TIME SERIES FORECASTING OF ELECTRICITY CONSUMPTION : NEW SOUTH WALES, AUSTRALIA

MSc Research Project
MSc Data Analytics

Gokulkrishnan Mohankumar
Student ID: 22135588

School of Computing
National College of Ireland

Supervisor: Jorge Basilio

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Gokulkrishnan Mohankumar
Student ID:	22135588
Programme:	MSc Data Analytics
Year:	2023
Module:	MSc Research Project
Supervisor:	Jorge Basilio
Submission Due Date:	14/12/2023
Project Title:	TIME SERIES FORECASTING OF ELECTRICITY CONSUMPTION : NEW SOUTH WALES, AUSTRALIA
Word Count:	6423
Page Count:	21

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Gokulkrishnan Mohankumar
Date:	14th December 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Time Series Forecasting of Electricity Consumption: New South Wales, Australia

Gokulkrishnan Mohankumar
22135588

Abstract

This research examines how machine learning algorithms, particularly AutoRegressive Integrated Moving Average (ARIMA), Auto ARIMA, Seasonal AutoRegressive Integrated Moving Average (SARIMA) and Prophet, can be used to predict electricity consumption in New South Wales(NSW), Australia. The dataset covers the years 2018 to 2022, providing insights into consumption trends. The goal of the study is to provide valuable insights into energy demand, aiding in informed decision-making and sustainable resource management.

1 Introduction

1.1 Research brief

Forecasting electricity consumption is an essential component of environmentally friendly resource management, especially in areas where consumer preferences and the economy are changing. With a focus on ARIMA, Auto ARIMA, SARIMA, as well as the Prophet model, this chapter presents the research focusing on using machine-learning algorithms to predict the consumption of electricity in New South Wales(NSW), Australia (Ahmed and Agalgaonkar (2012)). The dataset, which covers the years 2018 through 2022, offers a thorough investigation of spending trends and patterns.

The use of algorithms based on machine learning not only improves the accuracy of predicted electricity consumption but also enables proactive and effective resource allocation. The use of ARIMA, Auto ARIMA, SARIMA, including the Prophet model enables a comprehensive approach that takes into account seasonality, trends, and unique occurrences (Preniqi and Konur (2020)). Because the study focuses on Australia's New South Wales, a region characterized by dynamic socioeconomic and environmentalist shifts, the findings are expected to give significant insights to legislators, utility suppliers, and environmental advocates. Authorities may arrive at informed judgments by staying tuned in to consumption patterns, supporting sustainable energy habits, and aligning resource allocation with evolving requirements and standards of the community.

1.2 Research background

The dynamic landscape of electricity consumption, influenced by economic factors and changing consumer habits, emphasizes the need for advanced forecasting models (Blázquez-García and Barrio (2020a)). Understanding the historical background of the research provides the basis for navigating the complexities of power usage prediction unique to Australia's New South Wales. The need for advanced forecasting is growing as long as consumer behavior and economic factors continue to have an impact. This investigation aims to clarify the complex interplay of factors that influence change by examining the various facets of the dynamics of electricity consumption. The concentration on the state of New South Wales allows for a more in-depth analysis of these variables in a particular geographic setting, which adds to a comprehensive comprehension of the difficulties associated with forecasting power consumption (Preniqi and Konur (2020)).

The study recognizes the importance of economic considerations in driving power consumption patterns. Understanding the relationship between economic changes and power use is becoming increasingly important as the world shifts toward more sustainable energy practices (Blázquez-García and Barrio (2020a)). The study attempts to reveal not only overarching patterns but also the intricate effects of economic variations on energy demand by studying historical data ranging from 2018 to 2022. Consumer behaviors, another important influencer, are changing because of technological improvements and increased environmental consciousness. This study aims to decipher the complex relationship between consumer choices and electricity usage, offering insight into how developing patterns may influence future energy demands. As the inquiry focuses on New South Wales, its results promise to provide region-specific perspectives, allowing stakeholders to develop tailored solutions-strategies that handle the specific difficulties and opportunities given by the local socioeconomic and consumer situation(Ahmed and Agalgaonkar (2012)).

1.3 Research aim

With the primary objective of accurately predicting electricity consumption, this research focuses on implementing machine-learning algorithms, specifically the Prophet model, SARIMA, ARIMA and Auto ARIMA. Australia's New South Wales serves as the focus, and a dataset covering the period from 2018 to 2022 is used in the analysis. Through an exploration of these sophisticated algorithms, the research aims to improve forecasting and offer insightful information about the complex workings of electricity spending habits in this particular area. With a temporal scope of 2018 to 2022, the dataset guarantees a thorough investigation of patterns and developments, providing the foundation for informed choices in the management of renewable resources.

1.4 Research objectives

- To determine the main key factors influencing the accuracy and relevance of time series forecasting models for electricity consumption.
- To evaluate how implementation of machine learning techniques affect the rate of electricity consumption.
- To make recommendations on strategies that can be employed to enhance the precision and applicability of machine learning models in electricity consumption forecasting.
- To discuss the challenges that currently exist in machine learning models for predicting electricity consumption rates and how can these challenges be addressed.

1.5 Research questions

1. What are the main key factors influencing the accuracy and relevance of time series forecasting models for electricity consumption?

2. What impact does the rate of consumption of electricity have when machine-learning techniques are implemented?

3. How can machine-learning models for electricity consumption forecasting be improved in terms of accuracy and usefulness?

4. What are the challenges that currently exist in machine learning models for predicting electricity consumption rates and how can these challenges be addressed?

1.6 Research rationale

The research is being driven by the urgent need to attain electricity consumption prediction, a requirement that is made more pressing by dynamic changes in the economy and changing consumer preferences. A deeper comprehension of the role artificial intelligence plays in this context is imperative due to the complex interplay of these variables, which highlights the intricate nature of forecasting models. To fully understand the complexities of energy demand forecasting, it is essential to look at the factors underlying these models (Mittal and Xu (2020)). The amount of electricity consumed in modern society is not only determined by basic needs. It also reflects the economy's pulse and the constantly changing patterns of consumer behavior. The energy industry is facing increasing difficulty in precisely forecasting demand as economies change and consumer preferences shift (Preniqi and Konur (2020)). This investigation explores the center of this difficulty, emphasizing the distinct setting of Australia's New South Wales.

The cornerstones of this research are machine-learning algorithms, more especially ARIMA, Auto ARIMA, SARIMA, as well as the Prophet model. These sophisticated computational tools have been crucial in helping to extract patterns from large, complicated datasets, giving rise to a more sophisticated understanding of trends in electricity consumption (Arslan (2022a)). The selection of these algorithms demonstrates a dedication to utilizing cutting-edge techniques to improve prediction accuracy and dependability. Australia's New South Wales study zone provides an intriguing microcosm for this investigation. Regional dynamics, legal frameworks, and geographic factors in addition to global economic forces influence the energy landscape of the region. The dataset under examination covers the years 2018 through 2022, capturing a time of significant economic events and changes in society. A thorough analysis that captures the peaks and valleys of trends in electricity consumption is made possible by this temporal breadth.

This research aims to significantly advance the field of demand for energy forecasting by identifying the variables influencing forecasting models. Changes in the economy, from recessions to expansions, have a significant impact on patterns of energy use. Similarly, new layers of complexity are introduced by shifting consumer preferences brought about by things like environmental consciousness or technological advancements. Comprehending the complex interplay between these variables is essential for improving forecasting models to match the changing dynamics of modern energy consumption. In this situation, machine learning as a tool transforms conventional forecasting techniques (Preniqi and Konur (2020)). Its capacity to identify complex patterns from large datasets and its ability to self-optimize and adjust to changing circumstances make it a powerful ally in the prediction of electricity consumption. Rather than just praising machine learning's abilities, this study critically looks at its drawbacks and future problems.

The study recognizes the need for a comprehensive grasp of the benefits and drawbacks of using machine learning techniques in energy forecasting, ranging from problems with data quality to the comprehension of complicated models. As the investigation progresses, it not only clarifies the situation about energy demand forecasting at the moment but also looks into possible ways to make it more effective (Blázquez-García and Barrio (2020a)). Using machine neural networks is not a magic bullet; to be effective, it needs to be carefully integrated with domain knowledge, data pre-processing done with care and constant model validation and improvement. This study considers these aspects and guides best practices and tactics for navigating the dynamic field of demand for energy prediction.

