International Research Journal of Multidisciplinary Scope (IRJMS), 2026; 7(2): 639-651

Original Article | ISSN (O): 2582-631X **DOI: 10.47857/irjms.2026.v07i02.09273**

Monthly Electricity Prediction with Time Series Model

Hedi[1]*, Ahmad Deni Mulyadi[1], Ika Yuliyani[1], Anie Lusiani[2], Agus Binarto[3]

[1]Department of Energy Conversion Engineering, Energy Conversion Engineering Study Program, Politeknik Negeri Bandung, Bandung, Indonesia, [2]Department of Mechanical Engineering, Politeknik Negeri Bandung, Bandung, Indonesia, [3]Department of Electrical Electronic Engineering, Politeknik Negeri Bandung, Bandung, Indonesia. *Corresponding Author's Email: hedi@polban.ac.id

Abstract

Electricity consumption exhibits a sustained upward trend driven by population growth, industrial expansion and increasing household demand for electrical appliances. As urban development intensifies, accurate electricity demand prediction becomes essential to support capacity planning and ensure a reliable and sustainable power distribution system. This study aims to enhance the accuracy of monthly electricity consumption prediction through a comparative analysis of three time series models, namely the Autoregressive Integrated Moving Average (ARIMA), Autoregressive Integrated Moving Average with Exogenous Variables (ARIMAX) and Autoregressive Fractionally Integrated Moving Average (ARFIMA) models. Historical monthly electricity consumption data, incorporating the number of working days as an exogenous variable, were analyzed through data transformation, stationarity testing and model identification stages. Model performance was evaluated using statistical indicators, including the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The results indicate that the ARIMA (4,1,4), ARIMAX (2,1,4) and ARFIMA (4,0.2996,4) models satisfy the AIC and BIC criteria; however, based on RMSE, MAE and MAPE, the ARIMAX (2,1,4) model provides the highest prediction accuracy. The inclusion of the number of working days as an exogenous variable significantly improves predictive performance compared to ARIMA and ARFIMA. These findings underscore the importance of incorporating relevant external factors to enhance electricity demand prediction and offer practical implications for electricity utilities and local governments in demand planning, distribution optimization and sustainable energy policy formulation.

Keywords: ARIMAX, Electricity Consumption, Exogenous Variables, Model Comparison, Time Series Models.

Introduction

Electricity consumption is one of the key indicators of economic development and societal well-being (1). Along with population growth and industrial expansion, the demand for electrical energy in various regions of Indonesia continues to increase, including in Tasikmalaya City. The rapid expansion of urban infrastructure, commercial activities and public services has intensified electricity utilization across residential, industrial and governmental sectors. This increasing dependency on electricity not only reflects economic progress but also highlights the strategic importance of reliable energy management systems at the regional level. This condition requires a reliable energy planning and management system to ensure electricity availability and meet public demand sustainably. In addition, monitoring and analyzing electricity consumption patterns are crucial for promoting energy use efficiency and formulating appropriate energy policies (2, 3). A systematic understanding of consumption dynamics enables policymakers to detect structural changes, anticipate shifts in demand behavior and align infrastructure investment with long-term development goals. Predicting future electricity consumption trends also plays a crucial role in assisting local governments and energy providers in designing optimal electricity supply strategies, anticipating demand surges and promoting the implementation of sustainable energy policies (4). Accurate energy demand planning requires the ability to generate high-precision predictions of future consumption patterns (5). Reliable prediction plays a crucial role in supporting energy providers and local governments in formulating supply strategies, determining generation capacity and optimizing distribution systems. In the context of decentralized regional development, accurate local-level forecasting becomes even more

(Received 18th November 2025; Accepted 12th March 2026; Published 09th April 2026)

important, as electricity planning decisions are increasingly tailored to specific demographic, economic and geographic characteristics. Through accurate prediction, the risk of an imbalance between supply and demand can be reduced, operational efficiency improved and production costs minimized through more effective resource planning (6, 7). Furthermore, precise forecasting contributes to environmental sustainability by preventing unnecessary electricity overproduction, thereby reducing fuel consumption and carbon emissions associated with power generation.

Electricity consumption patterns are influenced by various factors, such as differences between weekdays and holidays, long-term trends and seasonal fluctuations. In addition, electricity consumption data often exhibit long-memory characteristics, indicating strong dependence on previous periods. This persistence suggests that shocks or structural changes in electricity demand may have prolonged effects over time, making modeling more complex and requiring approaches capable of capturing both short-term adjustments and long-term dependencies. Therefore, time series models capable of capturing trends, seasonality, volatility and long-term dependence structures are required to produce predictions that more accurately represent the dynamics of future electricity consumption (8). Selecting an appropriate modeling framework is thus a critical step in ensuring that forecast results are statistically robust and practically relevant for decision-making processes.

Although numerous studies have examined electricity demand prediction, many focus on annual or nationally aggregated data, which may overlook short-term variability and more pronounced seasonal dynamics at the monthly level. Monthly electricity consumption data typically display stronger seasonal fluctuations, calendar effects and higher volatility, making prediction more challenging. For instance, variations in the number of working days, public holidays and local economic activities may significantly alter monthly demand patterns, yet these effects are often underrepresented in aggregate analyses. However, studies that explicitly address these complexities within a structured modeling framework while simultaneously evaluating model robustness using comprehensive accuracy measures remain limited. Consequently, the reliability of monthly-level prediction models is not always clearly justified. This limitation indicates the necessity of conducting a rigorous comparative evaluation that integrates both statistical goodness-of-fit criteria and forecasting accuracy indicators.

Various factors, such as differences between working days and holidays, long-term trends and seasonal fluctuations, influence electricity consumption patterns. Moreover, electricity consumption data often exhibit long-memory characteristics, indicating strong dependence on past periods. While time series models such as ARIMA and ARIMAX are widely applied due to their ability to capture linear relationships between past and current values (9–11), these models may not adequately represent long-range dependence in monthly data. Traditional ARIMA assumes integer-order differencing, which may oversimplify persistent stochastic structures observed in real-world electricity consumption series. In addition, previous studies often emphasize model estimation without providing a systematic comparison between competing models or without incorporating exogenous calendar-related variables that may significantly affect electricity demand. The absence of such comparative analysis limits the ability to identify the most reliable forecasting framework for regional energy planning applications. These limitations highlight the need for a more comprehensive modelling approach that considers long-memory characteristics, external influencing factors and rigorous model performance evaluation.

