Forecasting Electricity Consumption in the Philippines Using ARIMA Models

Samuel John E. Parreño

Abstract—**The electricity demand has been steadily increasing throughout the years. A robust predictive model is required to prepare for future electricity consumption. This paper applied the ARIMA models to forecast electricity consumption in the Philippines. Dataset used was retrieved from the Philippine Institute for Development Studies website. It contains 48 data points, of which 43 were used in model building, and the remaining 5 data points were used in forecast evaluation. The order of the ARIMA (*p,d,q*) model was based on the ACF and PACF plots. The model with the most negative AIC value was chosen among the candidate models, and the best-fitting model was identified. Based on the analysis results, ARIMA (0,2,1) is the statistically appropriate model to forecast electricity consumption in the Philippines. It is predicted that by 2030, the Philippines will consume 163,639.9 GWh of electricity. The R statistical software was used to do all of the calculations.**

Index Terms—**ARIMA, electricity consumption, time series forecasting.**

I. Introduction

One of the most critical energy sources is electricity. It is considered an essential part of modern life. People use it daily in different activities, routines, and work [1]. In the past years, due to increased population and economic growth, electricity consumption across many countries in the world increased continuously. The growth occurred almost annually during the last 50 years, except in the early 1980s and 2009 following the global financial crises. It is estimated that global electricity consumption grows around 1% to 2% per year [2]. The growth is notably strong in countries with developing economies [3].

The Philippine economy has been growing steadily at around 6% every year. In 2020, the country's gross domestic product (GDP) reached about 361.49 billion U.S. dollars [4]. Historical data in the Philippines has shown that expanding economy or a positive GDP growth rate was directly proportional to electricity consumption. Therefore, continuous GDP growth entails a consistently rising demand and electricity consumption [5].

The growing reliance on technological and electrical equipment necessitates future demand forecasts [6]. Various industries are implementing information technology to improve their operations. They provide appropriate storage units to keep up-to-date information and apply different approaches to use it to the greatest extent possible [7]. The forecasts can serve as a basis in designing power and grid distribution systems. In addition, electricity demand forecasts can help determine the number of resources needed to generate the forecasted electricity demand. Decisions based on accurate forecasts can help avoid power shortages and forced outages, leading to productivity and economic consequences [8].

The ARIMA models have been widely applied in forecasting different areas, including economies, stock market, marketing, industrial production, and social processes. It is a statistical analysis model appropriate for short-term forecasting that requires a minimum of 40 historical data point values. It is more efficient, robust, and accurate than traditional forecasting techniques [9], [10]. A study conducted in 2019 has found that ARIMA (1,2,1) could forecast the allocation of the electricity consumption of residential, commercial, and industrial use in the Philippines. Only 15 data points, from the years 2003 to 2017, were used to build the model. Since the data was insufficient to allow out-of-sample testing, the appropriate model was chosen using AIC values, MAPE, and MASE [11]. The small dataset used and the lack of out-of-sample testing in the previous study raise the issue of accuracy and efficiency of the model.

A large dataset is required to develop an accurate prediction model. A model built from a small dataset may be considered unreliable, and the predictions produced would be fragile. A large dataset allows some data points to be withheld for testing [12], [13].

This paper aims to develop a model that will accurately forecast electricity consumption in the Philippines. This paper used the Box-Jenkins method to forecast the electricity consumption for 2021 to 2030 using the historical data on electricity consumption from 1973 to 2020. The first 43 data points, from 1973 to 2015, have been used for model building, and the remaining 5 data points, 2016 to 2020, have been used for forecast evaluation. Candidate models were identified and observed, and the best model was selected.

II. Review of Related Literature

The Philippines electricity market has a rapidly developing structure because of the increasing population and government-initiated infrastructure boom. However, energy sources are scarce, and the country heavily relies on import sources for its electricity generation. In 2020, 76% or 17 gigawatts (GW) of the country's total electricity capacity was generated by coal, natural gas, and oil. Geothermal and other renewable sources generated the remaining 6 GW. Out of the total coal consumed by power plants, 87% or 25,921,662 metric tons mainly were imported from China. A reduction of about 4% or 4,286 gigawatt-hours (GWh) in power generation was observed from 2019 to 2020. The decrease was caused by lower demand, which can be significantly

Manuscript received February 7, 2022; revised April 9, 2022.
S. J. E. Parreño is with the University of Mindanao Digos College, Philippines (e-mail: samueljohnparr@gmail.com).

doi: 10.18178/ijmlc.2022.12.6.1112

attributed to the effect of the COVID-19 pandemic. The residential sector has the largest electricity consumption with a share of 33.7%, followed by the industrial (25.1%) and the commercial (20.4%) sectors [14]-[16].

The nonlinear, time-dependent, and unpredictable features of electricity data make the changes in consumption challenging to predict [17]. The non-storability of electricity makes it more volatile than other commodities. This indicates that the supply must always match the predicted values to meet the demands. Unforeseen events like power outages may result in capacity shortages [18]. Generating electricity requires other resources, including fossil fuels, nuclear energy, and other renewable sources like geothermal. The number of generation units may also affect the electricity supply.