1.7 Summary

To sum up, this introduction establishes the groundwork for an extensive investigation into time series predicting for consumption of electricity in Australia's New South Wales. To support informed choices and sustainable resource management, the study intends to provide significant insights into patterns of energy demand by addressing the research's context, aim, objectives, questions, and justification. The approaches, models, and analyses used to accomplish the research goals will be covered in detail in the upcoming chapters.

2 Related Work

A crucial component of making well-informed decisions in the ever-changing fabric of consumer preferences and economic fluctuations is the accurate forecasting of electricity consumption. This section acts as an important introduction, setting the stage for later talks about forecasting techniques and the critical function of machine learning. Understanding the significant influence that precise forecasts have on managing these oscillations emphasizes the need for sophisticated instruments and procedures (Preniqi and Konur (2020)). The project aims to unravel the complexities surrounding the continuously evolving environment of energy consumption, and forecasting techniques, to provide a more nuanced understanding of the interplay between financial conditions and shifting consumer behaviors.

2.1 Role of machine learning in forecasting of electricity rates

Machine learning (ML) has the potential to significantly alter the course of electricity rates. Its impact is very significant because it can examine large datasets with a level of sophistication never seen before. Machine learning models demonstrate a noteworthy capacity to identify complex patterns in these data environments, offering a thorough comprehension of the subtle variables affecting energy usage. The addictiveness of ML models—, which allow them to dynamically adapt to changing conditions and continuously optimize predictions—is what sets them apart from other forms of influence. The incorporation of machine learning (ML) into forecasting models improves their accuracy and increases their responsiveness to the dynamically changing patterns of electricity consumption (Chadalavada and Rekha (2020)). This improvement in forecasting capabilities represents a paradigm shift rather than merely a technology advancement, allowing energy sector stakeholders to better navigate the complexity of demand.

Machine learning (ML) is playing a revolutionary role in affecting power tariffs, ushering in an important change in the energy sector. The potential impact of machine learning is enormous, owing to its ability to examine massive datasets with increasing sophistication. ML models, in contrast to traditional approaches, have a remarkable ability to discover detailed patterns within various data settings, providing a deep insight into the numerous aspects driving energy usage. One defining aspect of machine learning (ML) models is their adaptability, which allows them to adjust dynamically to changing circumstances and continuously optimize predictions. This versatility distinguishes ML from traditional approaches, allowing it to record and adjust to the changing dynamics of electricity usage. Because ML is adaptive, predictive algorithms not only keep up with but also stay a step ahead of the competition's continuously changing patterns of energy demand.

There is a noticeable gain in accuracy as well as responsiveness when ML is used in forecasting models. ML models improve their ability to forecast by learning from prior data and detecting subtle correlations that traditional approaches may miss. This increase in precision is more than just a technological achievement; it signifies an essential change in how the energy sector understands and navigates the complexities of demand. Energy stakeholders, including energy companies and politicians, stand to benefit considerably from ML integration (Arslan (2022b)). Forecasting models with greater accuracy and adaptability enable decision-makers to make educated decisions about resource allocation, infrastructure design, and energy pricing methods. The impact of ML on the price of electricity is more than just a result of enhanced technology; it also represents a strategic breakthrough that links the renewable energy sector with changing consumer patterns, encouraging a more robust and responsive strategy for managing demand challenges.

2.2 Strategies in machine learning to predict electricity consumption

Applying machine-learning techniques strategically increases the predictive power of models. In-depth techniques are covered in this section, along with a toolkit that can be used to improve the precision and resilience of machine learning algorithms used to forecast electricity consumption. Model input is optimized through feature engineering, a sophisticated process of variable selection and transformation (Blázquez-García and Barrio (2020b)). Model assembling—the skillful fusion of various models—captures the strengths of the group, and hyperparameter tuning refines model arrangements. Together, these calculated moves create a complex framework that protects the machine learning algorithms from the complexities and uncertainties present in the ever-changing field of power consumption prediction.

Strategic techniques are critical in enhancing model accuracy and flexibility when estimating electricity usage using machine learning. Feature engineering is a critical method that entails a rigorous process of choosing and modifying variables. This method ensures that artificial intelligence models have the most pertinent characteristics, increasing their ability to grasp the complicated patterns that accompany electricity usage. Model assembling presents a strategic approach through the smart combination of multiple models. This ensemble technique capitalizes on the capabilities of individual models to produce a collective effect that frequently outperforms the forecasting accuracy of any one model(Parizad and Hatziadoniu (2021)). This technique strengthens the entire forecasting framework's durability and adaptability by exploiting the diversity of numerous models. Another strategic approach aimed at optimizing machine learning model setup is hyperactive parameter tuning. This technique entails fine-tuning a model's hyper parameters to ensure it is perfectly matched to the individual intricacies of power usage patterns. This painstaking tweaking improves the models' accuracy and reliability, making them more capable of reflecting the changing dynamics of energy usage.

These strategic movements establish an extensive framework that strengthens machine-learning techniques against the challenges and unpredictability buried in the constantly shifting environment of electrical consumption prediction. Practitioners can build durable and highly effective predictive systems by focusing on feature optimization, model ensemble approaches and hyperparameter tuning.

2.3 ARIMA Model

Combining auto-regression, differentiation, and moving averages, ARIMA (An autoregressive integrated continuing Average) is a powerful time series forecasting model. With the help of this dynamic trio, ARIMA can effectively identify patterns and abnormalities in non-stationary data. Through the identification of past trends, ARIMA offers a basis for accurate future value prediction, providing priceless insights into the changes in time influencing the phenomenon in question.

ARIMA's power rests in its capacity to catch detailed temporal patterns and adapt to different types of input. ARIMA's differencing method is extremely useful in non-stationary contexts where statistical features alter with time. This modification stabilizes the average and variance, allowing for reliable modeling. The use of auto-regression recognizes pattern persistence throughout time, resulting in a deeper comprehension of temporal connections. Furthermore, ARIMA excels at anomaly identification by detecting anomalies or irregularities in time series data (Al-Musaylh and Li (2018)). Anomalies frequently reflect changes in underlying structures and ARIMA's ability to detect such deviations makes it a useful tool for identifying unforeseen occurrences or fluctuations. This function is very useful in estimating electricity consumption because rapid swings can be symptomatic of external events such as economic shifts, environmental legislation, or unanticipated catastrophes are all possibilities.

ARIMA's capacity to discern historical trends is invaluable as a basis for future value prediction. ARIMA presents a solid foundation for making accurate forecasts about future values by understanding previous trends and the effects they have on time series. This is especially important in the forecasting of energy use, where recognizing historical consumption trends is critical for projecting future demands and maximizing resource allocation (Al-Musaylh and Li (2018)). ARIMA is a strong and versatile time series forecasting model due to its use of auto-regression, differentiation, and moving averages. Its capacity to adapt to unpredictable data, anomaly detection capacity, and emphasis on historical patterns provide it with the instruments needed to deliver crucial insights into the temporal changes influencing the occurrence under examination. When it comes to power, ARIMA appears as a great asset for precise and informed forecasts in consumption planning, where the relationship of many components needs a sophisticated approach.
Details regarding the number of yearly publications based on the rate of electricity consumption using the ARIMA model [1] are shown in Figure 1.

