International Journal of Energy Economics and Policy

ISSN: 2146-4553

available at http: www.econjournals.com

International Journal of Energy Economics and Policy, 2023, 13(5), 544-552.

Forecasting the Total Non-coincidental Monthly System Peak Demand in the Philippines: A Comparison of Seasonal Autoregressive Integrated Moving Average Models and Artificial Neural Networks

Samuel John E. Parreño*

Department of Teacher Education, University of Mindanao Digos College, Digos City, Philippines. *Email: samueljohnparr@gmail.com

Received: 25 March 2023 **Accepted:** 16 July 2023 **DOI:** https://doi.org/10.32479/ijeep.14240

ABSTRACT

This paper aims to determine suitable seasonal autoregressive integrated moving average (SARIMA) and feed-forward neural network (FFNN) models to forecast the total non-coincidental monthly system peak demand in the Philippines. To satisfy the stationary requirement of the SARIMA model, seasonal differencing, and first-differencing were applied. The findings reveal that SARIMA (0,1,1) $(0,1,1)_{12}$ is the appropriate SARIMA model. All the model parameters were statistically significant. Also, the residuals were normally distributed. For the feed-forward neural networks, the NNAR $(10,1,6)_{12}$ was found to be the appropriate model. The evaluation statistics indicate that the models developed are suitable for forecasting. A comparison of the models has been performed by examining their respective root mean square error (RMSE), mean absolute error (MAE), and mean absolute percent error (MAPE) values. It was found that the FFNN performs better and is the most suitable model to forecast peak demand.

Keywords: Electricity Peak Demand, Time Series Analysis, SARIMA, Artificial Neural Networks
JEL Classifications: C10, C53, Q47

1. INTRODUCTION

Human development has been centered on mathematics over the years. It became crucial to human civilization in the past and continues to be so in the present. Mathematics plays a very evident and significant part in human activities from the most basic to the most complex ones. To explain and resolve issues and events in the actual world, mathematics has evolved from basic arithmetic to complex concepts, algorithms, and applications. One economic problem that the world face is the scarcity of resources such as food, land, shelter, and energy. With this, the government and concerned authorities should formulate concrete and research-based solutions to address and avoid the worst problems that may happen in the future, particularly in resource distribution.

Resources such as energy, particularly electricity, play an important role in various activities and routines in human life. Electricity, because of its significant role, has become a basic necessity of modern human life. It can also be a key component for a country's economic growth as well as its political and social security. It has various advantages when compared to other forms of energy. It is convenient, clean, reliable, transfer efficient, and easier to control. (Zohuri, 2016; Parreño, 2022a). In previous years, electricity demand and consumption across different countries in the world have been steadily increasing. Its rise is attributed to population and economic growth wherein the increase is particularly strong in developing countries. The growth happens almost annually, and it is estimated that the electricity demand and consumption will increase by 1-2% annually (Ritchie et al., 2022; World Energy Outlook, 2019). The growing reliance on electricity warrants future consumption and demand forecasts. Forecasts are particularly important to power utility holders, power system administrators, energy system operators, and planners. Also, it is beneficial to the government and concerned authorities as it could be used as a scholarly basis for developing policies and strategies that could address issues in energy.

Various electricity generation units were adopted by power plants to meet the particular electricity demand/load categories. Peak load units have the least efficiency and the highest price when compared to the different electricity generation units. According to estimates, a 5-15% decrease in peak load would have significant positive consequences on resource conservation and real-time power price reduction (Sheffrin et al., 2008). Hence, employing efficient peak load management approaches has become a necessity.

To achieve that, power stations should be able to forecast the size and time of occurrence of peak/load demand. This enables power plants to avoid congestion in the grid by giving them ample start-up time. They can serve as the basis for power station operations in making sure that future electricity generations meet future energy demands. In this respect, a 1% decrease in forecast errors can significantly decrease the power generation costs and thus secure its supply (Alfares and Nazeerddin, 2002). Hence, it is clear that the forecasts are crucial in guaranteeing benefits to the economy as well as the grid's security and stability.

The economy of the Philippines has been steadily growing at around six percent per annum. As of 2021, the gross domestic product (GDP) of the country reached 394.09 billion US dollars (Statista, 2023). Previous data revealed that a positive increase in the gross domestic product was directly proportional to energy consumption. Hence, a steady increase in GDP implies a consistent increase in energy demand and consumption (Department of Energy, 2019). The energy market in the Philippines is rapidly evolving due to the rising population and government-initiated infrastructure expansion. However, the country heavily relies on imports to generate electricity because energy resources are scarce. As of 2020, 76% of the total electricity capacity was produced from nonrenewable resources such as coal, natural gas, and oil. The remaining 24% were produced from renewable resources. Eighty-seven percent of the overall coal used by the power plants was imported from China (Philippines Energy Market, 2020; Department of Energy, 2020). There are three electrical grids in the Philippines: one in Luzon, one in the Visayas, and one in Mindanao.

In such a case, it is clear that an accurate peak load demand forecast is crucial to the operation of the power stations. It helps power stations avoid overestimation or underestimation that may result in increased operating costs for the supply-since a large amount of the resource used in electricity generation is imported. This will lead to positive consequences for the economy.

Forecasting methods can be categorized based on the objectives and methodology being applied. It can be classified as short-, medium-, and long-term forecasting depending on the time range. Also, it can be categorized based on the methodology-traditional statistical models or non-statistical AI-based models. Traditional models apply a mathematical combination of past data, wherein parameter estimates of these models are easier to interpret. While non-statistical AI-based models are mostly adaptive and provide high precision and robustness to non-stationary data since the method is nonlinear and nonparametric (Amjady, 2001).

Several studies have been performed to forecast electricity consumption and peak load demand using traditional statistical models. Parreño (2022c) predicted the electricity demand of the Philippines using autoregressive integrated moving average (ARIMA). The data was divided into two portions-the first portion was used in model building, while the other was used in forecast evaluation. In the model selection stage, the model with the smallest Akaike information criterion (AIC) value was chosen. The analysis revealed that the ARIMA (0,2,1) was the most accurate and reliable model. Similarly, Wahid et al. (2020) applied the ARIMA model in forecasting electricity consumption in Pakistan. According to the forecasting results, electricity consumption in Pakistan will increase. Moreover, the study of Kim et al. (2019) utilized the ARIMA model, together with ARIMA-GARCH, multiple seasonal exponential smoothing models, and the artificial neural network to forecast the electricity load in institutional buildings. The RMSE and MAPE were used in comparing the performances of the models.

