European Journal of Computer Science and Information Technology

Vol.11, No.2, pp.37-69, 2023

Print ISSN: 2054-0957 (Print),

Online ISSN: 2054-0965 (Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

Forecasting Peak Electricity Consumption Demand in Luzon by utilizing ARIMA Model

Alexandrea Q. Albiento, Jay Zelle L. Nopia, Melanie C. Umpacan, Romie C. Mabborang, Marivic G. Molina

Pamantasan ng Lungsod ng Maynila (University of the City of Manila)

doi: https://doi.org/10.37745/ejcsit.2013/vol11n23769 Published June 5, 2023

Citation: Albiento A.Q., Nopia J.Z.L., Umpacan M.C., Mabborang R.C., Molina M.G. (2023) Forecasting Peak Electricity Consumption Demand in Luzon by utilizing ARIMA Model, *European Journal of Computer Science and Information Technology,* Vol.11, No.2, pp.37-69

ABSTRACT: *This research study focuses on forecasting the future values of peak demand in electricity consumption for Luzon, Philippines based on monthly historical data spanning from 2001 to 2020. The data was obtained from the official website of the Philippines Department of Energy (DOE). The primary objective of this study is to employ the ARIMA (Autoregressive Integrated Moving Average) model-building procedure developed by Box and Jenkins to accomplish accurate peak demand forecasting. The methodology involved conducting various tests and evaluations to identify the ARIMA model with the least Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE). After careful analysis, the best-fitting ARIMA model was determined to be ARIMA (11, 1, 12). The findings of this study indicate that, according to the ARIMA (11, 1, 12) model, Luzon's peak demand is projected to reach 10,497.65 megawatts by December 2021. Furthermore, the model predicts that by the end of 2022, 2023, and 2024, Luzon's peak demand will be approximately 10,738.34 MW, 10,953.98 MW, and 11,148.43 MW per electrical grid, respectively. The accuracy of the ARIMA (11, 1, 12) model is found to be satisfactory, with a low MAPE value of 3.639% and the most negligible RMSE value of 517.132. The implications of these forecasted peak demand values are significant for decision-makers in the energy and utilities sector. The accurate predictions provided by the ARIMA model can aid in resource allocation, infrastructure planning, and overall operational strategies to effectively meet the anticipated high-demand periods. In conclusion, this study successfully forecasts Luzon's future values of peak demand in electricity consumption using the ARIMA (11, 1, 12) model. The findings highlight the importance of accurate peak demand forecasting and provide valuable insights for energy industry professionals.*

KEYWORDS: ARIMA model, peak demand, electricity consumption, forecasting, seasonality, time series analysis.

INTRODUCTION

The world has changed since humankind began striving for greatness. Humans have always been wanderers, even in civilizations with advanced technology and other ingenious innovations that have been around for many generations. One of the most coveted resources of this modern civilization is energy. Energy is divided into two groups, renewable and non-renewable energy sources, and it can change its form. In addition, it occurs in both potential and kinetic forms. Electricity is one of the different types of energy.

The topic of global power outages has received a lot of media attention lately. According to the Philippine news source ABS-CBN, some areas of Luzon have experienced power disruptions. A National Grid Corporation of the Philippines (NGCP) representative explained, "a red alert status is issued when the power supply is insufficient to meet consumer demand and the transmission grid's regulating requirement. A yellow alert is issued when the operating margin fails to meet the transmission grid's regulating and contingency requirement." In Luzon, there is a red and yellow alert status. The same representative stated that the Luzon grid has a 12,186 MW capacity, and the estimated peak demand is 12,468 MW. Five provinces, namely Guimaras, Aklan, Antique, Capiz, and Iloilo, have been affected by the power outages in some areas of Visayas in the Philippines.

An archipelago of 7,107 islands known as the Republic of the Philippines lies on the western edge of the Pacific Ocean, north of the equator. The Philippines experiences heavy rainfall, high relative humidity, and high average temperatures. El Niña and La Niña, two climatological phenomena, have caused unprecedented droughts and floods in the nation, which are both potential causes of power outages in the Philippines.This paper discusses the impact of peak demand on. For governments, energy companies, and consumers, the urgency is growing. *Peak demand* is defined as the point in time at which it reaches its peak. Peak demand usually occurs during periods of excessive heat or high-energy activity. Peak energy demand places tremendous pressure on energy infrastructure, increasing the likelihood of power outages, blackouts, and other disruptions to energy supplies.

In statistics and econometrics, the Autoregressive Integrated Moving Average (ARIMA) model is used to quantify long-term events. It is used to understand historical data or to predict new data in a series. It is used when a metric is measured at intervals, such as daily, weekly, or monthly time intervals. The Box-Jenkins method, of which ARIMA is a subset, was developed. The two primary approaches are univariate and multivariate. The former only takes into account the values of the time series in the past when predicting the future. However, the latter creates a forecast using variables and values outside the system.

This study attempts to predict future needs to avoid disruptions that could impact everyone's way of life. In addition, it provides a comprehensive awareness of the challenges associated with regulating during periods of high demand.

A. Statement of the Problem

Peak energy demand is increasing, stressing energy infrastructure, including power generation, transmission, and distribution, and raising environmental concerns. Peak demand underscores the need for efficient energy management strategies that balance supply, demand, and environmental impacts of energy production and consumption.

In this research study, we aim to explore and analyze the forecasting of peak demand using the ARIMA (Autoregressive Integrated Moving Average) model. Peak demand forecasting is a crucial aspect in various industries, such as energy and utilities, as it helps in planning and optimizing resources to meet the anticipated high-demand periods. To achieve our research objectives, we have formulated four specific research questions. First, we seek to determine the best ARIMA model for forecasting peak demand. This involves examining different variations of the ARIMA model and identifying the one that yields the most accurate predictions.

Next, we aim to assess the accuracy of the best ARIMA model in forecasting peak demand. This evaluation will help us understand the reliability and precision of the chosen model and its suitability for real-world applications.Furthermore, we endeavor to forecast the monthly peak demand for the years 2021, 2022, 2023, and 2024. By estimating the future peak demand values, we can provide insights into the expected patterns and trends in demand over the upcoming years.

Lastly, we intend to investigate the probable implications of the forecasted peak demand values. Understanding the potential consequences and effects of these forecasted values will aid decision-makers in making informed choices regarding resource allocation, infrastructure planning, and overall operational strategies.Through this research, we hope to contribute to the field of peak demand forecasting by providing valuable insights and recommendations based on the analysis of ARIMA models.

B. Conceptual Framework

This section in this chapter discusses the conceptual framework; it includes an illustration of how this research will be conducted and briefly shows the concepts encompassing the study of forecasting peak demand in the Philippines using the best ARIMA model.

European Journal of Computer Science and Information Technology
Vol.11, No.2, pp.37-69, 2023
Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

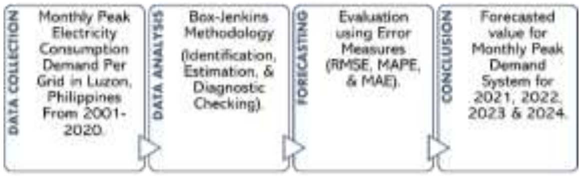

Figure 1: Conceptual Framework for Forecasting Peak Demand in Electricity Consumption using ARIMA Model

Figure 1 illustrates the conceptual framework of the study. The researchers used the 2020 Power Statistics from the Department of Energy, available online, to compile information on monthly peak system demand in Luzon, Philippines, from 2001 to 2020. The researchers used Python programming to select the best Autoregressive Integrated Moving Average Model (ARIMA) for forecasting. To ensure that the model best fits the time series data, the selected model undergoes statistical tests, selection criteria, and error measures. The projected statistics provided significant and conclusive insights.

RELATED LITERATURE

A. Energy consumption in the Philippines

Energy is a widely discussed topic around the globe, primarily because most things around are heavily reliant on energy, such as home appliances, gadgets, manufacturing equipment, cars, and more.

According to the press release from a joint survey conducted by the National Statistics Office and the Department of Energy (2013), electricity remains the most common energy source households use in the Philippines. About 87 percent of 21.0 million households used electricity from March to August 2011. The other sources used by a significant proportion of households include fuelwood, charcoal, LPG, and kerosene, with at least one-third of the total households using any of these fuel types in 2011.

Electricity was also popularly used for recreation and space cooling, with about 79% and 66% of the household's using electricity, respectively. For recreation, about three in four households (75%) used electricity for TV viewing, one in three households (34%) used it for their radio/tape recorder/stereo, and one in four households (26%) for their VCR/karaoke/videoke. At least one in three households with an average monthly income of P10,000 or more consumed electricity for their VCR/karaoke/video.

For space cooling, two in three households (65%) used electricity for electric fans, while one in 10 households (9%), for an air conditioning unit. The other uses of electricity were for ironing clothes (46% of total households), laundry (29%), cooking and food preparation (20%), computer activity (15%), water heating for bathing (4%), and water pumping (3%). In conclusion, even in a low-income household, electricity is one of the most sought.

In 2020, the Department of Energy issued a Power Situation Report in the Philippines, highlighting the power demand and supply and the current status of amidst the eruption of Taal volcano and the start of the COVID-19 pandemic. Early in the year, the eruption of the Taal volcano temporarily halted several manufacturing activities in different parts of the Luzon region. However, before that, electricity consumption slightly increased from December 2019 to January 2020.

Electricity consumption noticeably declined in February and March 2020 and showed an all-time low in April 2020. Recalling the past events from 2020, March of the same year was the pandemic's start. And the beginning of the implementation of quarantines in different parts of the Philippines. The 2020 Power Situation Report showed that there was a decrease in consumption across all sectors, except residential, which further increased by 12.2%; the recurring imposition of various community quarantine restrictions and physical distancing measures caused a significant shift in favor of the residential sector as the majority of economic sectors adopt alternative work arrangements and various “new normal” measures. Companies and organizations implemented the work-from-home scheme, online businesses, and enterprises flourished, schools shifted to online classes, discussions, and events were held virtually, and contactless transactions were widely adopted. All these activities, among others, triggered the sharp increase in electricity consumption in the residential sector.

As different sectors slowly recovered from the pandemic energy consumption shows increased demand.

B. Peak demand in electricity consumption

When electricity demand at a particular time or period exceeds the electrical network's average capacity, demand peaks are created. Electrical peak demand occurs on several different timescales and is categorized as daily, monthly, seasonal, annual, and event associated. Although peak demand occurs at around 1% of annual hours, it affects networks' stability, security, reliability, and, most importantly, the cost of energy (Palmer, 2014).