[1] https://www.mdpi.com/1996-1073/14/23/7952

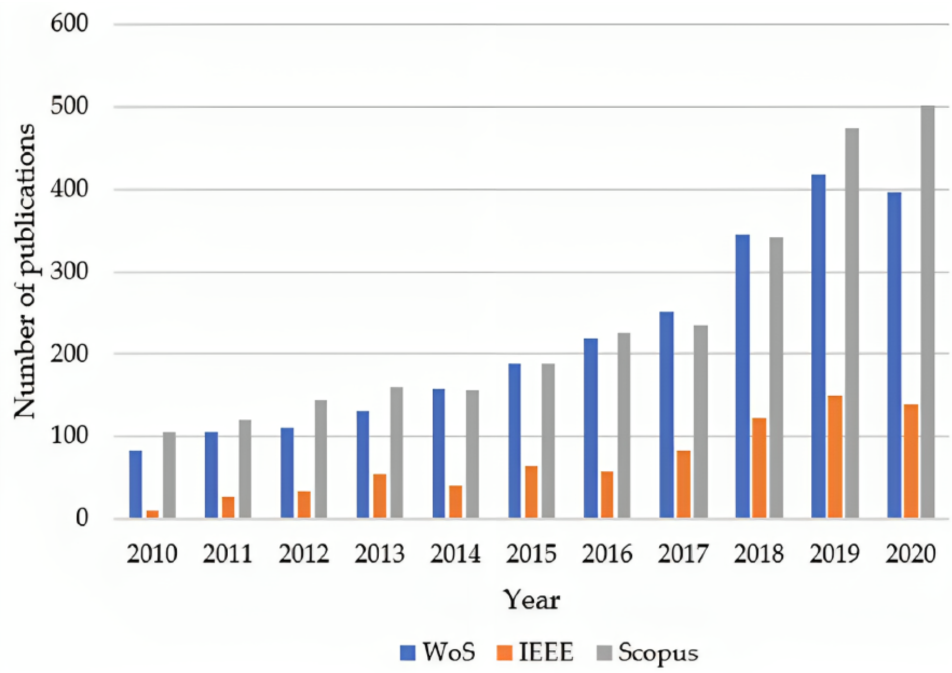

Figure 1: Publications based on the rate of electricity consumption using the ARIMA model

2.4 SARIMA Model

By smoothly incorporating seasonality through its predictive powers, SARIMA (Seasonal ARIMA) expands the capabilities of ARIMA. When recurrent patterns over predetermined intervals have a substantial impact on time series data, this model performs exceptionally well. SARIMA's incorporation of seasonally induced variations improves its adaptability and provides a more nuanced understanding of the temporal dynamics influencing observed values.

The strength of SARIMA is its ability to properly handle time-series information with recurring patterns over set intervals. This makes it especially useful in scenarios where seasonality is important, such as forecasting electricity usage. The predictive power of the model is used to capture and explain the impact of seasonally driven fluctuations, offering a full knowledge of its fundamental temporal dynamics. SARIMA, for example, can account for seasonal surges in the consumption of energy during adverse temperatures or holidays when forecasting electricity usage. Because of its precise inclusion of seasonality, SARIMA is an invaluable tool for acquiring subtle insights concerning the temporal nuances affecting observed values.

2.5 Prophet Model

The team 'Meta' has created the Prophet model, which shows promise as a flexible tool for time series forecasting especially when using everyday observations. Prediction accuracy is increased by its inherent flexibility, which makes it possible to seamlessly accommodate irregularities like holidays and outliers. Prophet's ability to adapt to a variety of data patterns makes it a useful tool in situations where unpredictability and variability are prevalent. In forecasting, the Prophet model is a useful tool for navigating through complex temporal intricacies and predicting unpredictable changes(Chadalavada and Rekha (2020)).

The Prophet model developed by Facebook marks a substantial advancement in time series projections, particularly when considering everyday observations. The framework's built-in adaptability is a prominent feature, improving forecast accuracy by accepting oddities like vacations and outliers effortlessly. Prophet's capacity to adapt to a wide range of data patterns makes it a powerful tool in settings defined by uncertainty and variability. Its applicability extends to instances where conventional forecasting models may struggle, providing a reliable answer for negotiating complicated temporal subtleties. The Prophet model appears as a trustworthy and versatile ally in the area of forecasting, particularly where unanticipated changes are common, helping analysts to make better-informed judgments by extracting valuable insights from different and dynamic information. A sample energy demand forecast plot using Prophet [2] is shown in Figure 2

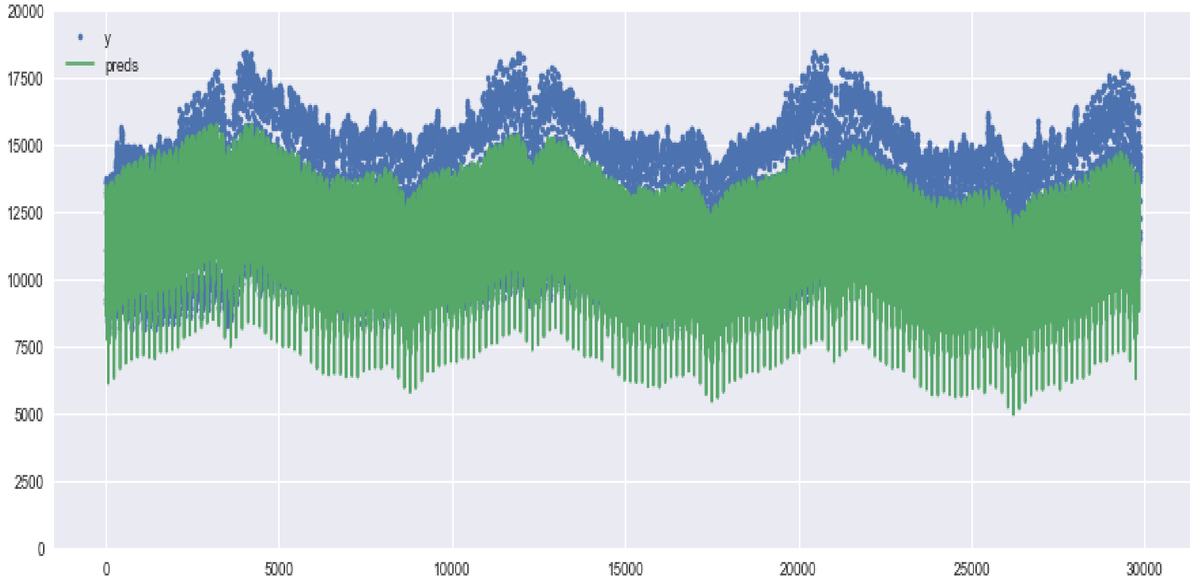

Figure 2: A Sample energy demand forecast plot using Prophet(Actual-Blue, Forecast-Green)

2.6 Characteristics of the Arima, Sarima, and Prophet Model

ARIMA's ability to capture underlying trends is based on its reliance on auto-regression, distinction, and moving averages. These capabilities are expanded by SARIMA to take seasonal patterns into account. Prophet incorporates outliers and holiday effects into its daily observation-forecasting model, further augmenting its flexibility.

ARIMA model: The ARIMA model is built around three important components: auto-regression (AR), distinction (I), and the concept of moving averages (MA). Auto-regression predicts future outcomes based on previous findings, differencing converts data that is not stationary into a more regular form and moving averages smoothing out volatility (Al-Musaylh and Li (2018)). These properties enable ARIMA to detect and capture hidden patterns of time series data. It works especially well in situations when the previous sequence of measurements influences future values. ARIMA's adaptability

[2] https://www.linkedin.com/pulse/forecasting-energy-consumption-machine-learning-facebook-sharma/

to a wide range of information patterns makes it a flexible model for a wide range of applications.

SARIMA model: SARIMA extends ARIMA's capabilities by introducing seasonality. SARIMA adds seasonal characteristics to auto-regression, distinguishing and moving averages, resulting in it being well-suited for sequences of data influenced by recurring patterns over periods specified (Abu-Salih and Huneiti (2022)). This model is useful in situations where seasonal fluctuations have a substantial impact on observable data, such as forecasting power use. The properties of SARIMA increase its adaptability, permitting it to document the intricacies of seasonally driven oscillations.

Prophet model: The Prophet presents novel features to handle difficulties that traditional models cannot fully address. Its capacity to account for outliers and vacation impacts in its regular observation-forecasting framework is one of its distinguishing features. This feature improves Prophet's adaptability and versatility, making it more resilient to anomalies and unforeseen occurrences in the data (Chadalavada and Rekha (2020)). The model's emphasis on capturing daily fluctuations is ideally suited to applications that need high-frequency data analysis. Prophet's qualities make it appropriate for circumstances in which the data demonstrates irregular patterns and the impact of vacations and outliers is significant.

2.7 Literature gap

The extant literature provides insightful analyses of electricity consumption forecasting; however, there is a conspicuous void in the discussion of the combined effects of ARIMA, SARIMA, Auto ARIMA and the Prophet Model, as well as the incorporation of machine learning techniques. By offering a comprehensive viewpoint, this study seeks to close this gap and add to the continuing conversation about improving estimates of electricity consumption. This study aims to provide a comprehensive understanding by combining the advantages of both modern algorithms for machine learning and traditional time series models. This will help forecasting methodologies better navigate the complexities of changing consumer behaviour and dynamic economic shifts.

