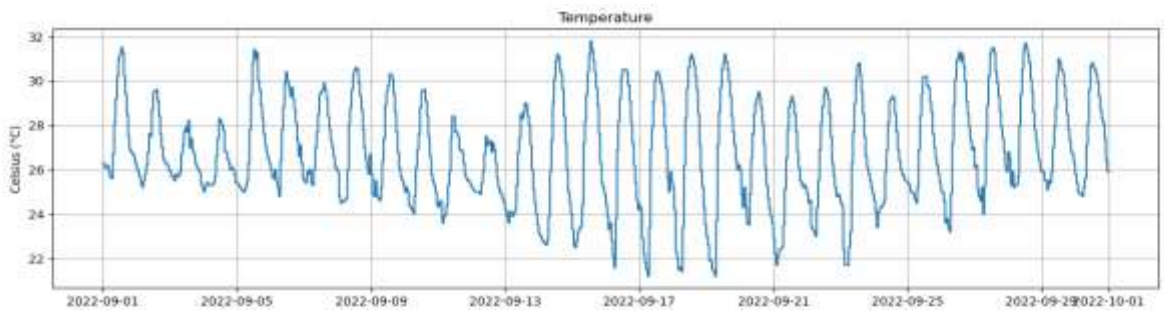

Figure 4-2 Temperature patterns from 1st September to 30th September 2022 for the local area of Boshan campus in Taiwan

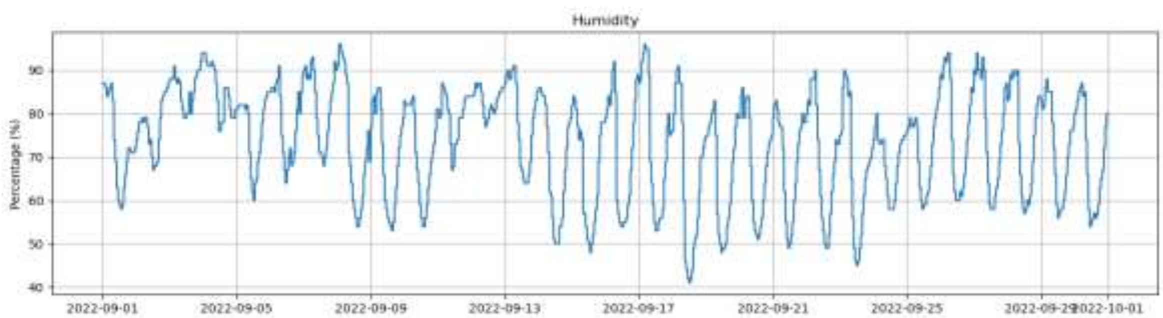

Figure 4-3 Humidity patterns from 1st September to 30th September 2022 for the local area of Boshan campus in Taiwan

Figure 4-4 Humidity patterns from 1st September to 30th September 2022 for the local area of Boshan campus in Taiwan

4.2 Comparative Analysis for Meteorological Parameters

Figure 4-5 below combines multiple sub-plots for the weather variables from the specific area of Taiwan's Boshan campus. Each sub-plot shows the trend of the weather conditions throughout September 2022. The smart grid data was only a month long, so this weather data is also sourced for just one month only (briefly explained in the 'Data Acquisition' chapter in chapter 3). For ease, the figure's captions are written to show the unit in which the variables are presented, i.e., 'Temperature 2m (C)' states that the temperature variable is in degrees Celsius. The temperature is recorded for 2 meters above the ground level. Other variables are also presented in a similar way, which is also outlined in Chapter 3.

This first subplot displays daily temperature cycles, with greater daytime temperatures and lower night-time temperatures as predicted. The magnitude of the temperature fluctuations indicates a significant variation between day and night temperatures, which may be typical of the climate of the geographical area throughout September.

The second sub-plot shows the relative humidity levels, which generally behave inversely to temperatures, with higher humidity at night and lower humidity during the day. This is to be expected because the ability of the air to contain moisture varies with temperature.

The third sub-plot represents the temperature individuals feel, considering humidity and wind speed. The sub-plot typically follows the temperature pattern, but some variances indicate the effect of moisture and wind on how warm it feels.

The fourth sub-plot shows the occurrence of precipitation. Peaks signify rainfall events, and their influence on temperature and humidity is frequently seen as transient temperature decreases and humidity rises.

The fifth sub-plot shows the fraction of the sky covered by clouds. A higher cloud cover frequently corresponds with lower daytime temperatures (because less solar energy reaches the surface) and higher night-time temperatures (due to the cloud's insulating effect).

Meanwhile, the last sub-plot tracks the wind speed, which affects the perceived temperature. Due to the wind chill effect and evaporative cooling, higher wind speeds increase the impression of chilly when temperatures are low and decrease the feeling of heat when temperatures are high.

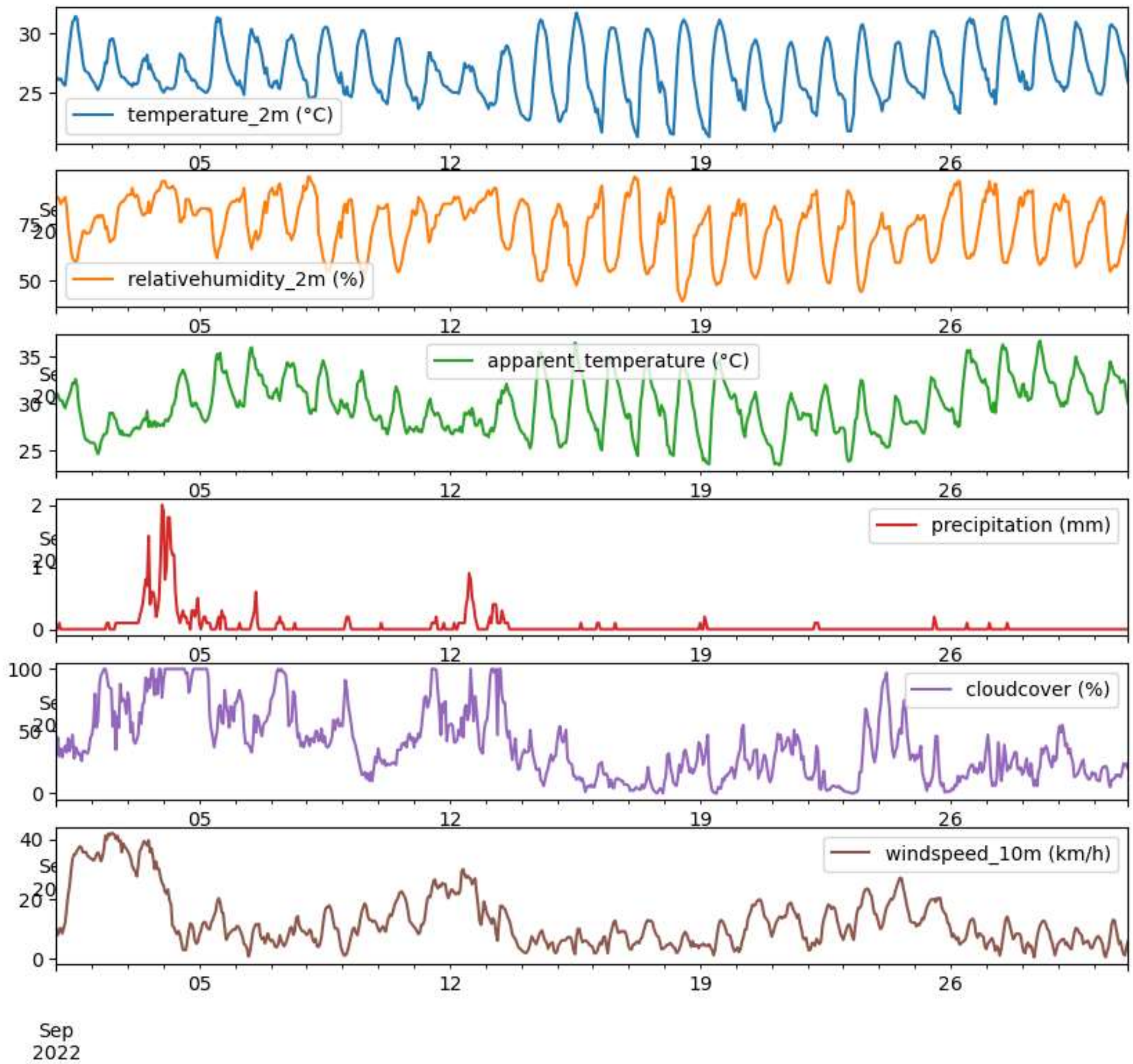

Figure 4-5 Comparative Analysis of Weather Variables for the Month of September 2022

4.3 Hypothesis Testing

Hypothesis testing is a fundamental statistical approach for determining the likelihood of a particular hypothesis regarding a data set being true. It begins with an initial assumption known as the null hypothesis (H0), which often states no impact or difference in the population. An alternative hypothesis (H1) is given as a counterclaim, implying an effect or difference [58].

This result is based on a hypothesis testing framework in which the null hypothesis states that the time series has a unit root. This statistical quality indicates non-stationarity and the presence of time-dependent structures in the data. The alternative hypothesis contends that the time series has a unit root and is stationary, showing no time-dependent structure.

When the ADF test on the time series data was run, the following results were obtained:

- For the regularly used confidence levels (1%, 5%, and 10%), the test statistic did not fall below the critical values. This result implies that the null hypothesis of the presence of a unit root cannot be rejected.
- The p-value was more than 0.05, suggesting that the null hypothesis cannot be discarded [43].

From Figure 4-6, the rolling mean shows a pattern by averaging the data across time and smoothing out short-term variations. It varies over the month, showing that the average time series value changes. The rolling standard deviation depicts the shifting volatility of the data; higher values imply more significant variability within the rolling frame.

Periodic peaks and troughs in the original data show a cyclical component within the time series. The rolling mean and standard deviation do not remain constant, indicating that the time series is non-stationary — that is, the series statistical features, i.e., mean and variance, do not remain constant across time.

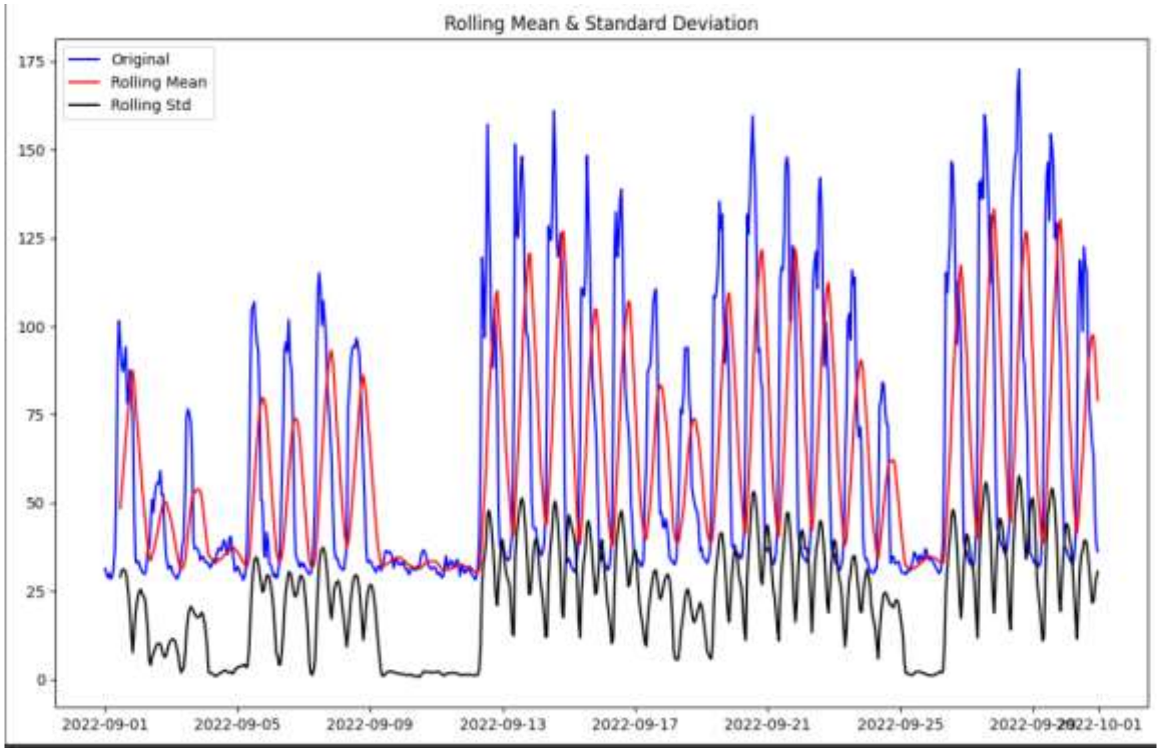

Figure 4-6 Rolling Mean and Standard Deviation

4.4 Daily Sum of Power Generated from the Smart Grid Batteries

Figure 4-7 shows the daily sum of power generated by the batteries. This figure can cause some confusion as it was observed in the EDA of the variables and especially kPt that the power generated was in some kilowatts ranging from not more than 200 per day. But this sum shows a lot of power generation in batteries daily. But the reason for this inconsistency is that the data is just for one month, and there is no clue for the power generation data before and after this very month; it can be said that much power is already available in the batteries before September 2022. Still, this analysis aims to determine if the power generation sees a positive or negative trend. This shows that it follows an upward trend, minimising the chance of blackouts until a sudden surge in power demand.