In 2020, the power situation report indicates that peak demand and electricity sales in the Philippines were lower compared to 2019. This decline can be attributed to the quarantine restrictions imposed due to the COVID-19 pandemic, resulting in a slowdown of economic activities in Luzon. The highest peak demand occurred on March 9, 2021, just before the first community quarantine was imposed. However, the subsequent Luzon-wide Enhanced Community Quarantine (ECQ) from March 15 to April 30, 2020, significantly decreased peak demand and electricity consumption.

During the ECQ period, peak demand reached an all-time low at 8,377 MW, with electricity consumption dropping to 5,314 GWh, both recorded in April. These figures reflect the direct impact of quarantine measures on energy usage, as businesses closed or operated at limited capacity and people stayed at home. However, peak demand and electricity sales began to recover in June as quarantine levels were eased to General Community Quarantine (GCQ) and Modified General Community Quarantine (MGCQ). This indicates that the gradual relaxation of restrictions allowed economic activities to resume, increasing energy demand.

Among the three primary island groups, Luzon had the highest decrease (241 MW) in peak demand during the 2020 Power Situation Report.

C. Autoregressive Integrated Moving Average (ARIMA) in forecasting

Box - Jenkins Analysis is a systematic method of identifying, fitting, checking, and using integrated autoregressive, moving average (ARIMA) time series models. The method is appropriate for time series of medium to long lengths (at least 50 observations).

According to Young, W. (1977), there are three stages in the Box-Jenkins approach to time series analysis and forecasting — identification, estimation and diagnostic checking, and the forecasts themselves. At the identification stage, we first choose a set of temporary values for the parameters p, d, and q based on an identification procedure. Diagnostic checks are then made to determine the model's representativeness vis-a-vis the data set. If an alternative model is suggested as a result of these checks, then the cycle is repeated up to this point. Forecasts are made based on the final model specification obtained from the estimation process and its associated choice of model criterion.

The primary tools in the Box-Jenkins identification procedure consist of the autocorrelation and partial autocorrelation functions. The Autocorrelation and Partial Autocorrelation Functions provide a valuable measure of the degree of dependence between values of a time series at specific intervals of separation and thus play an essential role in predicting future values of a time series. (Bolland, J.) The ACF and the PACF plot determine the values of the q and p models, respectively. The autoregressive process assumes It is a linear function of the initial values. The integrated process shows that the behavior of the time series may be affected by the cumulative effect of some processes. The moving average process shows that the current value of a moving averaging process is a linear combination of the current disturbance with one or more previous perturbations.

The moving average order indicates the number of previous periods embedded in the current value. (Fattah, J., Ezzine, L., Aman, Z., et al, 2018).

D. Related Studies

1. Energy consumption in the Philippines

(Cabauatan & Tatlonghari, 2017) Emphasizes the importance of establishing a causal relationship between and economic activities in the Philippines. Validating this relationship through empirical evidence would support prioritizing preferential support to the energy sector, aiming to increase domestic energy supply at the lowest cost while minimizing negative impacts on the public and other business sectors. Even if the evidence suggests otherwise, indicating that it does not directly drive economic growth, it still underscores the significance of expanding energy use to sustain a rapidly growing economy. Therefore, regardless of the specific findings, ensuring an adequate and stable energy supply, whether from domestic sources or imports, is considered favorable.

Additionally, a survey conducted in October 2014 revealed that most households in the study used electricity, with 87.6 percent utilizing it between October 2003 and September 2004, showing an increase compared to previous years. The number of households using LPG doubled during the same period. Gasoline and diesel were also commonly used by households for power generation and transportation. Conversely, the popularity of kerosene declined over time.

Furthermore, a recent study utilized data from the 2011 Household Survey (HECS) to analyze the Philippines' energy use and income profiles. The study identified three income classes: low-income, middle-income, and high-income. The results showed that low-income households had the lowest average monthly electricity consumption, while high-income households had the highest. The pattern was similar for LPG consumption, with middle-income households falling in between. Low-income households had higher biomass usage, while middle-income households consumed more charcoal and fuelwood. High-income households had minimal biomass consumption and higher charcoal consumption.

2. Peak demand in electricity consumption

Peak load refers to a period of high electricity demand requiring a steady power supply. In the context of India, where power shortages and regional imbalances are prevalent, accurate forecasting of electricity demand is crucial. However, the complexity of the electricity market, characterized by uncertainty, randomness, seasonality, and non-stationarity, poses challenges for informed decision-making. As India's power market becomes more competitive, sophisticated load forecasting techniques are needed to meet industry demands.

A recent study emphasizes the significance of incentivizing demand flexibility in households with heat pumps and electric vehicles. By reducing the need for additional electricity generation and transmission capacities, incentives can help address the concentration of these devices in specific areas. Grid upgrades, local production, or battery storage may be necessary alternatives. Economic incentives, market structure changes, and barrier removal are essential to promote demand flexibility and battery storage adoption. These measures facilitate the provision of flexible services to multiple markets, alleviating the reliance on grid reinforcements (Andersen et al., 2017).

3. Autoregressive Integrated Moving Average (ARIMA) in forecasting.

The increasing global energy demand calls for intelligent forecasting models and algorithms (Barak, Sadegh). One popular model for time series forecasting is ARIMA, which incorporates moving average (MA), autoregressive (AR), and differencing (I) components. This study examines the fundamental concepts of ARIMA, such as stationarity, order selection, and model estimation. It also explores the advantages and disadvantages of ARIMA and provides examples of its application in different contexts. The findings highlight the importance of understanding ARIMA's assumptions and limitations before using it for forecasting.

The time series approach is known for its flexibility and minimal assumptions (Ho & Xie, 1998). It has solid theoretical and statistical foundations and does not require a priori postulation of models when analyzing failure data. Ma et al. (2018) describe the ARIMA model forecast as a highly sophisticated technique for time series prediction that realistically captures dynamic change rules. Under specific conditions, it can be used for statistical analysis and forecasting of time series, especially for short-term predictions. In a study conducted by Pappas and Ekonomou, the proposed ARIMA model for forecasting Greek was compared to three other analytical time-series models. The results demonstrated that the ARIMA model outperformed the other models regarding efficiency. However, it should be noted that statistical forecasting techniques are often required for regular data, while big data sets or seasonal data patterns may present challenges (references [2, 9, 10]). Although real-world time series data rarely exhibit linear structures, ARIMA models can still provide forecasts based on both economic and non-economic indices.

E. Synthesis

Energy is a crucial and widely discussed topic worldwide as it plays a vital role in powering various aspects of modern life, including home appliances, gadgets, manufacturing equipment, and transportation. In the context of the Philippines, it is essential to establish a clear relationship between and economic activities. While empirical evidence may not directly establish a causal link between economic growth, expanding energy use remains essential for sustaining a rapidly growing economy. Therefore, it is crucial to prioritize support for the energy sector, intending to increase the domestic energy supply at the lowest cost while minimizing any negative impacts on the public and other business sectors. Regardless of the specific findings, ensuring a sufficient and stable energy supply, whether from domestic sources or imports, is considered favorable. This approach will contribute to overall economic development and improve the quality of life for the population.

In addition to the significance of, it is essential to address the peak demand for electricity. Peak demand occurs when the electricity demand exceeds the average capacity of the electrical network during specific times or periods. Such peaks in demand can have significant implications, affecting stability, security, reliability, and energy cost. In countries like India, where power shortages and regional imbalances are prevalent, accurately forecasting electricity demand becomes crucial. The complexity of the electricity market poses challenges for informed decision-making, as it is characterized by uncertainty, randomness, seasonality, and non-stationarity. Sophisticated load forecasting techniques must address these challenges and fulfill the industry's demands. These techniques enable power providers to predict high-demand periods and effectively allocate resources to maintain a steady power supply. By improving demand forecasting through advanced methods, power systems can be managed more efficiently in an evolving market.

One such method for analyzing time series data and forecasting energy demand is Box-Jenkins Analysis. It is a systematic approach that involves identifying, fitting, checking, and utilizing Autoregressive Integrated Moving Average (ARIMA) models. ARIMA, a popular time series forecasting model, combines moving average, autoregressive, and differencing techniques to capture the patterns and trends in the data. Understanding the fundamental concepts of ARIMA, such as stationarity, order selection, and model estimation, is crucial for its practical application in forecasting energy demand. Moreover, it is essential to recognize ARIMA's advantages, disadvantages, assumptions, and limitations before relying on it for accurate forecasts and informed decisions regarding energy demand. By considering these factors, stakeholders can ensure reliable and effective forecasting, leading to better planning and management of energy resources.

METHODOLOGY

A. Research Design

In this study, the researchers employed the Autoregressive Integrated Moving Average (ARIMA) process to forecast the acquired time series data. ARIMA belongs to the class of nonstationary methods. Unlike stationary approaches, ARIMA does not assume that the process remains in a state of statistical equilibrium or that the probabilistic characteristics of the process remain constant. In simpler terms, ARIMA does not assume a fixed, consistent mean level or a constant variance for the process. The ARIMA process breaks down the obtained time series data into Autoregressive (AR), Integrated (I), and Moving Average (MA) processes. The Autoregressive

Vol.11, No.2, pp.37-69, 2023

Print ISSN: 2054-0957 (Print),

Online ISSN: 2054-0965 (Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

(AR) part of ARIMA represents the weighted moving average over past observations from the data. The Integrated (I) part represents the linear or polynomial trends of the time series data, and the Moving Average part represents the weighted moving average from past errors. AR, I, and MA parts of the ARIMA process are indicated by the p,d, and q parameters, respectively, of the ARIMA model wherein:

p = order of autocorrelation
d = order of integration (differencing)
q = order of moving averages
Hence, the ARIMA model is expressed as:

$$X_t = \theta_0 + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} ++ \varphi_p X_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \theta_q e_{t-q}$$

ARIMA consists of four significant steps for model building: Identification, Estimation, Diagnostics, and Forecast. For the first step, the fitting ARIMA model needs to be identified for the particular datasets, and the parameters should have the smallest possible values to analyze the data correctly and forecast accordingly. Figure 2 shows Box and Jenkins's four-stage procedure for finding the best model.