To address these research gaps, this study applies three time series models, ARIMA, ARFIMA and ARIMAX, to monthly electricity consumption data in Tasikmalaya City. The inclusion of ARFIMA enables the assessment of fractional integration properties and long-term persistence, while ARIMAX facilitates the incorporation of relevant exogenous information, specifically the number of working days per month, to capture calendar-related effects. The models are compared using multiple forecasting accuracy measures to determine the most reliable approach for energy planning purposes. Model performance is evaluated using both information criteria and error-based accuracy indicators to ensure a balanced assessment between model parsimony and predictive capability. The prediction results are expected to serve as a foundation for formulating more efficient energy planning strategies to ensure a stable and sustainable electricity supply in Tasikmalaya City. By providing an empirical comparison of short-memory and long-memory models within a unified analytical framework, this study contributes to strengthening the methodological basis of monthly electricity demand forecasting at the regional level. In line with this objective, the research hypothesis states that the ARIMAX model, which incorporates an exogenous variable representing the number of working days per month, provides higher forecasting accuracy compared to the ARIMA and ARFIMA models, which rely solely on endogenous variables.

Methodology

Research Approach

This study employs a quantitative approach using time series analysis to predict monthly electricity consumption in Tasikmalaya City, West Java Province, Indonesia (approximate GPS coordinates' range 7°19′S–7°26′S and 108°10′E–108°15′E). This approach was selected because it can systematically identify long-term trend patterns as well as random fluctuations in historical data. Time series analysis enables researchers to understand the dynamics of changes in electricity consumption over time and reveal intertemporal dependencies among observations. By utilizing available historical data, this approach can be used to construct mathematical models that represent the characteristics of electricity consumption.

Various time series models have been employed in energy consumption forecasting studies, ranging from statistical approaches to artificial intelligence–based techniques. Artificial Neural Networks (ANN), for example, have been used to forecast long-term electricity load and have been shown to deliver higher accuracy than ARIMA models (12) . The strength of ANN lies in its ability to capture nonlinear relationships and complex patterns in electricity consumption data (13). However, its implementation is relatively more complex, requiring intensive training processes and its performance may decline when the data exhibit instability or high volatility. On the other hand, the ARIMA model is one of the most widely used statistical approaches in energy consumption analysis, but it has limitations in capturing long-memory characteristics that often appear in energy time series data.

To address the limitations of ARIMA in representing long-memory dependency structures, the ARFIMA model is employed because it allows for fractional differencing, enabling a more accurate characterization of persistent temporal dynamics (14). The effectiveness of this approach in capturing long-memory behavior has been demonstrated in various empirical studies. One such study shows that sectoral energy consumption in the United States exhibits significant long-term persistence and is more appropriately modeled using fractionally integrated processes rather than conventional ARIMA models (15).

In addition, the ARIMAX model is applied to incorporate exogenous variables that may enhance forecasting performance. Empirical findings from Accra Airport, Ghana, indicate that ARIMAX yields higher prediction accuracy compared to models without external variables (16). Consequently, ARIMAX is regarded as a more comprehensive approach for analyzing electricity consumption influenced by external factors such as the number of working days, meteorological conditions and economic activity (17) .

Data

This study utilizes secondary data obtained from official institutions to ensure the validity and reliability of the analysis results. The primary dataset consists of monthly electricity consumption in Tasikmalaya City, measured in Gigawatt-hours (GWh), sourced from the State Electricity Company for the observation period from January 2009 to February 2025 (a total of 194 monthly observations). This dataset represents the total electricity used by residential, industrial and commercial customers.

In addition, the exogenous variable, namely the number of working days per month, was obtained from the national calendar, as this factor has the potential to influence electricity consumption patterns. All data were then converted into a monthly time series format and processed using the EViews statistical software for model identification, estimation and forecasting. Monthly data were selected to more accurately capture long-term trend variations in the context of the city's energy demand.

A stationarity test was conducted to ensure that the time series data satisfy the basic assumptions of the ARIMA (p, d, q) and ARIMAX (p, d, q) models. Stationary data are characterized by a constant mean, variance and covariance over time, allowing stable and consistent modeling of intertemporal relationships. In this study, the Augmented Dickey-Fuller (ADF) test was employed to detect the presence of unit roots, which indicate non-stationarity in the data (18, 19).

If the test results indicate that the data are non-stationary at the level form, differencing is applied one or more times until a differencing order (d) that yields stationarity is obtained. This process is essential to avoid biased parameter estimates and to ensure that the developed time-series models accurately, stably and reliably represent electricity consumption patterns.

ARIMA and ARIMAX Models

The ARIMA modeling process begins with the identification stage, which involves analyzing serial correlations using the AC and PAC plots to determine the appropriate orders of the autoregressive (p) and moving average (q) parameters. The patterns observed in the AC and PAC plots serve as the basis for identifying the model structure that best fits the characteristics of the time-series data. Subsequently, several combinations of p and q values are tested to obtain the optimal model based on the lowest information criteria values, namely the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) (20). The selection is guided by the AIC and BIC formulations as presented in Equations [1, 2], respectively, which balance model goodness-of-fit and parameter complexity.

$$\mathrm{AIC} = \log(\mathrm{MSE}) + \frac{2k}{n} + c \qquad [1]$$

$$\mathrm{BIC} = \log(\mathrm{MSE}) + k\,\frac{\log(\mathrm{n})}{n} \qquad [2]$$

Where: k= number of estimated parameters, n= number of data observations.

