JOURNAL OF STATISTICAL MODELING & ANALYTICS (JOSMA)
(ISSN: 2180-3102)
UNIVERSITI MALAYA

The Box-Jenkins Model Approach in Forecasting Electricity Demand

Balqis Husna Md Said, Siti Roslindar Yaziz*, Nor Azuana Ramli, Syarranur Zaim & Roslinazairimah Zakaria

Centre of Mathematical Sciences, Universiti Malaysia Pahang Al-Sultan Abdullah, Lebuh Persiaran Tun Khalil Yaakob 26300 Kuantan, Pahang

Corresponding author: roslindar@umpsa.edu.my

Received 15 April 2026
Accepted 01 Sept 2026
Published 20 Nov 2026

RESEARCH ARTICLE

Abstract

Accurate forecasting of electricity demand plays a critical role in ensuring system reliability, guiding infrastructure development, and enhancing the efficiency of national energy planning. Recognizing the presence of strong weekly seasonality in Malaysia's electricity consumption patterns, this study applies the Box-Jenkins SARIMA framework to develop a model tailored to short-term forecasting needs. Using total weekly electricity demand data from 2022 to 2024, the study evaluates several SARIMA models. The SARIMA$(0,1,0)(0,1,1)_{52}$ model emerged as the optimal structure, achieving a Mean Absolute Percentage Error (MAPE) of 5.57%. The result meets widely accepted the forecasting accuracy standards and reflects the model's robustness in handling seasonal variability. These findings support the applicability of SARIMA models in operational planning and grid management, especially in contexts where timely and interpretable forecasts are essential. This study reinforces the value of seasonal time series approaches in addressing Malaysia's evolving electricity demand and paves the way for future enhancements using hybrid forecasting frameworks.

Keywords: Box-Jenkins; Forecast; Electricity Demand; SARIMA; Time Series

1. Introduction

Electricity is one form of energy that exists in this world, and for Malaysia, electricity first made its appearance at the turn of the 20th century. The earliest known record of power generation dates to 1894 in a small mining town in Rawang, Selangor, where both Loke Yew and Thanboosamy Pillai installed an electric generator to operate their tin mines (The Sun, 2024). This marked a pivotal legal and industrial precedent in Malaysia's infrastructure history. From a regulatory standpoint, this milestone represented the informal start of energy governance before formal electricity regulations were institutionalized. Analytically, this small act of electrification in Rawang laid the groundwork for a progressive expansion of power generation infrastructure across the nation. As demand grew and urbanization accelerated, electricity quickly transformed from a niche mining utility into a national public good. The exponential growth in consumption, powered by industrialization, urban advancement, and digital transformation, has made electricity a vital force in society's development supporting education, commerce, healthcare, and national security.

Originally, electricity in Malaysia was predominantly generated by petrol-based generators, which, while innovative at the time, posed significant limitations in terms of scale efficiency, and environmental impact. The use of fossil fuels introduced cost volatility, carbon emissions, and mechanical unreliability, highlighting the need for innovation in the energy sector. With the advancement of technology and global shifts toward sustainability, Malaysia embraced more efficient and greener energy sources. One prominent leap forward was the commissioning of the Bakun Hydroelectric Plant in Sarawak. As one of the most powerful hydroelectric power facilities in Malaysia, it boasts a total installed generation capacity of 2400 MW, reinforcing its role as a cornerstone of national energy supply (Nabipour Afrouzi et al., 2021). This plant exemplifies the nation's strategic pivot toward renewable energy. The Bakun Dam not only reduces Malaysia's dependence on imported fossil fuels but also contributes to its carbon reduction commitments under international environmental accords. The transition from small-scale petrol-based generators to large-scale hydroelectric installations like Bakun symbolizes Malaysia's evolution toward energy self-sufficiency and environmental stewardship. In essence, this historical development captures the legal recognition of electricity as a national priority, the economic imperative of energy security, and the environmental rationality behind renewable adoption. It sets the contextual foundation for understanding why robust electricity demand forecasting is necessary to sustain the integrity and reliability of a complex and rapidly evolving energy ecosystem.

Building on this trajectory, Malaysia has advances in innovation and technological integration have made electricity consumption significantly sophisticated and multifaceted. The key issue here is not merely the increase in electricity usage, but the transformation in demand itself, leading to new challenges in managing energy systems effectively. This transformation is particularly evident in the domestic sector, where daily life has shifted from using energy for basic functions, to supporting digital lifestyles, and interconnected smart devices. Today, households across Malaysia are increasingly adopting smart home technologies, ranging from automated lighting, climate control, and household devices to AI-enabled security systems and real-time energy monitoring tools. These smart systems, while enhancing convenience and efficiency, also introduce substantial variability and load volatility into household electricity demand (Jiding et al., 2024).

The underlying rule in energy planning is that energy systems must be both robust and adaptable to ensure a reliable and continuous supply. As homes become digitally automated and reliant on instantaneous energy delivery, the predictability of traditional domestic load profiles is disrupted. For instance, smart appliances adjust their operation in response to user behaviour, weather conditions, or pricing signals. Thus, altering energy consumption in real time introduces new challenges in demand forecasting, and grid integration that requires a shift from static modelling to predictive, behaviour-sensitive systems. These shifts are mirrored in the industrial and urban infrastructure sectors, where Malaysia's ambition to develop smart cities, and encouraged electric vehicle (EV) adoption further amplifies the complexity. As EVs become more common, particularly in urban areas, they add significant peak loads during charging periods. Moreover, smart infrastructures such as intelligent traffic systems, urban sensors, and data centre consume electricity in non-linear and sometimes unpredictable patterns.

Ahmad & Tahar (2014) emphasize that for Malaysia to meet its future energy needs, it must strike a delicate balance between rising consumption and the development of sustainable energy sources. This principle is especially important in the context of high-tech infrastructure, where energy stability and environmental sustainability must co-exist. The reality is that electricity demand today is no longer just a matter of quantity. This evolution towards a smart, connected society has fundamentally altered the landscape of electricity demand, where stability and reliability of electricity supply have become more critical and more difficult to maintain. These developments highlighted the urgent need for advanced demand forecasting models that can interpret and respond to these emerging complexities. Thus, the

increasing sophistication of electricity consumption in Malaysia provides strong justification for adopting high-resolution, seasonally aware forecasting frameworks to support both short-term operations and long-term strategic planning.

In response to the escalating demand for a consistent and dependable electricity supply in Malaysia, precise forecasting is needed. Electricity, unlike other commodities, cannot be stored easily on a scale, meaning supply must be matched to real-time demand to prevent instability, wastage, or outages. This makes forecasting an essential tool in the management of electricity grids, aiding generation schedules, distribution logistics, and supporting informed projections that enhance balance and efficiency across the energy value chain, from production to end-user delivery. Residential, commercial, and industrial sectors are reshaping consumption patterns, introducing variability and real-time responsiveness that challenge traditional load forecasting models. Accurate demand forecasting helps utilities align generation with fluctuating consumption, as seen during the heatwave of mid-2023, when unexpected demand surges due to widespread air-conditioner use caused temporary power supply issues in parts of Penang and Klang Valley. Ishak et al. (2022) accentuated that such anomalies must be anticipated in future forecasts, especially as energy use becomes more climate sensitive. Furthermore, the anticipated 20% rise in Malaysia's electricity demand over the next decade reflects the growing pressure on energy systems to maintain balance and reliability (The Sun, 2024) and highlights the urgency of expanding and upgrading national energy infrastructure, while investing in sustainable energy sources. In short, electricity demand forecasting stands at the intersection of real-time supply management, strategic planning, and environmental responsibility, making it essential for maintaining grid stability and guiding the future of Malaysia's energy sector.

It is necessary for electricity generators, distributors, and system operators to undertake systematic forecasting studies to anticipate future electricity usage across various sectors. Additionally, accurate forecasting supports optimal generation planning, cost reduction, and mitigates the risks of blackouts or shortfalls. Moreover, well-informed forecasts support the design of effective demand-side management programs, which are critical for promoting energy conservation among consumers and reducing peak load stress on the grid. A wide range of statistical and machine learning models have been developed for this purpose. Traditional models such as the Seasonal Autoregressive Integrated Moving Average (SARIMA) are widely used for their robustness in capturing linear trends and seasonality in time-series data. Alongside SARIMA, exponential smoothing methods, including Simple, Holt's, and Brown's models, are frequently employed for their simplicity and effectiveness in short-term forecasting. With advancements in computing power, more complex methods such as Fuzzy Time Series (FTS), Least Squares Support Vector Machines (LSSVMs), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), and Artificial Neural Networks (ANN) have gained popularity due to their ability to model nonlinear patterns and adapt to highly variable consumption behaviours (Zhang et al., 1998; Pai & Hong, 2005). These models, especially when hybridized, offer flexible frameworks for addressing the complex and dynamic nature of electricity demand.

Positioned within the dynamic and evolving field of electricity demand forecasting, this study contributes to a growing body of research focused on refining predictive accuracy through advanced statistical methodologies. Understanding the place of the Box-Jenkins approach within this landscape requires a critical examination of prior literature, particularly studies that have employed time series models across diverse geographical and operational contexts. As such, this introduction synthesizes findings from key academic works relevant to demand forecasting, with an emphasis on applications within Malaysia and similar regions. By doing so, it accounts for shifting consumption patterns driven by technological adoption, climate variability, and major global disruptions such as the COVID-19 pandemic. This contextual review not only outlines established methodologies but also underscores the specific gaps that this research seeks to address, thereby affirming its relevance and potential

contribution to both theory and practice in the energy forecasting domain (Box et al., 2015; Lee et al., 2022).

Sim et al. (2019) explored the application of the SARIMA model for forecasting monthly electricity consumption at Universiti Tun Hussein Onn Malaysia (UTHM). This research uses UTHM electricity consumption data from January 2009 to December 2018, comprising 120 observations that clearly display seasonality. The SARIMA method was applied using 2 approaches, using IBM SPSS's Expert Modeler and the Box–Jenkins methodology. Following the Box-Jenkins' five step process, the researchers identified that the time series data was non-stationary, requiring both first differencing and first seasonal differencing to achieve stationarity and subsequently detected six initial SARIMA models. Through parameter estimation and diagnostic checks, they identified SARIMA$(0,1,1)(0,1,1)_{12}$ as the optimal structure and achieved a MAPE of 8.4%. The Expert Modeler also generated the same best model. Their analysis highlights that SARIMA can effectively capture both trend and seasonal components in institutional energy data; however, the monthly aggregation and university-specific dataset limit the model's responsiveness to finer fluctuations and its applicability to broader, national-scale patterns.

