JSCE (Journal of System and Computer Engineering)
Vol.6, No.1, January 2025, pp. 131-136
ISSN (online): 2723-1240
DOI : https://doi.org/10.61628/jsce.v6i1.1666

ARIMA Method Implementation for Electricity Demand Forecasting with MAPE Evaluation

Supriyadi La Wungo [1,a]; Firman Aziz [2,b*]; Jeffry [3,c] ; Mardewi [1,d]; Nasruddin [2,e]
[1] *STMIK Kreatindo, Kali Bambu Street, Manokwari 98312, Indonesia*
[2] *Universitas Pancasakti, Andi Mangerangi Street 73, Makassar 90121, Indonesia*
[3] *Institut Teknologi Bacharuddin Jusuf Habibie, Jl. Balaikota No.1, Parepare 91122, Indonesia*
[a] *supriyadi.la.wungo@gmail.com* ; [b] *firman.aziz@unpacti.ac.id* ; [c] *jeffry@ith.ac.id* ;[d] *mardewi0004@gmail.com* ;[d] *darinasruddin@gmail.com*
* *Corresponding author*

Abstract

Electricity demand forecasting is critical for efficient energy management and planning. This study focuses on the development and implementation of the Autoregressive Integrated Moving Average (ARIMA) method for forecasting electricity demand in South Sulawesi's power system. The evaluation of forecasting accuracy was conducted using the Mean Absolute Percentage Error (MAPE), which measures the percentage error between predicted and actual values. Two experiments were conducted with different ARIMA models: ARIMA(5,1,0) and ARIMA(2,0,1). Results showed that the ARIMA(5,1,0) model achieved a MAPE of 2.15%, while the ARIMA(2,0,1) model performed slightly better with a MAPE of 1.91%, indicating highly accurate predictions. The findings highlight the effectiveness of the ARIMA method in forecasting electricity demand, providing a reliable tool for energy providers to optimize resource allocation and enhance operational efficiency. Future research may explore integrating ARIMA with other advanced methods to further improve forecasting performance.

Keywords*—ARIMA, Electricity Demand Forecasting, Time Series Analysis, MAPE, Forecasting Models.*

1. Introduction

The increasing demand for electricity necessitates effective planning strategies, one of which involves forecasting electricity usage based on historical data (Ahmad & Zhang, 2020). The Autoregressive Integrated Moving Average (ARIMA) method has been a widely used approach in time series forecasting due to its ability to analyze historical patterns and provide accurate predictions (Schaffer et al., 2021). Several studies published in reputable journals support the effectiveness of ARIMA for this application (Song & Cao, 2022).

Research (Nepal et al., 2020) in Energy Reports applied ARIMA to forecast short-term electricity demand in Europe. The study demonstrated that ARIMA achieved a Mean Absolute Percentage Error (MAPE) of 6.5%, indicating good predictive accuracy (Mystakidis et al., 2024). However, the model's performance declined when handling more complex seasonal patterns. Similarly, Chen et al. (2021) in Applied Energy combined ARIMA with Long Short-Term Memory (LSTM) to improve electricity forecasting in urban areas of China (Alizadegan et al., 2024). This hybrid approach reduced the MAPE to 3.2%, but it required significant computational resources, making it less suitable for small-scale applications (Kondaiah et al., 2022).

Furthermore, Rahman et al. (2022) in Renewable and Sustainable Energy Reviews evaluated the performance of ARIMA for electricity demand forecasting in South Asia. The findings revealed that ARIMA struggled with high fluctuations in daily data, often caused by sudden changes in weather or economic activities. Meanwhile, Al-Turki and Basheer (2023) in the

Journal of Energy Systems highlighted the need for integrating ARIMA with other methods to enhance its performance in handling seasonal variability.

Additionally, the application of machine learning and hybrid models to forecast trends and consumer behavior has gained traction, as highlighted by Jeffry, Usman, & Aziz (2023). Their research demonstrated the effectiveness of ensemble methods, such as logistic regression, in analyzing customer behavior patterns. Although their work primarily focused on consumer trends, the integration of ensemble methods into forecasting models offers valuable insights for enhancing prediction accuracy in dynamic systems, including electricity demand forecasting.

These studies indicate that ARIMA is a reliable method for electricity demand forecasting but has limitations in capturing complex seasonal patterns and managing dynamic real-time data. To address these gaps, this study aims to optimize ARIMA by adapting it to handle seasonal fluctuations, incorporating regional analyses, and evaluating its performance using real-time data. Inspired by previous works, including the integration of advanced methods like those highlighted by Jeffry et al. (2023), this approach is expected to contribute significantly to improving the accuracy of electricity demand forecasting, thereby supporting more efficient and sustainable energy management.