Parameter estimation is conducted using the Maximum Likelihood Estimation (MLE) method, which yields consistent, asymptotically efficient and approximately unbiased estimators under standard assumptions. In general, the ARIMA $(\mathrm{p}, \mathrm{d}, \mathrm{q})$. The model can be written as shown in Equation [3], which represents the combined autoregressive, differencing and moving average structure of the time-series process:

$$\phi_\mathrm{p}(\mathrm{L})(1-\mathrm{B})^\mathrm{d}(\mathrm{Y_t}-\mu) = \theta_\mathrm{q}(\mathrm{L})\varepsilon_\mathrm{t} \qquad [3]$$

In Equation [3], L denotes the lag operator ($\mathrm{L^k Y_t} = \mathrm{Y_{t-k}}$), B is the backward shift operator, d represents the order of differencing, μ is the process mean and ε_t is a white-noise error term with zero mean and constant variance.

The autoregressive component is defined through the polynomial $\phi_\mathrm{p}(\mathrm{L})$ as presented in Equation [4], while the moving average component is defined through the polynomial $\theta_\mathrm{q}(\mathrm{L})$ as shown in Equation [5]:

$$\phi_\mathrm{p}(\mathrm{L}) = \sum_{\mathrm{i}=1}^{\mathrm{p}}(1-\phi_\mathrm{i}\mathrm{L}^\mathrm{i}) \qquad [4]$$

$$\theta_\mathrm{q}(\mathrm{L}) = \sum_{\mathrm{j}=1}^{\mathrm{p}}\left(1-\theta_\mathrm{j}\mathrm{L}^\mathrm{i}\right) \qquad [5]$$

Thus, Equations [3-5] collectively describe the general mathematical formulation of the ARIMA model, integrating the autoregressive, differencing and moving average components into a unified framework.

For the ARIMAX model, the ARIMA structure is extended by incorporating the influence of one or more exogenous variables $\mathrm{X_{k,t}}$. (21). The general formulation of the ARIMAX model is presented in Equation [6], where the dynamic behavior of the dependent variable is explained not only by its past values and past errors but also by external predictors:

$$\phi_\mathrm{p}(\mathrm{L})(1-\mathrm{L})^\mathrm{d}(\mathrm{Y_t}-\mu) = \theta_\mathrm{q}(\mathrm{L})\varepsilon_\mathrm{t} + \sum_{\mathrm{k}=1}^{\mathrm{r}}\beta_\mathrm{k}\,\mathrm{X_{k,t}} \qquad [6]$$

As shown in Equation [6], β_k represents the coefficient that measures the magnitude and direction of the influence of the k-th exogenous variable $X_{k,t}$ on electricity consumption, while $r$denotes the number of external variables included in the model.

ARFIMA Model

Meanwhile, the ARFIMA model is employed to capture long-memory characteristics by allowing the differencing degree d to take fractional values, as introduced and further developed in earlier studies (22).

$$\phi_p(B)(1-B)^d(Y_t-\mu)=\theta_q(B)\varepsilon_t \quad [7]$$

The general form of the ARFIMA (p,d,q). The model is presented in Equation [7], which extends the ARIMA framework by permitting fractional integration. As indicated in Equation [7], a fractional differencing parameter $(-0.5<d<0.5)$ indicates the presence of long-term dependence within the data (23). The estimated model is then tested for statistical significance and assessed using model selection criteria, such as the AIC and BIC, to identify the most appropriate model for the forecasting stage.

Model Performance Evaluation

The performance of each forecasting model was evaluated using several statistical accuracy metrics, including RMSE, MAE and MAPE (24). These measures are formally defined in Equations [8-10], respectively and provide a comprehensive assessment of how closely the predicted values match the observed data.

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(Y_t-\hat{Y}_t\right)^2} \quad [8]$$

$$\mathrm{MAE}=\frac{1}{n}\sum_{t=1}^{n}\left|Y_t-\hat{Y}_t\right| \quad [9]$$

$$\mathrm{MAPE}=\frac{100}{n}\sum_{t=1}^{n}\left|\frac{Y_t-\hat{Y}_t}{Y_t}\right| \quad [10]$$

Where Y_t represents the actual value, $\hat{Y}_t$ is the predicted value and n denotes the total number of observations.

RMSE measures the magnitude of prediction errors by giving greater weight to large deviations, whereas MAE is the average absolute difference between the observed and predicted values. MAPE expresses prediction accuracy in percentage terms, which allows for easier comparison between models with different data scales. In addition to error-based measures, the AIC and BIC were also considered to evaluate the trade-off between model fit and complexity. The model with the lowest AIC, BIC, RMSE, MAE and MAPE values was selected as the best-performing model for forecasting monthly electricity consumption in Tasikmalaya City.

Results

Figure 1 displays the monthly electricity consumption data in Tasikmalaya City from January 2009 to February 2025. The figure clearly shows a strong upward trend over time, indicating a continuous increase in electricity demand during the observation period. In addition, recurring seasonal fluctuations are visible, reflecting monthly consumption patterns that repeat systematically each year.

The modeling process is carried out by dividing the data into two parts, namely an 80% in-sample portion covering the period from January 2009 to February 2023 (170 months) for model training and estimation and a 20% out-of-sample portion covering the period from March 2023 to February 2025 (24 months) for evaluating the model's performance and predictive ability (25). This separation allows for a more robust evaluation of the forecasting accuracy using unseen data.

The variables used in the modeling consist of monthly electricity consumption as the dependent variable (Y_t) and the number of monthly working days as the exogenous variable (X_t), which is assumed to influence fluctuations in electricity consumption in Tasikmalaya City. The presence of increasing variability toward the end of the series further justifies the application of variance-stabilizing transformations before model estimation.