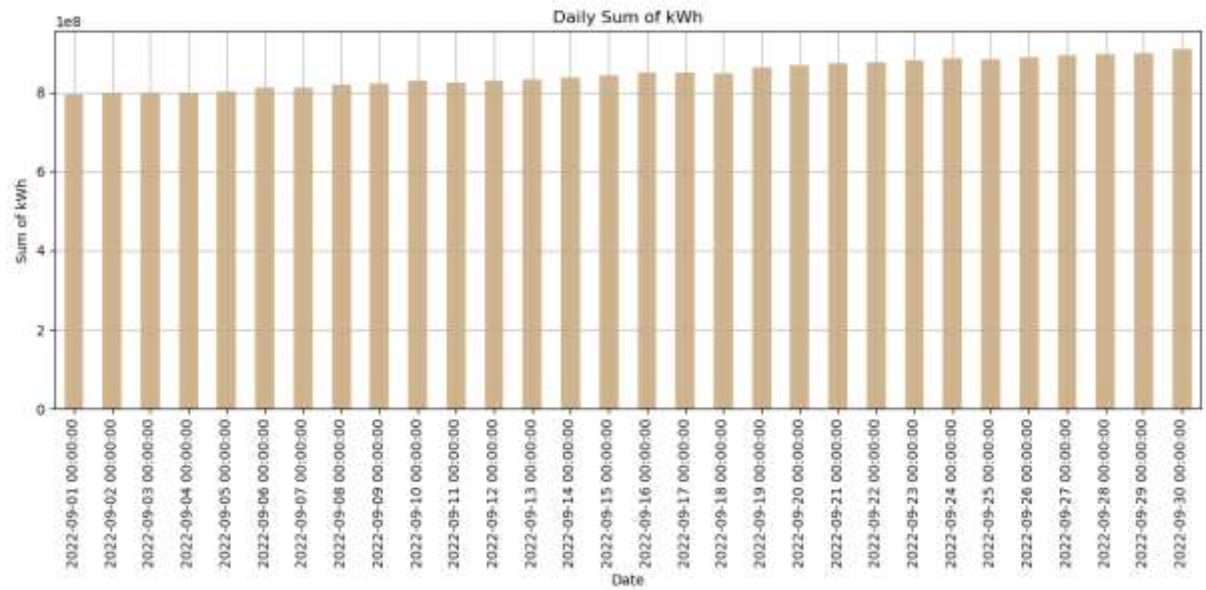

Figure 4-7 Daily Sum of energy stored in smart grid batteries

4.5 Evaluation of the Facebook Prophet Model (On Smart Grid Data Only)

The model has been trained (using training data), and its performance is tested using a different set of data (using testing data), a standard approach for validating the robustness of prediction models.

The Prophet model effectively caught the cyclical patterns within the training data set, as seen by the anticipated output's unity with the known test data. This indicates that the model is sensitive to the intrinsic periodicities in power generation data, such as daily or weekly cycles. The tight alignment of the anticipated data (green dashed line in Figure 4-8) with the test data suggests that Prophet's components, including trend, seasonality, and holidays, effectively model the power generation time series.

The prediction for the next 24 hours (figure 4-9) shows predicted generation and includes a measure of uncertainty. Based on previous trends, the model anticipates a specific pattern of power output, which is altered by factors such as daylight hours if the power source is solar. The breadth of the confidence interval in the exact figure provides an estimate of variability, which can be impacted by unmodeled causes or intrinsic instability in power generation. A narrower confidence interval implies better predictability in power output at specific periods. In contrast, a more extensive range may indicate more significant uncertainty, potentially owing to less predictable variables such as weather conditions impacting solar or any renewable energy generation.

When smart grid power generation data is anticipated using the Prophet model, the analysis gives a more detailed knowledge of the model's prediction accuracy and possible consequences for energy management using error evaluation metrics in Table 9. The MAE of 26.37 implies that the model's predictions differ from the actual kilowatt-hour production by this value on average. At the same time, this degree of error is tolerable for large-scale operations. The Mean Squared Error (MSE) and Root Mean Squared Error (RMSE), at 1197.91 and 34.61, respectively, indicate that the forecast deviates considerably from the actual generation numbers.

The Mean Absolute Percentage Error (MAPE) of 0.69% shows good relative accuracy, indicating that the model effectively captures the general trend and power generation amount. This low percentage error is encouraging for smart grid operating efficiency, as predicted accuracy is critical for maintaining supply and demand balance, scheduling maintenance, and guaranteeing stability.

Table 9 Error evaluation for FBProphet Model on smart grid data only for forecasting the next 24 hours of power generation

Error Type	**Value**
Mean Absolute Error	26.37
Mean Squared Error	1197.91
Root Mean Squared Error	34.61
Mean Absolute Percentage Error	0.69

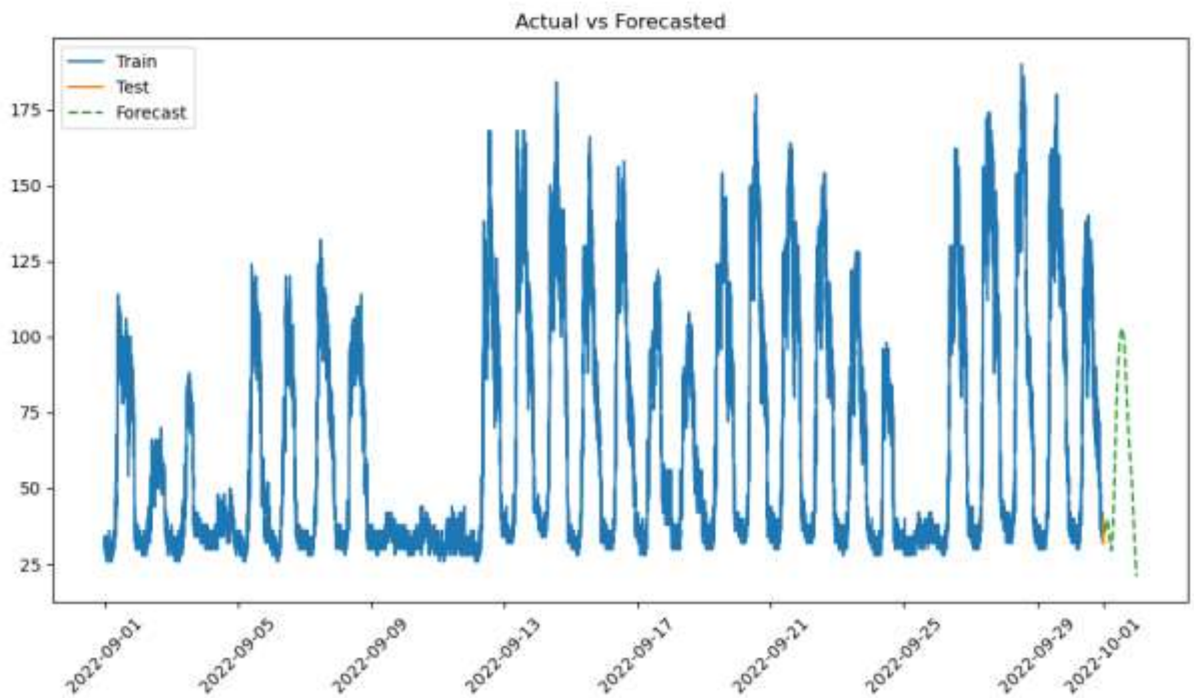

Figure 4-8 FBProphet Model Forecast (Actual vs. Predicted) on Grid Data (Next 24 Hours)

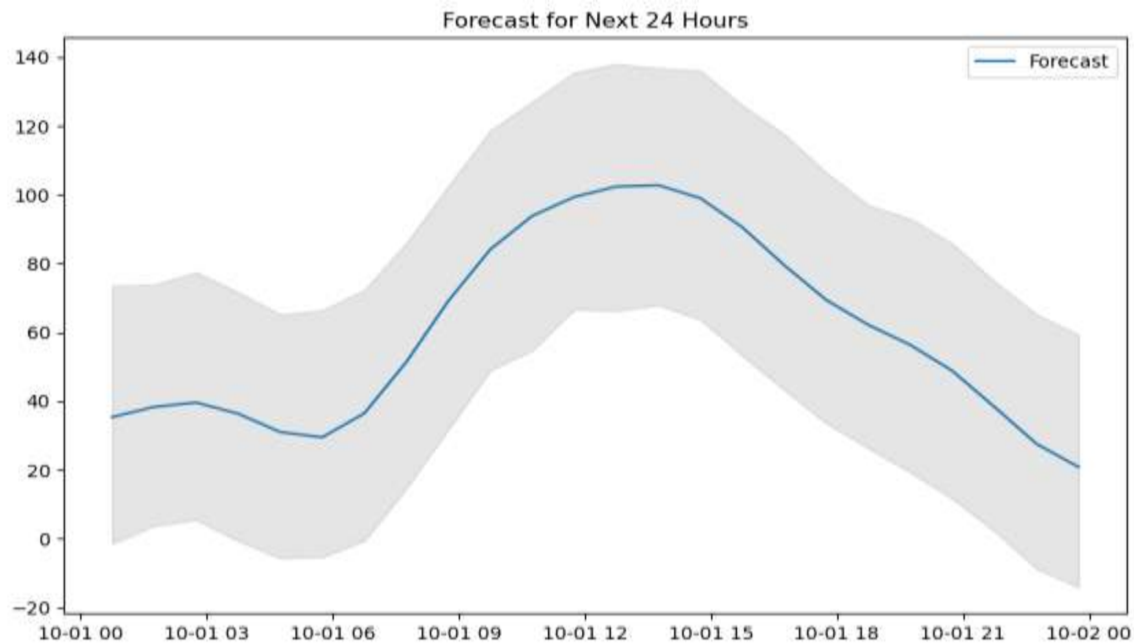

Figure 4-9 FBProphet Forecast for the Next 24 Hours

4.6 Evaluation of the SARIMAX model

The SARIMAX model forecasts the next 24 hours of data following successful training. The projection considers the testing sets exogenous properties to achieve accuracy. The two figures (4-10 and 4-11) show how a SARIMAX (Seasonal Autoregressive Integrated Moving Average with eXogenous variables) model predicts kilowatt-hour (kPt) production for a smart grid system.

The absolute kPt values are presented with the SARIMAX predicted values across multiple days in Figures 4-10. The actual data displays substantial changes, with peaks and troughs indicating fluctuating power production due to demand cycles or variations in renewable energy supply. The SARIMAX forecast, shown by a dashed line with cross symbols, matches these actual values closely, with a shaded forecast region reflecting the confidence interval around the estimates. The overlap between anticipated and actual values indicates that the model successfully captured time series dynamics up to the end of the observed data.

Figure 4-11 focuses on a shorter period and compares actual and anticipated kPt values more thoroughly. The projection closely resembles the data trend, accurately tracking the rise and decrease in power production.