2. Method

The methodology for this study was systematically designed to ensure accurate forecasting of electricity demand using the Autoregressive Integrated Moving Average (ARIMA) model, with rigorous evaluation through the Mean Absolute Percentage Error (MAPE). The methodological steps are described below in detail:

2.1 Data Collection

Data collection for this study involved sourcing historical electricity consumption and system attributes from reliable sources. Key data included the RUPTL 2021-2030 Document, which provides projections for electricity consumption growth in South, Southeast, and West Sulawesi, and power balance data for 2021-2022, detailing monthly peak load data from January 2020 to December 2022. The study focused on four key attributes: DMP (total electrical energy produced), BP (peak load), Cad (reserve capacity), and Status (system condition). Data was recorded monthly to align with operational planning and industrial demand trends. A literature review also supported the analysis, focusing on demand forecasting and ARIMA-based methods.

2.2 Preprocessing

Preprocessing ensures the dataset is suitable for ARIMA modeling. Key steps include:

2.2.1 Handling Missing Data:

Missing values were identified and addressed through:

- Linear Interpolation: For small gaps in continuous data.
- Seasonal Mean Imputation: For significant missing periods, the average demand from the same seasonal period was used.

2.2.2 Outlier Detection and Handling:

Outliers were identified using the Interquartile Range (IQR) and Z-score methods. Depending on their nature, outliers were either smoothed or removed.

2.2.3 Stationarity Testing:

Stationarity is a prerequisite for ARIMA.

- The Augmented Dickey-Fuller (ADF) Test was applied, with p-values below 0.05 indicating stationarity.
- If non-stationary, differencing was applied to stabilize the mean, while transformations like logarithmic scaling were used to reduce variance.

2.2.4 Seasonality Decomposition:

The dataset was decomposed into trend, seasonal, and residual components using Seasonal Decomposition of Time Series (STL) for a deeper understanding of patterns.

2.3 ARIMA Model Development

The ARIMA model development followed a structured workflow:

2.3.1 Model Identification:

Initial parameters *(p,d,q)* were determined based on the analysis of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF).

- **p** (autoregressive order): Determined from PACF plot.
- **d** (degree of differencing): Determined through stationarity testing.
- **q** (moving average order): Determined from ACF plot.

2.3.2 Parameter Optimization:

- Multiple ARIMA models with varying *(p,d,q)* combinations were tested.
- The best-fitting model was selected based on the lowest Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values.

2.3.3 Cross-Validation:

- The dataset was split into training (80%) and testing (20%) subsets.
- Time-based cross-validation was performed to evaluate model stability and prevent overfitting.

2.4 Forecasting and Validation

The optimal ARIMA model, identified as ARIMA *(p,d,q)*, was used to generate electricity demand forecasts. Validation involved:

1. Testing on Held-Out Data: Predicted values were compared against actual values from the test set.
2. Error Metrics:
   - Mean Absolute Percentage Error (MAPE) was calculated to measure accuracy:

$$\mathrm{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \times 100$$

   - Additional metrics like Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were computed for robustness.
3. Residual Analysis: Residual plots and statistical tests (e.g., Ljung-Box test) ensured that residuals were white noise, confirming model validity.

2.5 Performance Evaluation

Performance was evaluated based on MAPE, categorized as:

- **< 10%**: Very good.
- **10–20%**: Good.
- **20–50%**: Fair.

- **> 50%**: Poor.

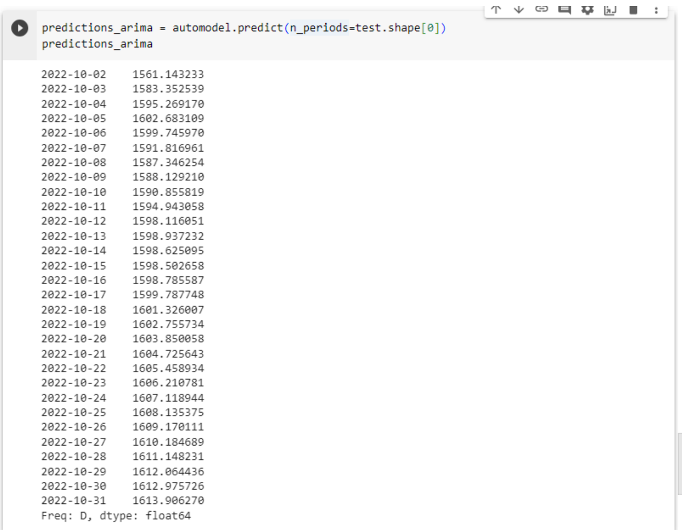
Forecast accuracy was compared against baseline models such as naive and seasonal naive models to assess ARIMA's relative performance.

