Available online at https://ejournal.corespub.com/index.php/ijmsc/index

International Journal of Mathematics, Statistics, and Computing

e-ISSN 3025-0803

Vol. 3, No. 3, pp. 88-94, 2025

Prediction of The Electricity Capacity Ready to Sell in DKI Jakarta Using Holt's Linear Exponential Smoothing and Arima Methods

Dhela Asafiani Agatha[1*], Wiwik Wiyanti[2]

[1,2]*Department of Statistics, Faculty of Science, Computer, and Mathematics, Matana University Tangerang,, Banten, Indonesia*

**Corresponding author email: dhela.agatha@student.matanauniversity.ac.id*

Abstract

The point of this study is to look at how well Holt's Linear Exponential Smoothing and Autoregressive Integrated Moving Average (ARIMA) can predict time series data that have trend and non-seasonal characteristics. The information on the power capacity available for sale (kWh) at DKI Jakarta serves as the case study. It is anticipated that this study will serve as a guide for choosing efficient techniques for data types with trend and non-seasonal characteristics. This study uses a quantitative methodology with the application of Holt's Linear Exponential Smoothing and Autoregressive Integrated Moving Average (ARIMA). A total of 36 data points—monthly data from January 2020 to December 2022—were used in this study. From the analysis results, the error accuracy level was obtained based on the MAPE calculation, namely 3.18% for Holt's Linear Exponential Smoothing. Meanwhile, the best model with the ARIMA method is ARIMA (3,1,1) with a MAPE value of 3.124%. Based on the forecast results from January to March 2023, the predictions with the best model, namely ARIMA (3,1,1), are 3,140,106,571 kWh, 3,149,746,276 kWh and 3,154,664,915 kWh.

Keywords: Arima, Exponential Smoothing, Linier Holt, Predict, Electricity.

1. Introduction

The availability of ready-to-use electricity is very important in ensuring the stability and reliability of electricity supply for consumers. In Indonesia, until 2024, the government is focusing on infrastructure development. Reported from (Trianto, 2024), throughout 2023, the State Electricity Company of Indonesia will continue to massively intensify the development of electricity infrastructure to support economic growth and national equality. Throughout 2023, the State Electricity Company of Indonesia succeeded in increasing its power generation capacity to 4,182.2 Mega Watt (MW). Based on data from the *Badan Pusat Statistik (BPS)*, the largest distribution of electricity supply from 2021 and 2022 in the largest order is West Java, East Java, and DKI Jakarta, which has now changed its name to DKI Jakarta (Statistik, 2024). It's interesting to note that the special region of Jakarta has up to now developed into the hub of industry, the economy, and numerous other industries that depend on energy. Therefore, it is important to consider the importance of development, especially regarding electricity supply in Jakarta. Quoted based on data from the *Badan Pusat Statistik (BPS)*, DKI Jakarta is the third largest province in terms of electricity distribution from the State Electricity Company of Indonesia amounting to 34,578,291,710 kWh in 2022 (Statistik, 2024).

The demand for energy in DKI Jakarta keeps rising due to the city's dynamic economic development and fast population growth. PLN, as the main electricity provider in this area, faces great pressure to ensure adequate electricity supply to meet the increasing needs, especially in the industrial, commercial, and urban sectors (Murdifi, 2021). Meanwhile, for households, electricity consumption increased by 4.39% from the beginning of the year to March 2022 compared to the previous year. On the same website, PLN also reported an increase in the number of household customers in Jakarta and its surroundings by 3.3% compared to the previous year. In another report, or the Jakarta Distribution Unit, electricity consumption increased by 2% in 2022, affecting the available saleable electricity capacity (Christian, 2022). Despite various efforts to improve electricity infrastructure, DKI Jakarta still faces significant challenges in providing stable and high-quality electricity (Murdifi, 2021).

The increase in demand for electricity each year can actually be predicted. This prediction can be used as a reference for making policies for the state electricity provider, in this case PLN. For example, what is the estimated amount of electricity supply that must be met. Therefore, there is a need for a reliable prediction method for analysing

PLN electricity capacity data. There are many forecasting methods that can be used for time series statistical data. These methods are used according to the characteristics of the statistical data available. This is because it is necessary to obtain predictions with a small error level.