The error metrics presented provide a quantitative evaluation of the model's prediction performance in Table 10. The error metrics for the SARIMAX model's forecast of smart grid power generation show predictive solid performance: a Mean Absolute Error of 3.54 indicates that the model's forecasts are only 3.54 kilowatt-hours off from the actual figures on average, indicating a high level of forecast precision. Although the Mean Squared Error is moderate at 22.97, it indicates the presence of specific predictions with significant errors due to the squaring of error factors. This is supported by a Root Mean Squared Error of 4.79, indicating that the model's forecasts may occasionally deviate from reality. Nonetheless, a Mean Absolute Percentage Error of 4.92% indicates a relatively good model, with projections often departing from actual values by less than 5%.

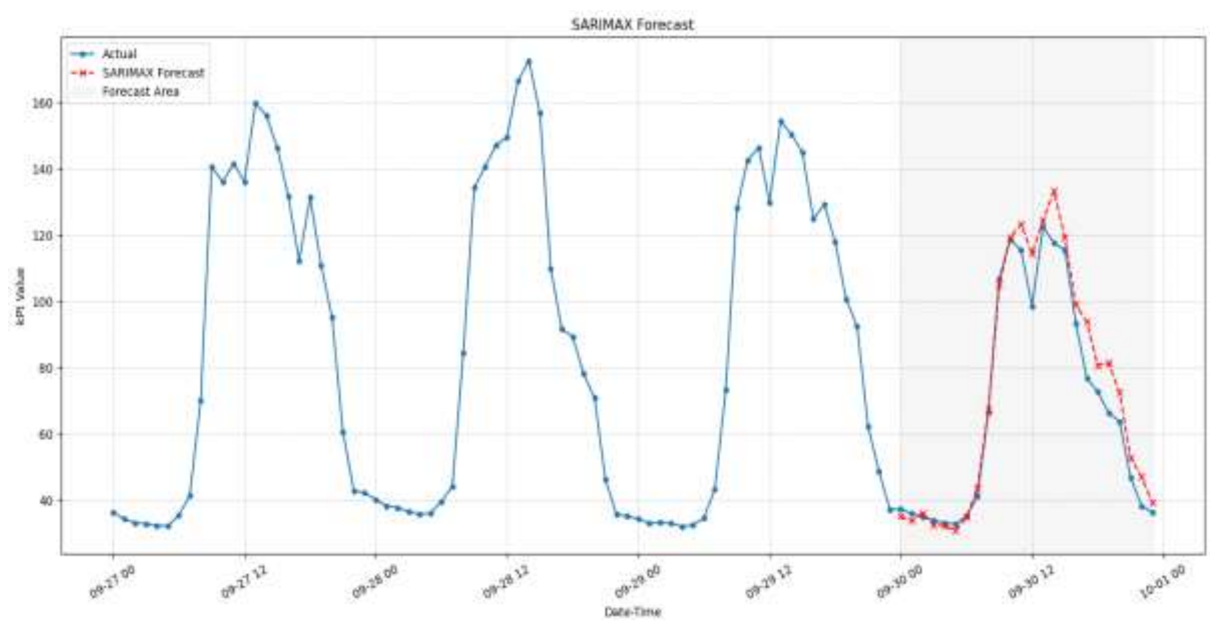

Figure 4-10 SARIMAX forecasting for the next 24 hours.

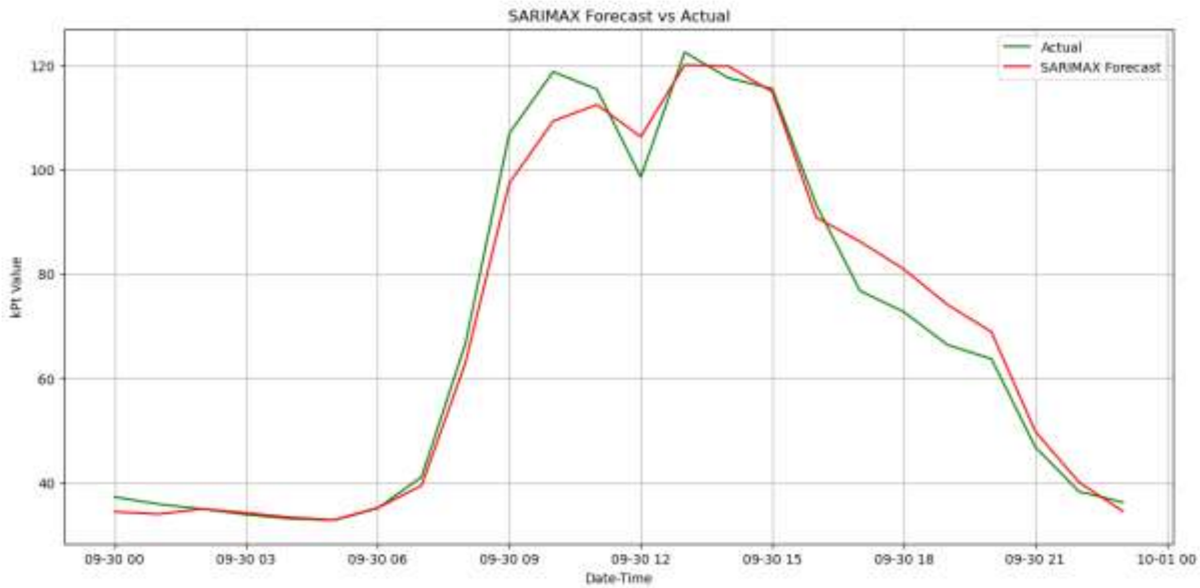

Figure 4-11 SARIMAX fit on the data for the next 24 hours (Actual vs. Forecast)

Table 10 Error evaluation for the SARIMAX model for the next 24 hours

Error Type	**Value**
Mean Absolute Error	3.54
Mean Squared Error	22.97
Root Mean Squared Error	4.79
Mean Absolute Percentage Error	4.92

4.7 Evaluation of FBProphet Model (Joined with Weather data)

The two charts (4-12 and 4-13) and error metrics (table 11) are related to the use of the Prophet forecasting model in estimating kilowatt-hour (kPt) generation for a smart grid system, keeping in view the metrological data too. Figure 4-12 shows the actual and anticipated kPt values over a month, with the Prophet forecast closely mirroring the observed values. Though there are instances of substantial variation, the prediction line reflects the rhythm and peaks of power output.

Figure 4-13 charts break down the prediction into its constituent parts: trend, weekly seasonality, yearly seasonality, daily patterns, and various regressors that may account for holidays or other unique occurrences. The trend component exhibits a declining pattern during the month, which might indicate a seasonal decrease in power generation or an underlying negative tendency in the data. Weekly seasonality follows a consistent pattern, suggesting regular variations within days of the week. Yearly and daily components reveal the effect of longer-term cycles and within-day patterns on electricity generation.

A Mean Absolute mistake of 6.17 indicates that the forecast deviates from the actual values by this amount on average, which, given the kPt value scale, may reflect a considerable level of mistake. The Mean Squared Error of 62.9, a higher number owing to error squaring, indicates that more significant errors have a greater influence on the model's performance. Substantial forecast errors are confirmed by a Root Mean Squared Error of 7.93, more than the MAE. The Mean Absolute Percentage Error of 9.22% indicates that the model's predictions are often within 9.22% of the actual values. This shows that the model is moderately accurate in relative terms but might benefit from more accuracy for more precise energy management methods.

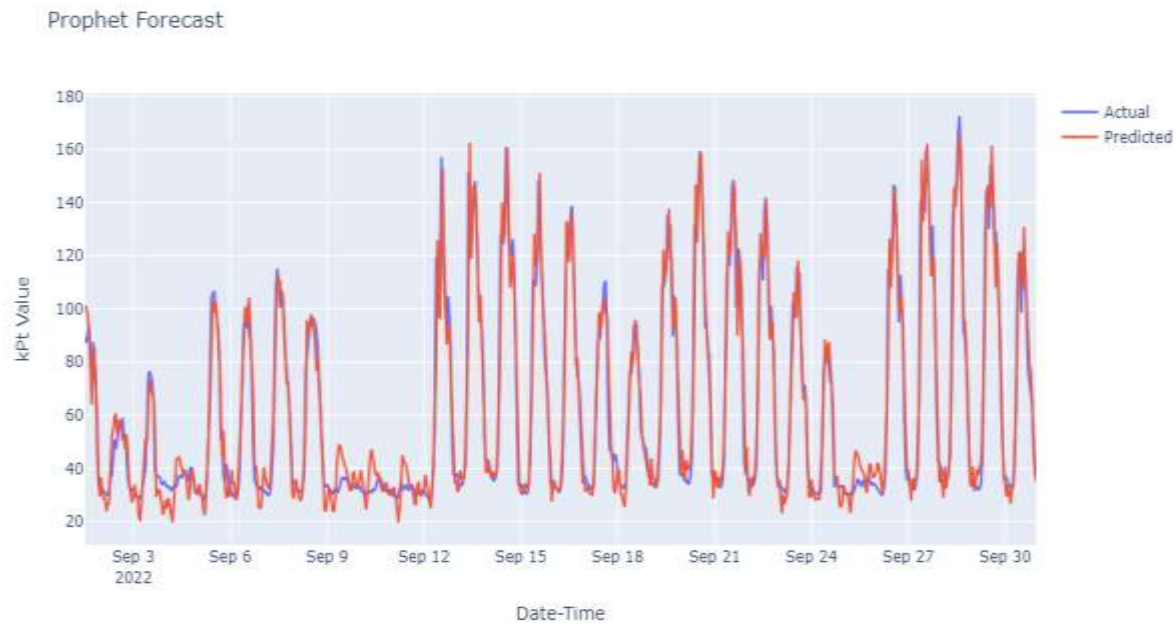

Figure 4-12 Prophet Forecasting (Actual vs. Predicted)

Table 11 Error evaluation for Facebook Prophet Model for the next 24 hours trained on smart grid and weather data

Error Type	**Value**
Mean Absolute Error	6.17
Mean Squared Error	62.9
Root Mean Squared Error	7.93
Mean Absolute Percentage Error	9.22