A study by Ping & Kamarudin (2022) applying a non-seasonal ARIMA model to forecast national electricity demand in Malaysia, it utilized univariate electricity demand data for three sectors, residential, commercial, and industrial, ranged from 1990 to 2018. Eighty percent of the data exploited for model construction and the other 20% for prediction and evaluation. The four-stage Box-Jenkins approach was applied into this study, where in identification stage Augmented Dickey-Fuller (ADF) test was employed to test for data stationarity, leading to the application of differencing, first differencing for residential, and commercial, and third differencing for the industrial sector. Then from Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots, the components of p, d, q were identified. Secondly, model parameters were estimated. Next, for model selection, Akaike Information Criterion (AIC) was applied with the lowest value that signifies the best model, which resulted in ARIMA(1,3,1) for industrial and ARIMA(0,1,0) with drift for both residential and commercial sectors. Lastly, in diagnostic checking stage, Ljung-Box Q (LBQ) test was performed to ensure residual independently distributed, thus confirming the adequacy of the selected models. The models' performance then evaluated using MAPE and RMSE. All selected ARIMA models demonstrated high forecasting accuracy as their MAPE values were below 10%. While the model effectively captured linear trends and autocorrelation, its lack of seasonal components restricted its ability to account for predictable periodic variations, such as those triggered by weather extremes or cultural events. Although the ARIMA model produced acceptable results for long-range forecasts, its limited sophistication weakens its utility for detailed policy planning or short-term operational forecasting where seasonality plays a critical role.

Zaim et al. (2025) reports the essential need for a precise electricity demand forecasting in Malaysia to optimize planning efficiency within power generation sector. This research highlights that this country's electricity demand data inherently exhibits seasonal patterns, making the evaluation of forecasting capabilities for weekly peak electricity demand crucial. Weekly forecasting provides a useful balance between the detail of daily data and the stability of monthly aggregation. Daily data often involves multiple seasonal patterns, day-of-week, annual, etc., and high volatility, complicating model estimation and reducing forecast reliability (Hyndman & Athanasopoulos, 2018; Box et al., 2015). On the other hand, monthly data smooths over intra-week variations, potentially masking operationally important patterns such as weekend vs weekday loads. Prior energy forecasting work has shown that aggregated demand tends to be more predictable, that is lower forecast errors, compared to highly granular datasets (Merza et al., 2018). Furthermore, recent models specifically developed for weekly time-series forecasting have demonstrated good accuracy and practicality (Godahewa et al., 2023). Thus, choosing weekly data in this study allows capturing meaningful load dynamics with manageable

noise, achieving better forecast reliability for short-term planning. This study finds out that SARIMA(4,1,0)$(0,1,0)_{52}$ to model weekly peak electricity demand in Malaysia from 2005 to 2016, achieving a robust MAPE of 2.95%. Their rigorously structured Box–Jenkins methodology confirmed SARIMA's effectiveness in short-term peak forecasting. Nevertheless, the study focused exclusively on peak values rather than total weekly consumption, and the data taken did not include numbers from 2017 to 2023 when choosing SARIMA model, omitting the significant behavioural shifts and remote work trends that emerged post-2020 due to COVID-19. Evidence from Malaysian studies suggests that electricity demand patterns in the post-COVID period behave differently compared to pre-pandemic norms, with higher residential loads, altered daily load profiles due to hybrid work, and more volatile grid demand (Mustapa et al., 2021; Mohamad, 2022; Januri et al., 2022). Moreover, the assumption of normality for the SARIMA models' residuals was not satisfied due to continued presence of outliers.

In addition to Malaysian literature, international studies further validate SARIMA's relevance in short-term electricity forecasting. Kien et al. (2023) concentrates on the significant interest in load forecasting globally due to its indispensable role in electricity planning, investment strategies for power systems and developing business and system operation plan within the power sector. This research applied SARIMA models to hourly electricity demand in Hanoi, selecting SARIMA$(0,1,7)(0,1,1)_{24}$ as the optimal model after rigorous ACF and PACF diagnostics and residual testing. Their study emphasized the model's robustness in capturing both daily and weekly cycles under climate variability. These international findings underscore SARIMA's strength in handling complex consumption data and reinforce its applicability to Malaysia's evolving electricity landscape. Despite all of that, the study also acknowledges that the approach of using the first difference for non-stationary series imposes certain restrictions on the predicting range. A more in-depth study of historical data fluctuations would be feasible with a longer period of real statistical data.

Sari et al. (2023) addresses the critical issue of predicting electricity loads in East Kalimantan, focuses on religious holidays. This investigation focused on two sample religious holidays, Eid al-Fitr and Christmas Day, applying historical hourly electricity data from Indonesia's State Electricity Company (PT PLN) East Kalimantan from 2015 to 2018. The SARIMA model was employed due to its ability to accommodate seasonal patterns, which were evident in the hourly data. The methodology involved several steps. Firstly, differencing was applied to the non-stationary time series data, specifically, two non-seasonal differencing ($d=2$) and one seasonal differencing ($D=1$) with seasonal period of 24. Secondly, SARIMA orders were identified using ACF and PACF plots. Then, a significance test was performed on alternative models to identify valid candidates. Next was diagnostic checking conducted on the residuals to ensure model adequacy, LBQ test and Kolmogorov-Smirnov test were used. Lastly, the best model for Eid al-Fitr is SARIMA$(0,2,1)(0,1,0)_{24}$ with MAPE of 16.99% and for Christmas Day best model is SARIMA$(0,2,1)(0,1,0)_{24}$ with MAPE of 8.68%. MAPE values indicate adequate to high accuracy for forecasting. However, this study is only limited to specific two holidays, while SARIMA model proved effective for these significant events, but its accuracy for other holidays in East Kalimantan would require further investigation. Additionally, there was a noticeable difference between MAPE value for both holidays.

Andoh et al. (2021) investigates how SARIMA model could be applied to forecast electricity demand in the Western Regions of Ghana. The secondary data specifically uses monthly consumption data for three consumer categories, namely Special Load Tariff, Non-Special Load Tariff, and Prepaid Tariff spreads from January 2008 to December 2013. The Box-Jenkins methodology was adopted, involving the estimation of non-seasonal and seasonal parameters, with monthly seasonality. The ACF and PACF were utilized to confirm parameter acceptability, and models were validated against historical data from early 2014 to end of 2020. Using the F-criterion for model selection, aiming for the minimum sum of squared estimate of error (SSE). The study successfully formulated models considered viable for the future consumption forecasts. The predicted value closely aligned with the actual data with an

error range of -6.00% to 6.00%. The developed models were deemed reliable with a 95% confidence level, useful for estimating future electricity usage and revenue projections.

The uniqueness of this study lies in its use of total weekly electricity consumption data from 2022 to 2024, which enables the capture of both short-term fluctuations and longer seasonal cycles in Malaysia's energy demand. Weekly data strikes an optimal balance, avoiding the noise inherent in daily data while retaining more actionable granularity than monthly data, making it particularly well-suited for operational planning and real-time decision-making. Compared to monthly data, that smoothen over critical short-term variations relevant for operational planning (Hyndman & Athanasopoulos, 2018; Almeida-Dias et al., 2010). As highlighted by Box et al. (2015), weekly aggregation allows forecasters to detect trends without excessive volatility while still preserving the seasonality and structural shifts needed for robust modelling. This temporal resolution is especially valuable in a post-COVID context, where electricity demand is increasingly influenced by dynamic factors such as work-from-home arrangements, irregular holiday patterns, and climate-related events create non-stationary but recurring weekly patterns.

The COVID-19 pandemic fundamentally altered global electricity demand patterns, including in Malaysia. Electricity usage declined sharply during movement control orders and lockdowns due to the closure of commercial and industrial sectors, while residential use increased due to remote work. A study by Energy Agency (2020) noted a 5% global drop in electricity demand in 2020, with Malaysia experiencing notable shifts in load profiles. Mohd Zain (2021) also recorded reductions in demand at Kuantan Community College, highlighting the long-term impact of COVID on consumption behaviour. Therefore, including 2022–2024 data allows this study to model electricity demand under new normal conditions, incorporating digital transformation, hybrid work culture, and post-pandemic industrial recovery. Existing models trained on pre-2020 data risk underfitting or misrepresenting new behavioural trends.

The SARIMA model has consistently demonstrated adaptability across applications involving seasonally varying, autocorrelated time series such as electricity load, economic indicators, and weather-dependent data. It was selected for this study due to its balance between accuracy, interpretability, and modest data requirements. Unlike machine learning methods such as LSTM or ANN, which require large training datasets and often lack transparency, SARIMA provides a statistically rigorous framework for capturing linear trend and seasonality with explainable parameters (Box et al., 2015; Hyndman & Athanasopoulos, 2018). Moreover, as a benchmark model in the energy forecasting literature, SARIMA facilitates transparent comparison with prior Malaysian studies, making it a logical starting point for assessing post-COVID electricity demand behaviour.

The SARIMA model has consistently demonstrated adaptability across applications involving seasonally varying and autocorrelated time series whether in macroeconomic indicators, hospital admissions, or electricity load. It was selected due to its balance between accuracy, interpretability, and modest data requirements. Unlike machine learning methods that require large training datasets and often lack transparency, SARIMA provides statistically rigorous framework for capturing linear trend and seasonality with explainable parameters (Box et al., 2015; Hyndman & Athanasopoulos, 2018). This makes it suitable for policymakers and utility managers seeking not only accuracy but also clarity in decision-making frameworks. The structured seasonal modelling SARIMA offers makes it the optimal candidate for modelling Malaysia’s weekly electricity consumption.

As Malaysia continues its journey toward smart living, electric mobility, and sustainable urbanization, accurate and high-frequency demand forecasting will be central to energy policy, infrastructure planning, and environmental stewardship. Integrating real-world disruptions like COVID-19, seasonality from cultural and climate cycles, and digital behavioural shifts is essential for modern forecasting. In this study, post-COVID effects are displayed through ongoing hybrid work trends, changes in residential consumption, and the stabilization of industrial recovery, all observable in the

2022-2024 weekly demand patterns. Specifically, this research aims to address these keys questions: (1) How SARIMA model can accurately forecast total weekly electricity demand in Malaysia within the context of post-COVID? (2) Why does weekly data provide a more balanced and operationally relevant forecasting resolution compared to daily or monthly data? and (3) What is the limit for SARIMA framework remain effective compared to existing benchmark studies in Malaysia's electricity demand forecasting? This effort filled in any gaps in existing literature while contributing to more reliable, data-driven electricity planning nationwide.

2. Methodology

2.1 Box-Jenkins: SARIMA Model

As said by Zaim et al. (2025), the electrical demand pattern exhibits evident periodic fluctuations attributed to seasonal changes. This can be addressed by using the SARIMA model recommended by Box-Jenkins, where the model accommodates seasonality by integrating seasonal autoregressive (SAR), seasonal differencing (SD), and seasonal moving average (SMA) components into the ARIMA structure. A SARIMA model is commonly expressed in the notation $SARIMA(p,d,q)(P,D,Q)_S$, where the mathematical model is given by Equation 1,

$$\Phi_P(B^s)\phi_p(B)\nabla^d\nabla_s^D y_t = \Theta_Q(B^s)\theta_q(B)\varepsilon_t \quad (1)$$

In this expression, y_t denotes the value of the time series at time, t, while B represents the backshift operator such that $By_t = y_t - 1$. The non-seasonal differencing of order d is given by $\nabla^d = (1-B)^d$, and the seasonal differencing of order D with seasonality length s is represented by $\nabla_s^D = (1-B^s)^D$. The non-seasonal autoregressive and moving average components are described by the polynomials $\phi_p(B)$ and $\theta_q(B)$, respectively, while the seasonal autoregressive and moving average components are represented by $\Phi_P(B^s)$ and $\Theta_Q(B^s)$. The error term ε_t is assumed to be white noise, normally distributed with zero mean and constant variance.