Other statistical methods that were used in forecasting electricity consumption and peak demand were the exponential smoothing model (Taylor, 2012), support vector machines (Zhang et al., 2016), and grey models (Xu et al., 2017). While examples of non-statistical AI-based models used in forecasting electricity use are long short-term memory (LSTM), the artificial neural network (ANN), and support vector regression (SVR) (Wang et al., 2019).

This study proposes using seasonal autoregressive integrated moving average (SARIMA) and artificial neural networks (ANN), specifically the feed-forward neural network (FFNN) model, to forecast the total non-coincidental monthly peak demand in the Philippines. This study will also compare the performances of the models to determine the most appropriate model. The contributions of this study are described as follows.

1. To the best of the author's knowledge, no study forecasts the total non-coincidental monthly peak demand in the Philippines. The findings may provide valuable insights into the current situation of electricity demand in the country. These insights may serve as a scholarly basis that may be beneficial to the government and appropriate agencies in formulating energy policies.
2. Also, there was no study found that compares the performances of statistical and non-statistical models when forecasting the total non-coincidental monthly electricity peak demand in the country. The performances of the models considered may be used as a benchmark when exploring other models that were not considered in this paper.

2. MATERIALS AND METHODS

2.1. Data Source

The data considered in this study were retrieved from the Department of Energy, Philippines. The monthly peak demands (in megawatts) of Luzon, Visayas, and Mindanao grids were recorded during the period from January 2001 to December 2020. Since this study aims to forecast the total non-coincidental monthly peak demand of the Philippines, the sum of the monthly peak demands of Luzon, Visayas, and Mindanao grids was calculated. The total number of observations in the dataset over the 20-year period is 240. As suggested by Hyndman and Athanasopoulos (2018), the data were divided into two segments: The first 80% or 192 data points were used for model building while the remaining 20% or 48 data points were used for model validation.

2.2. Augmented Dickey-Fuller Test

The Augmented Dickey-Fuller test is a unit root test. The test was implemented to the raw peak demand observations, the log-transformed data, and the differencing. The test permits for a higher-order autoregressive process by including Δy_{t-p} in the Dickey-Fuller model. The test employs the following model:

$$\Delta y_{t-p} = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \ldots + \delta_{p-1} \Delta y_{t-p+1} + \epsilon_t \quad (1)$$

Where α is the constant, β is the time trend coefficient, p is the lag-order of the autoregressive process, and the Δy_{t-p+1} are the differencing terms. In this test, we want to check whether $\gamma \neq 0$. If $\gamma = 0$, then the null hypothesis is not rejected and we have a random walk non-stationary process. Otherwise, if $\gamma \neq 0$ and $-1 < 1 + \gamma < 1$, then we have a stationary process. We want to reject the null hypothesis to apply the autoregressive integrated moving average (ARIMA) model (Holmes et al., 2021).

2.3. Akaike Information Criterion

The Akaike information criterion is a ubiquitous tool in model selection. This tool adds a penalizing term based on the number of variables to the log-likelihood as the fit of the models improves by adding more variables. The AIC value of the model is the following:

$$AIC = 2 \times Q + 2p \quad (2)$$

Where p is the number of unknown parameters of the model Q. The lower the value of the AIC, the better the model fitting (Lord et al., 2021).

2.4. Seasonal Autoregressive Integrated Moving Average Model

The seasonal autoregressive integrated moving average (SARIMA) model can be written as

$$\Phi^* (B^S)\, \Phi (B)\, (1-B)^d\, (1-B^S)^D\, y_t = \delta + \Theta^* (B^S)\, \Theta (B)\, \epsilon_t \quad (3)$$

Where $\Phi (B) = 1-\Phi_1 B-\Phi_2 B^2-\ldots-\Phi_p B^p$ is the autoregressive operator of order p; $\Theta (B) = 1-\theta_1 B-\theta_2 B^2-\ldots-\theta_q B^q$ is the moving average operator of order q; $(1-B)^d$ is the difference d; $\Phi^* (B^S)$ is the seasonal autoregressive operator of order P; $\Theta^* (B^S)$ is the seasonal moving average operator of order Q; $(1-B^S)^D$ is the seasonal difference D; S is the seasonal lag; B is the backshift operator; and ϵ_t is the error term (Montgomery et al., 2011). Here the p, d, q, P, D, and Q are integers and the orders of the model. To build the model, the following three-step iterative procedure was implemented: Model identification, parameter estimation, and diagnostic checking (Box and Jenkins, 1976). And as proposed by authors Hyndman (2001) and Adhikari and Agrawal (2013), another stage was added, which was forecast evaluation.

This paper follows the steps employed by Parreño (2022b) in building the appropriate SARIMA model. It is described as follows:

Stage 1: The dataset is divided into two segments; the first 80% or 192 data points were used for model building while the remaining 20% or 48 data points were used for model validation. The model building set will be used in stages 1 to 3. Check the stationarity of the model building set by using data visualization and applying the augmented Dickey-Fuller (ADF test), as the ARIMA model requires stationary data. Also, identify the seasonality of the historical data. A stationary time series has a constant mean and variance. If the data are non-stationary, transformation is applied. If the data are still non-stationary, differencing is applied and repeated until the data become stationary. The appropriate orders for the autoregressive (p), seasonal autoregressive (P), differencing (d), seasonal differencing (D), moving average (q), and seasonal (Q) parameters for the SARIMA model are identified by inspecting the autocorrelation plot (ACF) and partial autocorrelation plot (PACF).

Stage 2: The steps in stage 1 yield several tentative models, and to choose the best model, the Akaike information criterion (AIC) values of each tentative model are examined. The model with the least AIC value is selected in this stage. Then, the Wald z-test is performed to estimate the parameters of the model chosen. If there exists a coefficient that is not statistically significant, the model will be refitted by choosing the model with the second least AIC value, and the Wald z-test will be performed again.

Stage 3: Test the suitability of the model by visualizing its residuals' ACF and PACF. If the spikes on the ACF and PACF plots are within acceptable limits, then the model is deemed adequate. Moreover, the Ljung-Box test will be utilized to test if the model does not show a lack of fit. If the assumptions are violated that is, the spikes on the ACF and PACF plots are outside the acceptable limits and the Ljung-Box test result shows that the model shows a lack of fit, the model with the second smallest AIC value will be selected and the model parameter estimation (stage 2) process is repeated.