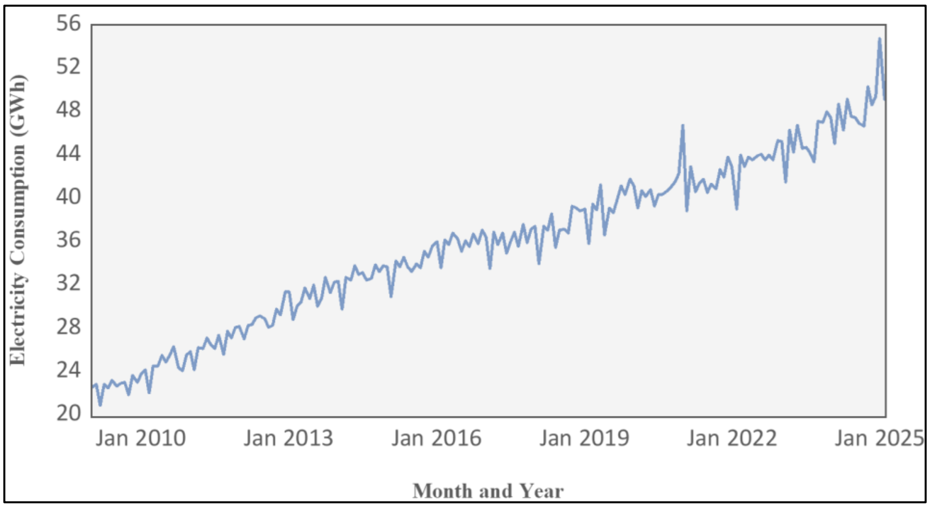


Figure 1: Monthly Electricity Consumption in Tasikmalaya City, 2009–2025

The initial examination showed that the data contained no missing values or outliers, indicating that it was suitable for further analysis. Based on 170 months of in-sample data, a natural logarithm transformation was applied to the variable Y_t, converging into $\log(Y_t)$ to stabilize the variance and improve the data distribution. Figure 1 displays a linear trend, indicating that the dependent variable was not yet stationary. Therefore, a first differencing of $\log(Y_t)$was performed. This result is supported by the statistical tests summarized in Table 1, which confirm that the stationarity criteria were satisfied after differencing. As shown in Table 1, the ADF test statistic (-3.5844) is lower than the 1% and 5% critical values and the corresponding probability value (0.0071) is below the 5% significance level. These findings indicate that the null hypothesis of a unit root can be rejected, confirming that the differenced series D(LOGY) is stationary.

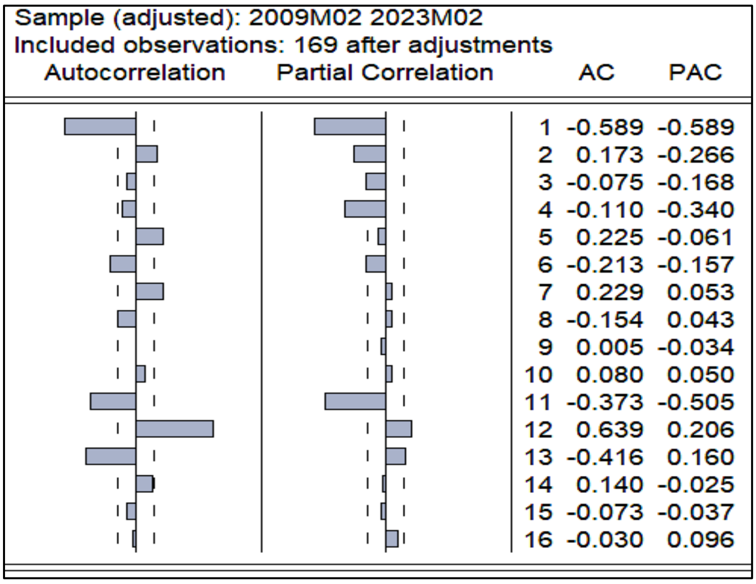
Table 1: Stationarity Test Results Using ADF

Null Hypothesis: D(LOGY) has a unit root

		t-statistic	Probability
ADF test statistic		-3.5844	0.0071
Test critical values	1% level	-3.4725	
	5% level	-2.8800	

In Figure 2, the PAC shows significant spikes at the first few lags, particularly at lag 1, followed by a gradual decline, while the AC also exhibits a strong negative spike at lag 1 and additional significant spikes at seasonal lag 12. The significant spike at lag 1 in both ACF and PACF suggests the presence of short-term dependence, indicating potential AR (1) and/or MA (1) components.

Moreover, the pronounced spike at lag 12 indicates possible seasonal effects in the series, which may require consideration of seasonal AR or MA components in the modeling process. Although a seasonal spike is observed at lag 12, preliminary seasonal model estimations did not provide significant improvement in AIC and BIC values. Therefore, the final model selection focused on non-seasonal ARIMA specifications. These patterns suggest the potential inclusion of AR orders $p = 1, 2, 3, 4$ and MA orders $q = 1,2,3,4$, while also motivating further evaluation of seasonal specifications.



```
Sample (adjusted): 2009M02 2023M02
Included observations: 169 after adjustments
  Autocorrelation    Partial Correlation      AC     PAC

                                         1 -0.589 -0.589
                                         2  0.173 -0.266
                                         3 -0.075 -0.168
                                         4 -0.110 -0.340
                                         5  0.225 -0.061
                                         6 -0.213 -0.157
                                         7  0.229  0.053
                                         8 -0.154  0.043
                                         9  0.005 -0.034
                                        10  0.080  0.050
                                        11 -0.373 -0.505
                                        12  0.639  0.206
                                        13 -0.416  0.160
                                        14  0.140 -0.025
                                        15 -0.073 -0.037
                                        16 -0.030  0.096
```

Figure 2: AC and PAC Plots of First-Differenced Log Electricity Consumption (2009–2023)

Table 2: Model Selection Criteria (AIC and BIC) for ARIMA Models

Model	AIC	BIC
(4,1,4)	-3.9299	-3.7447
(3,1,4)	-3.9283	-3.7616
(2,1,3)	-3.9236	-3.7939
(3,1,3)	-3.9119	-3.7637
(1,1,1)	-3.8668	-3.7186

Using the AIC and BIC criteria for evaluation, sixteen ARIMA candidate models were identified and the five models with the lowest AIC and BIC values are presented in Table 2. The ARIMA (4,1,4) model was selected as the best model for the electricity consumption time-series data.

The parameter estimates are presented in Table 3. Three AR components are statistically significant, each contributing strongly to explaining the dynamics of changes in electricity consumption, while three MA components indicate the presence of short-term shock patterns at their respective orders. The constant value of 0.0041 is also significant.