Holt and ARIMA are two prediction techniques that can be applied to describe data that exhibits trends (Makridakis & Wheelwright, 1997). Predictions using Holt's can be made in time series data in various fields. Holt's method has been applied to many time series data with the character of a linear trend and obtained forecast error values that can be said to be small. As in health research on diabetes mellitus (Evania et al., 2024). Research on dengue fever prediction utilizing Holt's through Pegel's classification (Wiyanti & Siregar, 2023). Research in the field of economics for stock price prediction obtained a prediction error of around 4%-6% compared to the ARCH-GARCH method, which has a larger error of 9.81%-12.4% (Nugroho et al., 2024).

Another prediction method that can be relied on to use is ARIMA. Research on ARIMA is also widely used for predicting time series data. For example, research on the ARIMA method for electrical energy needs was applied by (Wahyuni et al., 2024) with a case study in the city of Makassar, with the research results showing that the best model was ARIMA (1,0,0) with a MAPE value of 0.4735%, which means that the mathematical model is very good and suitable for use in forecasting. Research (Melantika & Wiyanti, 2024) obtained the best model for forecasting using ARIMA, which is ARIMA(2,2,2), with a forecast error value of 0.5404588%. In health research to predict the number of mental disorder patients, the ARIMA method is the best method from ARCH, GARCH with the smallest error value using MAPE, which is 9.7% (Arifin & Wiyanti, 2024).

Based on this background, this study is to compare Holt's method with ARIMA to predict the electricity capacity ready for sale. This study's primary goals are to ascertain how well the two forecasting techniques estimate the sale able electricity capacity in DKI Jakarta, assess the expected error values, and comprehend the prediction patterns that may serve as the foundation for future policy decisions pertaining to electricity supply management.

2. Literature Review

The theoretical basis of this study is to obtain the right method for predicting the State Electricity Company of Indonesia time series data. Many data characters cannot be predicted using any method. In statistics, time series data certainly has characters according to the data patterns they have. For example, seasonal, trend, linear, non-linear or cyclical. In this study, the research idea came from several previous researchers who used the Hol't and ARIMA methods which can be applied to various time series data, and the results have a calculation error that can be said to be small.

Numerous disciplines, including economics and health, use Hol't research, which is a component of the exponential smoothing method. Specifically, the MAPE computation yields a prediction error score of 2.4% when predicting diabetes mellitus, one of the highly expected killer diseases (Evania et al., 2024). Still in the health sector, the use of Holt's method, which is covered in Pegel's classification, produces a fairly small prediction error in the case of dengue fever data (Wiyanti & Siregar, 2023). Holt's method is also powerful for economic data, namely predicting stock data, where the prediction results get an error close to zero (Nugroho et al., 2024).

In addition to Holt's method, this study uses the ARIMA method, or in other words, the Box-Jenkins method. The theoretical basis for using this method is because, in several previous studies, the ARIMA method is a powerful method compared to other methods used. Research on the ARIMA method for electrical energy needs was applied by (Wahyuni et al., 2024) with a case study in the city of Makassar, with the research results showing that the best model was ARIMA (1,0,0) with a MAPE value of 0.4735%. Research (Melantika & Wiyanti, 2024) obtained the best model for forecasting using ARIMA, which is ARIMA(2,2,2), with a forecast error value of 0.5404588%. ARIMA is also used for health data prediction, namely predicting mental disorder patients by comparing it with ARCH and GARCH, with the smallest error value using MAPE, which is 9.7% (Arifin & Wiyanti, 2024).

3. Material and Methods

This research is a quantitative type of research, namely research based on collected data, then analyzed to obtain conclusions, which are based on the analyzed data. The type of data used in this study is secondary data, which was obtained from the official website of the *Badan Pusat Statistik (BPS).* The data used in this study is data on available electricity capacity (kWh) in DKI Jakarta from 2020 to 2022. Monthly data is what is displayed on the BPS website.