Stage 4: Compute the forecasts from the model produced from stages 1 to 3 and calculate the forecast errors by comparing them to the model validation set. If the ACF, PACF, and the normality of the forecast errors show that they exhibit Gaussian white noise properties, then the model is considered efficient and appropriate.

2.5. Feed-forward Neural Network Model

The feed-forward neural network model is one of the reputable neural network models for forecasting time series. It is based on simple mathematical representations of the human brain. The general form of the model is a black-box type model that can be used to model high-dimensional and nonlinear data. The model has a network of "neurons" which are organized in several layers-can be either the original or the constructed variables. The usual architecture of the neural network includes three layers: (1) the predictors (or inputs) which are the original predictors; (2) the hidden layers which are the set of constructed variables; and the (3) forecasts (or outputs). Figure 1 shows the architecture of the FFNN with a single hidden layer.

The feed-forward neural network involves a linear combination of input variables (Eq. 4) and activation or transfer function (Eq. 5). They are defined as follows,

$$a_i = \sum_{j=1}^{p} w_{1ij} x_j + \theta_i, \quad (4)$$

$$s(z) = \frac{1}{1+e^{-z}} \quad (5)$$

Where w_{1ij} are the weights (unknown parameters being estimated) and θ_i is the bias node.

2.6. Model Accuracy Measures

To measure the performances of the models considered in this paper, the following model accuracy measures were used. These accuracy measures are generally used in peak demand forecasting to show error

$$\text{Root Mean Squared Error} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\overline{y_t} - y_t\right)^2}, \text{ and} \quad (6)$$

$$\text{Mean Absolute Error} = \sum_{t=1}^{n}\left|\frac{y_t - \overline{y_t}}{y_t}\right| \quad (7)$$

$$\text{Mean Absolute Percent Error} = \sum_{t=1}^{n}\left|\frac{y_t - \overline{y_t}}{y_t}\right| \times 100\% \quad (8)$$

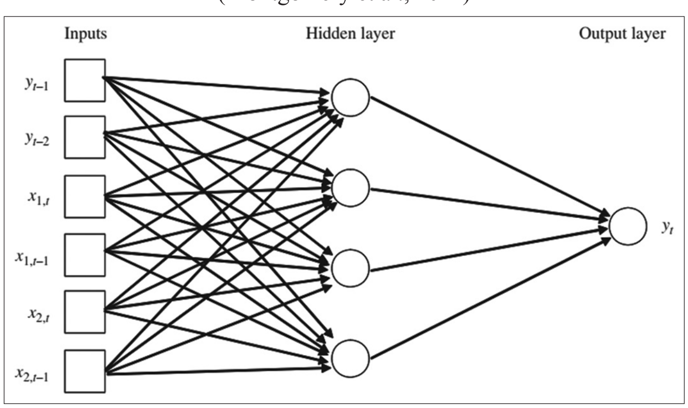
Where y_t is the actual observation and $\overline{y_t}$ is the forecasted peak demand at time *t*.

3. RESULTS AND DISCUSSION

3.1. Data Preliminaries

Table 1 shows a summary of the descriptive statistics for the total non-coincidental monthly system peak demand in the Philippines. The mean, median, minimum, maximum, and standard deviation are reported in Table 1.

Figure 1: An feed-forward neural network with a single hidden layer (Montgomery et al., 2011)



It can be observed from Table 1 that the peak demand has a median of 9679 MW which is not close to the mean of 10147 MW. Also, the minimum monthly peak demand is 6729 MW which occurred on February 2002. And the highest peak demand is 15565 MW which was recorded on June 2019. The standard deviation for the monthly peak demand is 2366.637 which indicates that the data are more spread out. Finally, the kurtosis of 2.149749 indicates that the data are leptokurtic or have observations concentrated about the mean.

The plot of the time series of the monthly peak demand is shown in Figure 2. It can be observed from the figure that the time series shows seasonality with low peak demands occurring in January and high peak demands occurring in May. Also, an upward trend is evident in the figure. In addition, it can be observed that there was a sudden drop in peak demand in April 2020. This unexpected drop can be attributed to the country's policy changes in response to the COVID-19 pandemic. During this period, the government enforced a lockdown and temporary closure of establishments and factories which were the main consumers of electricity in the country.

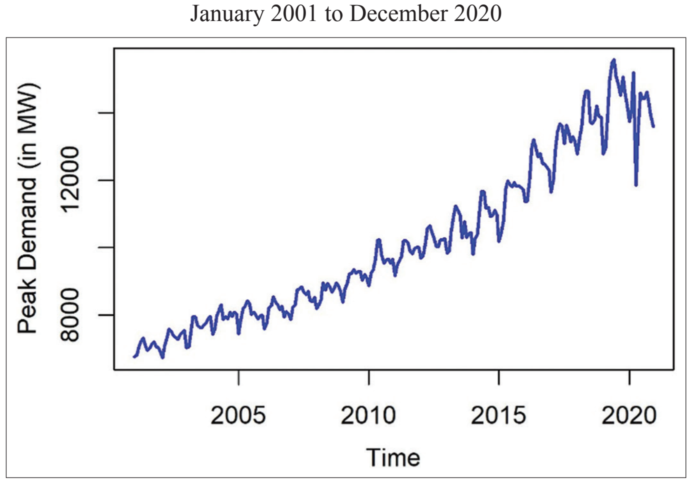
3.2. Forecasting Using Seasonal ARIMA Model

The data were divided into two segments: the first 192 data points, from January 2001 to December 2016, were used as model building set; while the remaining 48 data points, from January 2017 to December 2020, were used as validation set. The Augmented Dickey-Fuller test was applied to the model building set. The result indicates that there is no unit root in the time series ($P=0.03189 < 0.05$), however by inspecting Figure 2 and the ACF plot of the training set in Figure 3, it is clear that the time series is nonstationary. Thus, seasonal differencing was applied first as recommended by Hyndman and Athanasopoulos (2018). The ADF test was applied to the seasonally differenced training set and the result indicates that it does not have a unit root (P=0.0478

Table 1: Descriptive statistics for the total non-coincidental monthly system peak demand in the Philippines from January 2001 to December 2020