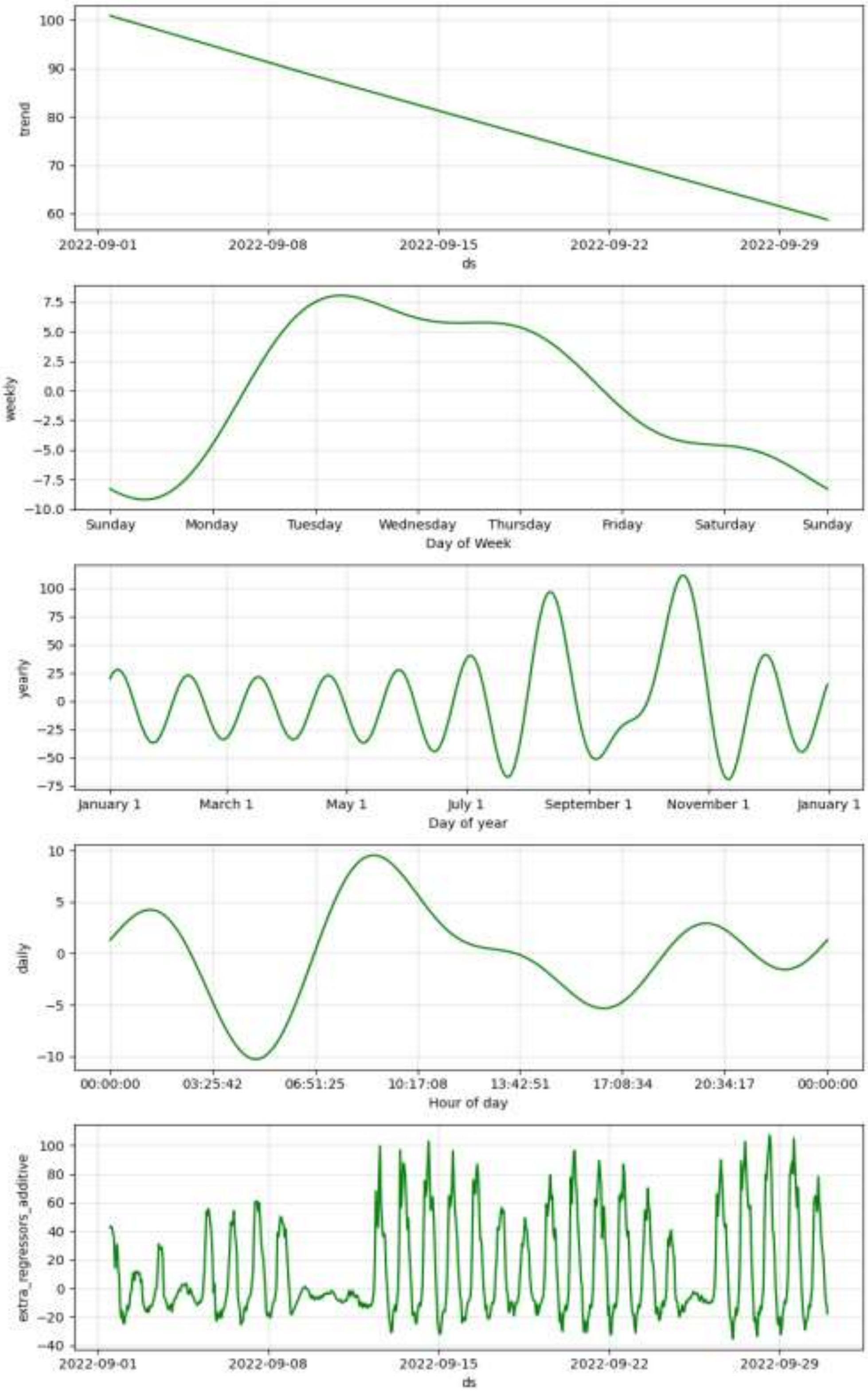

Figure 4-13 Forecast plots for Smart Grid Power Generation

4.8 Evaluation of Holt-Winters Exponential Smoothing Model

The performance of a Holt-Winters Exponential Smoothing model applied to smart grid power generation data and weather variables is depicted in Figures 4-14 and 4-15, and the accompanying error metrics are in Table 12. Because it extends exponential smoothing to capture seasonality in addition to trends, the Holt-Winters approach is particularly well-suited to time series data having a seasonal pattern.

The actual power-generating data, shown as a solid blue line in Figure 4-14, demonstrates a distinct periodic pattern, typical of power use cycles primarily impacted by consumer behaviour, industrial demand, or the availability of renewable energy sources. The orange line represents the model's forecast, which closely corresponds with the actual data during the training phase and extends into the test phase with green, indicating the model's ability to predict future values based on learned previous patterns. Figure 4-15 emphasises the model's predicting abilities, with projected values nearly mirroring the actual data, with occasional variations indicated. The projected values represent the model's effort to match the accurate kPt generation data's cyclical peaks and troughs.

Also, a Mean Absolute Error of 5.16 means that the model's predictions differ from the actual values by an average of 5.16 kPt, indicating a moderate degree of accuracy in the context of smart grid power production. The Mean Squared Error of 68.11, which is more than the MAE, suggests the presence of some significant errors due to the model's susceptibility to abrupt changes in power generation that the seasonal and trend components of the model do not capture. A Root Mean Squared Error of 8.25, in the same units as the power generation data, indicates the presence of more considerable errors while still indicating that the model is typically dependable in its predictions. Last, the Mean Absolute Percentage Error of 5.58% suggests that the model's estimates are usually within 5.58% of the actual data, sufficient for many practical applications in smart grid management.

Table 12 Error evaluation for Holt-Winters Exponential Smoothing Model

Error Type	**Value**
Mean Absolute Error	5.16
Mean Squared Error	68.11
Root Mean Squared Error	8.25
Mean Absolute Percentage Error	5.58

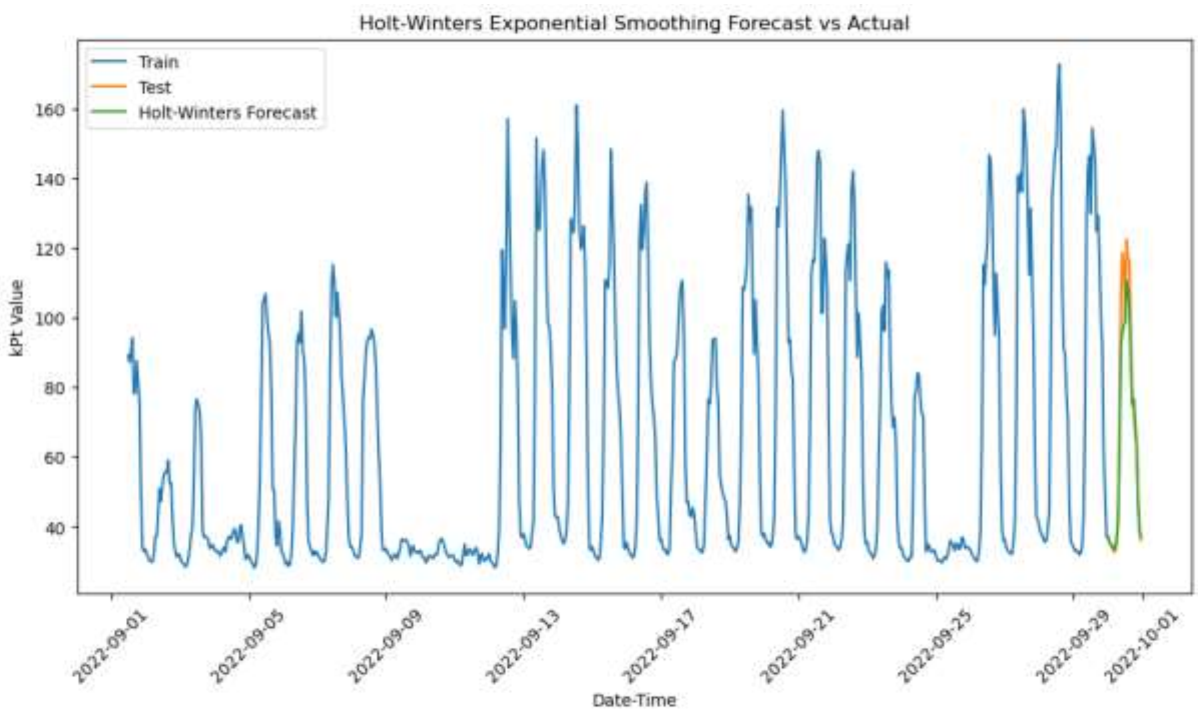

Figure 4-14 Holts-Winter Exponential Smoothing Model's forecasting for the next 24 hours

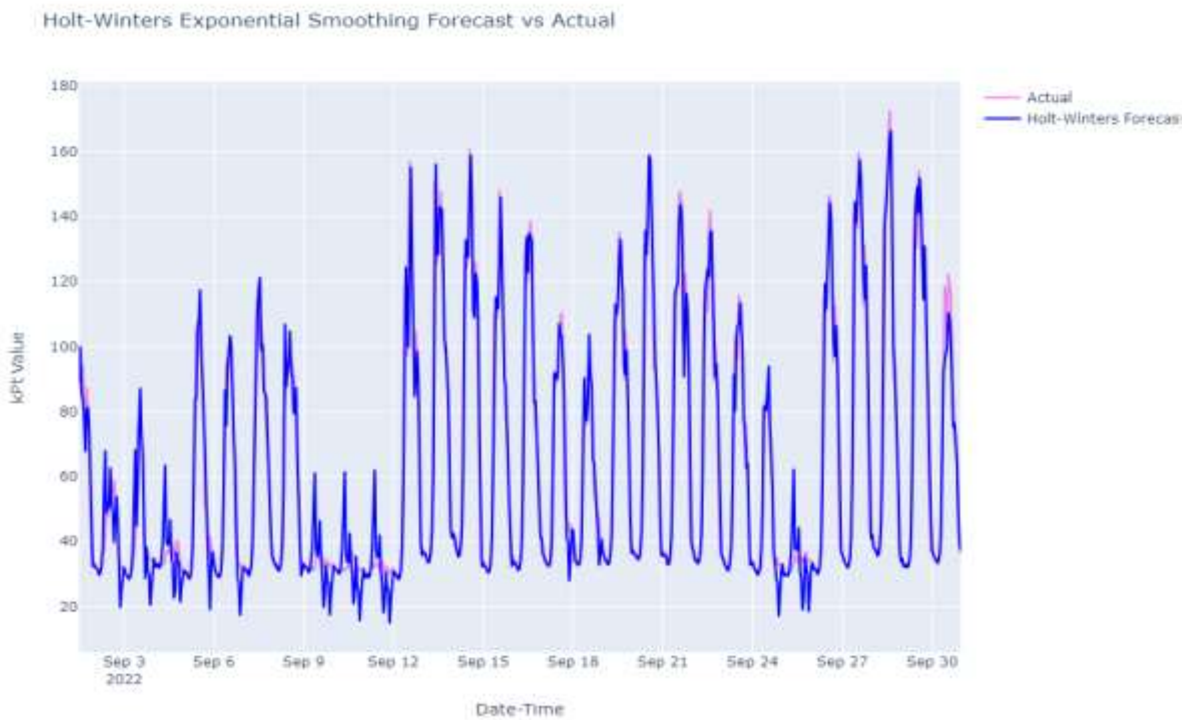

Figure 4-15 Holts-Winter Exponential Smoothing Model (Forecasting vs. Actual)

4.9 Evaluation of the XGBoost Model

The XGBoost model, as shown in Figures 4-16 and 4-17, has been used to anticipate power generation in a smart grid system, keeping the weather patterns in view. The projection in Figure 4-16 goes beyond the previous data utilised for training and testing. While the anticipated numbers may not fully match the actual data, they show some unity, particularly in reflecting power generation's overall pattern and periodicity.

Figure 4-17 further emphasises the model's efficacy, which compares the model's full month forecast to actual data. Despite minor variances, the anticipated pattern matches the actual data, indicating that the model has sufficiently internalised the cyclic behaviour of the power-generating process.

Error measures examine the model's predicted accuracy quantitatively, as shown in Table 13.

The model's average deviation from the actual data is moderate, as indicated by the Mean Absolute Error of 8.24, which may be acceptable within specified operating tolerances. A Mean Squared Error of 181.17 suggests the presence of several significant individual forecasting errors, emphasising potential outlier occurrences or model sensitivity concerns. The Root Mean Squared Error of 13.46, which is somewhat more significant than the MAE, indicates that these more considerable errors significantly influence model performance. A Mean Absolute Percentage Error of 9.78% represents a moderate relative error margin, indicating that while the model is usually effective at monitoring the data trend, accuracy might be improved in some areas.