2.2 Framework for Box-Jenkins Model Approach in Forecasting Electricity Demand

Figure 1 shows the methodological framework for applying the SARIMA model to forecast Malaysia's total weekly electricity demand data. The structured flow ensures systematic application of the SARIMA model after seasonal patterns are observed in the time series data. The four stages – identification, estimation, diagnostic checking and forecasting, are important to attain an accurate forecast. This framework was inspired by Zaim et al. (2025), with modification suited to this study, where they will proceed either with SARIMA model or SARIMA-GARCH model after checking for the heteroscedasticity. This is due to the inherently volatile nature of weekly peak demand, while this study focuses on total weekly electricity consumption. Total demand aggregates variations across residential, commercial, and industrial sectors, resulting in a smoother series with lower volatility. Since the SARIMA model primarily captures mean dynamics rather than variance behaviour, tests for heteroscedasticity are not strictly required as long as the residuals meet white noise assumptions (Box et al., 2015; Hyndman & Athanasopoulos, 2018) Diagnostic checks in this study, including the Ljung–Box test and normality evaluation, confirm that residuals are independently and randomly distributed, satisfying SARIMA's adequacy criteria for mean forecasting.

As per Box et al. (2015), the usual Box-Jenkins framework comprises four repetitive stages: Stage I – Model Identification, Stage 2 – Parameter Estimation, Stage 3 – Diagnostic Checking, and Stage 4 – Forecasting. These stages will be the basis of the proposed framework in Figure 1.

Stage 1: Model Identification

After obtaining and cleaning the total weekly electricity demand data from Grid System Operator (GSO). The data then plotted using Python programming in Jupyter Notebook, which obviously display seasonality. To confirm it further, a data decomposition was plotted which illustrate the trends, the seasonality and its residuals.

Following recognizing seasonality in the weekly electricity demand series, several statistical tests can be conducted to determine the SARIMA Model. Below are the tests administered to construct the SARIMA model:

(1) Stationarity Evaluation: Visually inspect the data plot to determine stationary in-variance. If data spreads or diminishes over time, transform the data by using Box-Cox transformation to stabilize the variance. Tests like ADF and KPSS are used to check if the data is stationary in-mean. If the data is not stationary, apply non-seasonal differencing to stabilize the data.

(2) ACF and PACF of the Stationary Data for Non-Seasonal portions: By plotting the ACF and PACF, it can be used to identify Autoregressive (AR) order value, p, and Moving Average (MA) order value, q.

(3) Seasonality allocation: Determine the seasonal time allotment by observing the seasonality of the data plot.

(4) ACF and PACF of the Stationary Data for Seasonality portion: By plotting the ACF and PACF for this part, it can be used to detect the Seasonal AR (SAR) order value, P and Seasonal MA (SMA) order value, Q.

Repetition of the steps might be needed to recognize the parameters for the SARIMA model that was later used to determine the best model of the data.

Stage 2: Parameter Estimation

Building on Stage 1, the next step involved stationarity and seasonality assessment. The time series was first analyzed to measure its stationarity and the presence of seasonal patterns. This involved applying unit root tests and visually examining seasonal structures through decomposition and ACF/PACF plots.

Subsequently, the non-seasonal (p, d, q) and seasonal (P, D, Q) orders were identified by analyzing the transformed series. Significant spikes observed in the ACF and PACF plots at both non-seasonal and seasonal lags guided the initial specification of AR and MA components.

(1) Stationarity and Seasonality Assessment: The time series was first analyzed to measure its stationarity and the presence of seasonal patterns. This involved applying unit root tests and visually examining seasonal structures through decomposition and ACF/PACF plots.

(2) Identification of Model Orders: The non-seasonal (p, d, q) and seasonal (P, D, Q) orders were identified by analyzing the different series. Significant spikes observed in the ACF and PACF plots at both non-seasonal and seasonal lags guided the initial specification of AR and MA components.

Afterwards is the model fitting and optimization, where the selected orders were used to initialize the SARIMA model using the *auto_arima* function from the *pmdarima* Python library. The model parameters (AR, MA, and seasonal coefficients) were estimated using Maximum Likelihood Estimation (MLE), which selects parameter values that maximize the probability of observing the given dataset under the specified SARIMA structure. MLE is widely used in time series forecasting because it produces consistent and asymptotically efficient parameter estimates (Hyndman & Athanasopoulos, 2018), to refine parameter estimates and improve model accuracy. It automatically does a stepwise search through various models to find the optimal model that has the lowest AIC or Bayesian Information Criterion (BIC), where the model with the lowest value was selected as the most suitable candidate.

Stage 3: Diagnostic Checking

Diagnostic Checking, a vital step in the modelling process, where such checks were designed to test any suspected departures from the fit, and to show up any other serious discrepancies (Young, 1977). These tools are essential to ensure that the SARIMA model not only fits the historical data well but also provides accurate and unbiased forecasts. As noted by Box et al. (2015), model adequacy checking is a critical component of the Box–Jenkins methodology to avoid the risk of overfitting or misrepresenting the data's true structure

It involves analysing the model residuals, ensuring they behave like white noise, namely, zero mean, and absence of autocorrelation. Residual plots are examined to detect any systematic patterns or significant correlations that may indicate model misspecification. Additionally, formal statistical tests such as the LBQ test are applied to determine whether the residuals are independently distributed. A non-significant result from the test suggests that the model has adequately captured the underlying structure of the time series. Should the residuals fail to meet the assumptions of independence, normality, and homoscedasticity, model refinement may be undertaken through re-specifying SARIMA orders (p,d,q)(P,D,Q), applying additional differencing, adjusting seasonal lags, or incorporating intervention analysis to address outliers and structural breaks (Box et al., 2015; Wei, 2006).

Stage 4: Forecasting

Forecasting requires generating future electricity demand values based on the previously estimated SARIMA model. After the model has been validated through rigorous diagnostic checks, it is applied to produce one-step-ahead forecasts, in which each subsequent value is predicted using all available historical information up to the current point in time. This forecasting prospect is particularly advantageous for short-term operational planning, as it enables utility providers to respond promptly to fluctuations in electricity demand and allocate resources more efficiently.

The accuracy of the forecasting model is quantitatively assessed using three commonly applied statistical error metrics. To evaluate the predictive performance of the model, standard error metrics are employed, namely the Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE) as defined in Equations 2, 3 and 4, respectively.

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y_t}| \tag{2}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \hat{y_t})^2}{n}} \tag{3}$$

$$MAPE = \frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y_t}}{y_t}\right| \tag{4}$$

MAE measures the average magnitude of forecast errors in absolute terms, providing a straightforward interpretation of model performance. RMSE penalizes larger errors more heavily than MAE due to the squaring of residuals, making it useful when large deviations are particularly undesirable. MAPE is widely used due to its scale-independent nature and ease of interpretation, especially in contexts where communicating forecast performance to non-technical audiences is necessary.

According to Vivas et al. (2020), a MAPE value below 10% is generally considered indicative of highly accurate predictions, while values between 10–20% reflect good forecasting performance. Out of the commonly used error metrics, MAPE is often favored for its interpretability, as it expresses forecast errors as a percentage of actual values. This characteristic makes it especially suitable for

communicating model performance to stakeholders who may not have a technical background. A lower MAPE value generally reflects higher forecast accuracy, and it is widely regarded as a benchmark indicator in time series forecasting applications, particularly in the energy sector.

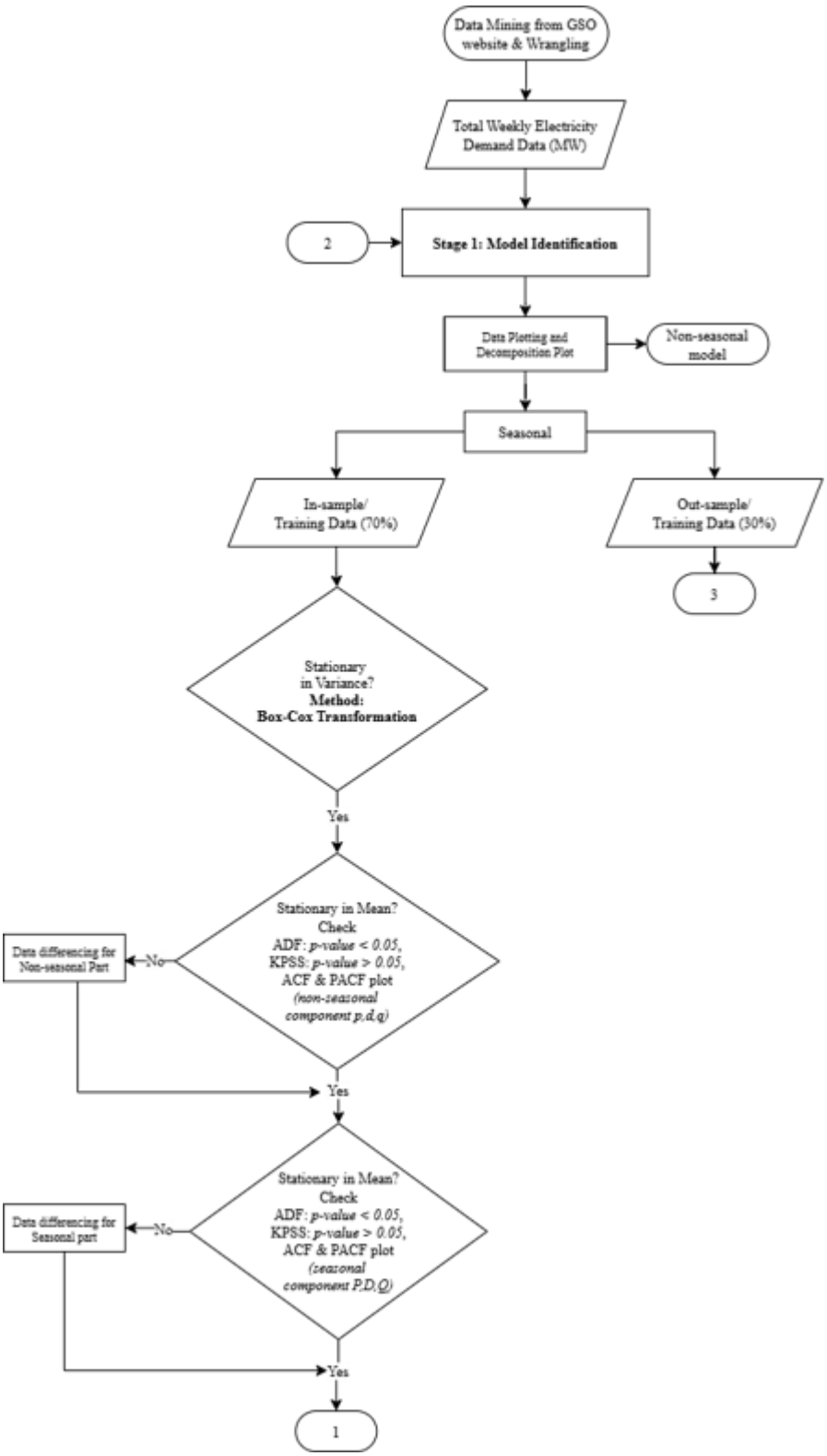

Figure 1: Framework for Box-Jenkins SARIMA Model Approach in Forecasting Seasonal Electricity Demand