Mean	Median	Minimum	Maximum	SD	Kurtosis
10147	9679	6729	15565	2366.637	2.149749

Figure 2: Time Series plot of peak demand in the Philippines from January 2001 to December 2020



< 0.05). However as shown in Figure 4, spikes on the ACF plot are slowly decaying and outside the acceptable limits. Hence, the time series remains nonstationary. Thus, first-differencing was applied. The ADF result of the first-differenced time series reveals that it does not have a unit root (P<0.01). Also, the ACF and PACF plots in Figures 5 and 6, respectively, indicate that the time series has become stationary. Hence, the series was ready for model identification. The ACF and PACF plots have damped sinusoidal behaviors which indicate an ARIMA (p, d, q) model. In addition, the ACF plot displays a sharp cut -off in lag 1 which implies q=1. Also, the PACF plot displays a cut-off at lag 2 which implies p=2. For the seasonal part, the ACF plot shows a spike at lag 12 which indicate Q=1. The PACF plot shows spikes at lag 2 implying p=2. Table 2 presents the candidate SARIMA models with corresponding AIC values.

Figure 3: Autocorrelation plot of the training set

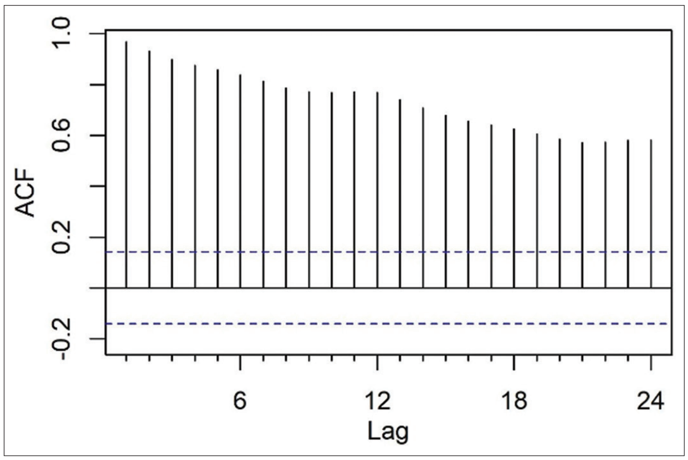


Figure 4: Autocorrelation plot of the seasonally differenced training set

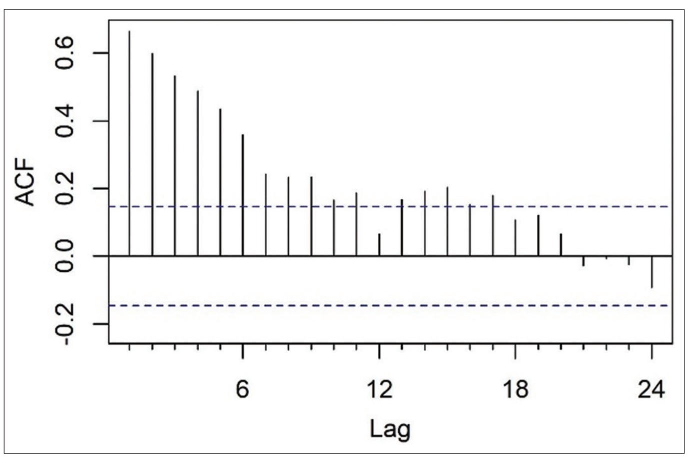
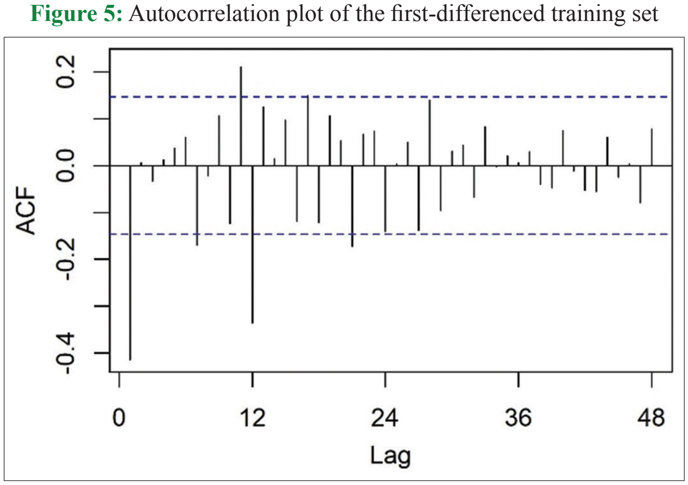


Figure 5: Autocorrelation plot of the first-differenced training set



It can be observed from Table 2 that SARIMA (0,1,1) $(0,1,2)_{12}$ has the least AIC value indicating that the model has a better model fitting. The coefficients of SARIMA (0,1,1) $(0,1,2)_{12}$ were then tested using Wald z-test. The results show that MA(1) is statistically significant (P<0.05), SMA(1) is statistically significant (P<0.05), and SMA(2) is not statistically significant (P=0.08019). Since there exists a statistically not significant coefficient, thus the model with the second least AIC value is selected, SARIMA (0,1,1) $(0,1,1)_{12}$. The Wald z-tests were applied to SARIMA (0,1,1) $(0,1,1)_{12}$. Table 3 displays the results of the Wald z-test and it can be observed that the coefficients are statistically significant. Hence, this model was used in forecasting the monthly system peak demand.

Figure 6: Partial autocorrelation plot of the first-differenced training set

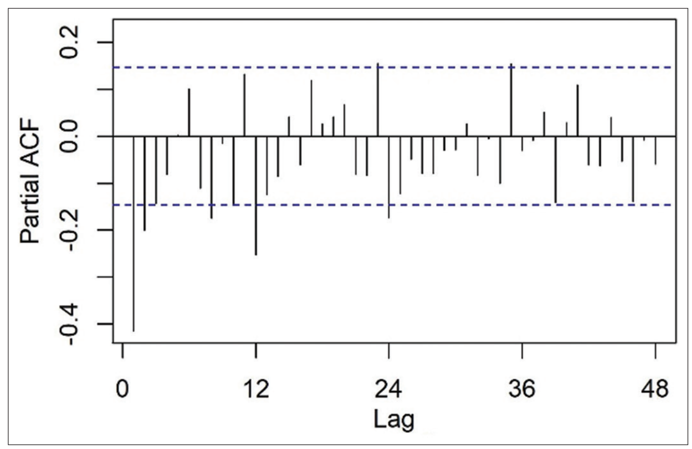


Table 2: Candidate models for monthly system peak demand

Candidate models	AIC
SARIMA $(0,1,1)(0,1,1)_{12}$	2368.617
SARIMA $(0,1,1)(0,1,2)_{12}$	2367.752
SARIMA $(0,1,2)(0,1,1)_{12}$	2369.789
SARIMA $(0,1,2)(0,1,2)_{12}$	2368.649
SARIMA $(0,1,3)(0,1,2)_{12}$	2369.661
SARIMA $(2,1,1)(0,1,1)_{12}$	2370.428
SARIMA $(2,1,1)(1,1,1)_{12}$	2369.933
SARIMA $(2,1,2)(1,1,1)_{12}$	2371.519
SARIMA $(2,1,1)(2,1,1)_{12}$	2368.803