In this study, data analysis was carried out by applying two forecasting methods. The two methods are Holt's Linear Exponential Smoothing and Autoregressive Integrated Moving Average (ARIMA) methods. The research flow can be seen in Figure 1. The first step that needs to be taken is to collect the data of research. Then next, import the data into the *R* software. The data must then be split into two sections. Holt's Linear apply to the first data, also known as ARIMA training data, is the first set of data. The remaining data, however, is ARIMA and Holt's Linear testing data. The prediction results obtained from calculations using Holt's Linear and ARIMA, then the prediction error is calculated using MAPE. The results of MAPE are then compared; the smallest is the best method for forecasting.

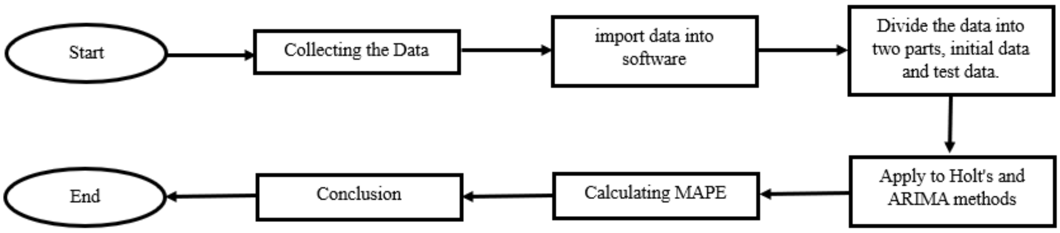

Figure 1: Research Flow Diagram

3.1. Holt's Linear Exponential Smoothing

Holt's Linear Exponential Smoothing is a method that uses level components. Holt's linear exponential smoothing method is based on two smoothing equations, one for the level and one for the trend (Makridakis & Wheelwright, 1997). The terms and formulas for the Holt's Linear Exponential Smoothing method used in this study can be seen from (Makridakis & Wheelwright, 1997).

Level

$$Lt = \alpha \cdot y_t + (1 - \alpha) \cdot (L_{t-1} + b_{t-1}) \tag{1}$$

where y_t is the actual value of the series at time t, α is the smoothing parameter ($0 \leq \alpha \leq 1$), and b_{t-1} is the trend component at time $t-1$. L_t is the level estimate at time series t.

Trend

$$b_t = \beta \cdot (L_t - L_{t-1}) + (1 - \beta) \cdot b_{t-1} \tag{2}$$

where β is the smoothing parameter for the trend ($0 \leq \beta \leq 1$). b_t is the slope estimate at time series t.

Forecast

$$F_{t+m} = L_t + b_t \cdot m \tag{3}$$

where m is the future time period m. F_{t+m} is the predicted value to m.

3.2. Autoregressive Integrated Moving Average (ARIMA)

Autoregressive integrated moving average is referred to as ARIMA. Autoregressive (AR), Integrate (I), dan Moving Average (MA) or Or it can usually be written as ARIMA (p, d, q). Each representing the autoregression, the difference, and the moving average (Makridakis & Wheelwright, 1997).

Autoregression (AR): This component describes the relationship between previous values in a time series and the current value. In other words, the AR model uses previous values as predictors of the current value.

Integration (I): This component reflects the differentiating process applied to the data to make it stationary. Stationary data has statistical characteristics that are constant over time, such as mean, variance, and covariance.

Moving Average (MA): This component involves calculating the average of previous values in a time series to predict the current value. The MA model states that the current value is affected by random noise from previous periods.

3.2.1. White Noise

White noise in the context of ARIMA refers to the residuals remaining after modeling the autoregressive and moving average patterns of a time series, which should not exhibit any predictable pattern or structure. The similarity of the ARIMA model residuals to white noise indicates that no further information can be extracted from the residuals. In ARIMA analysis, it is important to examine the residuals of the model to ensure that there are no patterns remaining in the data.

$$Y_t = c + e_t \tag{4}$$

Y_t represents the white noise, which consists of two parts: the overall level, which is c, and the error e_t or residual at time t. Any good forecasting model should have a forecast error that follows the white noise (Makridakis & Wheelwright, 1997).