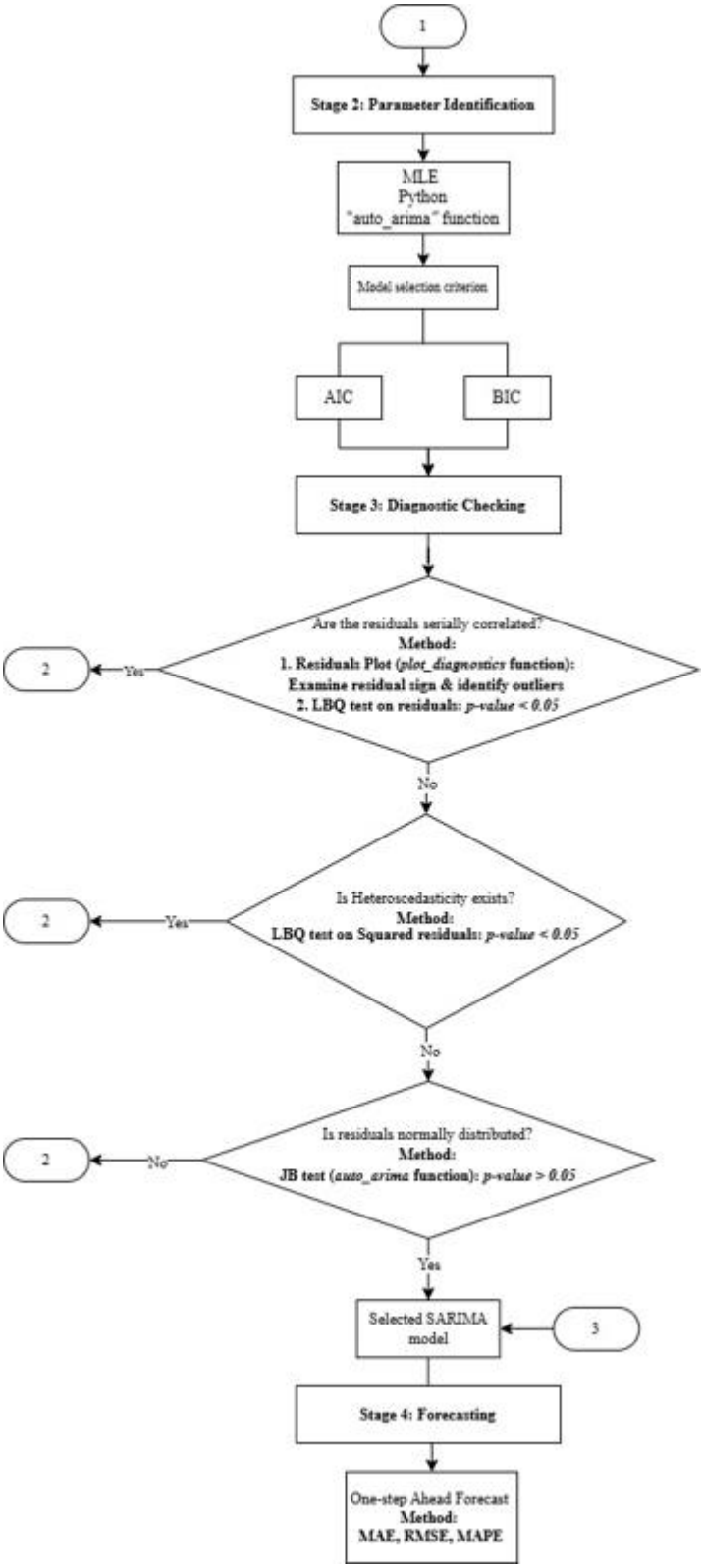

Figure 1: Framework for Box-Jenkins SARIMA Model Approach in Forecasting Seasonal Electricity Demand (Continued)

3. Results and Discussion

3.1 Dataset

The total weekly electricity demand data used in this study, measured in megawatts (MW), spans from 2022 to 2024. Data was obtained from GSO, an authorised organisation under Tenaga Nasional Berhad (TNB), responsible for the secure, reliable and real-time operation of the national electricity power grid Categorized as secondary data, it was extracted from the official GSO website, which regularly publishes grid operation statistics and national electricity demand records. This dataset contains 10-minute interval observations of daily peak electricity demand observations across Peninsular Malaysia, reflecting consumption patterns from both residential and industrial.

The raw dataset, presented in 10-minute intervals from 1st January 2022 to 31st December 2024, consists of 157,726 observations. The dataset was then aggregated to weekly frequency using the *.resample('W').sum()* function in Python, which groups the datetime index and computes the total weekly electricity demand. Table 1 presents the data split for the weekly total electricity demand with a total of 156 data points collected over a three-year period. The dataset was first explored through time series plotting and decomposition plot, as in Figure 3, to uncover underlying patterns such as trend and seasonality.

Following this, the data was partitioned using 70% training and 30% testing sets, providing a good basis for both model training and performance evaluation. Although a 67:33 ratio would preserve exactly 52 weeks (1 year) of testing data, aligning with best practices for matching forecast horizons, the 70:30 ratio was selected to prioritize greater training stability and improve model learning, as recommended in time series literature, where more data often leads to better parameter estimation (Chatfield, 2001; Hyndman & Athanasopoulos, 2018). As observed in Figure 3, the seasonal cycle repeats every 52 weeks, and this cycle length was applied in the SARIMA modelling framework.

Table 1: Total Weekly Electrical Demand Data Splitting

Duration	Number of Data	In-Sample Data (70%)	Out-Sample Data (30%)
2022 – 2024	156	1-109	110-156

3.2 Modelling and Forecasting using SARIMA Model

Based on Figure 1, which illustrates the SARIMA framework, the modelling and forecasting of electricity demand are conducted accordingly. Figure 2 illustrates the weekly fluctuations in electricity demand in Malaysia based on in-sample (training) data. The plot exhibits distinct seasonal patterns, with demand generally peaking during warmer months and declining toward the end of each year and the beginning of the next. The series shows an overall upward trend, with notable rises in peak demand occurring in early 2024. The consistent presence of high-frequency variations highlights the dynamic and complex nature of electricity consumption, which is likely shaped by a combination of meteorological condition, economic activity, and behavioural factors. These observed patterns support the need for a seasonal time series model such as SARIMA, which capable of capturing both repetitive seasonal cycles and underlying trends for accurate forecasting.

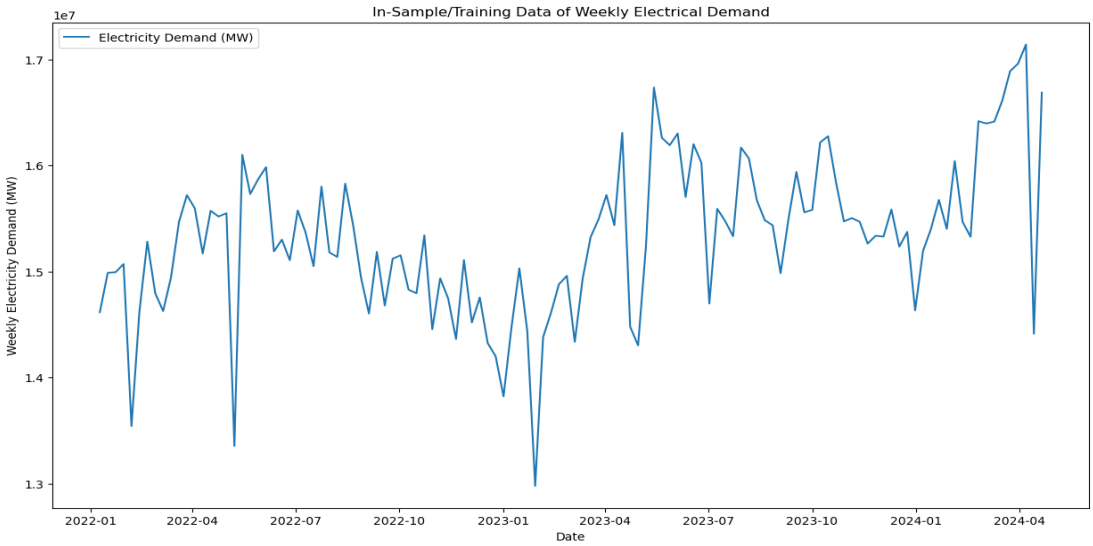

Figure 2: In-Sample Data of Weekly Total Electricity Demand from 2022 to 2024

Figure 3 displays the time series decomposition of total weekly electricity demand from 2022 to 2024 in Malaysia. The decomposition of the weekly electrical demand time series, breaking it down into its constituent components: observed data, trend, seasonal, and residual. The first subplot displays the raw weekly demand in MW, consistent with the previous visual. The Trend subplot clearly indicates a general upward trend in electricity demand over the observed period, suggesting an underlying growth in consumption.

The Seasonal subplot reveals a distinct cyclical pattern, confirming the presence of recurring weekly or yearly variations in demand. Finally, the Resid subplot, representing the residuals, shows the remaining irregular fluctuations after the trend and seasonal components have been removed. These residuals appear to be largely random and centred around zero, indicating that the decomposition has effectively captured the systematic patterns in the data, with no obvious remaining structure in the noise. The relatively consistent spread of residuals also suggests that the assumption of constant variance may be reasonable for this dataset.

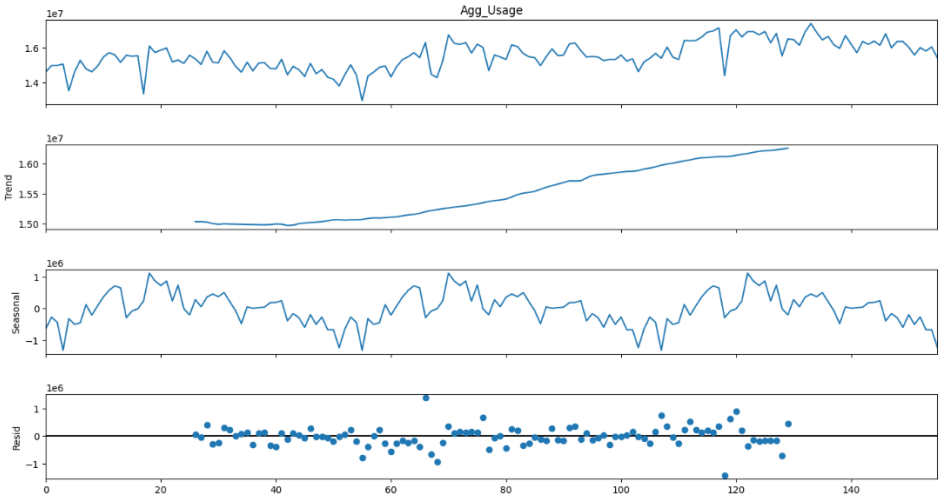

Figure 3: Decomposed Data of Total Weekly Maximum Electricity Demand from 2022 to 2024

Prior to modelling, the training data was evaluated for variance stationarity using the Box–Cox transformation. The lambda (λ) parameter for the Box–Cox transformation was optimized within the

range $-1 < \lambda < 1$, following standard practice to maintain interpretability of results and avoid numerical instability or distortion of the series. This restriction ensures that transformed values remain approximately linear and comparable to the original scale (Box et al., 2015; Hyndman & Athanasopoulos, 2018). The Box–Cox transformation was applied to evaluate whether variance stabilization was needed before model identification. The resulting $\lambda = 1.00$ indicated that no transformation was necessary, implying that the variance of the series was already stable with time. Consequently, the raw data were retained for subsequent model identification, satisfying the assumption of homoscedasticity required for Box–Jenkins estimation (Osborne, 2010; Chatfield, 2001).