SARIMA: Seasonal autoregressive integrated moving average, AIC: Akaike information criterion

Table 3: Wald z-test results for the coefficients of SARIMA $(0,1,1)(0,1,1)_{12}$

	Estimate	Std. Error	z-value	P-value
MA (1)	−0.502013	0.073433	−6.8363	8.124e-12
SMA (1)	−0.547292	0.076209	−7.1815	6.896e-13

SARIMA: Seasonal autoregressive integrated moving average

The residuals of SARIMA $(0,1,1)(0,1,1)_{12}$ were examined to determine if the model was able to capture the behavior of the training set. Figure 7 displays the residual plots of the model. It can be observed that the residuals behaved like white noise since the mean of the residuals was close to zero. Also, most of the spikes of the ACF plots are within acceptable limits. This implies that there is no significant correlation among the residuals. Further, the histogram suggests that the residuals are slightly normal. These results mean that the model appears to account for all available information. The claim was supported by the Ljung-Box test since the $P > 0.05$ ($P = 0.7119$). This means that the model does not show lack of fit.

The final step is forecast evaluation. A 48-month forecast was produced from SARIMA $(0,1,1)$ $(0,1,1)_{12}$ and was compared to the data validation set. Figure 8 displays the observations from the validation set and the forecasts from SARIMA $(0,1,1)$ $(0,1,1)_{12}$. It is apparent that the actual and forecasted values are relatively close to each other with the exception of the unexpected peak demand in April 2020. Further, the ACF, PACF, and normal Q-Q plots of the forecast errors are shown in Figures 9-11, respectively. The spikes on the ACF and PACF plots in Figures 9 and 10, respectively, are within the acceptable limits, hence there is no significant correlation among the forecast errors. The normal Q-Q plot in Figure 11 suggests that the forecast errors might not be normal because of the presence of an outlier. If we ignore the outlier, the errors may be normal. Overall, we conclude that SARIMA $(0,1,1)$ $(0,1,1)_{12}$ was indeed appropriate to forecast the monthly system peak demand in the Philippines.

Figure 7: Residual plots of seasonal autoregressive integrated moving average $(0,1,1)(0,1,1)_{12}$

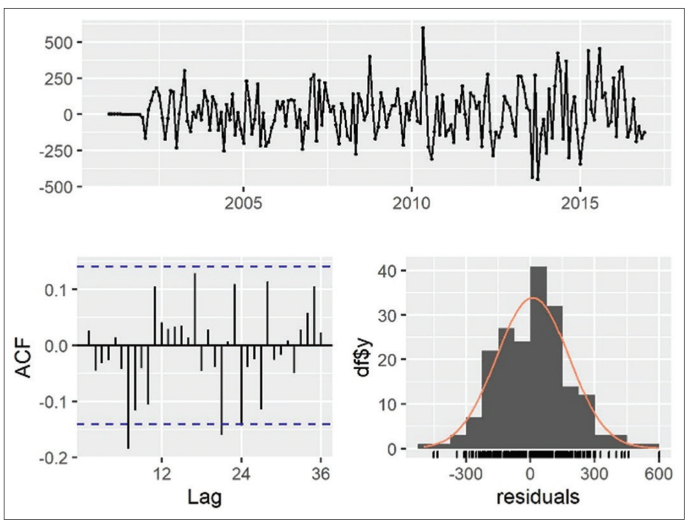


Figure 8: Plot of actual and forecasted values based on seasonal autoregressive integrated moving average $(0,1,1)$ $(0,1,1)_{12}$