3.2.2. Autocorrelation

Autocorrelation measures how closely time series values are related to each other. When positive autocorrelation occurs, high (or low) values in a time series tend to be followed by high (or low) values at certain time intervals.

Conversely, negative autocorrelation occurs when high (or low) values in a time series are followed by low (or high) values at certain time intervals. Autocorrelation plays an important role in time series analysis because it can influence the selection of statistical models and the interpretation of analysis results.

3.2.3. PACF

PACF is used to convey the correlation between time series data, helping to determine whether a decreasing lag pattern is needed in an AR (autoregressive) model or not.

3.2.4. ADF (Augmented Dickey Fuller)

The Augmented Dickey-Fuller (ADF) test is a statistical tool that helps assess whether a time series data set exhibits stationary properties. It helps identify whether there is a persistent trend or pattern by determining the presence of a unit root in the data (Makridakis & Wheelwright, 1997).

3.2.5. Durbin-Watson

The Durbin-Watson test is a statistical tool used to evaluate the presence of autocorrelation in the residuals of a regression model. The Durbin-Watson test identifies patterns in the regression residuals to determine whether autocorrelation is present. The Durbin-Watson test scale ranges from 0 to 4; values closer to 2 indicate no autocorrelation, while values closer to 0 or 4 indicate significant positive or negative autocorrelation (Makridakis & Wheelwright, 1997).

3.2.6. Mean Absolute Percentage Error (MAPE).

MAPE (Mean Absolute Percentage Error) is another useful indicator but provides relative information as opposed to the absolute information in MAE or MSE (Makridakis & Wheelwright, 1997).

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} \left|\frac{A_t - F_t}{A_T}\right| \times 100\% \tag{5}$$

where;

A_t : Actual value at the time t

F_t : The value of estimated by the model at time t

n : Number of observations in the data

4. Results And Discussion

The data used is the monthly electricity capacity ready for sale (kWh) for the last 3 years, from 2020 to 2022. Data were analyzed using R software. The R programming language to export the data is "library(readxl)" which functions to call up the desired Excel data. 23 data were used initially, spanning the period from January 2020 to November 2021. Thirteen data, spanning from December 2021 to December 2022, make up the test data. In dividing data into two parts, use the "subset()" function to perform sub setting or separate data into several parts. The data that has been divided can be seen in the graph shown in Figure 2.

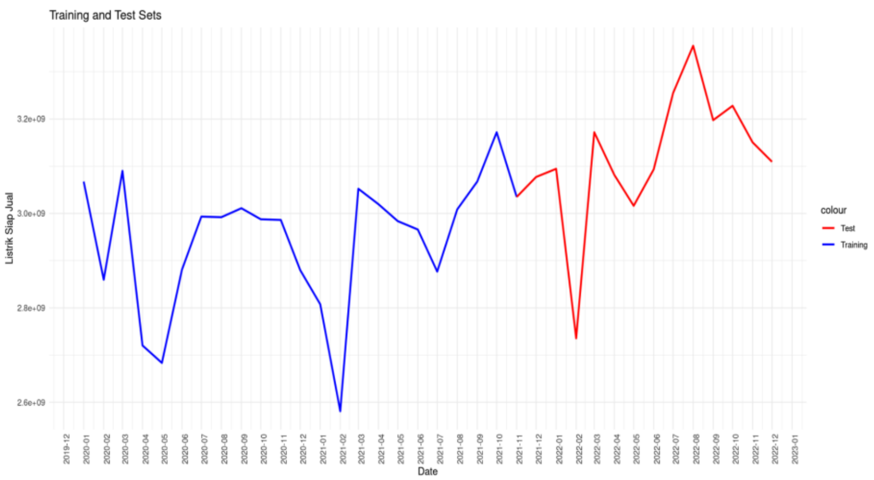

Figure 2: Plot of Training and Testing Data

4.1. Result Analysis of Holt's Linear Exponential Smoothing Method

The prediction results for the next 12 month, using Holt's, can be seen in Figure 3. The predicted value of the electricity capacity ready for sale for the next three months in January 2023 is estimated at 3,145,857,445 kWh, in February 2023 around 3,151,161,698 kWh, and in March 2023 around 3,156,465,951 kWh.