Table 3: Estimation Results of the ARIMA (4,1,4) Model

Variable	Coefficient	Std. Error	t-statistic	Probability
C	0.0041	0.0008	5.0716	0.0000
AR (1)	-0.7208	0.1288	-5.5958	0.0000
AR (2)	0.5669	0.1679	3.3755	0.0009
AR (3)	0.7419	0.1654	4.4860	0.0000
AR (4)	-0.1303	0.1191	-1.0939	0.2757
MA (1)	-0.1721	0.1033	-1.6661	0.0977
MA (2)	-1.1469	0.0855	-13.4210	0.0000
MA (3)	-0.3005	0.0762	-3.9456	0.0001
MA (4)	0.7806	0.0788	9.9060	0.0000

The selected ARIMA model is then used as the basis for building the ARIMAX model by adding the exogenous variable. The ARIMAX model is created by combining the dependent variable with the exogenous variable. Ten ARIMA candidate models for the Dlog (Y) variable are integrated with the ARMA model of the exogenous variable log(X). Using these combinations, the AIC and BIC values are calculated to select the best model, as shown in Table 4.

Table 4: Comparison of ARIMAX Models Based on AIC and BIC Values

Model	AIC	BIC
ARIMAX(2, 1, 4)	-3.9708	-3.8041
ARIMAX(3, 1, 3)	-3.9699	-3.8032
ARIMAX(3, 1, 4)	-3.9616	-3.7764
ARIMAX(4, 1, 4)	-3.9613	-3.7576
ARIMAX(2, 1, 3)	-3.9126	-3.7644

The comparison results indicate that the model with the lowest AIC and BIC values is the most optimal in describing the relationship between electricity consumption and the number of monthly working days. Accordingly, the ARIMAX (2,1,4) model is selected as the best model for forecasting electricity consumption in Tasikmalaya City. The parameter estimates of the ARIMAX model are presented in Table 5. The estimation results show that the exogenous variable log(X) has a positive and significant effect on Dlog (Y), with a coefficient of 0.0734. The autoregressive and moving-average components collectively demonstrate strong and significant effects, indicating dominant short-term dependence and the impact of shocks from several previous periods. Overall, the ARIMAX model indicates that the dependent variable is affected by both internal dynamics and the external factor log (X).

Table 5: Estimation Results of the ARIMAX (2, 1, 4) Model

Variable	Coefficient	Std. Error	t-statistic	Probability
LOG(X)	0.0734	0.0242	3.0296	0.0029
C	-0.2171	0.0729	-2.9762	0.0034
AR (1)	-1.7085	0.0269	-63.4804	0.0000
AR (2)	-0.9806	0.0221	-44.4530	0.0000
MA (1)	0.8870	0.0812	10.9177	0.0000
MA (2)	-0.3500	0.0927	-3.7771	0.0002
MA (3)	-0.7839	0.1036	-7.5654	0.0000
MA (4)	-0.0747	0.0835	-0.8952	0.3720

The identification of the ARFIMA model is carried out using the same approach as ARIMA, applied to data that have been transformed using the natural logarithm, namely log (Y). Next, the parameter d is estimated for various combinations of p and q in the ARFIMA(p, d, q) model. The smallest AIC and BIC values from the five combinations of p and q, along with the estimated d parameter, are presented in Table 6. The best model obtained is ARFIMA (4, 0.299560, 9).

Table 6: AIC and BIC Values of the ARFIMA (p, d, q) Model

ARFIMA	d	AIC	BIC
(4, d, 4)	0.2996	-3.7869	-3.5840
(3, d, 4)	0.3907	-3.7245	-3.5400
(2, d, 4)	0.4398	-3.7340	-3.5680
(4, d, 3)	0.3812	-3.7500	-3.5655
(4, d, 2)	0.4770	-3.5796	-3.4135

Table 7: Parameter Estimates of the ARFIMA (4, 0.2996, 4) Model

Variable	Coefficient	Std. Error	t-statistic	Probability
C	3.3916	18.8021	0.1804	0.8571
D	0.2996	0.2241	1.3365	0.1833
AR (1)	-1.6899	0.0606	-27.8970	0.0000
AR (2)	0.0794	0.0638	1.2449	0.2150
AR (3)	1.6869	0.0641	26.3002	0.0000
AR (4)	0.9146	0.0574	15.9386	0.0000
MA (1)	1.6960	0.2310	7.3429	0.0000
MA (2)	0.1080	0.4556	0.2371	0.8129
MA (3)	-1.1978	0.3155	-3.7966	0.0002
MA (4)	-0.5963	0.1123	-5.3117	0.0000

The parameter estimates of the selected model are presented in Table 7. Functionally, this model combines fractional differencing $(1-\mathrm{L})^{\mathrm{d}}$ with $d = 0.2996$ to capture long-memory characteristics, along with autoregressive and moving-average components up to the fourth order. The model's dynamic structure is primarily driven by three significant autoregressive coefficients and three significant moving-average coefficients.

Table 8 presents the prediction error values of the ARIMA, ARIMAX and ARFIMA models for the 2009–2023 period. The analysis results show that the ARIMAX model has the highest prediction accuracy compared to the other two models.

Table 8: Comparison of ARIMA, ARIMAX and ARFIMA Model Performance

Indicator	ARIMA (4, 1, 4)	ARIMAX (2,1,4)	ARFIMA (4, 0. 2996, 4)
RMSE	1.0977	1.0793	1.4597
MAE	0.8236	0.7891	1.2038
MAPE	2.43%	2.34%	3.55%

Based on the out-of-sample data for the 2023–2025 period (24 observations), Table 9 presents the prediction error values of the three forecasting models: ARIMA (4, 1, 4), ARIMAX (2, 1, 4) and ARFIMA (4, 0.2996, 4). All three error indicators show that the ARIMAX (2, 1, 4) model yields the lowest error values across all metrics, indicating that the inclusion of an exogenous variable substantially enhances forecasting accuracy by capturing external variations affecting the data. In contrast, the ARFIMA model, although intended to capture long-term dependencies, demonstrates the weakest performance in this dataset, with considerably higher RMSE and MAPE values.