Short-term electricity demand forecasting methods that were commonly used include the following: time series [19], regression analysis [20], support vector regression [21], neural network [22], naïve Bayes [23], fuzzy logic [24], and wavelet echo state network [25]. Each of these methods has a unique scenario wherein it is applicable. The following studies were selected from the literature to present the different methods used to forecast electricity demand. By using regression analysis, Guzman and Rey have forecasted the residential power consumption of Bogotá, Colombia. The model used includes the economic interpretation of coefficients [26]. Regression analysis can be used for prediction and explicating purposes, as it can explain which combination of variables is strongly associated with the outcome variable. However, it is not realistic to include all the potential predictors [27]. By using support vector regression with chaotic genetic algorithm, Hong *et al.* estimated the cyclic electrical load in Northeast China [21]. The SVR provides a unique solution, unlike other methods which can have numerous solutions and may not be robust over distinct samples. However, selecting an appropriate kernel function may be challenging [28]. By using an artificial neural network, Chae *et al.* have predicted the daily peak and sub-hourly electricity consumption of a commercial building [22]. The ANN can detect all possible relationships among the predictors. However, the method requires more extended computation, leading to overfitting [29]. By using a naïve Bayes classifier, Lin *et al.* estimated the energy consumption of heating ventilation and air conditioning (HVAC) [30]. The algorithm is flexible and produces reliable results. However, a large dataset is required to achieve reliable results [31]. By using fuzzy logic, Lakshmipriya and Enigo carried out a short-term forecast of electricity consumption in India [24]. One of the advantages of this method is the ease with which mathematical concepts can be applied to its fuzzy logic. However, developing rules and membership functions may be time-consuming [32]. By using a wavelet echo state network, Deihimi *et al.* have predicted the short-term electric load of a North-American electric utility under the effect of some exogenous variables [25]. One advantage of ESN is that the training phase computations have lower complexity than feed-forward or recurrent ANNs [33].

The Box-Jenkins model, or the ARIMA model, has been the most popular method for predicting and forecasting time series data. The model has been applied across different fields. The model has the following classifications: the moving average (MA), autoregressive (AR), and autoregressive integral moving average (ARIMA) [34]. The ARIMA model was applied in predicting electricity consumption in China, and it showed that ARIMA (1,1,1) was the appropriate model. The annual electricity consumption data from 2000 to 2012 that was used in the study was obtained from the Chinese statistical yearbook [35]. A similar study for forecasting energy consumption of Guandong province in China yielded the same model. Findings showed that the ARIMA (1,1,1) was the optimum model for predicting electricity consumption. It has high precision and stable predictions. The study used artificially generated and actual electricity consumption data sets [36].

Further, ARIMA and ARX models were used to model the electricity demand of faculties in a university in Thailand. The ARIMA model was used to model the demand. The Autoregressive with exogenous output (ARX) was applied to the number of students in the period. For predicting the electricity demand of two faculties, the ARIMA (0,1,1) model was shown to be the best fit. ARIMA (0,1,1), ARIMA $(0,1,1)(0,1,1)_{12}$, ARIMA $(0,1,1)(1,0,0)_{12}$, ARIMA $(0,0,1)(1,1,0)_{12}$, and ARIMA $(1,0,0)(1,0,0)_{12}$ were found to be the appropriate models for predicting electricity demand of other faculties [37]. Furthermore, ARIMA models were used to predict Ghana's electricity consumption. The study found that ARIMA (0,1,0) was the statistically appropriate model. It predicted that by 2030, Ghana would consume 9.5597 billion kWh of electricity [38].

Finally, using the Box-Jenkins ARIMA model, a study estimated the electricity demand in Zimbabwe up to 2025. The study found that ARIMA (1,1,6) was the most stable and best-fitting model for projecting electricity demand. The study estimated that the electricity demand in Zimbabwe would decrease [39].

III. Methodology

A. Source of Data

The dataset used in this paper was retrieved from the Philippine Institute for Development Studies website. The annual electricity consumption from 1973 to 2020 was in gigawatt-hours (GWh). The first 43 data points, from 1973 to 2015, have been used as a training set for model building. The remaining 5 data points, 2016 to 2020, have been used for forecast evaluation.

B. Methodology Used

The Box-Jenkins method was applied to analyze and predict the annual electricity consumption in the Philippines. Model-identification, parameter estimation, and diagnostic checking were used in the three-step iterative procedure in building the ARIMA model [40]. An additional stage was included, as proposed by the authors, which was the forecast evaluation [41].

The model-identification stage involves checking the stationarity of the variables, identifying seasonality in the time series (applying transformation and differencing if required), and visualizing the stationary time series with the autocorrelation (ACF) and partial autocorrelation (PACF) functions to identify the autoregressive (AR) and moving average (MA) components that would be included in the model.