3. Results And Discussion

Based on the conducted study, the evaluation of the forecasting results using the ARIMA method showed excellent accuracy based on the MAPE (Mean Absolute Percent Error) value. The following is a discussion of each experiment performed:

First Experiment Results

In the first experiment, the best model selected was **ARIMA(5,1,0)(0,0,0)[1]**. This model was used to forecast the peak electricity load in the South Sulawesi power system for October 2022. Based on the calculations, the MAPE value obtained was **2.15%**, which falls under the **Very Good** category according to the MAPE value criteria table.



```
predictions_arima = automodel.predict(n_periods=test.shape[0])
predictions_arima

2022-10-02    1561.143233
2022-10-03    1583.352539
2022-10-04    1595.269170
2022-10-05    1602.683109
2022-10-06    1599.745970
2022-10-07    1591.816961
2022-10-08    1587.346254
2022-10-09    1588.129210
2022-10-10    1590.855819
2022-10-11    1594.943058
2022-10-12    1598.116051
2022-10-13    1598.937232
2022-10-14    1598.625095
2022-10-15    1598.502658
2022-10-16    1598.785587
2022-10-17    1599.787748
2022-10-18    1601.326007
2022-10-19    1602.755734
2022-10-20    1603.850058
2022-10-21    1604.725643
2022-10-22    1605.458934
2022-10-23    1606.210781
2022-10-24    1607.118944
2022-10-25    1608.135375
2022-10-26    1609.170111
2022-10-27    1610.184689
2022-10-28    1611.148231
2022-10-29    1612.064436
2022-10-30    1612.975726
2022-10-31    1613.906270
Freq: D, dtype: float64
```

Figure 1. Forecasting Results of Experiment 1.

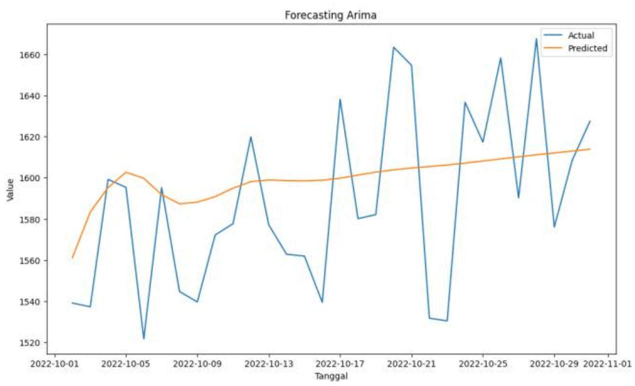


Figure 2. Chart of Forecasting Results using the auto ARIMA model in the First Experiment

The prediction result table revealed that most of the absolute percentage error values between the actual and predicted data were below 5%, indicating relatively accurate predictions. The forecasting graph also showed that the predicted pattern closely matched the actual data pattern. Thus, the ARIMA(5,1,0) model provided very good forecasting results in this experiment.

Second Experiment Results

In the second experiment, the best model selected was **ARIMA(2,0,1)(0,0,0)[1]**. This model was used to forecast the peak electricity load data for July 2023. Based on the calculations, the MAPE value obtained was **1.91%**, which also falls under the **Very Good** category. This value is lower than that of the first experiment, indicating that this model generated more accurate predictions.

```
predictions_arima = automodel.predict(n_periods=test.shape[0])
predictions_arima

2023-07-01    1596.605199
2023-07-02    1599.318388
2023-07-03    1600.416407
2023-07-04    1600.726908
2023-07-05    1600.653587
2023-07-06    1600.393353
2023-07-07    1600.042248
2023-07-08    1599.647120
2023-07-09    1599.230817
2023-07-10    1598.804484
2023-07-11    1598.373557
2023-07-12    1597.940686
2023-07-13    1597.507164
2023-07-14    1597.073621
Freq: D, dtype: float64
```

Figure 3. Forecasting Results of Experiment 2.