Next step was to evaluate stationarity in-mean. At first glance, the plotted weekly electricity demand data displayed a slight upward trend, suggesting that the series may not be stationary in-mean. To validate this observation, ADF and KPSS test were performed. This study adopts a dual-testing method to ensure the robustness from complementary perspective, as these tests evaluate stationarity from opposite null hypotheses. ADF tests for the presence of unit root (non-stationarity), while KPSS tests for stationarity of data around a deterministic trend (Kwiatkowski et al., 1992; Brooks, 2008). Using both reduces the risk of misclassification and improves diagnostic accuracy (Shrestha & Bhatta, 2018), and strengthens the choice of d in ARIMA/SARIMA modelling.

Stationarity was tested at each differencing stage using both ADF and KPSS tests, as shown in Table 2. The initial series and both first-order non-seasonal and seasonal differencing steps failed to satisfy the conditions for stationarity at the 5% significance level. Only after applying double differencing ($d = 1, D = 1$) did both tests shows stationarity, with ADF $p-$value < 0.01 and KPSS $p-$value > 0.1. Confirming the series was suitable for modelling. This outcome is further illustrated in Figure 5.

Figure 4 illustrates the results of first-order non-seasonal differencing. The transformation removed the trend, and series fluctuates around a constant mean. However, as reflected in Table 2, ADF and KPSS tests implies the series remained non-stationary, with both $p-$values failing the 5% significance level. The ACF plot exhibits a gradual decay and the PACF has a prominent spike at lag 1, suggesting potential autoregressive behaviour. As further differencing in the non-seasonal component could risk over-differencing, the analysis proceeded with seasonal differencing ($D = 1$) to account for weekly seasonal structure as recommended by Hyndman & Athanasopoulos (2018). The plot for seasonal differencing can be viewed in Figure 5.

Table 2: Stationarity Test using ADF-KPSS Test

Differencing Step	**ADF $p-$value**	**KPSS $p-$value**	**Stationarity**
Before differencing	0.1843	0.0484	Not stationary
Non-seasonal	0.1642	0.04167	Not stationary
Seasonal	0.8304	< 0.01	Not Stationary
Double Differencing	0.0093	> 0.1	Stationary

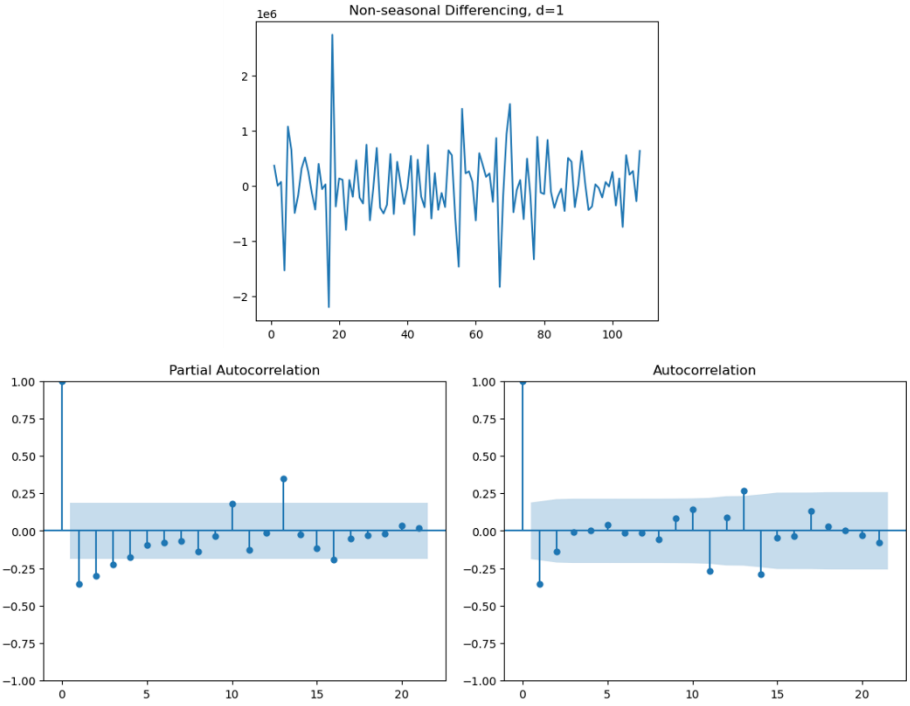

Figure 4: Non-seasonal Differenced of Total Weekly Electrical Demand Data

Figure 5 represents a seasonally differenced ($D = 1$) time series of electricity demand series introduced greater fluctuations but did not remove the upward trend, particularly in the later part of the series. Although the ACF and PACF plots were not reproduced for the seasonally differenced series in Figure 5, these were examined during the model identification stage to ensure adequacy of seasonal differencing. As Box et al. (2015) describe, seasonal differencing is a standard transformation to achieve stationarity, and its adequacy can be confirmed through unit root tests and subsequent residual diagnostics rather than by re-plotting ACF/PACF at each differencing stage (Wei, 2006; Hyndman & Athanasopoulos, 2018), the stationarity test results in Table 2, where both tests failed to meet the 5% significance threshold. To achieve full stationary, double differencing ($d = 1, D = 1$) was applied. The resulting series is displayed in Figure 6, where the plot exhibits fluctuations around a constant mean with no apparent trend or recurring seasonal pattern. This suggests that the series is now stationary. The ADF test yielded a $p-$value of 0.0093, while the KPSS test returned a $p-$value larger than 0.1, surpassing the lookup table range due to computational limitation libraries like *statsmodels* in Python, confirming the absence of unit roots and presence of stationarity.

The PACF plot in Figure 6 presents low autocorrelation values with a few significant spikes, while the ACF plot presents a sharp drop after lag 1. This pattern indicates a short-memory process and supports the inclusion of low-order AR and MA terms. These diagnostics inform the selection of appropriate model orders (p, q, P, Q) in the subsequent SARIMA modelling process.

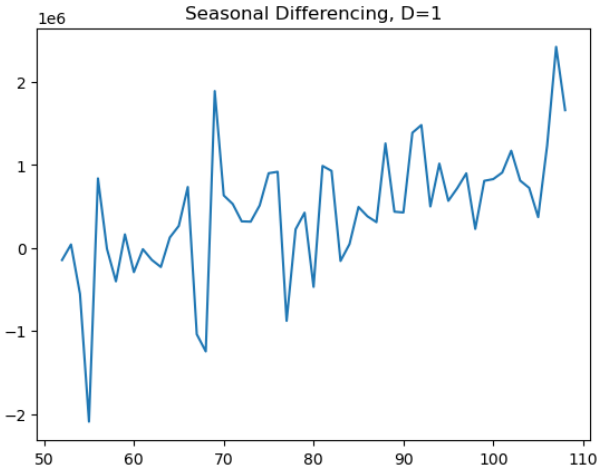

Figure 5: Seasonal Differenced of Total Weekly Electrical Demand Data

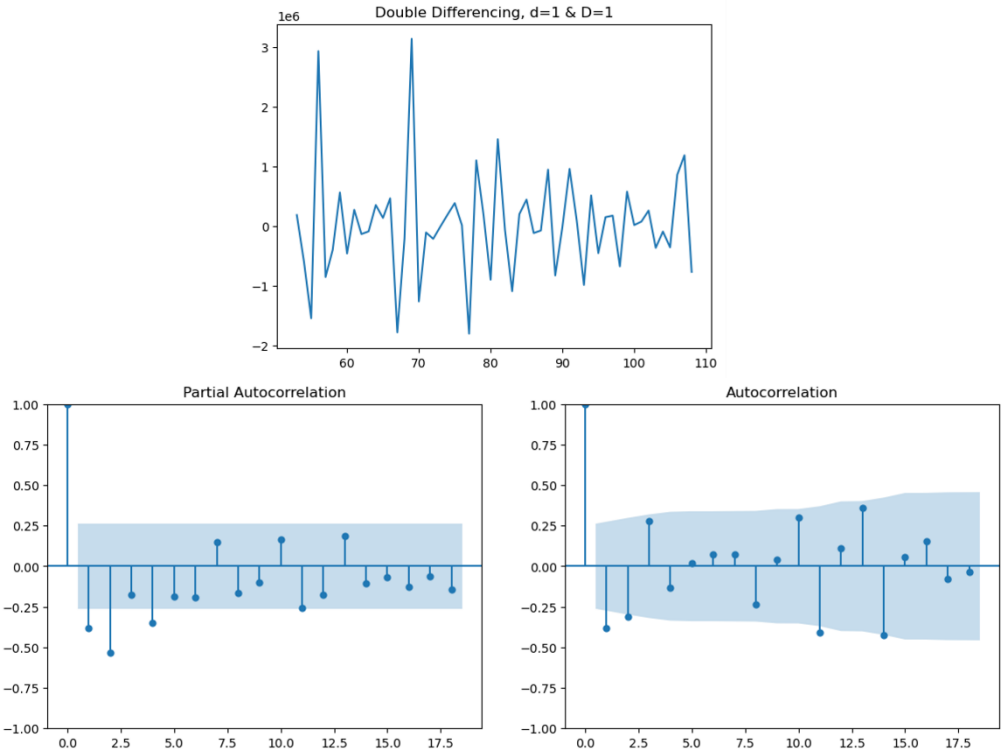

Figure 6: ACF and PACF for Double Differenced

The *auto_arima* function from *pmdarima* library in Python was used to identify the best-fitting ARIMA model for the data. The function makes a stepwise search based on the specific range of SARIMA parameters where the best model will be selected based on the smallest value of AIC.

According to Zaim et al. (2025), the model with the lowest AIC or BIC that still satisfactorily captures the time series, and volatility components should be selected. Then proceed with the LBQ test applied to assess the presence of serial correlation in the residuals. The null hypothesis endorses that the residuals are not serially correlated, with the results of $p-$value equals to 0.8794, exceeds the 5% significance level. Thus, proves that the residuals are not serially correlated and the model satisfies the white noise assumption. The absence of the serially correlation in residuals further justifies the application of Box-Jenkins model to this time series data, where the model satisfies one of the essential assumptions of adequacy in Box-Jenkins modelling.