Table 9: Prediction Errors of ARIMA, ARIMAX and ARFIMA (Out-of-Sample 2023–2025)

Indicator	ARIMA (4, 1, 4)	ARIMAX (2,1,4)	ARFIMA (4, 0.2996, 4)
RMSE	3.5105	1.1544	19.1425
MAE	1.4098	0.9922	3.7851
MAPE	2.89%	2.07%	8.21%

Figure 3 presents a comparison between the actual data and the predicted values from the three models for the 2023–2025 period. The ARIMAX model shows the closest pattern to the actual data, both in movement direction and fluctuation amplitude, confirming its ability to capture data dynamics more accurately, consistent with its lowest error values. The ARIMA model also follows the actual pattern, although some deviations appear in certain periods, particularly during value spikes. In contrast, the ARFIMA model displays substantial differences in pattern and fluctuations, making it less aligned with the actual data and resulting in lower accuracy. Overall, Figure 3 confirms that ARIMAX provides the most accurate predictions of the actual conditions.

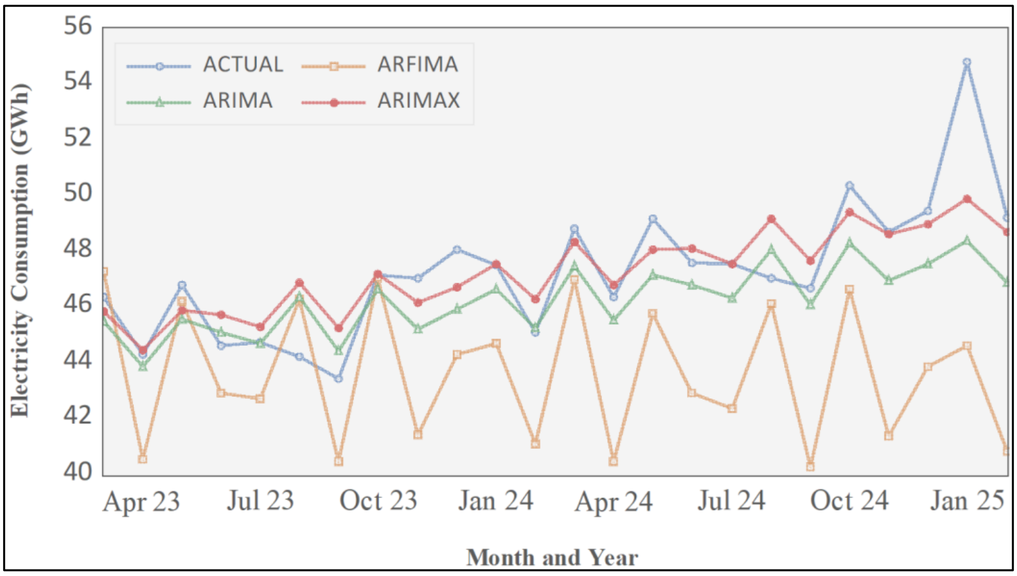


Figure 3: Actual vs. Predicted Values of ARIMA, ARIMAX and ARFIMA (2023–2025)

Figure 4 presents the forecasted electricity consumption for 2025–2026 using three models, ARIMA, ARIMAX and ARFIMA, compared with the actual patterns from previous years. In general, ARIMA and ARIMAX produce forecasts that follow the upward trend in electricity consumption, with monthly fluctuations consistent with the historical data. Among the three, ARIMAX shows the most stable pattern and is closest to the actual data, providing a smoother and more accurate representation of the trend. In contrast, ARFIMA yields lower, relatively flat predictions, failing to capture the increasing pattern observed in the actual data. These findings indicate that for the 2025–2026 prediction period, the ARIMA model, particularly ARIMAX, is more capable of representing the growth in electricity consumption, whereas ARFIMA tends to be less suitable and yields undervalued estimates.

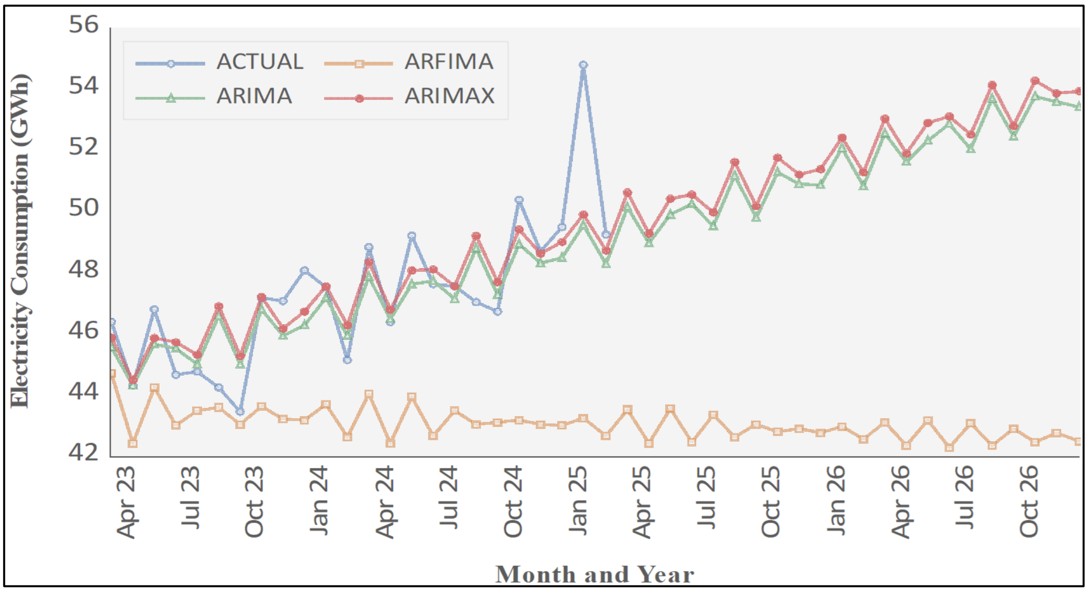


Figure 4: Electricity Consumption Forecasts: ARIMA vs ARIMAX vs ARFIMA

Discussion

The electricity consumption projections for the 2025–2026 period show different predictive patterns across the three models, with ARIMAX consistently demonstrating the most stable pattern and better alignment with electricity demand dynamics. In 2025, both ARIMA and ARIMAX indicate an increasing trend consistent with historical patterns; however, ARIMAX produces smoother forecasts and avoids excessive fluctuations. This result is consistent with recent studies comparing various time-series models in electricity load forecasting, which indicate that models incorporating exogenous variables tend to provide more stable and accurate predictions when demand patterns are influenced by external factors such as seasonality or economic activity, compared to traditional univariate models such as ARIMA alone (26).