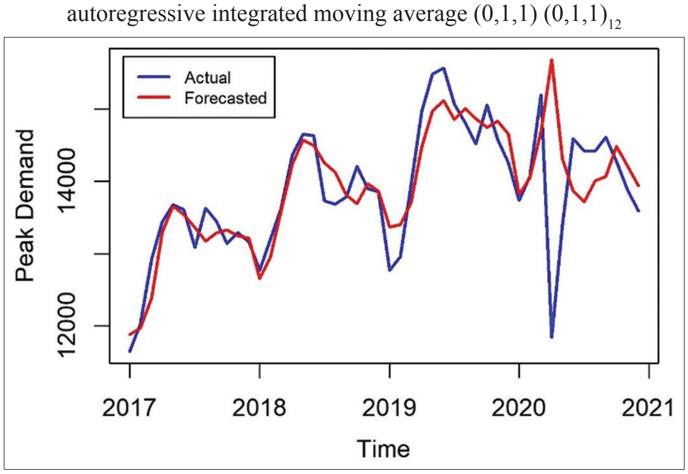


Figure 9: Autocorrelation plot of forecast errors

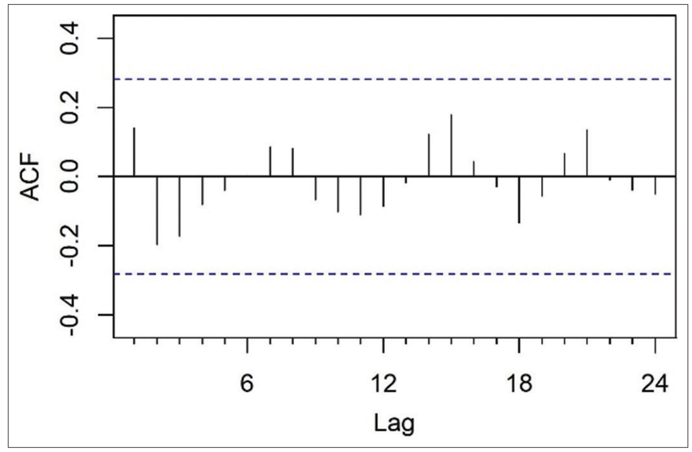


Figure 10: Partial autocorrelation plot of forecast errors

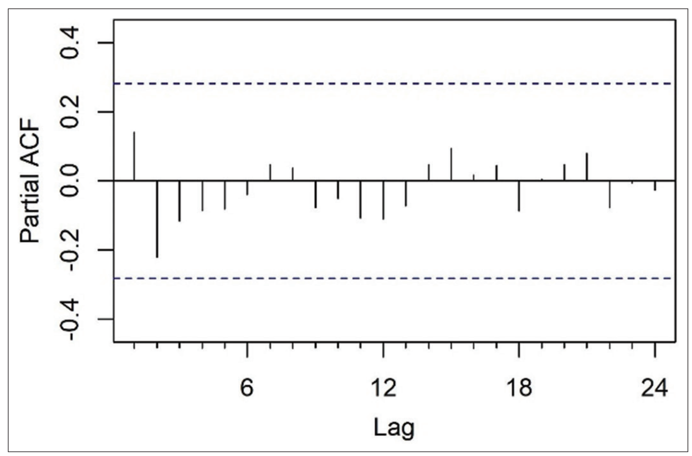
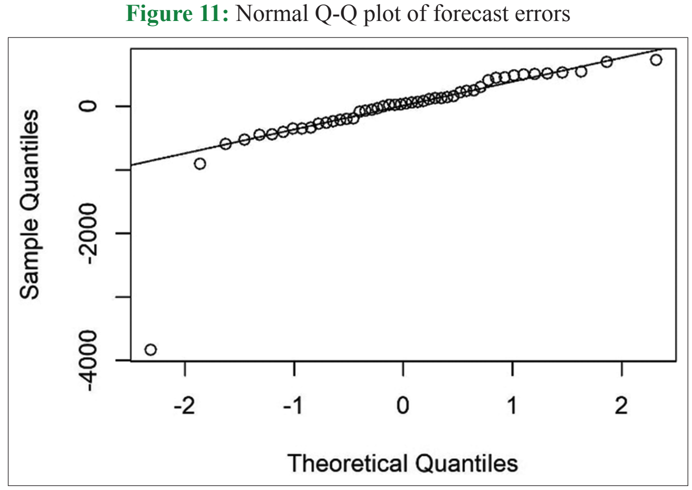


Figure 11: Normal Q-Q plot of forecast errors



3.3. Forecasting Using Feed-forward Neural Network

An FFNN with one hidden layer and with NNAR (p, P, k) notation was applied to the time series wherein p is the number of inputs that are lagged, P is the number of seasonal lags, and k is the number of nodes in the hidden layer. The best model obtained from an average of 20 networks, each of which is an 11-6-1 network with 79 weights options were linear output units, is NNAR $(10,1,6)_{12}$ and has σ^2 estimated as 25883. This results in 10 lags, a single seasonal lag, and 6 hidden layers. This means that the last 10 observations and the last observed value from the same season were used as predictors, and there are 6 neurons in the hidden layer.

The residual plots of NNAR $(10,1,6)_{12}$ are presented in Figure 12. Most of the residuals are close to zero with exception of one large negative residual which is a result of an unexpected drop in April 2020. This can also be seen on the histogram. The histogram suggests that the residuals may be normal if we ignore the outlier. Also, majority of the spikes on the ACF plot are inside the acceptable limits. Thus, if the single outlier is ignored, the residuals may be normal. Consequently, the forecasts produced from the model will be reliable. Figure 13 presents the actual and forecasted plot of the monthly peak demand. It is evident that the forecasted plot closely follows the pattern of the actual plot.

Figure 12: Residual plots of NNAR $(10,1,6)_{12}$

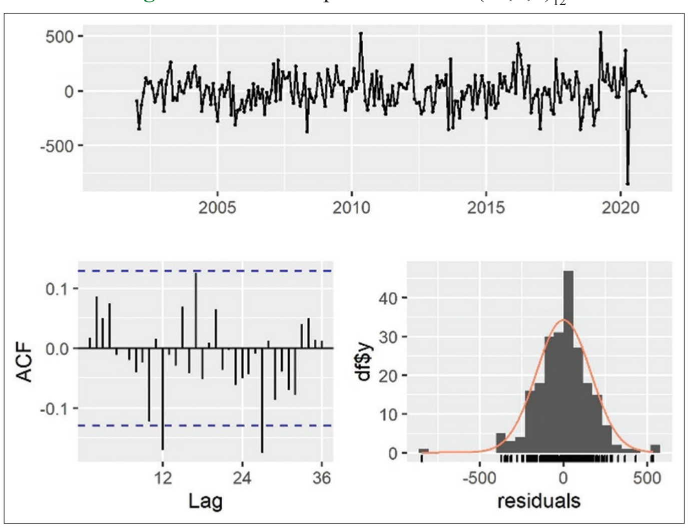
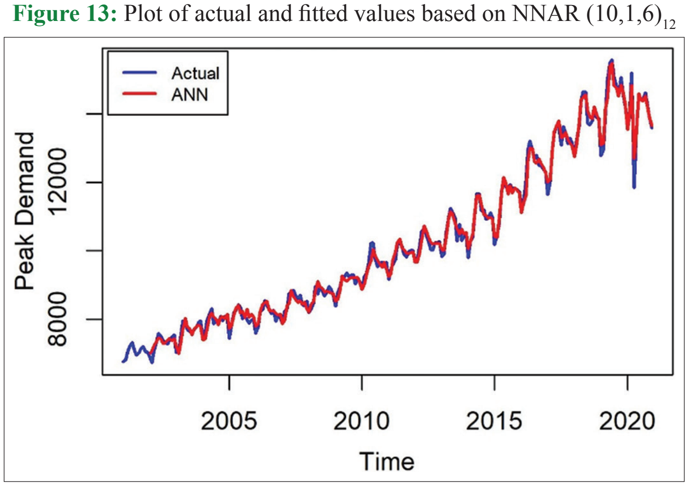


Figure 13: Plot of actual and fitted values based on NNAR $(10,1,6)_{12}$



3.4. Forecasts and Comparison of Accuracy Measures

Figures 14 and 15 present the plot of forecasts from SARIMA $(0,1,1)$ $(0,1,1)_{12}$ and NNAR $(10,1,6)_{12}$, respectively. It can be observed that the predicted values from SARIMA $(0,1,1)$ $(0,1,1)_{12}$ do not exceed the maximum value and as time increases, the confidence intervals grow wider. On the contrary, the predicted values from NNAR $(10,1,6)_{12}$ eventually exceed the maximum value. The highest predicted value based on the NNAR model will occur on March 2023 which is a compelling prediction as based on the historical data, the peak demand is at its highest during April or May. It can also be inferred that the FFNN was able to capture the positive trend as its values continue to increase each year. Table 4 shows the predicted values based on SARIMA (0,1,1,) $(0,1,1)_{12}$ and NNAR $(10,1,6)_{12}$.