In the model estimation stage, nonlinear least squares and the maximum likelihood estimation were used in estimating the model parameters. Generally, the maximum likelihood estimation is preferred. The nonlinear function maximization was used in solving the maximum likelihood equation, and backcasting was applied to find the initial residual estimates.

Diagnostic model checking would be the next step after model fitting. It is helpful to visualize the residuals' autocorrelation and partial autocorrelation functions to check if they show any patterns (large correlation values). The model was considered adequate if the spikes in the plots of autocorrelation and partial autocorrelation functions were within the acceptable limits. The forecasts could then be generated. If the spikes were large and outside the acceptable limits, the p and q values would be adjusted, and the model would be re-estimated. In this study, the Ljung-Box test was used to test if the time series model does not show a lack of fit.

The final stage was forecast evaluation. It involved computing the forecasts and calculating the forecast errors. The forecast errors would be used to assess the forecasting efficiency of the model. The ACF, PACF, and the normality of the forecast errors would be assessed. If the forecast errors exhibit properties of Gaussian white noise, the model would be considered efficient and useful.

C. Statistical Treatments

The ARIMA model is divided into three parts: 1) the autoregressive (AR) part uses past values of the data series as input of the linear regression equation to predict future values, 2) the integrated part shows how many differences are needed to make the data series stationary, and 3) the moving average (MA) part uses past prediction errors in a regression-like model to predict future values of the data series. There are four processes types: AR (p), MA (q), ARMA (p,q), and ARIMA (p,d,q). The general ARIMA model is given by

$$\Phi(B)(W_t - \mu) = \theta(B)\epsilon_t \tag{1}$$

where W_t is the response series difference, μ is the mean, ε_t is the random error, $\Phi(B) = 1 - \Phi_1 B - \dots - \Phi_p B^p$ and $\theta(B) = 1 - \theta_1 B - \dots - \theta_p B^q$ [40].

The Autocorrelation Function (ACF) plot was used to determine the stationarity of the series. The autocorrelation function plot of a stationary series will drop quickly to zero while the non-stationary series decays exponentially. Also, ACF helps identify the order of a moving average (MA) model. Given a time series Y_t, the sample ACF at lag k is given by

$$r_k = \frac{E[(Y_t-\mu)(Y_{t+k}-\mu)]}{E[(Y_t-\mu)^2]} \tag{2}$$

where E is the expected value operator [40].

The Partial Autocorrelation Function (PACF) was used in identifying the order of the autoregressive (AR) component. Given a time series y_t, the PACF of lag k, is the partial correlation between y_t and y_{t-k}. It is given by

$$PACF = (y_t, y_{t-k}) = \frac{Cov(y_t, y_{t-k} | y_{t-1}, \dots, y_{t-k+1})}{\sigma_{y_t | y_{t-1}, \dots, y_{t-k+1}} \sigma_{y_{t-k} | y_{t-1}, \dots, y_{t-k+1}}} \tag{3}$$

The Augmented Dickey-Fuller (ADF) test is a unit root test. However, interpretations can be translated to a test of stationarity of the time series. The test's null hypothesis is that the data is non-stationary i.e., a unit root is present. The test's alternative hypothesis is that the data is stationary. The ADF test statistic is

$$DF = \frac{\gamma}{SE(\gamma)} \tag{4}$$

where γ is the least square estimate and $SE(\gamma)$ is the standard error.

The Box-Cox Transformation was utilized if the variance of the series is non-constant or if the residuals of the model exhibit lack of fit. An exponent lambda (λ) is at the core of the Box-Cox Transformation. The optimal value of lambda (λ) will be selected and used. Given a time series Y_t, the formula is

$$W_t = \frac{Y_t^{\lambda} - 1}{\lambda} \tag{5}$$

if λ is not equal to zero. If $\lambda = 0$, then we have,

$$W_t = \log Y_t \tag{6}$$

The Shapiro-Wilk test is used to determine if the given random sample comes from a Gaussian distribution. The test's null hypothesis is that the data is a Gaussian distribution, while the alternative hypothesis is that the data is not a Gaussian distribution. The test gives a W value. Small W is evidence of departure from normality. The test statistic of the test is given by

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_i\right)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{7}$$

where the x_is are the ordered sample values and the a_is are the constants from covariances, variances, and order statistics of a sample size n from a Gaussian distribution.

The Ljung-Box Q-test was used in checking the model residuals assumptions. It is given by

$$Q = n(n+2) \sum_{k=1}^{h} \frac{r_k^2}{n-k} \tag{8}$$

where h is the maximum lag, r_k is the autocorrelation at lag k, and n is the number of observations.