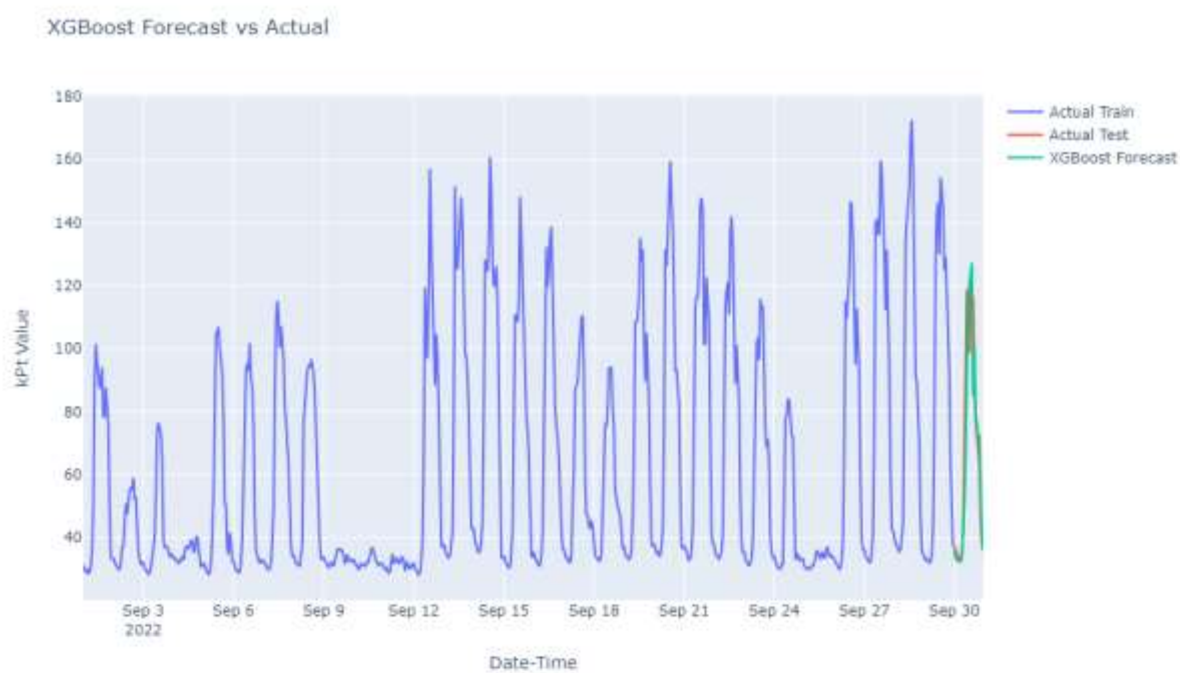

Figure 4-16 XGBoost Model's forecasting for the next 24 hours

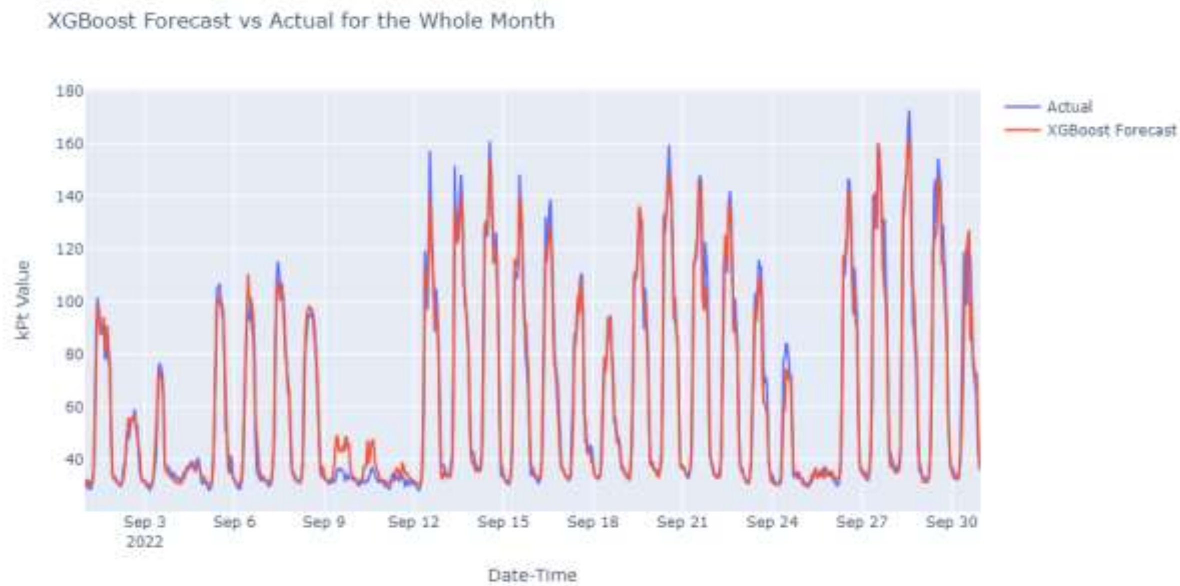

Figure 4-17 XGBoost Forecast vs. Actual data

Table 13 Error evaluation for XGBoost Model trained on the smart grid and weather data for forecasting the next 24 hours of power generation.

Error Type	**Value**
Mean Absolute Error	8.24
Mean Squared Error	181.17
Root Mean Squared Error	13.46
Mean Absolute Percentage Error	9.78

4.10 Evaluation of the LSTM Model

The LSTM model, a recurrent neural network well-suited for sequential data, predicts the purple kilowatt-hour (kPt) values alongside the blue actual kPt values. Figure 4-18 depicts the result of an LSTM (Long Short-Term Memory) model's power generation projection, keeping in view the weather variables, as compared to actual recorded data from the latter days of September 2022.

Forecasts from the model closely track actual kPt data, capturing peaks and troughs with excellent precision. However, there are certain anomalies when the forecast either underestimates or overestimates the actual numbers, which is especially visible during rapid fluctuations in power production.

The following error measures are used to quantify forecast performance shown in Table 14:

A mean Absolute Error of 5.06 suggests that the model's predictions differ from the actual values by an average of 5.06 kPt, suggesting a relatively tight forecast that roughly matches the actual results. Mean Squared Error of 62.92, this more significant error metric highlights the cumulative impact of forecast errors, emphasising the presence of specific projections that deviate significantly from actual values.

The Root Mean Squared Error, having a value of 7.93, which expresses the average error in the same units as the predicted variable (kPt), indicates that, while the model is typically trustworthy, there are times when its predictions considerably diverge from the actual data. According to the MAPE, the model's projections are within 6.91% of the actual values, indicating an acceptable degree of accuracy in terms of smart grid forecasting.

These insights from the performance of the LSTM model are critical for operational decision-making in a smart grid scenario. Despite its flaws, the model's capacity to estimate power generation with moderate accuracy may be instrumental in controlling supply and demand, scheduling maintenance, and optimising resource allocation. The LSTM model's low error rates imply that it might be a reliable tool for predictive analysis in smart grid operations, with additional tweaking possibly improving accuracy and dependability.

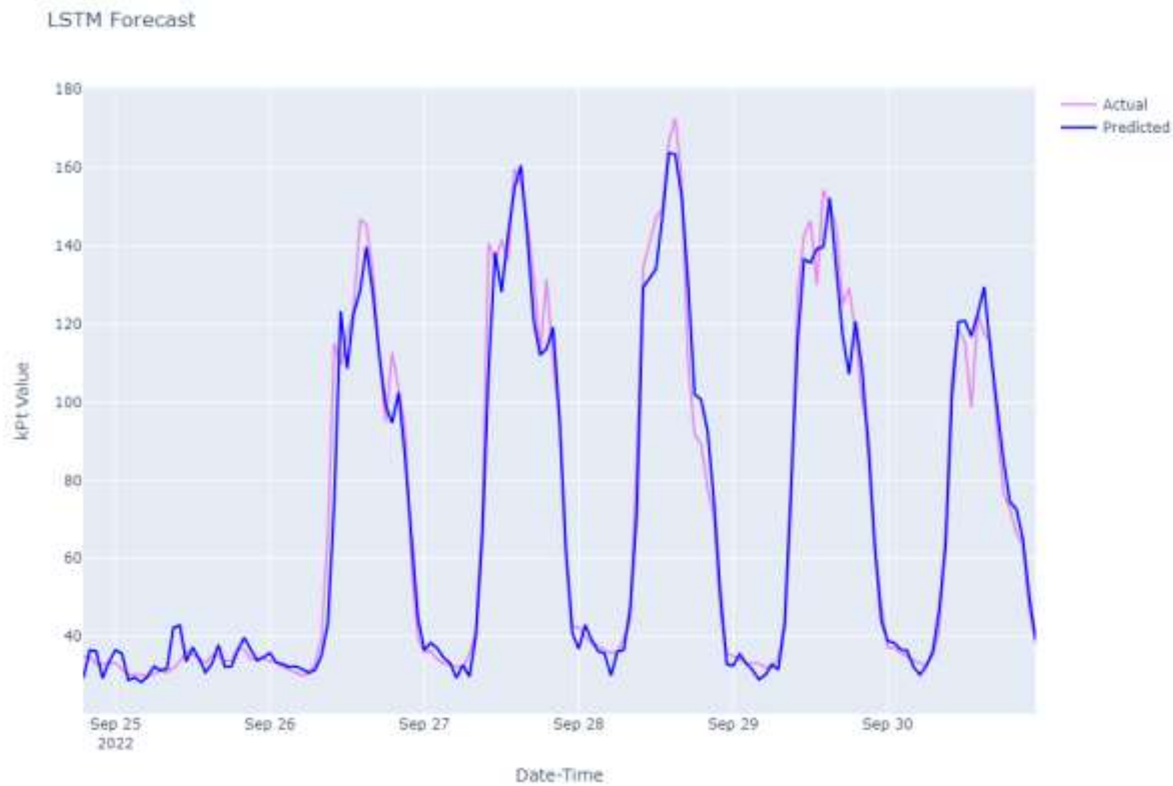

Figure 4-18 LSTM Forecast results (Actual vs. Predicted) for the next 24 hours trained on weather data and smart grid variables

Table 14 Error evaluation for Long Short-Term Memory (LSTM) Model combined trained on smart grid and weather data to forecast the power generation for the next 24 hours

Error Type	**Value**
Mean Absolute Error	5.06
Mean Squared Error	62.92
Root Mean Squared Error	7.93
Mean Absolute Percentage Error	6.91

4.11 Evaluation of Hybrid Time Series Model

The hybrid forecasting model represented in the pictures provides predicted insights for electricity generation in a smart grid over the next 12 and 24 hours. The model's performance is illustrated by its close tracking of real kilowatt-hour (kPt) values, as seen in the anticipated and actual data overlays. The 12-hour forecast in Figure 4-19 displays a precise hour-by-hour prediction pattern corresponding to the actual data's peaks and troughs, critical for real-time energy management and fast decision-making. The 24-hour forecast in Figure 4-20 broadens the predictive reach by providing insights into the whole daily cycle of power generation, which is critical for daily operational planning, including energy storage and load control.

The evaluation metrics evaluate the model's correctness. The projections are close to the actual values, with Mean Absolute Errors of 5.06 and 4.74, suggesting precise predictions. The Mean Squared Errors, 59.22 and 47.33, indicate differences in the model's predictions, particularly for more considerable deviations, which are penalised more harshly in this metric. The Root Mean Squared Errors of 7.70 and 6.88 show a similar pattern but are easier to read since they are in the same units as the output variable. The Mean Absolute Percentage Errors of 0.11% and 0.06% show that the forecasts have the lowest relative errors, with the 24-hour forecasts being much more accurate.

The residual plots, which illustrate the difference between the actual and anticipated values, demonstrate how mistakes are distributed. Positive numbers indicate that the model underpredicts the real kPt, whereas negative values indicate that it overpredicts. While there are some inaccuracies in prediction, they are extremely few and within a narrow range, indicating the model's usefulness.

The hybrid forecasting model's results accurately project electricity generation for a smart grid. The model captures the dataset's fundamental trends and cycles while maintaining constant precision in short-term and daily forecasting timeframes. The model's utility for smart grid energy forecasting is highlighted by residuals, demonstrating that the projections are well-calibrated, with most tiny differences.