The predictive capabilities of the two models are compared by the measures RMSE, MAE, and MAPE. It can be observed from Table 5 that the RMSE, MAE and MAPE values of the FFNN model are much lower. Thus, the FFNN manifested to be a better

Figure 14: Plot of forecasts from seasonal autoregressive integrated moving average $(0,1,1)$ $(0,1,1)_{12}$

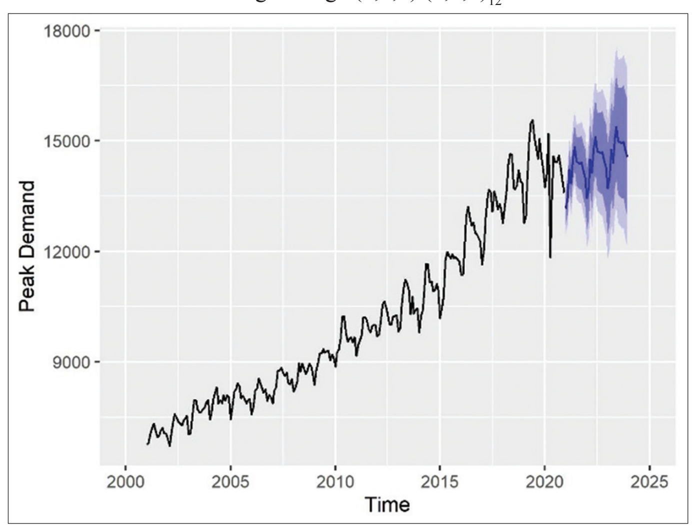


Figure 15: Plot of forecasts from NNAR $(10,1,6)_{12}$

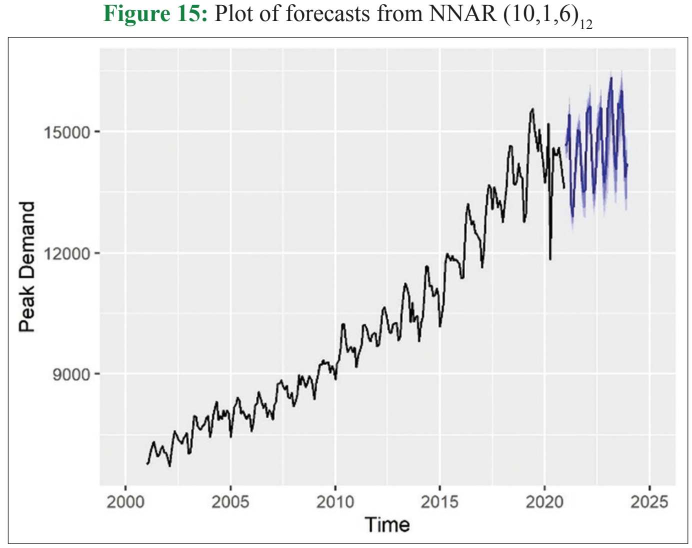


Table 4: Forecasted values of the monthly non-coincidental peak demand from January 2021 to December 2023 based on SARIMA and FFNN

Time	SARIMA	FFNN
January 2021	13165.34	14647.91
February 2021	13447.86	14755.73
March 2021	14219.65	15419.81
April 2021	13852.03	13178.47
May 2021	14529.35	12905.2
June 2021	14826.49	13818.63
July 2021	14435.28	14557.17
August 2021	14415.06	15040.53
September 2021	14382.63	15002.02
October 2021	14416.41	13869.78
November 2021	14189.22	13516.02
December 2021	14009.2	13549.59
January 2022	13443.49	15469.19
February 2022	13726.02	15577.41
March 2022	14497.81	15631.08
April 2022	14130.18	14272.27
May 2022	14807.5	13481.44
June 2022	15104.64	13792.04
July 2022	14713.43	15048.86
August 2022	14693.21	15153.67
September 2022	14660.78	15586.05
October 2022	14694.56	14396.1
November 2022	14467.38	13743.78
December 2022	14287.35	14151.34
January 2023	13721.64	15594.87
February 2023	14004.17	15974.59
March 2023	14775.96	16325.98
April 2023	14408.34	15163.39
May 2023	15085.65	14170.09
June 2023	15382.79	14078.19
July 2023	14991.58	15688.03
August 2023	14971.37	15595.12
September 2023	14938.93	16015.36
October 2023	14972.72	14908.91
November 2023	14745.53	13880.09
December 2023	14565.51	14211.93

SARIMA: Seasonal autoregressive integrated moving average, FFNN: Feed-forward neural network

Table 5: Accuracy measures for the fitted models

Model	RMSE	MAE	MAPE (%)
SARIMA $(0,1,1)(0,1,1)_{12}$	332.9824	179.703	1.66724
NNAR $(10,1,6)_{12}$	160.8824	118.9001	1.206689

SARIMA: Seasonal autoregressive integrated moving average, RMSE: Respective root mean square error, MAE: Mean absolute error, MAPE: Mean absolute percent error

model for forecasting the total non-coincidental monthly peak demand in the Philippines.

4. CONCLUSION

This study has estimated the total non-coincidental monthly peak demand in the Philippines. Based on the study, the peak demand is seasonal and is increasing. It was also observed that the policy changes and lockdowns made by the government in response to the COVID-19 pandemic may have affected the peak demand. It was seen that there was a sudden drop in peak demand in April 2020. These observations imply that the time series is nonstationary. Thus, to apply the SARIMA model to the time series, differencing must be applied. After the time series was stationary with seasonal difference $D = 1$ and first difference $d = 1$, the seasonal ARIMA model was constructed. It was found that SARIMA $(0,1,1)(0,1,1)_{12}$ is the most suitable SARIMA model to forecast the total non-coincidental monthly peak demand since it has significant coefficients, acceptable residuals, and normal forecast errors. For the artificial neural network, specifically the feed-forward neural network (FFNN), an average of 20 networks were produced to obtain the best model. The resulting best model found was NNAR $(10,1,6)_{12}$. The residuals of the neural network were analyzed and found that the model was able to capture the behaviors and patterns of the original time series. Hence, NNAR $(10,1,6)_{12}$ is appropriate to forecast the total non-coincidental monthly peak demand.