IV. Results and Discussions

The electricity consumption in the Philippines during the period 1973 to 2020 is depicted in Fig. 1. It can be observed that there was a 4% decrease in electricity consumption from 2019 to 2020. The decrease in consumption can be attributed to the changes brought by the COVID-19 pandemic, wherein changes in establishments and offices were imposed to prevent and manage the spread of the virus (e.g., ventilation adjustments). Moreover, some establishments were forced to close due to the pandemic temporarily.

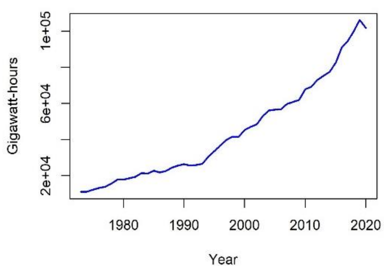

Fig. 1. Time series plot of electricity consumption in the Philippines.

Data from 1973 to 2015 was used for model building. It can be observed from Fig. 2 that the series has an upward trend. The presence of the trend made the data non-stationary, which can be confirmed by the ADF test (p=0.9824). Moreover, exponential growth is observed; transformation and differencing are needed to stabilize the series.

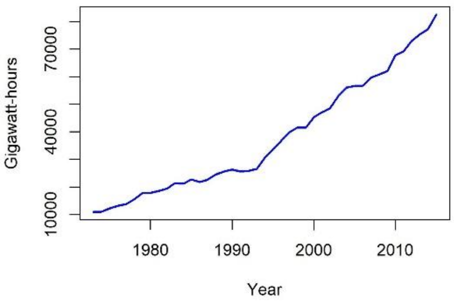

Fig. 2. Time series plot of electricity consumption of the Philippines from 1973 to 2015.

A. Model-Identification

Box-Cox transformation was implemented to make the variance of the time series stable. Fig. 3 depicts the Box-Cox transformed series plot with λ=0. It can be observed that the series remains non-stationary as mean levels in the series are changing over time. Hence, first differencing is applied to remove the linear trend.

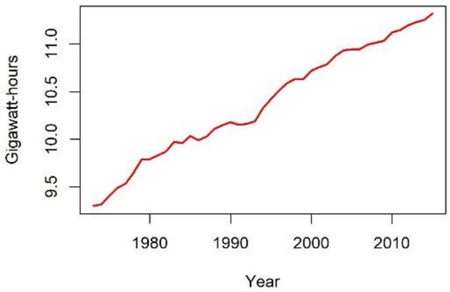

Fig. 3. Time series plot of the log-transformed series.

The plot of the time series after first-differencing is shown in Fig. 4. It can be observed from the plot that after using the method of differencing, the series has been stabilized. However, the ADF test is used to confirm stationarity, and the results show that it is still non-stationary (p=0.2757). Thus, there is a need for second-differencing.

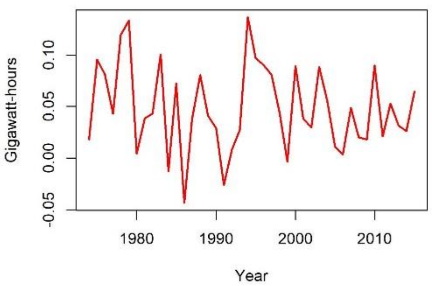

Fig. 4. Time series plot of the first-differenced series.

The plot of the second-differenced series is depicted in Fig. 5. To confirm the stationarity, the ADF test is used. The p-value is 0.01 and is less than α=0.05. Thus, the series is considered stationary. Then, identifying and listing the candidate models would be the next step.

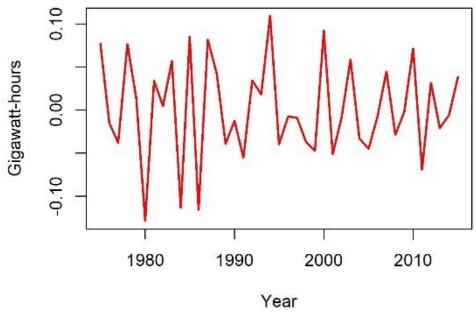

Fig. 5. Time series plot of the second-differenced series.

The candidate models were chosen using ACF and PACF plots [41].

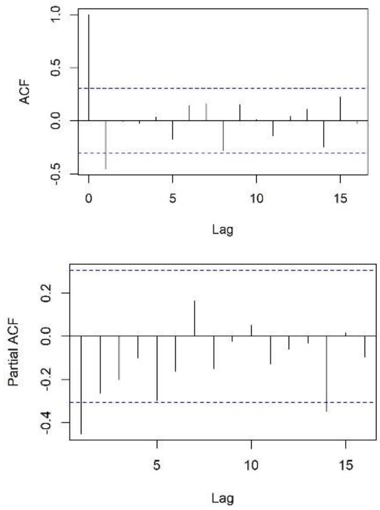

Fig. 6. ACF and PACF plots of the differenced series.