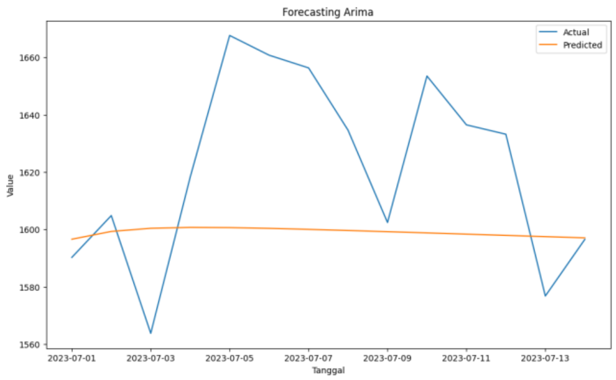


Figure 4. Chart of Forecasting Results using the auto ARIMA model in the Second Experiment

The prediction result table showed that there were four days with the difference between the predicted and actual data close to 0%. Additionally, most of the absolute error values were below 3%, indicating a very high accuracy level. The forecasting graph also showed that the predicted pattern was almost identical to the actual data, making the prediction validation very convincing.

Comparison and Discussion

A comparison between the two experiments showed that the ARIMA(2,0,1) model in the second experiment provided better results compared to ARIMA(5,1,0) in the first experiment, with a lower MAPE value (1.91% vs 2.15%). This indicates that the selection of the appropriate model parameters has a significant impact on the forecasting accuracy.

Overall, the use of MAPE as a performance indicator for the forecasting system provided an objective assessment of the prediction results. The low MAPE values in both experiments indicate that the ARIMA model used is highly suitable for the peak electricity load data characteristics in the South Sulawesi region.

With high prediction accuracy, these forecasting results can serve as a reference for decision-making in planning and managing the electricity power system, ultimately improving operational efficiency and reducing the risk of peak electricity load estimation errors.

4. Conclusions

The study successfully applied the ARIMA model to forecast peak electricity load in the South Sulawesi power system. The results demonstrated that the ARIMA models used provided highly accurate forecasts, with MAPE values indicating excellent prediction performance. The ARIMA(5,1,0)(0,0,0)[1] model in the first experiment achieved a MAPE of 2.15%, while the ARIMA(2,0,1)(0,0,0)[1] model in the second experiment achieved an even better MAPE of 1.91%. These low MAPE values suggest that the ARIMA models are highly effective for forecasting peak electricity load in the region.

The findings highlight the importance of selecting appropriate model parameters, as evidenced by the improved accuracy in the second experiment with the ARIMA(2,0,1) model. The results are valuable for energy system planning, as accurate demand forecasting can lead to better management of electricity generation and distribution, improving operational efficiency and reducing risks associated with inaccurate load predictions. Future studies may explore additional methods or models to further enhance forecasting accuracy and adapt to evolving electricity demand trends.

References

Ahmad, T., & Zhang, H, B. Y. (2020). A review on renewable energy and electricity requirement forecasting models for smart grid and buildings. *Elsevier*. https://www.sciencedirect.com/science/article/pii/S2210670720300391

Alizadegan, H., Rashidi Malki, B., Radmehr, A., Karimi, H., & Ilani, M. A. (2024). Comparative study of long short-term memory (LSTM), bidirectional LSTM, and traditional machine learning approaches for energy consumption prediction. *Energy Exploration and Exploitation*. https://doi.org/10.1177/01445987241269496

Jeffry, J., Usman, S., & Aziz, F. (2023). Analisis Perilaku Pelanggan menggunakan Metode Ensemble Logistic Regression. JURNAL TEKNOLOGI DAN ILMU KOMPUTER PRIMA (JUTIKOMP), 6(2), 90–97

Kondaiah, V. Y., Saravanan, B., Sanjeevikumar, P., & Khan, B. (2022). A review on short-term load forecasting models for micro-grid application. *The Journal of Engineering*, *2022*(7), 665–689. https://doi.org/10.1049/TJE2.12151

Mystakidis, A., Koukaras, P., Tsalikidis, N., & Energies, D. I. (2024). Energy Forecasting: A Comprehensive Review of Techniques and Technologies. *Mdpi.Com*. https://www.mdpi.com/1996-1073/17/7/1662

Nepal, B., Yamaha, M., Yokoe, A., & Yamaji, T. (2020). Electricity load forecasting using clustering and ARIMA model for energy management in buildings. *Japan Architectural Review*, *3*(1), 62–76. https://doi.org/10.1002/2475-8876.12135

Schaffer, A. L., Dobbins, T. A., & Pearson, S. A. (2021). Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions. *BMC Medical Research Methodology*, *21*(1). https://doi.org/10.1186/S12874-021-01235-8

Song, Y., & Cao, J. (2022). An ARIMA-based study of bibliometric index prediction. *Aslib Journal of Information Management*, *74*(1), 94–109. https://doi.org/10.1108/AJIM-03-2021-0072/FULL/HTML