Entering 2026, the upward trend in electricity consumption is more evident, with ARIMAX continuing to reflect a realistic, stable growth trajectory. In contrast, ARFIMA produces flatter predictions that do not fully capture the actual growth direction of demand. This difference reinforces findings from recent forecasting studies showing that although classical statistical models are effective for relatively stable patterns, models that incorporate external factors are generally more consistent in capturing medium- to long-term demand growth trends (27).

These findings are further supported by recent research evaluating the performance of modern energy forecasting approaches. Several studies report that statistical models incorporating external components (such as seasonal variables or economic indicators) generally achieve higher forecasting accuracy compared to purely univariate methods, particularly in cases involving strong seasonal variation and overall growth trends. This supports our conclusion that ARIMAX is more appropriate for medium-term projections than models that are less responsive to external dynamics (28).

Overall, the forecasting results indicate increasing pressure on the power system as electricity demand continues to grow through 2026. This condition calls for anticipatory measures in the form of expanding generation capacity, strengthening transmission and distribution networks and optimizing system operations to maintain reliable electricity delivery. The sustained rise in demand also provides an essential foundation for medium-term policy planning, ensuring that energy infrastructure development can be carried out in a targeted manner and aligned with future requirements.

Conclusion

This study provides strong empirical evidence that the ARIMAX model is the most effective forecasting framework for estimating monthly electricity consumption during the 2023–2025 period. Its superior performance, reflected in the lowest prediction error metrics compared with ARIMA and ARFIMA, demonstrates that the inclusion of exogenous variables substantially improves forecasting accuracy. These findings highlight that electricity consumption dynamics are shaped not only by historical patterns but also by influential external factors.

From a policy and planning standpoint, the results offer important practical implications for government agencies, electricity providers and energy policymakers in selecting reliable forecasting methods. Higher predictive accuracy supports more informed decisions related to resource allocation, supply capacity planning, distribution system optimization and long-term energy policy formulation. Overall, this study strengthens the methodological significance of the ARIMAX model for time-series forecasting involving external variables and provides a solid empirical foundation for its wider application in energy research and other multidisciplinary domains.

Despite these contributions, this study has several limitations. The analysis relies on the availability and quality of historical data and selected exogenous variables, which may not fully capture all structural changes or unexpected shocks affecting electricity consumption. In addition, the modelling framework is limited to linear time-series approaches, which may not entirely represent complex nonlinear dynamics.

Future research is therefore encouraged to incorporate a broader set of exogenous variables, explore hybrid or nonlinear modelling approaches and compare the performance of ARIMAX with advanced machine learning techniques. Such extensions would further enhance predictive robustness and provide deeper insights into electricity demand dynamics.

Abbreviations

AC: Autocorrelation, ADF: Augmented Dickey Fuller, AIC: Akaike Information Criterion, ARFIMA: Autoregressive Fractionally Integrated Moving Average, ARIMA: Autoregressive Integrated Moving Average, ARIMAX: Autoregressive Integrated Moving Average with Exogenous Variables, BIC: Bayesian Information Criterion, MAE: Mean Absolute Error, MAPE: Mean Absolute Percentage Error, RMSE: Root Mean Square Error.

Acknowledgement

We would like to express our sincere gratitude to Politeknik Negeri Bandung for the financial support provided through Research Agreement Letter No. 1 08.8/R7/PE.01.03/2025.

Author Contributions

Hedi: conceptualization, methodology, data analysis, writing of the original draft, Ahmad Deni Mulyadi: data collection, data curation, validation, Ika Yuliyani: formal analysis, visualization, writing–review and editing, Anie Lusiani: methodological review, critical revision of the manuscript, Agus Binarto: manuscript review, language editing, final approval. All authors reviewed and approved the final manuscript.

Conflict of Interest

The authors affirm that this research has no conflicts of interest.

Data Availability

The monthly electricity consumption data of Tasikmalaya City used in this study are available from the corresponding author upon reasonable request. Due to confidentiality agreements with the local electricity provider, the data cannot be publicly shared.

Declaration Of Generative AI And AI Assisted Technologies in the Writing Process

The authors declare that no generative AI or AI-assisted technologies were used in the writing process, study design, data analysis, or interpretation of the results. All intellectual contributions were made solely by the authors, who take full responsibility for the content of this manuscript.

Ethics Approval

This study was approved by Politeknik Negeri Bandung through an internal evaluation process in accordance with institutional ethics standards and policies.

Funding

None.

References

1. Bonsu MO, Assibey W, Wang Y. The triangular relationship between energy consumption, trade

openness and economic growth: new empirical evidence. Cogent Econ Financ. 2022;10(1). https://doi.org/10.1080/23322039.2022.2140520