Fig. 6 presents the differenced series's autocorrelation and partial autocorrelation function plots. It can be observed that the autocorrelation function cuts off at lag 1, and the partial autocorrelation function tails off. Hence, we have q=1 and p=0. Moreover, we considered p=0,1,2, q=0,1,2 and d=2. The Akaike Information Criterion (AIC) is used to select the appropriate model among the candidate models. The candidate ARIMA models for the annual electricity consumption in the Philippines and their corresponding AIC values are presented in Table I. The best ARIMA model found is ARIMA (0,2,1) because it has the most negative AIC value.

TABLE I: CANDIDATE MODELS WITH CORRESPONDING AIC VALUES

Model	AIC
ARIMA (0,2,0)	-116.4855
ARIMA (0,2,1)	-136.6084
ARIMA (0,2,2)	-134.8391
ARIMA (1,2,0)	-124.3067
ARIMA (1,2,1)	-134.8495
ARIMA (1,2,2)	-133.2334
ARIMA (2,2,0)	-125.7119
ARIMA (2,2,1)	-132.8622
ARIMA (2,2,2)	-130.8495

B. Model Parameter Estimation

The model parameters' estimates standard error, z-value, and p-value are shown in Table II. The p-value of the moving average component or MA (1) is less than α=0.05 level of significance (p<0.01). Therefore, the estimate of the parameter is significantly different from zero. This result indicates that the model is capable of making accurate predictions. Performing diagnostic checks on the model is the next step.

TABLE II: COEFFICIENT ESTIMATE OF ARIMA (0,2,1)

	Estimate	Std. Error	*z*-value	*p*-value
MA (1)	-0. 99997	0.1256	-7.9615	<0.01

C. Diagnostic Checking

To determine if the model is adequate, residual analysis is performed. Fig. 7 shows the plots of the results of the residual analysis. It can be observed in the plot of residuals versus time that there are no strong patterns. This indicates that the residuals are not time-dependent. The constancy of variance of the residuals is also observed. Thus, the constancy of error variance assumption was met. In the normality plot, the residuals are close to the theoretical line, which indicates the normality of the residuals.

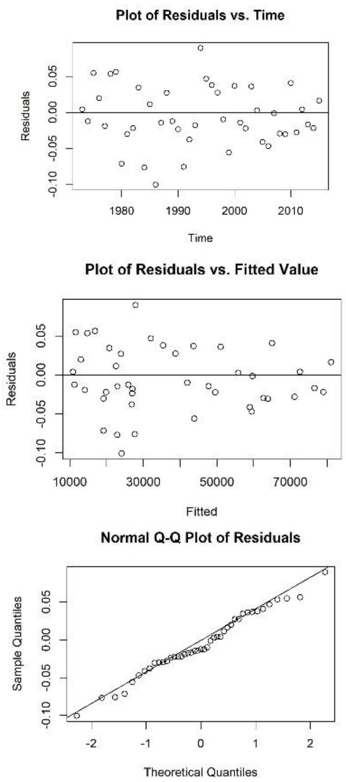

Fig. 7. Residual analysis plots.

TABLE III: SHAPIRO-WILK NORMALITY TEST

	Test Statistic	*p*-value
Shapiro-Wilk	0.98301	0.7652

TABLE IV: LJUNG-BOX TEST

Test Statistic	Df	*p*-value
5.8836	10	0.8249

Finally, it can be observed from Fig. 8 that the autocorrelation and partial autocorrelation are within acceptable limits. Table III presents the results of the Shapiro-Wilk test. From the table, the *p*-value (0.7652) is greater than α=0.05. Hence the residuals were approximately normally distributed. Thus, we can consider the residuals as white noise.

The Ljung-Box test is performed to formally test the model's lack of fit. The results are presented in Table IV. It can be observed from the table that the *p*-value (0.8249) is greater than α=0.05, which indicates that the model fits the data.

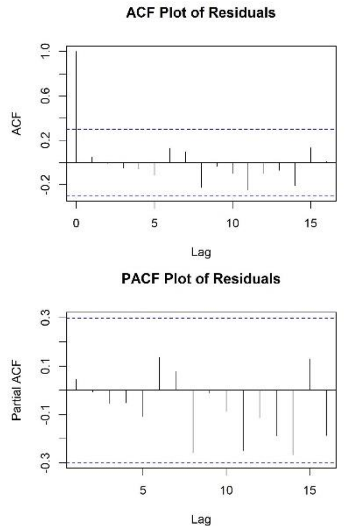

Fig. 8. ACF and PACF plots of the differenced series.

D. Forecast Evaluation

The ARIMA (0,2,1) model is used together with the validation set to obtain the forecasted values from 2016 to 2020. The forecasted values are then compared against the five values in the validation set. Table V shows the actual values, forecasted values, as well as errors of the ARIMA (0,2,1) model.