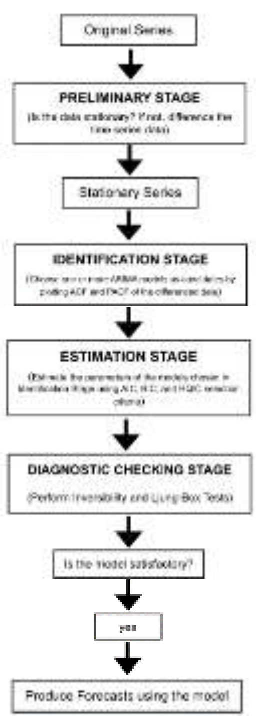

Figure 3: The Model Building Procedure

Figure 3 shows Box and Jenkins's four-stage procedure for finding the best model. As can be seen from the figure, a preliminary stage is taken to see if the time-series data is stationary or not. The gathered data must be stationary so that the researchers can assume that the process is in statistical equilibrium with probabilistic properties that remain constant over time. To test the stationarity of the data, the researchers used Augmented-Dickey-Fuller (ADF), Kwiatkowski, Phillips, Schmidt, and Shin (KPSS), and Phillips-Perron (PP) tests. If the data are nonstationary, differencing the time-series data must be necessary until the researcher can say through these tests that the data is stationary. For identifying the possible ARIMA models, the researchers used the ACF and PACF plots of the stationary data. In the estimation stage, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan-Quinn Information Criterion (HQIC) are used to find the best ARIMA models. After finding the best possible ARIMA models, the Inversibility and Ljung Box tests are performed for the diagnostic checking stage. Lastly, the model with the least Mean Absolute Percentage and Root Mean Square errors will be used for the forecasting procedure.

B. Data Collection Procedure

The time-series data was sourced from the official website of the Department of Energy of the Philippines. Their website provides various information about power statistics in the Philippines. The researchers retrieved the 2001-2020 Peak Demand per Grid file from this website since the study's objective is to forecast the peak demand in Luzon. The gathered data is then transferred to another Microsoft Excel file and saved in a CSV format. The organized CSV file is analyzed using the Python programming language in Visual Studio Code computer software. *Python* is a popular programming language released in 1991 by Guido van Rossum that is mainly used for web development, software development, system scripting, and mathematics. At the same time, Visual Studio Code computer software is a source-code editor made by Microsoft running for Windows, Linux, and macOS.

RESULTS AND DISCUSSIONS

B. Time Series Data Plot

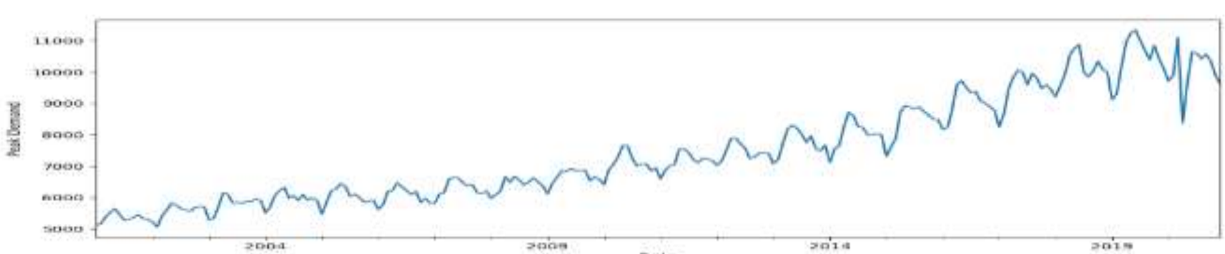

Figure 4.1: Peak Demand in Luzon, Philippines from years 2001-2020

European Journal of Computer Science and Information Technology
Vol.11, No.2, pp.37-69, 2023
Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

The graph consists of 240 observable data points in Luzon, Philippines, from 2000 to 2020. The data shows a drastic drop from the year 2020, one factor the researcher can deduct from the given data is that the year 2020 is when the COVID-19 pandemic started in the Philippines, and it influences how people consume energy from day to day. It also shows that June 2019 had the highest peak demand, and February 2002 showed the lowest peak demand. From the graph, the researchers can identify that the given data set is nonstationary and is not eligible for different tests and selection criteria for selecting the best fit ARIMA model.

C. Training and Test Data

The given dataset for peak demand in Luzon, Philippines, from 2001-2020 shows nonstationary; thus, before transforming the dataset to stationary, it is necessary to split the data into training and testing sets. The training set is used to develop models and feature sets; they are the substrate for estimating parameters, comparing models, and all the other activities required to reach the final model, while the test set is used only after these activities for estimating a final and unbiased assessment of the model's performance. This is necessary to avoid biases in the outcome.

The researchers opted to use 80% of the data as the training sets and 20% as the test set for the final forecasting model. The training set consists of 192 observable data points from 2001 to 2016, while the test set consists of 48 observations from 2017 to 2020.

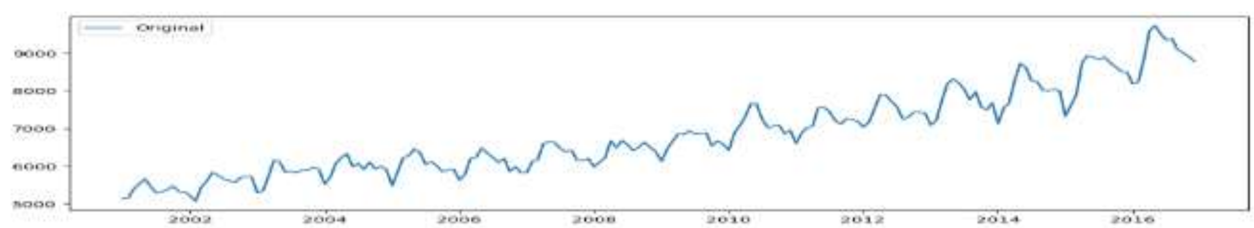

Figure 4.2: Training Set from the original data series

The figure above presents the newly introduced dataset, which was used for various tests, including differencing, ADF, PP, KPSS Tests, and others. The selection criterion was applied to identify the best-fit ARIMA model, which was subsequently used to forecast the monthly peak demand for the years 2021 to 2023.

European Journal of Computer Science and Information Technology
Vol.11, No.2, pp.37-69, 2023
Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

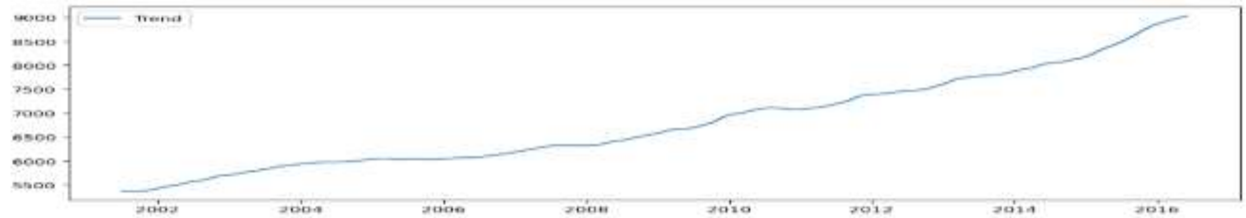

Figure 4.3: Trend from the Training Set

The graph derived from the training set illustrates the trend exhibited by the time series data. A trend refers to the linear growth or decline observed in the series over time. Upon examining the graph, it is evident that the time series follows an upward trend, displaying a consistent ascending pattern. This indicates a linear increase in values from the initial year until 2016.

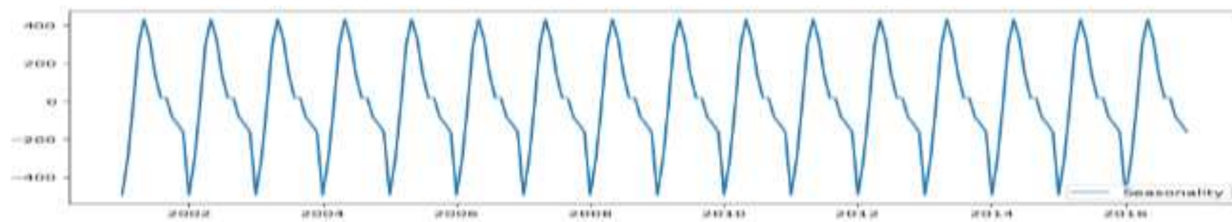

Figure 4.4: Seasonality of the Training Set

The training set was used to extract seasonality, as depicted in the figure above. The graph substantiates the presence of seasonality in the time series. To address seasonality in the time series analysis, the data underwent differencing.

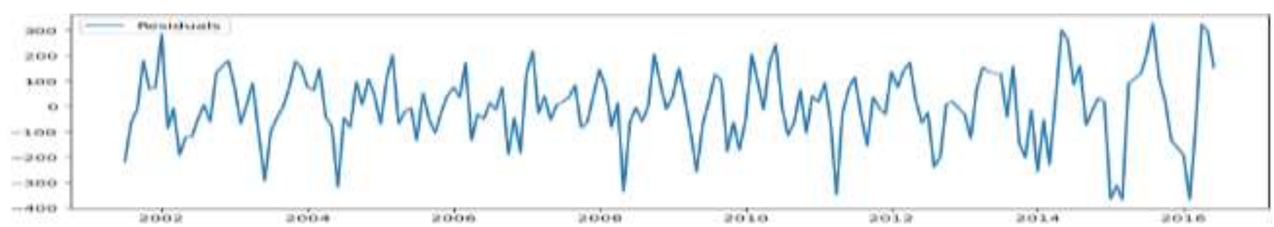

Figure 4.5 Residuals from the Training Set

The residual plot of the training set is depicted in Figure 4.5. To ensure a suitable forecasting method, it is necessary for the residuals to exhibit certain properties: uncorrelated residuals, zero mean residuals, constant variance in residuals, and normally distributed residuals.

C. Test for Stationarity (ACF, PACF Plots)

European Journal of Computer Science and Information Technology
Vol.11, No.2, pp.37-69, 2023
Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

Once the time series was plotted, it was necessary to examine the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the given time series data. These functions determined the stationarity of the series and influenced the selection of values for the ARIMA process. Specifically, the ACF determined the values for the moving average (MA) parameter, denoted as q, while the PACF helped determine the values for the autoregressive (AR) parameter, denoted as p, in the ARIMA models.