3 Methodology

3.1 Introduction

The purpose of this study is to investigate the feasibility of using machine learning algorithms, namely ARIMA, SARIMA, Auto ARIMA and the Prophet Model to forecast electrical consumption in the Australian state of New South Wales. Insights into consumption patterns are provided by the dataset, which spans the years 2018 to 2022. This research aims to shed light on energy consumption in order to help with sustainable resource management and well-informed decision-making. This chapter will explain the techniques and processes that have been used throughout the research work. However, this chapter will explain the data processing techniques and selected machine learning algorithms and the entire flow of work.

3.2 Brief about dataset

The selected dataset for this research work is associated with the New South Wales, Australia. This dataset contains data on demand and prices collected throughout time by the Australian Energy Market Operator (AEMO). Also the dataset contains daily weather data retrieved from the Bureau of Meteorology-

a. Observatory Hill is located in Greater Sydney at 066214.

b. An AWS signal station in Newcastle Nobbys (61055) serves the Newcastle and Lake Macquarie areas.

c. Shellharbour Airport, Albion Park, 68241 Illawarra

Data from the most recent available day has been used to fill up any gaps. There is now only one file including all of the weather station data.

Dataset Link:[3]

3.3 ARIMA model

ARIMA stands for Autoregressive Integrated Moving Average. Autoregressive captures the relationship between current value and previous values. Integration is the process of making data stationary. Stationary data means the properties like mean and variance remain constant over time (Al-Musaylh and Li (2018)). Moving average Smoothness sudden and extreme changes in data.

3.4 Auto-ARIMA model

Auto ARIMA is also a time series forecasting model with capabilities to figure out best hyper parameters for ARIMA models to achieve higher accuracy. It is self capable to figure out the best settings for the forecasting model.

3.5 SARIMA model

SARIMA stands for Seasonal AutoRegressive Integrated Moving Average. SARIMA is a time series forecasting model that extends the capabilities of the ARIMA model to handle seasonality in the data (Abu-Salih and Huneiti (2022)). While ARIMA is effective for non-seasonal data, SARIMA is designed to capture and model the seasonal patterns and variations that occur at regular intervals.

3.6 Prophet model

Prophet Model is a time series forecasting machine learning model developed by Facebook. The Prophet model has capabilities to capture weekly, monthly and yearly trends in the data-set (Chadalavada and Rekha (2020)). It is able to identify patterns in the data, adjust for holidays and special events, and handle unexpected changes.

3.7 Summary

In order to accomplish the research work, the above mentioned algorithms such as ARIMA, Auto-ARIMA, SARIMA and finally the Prophet model have been used. All these

[3] `https://www.kaggle.com/datasets/neosh11/electricity-nsw-australia-201801-202209`

above mentioned algorithms have been used throughout the research and finally the models have been compared to derive the best model for electricity consumption forecasting.

4 Design Specification

The below architecture shown in Figure 3 is associated with the forecasting of electricity consumption research work. Throughout the entire process, the best algorithm for electricity consumption forecasting has been derived. In order to do so, at the first step one dataset has been selected. After that, data preprocessing technique has been performed. After that statistical modelling has been performed with ARIMA, Auto-ARIMA, SARIMA. Autoarima has been used here to choose best parameters for the model so that overfitting can be avoided. Also on the other hand Prophet model has been used to forecast the electricity consumption rate. Then after completion of both of these steps, models have been evaluated with the help of Mean Absolute Error(MAE), Mean Squared Error(MSE), Root Mean Squared Error(RMSE). After complete evaluation of the models, they have been compared. And finally, the forecasting of electricity consumption is done with higher accurate model having minimum error rate.

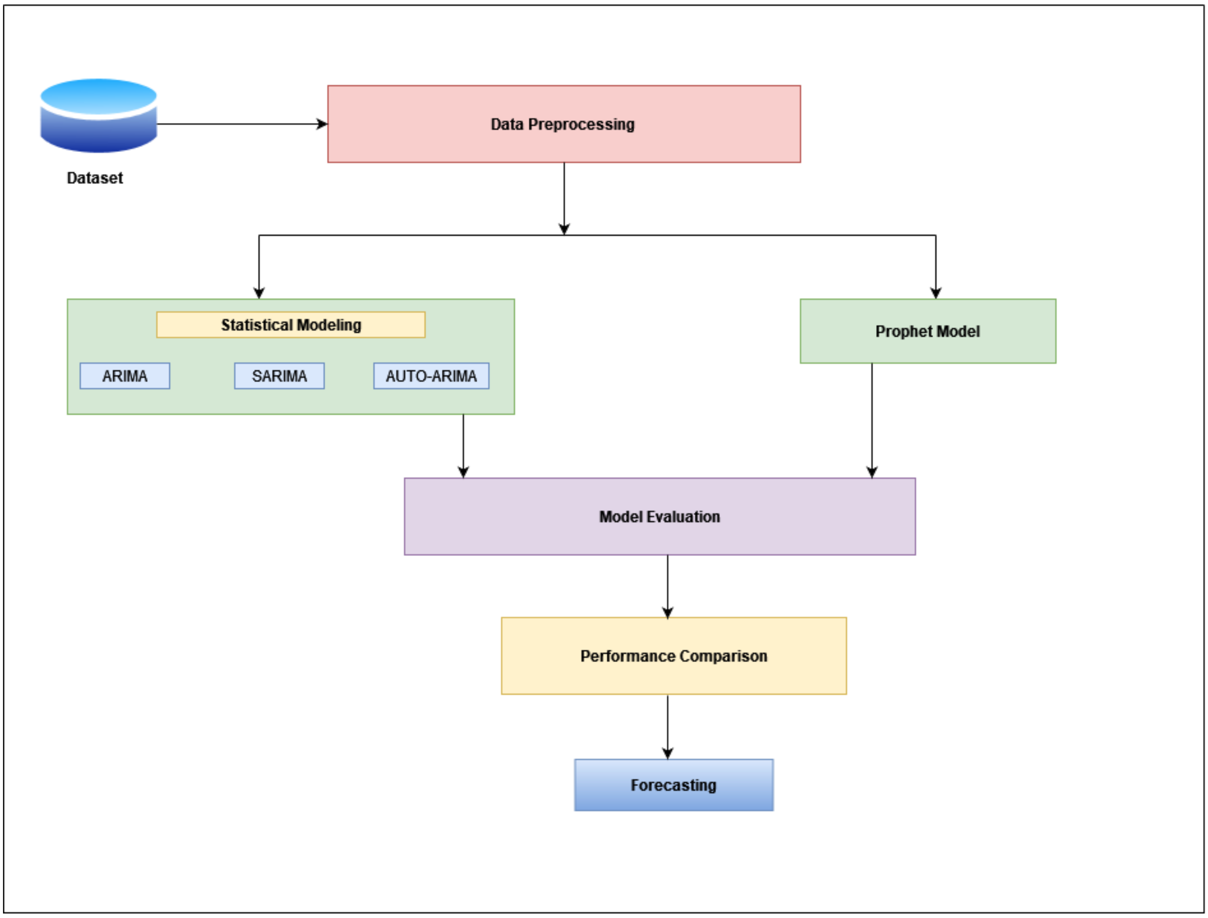

Figure 3: Overall architecture for forecasting of electricity consumption

5 Implementation

5.1 Data Preprocessing and Analysis

The graph shown in Figure 4 depicts the total demand and Regional Reference Price(RRP) over time. The total demand is higher than the RRP and the RRP is lower than the total demand. For a while in 2020, RRP surpasses the total demand but overall total demand is higher than the RRP.