Using the *auto_arima* function, which performs a stepwise search based on the minimization of AIC and BIC, 16 models were initially identified. Of these, seven were excluded because their AIC values were infinite, and one was removed as it contained an additional parameter (the intercept). The remaining eight models were then compared to determine the optimal specification. Between the two criteria, AIC served as the primary selection metric, while BIC was referenced for confirmation of model parsimony, following recommendations by Burnham & Anderson (2003) and Hyndman & Athanasopoulos (2018). In line with the principle of parsimony, when models demonstrate similar explanatory power, the simpler model is generally preferred.

The diagnostic summary of each model presents in Table 3, which reports the diagnostic checks based on serial correlation, the mean of residuals, and residual normality. Four model – SARIMA$(0,1,0)(0,1,1)_{52,}$ SARIMA$(1,1,0)(0,1,0)_{52,}$ SARIMA$(0,1,0)(0,1,0)_{52}$, and SARIMA$(0,1,0)(1,1,0)_{52}$ showed no serial correlation up to lag 20. In contrast, the remaining models displayed significant serial correlated at lags 1 through 5, indicating model inadequacy and suggesting a failure to capture all relevant information within the time series structure.

A model with residuals centred around zero is desirable to ensure unbiased forecasting. Among the tested model, SARIMA$(0,1,0)(0,1,1)_{52}$ recorded a mean residual of -5969.98, which was of lower magnitude compared the other two models. The relatively high mean residual observed in the remaining models indicate substantial deviations from the actual values, thereby reflecting poor model fit.

The Jarque-Bera (JB) test was employed to assess the normality of residuals. Across all models, the test produced statistically significant values, $p-$value = 0.00, leading to rejection of the null hypothesis of normality. For instance, the JB statistic for SARIMA$(0,1,0)(0,1,1)_{52}$ was 17.06. The rejection of normality may be due to the presence of outliers in the data, a common issue in real-world time series. Potential outliers were visually identified from residual and ACF plots, where several large residuals corresponded to atypical demand spikes during festive weeks and post-COVID recovery periods. These observations reflect genuine consumption shocks rather than measurement errors, consistent with similar findings in energy demand studies (Ghahramani et al., 2020); Suganthi & Samuel, 2012). As such, they were retained to preserve the integrity of the dataset and capture real operational volatility.

Formal robustness re-estimation excluding these points was not conducted, as the purpose of this study is to evaluate the SARIMA model's practical performance under realistic load conditions. This approach aligns with recent electricity forecasting literature that recommends retaining structurally valid outliers to ensure generalizable and operationally meaningful model evaluation (Tascikaraoglu & Uzunoglu, 2014; Nowotarski & Weron, 2018). The SARIMA$(0,1,0)(0,1,1)_{52}$ model thus represents forecasting accuracy under actual variability rather than under idealized, filtered data.

In summary, among the evaluated SARIMA models, only three models, SARIMA$(0,1,0)(0,1,1)_{52}$, SARIMA$(1,1,0)(0,1,0)_{52}$, and SARIMA$(0,1,0)(1,1,0)_{52}$, fulfilled the essential diagnostic conditions of white noise residuals and minimal forecast bias. However, none of the models met the assumption of normally distributed residuals. Despite this, the outcome of the LBQ-test justifies the consideration of the Box-Jenkins model, particularly in the case of SARIMA$(0,1,0)(0,1,1)_{52}$, where the model structure and residual behaviour align with the underlying assumptions of the Box-Jenkins methodology. This reinforces the model's suitability for application in forecasting weekly electricity demand.

Table 3: Diagnostic Checking on Identified SARIMA Models

SARIMA Model	Serially Correlated	Mean Residual	Normality test
SARIMA(0,1,0)(0,1,1)$_{52}$	Not Serially Correlated (up to Lag 20)	-5969.979	JB = 17.06 $p-$value = 0.00
SARIMA(1,1,0)(0,1,1)$_{52}$	Serially Correlated at Lag(s): 1, 2, 3, 4, 5	-71810.0206	JB = 17.78 $p-$value = 0.00
SARIMA(1,1,0)(1,1,0)$_{52}$	Serially Correlated at Lag(s): 1, 2, 3, 4, 5	73113.1179	JB = 20.91 $p-$value = 0.00
SARIMA(1,1,0)(0,1,0)$_{52}$	Not Serially Correlated (up to Lag 20)	51290.7029	JB = 28.45 $p-$value = 0.00
SARIMA(0,1,1)(0,1,1)$_{52}$	Serially Correlated at Lag(s): 1, 2, 3, 4, 5	-199928.3506	JB = 17.18 $p-$value = 0.00
SARIMA(1,1,1)(0,1,1)$_{52}$	Serially Correlated at Lag(s): 1, 2, 3, 4, 5	-175603.7082	JB = 16.83 $p-$value = 0.00
SARIMA(0,1,0)(0,1,0)$_{52}$	Not Serially Correlated (up to Lag 20)	59933.6024	JB = 27.82 $p-$value = 0.00
SARIMA(0,1,0)(1,1,0)$_{52}$	Not Serially Correlated (up to Lag 20)	44786.1629	JB = 20.43 $p-$value = 0.00

From Table 4, SARIMA(0,1,0)(0,1,1)52, SARIMA(1,1,0)(0,1,0)$_{52}$, and SARIMA(0,1,0)(1,1,0)$_{52}$ are named Model 1, Model 2 and Model 3 respectively. It presents the parameter estimates and model evaluation metrics for the selected SARIMA model. The models differ in their inclusion of non-seasonal and seasonal autoregressive or moving average components. Parameters were estimated using the maximum likelihood method, and statistical significance is indicated by $p-$values in parentheses.

Model 1 and Model 3 include only seasonal terms, while Model 2 includes both non-seasonal and seasonal components. The seasonal terms, Φ_1, in all three models are statistically significant at the 1% level, with $p-$values equal to 0.0000. Model 2's non-seasonal coefficient, φ_1, has a $p-$value of 0.079, indicating marginal significance at the 10% level, meaning the null hypothesis of no effect is rejected at a 90% confidence level. While the 5% threshold is more conventional, the 10% level is often considered acceptable in exploratory time series modelling when sample sizes are limited or when the inclusion of a parameter improves overall model fit and forecast accuracy (Gujarati & Porter, 2009; Hyndman & Athanasopoulos, 2018).

No constant term was estimated in any of the models. This is expected due to the application of both non-seasonal and seasonal differencing ($d = 1, D = 1$). In such cases, the constant component becomes unnecessary and is typically excluded from the model specification.

Model performance was assessed using AIC, BIC, and log-likelihood values. These criteria provide a basis for comparing model fitness and parsimony. As shown, all models produce distinct AIC and BIC values, which are considered in selecting the final forecasting model.

Table 4: Result of Possible SARIMA Model

Parameter	Model 1	Model 2	Model 3
C	-	-	-
φ_1	-	-0.0821 *(0.079)	-
Θ_1	- 0.1867 *(0.000)	-	-0.1153 *(0.000)
AIC	1692.613	1696.020	1747.897
BIC	1696.663	1700.071	1751.948
Log^{-l}	-844.306	-846.010	-871.949

*value in parentheses denote as $p-$value and Log-l abbreviation is log likelihood.

After fitting the best SARIMA models, diagnostic checks were performed to assess whether the residuals satisfied the assumptions of a white noise process. This is crucial for the validity and reliability of the chosen model. Figure 7 shows the diagnostic plot of the residual of the fitted model. The plots in Figure 7 consist of standardized residuals, histogram, normal Q-Q plot and correlogram plot.

The plot of standardized residuals over time shows no obvious patterns, trends, or changes in variance. The residuals appear to be randomly scattered around zero, suggesting the model has effectively captured the systematic variations in the original time series. The histogram, overlaid with a KDE curve, shows an approximately symmetric and bell-shaped distribution centred near zero, suggesting that the residuals are reasonably normally distributed. However, despite this approximate symmetry, the strict assumption of normality was not met, as previously noted due to the presence of outliers. The Q-Q plot further supports this, as most points lie close to the 45-degree reference line, with only minor deviations observed at the tails.

The ACF plot shows that all autocorrelation spikes fall within the 95% confidence interval, indicating the absence of significant autocorrelation across lags. This observation is consistent with the LBQ-test result, which yields a $p-$ value greater than 0.05. Hence, the null hypothesis of no serial correlation cannot be rejected, confirming that the residuals behave as white noise.

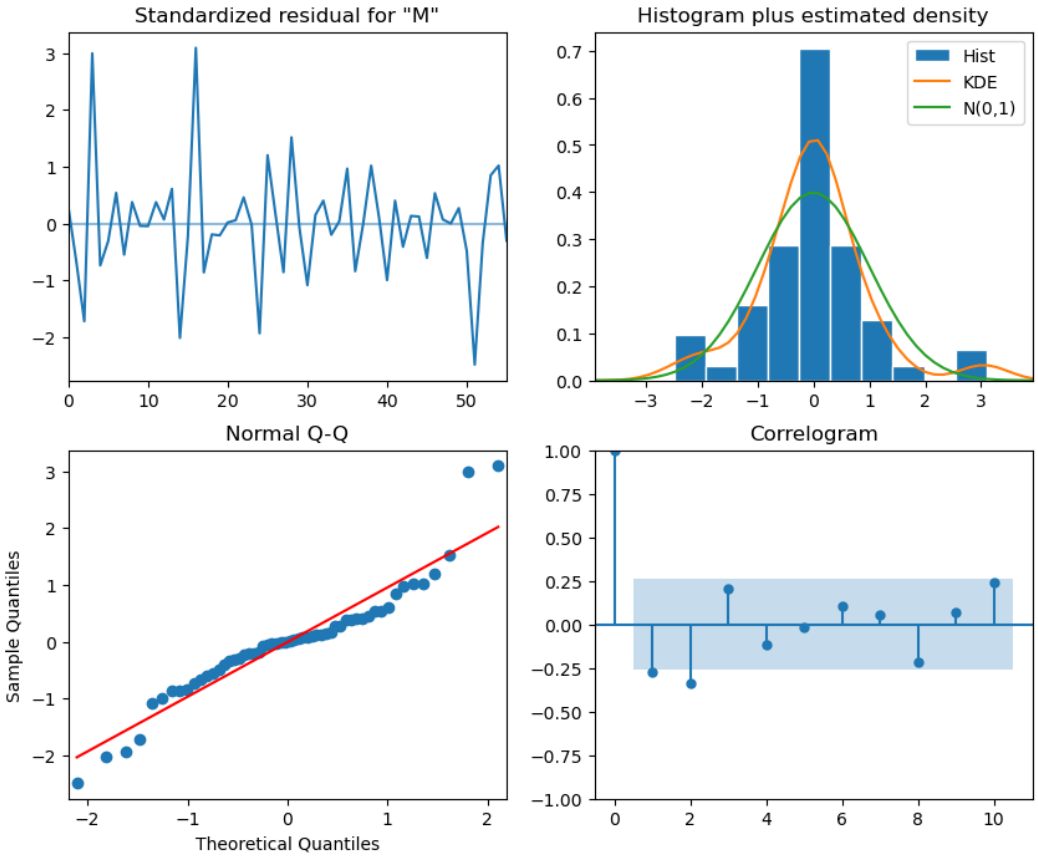

Figure 7: Diagnostic Plots of Residual of a Fitted Model

Table 5 presents the forecasting accuracy of the three candidate SARIMA models identified in Table 4, evaluated using a 30% holdout test dataset. Among them, the SARIMA(0,1,0)(0,1,1)$_{52}$ model achieved the lowest RMSE (1073867.58), MAE (894782.67), and MAPE (5.57%) values, indicating superior overall performance. For comparison, the alternative models exhibited higher MAPE values exceeding 6%, with wider confidence intervals around forecasted means, suggesting reduced stability

in predictions. Based on these results, SARIMA$(0,1,0)(0,1,1)_{52}$ was selected as the most statistically and operationally reliable model for subsequent analysis.