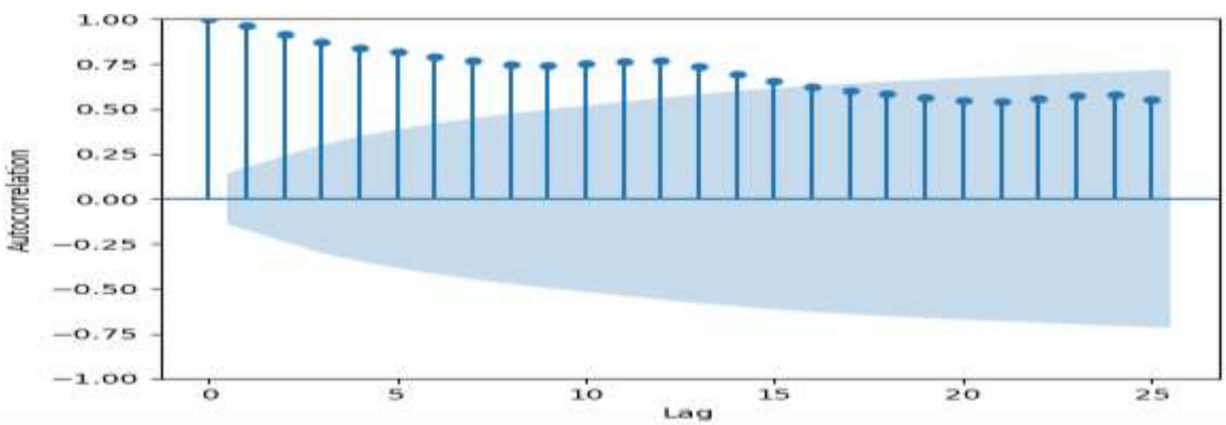

Figure 4.6: ACF of the Peak Demand Training Set

Figure 4.6 presented the ACF plot, which showed a gradual decrease in the plotted ACF data. This observation led the researchers to conclude that the training set was not stationary. Non-stationary data displayed a higher and positive autocorrelation coefficient, denoted by r_1. The blue area in the plot represented the 95% confidence interval, indicating the significance threshold. The ACF plot played a vital role in determining the appropriate value for q in the ARIMA model. Thus, it was imperative to verify the stationarity of the dataset.

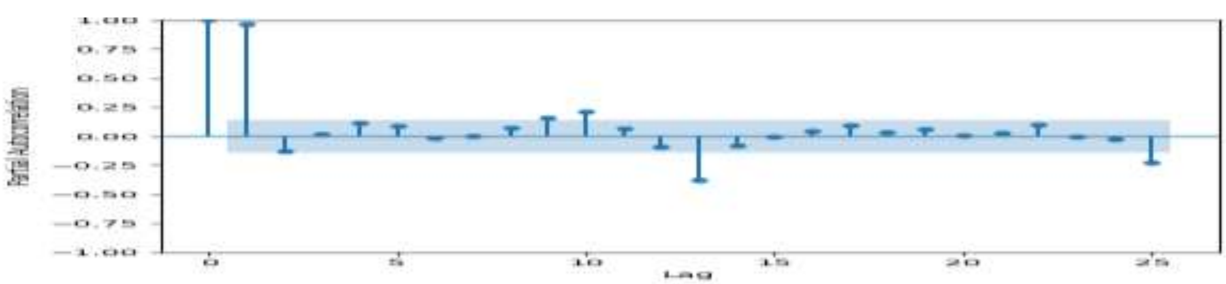

Figure 4.7 : PACF of the Peak Demand Training Set

European Journal of Computer Science and Information Technology

Vol.11, No.2, pp.37-69, 2023

Print ISSN: 2054-0957 (Print),

Online ISSN: 2054-0965 (Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

As the given dataset did not exhibit stationarity in the ACF plot, the same lack of stationarity was observed in the PACF plot. The PACF plot examines the correlation between the current and lagged variables while removing the influence of correlation from previous lags. Essentially, the PACF identifies and eliminates the lags responsible for autocorrelation. It is important to note that the absence of stationarity in the time series would render the ARIMA model ineffective, making it impossible to proceed with the framework.

D. Test for Stationarity (Augmented Dickey-Fuller's, Philipps-Perron's, Kwiatkowski-Phillips-Shmidt-Shin's tests)

To further assess the stationarity of the time series, additional tests such as the Augmented Dickey-Fuller Test, Philipps-Perron Test, and Kwiatkowski-Phillips-Schmidt-Shin's Test were conducted. The table below presents the results of these tests, obtained using Python programming. These tests provide valuable insights into the stationarity of the time series.

Table 4.1 : ADF Test of the Peak Demand Training Set

```
ADF :  -0.07847388785782463
P-Value:  0.9515489312172088
Num OF Lags :  15
Num of Observations Used for ADF Regression and Critical Values Calculation : 224
Critical Values :
         1% : -3.459884913337196
         5% : -2.8745310704320794
         10% : -2.573693840082908
```

The ADF test was performed using Python programming to assess the stationarity of the time series. The null hypothesis is rejected if the p-value is less than 0.05. After conducting the ADF test, it was found that the p-value for this time series was 0.951549, which exceeded the threshold of 0.05. Consequently, the null hypothesis was not rejected. Therefore, based on the ADF test results, it can be concluded that the given time series data was nonstationary.

Table 4.2: PP Test of the Peak Demand Training Set

```
     Phillips-Perron Test (Z-tau)
=====================================
Test Statistic                 -0.882
P-value                         0.794
Lags                               15
-------------------------------------

Trend: Constant
Critical Values: -3.47 (1%), -2.88 (5%), -2.57 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.
```

European Journal of Computer Science and Information Technology
Vol.11, No.2, pp.37-69, 2023
Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

The PP test was performed to assess the stationarity of the dataset, where the null hypothesis is rejected if the p-value is less than 0.05. To determine the dataset's stationarity, it was necessary for the resulting p-value to be below 0.05. Upon analyzing the test results, it was observed that the p-value was 0.794, exceeding the threshold of 0.05. Consequently, there was insufficient evidence to reject the null hypothesis. Thus, it can be concluded that the dataset had a unit root and was nonstationary.

Table 4.3: KPSS Test of the Peak Demand Training Set

```
KPSS :  1.9279933731079535
P-Value:  0.01
Num Of Lags :  9
Critical Values :
         10% : 0.347
         5% : 0.463
         2.5% : 0.574
         1% : 0.739
```

To ascertain the stationarity of the dataset, it was necessary to examine the p-value with a threshold of 0.05. Unlike the previous tests, the p-value in the KPSS test should be greater than the critical value of 5% in order to retain the null hypothesis. Upon conducting the KPSS test on the provided dataset, the resulting p-value was determined to be 0.01, which is lower than 0.05. As a result, sufficient evidence was found to suggest the presence of a unit root and nonstationarity in the time series.

E. Transformation of Time-Series Through Differencing

Following the initial tests, the researchers concluded that the time series data was nonstationary. As a result, the time series had to be differenced, which entailed calculating successive changes in the value of the time series.

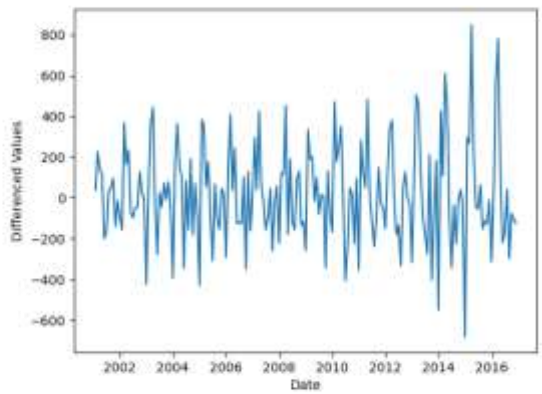

Figure 4.8: Peak Demand Training Set First Order Differencing

Figure 4.8 illustrates the first-order differencing of the time series of peak demand in electricity consumption. It also helps stabilize the mean of a time series by removing changes in the level and eliminating trends and seasonality. Compared to the previous model shown in Figure 4.2, the graph above appears to be stationary. A stationary time series shows no predictable pattern.

F. Test for Stationarity (ACF, PACF Plots)

In order to check whether the first-order differencing in the time series significantly changed the non-stationarity of the time series' training test, the ACF and PACF must be plotted. The figures below depict the ACF and the PACF of the differenced time series data of peak demand in electricity consumption.

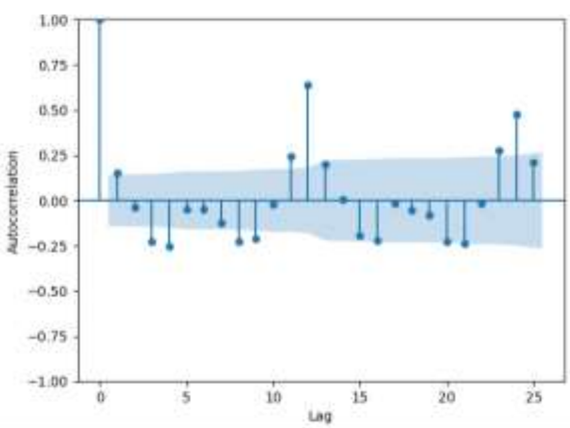

Figure 4.9 : ACF plot of the Training Set (First Order Differenced)

The above ACF plot exhibits notable lag values that are crucial in determining the moving average model's value. In contrast to the previous ACF plot, the depicted figure demonstrates stationarity

European Journal of Computer Science and Information Technology
Vol.11, No.2, pp.37-69, 2023
Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

as the ACF decays rapidly to zero. As for the q model, potential values include 1, 3, 4, 8, 9, 11, 12, 23, and 24.

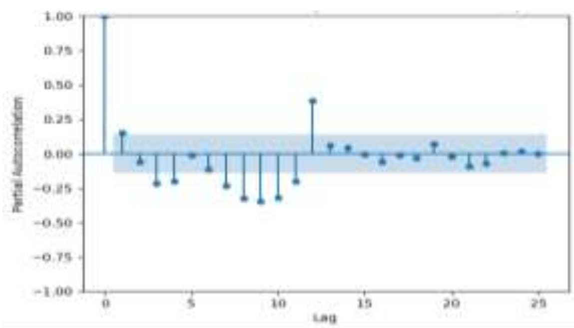

Figure 4.10 PACF plot of the Training Set (First Order Differenced)

The above PACF plot reveals significant lag values that play a vital role in determining the value of the autoregressive model. Based on the corresponding ACF plot, the figure above is deemed to exhibit stationarity. The potential values for the p model, therefore, include 1, 3, 4, 7, 8, 9, 10, 11, and 12.

G. Test for Stationarity (Augmented Dickey-Fuller's, Philipps-Perron's, Kwiatkowski-Phillips-Shmidt-Shin's tests)

To verify the stationarity of the differenced time series data, the researchers utilized the previous three stationarity tests as evidence of its stationarity. The values obtained for the stationarity tests are presented below using Python programming.