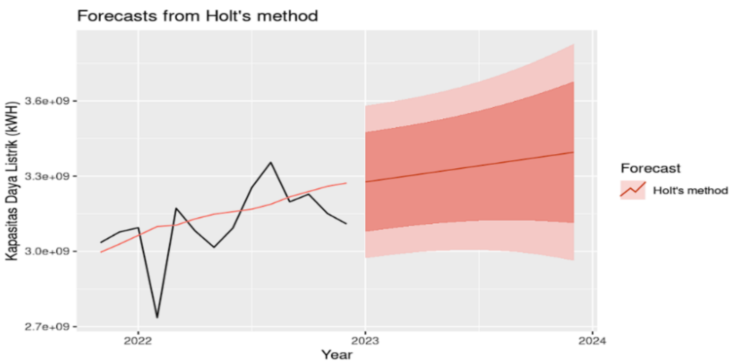

Figure 3: Plot of the 12-Year Ahead Prediction Using the HLT Method

The results of the prediction error calculation using MAPE can be seen in Figure 4. To calculate the error value using R software, "accuracy()" is used. The MAPE value obtained was 3.1%.

```
       ME      RMSE       MAE       MPE     MAPE
-33802614 130962160 96461392 -1.243435 3.185827
```

Figure 4: Accuracy Value

4.2. ARIMA Result

The first step in using the ARIMA method is to check stationarity. The results can be seen in Figure 5. To check the stationarity test, use the " library adf()." From the output results, the value of p>0.05 is obtained. This shows that the data is not stationary. So, the next data analysis is to stationaries the data. To stationarize data using the "diff()" function library. The ARIMA models formed are ARIMA (0,1,0), ARIMA(1,1,1), and ARIMA(3,1,1).

```
        Augmented Dickey-Fuller Test

data:  dataa1
Dickey-Fuller = -3.0938, Lag order = 3, p-value = 0.1476
alternative hypothesis: stationary
```

Figure 5: Check Data Stationarity

The prediction results using the ARIMA (0,1,0) model can be seen in Figure 6. The predicted value for the next 3 months is stable, namely 3,109,429,702 kWh for January, February, and March 2023.

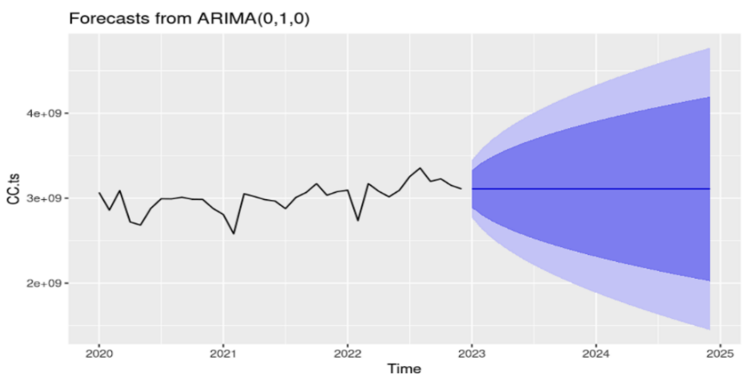

Figure 6: Prediction Results for ARIMA(0,1,0)

The prediction results of the ARIMA (1,1,1) model can be seen in Figure 7. The predicted results of the electricity capacity ready for sale for the next three months in January 2023 are estimated at 3,140,707,705 kWh, in February 2023 around 3,146,256,286 kWh, and in March 2023 around 3,147,240,580 kWh.