The RMSE reflects the square root of the average squared differences between predicted and actual electricity demand values. As RMSE is sensitive to large errors, the relatively high value may indicate the presence of a few significant deviations in the forecast. However, in electricity demand forecasting, the interpretation of error magnitude must be considered relative to the overall scale of the data (Hyndman & Athanasopoulos, 2018). The MAE offers a straightforward measure of the average forecast error, expressed in the same units as the original data. Unlike RMSE, MAE is less affected by outliers and offers a more stable indication of typical prediction error.

The MAPE, at 5.57%, expresses the average absolute error as a percentage of actual values and serves as a scale-independent measure of accuracy. According to Lewis (1982), a MAPE below 10% is generally considered highly accurate for time series forecasting. In comparison, Zaim et al. (2025) reported a lower MAPE of 2.95% when forecasting peak demand prior to the COVID-19 period. The higher error in this study reflects differences in scope, total demand versus peak demand, and the inclusion of post-COVID data (2022–2024), which exhibits greater volatility due to shifts in consumption behaviour. Nonetheless, a MAPE of 5.57% remains competitive relative to benchmarks in electricity demand forecasting, where values between 5–8% are commonly reported in similar Malaysian studies (Sim et al., 2019; Ping & Kamarudin, 2022). Therefore, the SARIMA$(0,1,0)(0,1,1)_{52}$ model provides an adequate level of forecasting accuracy, making it suitable for electricity demand forecasting and practical decision-making in energy planning and resource management.

While Zaim et al. (2025) concentrated on weekly peak demand, this study emphasizes total weekly electricity consumption, which offers broader operational benefits. Forecasting total demand provides utilities with a comprehensive understanding of cumulative energy usage, supporting generation scheduling, fuel procurement, and maintenance planning throughout the week rather than only during peak load events. It also helps identify persistent consumption trends that affect overall system efficiency, grid stability, and reserve margin requirements—factors increasingly relevant in Malaysia's post-COVID energy environment (Ugbehe et al., 2025; (Mir et al., 2020). In contrast, peak-only forecasts are mainly useful for short-term capacity planning but do not capture sustained energy usage patterns that inform medium-term operational strategies and policy formulation.

The 2022–2024 dataset captures Malaysia's electricity consumption during the post-COVID transition period, which is characterized by heightened volatility and altered usage behaviour. Weekly patterns reveal irregular fluctuations driven by hybrid work arrangements, elevated residential loads during work-from-home periods, and intermittent industrial recovery cycles (Zulkifli, 2022; Jiang et al., 2021; Nasir et al., 2023). Seasonal SARIMA structures, seasonal differencing and moving-average terms, can effectively accommodate repeated yet non-uniform weekly cycles in aggregated consumption data (Hyndman & Athanasopoulos, 2018). At the same time, sectoral and recent studies note that post-pandemic series may contain structural breaks and atypical shocks that sometimes motivate hybrid or intervention approaches in future work (Li et al., 2022; Lazo et al., 2022).

Table 5: Forecast Accuracy of SARIMA Model

Model	Forecast Accuracy (Test Set Evaluation)		
	RMSE	**MAE**	**MAPE (%)**
SARIMA$(0,1,0)(0,1,1)_{52}$	1073867.58	894782.67	5.57
SARIMA$(1,1,0)(0,1,0)_{52}$	1137123.93	955978.48	5.95
SARIMA$(0,1,0)(1,1,0)_{52}$	1105211.95	924845.26	5.76

Figure 8 demonstrates a comparative analysis between the training data, actual testing data, and predicted values using the SARIMA forecasting model. The green line represents in-sample training data, capturing historical weekly maximum electricity demand. It shows notable fluctuations with a discernible upward trend, confirming the presence of seasonality and supporting the rationale for applying the seasonal Box-Jenkins model.

In contrast, the blue line reflects the actual testing data, capturing the observed electricity demand in the holdout sample. The red dashed line presents the one-step-ahead forecast values generated from the selected SARIMA model. Visually, the predicted values appear to track the actual testing data with a relatively close alignment, despite certain periods where discrepancies are evident.

This alignment is further strengthened by the observation that the predicted series generally follows the same seasonal pattern and overall level as the actual values. The confidence intervals implicitly bound predictions, as seen by the degree of deviation between predicted and observed values. This echoes the findings of Zaim et al. (2025), achieved strong performance with MAPE below the 5% threshold, suggesting the forecasting reliability of the seasonal SARIMA framework for Malaysian electricity demand.

While the visual comparison in Figure 8 indicates that the model reasonably reflects the overall trend and short-term seasonal variation, such observations are only indicative and should be interpreted alongside quantitative performance metrics rather than as conclusive evidence of model accuracy. This discrepancy could be attributed to abrupt or unexpected changes in electricity usage due to external factors such as weather anomalies or economic activities, which are not directly captured in the SARIMA structure.

Overall, the prediction performance appears robust, affirming the SARIMA model's suitability for one-step-ahead forecasting of weekly peak electricity demand. These findings are consistent with the original study's conclusion that the seasonal Box-Jenkins model exhibits high accuracy and reliability when properly specified and validated.

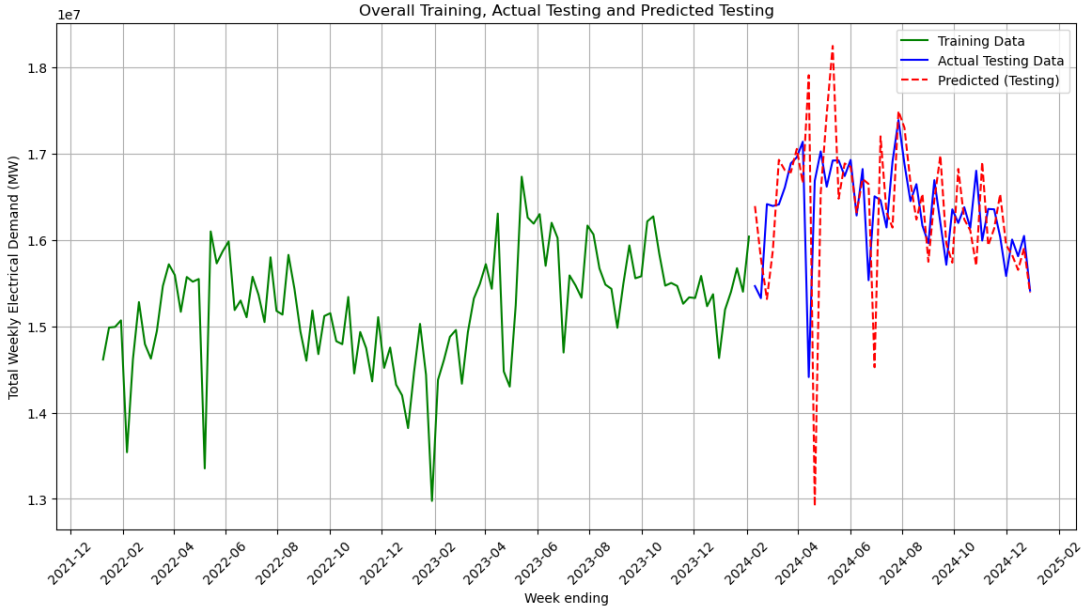

Figure 8: Comparison of Training Data, Actual Testing Data, and SARIMA Model Forecasted Values

4. Conclusion

This study concludes that the SARIMA$(0,1,0)(0,1,1)_{52}$ model demonstrates adequate and reliable forecasting performance for Malaysia's weekly electricity demand. While this value is higher than the

2.95% reported by the previous work, where for weekly peak demand forecasting, the performance difference can be attributed to the broader forecasting scope of total weekly demand and the inclusion of post-COVID consumption data, which exhibits greater volatility and behavioural variability. This acknowledgment highlights that the model performs robustly under more complex, realistic conditions rather than idealized, pre-pandemic demand patterns. Future research could explore hybrid or nonlinear approaches, such as SARIMA–TCN or SARIMA–GARCH models, to further enhance accuracy under such dynamic conditions. Its simplicity, compliance with key Box–Jenkins assumptions, particularly the white noise condition and residual independence, and its forecasting performance (MAPE = 5.57%) support its appropriateness for short-term electricity demand forecasting. While this error is higher than the 2.95% reported by previous author, for weekly peak demand forecasting, it remains below the 10% threshold commonly considered "highly accurate" for time series models and is comparable to other Malaysian electricity studies, which reported MAPEs within the 5–8% range. Therefore, rather than claiming universal superiority, this study positions the SARIMA$(0,1,0)(0,1,1)_{52}$ model as a statistically sound and contextually reliable approach for forecasting total weekly electricity demand in Malaysia's post-COVID energy landscape.

The seasonal Box-Jenkins approach implemented here provides a clear, replicable structure for forecasting, which is critical for real-time grid operation and resource planning. Beyond statistical performance, this study's contribution lies in demonstrating the operational benefits of forecasting total weekly electricity demand rather than only peak demand. Total demand forecasting provides a more comprehensive view of energy consumption patterns across residential, commercial, and industrial sectors, enabling utilities to optimize generation scheduling, fuel procurement, and maintenance planning throughout the week. Unlike peak-only models, which primarily aid short-term capacity allocation, total demand forecasts support medium-term grid stability assessments, load factor management, and resource utilization efficiency. These practical benefits strengthen the relevance of total demand modelling for Malaysia's post-COVID energy management and policy development. By leveraging weekly granularity and including post-COVID consumption behaviour, this model improves upon earlier methods that relied solely on monthly or peak data. The findings also emphasize the value of integrating seasonal components to capture recurring demand patterns effectively.

However, several limitations exist. The assumption of residual normality was not fully satisfied, likely due to the presence of outliers inherent in real-world electricity demand data. Additionally, the model's forecasting accuracy, MAPE = 5.57%, while within the 'highly accurate' range, is higher than the 2.95% reported in the reference study by Zaim et al. (2025) for weekly peak demand forecasting. This performance difference can be attributed to the broader scope of total demand modelling and the greater volatility of post-COVID consumption patterns. Furthermore, although the SARIMA$(0,1,0)(0,1,1)_{52}$ model performs well for forecasting, its extension to multistep or probabilistic forecasting remains unexplored. Future work should therefore focus on hybrid frameworks that combine SARIMA with machine learning or volatility-based models to enhance adaptability. Incorporating exogenous variables like temperature, economic indicators, or policy-related factors may further improve predictive robustness and interpretability.