Table 4.4 ADF Test of the Peak Demand Training Set (First Order Differenced)

```
ADF :  -3.765074787474299
P-Value:  0.003285554254725
Num Of Lags :  13
Num of Observations Used for ADF Regression and Critical Values Calculation : 177
Critical Values :
         1% : -3.467845319799907
         5% : -2.878011745497439
         10% : -2.575551186759871
```

To reject the null hypothesis for the ADF test, the p-value had to be less than 0.05. According to the results obtained from the Python programming, the p-value was found to be 0.003286, which

European Journal of Computer Science and Information Technology

Vol.11, No.2, pp.37-69, 2023

Print ISSN: 2054-0957 (Print),

Online ISSN: 2054-0965 (Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

was lower than 0.05. As a result, the null hypothesis was rejected, indicating that the differenced time series was stationary based on the ADF test.

Table 4.5 PP Test of the Peak Demand Training Set (First-Order Differenced)

```
     Phillips-Perron Test (Z-tau)
=====================================
Test Statistic                -13.381
P-value                         0.000
Lags                               15
-------------------------------------

Trend: Constant
Critical Values: -3.47 (1%), -2.88 (5%), -2.57 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.
```

The null hypothesis in the PP test is rejected when the p-value is less than 0.05. After analyzing the data using Python programming, the resulting p-value was $4.99\ x\ 10^{-25}$, which is below the significance level of 0.05. Consequently, based on this evidence, the null hypothesis is rejected, providing sufficient grounds to conclude that the differenced time series data is stationary.

Table 4.6: KPSS Test of the Peak Demand Training Set (First-Order Differenced)

```
KPSS :  0.031585666663892812
P-Value:  0.1
Num Of Lags :  8
Critical Values :
        10% : 0.347
        5% : 0.463
        2.5% : 0.574
        1% : 0.739
```

In order to ascertain the stationarity of the time series, it was necessary for the p-value to be greater than 0.05, unlike the previous two tests where the null hypothesis was rejected. According to the results of the KPSS test, the p-value was found to be 0.1, which exceeded the critical value of 0.05. Therefore, it can be concluded that the time series did not possess a unit root and was stationary.

The three conducted tests, affirming the stationarity of the time series data, confirm its suitability for undergoing procedures aimed at formulating a forecasting model for peak demand in electricity consumption.

H. Identifying Best Arima Model Using AIC, BIC, & HQIC

European Journal of Computer Science and Information Technology
Vol.11, No.2, pp.37-69, 2023
Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

Based on the values obtained from the PACF and ACF plots, a total of 81 models were identified. The researchers then selected the eight best models, considering the lowest values of AIC, BIC, and HQIC, in order to formulate the forecasting model for the test data concerning peak demand in electricity consumption.

Table 4.7: **Best ARIMA Model using selection criterion (AIC, BIC, & HQIC)**

Model	AIC	BIC	HQIC
12, 1, 12	2494.787	2576.093	2527.720
12, 1, 1	2498.684	2544.216	2517.126
12, 1, 11	2500.043	2578.096	2531.658
12, 1, 9	2501.531	2573.081	2530.512
12, 1, 3	2501.866	2553.902	2522.943
12, 1, 4	2502.609	2557.898	2525.004
12, 1, 24	2508.338	2628.672	2557.079
11, 1, 12	2509.558	2587.611	2541.172

Table 4.7 presents a summary of the selected eight models, ranked based on their lowest values for AIC, BIC, and HQIC, respectively. Upon examining the table, it was observed that ARIMA (12, 1, 12) had the lowest AIC, ARIMA (12, 1, 1) had the lowest BIC, and ARIMA (12, 1, 1) had the lowest HQIC value. However, the results exhibited inconsistencies. Consequently, the researchers chose to include all eight models and further investigated using additional tests to determine the best-fitting model among them.

I. Diagnostic Test (Invertibility - ARMA Structure)

To ensure the accuracy of the ARIMA model, diagnostic tests were performed as a crucial step in line with the Box-Jenkins Methodology to assess its suitability for forecasting. The verification of invertibility involves examining whether the autoregressive (AR) and moving average (MA) roots lie outside the unit circle. In Python programming, it is convenient to determine invertibility by finding the inverse roots of the AR and MA models. For invertibility, it is necessary that the inverse roots of the AR and MA models lie within the unit circle. Provided below are the invertibility tests conducted for all eight models.

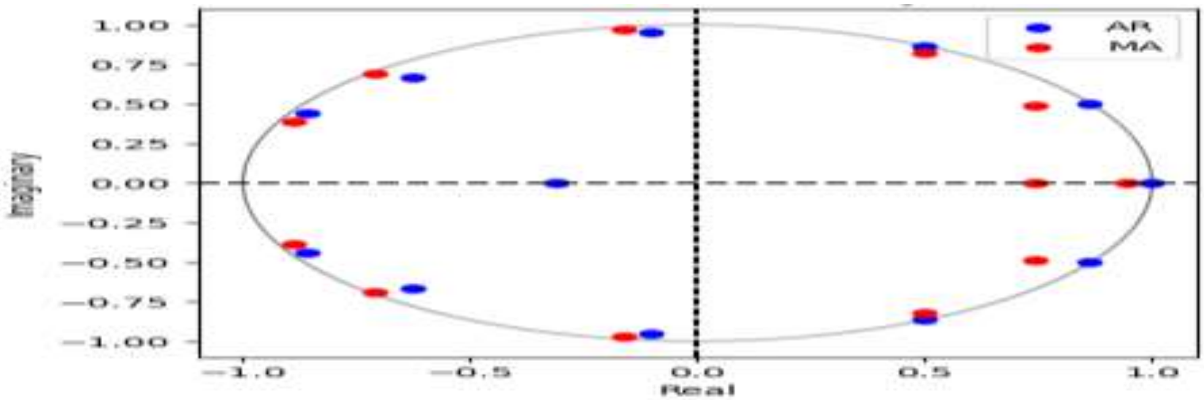

Figure 4.11: Inverse Roots of ARIMA Model (12, 1, 12)

Figure 4.11 illustrates the inverse roots of the ARMA structure. The depicted figure reveals that all 12 AR and MA roots are positioned inside the unit circle, indicating that the model is well-suited for forecasting purposes.

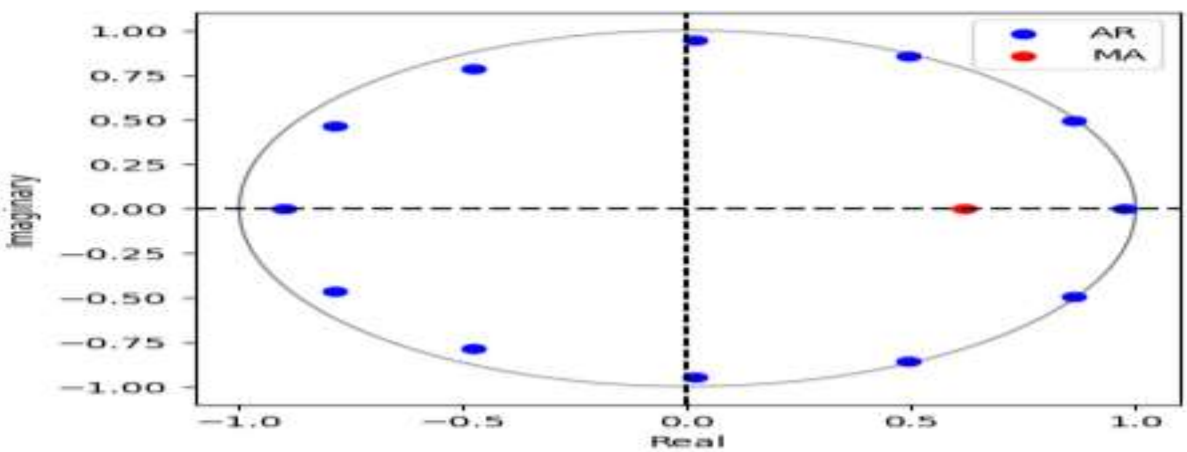

Figure 4.12: Inverse Roots of ARIMA Model (12, 1, 1)
Figure 4.12 depicts the inverse roots of the ARMA structure. From the presented figure, it can be observed that all 12 AR roots and 1 MA root are situated within the unit circle. This indicates that the model is suitable for forecasting purposes.

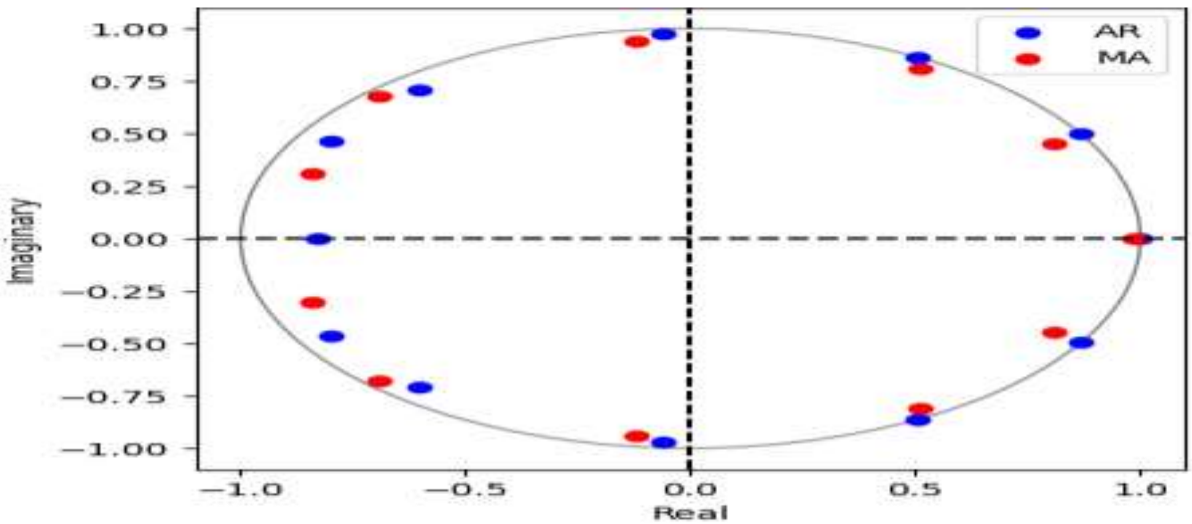

Figure 4.13 Inverse Roots of ARIMA Model (12, 1, 11)

Figure 4.13 illustrates the inverse roots of the ARMA structure. The depicted figure demonstrates that all 12 AR roots and 11 MA roots are positioned within the unit circle. This indicates that the model is well-suited for forecasting purposes.