When comparing the performances of the SARIMA and FFNN, it was found that the FFNN outperforms the SARIMA model. The FFNN has better RMSE, MAE, and MAPE than the SARIMA. Also based on the forecasts, the FFNN was able to capture the positive trend of the historical time series. Therefore, we conclude that the FFNN is more efficient and accurate than the SARIMA model in forecasting the total non-coincidental monthly peak demand in the Philippines.

The forecasts of the SARIMA model closely follow the seasonal patterns of the historical time series. It is also observed that the values do not exceed the maximum value found in the historical time series. Also, the forecasts of the FFNN closely follow the seasonal patterns but the values do exceed the maximum data of the historical time series. It can also be observed from the predictions of the FFNN that the highest peak demand will occur every March, with the highest to occur on March 2023. It is noteworthy to mention that based on the historical time series, the highest peak demand is usually recorded in April or May. Therefore, electric grids should not only anticipate a high peak demand in April or May, but also in March. These forecasts may serve as a basis for electric grids to make necessary adjustments and decisions to avoid power outages. In addition, these results may be beneficial to policymakers in formulating decisions and appropriate strategic plans regarding the electricity markets. Future works could examine and compare the performances of other models that were not considered in this paper such as exponential smoothing models, Naïve Bayes, and mixed models.

5. ACKNOWLEDGMENT

The author would like to acknowledge the support of the University of Mindanao for funding the publication of this paper.

REFERENCES

Adhikari, R., Agrawal, R. (2013), An Introductory Study on Time Series Modeling and Forecasting. Germany: LAP Lambert Academic Publishing.

Alfares, H., Nazeeruddin, M. (2002), Electric load forecasting: Literature survey and classification of methods. International Journal of Systems Science, 33(1), 23-34.

Amjady, N. (2001), Short-term hourly load forecasting using time-series modeling with peak load estimation capability. IEEE Transactions on Power Systems, 16(3), 498-505.

Box, G., & Jenkins, G. (1976), Time Series Analysis, Forecasting and Control, San Francisco, CA: Holden-Day.

Department of Energy. (2019), 2019 Power Situation Report. Available from: https://www.doe.gov.ph/sites/default/files/pdf/electric_power/2019-power-situation-report.pdf

Department of Energy. (2020), 2020 Power Situation Report. Available from: https://www.doe.gov.ph/sites/default/files/pdf/electric_power/2020_power-situation report_as_of_09-september-2021.pdf?withshield=2

Holmes, E., Scheuerell, M., Ward, E. (2021), Applied Time Series Analysis for Fisheries and Environmental Data. Edition 2021. Washington: University of Washington.

Hyndman, R. (2001), Box-Jenkins modeling. In: Regional Symposium on Environment and Natural Resources.

Hyndman, R., Athanasopoulos, G. (2018), Forecasting: Principles and Practice. 2nd ed. Melbourne, Australia: OTexts. Available from: https://OTexts.com/fpp2

Kim, Y., Son, H., Kim, S. (2019), Short term electricity load forecasting for institutional buildings. Energy Reports, 5, 1270-1280.

Lord, D., Qin, X., Geedipally, S. (2012) Fundamentals and data collection. In: Highway Safety Analytics and Modeling. Ch. 2. Netherlands: Elsevier. p17-57.

Montgomery, D., Jennings, C., Kulachi, M. (2011), Introduction to Time Series Analysis and Forecasting. 2nd ed. United States: Wiley.

Parreño, S. (2022a), Analysis and forecasting of electricity demand in Davao del Sur, Philippines. International Journal on Soft Computing, Artificial Intelligence and Applications, 11(1/2), 25-33.

Parreño, S. (2022b), Application of time series analysis in forecasting coal production and consumption in the Philippines. ICTACT Journal on Soft Computing, 13(1), 2798-2804.

Parreño, S. (2022c), Forecasting electricity consumption in the Philippines using ARIMA models. International Journal of Machine Learning and Computing, 12(6), 279-285.

Philippines Energy Market. (2020), Available from: https://www.trade.gov/marketintelligence/philippines-energy-market

Ritchie, H., Roser, M., Rosado, P. (2022), Energy. Available from: https://ourworldindata.org/energy [Online Resource]

Sheffrin, A., Yoshimura, H., LaPlante, D., Neenan, B. (2008), Harnessing the power of demand. The Electricity Journal, 21, 39-50.

Statista. (2023), Gross Domestic Product (GDP) in the Philippines 2027. Available from: https://www.statista.com/statistics/578709/gross-domestic-product-gdp-in-philippines

Taylor, J. (2012), Short-term load forecasting with exponentially weighted methods. Power Systems, IEEE Transactions, 27(1), 458-464

Wahid, F., Ullah, H., Ali, S., Jan, S.A., Ali, A., Khan, A., Khan, I.A., Bibi, M. (2020), The determinants and forecasting of electricity consumption in Pakistan. International Journal of Energy Economics and Policy, 11(1), 241-248.

Wang, W., Hong, T., Xu, X., Chen, J., Liu, Z., Xu, N. (2019), Forecasting district-scale energy dynamics through integrating building network and long short-term memory learning algorithm. Applied Energy, 15, 217-230.

World Energy Outlook. (2019), Available from: https://www.iea.org/reports/world-energy-outlook-2019

Xu, N., Dang, Y., Gong, Y. (2017), Novel grey prediction model with nonlinear optimized time response method for forecasting of electricity consumption in China. Energy, 118, 473-480.

Zhang, F., Deb, C., Lee, S., Yang, J., Shah, K. (2016), Time series forecasting for building energy consumption using weighted support vector regression with differential evolution optimization technique. Energy and Buildings, 126(15), 94-103.

Zohuri, B. (2016), Application of Compact Heat Exchangers for Combined Cycle Driven Efficiency in Next Generation Nuclear Power Plants. New York: Springer.