TABLE V: ACTUAL AND FORECASTED VALUES

Date	Actual	Forecast	Error
2016	90798	86527.10	4270.8972
2017	94370	95372.85	-1002.8484
2018	99765	99120.09	644.9127
2019	106041	104807.78	1233.2249
2020	101756	111307.40	-9551.4006

Fig. 9 shows the forecasted electricity consumption for 2016 to 2020 and the actual data. A large forecast error is observed in the year 2020. The forecasted consumption is 111,307.40 GWh, while the actual consumption is 101,756 GWh. The model cannot anticipate the decrease in electricity consumption caused by the policy adjustments and changes in response to the COVID-19 pandemic.

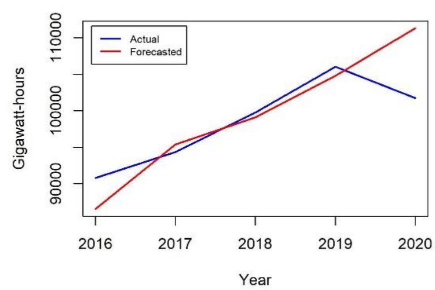

Fig. 9. Actual vs. forecasted values.

It was observed from Fig. 10 that the individual values of ACF and PACF of forecast errors are within acceptable limits, which indicates that correlations are significantly close to zero. Therefore, the forecast errors have a white noise-like behavior.

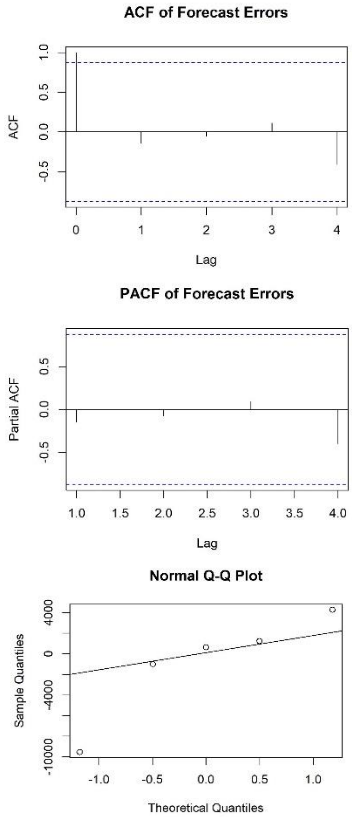

Fig. 10. Plots of the forecast errors.

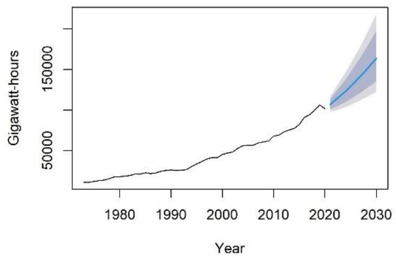

Fig. 11. Forecasted electricity consumption of the Philippines for 2021-2030.

TABLE VI: POINT FORECAST AND 95% CONFIDENCE INTERVAL

Date	Forecast	Lower 95% C.I.	Upper 95% C.I.
2021	106707.0	98153.41	116006.0
2022	111898.9	99306.06	126088.7
2023	117343.4	101230.38	136021.2
2024	123052.9	103581.70	146184.2
2025	129040.1	106235.58	156739.8
2026	135318.6	109133.17	167787.0
2027	141902.6	112242.89	179399.9
2028	148807.0	115546.67	191641.4
2029	156047.3	119034.06	204569.7
2030	163639.9	122699.26	218241.1

The Shapiro-Wilk test is used to test the normality of the forecast errors formally. Since the *p*-value (0.2733) is greater than α=0.05, the forecast errors are approximately normally distributed. Therefore, we can consider the forecast errors as Gaussian white noise.

Fig. 11 shows the plot of the forecasted electricity consumption and forecast intervals of the Philippines from the ARIMA (0,2,1) model. It can be observed that there is an increasing trend from the historical years of 1973 to 2020 up to the predicted years of 2021 to 2030. It is expected that by 2030, the Philippines will consume 163,639.9 gigawatt-hours of electricity. Table VI further shows the forecasted electricity consumption for 2021 to 2030.

V. CONCLUSIONS AND RECOMMENDATIONS

A. Conclusions

In this paper, the implementation of a time series analysis model for electricity consumption in the Philippines was presented. The study aimed to understand the characteristics or behavior and efficiently forecast electricity consumption in the Philippines based on the available historical data. It was observed that the historical data of electricity consumption shows an upward trend, it is because of the country's steadily growing economy, which entails a consistently rising demand and consumption of electricity. There was a slight 4% decrease in electricity consumption from 2019 to 2020. These changes were caused by the changes brought by the COVID-19 pandemic. With the consistent increase in electricity consumption, developing plans for future electricity needs is imperative. A forecasting model was developed to measure the expected electricity consumption to assist in decision-making. The forecasting model considers the characteristics of the electricity consumption historical data, including the fact that the data had an upward trend. The best model to forecast the annual electricity consumption in the Philippines is ARIMA (0,2,1). The model has undergone several validation processes to ensure reliable and accurate forecasts. In general, the model was useful for forecasting. Based on the generated forecasts, the Philippines in the year 2030 will consume 163,639.9 gigawatt-hours of electricity. The electricity consumption can be as low as 122,699.26 gigawatt-hours, or it can be as high as 218,241.1 gigawatt-hours, considering the 95% confidence interval.