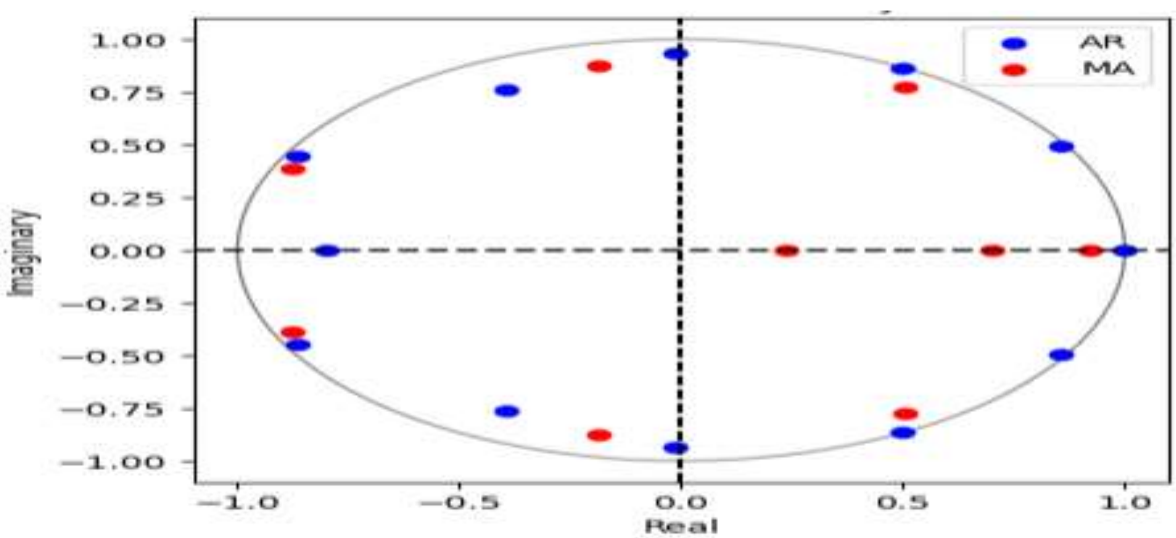

Figure 4.14: Inverse Roots of ARIMA Model (12, 1, 9)

The inverse roots of the ARMA structure are presented in Figure 4.14. The figure indicates that all 12 AR roots and 9 MA roots lie within the unit circle. This observation suggests that the model is suitable for accurate forecasting.

Figure 4.15: **Inverse Roots of ARIMA Model (12, 1, 3)**

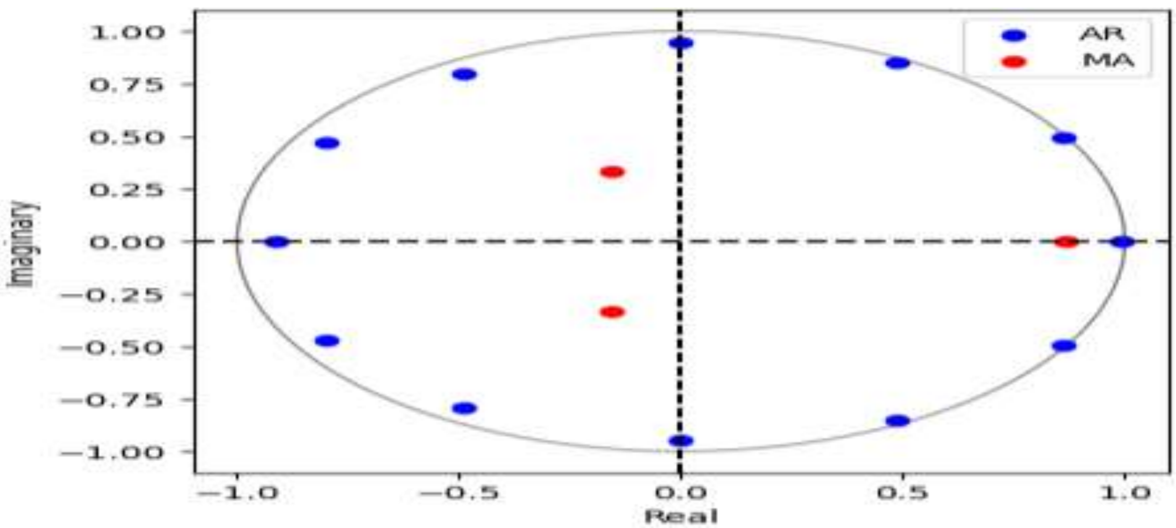

Figure 4.15 displays the inverse roots of the ARMA structure. From the above figure, it is evident that all 12 AR roots and 3 MA roots reside within the unit circle. This indicates that the model is well-suited for effective forecasting.

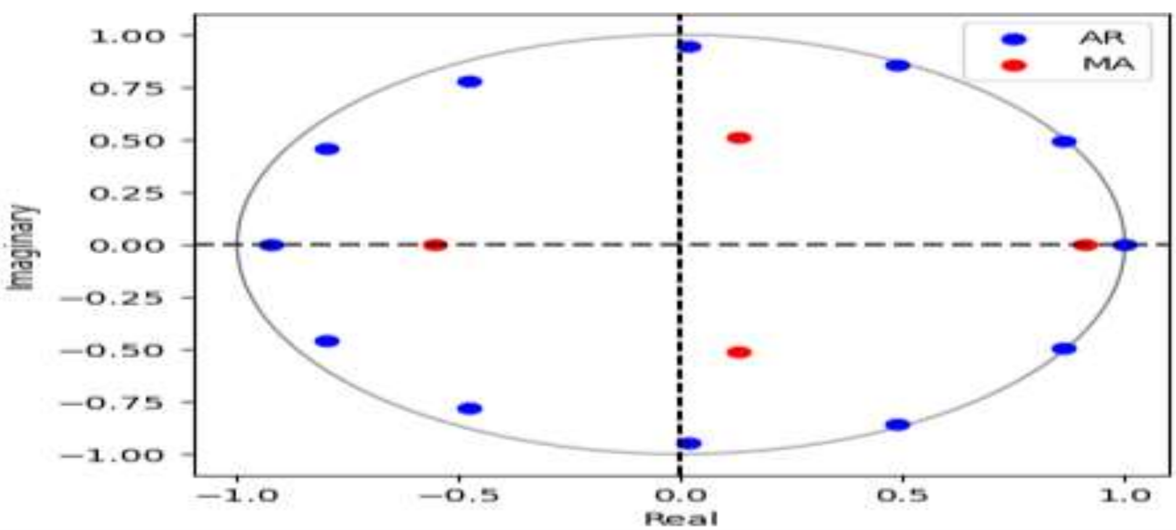

Figure 4.16Inverse Roots of ARIMA Model (12, 1, 4)

The inverse roots of the ARMA structure are depicted in Figure 4.16. As observed in the figure, all 12 AR roots and 4 MA roots are located inside the unit circle. This suggests that the model is suitable for accurate forecasting purposes.

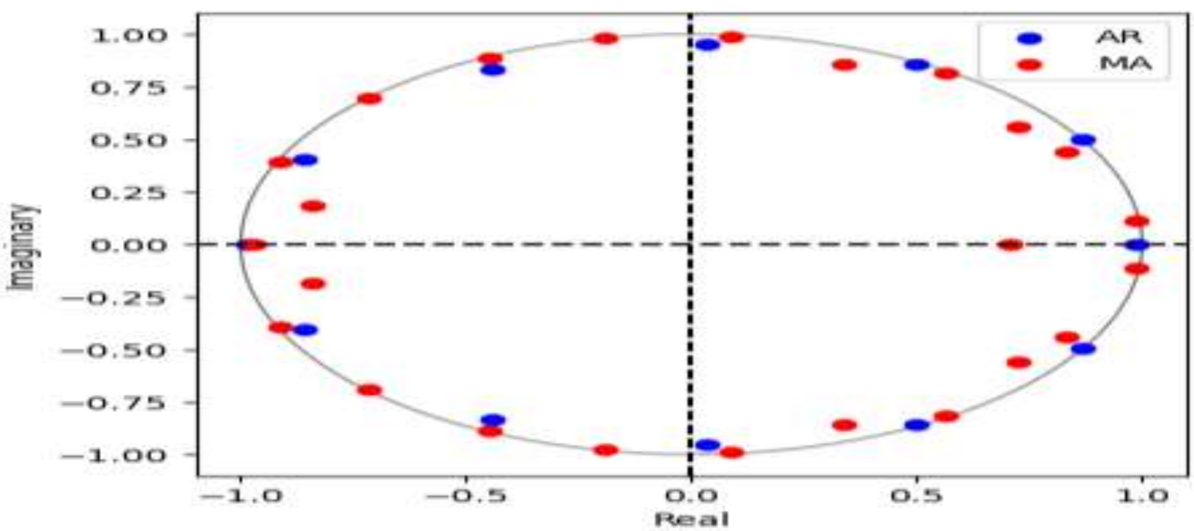

Figure 4.17 Inverse Roots of ARIMA Model (12, 1,24)

Figure 4.17 provides a visualization of the inverse roots of the ARMA structure. From the above figure, it is evident that all 12 AR roots and 24 MA roots reside within the unit circle. This indicates that the model is suitable for accurate forecasting purposes.

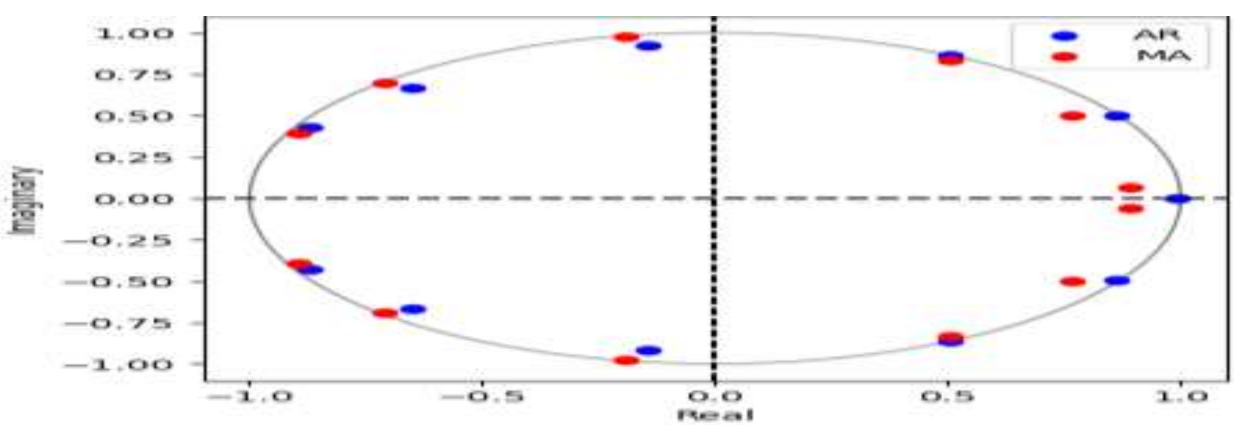

Figure 4.18Inverse Roots of ARIMA Model (11, 1,12)

Vol.11, No.2, pp.37-69, 2023
Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

Figure 4.18 displays the inverse roots of the ARMA structure. As depicted in the figure, all 11 AR roots and 12 MA roots are situated within the unit circle. This observation indicates that the model is well-suited for accurate forecasting.