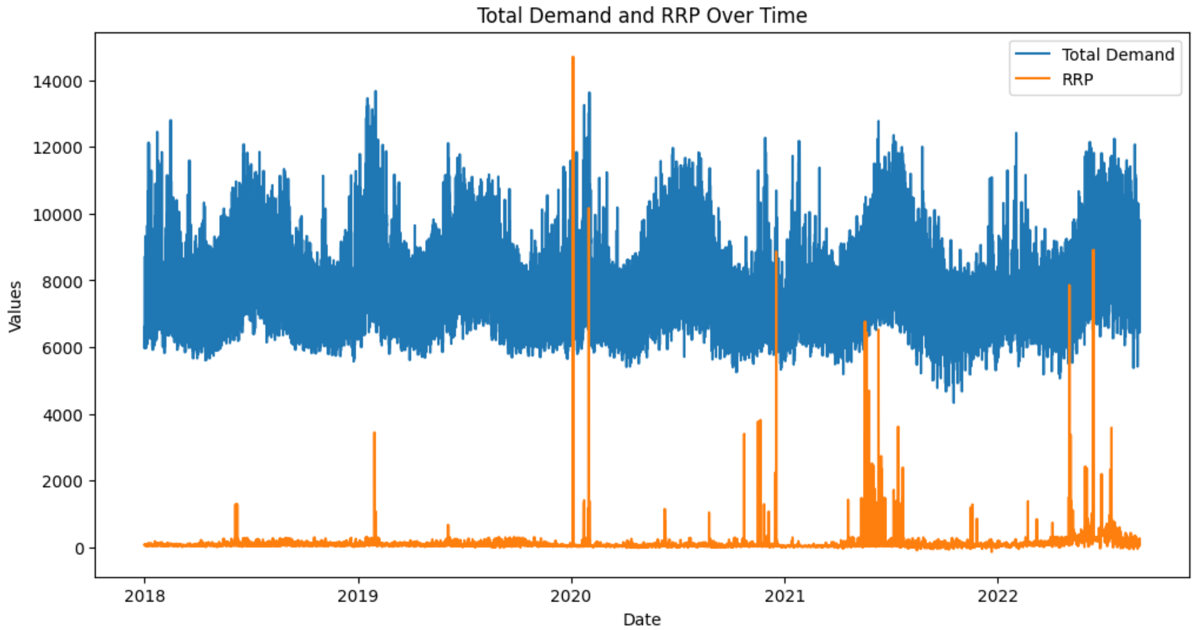

Figure 4: Total Demand and Regional Reference Price(RRP) Over Time

Trend:

The trend component shows the overall direction of the data over time. In this case, the trend is increasing, which means that electricity consumption in NSW is increasing overall.This can be observed in Figure 5.

Seasonal:

The seasonal component shows the cyclical fluctuations in the data that occur over a regular period of time. In this case, the seasonal component shows a clear annual cycle, with electricity consumption being higher in the summer months and lower in the winter months. This can be observed in Figure 5.

Residual:

The residual component shows the remaining variation in the data after the trend and seasonal components have been removed. This component is typically caused by random events or factors that are not easily explained. This can be observed in Figure 5.

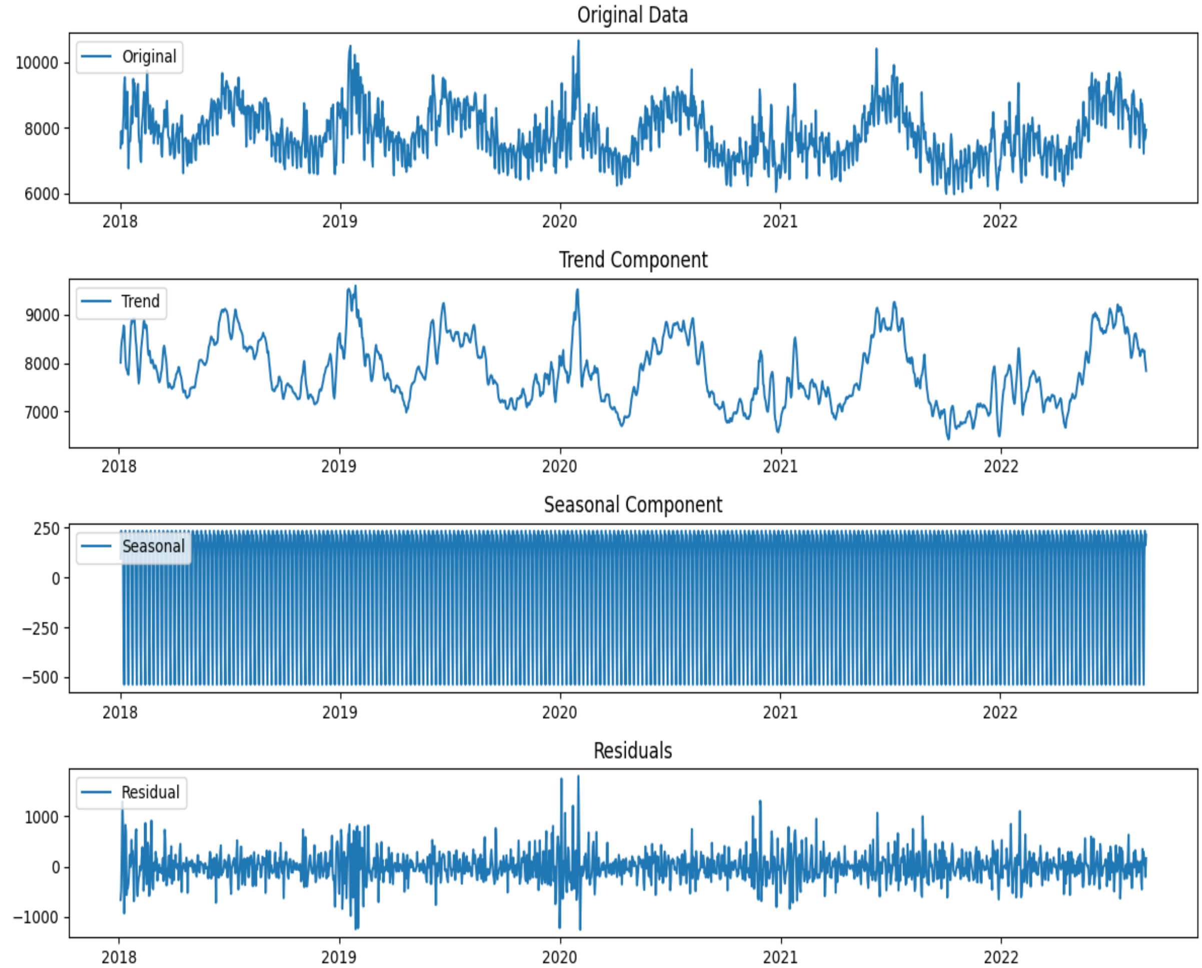

Figure 5: Original Vs. Trend Vs. Seasonal Vs. Residual

The Autocorrelation Function(ACF) plot shown in Figure 6 depicts the correlation between the TOTALDEMAND and itself at different time lags. The ACF plot in the image shows that the signal is highly correlated with itself at lags of 1, 2, and 3. This indicates that the signal has a strong seasonal pattern.

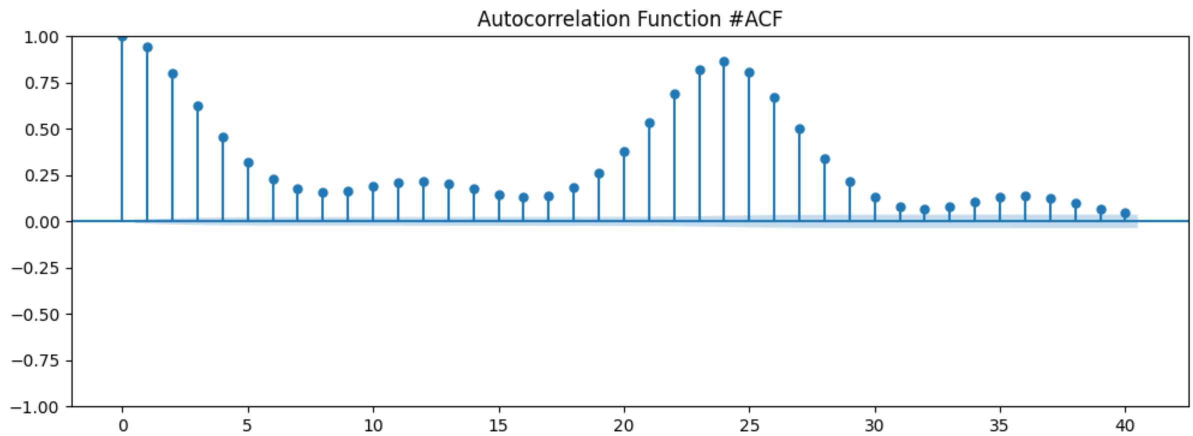

Figure 6: Autocorrelation Function ACF

The TOTALDEMAND plot in Figure 7 shows that the signal is significantly correlated with itself at a lag of 1. This indicates that there is a strong short-term relationship in the TOTALDEMAND.