2. Caldera M, Hussain A, Romano S. Energy-consumption pattern-detecting technique for household appliances for smart home platform. Energies. 2023;16(2):824. https://doi.org/10.3390/en16020824
3. Duarte JE, Rosero-Garcia J. Analysis of variability in electric power consumption: a methodology for setting time-differentiated tariffs. Energies. 2024;17(4):842. https://doi.org/10.3390/en17040842
4. Ugbehe PO, Diemuodeke OE, Aikhuele DO. Electricity demand forecasting methodologies and applications: a review. Sustain Energy Res. 2025; 12(19):1–32. https://doi.org/10.1186/s40807-025-00149-z
5. Matos M, Almeida J, Gonçalves P, Baldo F, Braz FJ, Bartolomeu PC. A machine learning-based electricity consumption forecast and management system for renewable energy communities. Energies. 2024;17(3):1–25. https://doi.org/10.3390/en17030630
6. Zhang W, Wu J. A stochastic MPC-based flexibility scheduling strategy for community integrated energy system considering multi-temporal-spatial-scale and inertia components. Processes. 2024;12(3):457. https://doi.org/10.3390/pr12030457
7. Navale S, Mishra N, Borhade S. Deep learning approaches for energy consumption forecasting: analyzing stress factors and optimizing models for future demand. Discover Appl Sci. 2025;7:1092. https://doi.org/10.1007/s42452-025-07718-3
8. Zhou L, Zhou Y, Liu L, Zhao X. A comparative study of different deep learning methods for time-series probabilistic residential load power forecasting. Front Energy Res. 2024;12:1490152. https://doi.org/10.3389/fenrg.2024.1490152
9. Chen H, Bao, Pei L, Ling, Zhao Y, Feng. Forecasting seasonal variations in electricity consumption and electricity usage efficiency of industrial sectors using a grey modeling approach. Energy. 2021;222: 119952. https://doi.org/10.1016/j.energy.2021.119952
10. Chodakowska E. ARIMA models in electrical load forecasting and their applications. Energies. 2021;14(23):7952. https://doi.org/10.3390/en14237952
11. Maku TO, Adehi MU, Adenomon MO. Modeling and forecasting electricity consumption in Nigeria using ARIMA and ARIMAX time series models. Sci World J. 2023;18(3):414–421. https://doi.org/10.4314/swj.v18i3.14
12. Tarmanini C, Sarma N, Gezegin C, Ozgonenel O. Short term load forecasting based on ARIMA and ANN approaches. Energy Rep. 2023;9:550–557. https://doi.org/10.1016/j.egyr.2023.01.060
13. Vanting NB, Ma Z, Jørgensen BN. A scoping review of deep neural networks for electric load forecasting. Energy Inform. 2021;4(Suppl 2). https://doi.org/10.1186/s42162-021-00148-6
14. Aron S, Lucas M. Comparing ARFIMA and ARIMA models in forecasting under five mortality rate in Tanzania. Asian J Probab Stat. 2025;27(1):107–121. https://doi.org/10.9734/ajpas/2025/v27i1707
15. Adekoya OB. Modeling of persistence and seasonality in sectoral energy consumption in the USA using fractionally integrated processes: implications for economic policy. Nat Resour Res. 2020;29(4):2787–2800. https://doi.org/10.1007/s11053-019-09599-x
16. Moslemi Z, Rehome S, Clark L, *et al.* Comprehensive forecasting of California’s energy consumption: a multi-source and sectoral analysis using ARIMA and ARIMAX models. World J Adv Res Rev. 2024;22(2):484–497. https://doi.org/10.48550/arXiv.2402.04432
17. Djimasbe R, Gyamfi S, Iweh CD, Ribar BN. Development of an ARIMAX model for forecasting airport electricity consumption in Accra-Ghana: the role of weather and air passenger traffic. e-Prime Adv Electr Eng Electron Energy. 2024;9(July):100691. https://doi.org/10.1016/j.prime.2024.100691
18. Dickey DA, Fuller WA. Distribution of the estimators for autoregressive time series with a unit root. J Am Stat Assoc. 1979;74(366):427–431. https://doi.org/10.1080/01621459.1979.10482531
19. Hossain ML, Shams SMN, Ullah SM. Time-series and deep learning approaches for renewable energy forecasting in Dhaka: a comparative study of ARIMA, SARIMA and LSTM models. Discover Sustain. 2025;6:775. https://doi.org/10.1007/s43621-025-01733-5
20. Zhang X, Cao W. Research on time series forecasting method based on autoregressive integrated moving average model with zonotopic Kalman filter. Sustainability. 2025;17(7):2993. https://doi.org/10.3390/su17072993
21. Bennett C, Stewart RA, Lu J. Autoregressive with exogenous variables and neural network short-term load forecast models for residential low voltage distribution networks. Energies. 2014;7:2938–2960. https://doi.org/10.3390/en7052938
22. Devianto D, Ramadani K, Asdi Y. The hybrid model of autoregressive integrated moving average and fuzzy time series Markov chain on long-memory data. Front Appl Math Stat. 2022;8:1045241. https://doi.org/10.3389/fams.2022.1045241
23. Boutahar M, Marimoutou V, Nouira L. Estimation methods of the long memory parameter: Monte Carlo analysis and application. J Appl Stat. 2007;34(3):261–301. https://doi.org/10.1080/02664760601004874
24. Hodson TO. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. Geosci Model Dev. 2022;15(14):5481–5487. https://doi.org/10.5194/gmd-15-5481-2022
25. Ungureanu S, Topa V, Cziker AC. Analysis for non-residential short-term load forecasting using machine learning and statistical methods with financial impact on the power market. Energies. 2021;14(21):6966. https://doi.org/10.3390/en14216966
26. Kim H, Park S, Kim S. Time-series clustering and forecasting household electricity demand using smart meter data. Energy Rep. 2023;9:4111–4121. https://doi.org/10.1016/j.egyr.2023.03.042
27. Mustafa AT, Al-Yozbaky OSA. Forecasting energy demand and generation using time series models: a

comparative analysis of classical, grey, fuzzy and intelligent approaches. Franklin Open. 2025;12:100350. https://doi.org/10.1016/j.fraope.2025.100350

28. Zhang H, Liu X, Gao X, Liu Z. Hybrid ARIMAX model integrating optimal time-lag selection and Markov chain error correction for residential heating load forecasting. Energy Rep. 2025;14:5259–5273. https://doi.org/10.1016/j.egyr.2025.11.082

How to Cite: Hedi, Mulyadi AD, Yuliyani I, Lusiani A, Binarto A. Monthly Electricity Prediction with Time Series Model**.** Int Res J Multidiscip Scope. 2026; 7(2): 639-651. DOI: 10.47857/irjms.2026.v07i02.09273