J. Diagnostic Test (Independence of Residuals - Ljung Box Test)

Another diagnostic test used to assess the model is the Ljung-Box Test. To fully comprehend this test, it is important to establish the hypothesis as follows:

H_0: The residuals are independently distributed.

H_A: The residuals are not independently distributed; they exhibit serial correlation.

The purpose of this test is to examine whether the residuals exhibit independence in their distribution. To retain the null hypothesis, the decision rule states that the p-value should be greater than 0.05. In other words, if the p-value is above this threshold, we fail to reject the null hypothesis, indicating that the residuals are likely independently distributed.

MODEL (p,d,q)	LJUNG BOX TEST STATISTIC	LJUNG BOX TEST P-VALUE
12, 1, 12	1.771625	0.999887
12, 1, 1	1.923815	0.999818
12, 1, 11	1.358357	0.999976
12, 1, 9	1.715938	0.999906
12, 1, 3	1.822958	0.999866
12, 1, 4	1.806734	0.999873
12, 1, 24	1.734264	0.9999
11, 1, 12	2.054857	0.999736

Table 4.8 Best ARIMA Model using Ljung-Box Test

In the Ljung-Box Test, it is not necessary for the model to have the lowest p-value; rather, it must satisfy the condition of having a p-value greater than 0.05 to pass the test. All eight models in our study exhibit p-values that exceed 0.05. Therefore, we can conclude that the residuals of all eight models are independently distributed.

K. Forecasting Evaluation (Error Measures)

The evaluation of the models involved assessing their forecasting accuracy using error measures such as MAPE, MAE, and RMSE.

European Journal of Computer Science and Information Technology
Vol.11, No.2, pp.37-69, 2023
Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

MAPE, which represents the mean of all absolute percentage errors between the predicted and actual values, serves as an important indicator of a model's accuracy. A desirable outcome is to have a MAPE value below 10%, indicating a high level of accuracy. Based on the results obtained, all eight models achieved a MAPE value below 10%, which is considered very good according to the value interpretation table.

The MAE score of a model provides insight into its accuracy, where a lower MAE score indicates better performance.

Furthermore, the RMSE value is another valuable error measure that helps determine the model's suitability for forecasting. Lower RMSE values indicate more precise predictions, whereas higher RMSE values suggest increased errors and less accurate predictions. The models are evaluated using error measures, such as the MAPE, MAE, and RMSE, to verify the forecasting accuracy of the models.

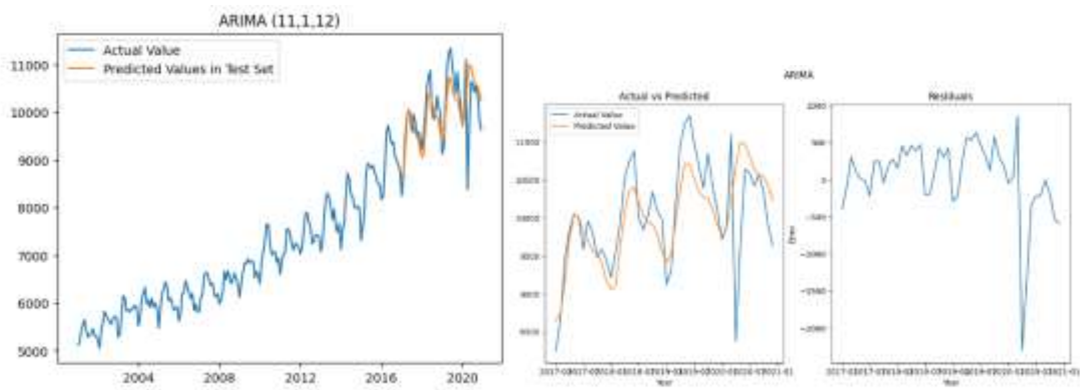

Figure 4.19 Mapping comparison of ARIMA (11, 1, 12) to the actual value

The accuracy of the predicted values is assessed by comparing them to the actual values using the test set obtained from data splitting. In this study, a splitting ratio of 80% for training data and 20% for testing data was employed. The decision to split the data into training and test sets was made to avoid biased outcomes that may arise from using the entire dataset for all ARIMA procedures and forecasting.

Figure 4.19 presents the outcomes of the sample forecast for ARIMA (11, 1, 12), along with the corresponding residuals.

L. Sample Forecast

European Journal of Computer Science and Information Technology
Vol.11, No.2, pp.37-69, 2023
Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

The utilization of ARIMA (11, 1, 12) resulted in the prediction depicted in the figure below. The orange line represents the ARIMA prediction, while the blue line represents the observed values. Additionally, a 95% confidence interval is displayed to provide a measure of the prediction's uncertainty.

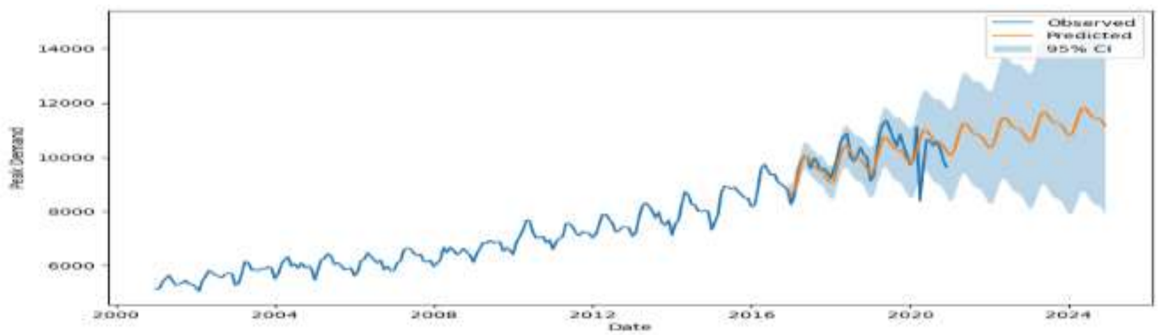

Figure 4.20: ARIMA (11, 1, 12) Prediction with 95% Confidence Intervals

Figure 4.20 showcases the ARIMA (11, 1, 12) prediction along with a 95% confidence interval. The graph of the predicted value is noteworthy as it falls within the significance threshold. Furthermore, among the various models considered, ARIMA (11, 1, 12) demonstrates the lowest error measures, indicating its strong performance in accurately mapping the comparison with the actual value.Figure 4.20 illustrates ARIMA (11, 1, 12) prediction with 95% confidence interval. The graph for the predicted value shows significance as it is within the significance threshold, also among the models, ARIMA (11, 1, 12) has the least error measures, therefore, showed great result in mapping comparison with the actual value

Table 4.10 Peak Demand Forecast for 2021 with 95% Confidence Interval

Month	Forecasted Value	Lower Peak Demand	Upper Peak Demand
January	10058.275054	8385.207522	11731.342586
February	10159.383316	8441.685758	11877.080873
March	10508.931741	8747.600152	12270.263330
April	10947.347377	9148.669503	12746.025251
May	11235.385522	9398.561191	13072.209853
June	11224.732576	9353.733011	13095.732141
July	11066.368780	9163.524397	12969.213162
August	10886.547324	8956.888168	12816.206480
September	10831.366196	8879.558610	12783.173782
October	10805.245420	8830.693365	12779.797475
November	10696.133534	8695.159088	12695.107979
December	10497.649019	8462.866485	12532.431553

In Table 4.10, the forecasted values for peak demand in 2021 are presented, along with the corresponding lower and upper bounds. The results indicate that the highest forecasted peak demand value for 2021 occurs in May, reaching 11235.39 MW, while the lowest forecasted value is observed in January, at 10058.28 MW. Additionally, January exhibits the lowest lower peak demand value at 8385.21 MW, whereas June showcases the highest upper peak demand value at 13095.73 MW.

Table 4.11 Peak Demand Forecast for 2022 with 95% Confidence Interval

Month	Forecasted Value	Lower Peak Demand	Upper Peak Demand
January	10332.577075	8256.766408	12408.387743
February	10420.597753	8300.532053	12540.663453
March	10746.003083	8583.776383	12908.229783
April	11176.841730	8975.804869	13377.878591
May	11444.697001	9205.932449	13683.461553
June	11447.390602	9173.730449	13721.050756
July	11282.817966	8975.726616	13589.909287
August	11108.927911	8773.856126	13443.999696
September	11061.200195	8701.447180	13420.953209
October	11033.949066	8650.341834	13417.550199
November	10938.536206	8527.740246	13349.332166
December	10738.336526	8292.468518	13184.203534

Table 4.11 presents the forecasted values for peak demand in 2022, including the corresponding lower and upper bounds. The results indicate that the highest forecasted peak demand value for 2022 occurs in June, reaching 11447.39 MW, while the lowest forecasted value is observed in January, at 10332.58 MW. Additionally, January exhibits the lowest lower peak demand value at 8256.77 MW, whereas June showcases the highest upper peak demand value at 13721.05 MW.