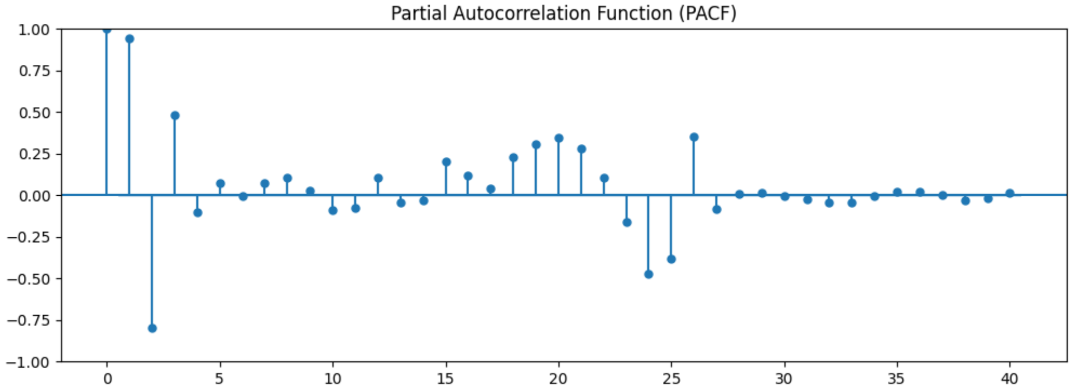

Figure 7: Partial Autocorrelation Function (PACF)

The weekly electricity demand time series shown in Figure 8 depicts a clear seasonal pattern, with higher demand in the summer months and lower demand in the winter months. The trend is increasing, indicating that overall demand is growing over time. The residual component shows some spikes, which could be due to major events such as heatwaves or cold snaps.

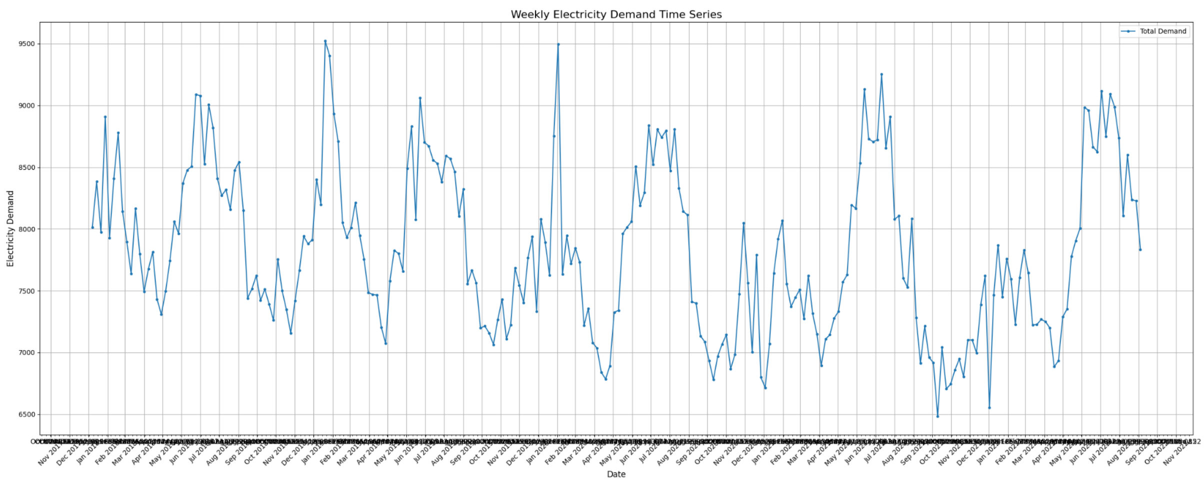

Figure 8: Weekly electricity demand time series

The monthly electricity consumption time series in NSW, Australia shown in Figure 9 depicts a clear seasonal pattern, with higher consumption in the summer months and lower consumption in the winter months. The trend is increasing, indicating that overall consumption is growing over time. The residual component shows some spikes, which could be due to major events such as heatwaves or cold snaps.

In short, the monthly electricity consumption time series is increasing overall, with a clear seasonal pattern.

Monthly electricity consumption in NSW, Australia is increasing overall, with a clear seasonal pattern.

This analysis captures the two most important insights from the graph: the trend and the seasonality.

Figure 9: Original Vs. Trend Vs. Seasonal Vs. Residual

The graph shown in Figure 10 depicts the median monthly electricity consumption in NSW, Australia from 2018 to 2022.

The graph shows a clear seasonal pattern, with higher consumption in the summer months and lower consumption in the winter months. The trend is increasing, indicating that overall consumption is growing over time.

Mean monthly electricity consumption in NSW, Australia is increasing overall, with a clear seasonal pattern.

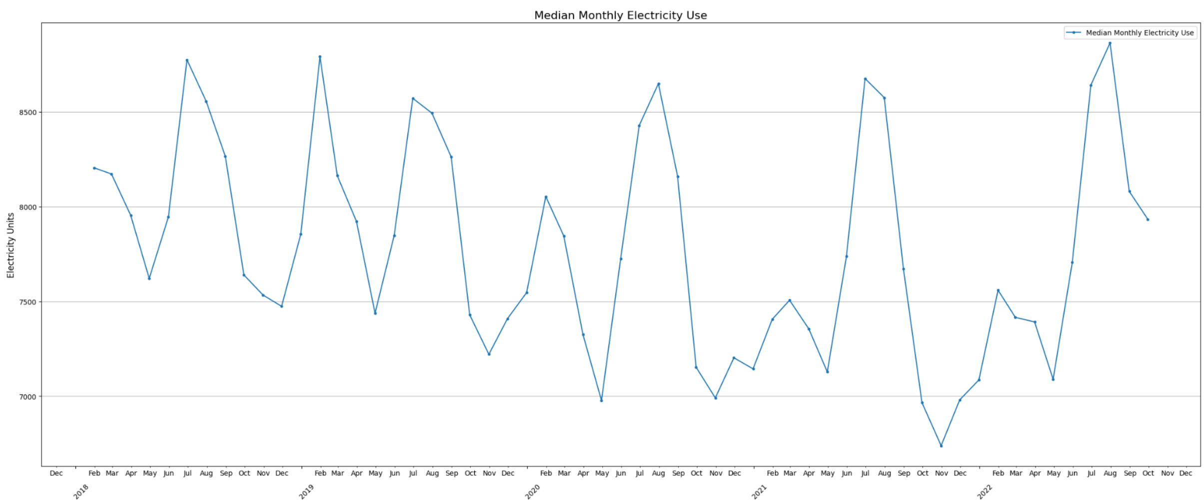

Figure 10: Median monthly electricity consumption

6 Evaluation

The purpose of this section is to provide a comprehensive analysis of the results and main findings of the study as well as to find the best accurate model for electricity forecasting.

6.1 Experiment of ARIMA Model

Training Set Results:

MAE (Mean Absolute Error): On average, our prediction for TOTALDEMAND was off by about 462.99 units.

MSE (Mean Squared Error): When squaring the errors, summing them up, and then finding the average, a MSE of 1685493.32 is obtained. This metric emphasizes larger errors. In simpler terms, it penalizes bigger mistakes more.

RMSE (Root Mean Squared Error): Taking the square root of MSE, this study calculates RMSE, which is about 1298.27. This is like saying, "On average, the prediction was off by approximately 1298.27 units." It gives us a sense of the typical size of our prediction errors.

Test Set Results:

MAE (Mean Absolute Error): In the real world (the test set), our prediction was off by about 550.17 units, on average. This is like saying, "In actual scenarios, the average error was around 550.17 units."

MSE (Mean Squared Error): Squaring the errors, summing them up, and averaging gives us a MSE of 423498.21. Again, it highlights larger errors more than smaller ones.

RMSE (Root Mean Squared Error): The square root of MSE is RMSE, which is about 650.77. This metric gives us a sense of the typical size of errors in the real-world predictions.

The comparison between the actual and predicted values is shown in Figure 11.