B. Recommendations

The recommendations that follow are based on the findings of the analysis.

1) Based on the predicted electricity consumption, preparation is necessary for the energy supply system and substitute to ensure sustainable, reliable, secure, sufficient, and accessible energy.
2) Future researchers need to examine the monthly electricity consumption in the Philippines to detect and explore the seasonality of the electricity consumption.
3) Future studies could compare the presented results with other models, such as, for instance, Artificial Neural Networks (ANNs) and hybrid models.

CONFLICT OF INTEREST

The author declares no conflict of interest.

AUTHOR CONTRIBUTIONS

The author confirms sole responsibility for study conception, data analysis, interpretation of results, and manuscript preparation.

REFERENCES

[1] R. Gumaru and R. Banta, "A comprehensive study of electricity consumption of people in each region in the Philippines," *American Journal of Mechanics Application*, vol. 7, pp.45-48, 2019.

[2] H. Ritchie and M. Roser. Energy *Production and Consumption*. [Online]. Available: https://ourworldindata.org/energy

[3] World Energy Outlook. (2019). [Online]. Available: https://www.iea.org/reports/world-energy-outlook-2019

[4] Gross Domestic Product (GDP) in the Philippines 2026. [Online]. Available: https://www.statista.com/statistics/578709/grossdomestic-product-gdp-in-philippines/

[5] 2019 Power Situation Report. [Online]. Available: https://www.doe.gov.ph/sites/default/files/pdf/electric_power/2019-power-situation-report.pdf

[6] A. Azadeh and Z. Faiz, "A meta-heuristic framework for forecasting household electricity Consumption," *Applied Software Computing*, vol. 11, pp. 614-620, 2011.

[7] P. Chujai, N. Kerdprasop, and K. Kerdprasop, "Time series analysis of household electric consumption with ARIMA and ARIMA models," in *Proc. the International Multiconference of Engineers and Computer Scientist*, Hong Kong, 2013.

[8] J. Steinbuks, J. de Wit, A. Kochnakyan, and V. Foster, "Forecasting electricity demand: An aid for practitioners," *Live Wire*, World Bank, 2017.

[9] F. Mahia, A. Dey, M. Masud, and M. Mahmud, "Forecasting electricity consumption using ARIMA model," in *Proc. 2019 International Conference on Sustainable Industry 4.0(STI)*, 2019, pp. 1-6.

[10] M. Poole, R. McPhee, and D. Canary, "Hypothesis testing and modeling perspectives inquiry," in *Handbook of interpersonal communication*, 3rd ed. M. Knapp and J. Daly, Ed., SAGE, 2002, pp. 23-72.

[11] A. Delima, "Application of time series analysis in projecting Philippines' electric consumption," *International Journal of Machine Learning and Computing*, vol. 9, pp. 694-699, 2019.

[12] M. Lateh, A. Mudah, Z. Yusof, N. Muda, and M. Azmi, "Handling a small dataset problem in prediction model by employ artificial data generation approach: A review," *Journal of Physics: Conference Series 892 012016*, 2017.

[13] D. Li, C. Sen Wu, T. Tsai, and F. Chang, "Using mega-fuzzification and data trend estimation in small data set learning for early FMS scheduling knowledge," *Comput. Oper. Res.*, vol. 33, no. 6, pp. 1857–1869, 2006.

[14] Philippines Energy Market. [Online]. Available: https://www.trade.gov/market-intelligence/philippines-energy-market

[15] 2020 Power Situation Report. [Online]. Available: https://www.doe.gov.ph/sites/default/files/pdf/electric_power/2020_power-situation-report_as_of_09-september-2021.pdf?withshield=2

[16] 2020 Coal Statistics. [Online]. Available: https://www.doe.gov.ph/sites/default/files/pdf/energy_resources/2020-coal-statistics.pdf

[17] J. Wang, Z. Ji, M. Shi, and F. Huang, "Scenario analysis and application research on big data in smart power distribution and consumption systems," in *Proc. the CSEE*, China, pp. 1829-1836, 2015.

[18] K. Carpio, A. Go and C. Roncal, "Forecasting day-ahead electricity prices of Singapore through ARIMA and wavelet-ARIMA," *DLSU Business & Economics Review*, vol. 22, pp. 97-118, 2012.

[19] A. Clements, A. Hurn, and Z. Li, "Forecasting day-ahead electricity load using a multiple equation time series approach," *European Journal of Operational Research*, vol. 251, 522-530, 2016.

[20] C. Li, X. Zheng, Z. Yang, and L. Kuang, "Predicting short-term electricity demand by combining the advantages of ARMA and XGBoost in fog computing environment," *Wireless Communications and Mobile Computing*, 2018.