Table 4.12 Peak Demand Forecast for 2023 with 95% Confidence Interval

Month	Forecasted Value	Lower Peak Demand	Upper Peak Demand
January	10579.116421	8093.051589	13065.181252
February	10649.573409	8119.765735	13179.381084
March	10958.933850	8387.830880	13530.036821
April	11375.525956	8765.494420	13985.557498
May	11632.341212	8985.065806	14279.616619
June	11640.539617	8957.708076	14323.371159
July	11471.981994	8755.298660	14188.707330
August	11309.608512	8563.802025	14055.415000
September	11259.636834	8487.742779	14031.530889
October	11239.289890	8442.515839	14036.063741
November	11152.842210	8327.682655	13978.001765
December	10953.982270	8094.266859	13813.697682

European Journal of Computer Science and Information Technology

Vol.11, No.2, pp.37-69, 2023

Print ISSN: 2054-0957 (Print),

Online ISSN: 2054-0965 (Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

Table 4.12 displays the forecasted values for peak demand in 2023, along with the corresponding lower and upper bounds. The results indicate that the highest forecasted peak demand value for 2023 occurs in June, reaching 11640.54 MW, while the lowest forecasted value is observed in January, at 10579.12 MW. Additionally, January exhibits the lowest lower peak demand value at 8093.05 MW, whereas June showcases the highest upper peak demand value at 14323.37 M

Table 4.13 Peak Demand Forecast for 2024 with 95% Confidence Interval

Month	Forecasted Value	Lower Peak Demand	Upper Peak Demand
January	10799.296749	7899.678535	13698.914964
February	10851.501529	7909.199809	13790.803249
March	11149.934798	8167.183473	14132.686123
April	11548.530832	8527.107890	14569.953774
May	11798.181275	8740.093140	14856.269410
June	11808.528949	8714.567083	14902.490814
July	11640.150866	8512.323293	14767.978039
August	11485.962492	8328.253010	14643.671974
September	11433.973947	8249.298712	14618.649181
October	11422.854662	8212.517179	14633.192146
November	11341.050442	8101.787036	14580.313847
December	11148.432042	7875.022758	14421.841314

Table 4.13 presents the forecasted values for peak demand in 2024, including the corresponding lower and upper bounds. The findings indicate that the highest forecasted peak demand value for 2024 occurs in June, reaching 11808.52 MW, while the lowest forecasted value is observed in January, at 10799.30 MW. Furthermore, December displays the lowest lower peak demand value at 7875.02 MW, whereas June exhibits the highest upper peak demand value at 14902.49 MW.

Implications

The findings of the study have significant implications for various sectors and provide valuable contributions to the body of knowledge. The results suggest that there will be an increase in peak demand for electricity consumption in Luzon, Philippines in the coming years. This information is crucial as it highlights the need for proactive measures and adjustments in different sectors to address the impacts of this growing demand.

The industrial sector, encompassing manufacturing and mining activities, plays a vital role in energy consumption. The study's implications suggest that industrial facilities will need to prioritize energy efficiency, explore cleaner technologies, and consider renewable energy options to meet the escalating energy demand. This contributes to the understanding of how industries can adapt and transition towards more sustainable practices.

In the transportation sector, the rising number of electric vehicles poses challenges in terms of production, sale, and infrastructure requirements. The study indicates the necessity for expanding charging infrastructure to support the increasing demand for electric vehicles. This insight contributes to the knowledge base surrounding the development and implementation of electric vehicle infrastructure.

The study also underscores the need for expansion and investments in the energy sector to meet the growing electricity demand. Power companies and utilities may need to construct new power plants or upgrade existing ones, leading to investments in both conventional and renewable energy sources. Additionally, improvements in transmission lines and grid infrastructure become vital. These findings contribute to understanding the implications and actions required to ensure a reliable and sustainable energy supply.

Considering the environmental impact, the study highlights the potential rise in greenhouse gas emissions and environmental consequences associated with increased electrical energy demand. This emphasizes the importance of transitioning towards cleaner energy sources, particularly renewable energy, to mitigate climate change and minimize environmental harm. The study contributes to the understanding of the environmental implications and the role of renewable energy in achieving sustainability goals.

Overall, the results of the study provide valuable insights into the expected increase in peak demand for electricity consumption, prompting necessary adjustments in various sectors. This knowledge contributes to informed decision-making, policy formulation, and the pursuit of sustainable energy practices to address the challenges and implications of growing energy demand.

CONCLUSION

The primary objective of this study was to forecast the future values of peak demand in electricity consumption for Luzon, based on monthly historical data from 2001 to 2020 obtained from the official website of the Philippines Department of Energy (DOE). To achieve this objective, the researchers followed the ARIMA model-building procedure developed by Box and Jenkins.

Through conducting various tests and evaluating models based on the Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE), the best-performing ARIMA model was determined to be ARIMA(11, 1, 12). According to this model, the forecasted peak demand for Luzon by December 2021 is estimated to be 10,497.65 megawatts. Furthermore, the ARIMA(11, 1, 12) model suggests that by the end of 2022, 2023, and 2024, Luzon's peak demand would be approximately 10,738.34 MW, 10,953.98 MW, and 11,148.43 MW per electrical grid, respectively.

The performance of the ARIMA(11, 1, 12) model was deemed satisfactory based on evaluation measures. The MAPE was found to be only 3.639%, indicating a high level of accuracy in the forecasts. Additionally, the RMSE value was observed to be minimal. These results highlight the reliability and precision of the ARIMA(11, 1, 12) model in predicting peak demand for electricity consumption in Luzon.

In conclusion, this study successfully achieved its objective of forecasting Luzon's future peak demand in electricity consumption. The ARIMA(11, 1, 12) model demonstrated strong performance, providing valuable insights into the expected energy demand in the upcoming years. The accuracy and reliability of the model's forecasts contribute to informed decision-making and planning in the energy sector.

Recommendations

Based on the findings of the study, the following recommendations can be made:

Energy Efficiency Initiatives: Given the projected increase in peak demand for electricity consumption in Luzon, it is crucial for various sectors to prioritize energy efficiency measures. Industrial facilities should optimize energy usage, adopt cleaner technologies, and explore renewable energy options. This will not only help meet the growing demand but also reduce energy costs and minimize environmental impact.

Expansion of Charging Infrastructure: With the expected rise in electric vehicle adoption, the transportation sector needs to focus on expanding charging infrastructure. Adequate charging stations should be established to cater to the growing demand for electric vehicles. This will facilitate the transition towards a sustainable transportation system and encourage further adoption of electric vehicles.

Investments in Energy Sector: Power companies and utilities should anticipate the increasing electricity needs and plan for infrastructure development accordingly. This may involve constructing new power plants or expanding existing ones to meet the rising demand. Investments should be made in both conventional and renewable energy sources to ensure a diversified and sustainable energy mix. Additionally, enhancing transmission lines and grid infrastructure will be essential for a robust and reliable energy supply.

Transition to Cleaner Energy Sources: The study highlights the environmental impact of increased electrical energy demand, including greenhouse gas emissions. To address climate change and minimize environmental damage, there should be a strong focus on shifting towards cleaner energy sources. Promoting and investing in renewable energy projects, such as solar, wind, and hydroelectric power, will contribute to reducing carbon emissions and fostering a sustainable energy future.

Continuous Monitoring and Research: As energy demand patterns evolve, it is crucial to continuously monitor and assess peak demand trends. Ongoing research and analysis will enable policymakers, energy planners, and stakeholders to make informed decisions, develop effective strategies, and adapt to changing energy needs. Regular updates to forecasting models, incorporating new data, and refining methodologies will improve the accuracy of future peak demand predictions.

REFERENCES

ABS-CBN News. (2023, May 8). *Power outages hit parts of Luzon as Masinloc plant trips*. Retrieved May 15, 2023, from https://news.abs-cbn.com/business/05/08/23/power-outages-hit-parts-of-luzon-as-masinloc-plant-trips

Andersen, F. M., Baldini, M., Hansen, L. G., & Jensen, C. L. (2017, December). *Households' hourly electricity consumption and peak demand in Denmark. Applied Energy,* 208, 607–619. https://doi.org/10.1016/j.apenergy.2017.09.094

ARIMA Modeling. (n.d.). *CORP-MIDS1* (MDS). https://www.mastersindatascience.org/learning/statistics-data-science/what-is-arima-modeling/

Boland. (n.d.). *Box-Jenkins Models.* https://lo.unisa.edu.au/pluginfile.php/1156111/mod_resource/content/1/Box-Jenkins.pdf

Cabauatan, R., & Tatlonghari, V. M. (2017, September). *Analysing in the Philippines using vector error correction. International Social Science Journal,* 67(225–226), 137–150. https://doi.org/10.1111/issj.12136

Department of Energy. (2021). *2020 POWER SITUATION REPORT*. In 2020_power-situation-report_as_of_09-september-2021.pdf.

Fattah, Ezzine, Aman, Moussami, & Lachhab. (n.d.). *Forecasting of demand using ARIMA model. International Journal of Engineering Business Management.* https://doi.org/10.1177/1847979018808673

Ho, S., & Xie, M. (1998, October). *The use of ARIMA models for reliability forecasting and analysis. Computers & Industrial Engineering*, 35(1–2), 213–216. https://doi.org/10.1016/s0360-8352(98)00066-7

Lena. (2023, February 20). *Power outage hits 5 provinces of Western Visayas. Republic of the Philippines - Philippine News Agency*. Retrieved May 15, 2023, from https://www.pna.gov.ph/articles/1195652.

Ma, L., Hu, C., Lin, R., & Han, Y. (2018, December 20). *ARIMA model forecast based on EViews software. IOP Conference Series: Earth and Environmenta*l Science, 208, 012017. https://doi.org/10.1088/1755-1315/208/1/012017

PAGASA. (n.d.). Retrieved May 15, 2023, from https://www.pagasa.dost.gov.ph/

Pappas, S., Ekonomou, L., Karamousantas, D., Chatzarakis, G., Katsikas, S., & Liatsis, P. (2008, September). *Electricity demand loads modeling using AutoRegressive Moving Average (ARMA) models.* Energy, 33(9), 1353–1360. https://doi.org/10.1016/j.energy.2008.05.008

Philippines Statistics Authority. (2005, August 31). October 2004 Household Survey.

Republic of the Philippines - Philippine Statistics Authority. (n.d.). *Electricity is the most common source of energy used by household.* Retrieved May 15, 2023, from https://psa.gov.ph/.

Sta. Maria, M.V., M. Estoque, M.A., & Villarin, J.V. (1999). *El Nino-La Nina Dynamics over Southeast Asia. Manila Observatory*. Retrieved May 15, 2023, from https://www.observatory.ph/publications/el-nino-la-nina-dynamics-over-southeast-asia/

The Philippine Embassy of Canberra, Australia. (n.d.). Retrieved May 15, 2023, from https://www.philembassy.org.au/the-philippines/map.

U.S. Energy Information Administration. (2022, December 27). Retrieved May 15, 2023, from https://www.eia.gov/energyexplained/what-is-energy/.

Young, W.L. (1977). *The Box-Jenkins approach to time series analysis and forecasting: principles and applications.* Nd. http://www.numdam.org/item/RO_1977__11_2_129_0.pdf