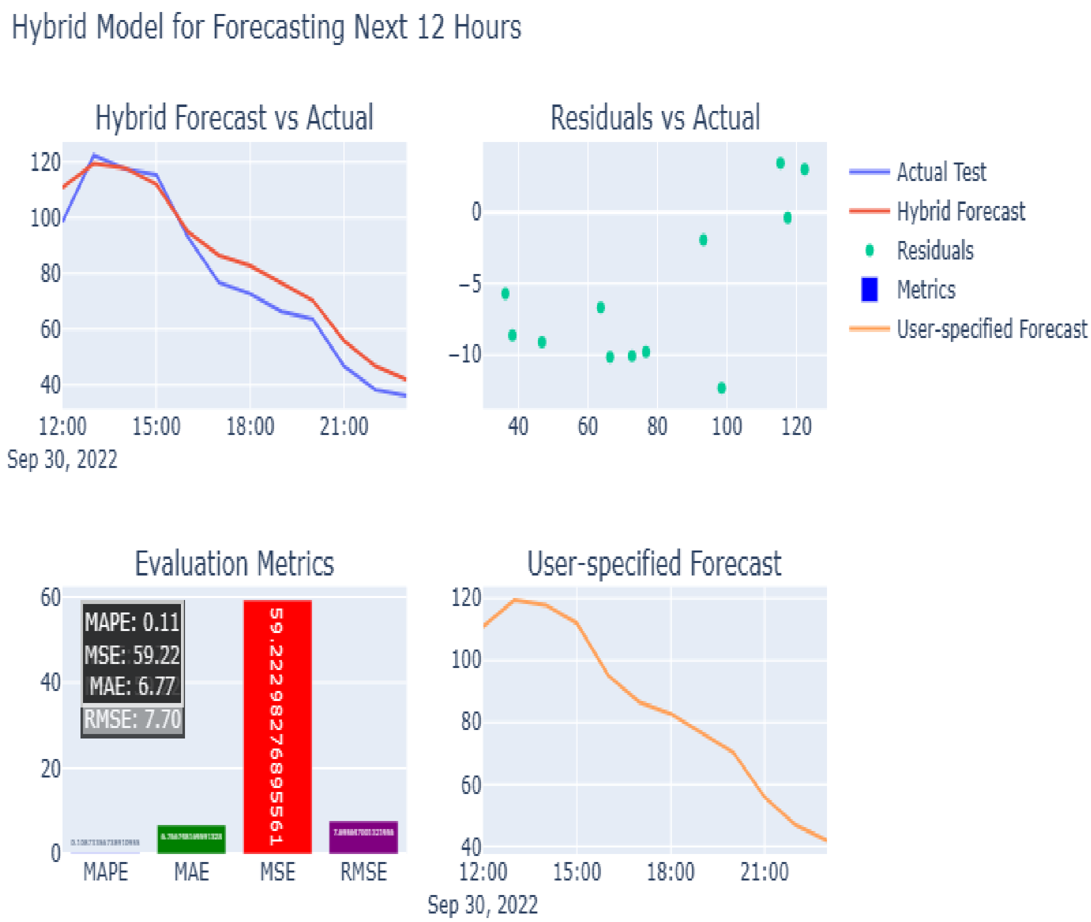

Figure 4-19 Dashboard for Hybrid Model Forecasting for the next 12 hours along with error evaluation metrics, residuals, and hybrid forecast vs actual plots

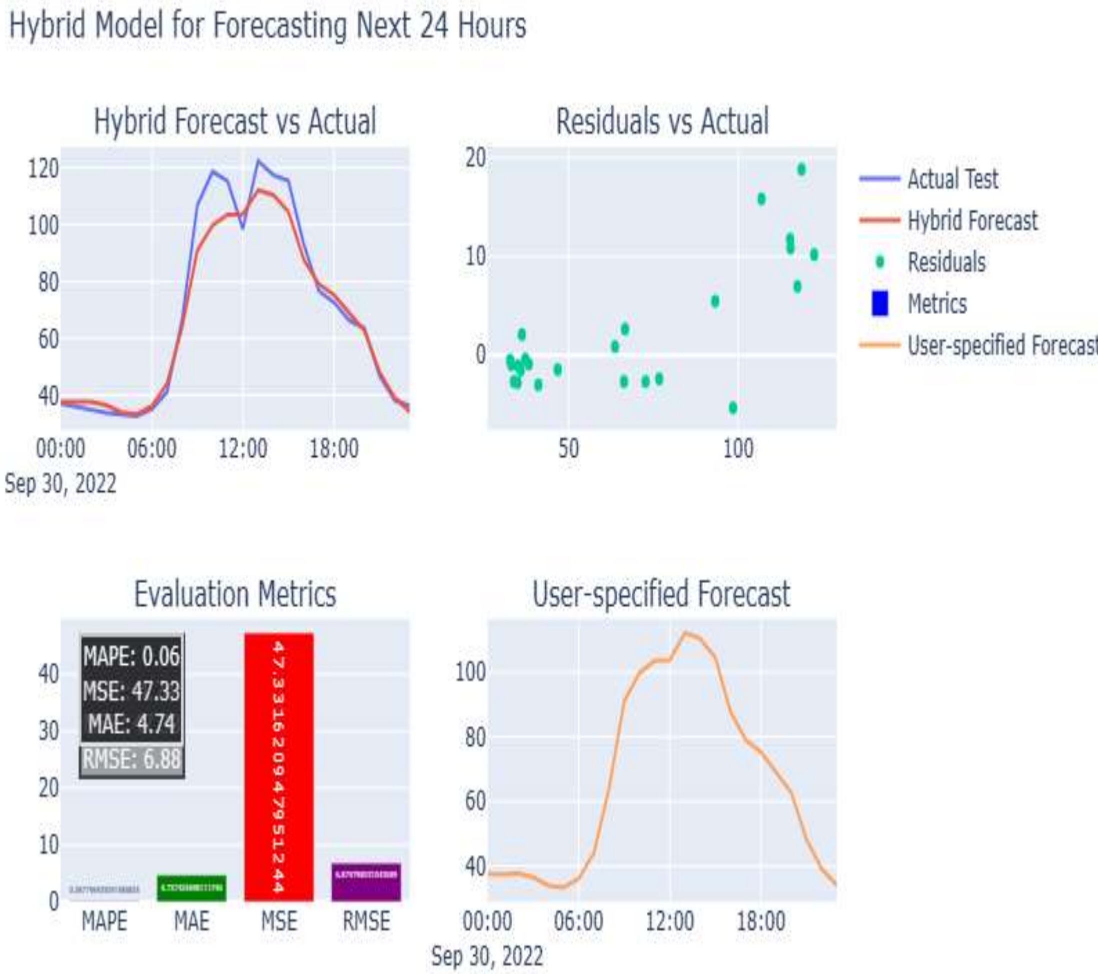

Figure 4-20 Dashboard for Hybrid Model Forecasting for the next 24 hours along with error evaluation metrics, residuals, and hybrid forecast vs actual plots

4.12 Comparison of Implemented Models Based on the Error Evaluation

The error metrics for multiple machine learning models used for a 24-hour power generation prediction for a smart grid system, along with weather variables, are summarised in Table 15. These indicators are critical for evaluating each model's performance and determining its predicted strengths and flaws.

The SARIMAX model has the lowest Mean Absolute Error (MAE), indicating that it delivers the closest average forecasts to the actual values, which is advantageous for precision in power supply control. However, its Mean Absolute Percentage Error (MAPE) is substantially bigger, suggesting that when it errs, it might be proportionally significant relative to the actual number.

Conversely, the Prophet model has a greater MAE than SARIMAX, indicating that its predictions are less accurate on average. This may influence the model's usability in applications requiring precise energy predictions. Its MAPE, conversely, is equivalent to SARIMAX, showing similar relative faults.

The Holt-Winters model has an even larger MAE, which may imply that it needs to be more accurate on average than SARIMAX and Prophet. It does, however, have a slightly higher MAPE than SARIMAX, which might imply that, while its predictions are further from the genuine values, they do not have as broad a range of proportional errors.

With the largest MAE and RMSE of any model, XGBoost's predictions may differ significantly from the actual values. As a result, it may be less suited for applications that need exact power output forecasts. It has the greatest MAPE, suggesting its proportional mistakes are similarly significant.

The LSTM model has a moderate MAE and RMSE, indicating that its predictions are in the middle of the pack compared to other models. Its MAPE is on the high side, implying that its forecasts may need to be more accurate than those of the other models.

Finally, the Hybrid model has the second-lowest MAE and the lowest MAPE, suggesting that it not only predicts values near the actuals on average but also has the most minor proportional errors of any model. This implies that the Hybrid model uses the qualities of other models to reach a high level of precision, both absolute and relative.

In conclusion, while each model has advantages, the Hybrid model stands out for its low error rates across all measures, indicating that it is a trustworthy tool for smart grid power forecasting. The SARIMAX model looks to be the most accurate regarding MAE, but its somewhat higher MAPE suggests that it may only sometimes provide proportionate accuracy. The LSTM and Holt-Winters models offer a good balance of the two forms of faults. While the XGBoost model is powerful for many machine learning tasks, it exhibits the most errors in this forecasting scenario. It may require further tweaking or ensemble approaches to increase its performance.

Table 15 Comparison of SARIMAX, FBProphet, Holt-Winters, XGBoost, LSTM, and Hybrid Time Series model on the combined data of smart grid batteries and weather variables for forecasting the next 24 hours of power generation

Model	**Mean Absolute Error (MAE)**	**Mean Squared Error (MSE)**	**Root Mean Squared Error (RMSE)**	**Mean Absolute Percentage Error (MAPE)**
SARIMAX	3.54	22.97	4.79	4.92
FBProphet	6.17	62.9	7.93	9.22
Holts-Winter	5.16	68.11	8.25	5.58
XGBoost	8.24	181.17	13.46	9.78
LSTM	5.06	62.92	7.93	6.91
Hybrid	4.74	47.33	6.88	0.06

Chapter 5 Discussion

In smart grids, combining demand response systems with hyperlocal weather predictions has increased the prediction of power generation. The granular insights provided by the hyperlocal meteorological data, which included precipitation, temperature, relative humidity, apparent temperature, cloud cover, and wind speed, substantially impacted power generation forecasts. The ability to anticipate meteorological conditions correctly resulted in more accurate forecasting of power demand and generation, allowing for more effective resource allocation and energy distribution.

Integrating hyperlocal weather predictions with a hybrid predictive model dramatically enhances smart grid power generation, providing a more sophisticated approach than prior research. This approach is distinguished by its utilisation of hyperlocal data, which offers precise insights into weather patterns at the micro-level, which is critical for accurate energy demand forecasts. The hybrid model, which combines several time series and machine learning methods, takes use of each's capabilities, a technique not completely explored in previous studies. This integration provides a thorough understanding of energy patterns, allowing for accurate forecasts under a variety of circumstances. The model's capacity to adapt to changing weather and energy consumption patterns is a significant advance above prior techniques, which frequently struggled with rapid meteorological changes [78]. The hybrid model improves resource allocation and energy distribution efficiency by precisely anticipating power demand and generation, solving grid inefficiencies observed in previous research [79-80]. Furthermore, because this technique is scalable, it is appropriate for growing urban energy demands, which is crucial as cities develop and adapt [81].

Five distinct models were trained, and a hybrid model was built to use the strengths of each model. With a MAPE of 0.06 and an MAE of 4.74 for a 24-hour forecast, the hybrid model outperformed the others, demonstrating the effectiveness of integrating several modelling methodologies in capturing the multidimensional character of the contributing elements and their dynamic interaction in determining power generation and demand.

Despite the positive outcomes, substantial setbacks occurred. While helpful, individual models have limitations that need to capture the data's complexities sufficiently. Specific models, for example, may struggle with non-linear trends, different seasonalities, or abrupt changes in meteorological circumstances. The hybrid method eliminated these limits while highlighting the importance of continuous development to increase flexibility and accuracy.

Another area for improvement was the quality and availability of data. There is no reliable open-source data accessible for academics to work with, forcing them to rely on data from energy firms or other commercial resources. While the time series data for Smart Grid batteries had enough variables for us to anticipate, the data was only accessible for one month, limiting many forecasting skills that these models would have covered. This method is only applicable to short-term load forecasting. However, if additional data is available, similar approaches may be used for long-term load forecasting with minor alterations and improvements to the methodology.

A similar attempt was conducted with the hyperlocal model, forecasting 500 hours; the results were identical to individual power projections without weather variables. The MAPE increased to 1.39%, but the prediction accuracy dropped drastically, showing that the model needs to capture a large percentage of the variation due to a lack of data. The model is only trained on 30 days of data, and whatever hours of input are provided, it trains on all data except those hours; for example, if it wishes to forecast 500 hours, the model will not train on the latest 500 hours of data. Many seasonal trends and other elements are lost over these 500 hours and must be captured by the model. Nonetheless, confining this approach to short-term predictions is the whole objective.

As a result, plans include dealing with data with more variables and span a more extended period. More forecast improvement may be noticed by experimenting with other time series, intense learning models, and their ensembles. These models might be utilised for commercial reasons since dashboards can be improved by discovering additional frameworks and libraries for data visualisation and having real-time interactive dashboards where users can participate and provide feedback on the dashboard. This can be accomplished with computationally affordable training models, saving customers from waiting 60 to 90 seconds for their projections.

Future smart grid predictive analytic research will combine ideas such as Federated Learning and approaches like Adversarial Networks. This subtle approach seeks to improve the depth and breadth of data analysis without subscribing to a single technique. The project might possibly examine decentralised data aggregation approaches by pulling inspiration from Federated Learning [82], tapping from a varied variety of data points from smart metres to augment the prediction models. This technique would provide a more thorough knowledge of energy trends in metropolitan settings while protecting data privacy and security.

Similarly, the research might investigate tactics like Adversarial Networks to improve forecast accuracy. This would include creating models capable of replicating a wide range of energy distribution and consumption situations, therefore teaching the system to anticipate and adapt to changing conditions. This strategy has the potential to considerably improve the robustness of predictive models, equipping them to deal with a wide variety of energy demand and supply changes.