[21] W. Hong, Y. Dong, W. Zhang, L. Chen, and B. Panigrahi, "Cyclic electric load forecasting by seasonal SVR with chaotic genetic algorithm," *International Journal of Electrical Power & Energy Systems*, vol. 44, pp. 604-614, 2013.

[22] Y. Chae, R. Horesh, R., Y. Hwang, and Y. Lee, "Artificial neural network model for forecasting subhourly electricity usage in commercial buildings," *Energy and Buildings*, vol. 111, pp. 184-194, 2016.

[23] S. Dhavale, S. Deshmukh, K. Shah, and P. Itkurkar, "Prediction of the future electricity consumption and production using the most efficient machine learning algorithm," *International Journal of Future Generation Communication and Networking*, vol. 13, pp. 1609-1621, 2020.

[24] P. Lakshmipriya and F. Enigo, "Fuzzy logic based short-term electricity demand forecast," *International Journal of Engineering and Technology*, vol. 10, pp. 529-534, 2018.

[25] A. Deihimi, O. Orang, and H. Showkati, "Short-term electric load and temperature forecasting using wavelet echo state networks with neural reconstruction," *Energy*, vol. 57, pp. 382-401, 2013.

[26] C. Guzman and J. Rey, "Forecasting residential electric power consumption for Bogotá Colombia using regression models," *Energy reports*, vol. 6, supp. 1, pp. 561-566, 2020.

[27] M. Lewis-Beck, A. Bryman, and T. Futing Liao, *The SAGE Encyclopedia of Social Science Research Methods*, Thousand Oaks, CA: Sage Publications, Inc., 2002.

[28] L. Auria and R. Moro, "Support vector machines (SVM) as a technique for solvency analysis," DIW Berlin discussion paper no. 8 111, 2008.

[29] J. Tu, "Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes," *J Clin Epidemiol*, vol. 49, 1996.

[30] C. Lin, S. Lin, H. Liu, and K. Tseng, "Applying the naïve Bayes classifier to HVAC energy prediction using hourly data," *Microsyst Technol*, vol. 28, pp. 121–135, 2022.

[31] B. Akaya and N. Çolakoğlu, "Comparison of Multi-class Classification Algorithms on Early Diagnosis of Heart Diseases," in *Proc. Conference: ISBIS Young Business and Industrial Statisticians Workshop on Recent Advances in Data Science and Business Analytics*, Istanbul, Turkey, 2019.

[32] M. Shamim, S. Emam, U. Qidwai, and S. Godil, "Fuzzy logic: A "simple" solution for complexities in neurosciences?" *Surgical Neurology International Surg Neurol Int*, 2011.

[33] G. Riberio, J. Sauer, N. Fraccanabbia, V. Mariani, and L. Coelho, "Bayesian optimized echo state network applied to short-term load forecasting," *Energies*, vol. 13, 2020.

[34] Q. Yang, J. Wang, H. Ma, and X. Wang, "Research on COVID-19 based on ARIMA model$^{\Delta}$ taking Hubei, China as an example to see the epidemic in Italy," *Journal of Infection and Public Health*, vol. 13, pp. 1415-1418, 2020.

[35] J. Miao, "The energy consumption forecasting in China based on ARIMA model," in *Proc. International Conference on Materials Engineering and Information Technology Applications*, 2015.

[36] F. Mahia, A. Dey, M. Masud, and M. Mahmud, "Forecasting electricity consumption using ARIMA model," in *Proc. 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI)*, 2019.

[37] K. Kandananond, "Electricity demand forecasting in buildings based on ARIMA and ARX models," in *Proc. the 8th International Conference on Informatics, Environment, Energy and Applications*, 2019.

[38] S. Sarkodie, "Estimating Ghana's electricity consumption by 2030: An ARIMA forecast," *Energy Sources, Part B: Economics, Planning, and Policy*, vol. 12, pp. 936-944, 2017.

[39] T. Nyoni, "Modeling and forecasting demand for electricity in Zimbabwe using the Box- Jenkins ARIMA technique," MPRA Paper 96903, University Library of Munich, Germany, 2019.

[40] D. Montgomery, C. Jennings, and M. Kulahci, *Introduction to time series and forecasting*, 2nd ed., New Jersey: John Wiley & Sons, 2015.

[41] M. Villa, and R. Ceballos, "Analysis and forecasting of fire incidence in Davao city," *Recoletos Multidisciplinary Research Journal*, vol. 8, pp. 35-50, 2020.

Samuel John E. Parreño was born in Davao City, Philippines. He is a faculty of the University of Mindanao Digos College. He received his bachelor's degree from the University of Mindanao Digos College, Digos City, in 2018. He is currently pursuing his master's degree in applied mathematics at the University of Southeastern Philippines, Davao City. He placed first in the Licensure Examination for Teachers in 2018.

Mr. Parreño is a member of the Philippine Association of Researchers and Statistical Software Users, Inc. His research interest is in data analytics, mathematical modeling, and machine learning.