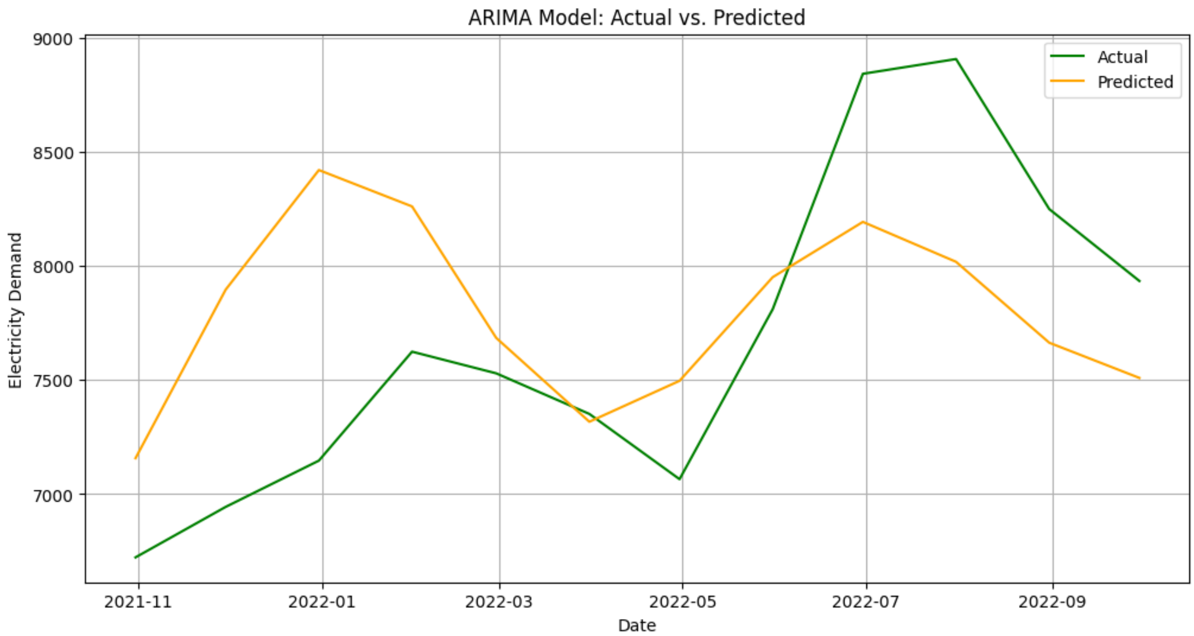

Figure 11: ARIMA Model: Actual Vs. Predicted

6.2 Experiment of Auto - ARIMA Model

Auto ARIMA model evaluation results:

Mean Absolute Error (MAE): 463.20

The MAE represents the average absolute difference between the actual and predicted values. In this case, the model's predictions, on average, deviate by approximately 463.20 units from the actual electricity demand.

Mean Squared Error (MSE): 304736.49

The MSE measures the average squared difference between actual and predicted values. A lower MSE indicates better accuracy. Here, the squared differences between predictions and actual values average to approximately 304736.49.

Root Mean Squared Error (RMSE): 552.03

The RMSE is the square root of the MSE, providing a measure of the average magnitude of prediction errors. In this context, the RMSE is around 552.03, indicating the typical error in the model's electricity demand predictions.

The comparison between the actual and predicted values is shown in Figure 12

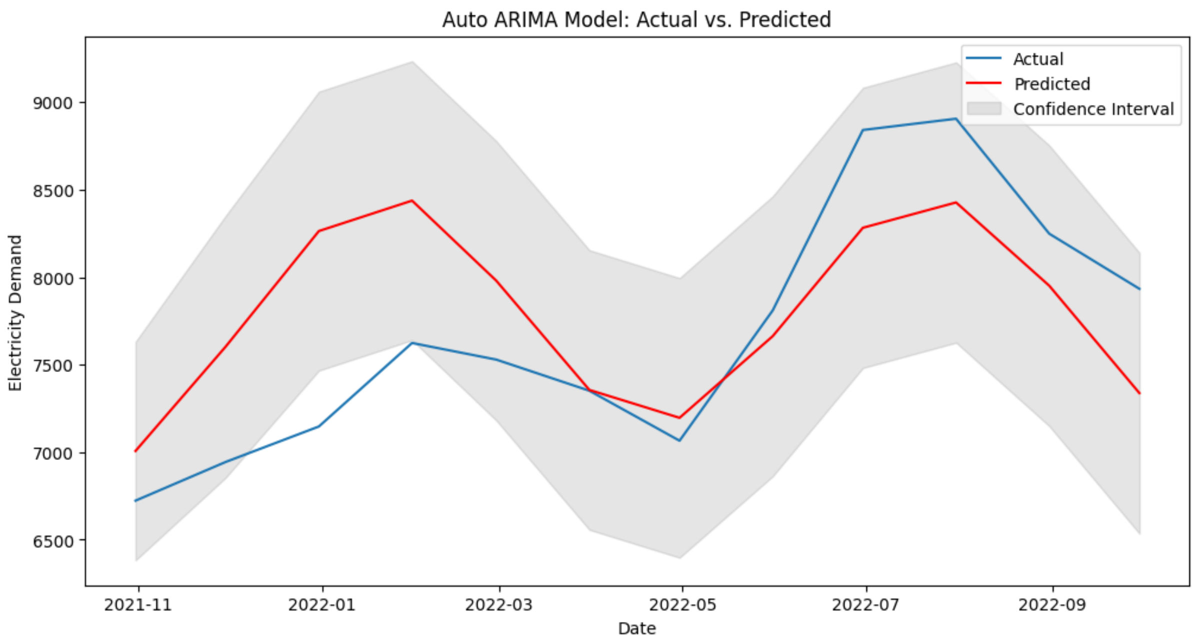

Figure 12: Auto ARIMA Model: Actual Vs. Predicted

6.3 Experiment with SARIMA Model

Training Set Metrics:

MAE (Mean Absolute Error): On average, the predictions on the training set are off by approximately 602.41 units.

MSE (Mean Squared Error): The average squared difference between the predictions and the actual values on the training set is about 2000218.66.

RMSE (Root Mean Squared Error): The RMSE, which is the square root of MSE, is around 1414.29. It provides a measure of how spread out the errors are.

Test Set Metrics:

MAE (Mean Absolute Error): The predictions on the test set have an average absolute error of approximately 324.36 units.

MSE (Mean Squared Error): The average squared difference between the predictions and the actual values on the test set is about 153552.35.

RMSE (Root Mean Squared Error): The RMSE for the test set is approximately 391.86, indicating the spread of errors in the predictions.

The comparison between the actual and predicted values is shown in Figure 13

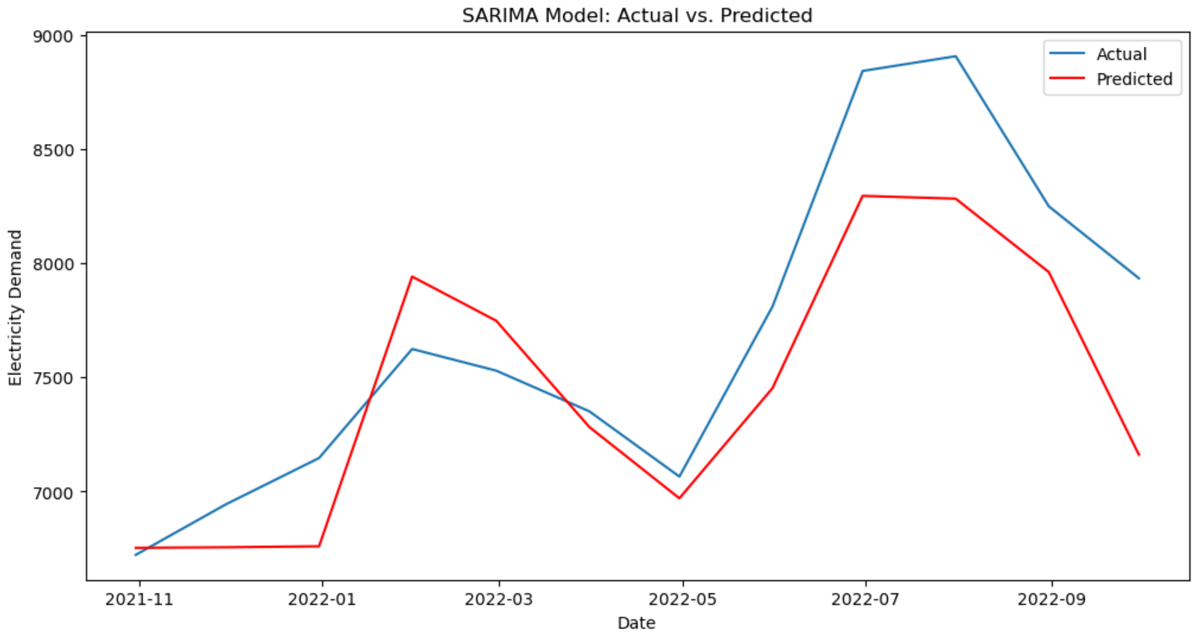

Figure 13: SARIMA Model: Actual Vs. Predicted

6.4 Experiment of Prophet Model

The Prophet model was trained and evaluated on the time series dataset. Here are the metrics for both the training and test sets:

Training Set Metrics:

Mean Absolute Error (MAE): 128.72

Mean Squared Error (MSE): 28644.52

Root Mean Squared Error (RMSE): 169.25

Test Set Metrics:

Mean Absolute Error (MAE): 2103.20

Mean Squared Error (MSE): 5969665.51

Root Mean Squared Error (RMSE): 2443.29

The comparison between the actual and predicted values is shown in Figure 14

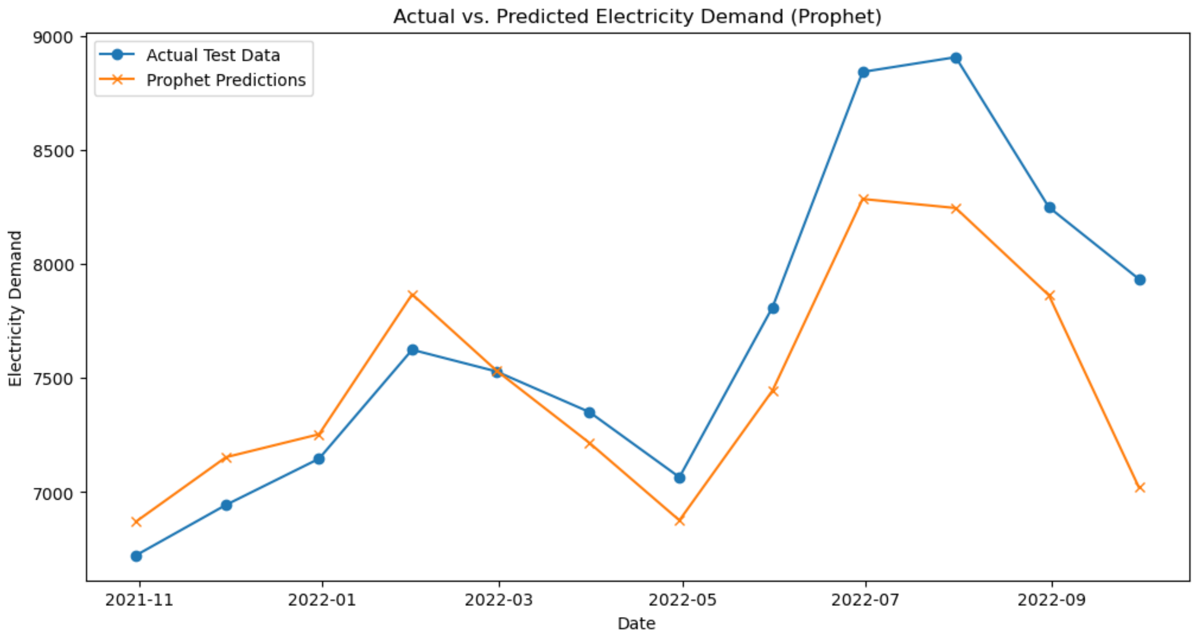

Figure 14: PROPHET Model: Actual Vs. Predicted

6.5 Discussion

The Auto ARIMA model excels at managing the complex linkages and seasonal patterns seen in the power consumption time series, which is why it produces better results. Researchers can reduce the possibility of biases and mistakes in the model construction using Auto ARIMA's automated parameter selection method.

7 Conclusion and Future Work

In conclusion, the machine learning models, especially ARIMA, Auto ARIMA, SARIMA, and Prophet, showcase varying degrees of accuracy in predicting electricity consumption in New South Wales. While each model has its strengths and limitations, they collectively contribute valuable insights for optimizing energy resource management and decision-making.

Among the evaluated models, the Auto ARIMA model stands out as the top performer in predicting electricity consumption trends in New South Wales. Its special ability lies in automating the selection of optimal hyperparameters, enabling it to adapt dynamically to the dataset's characteristics. This capability contributes to higher accuracy by fine-tuning the model's settings for improved predictions. This is the main key factor enhancing the accuracy of timeseries forecasting for electricity consumption.

Forecasting is the prediction of future energy needs. Predicting the energy consumption of different entities over a certain time frame is what this process is all about. These entities might be families, businesses, or even entire regions. The weather, economic activity, demographic changes, and technology developments are just a few of the factors that might impact these forecasts.

Energy demand forecasting is being revolutionized by machine learning. Unlike people or more conventional statistical models, it can sift through mountains of data in search of hidden patterns. AI has the ability to learn from data and conduct simultaneous analyses of various factors, resulting in progressively improved accuracy. This allows for more accurate predictions of future energy consumption, which in turn may lead to greater efficiency, lower carbon emissions, and significant savings.

In energy-related operations, improving machine learning models can be challenging due to data quality difficulties and the absence of standardized datasets. Making machine learning models understandable and explicable is another obstacle. Another difficulty in applying machine learning to energy-related tasks is the potential, social and ethical consequences. Another obstacle is the need for integration with pre-existing energy infrastructure and regulatory frameworks. When it comes to machine learning for energy, data privacy and security are major factors to think about. It is quite difficult to do global optimization with building energy modeling software tools since they demand the manual(human) entry of particular inputs. There are several obstacles to overcome in order to train, evaluate, embed and generate trustworthy machine learning models using physics-based simulation data.

The Auto ARIMA model's superior performance can be attributed to its adeptness in handling the intricate relationships and seasonality patterns present in the electricity consumption time series. By automating the process of parameter selection, Auto ARIMA minimizes human intervention, reducing the risk of biases and errors in the model configuration. In future research work, deep learning concepts can be used and a hybrid ML approach can also be considered.

8 Acknowledgment

I would want to express my appreciation to the National College of Ireland for offering such wonderful tools. A special appreciation goes to my thesis supervisor, Mr. Jorge Basilio, for his constant input and helpful direction during the process. Finally, I would want to sincerely thank my parents for their consistent support, which has allowed me to realise all of my potential. I sincerely hope that this effort will demonstrate the knowledge I have gained from this course and uphold the values of spreading reliable information in my research.

References

Abu-Salih, B., W. P. M. G. C. K. A.-O. M. and Huneiti, A. (2022). Short-term renewable energy consumption and generation forecasting: A case study of western australia., *Heliyon,* **8**(3).

Ahmed, T., M. K. and Agalgaonkar (2012). Climate change impacts on electricity demand in the state of new south wales, australia., *Applied Energy* (98): pp.376–383.

Al-Musaylh, M.S., D. R. A. J. and Li, Y. (2018). Short-term electricity demand forecasting with mars, svr and arima models using aggregated demand data in queensland, australia., *Advanced Engineering Informatics,* (35): pp.1–16.

Arslan, S. (2022a). A hybrid forecasting model using lstm and prophet for energy consumption with decomposition of time series data., *PeerJ Computer Science* (8): p.e1001.

Arslan, S. (2022b). A hybrid forecasting model using lstm and prophet for energy consumption with decomposition of time series data., *PeerJ Computer Science* (8): p.e1001.

Blázquez-García, A., C. A. M. A. S. R. and Barrio, I. (2020a). Short-term office building elevator energy consumption forecast using sarima, *Journal of Building Performance Simulation* **13**(1): pp.69–78.

Blázquez-García, A., C. A. M. A. S. R. and Barrio, I. (2020b). Short-term office building elevator energy consumption forecast using sarima., *Journal of Building Performance Simulation,* **13**(1): pp.69–78.

Chadalavada, R.J., R. S. and Rekha, V. (2020). Electricity requirement prediction using time series and facebook's prophet., *Indian Journal of Science and Technology,* **13**(47): pp.4631–4645.

Mittal, D.A., L. S. and Xu, G. (2020). November. electricity price forecasting using convolution and lstm models., *In 2020 7th International Conference on Behavioural and Social Computing (BESC)* pp. pp. 1–4.

Parizad, A. and Hatziadoniu, C. (2021). Using prophet algorithm for pattern recognition and short term forecasting of load demand based on seasonality and exogenous features., *In 2020 52nd North American Power Symposium (NAPS)* pp. pp.1–6.

Preniqi, Vjosa, B. K. M. D. T. E. F. G. M. A. A. and Konur, S. (2020). Comparative study of shortterm electricity price forecasting models to optimise battery consumption., *In 2020 International Conferences on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics)* pp. 342–349.