In terms of hyperlocal weather prediction, research will continue to focus on improving the models' capacity to read and use detailed local weather data. The objective is to increase the sensitivity of the algorithms to the minor effects of localised weather fluctuations on energy use and generation.

Future work will focus on identifying and implementing new, more efficient libraries and frameworks for data visualisation and dashboard capabilities. The goal is to discover solutions that have extensive analytical capabilities comparable to current tools but are more computationally efficient. This is especially important given the computing requirements of advanced models. The ideal visualisation tools would give real-time, interactive dashboards, allowing users to interact dynamically with data [83] and provide important feedback.

Conclusion

This research effectively showed the integration of hyperlocal weather forecasting with demand response mechanisms in smart grids, representing a big step forward in maximising power-generating efficiency. The research has given a sophisticated method for regulating and anticipating energy consumption in smart urban ecosystems by exploiting the predictive capabilities of powerful machine learning algorithms. The study rigorously analysed five independent forecasting models, each fed with hyperlocal meteorological data: SARIMAX, Prophet, XGBoost, LSTM, and Holt-Winters. SARIMAX's outstanding performance, evidenced by its remarkable MAPE and MAE, demonstrated its capacity to detect complicated seasonal patterns and autocorrelations in energy data. However, the hybrid model—a combination of SARIMAX, Prophet, and Holt-Winters—outperformed the individual models regarding predicted accuracy. The hybrid model demonstrated the transformational potential of merging machine learning algorithms with real-time data analytics, with a MAPE of only 0.06% and an MAE of 4.43. A vital component of this study is the creation of a user-centric dashboard that provides real-time insights on anticipated data and model performance metrics. With its customisation possibilities, this dashboard enables bespoke forecasting that helps informed decision-making in real-time circumstances. The actual deployment of this technology shows the thesis's compatibility with smart grid operational demands and the broader goals of smart cities. The hybrid model's remarkable performance highlights the possibility of merging machine learning algorithms with demand response tactics to achieve more flexibility and efficiency in energy management. This strategy is essential as the energy industry changes to meet the needs of a more unpredictable environment and the complicated energy consumption habits of urban populations expanding in size.

References

[1] Fu, Xueqian. "Statistical machine learning model for capacitor planning considering uncertainties in photovoltaic power." *Protection and Control of Modern Power Systems* 7, no. 1 (2022): 5.

[2] Chen, Yan, W. Sabrina Lin, Feng Han, Yu-Han Yang, Zoltan Safar, and KJ Ray Liu. "Incentive compatible demand response games for distributed load prediction in smart grids." *APSIPA Transactions on Signal and Information Processing* 3 (2014): e9.

[3] Hussain, SM Suhail, Furquan Nadeem, Mohd Asim Aftab, Ikbal Ali, and Taha Selim Ustun. "The emerging energy internet: Architecture, benefits, challenges, and future prospects." *Electronics* 8, no. 9 (2019): 1037.

[4] Komninos, Nikos, Eleni Philippou, and Andreas Pitsillides. "Survey in smart grid and smart home security: Issues, challenges and countermeasures." *IEEE Communications Surveys & Tutorials* 16, no. 4 (2014): 1933-1954.

[5] Abdella, Juhar, and Khaled Shuaib. "Peer to peer distributed energy trading in smart grids: A survey." *Energies* 11, no. 6 (2018): 1560.

[6] Mets, Kevin, Tom Verschueren, Wouter Haerick, Chris Develder, and Filip De Turck. "Optimizing smart energy control strategies for plug-in hybrid electric vehicle charging." In *2010 IEEE/IFIP network operations and management symposium workshops*, pp. 293-299. Ieee, 2010.

[7] Xia, Min, Haidong Shao, Xiandong Ma, and Clarence W. de Silva. "A stacked GRU-RNN-based approach for predicting renewable energy and electricity load for smart grid operation." *IEEE Transactions on Industrial Informatics* 17, no. 10 (2021): 7050-7059.

[8] Carrera, Berny, and Kwanho Kim. "Comparison analysis of machine learning techniques for photovoltaic prediction using weather sensor data." *Sensors* 20, no. 11 (2020): 3129.

[9] Wang, Yizhen, Ningqing Zhang, and Xiong Chen. "A short-term residential load forecasting model based on LSTM recurrent neural network considering weather features." *Energies* 14, no. 10 (2021): 2737.

[10] Arnold, George W., David A. Wollman, Gerald FitzPatrick, Dean Prochaska, David Holmberg, David H. Su, Allen R. Hefner Jr et al. "NIST framework and roadmap for smart grid interoperability standards, release 1.0." (2010).

[11] Judge, Malik Ali, Asif Khan, Awais Manzoor, and Hasan Ali Khattak. "Overview of smart grid implementation: Frameworks, impact, performance and challenges." *Journal of Energy Storage* 49 (2022): 104056.

[12] Ahmadzadeh, Sahar, Gerard Parr, and Wanqing Zhao. "A review on communication aspects of demand response management for future 5G IoT-based smart grids." *IEEE Access* 9 (2021): 77555-77571.

[13] Kumari, Sanju, Neeraj Kumar, and Prashant Singh Rana. "A Big Data Approach for Demand Response Management in Smart Grid Using the Prophet Model." *Electronics* 11, no. 14 (2022): 2179.

[14] He, Zheyu, Rongheng Lin, Budan Wu, Xin Zhao, and Hua Zou. "Pre-Attention Mechanism and Convolutional Neural Network Based Multivariate Load Prediction for Demand Response." *Energies* 16, no. 8 (2023): 3446.

[15] Ali, Abdelfatah, Mostafa F. Shaaban, and Hatem F. Sindi. "Optimal Operational Planning of RES and HESS in Smart Grids Considering Demand Response and DSTATCOM Functionality of the Interfacing Inverters." *Sustainability* 14, no. 20 (2022): 13209.

[16] Aliberti, Alessandro, Lorenzo Bottaccioli, Giansalvo Cirrincione, Enrico Macii, Andrea Acquaviva, and Edoardo Patti. "Forecasting Short-term Solar Radiation for Photovoltaic Energy Predictions." In *Smartgreens*, pp. 44-53. 2018.

[17] Asadinejad, Ailin, and Kevin Tomsovic. "Optimal use of incentive and price based demand response to reduce costs and price volatility." *Electric Power Systems Research* 144 (2017): 215-223.

[18] Ali, Abdelfatah, Mostafa F. Shaaban, and Hatem F. Sindi. "Optimal Operational Planning of RES and HESS in Smart Grids Considering Demand Response and DSTATCOM Functionality of the Interfacing Inverters." *Sustainability* 14, no. 20 (2022): 13209.

[19] Bakare, Mutiu Shola, Abubakar Abdulkarim, Mohammad Zeeshan, and Aliyu Nuhu Shuaibu. "A comprehensive overview on demand side energy management towards smart grids: challenges, solutions, and future direction." *Energy Informatics* 6, no. 1 (2023): 1-59.

[20] Ahmadzadeh, Sahar, Gerard Parr, and Wanqing Zhao. "A review on communication aspects of demand response management for future 5G IoT-based smart grids." *IEEE Access* 9 (2021): 77555-77571.

[21] Inagata, Tomoya, Keita Matsunaga, Yuji Mizuno, Fujio Kurokawa, Masaharu Tanaka, and Nobumasa Matsui. "An Improvement of Power Demand Prediction Method using Weather Information and Machine Learning: A Case of a Clinic in Japan (II)." In *2023 11th International Conference on Smart Grid (icSmartGrid)*, pp. 1-6. IEEE, 2023.

[22] Kadlec, Miroslav, Barbora Bühnová, Jan Tomšík, Jan Herman, and Kateřina Družbíková. "Weather forecast based scheduling for demand response optimization in smart grids." In *2017 Smart City Symposium Prague (SCSP)*, pp. 1-6. IEEE, 2017.

[23] Rao, Challa Krishna, Sarat Kumar Sahoo, and Franco Fernando Yanine. "Demand Response for Renewable Generation in an IoT based Intelligent Smart Energy Management System." In *2021 Innovations in Power and Advanced Computing Technologies (i-PACT)*, pp. 1-7. IEEE, 2021.

[24] Banerjee, Kamalika, Sawan Sen, Sandip Chanda, and Samarjit Sengupta. "A review on demand response techniques of load management in smart grid." In *2021 IEEE International Power and Renewable Energy Conference (IPRECON)*, pp. 1-8. IEEE, 2021.

[25] Guo, Jun, Yongfeng Wu, Ruizhe Yang, Ying Xu, and Faguo Zhu. "A Day-ahead Dispatch Strategy of High-Penetration PV Grids with Virtual Power Plants." In *2022 IEEE 6th Conference on Energy Internet and Energy System Integration (EI2)*, pp. 162-167. IEEE, 2022.

[26] Arumugham, Vinothini, Hayder MA Ghanimi, Denis A. Pustokhin, Irina V. Pustokhina, Vidya Sagar Ponnam, Meshal Alharbi, Parkavi Krishnamoorthy, and Sudhakar Sengan. "An Artificial-Intelligence-Based Renewable Energy Prediction Program for Demand-Side Management in Smart Grids." *Sustainability* 15, no. 6 (2023): 5453.

[27] Singhal, Rishi, Poonam Singhal, and Shailender Gupta. "Solar-Cast: Solar Power Generation Prediction from Weather Forecasts using Machine Learning." In *2022 IEEE 10th Power India International Conference (PIICON)*, pp. 1-6. IEEE, 2022.

[28] Islam, Md Aminul, Muhammad Ismail, Osman Boyaci, Rachad Atat, and Susmit Shannigrahi. "Graph Neural Network Based Prediction of Data Traffic in Cyber-Physical Smart Power Grids." In *2022 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pp. 322-327. IEEE, 2022.

[29] Dayaratne, Thusitha, Carsten Rudolph, Ariel Liebman, and Mahsa Salehi. "Robust demand response for device scheduling under false data injection attacks in smart grids." In *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, pp. 294-298. IEEE, 2020.

[30] Ibrahim, Bibi, Luis Rabelo, Edgar Gutierrez-Franco, and Nicolas Clavijo-Buritica. "Machine learning for short-term load forecasting in smart grids." *Energies* 15, no. 21 (2022): 8079.

[31] Das, Dipankar, and Satyajit Chakrabarti. "Forecasting non-linear WPI of manufacture of chemicals and chemical products in India: an MLP approach." *International Journal of Advanced Technology and Engineering Exploration* 8, no. 82 (2021): 1193.

[32] Lindley, Sarah J., Penny A. Cook, Matthew Dennis, and Anna Gilchrist. "Biodiversity, physical health and climate change: a synthesis of recent evidence." *Biodiversity and health in the face of climate change* (2019): 17-46.

[33] Sarkar, Pranta Kumar, Pallab Kumar Sarkar, and Md Monir Ahammod Bin Atique. "Prediction of Power Consumption in Smart Grid: A Reliable Path to a Smart City Based on Various Machine Learning Models." In *2022 International Conference on Recent Progresses in Science, Engineering and Technology (ICRPSET)*, pp. 1-6. IEEE, 2022.

[34] Fang, Xi, Satyajayant Misra, Guoliang Xue, and Dejun Yang. "Smart grid—The new and improved power grid: A survey." *IEEE communications surveys & tutorials* 14, no. 4 (2011): 944-980.

[35] Liu, Jiaqi, Hongji Hu, Samson S. Yu, and Hieu Trinh. "Virtual Power Plant with Renewable Energy Sources and Energy Storage Systems for Sustainable Power Grid-Formation, Control Techniques and Demand Response." *Energies* 16, no. 9 (2023): 3705.

[36] Ahmad, Ashfaq, Nadeem Javaid, Mohsen Guizani, Nabil Alrajeh, and Zahoor Ali Khan. "An accurate and fast converging short-term load forecasting model for industrial applications in a smart grid." *IEEE Transactions on Industrial Informatics* 13, no. 5 (2016): 2587-2596.

[37] Bakare, Mutiu Shola, Abubakar Abdulkarim, Mohammad Zeeshan, and Aliyu Nuhu Shuaibu. "A comprehensive overview on demand side energy management towards smart grids: challenges, solutions, and future direction." *Energy Informatics* 6, no. 1 (2023): 1-59.