In brief, SARIMA$(0,1,0)(0,1,1)_{52}$ demonstrates reliability in forecasting Malaysia's electricity demand and offers a strong foundation for future methodological enhancements and multistep applications.

5. Acknowledgement

The authors express their gratitude to Ministry of Higher Education Malaysia for supporting this work through the Fundamental Research Grant Scheme, referenced as FRGS/1/2024/STG06/UMP/02/5 (Universiti Malaysia Pahang Al-Sultan Abdullah reference: RDU240120)

6. References

Ahmad, S., & Tahar, R. M. (2014). Selection of renewable energy sources for sustainable development of electricity generation system using analytic hierarchy process: A case of Malaysia. *Renewable Energy*, *63*, 458–466. https://doi.org/10.1016/j.renene.2013.10.001

Almeida-Dias, J., Figueira, J. R., & Roy, B. (2010). Electre Tri-C: A multiple criteria sorting method based on characteristic reference actions. *European Journal of Operational Research*, *204*(3), 565–580. https://doi.org/10.1016/j.ejor.2009.10.018

Andoh, P. Y., Sekyere, C. K., Mensah, L. D., & Dzebre, D. E. (2021). Forecasting electricity demand in Ghana with the SARIMA model. *Journal of Applied Engineering and Technological Science*, *3*(1), 1–12. https://doi.org/10.37385/jaets.v3i1.316

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). John Wiley & Sons.

Brooks, C. (2008). *Introductory econometrics for finance* (2nd ed.). Cambridge University Press.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). Springer.

Chatfield, C. (2000). *Time-series forecasting*. Chapman & Hall/CRC. https://doi.org/10.1201/9781420036206

Energy Agency, I. (2020). *Global energy review 2020*. International Energy Agency. https://www.iea.org/reports/global-energy-review-2020

Ghahramani, M., Qiao, Y., Zhou, M. C., Hagan, A. O., & Sweeney, J. (2020). AI-based modeling and data-driven evaluation for smart manufacturing processes. *IEEE/CAA Journal of Automatica Sinica*, *7*(4), 1026–1037. https://doi.org/10.1109/JAS.2020.1003114

Godahewa, R., Bergmeir, C., Webb, G. I., & Montero-Manso, P. (2023). An accurate and fully-automated ensemble model for weekly time series forecasting. *International Journal of Forecasting*, *39*(2), 641–658. https://doi.org/10.1016/j.ijforecast.2022.01.008

Gujarati, D. N., & Porter, D. C. (2009). *Basic econometrics* (5th ed.). McGraw-Hill.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). OTexts. https://otexts.com/fpp2/

Ishak, I., Othman, N. S., & Harun, N. H. (2022). Forecasting electricity consumption of Malaysia's residential sector: Evidence from an exponential smoothing model. *F1000Research*, *11*, Article 54. https://doi.org/10.12688/f1000research.74877.1

Januri, S. S., Nasir, N., Ab Malek, I., & Md Yasin, Z. A. M. (2022). Forecasting unemployment rate in Malaysia using Box-Jenkins methodology during COVID-19 outbreak. *International Journal of Academic Research in Business and Social Sciences*, *12*(7), 1836–1848. https://doi.org/10.6007/ijarbss/v12-i7/13164

Jiang, P., Fan, Y. V., & Klemeš, J. J. (2021). Impacts of COVID-19 on energy demand and consumption: Challenges, lessons and emerging opportunities. *Applied Energy*, *285*, Article 116441. https://doi.org/10.1016/j.apenergy.2021.116441

Jiding, J., Meryam, S., Musa, S., Kasim, N., & Zainal, R. (2024). The challenges of developing smart house. *Research in Management of Technology and Business*, *5*(2), 549–564. https://doi.org/10.30880/rmtb.2024.05.02.051

Kien, D. T., Huong, P. D., & Minh, N. D. (2023). Application of SARIMA model in load forecasting in Hanoi city. *International Journal of Energy Economics and Policy*, *13*(3), 164–170. https://doi.org/10.32479/ijeep.14121

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, *54*(1–3), 159–178. https://doi.org/10.1016/0304-4076(92)90104-Y

Lazo, J., Aguirre, G., & Watts, D. (2022). An impact study of COVID-19 on the electricity sector: A comprehensive literature review and Ibero-American survey. *Renewable and Sustainable Energy Reviews*, *158*, Article 112135. https://doi.org/10.1016/j.rser.2022.112135

Lee, M. H. L., Ser, Y. C., Selvachandran, G., Thong, P. H., Cuong, L., Son, L. H., Tuan, N. T., & Gerogiannis, V. C. (2022). A comparative study of forecasting electricity consumption using machine learning models. *Mathematics*, *10*(8), Article 1329. https://doi.org/10.3390/math10081329

Lewis, C. D. (1982). *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*. Butterworth Scientific.

Li, Z., Ye, H., Liao, N., Wang, R., Qiu, Y., & Wang, Y. (2022). Impact of COVID-19 on electricity energy consumption: A quantitative analysis on electricity. *International Journal of Electrical Power & Energy Systems*, *140*, Article 108084. https://doi.org/10.1016/j.ijepes.2022.108084

Merza, F., Fawzy, A., AlNashef, I., Al-Zuhair, S., & Taher, H. (2018). Effectiveness of using deep eutectic solvents as an alternative to conventional solvents in enzymatic biodiesel production from waste oils. *Energy Reports*, *4*, 77–83. https://doi.org/10.1016/j.egyr.2018.01.005

Mir, A. A., Alghassab, M., Ullah, K., Khan, Z. A., Lu, Y., & Imran, M. (2020). A review of electricity demand forecasting in low and middle income countries: The demand determinants and horizons. *Sustainability*, *12*(15), Article 5931. https://doi.org/10.3390/su12155931

Mohamad, A. S. (2022). The COVID-19 total lockdown effects on grid energy demand: A Malaysia experience. In *2022 8th IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA)* (pp. 228–231). IEEE. https://doi.org/10.1109/ICSIMA55652.2022.9928918

Mohd Zain, A. Z. (2021). *Electricity demand forecasting at Kuantan Community College as impact of COVID-19 pandemic* [Unpublished manuscript or report].

Mustapa, S. I., Rasiah, R., Jaaffar, A. H., Abu Bakar, A., & Kaman, Z. K. (2021). Implications of COVID-19 pandemic for energy-use and energy saving household electrical appliances consumption behaviour in Malaysia. *Energy Strategy Reviews*, *38*, Article 100765. https://doi.org/10.1016/j.esr.2021.100765

Nabipour Afrouzi, H., Pamodha Wimalaratna, Y., Ahmed, J., Mehranzamir, K., Chuin Liew, S., Wooi, C.-L., & Mobin Siddique, B. (2021). A comprehensive review on available/existing renewable energy systems in Malaysia and comparison of their capability of electricity generation in Malaysia. In *Renewable Energy - Resource Assessment and Renewable Energy Technologies*. IntechOpen. https://doi.org/10.5772/intechopen.96586

Nasir, S. N. S., Ludin, N. A., Radzi, A. A. S. M., Junedi, M. M., Ramli, N., Marsan, A., Mohd, Z. F. A., Roslan, M. R., & Taip, Z. A. (2023). Lockdown impact on energy consumption in university building. *Environment, Development and Sustainability*, *25*(10), 12051–12070. https://doi.org/10.1007/s10668-022-02507-z

Nowotarski, J., & Weron, R. (2018). Recent advances in electricity price forecasting: A review of probabilistic forecasting. *Renewable and Sustainable Energy Reviews*, *81*, 1548–1568. https://doi.org/10.1016/j.rser.2017.05.234

Osborne, J. W. (2010). Improving your data transformations: Applying the Box-Cox transformation. *Practical Assessment, Research, and Evaluation*, *15*, Article 12. https://doi.org/10.7275/219z-1422

Pai, P. F., & Hong, W. C. (2005). Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms. *Electric Power Systems Research*, *74*(3), 417–425. https://doi.org/10.1016/j.epsr.2005.01.006

Ping, Y. Y., & Kamarudin, A. N. (2022). Forecasting the electricity demand in Malaysia using ARIMA model. *Journal of Statistical Modeling and Analytics*, *8*(1), 1–12.

Sari, C. A., Hasanah, P., Ali, S., & Dinata, W. (2023). Electricity load forecasting in East Kalimantan on religious holidays using SARIMA. *Mathematical Journal of Modelling and Forecasting*, *1*(2), 54–62.

Shrestha, M. B., & Bhatta, G. R. (2018). Selecting appropriate methodological framework for time series data analysis. *The Journal of Finance and Data Science*, *4*(2), 71–89. https://doi.org/10.1016/j.jfds.2017.11.001

Sim, S. E., Tay, K. G., Huong, A., & Tiong, W. K. (2019). Forecasting electricity consumption using SARIMA method in IBM SPSS software. *Universal Journal of Electrical and Electronic Engineering*, *6*(5), 103–114. https://doi.org/10.13189/ujeee.2019.061614

Suganthi, L., & Samuel, A. A. (2012). Energy models for demand forecasting—A review. *Renewable and Sustainable Energy Reviews*, *16*(2), 1223–1240. https://doi.org/10.1016/j.rser.2011.08.014

Tascikaraoglu, A., & Uzunoglu, M. (2014). A review of combined approaches for prediction of short-term wind speed and power. *Renewable and Sustainable Energy Reviews*, *34*, 243–254. https://doi.org/10.1016/j.rser.2014.03.033

The Sun. (2024, October 11). *Malaysia's power demand to soar 20% in next decade*. Tenaga Nasional Berhad Repository. https://irepository.uniten.edu.my/handle/123456789/33835

Ugbehe, P. O., Diemuodeke, O. E., & Aikhuele, D. O. (2025). Electricity demand forecasting methodologies and applications: A review. *Sustainable Energy Research*, *12*(1), Article 149. https://doi.org/10.1186/s40807-025-00149-z

Vivas, E., Allende-Cid, H., & Salas, R. (2020). A systematic review of statistical and machine learning methods for electrical power forecasting with reported MAPE score. *Entropy*, *22*(12), Article 1412. https://doi.org/10.3390/e22121412

Wei, W. W. S. (2006). *Time series analysis: Univariate and multivariate methods* (2nd ed.). Pearson.

Young, W. L. (1977). The Box-Jenkins approach to time series analysis and forecasting: Principles and applications. *Accounting and Business Research*, *7*(27), 228–243. https://doi.org/10.1080/00014788.1977.9729102

Zaim, S., Yaziz, S. R., Zakaria, R., Mohamad, N. N., & Ahmad Radi, N. F. (2025). Electricity demand forecasting in Malaysia using seasonal Box-Jenkins model. *Jurnal Teknologi*, *87*(2), 239–252. https://doi.org/10.11113/jurnalteknologi.v87.21867

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, *14*(1), 35–62. https://doi.org/10.1016/S0169-2070(97)00044-7

Zulkifli, Z. (2022). *Malaysia country report* (ERIA Research Project Report FY2022 No. 18). Economic Research Institute for ASEAN and East Asia.