[38] Salazar, Eduardo J., Mauro Jurado, and Mauricio E. Samper. "Reinforcement Learning-Based Pricing and Incentive Strategy for Demand Response in Smart Grids." *Energies* 16, no. 3 (2023): 1466.

[39] Chen, Yan, W. Sabrina Lin, Feng Han, Yu-Han Yang, Zoltan Safar, and KJ Ray Liu. "Incentive compatible demand response games for distributed load prediction in smart grids." *APSIPA Transactions on Signal and Information Processing* 3 (2014): e9.

[40] Feng, Ding, Sheng Lin, Zhengyou He, Xiaojun Sun, and Wei-Jen Lee. "Optimization method with prediction-based maintenance strategy for traction power supply equipment based on risk quantification." *IEEE transactions on transportation electrification* 4, no. 4 (2018): 961-970.

[41] Ungureanu, Stefan, Vasile Ţopa, and Andrei Cziker. "Industrial load forecasting using machine learning in the context of smart grid." In *2019 54th International Universities Power Engineering Conference (UPEC)*, pp. 1-6. IEEE, 2019.

[42] Kumari, Aparna, Darshan Vekaria, Rajesh Gupta, and Sudeep Tanwar. "Redills: Deep learning-based secure data analytic framework for smart grid systems." In *2020 IEEE*

international conference on communications workshops (ICC Workshops), pp. 1-6. IEEE, 2020.

[43] Shi, Hanhong, Lei Wang, Rafał Scherer, Marcin Woźniak, Pengchao Zhang, and Wei Wei. "Short-term load forecasting based on adabelief optimized temporal convolutional network and gated recurrent unit hybrid neural network." *IEEE Access* 9 (2021): 66965-66981.

[44] Xia, Min, Haidong Shao, Xiandong Ma, and Clarence W. de Silva. "A stacked GRU-RNN-based approach for predicting renewable energy and electricity load for smart grid operation." *IEEE Transactions on Industrial Informatics* 17, no. 10 (2021): 7050-7059.

[45] Singh, Upma, Mohammad Rizwan, Muhannad Alaraj, and Ibrahim Alsaidan. "A machine learning-based gradient boosting regression approach for wind power production forecasting: A step towards smart grid environments." *Energies* 14, no. 16 (2021): 5196.

[46] Zhang, Yufan, Honglin Wen, Qiuwei Wu, and Qian Ai. "Optimal adaptive prediction intervals for electricity load forecasting in distribution systems via reinforcement learning." *IEEE Transactions on Smart Grid* (2022).

[47] Bouquet, Pierre, Ilya Jackson, Mostafa Nick, and Amin Kaboli. "AI-based forecasting for optimised solar energy management and smart grid efficiency." *International Journal of Production Research* (2023): 1-22.

[48] Ucar, Ferhat. "A Comprehensive Analysis of Smart Grid Stability Prediction along with Explainable Artificial Intelligence." *Symmetry* 15, no. 2 (2023): 289.

[49] Hayes, Barry, Jorn Gruber, and Milan Prodanovic. "Short-term load forecasting at the local level using smart meter data." In *2015 IEEE Eindhoven PowerTech*, pp. 1-6. IEEE, 2015.

[50] Patil, P. "Time Series Data of Battery Energy Storage System." IEEE DataPort. 2023. https://ieee-dataport.org/documents/time-series-data-battery-energy-storage-system (accessed October 2, 2023).

[51] "Historical Weather API." Open-Meteo. No date. https://open-meteo.com/en/docs/historical-weather-api (accessed October 2, 2023).

[52] Adineh, Amir Hossein, Zahra Narimani, and Suresh Chandra Satapathy. "Importance of data preprocessing in time series prediction using SARIMA: A case study." *International Journal of Knowledge-Based and Intelligent Engineering Systems* 24, no. 4 (2020): 331-342.

[53] Chiranjeevi, Madderla, Skandha Karlamangal, Tukaram Moger, Debashisha Jena, and Ayush Agarwal. "Preprocessing Techniques of Solar Irradiation Data." In *2023 IEEE Renewable Energy and Sustainable E-Mobility Conference (RESEM)*, pp. 1-6. IEEE, 2023.

[54] Salzillo, Giovanni, Ciro Natale, Giovanni B. Fioccola, and Enrico Landolfi. "Evaluation of driver drowsiness based on real-time face analysis." In *2020 IEEE international conference on systems, man, and cybernetics (SMC)*, pp. 328-335. IEEE, 2020.

[55] Vambol, Sergij, Ramzan Soomro, Saghir Pervaiz Ghauri, Azhar Ali Marri, Hoang Thi Dung, Nazish Manzoor, Shella Bano, Sana Shahid, Ahmed Farooq, and Yurii Lutsenko. "Viable forecasting monthly weather data using time series methods." *Ecological Questions* 34, no. 1 (2023): 1-16.

[56] Petrică, Andreea-Cristina, Stelian Stancu, and Virgil Ghițulescu. "Stationarity–The Central Concept in Time Series Analysis." (2017).

[57] Wang, Jin, Tianyao Ji, and Mengshi Li. "A Combined Short-Term Forecast Model of Wind Power Based on Empirical Mode Decomposition and Augmented Dickey-Fuller Test." In *Journal of Physics: Conference Series*, vol. 2022, no. 1, p. 012017. IOP Publishing, 2021.

[58] Pasari, Sumanta, and Aditya Shah. "Time series auto-regressive integrated moving average model for renewable energy forecasting." In *Enhancing Future Skills and Entrepreneurship: 3rd Indo-German Conference on Sustainability in Engineering*, pp. 71-77. Cham: Springer International Publishing, 2020.

[59] Ogbonna, C. J., C. J. Nweke, E. C. Nwogu, and I. S. Iwueze. "Wavelet Transform as an Alternative to Power Transformation in Time Series Analysis." *Bulletin of Mathematical Sciences and Applications Vol* 17 (2016): 57-74.

[60] Kurniawan, Mahrus Lutfi Adi, and Indanazulfa Q. A'yun. "Dynamic Analysis On Export, FDI and Growth in Indonesia: An Autoregressive Distributed Lag (ARDL) Model." *Journal of Economics, Business, & Accountancy Ventura* 24, no. 3 (2022): 350-362.

[61] Chalkiadakis, Ioannis, Gareth W. Peters, and Matthew Ames. "Hybrid ARDL-MIDAS-Transformer time-series regressions for multi-topic crypto market sentiment driven by price and technology factors." *Digital Finance* 5, no. 2 (2023): 295-365.

[62] Wang, Jen Chun. "Analysis of energy use intensity and greenhouse gas emissions for universities in Taiwan." *Journal of Cleaner Production* 241 (2019): 118363.

[63] Elshewey, Ahmed M., Mahmoud Y. Shams, Abdelghafar M. Elhady, Samaa M. Shohieb, Abdelaziz A. Abdelhamid, Abdelhameed Ibrahim, and Zahraa Tarek. "A Novel WD-SARIMAX Model for Temperature Forecasting Using Daily Delhi Climate Dataset." *Sustainability* 15, no. 1 (2022): 757.

[64] Elshewey, Ahmed M., Mahmoud Y. Shams, Abdelghafar M. Elhady, Samaa M. Shohieb, Abdelaziz A. Abdelhamid, Abdelhameed Ibrahim, and Zahraa Tarek. "A Novel WD-SARIMAX Model for Temperature Forecasting Using Daily Delhi Climate Dataset." *Sustainability* 15, no. 1 (2022): 757.

[65] Almazrouee, Abdulla I., Abdullah M. Almeshal, Abdulrahman S. Almutairi, Mohammad R. Alenezi, and Saleh N. Alhajeri. "Long-term forecasting of electrical loads in kuwait using prophet and holt–winters models." *Applied Sciences* 10, no. 16 (2020): 5627.

[66] Wang, Yi, Qixin Chen, Tao Hong, and Chongqing Kang. "Review of smart meter data analytics: Applications, methodologies, and challenges." *IEEE Transactions on Smart Grid* 10, no. 3 (2018): 3125-3148. [67] Hochreiter, Sepp and Jürgen Schmidhuber. “Long Short-Term Memory.” Neural Computation 9 (1997): 1735-1780.

[68] Lu, Jixiang, Qipei Zhang, Zhihong Yang, and Mengfu Tu. "A hybrid model based on convolutional neural network and long short-term memory for short-term load forecasting." In *2019 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1-5. IEEE, 2019.

[69] Chen, Tianqi, and Carlos Guestrin. "Xgboost: A scalable tree boosting system." In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785-794. 2016.

[70] Chernozhukov, Victor, Whitney Newey, Víctor M. Quintas-Martínez, and Vasilis Syrgkanis. "Riesznet and forestriesz: Automatic debiased machine learning with neural nets and random forests." In *International Conference on Machine Learning*, pp. 3901-3914. PMLR, 2022.

[71] Łoś, H., G. Sousa Mendes, David Cordeiro, Nuno Grosso, Hugo Costa, Pedro Benevides, and Mário Caetano. "Evaluation of XGBoost and LGBM performance in tree species classification with sentinel-2 data." In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pp. 5803-5806. IEEE, 2021.

[72] Dudek, Grzegorz, Slawek Smyl, and Paweł Pełka. "Recurrent Neural Networks for Forecasting Time Series with Multiple Seasonality: A Comparative Study." *arXiv preprint arXiv:2203.09170* (2022).

[73] Livieris, Ioannis E., Stavros Stavroyiannis, Lazaros Iliadis, and Panagiotis Pintelas. "Smoothing and stationarity enforcement framework for deep learning time-series forecasting." *Neural Computing and Applications* 33, no. 20 (2021): 14021-14035.

[74] Kramar, Vadim, and Vasiliy Alchakov. "Evaluation of the Machine Learning Techniques for Forecasting the Seasonal Time Series." In *2023 IEEE Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)*, pp. 262-265. IEEE, 2023.

[75] Taylor, Sean J., and Benjamin Letham. "Forecasting at scale." *The American Statistician* 72, no. 1 (2018): 37-45.

[76] Hyndman, Rob J., and George Athanasopoulos. *Forecasting: principles and practice*. OTexts, 2018.

[77] Ma, Jie, Yuelong Jia, and Binglin Lv. "A hybrid forecasting model for monthly electricity consumption combining prophet and SARIMAX models." In *Third International Conference on Advanced Algorithms and Neural Networks (AANN 2023)*, vol. 12791, pp. 626-631. SPIE, 2023.

[78] Qi, Le, Mingyuan Chen, Jifeng Song, Minyu Zhang, Wenbin Zheng, and Qinghua Shang. "Short-Term Reactive Load Forecasting Based on a Hybrid Machine Learning Model." In *2022 Power System and Green Energy Conference (PSGEC)*, pp. 754-759. IEEE, 2022.

[79] Chadoulos, Spiros, Iordanis Koutsopoulos, and George C. Polyzos. "One model fits all: Individualized household energy demand forecasting with a single deep learning model." In *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*, pp. 466-474. 2021.

[80] Gupta, Rahul, Anil Kumar Yadav, S. K. Jha, and Pawan Kumar Pathak. "Time series forecasting of solar power generation using Facebook prophet and XG boost." In *2022 IEEE Delhi section conference (DELCON)*, pp. 1-5. IEEE, 2022.

[81] Rahman, Saidur, Md Golam Rabiul Alam, and Md Mahbubur Rahman. "Deep learning based ensemble method for household energy demand forecasting of smart home." In *2019 22nd International Conference on Computer and Information Technology (ICCIT)*, pp. 1-6. IEEE, 2019.

[82] Bondok, Atef H., Mohamed Mahmoud, Mahmoud M. Badr, Mostafa M. Fouda, Mohamed Abdallah, and Maazen Alsabaan. "Novel Evasion Attacks against Adversarial Training Defense for Smart Grid Federated Learning." *IEEE Access* (2023).

[83] Toasa, Renato, Marisa Maximiano, Catarina Reis, and David Guevara. "Data visualization techniques for real-time information—A custom and dynamic dashboard for analyzing surveys' results." In *2018 13th Iberian Conference on Information Systems and Technologies (CISTI)*, pp. 1-7. IEEE, 2018.