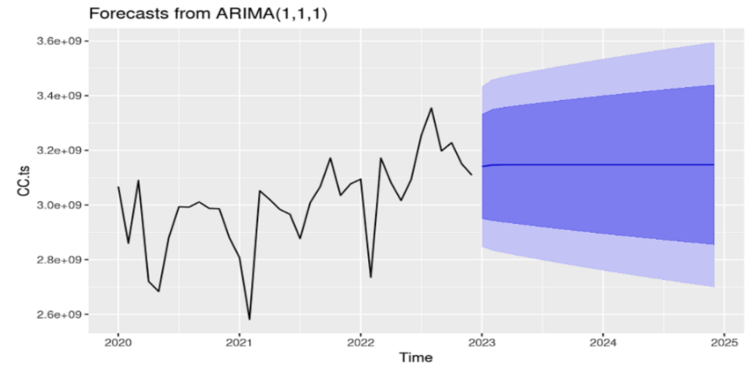

Figure 7: Prediction Results for ARIMA(1,1,1)

The prediction results using the ARIMA (3,1,1) model can be seen in Figure 8. The predicted value of the electricity capacity ready for sale for the next three months, namely January 2023, is 3,140,106,571 kWh; in February 2023 it is 3,149,746,276 kWh; and in March 2023 it reaches 3,154,664,915 kWh.

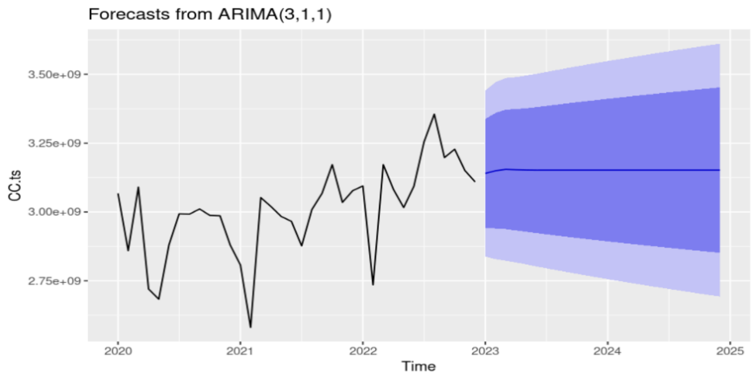

Figure 8: Prediction Results for ARIMA(3,1,1)

Next is the MAPE calculation of the ARIMA models above. The summary of the MAPE calculation results can be seen in Table 1. From the MAPE calculation results, it can be seen that the ARIMA (1,1,1) model has the smallest MAPE value among the other ARIMA models.

Table 1: MAPE of prediction

Method	MAPE	January 2023 Forecast (kWH)	February 2023 Forecast (kWH)	March 2023 Forecast (kWH)
Holt's Linear Exponential Smoothing	3.185827 %	3145857445	3151161698	315646595
ARIMA(0,1,0)	3.864937%	3109429702	3109429702	3109429702
ARIMA(1,1,1)	3.188198 %	3140707705	3146256286	3147240580
ARIMA(3,1,1)	3.124985 %	3140106571	3149746276	3154664915

5. Conclusion

From the analysis results, it was obtained that Holt's Linear Exponential Smoothing and ARIMA (1,1,1) methods are the best models with the smallest prediction error among other models. Although Holt's error rate is smaller than ARIMA (1,1,1), the difference is very small and can be said to be insignificant. In other words, Holt's Linear Exponential Smoothing and ARIMA (Autoregressive Integrated Moving Average) methods can be relied upon to predict non-seasonal trend data such as available electricity capacity (kWh).

Acknowledgments

The author would like to express special thanks to Mrs. Wiwik, my supervisor, who has helped the author in completing this research. The author would also thank to the Statistics program study of Matana University for the support in the preparation of this research. Special thanks also given to *Badan Pusat Statistik (BPS)*, which has provided the data used in the analysis.

References

Arifin, A., & Wiyanti, W. (2024). Prediction of the Number of Mental Disorder Patient Visits at Pelayanan Mentari Health Center, Kalideres District Using ARCH,. 7(2), 437–450.

Christian, D. (2022). Electricity Consumption Rises, PLN Records Revenue of IDR 25.13 Trillion in January 2022. February 21. https://web.pln.co.id/cms/media/siaran-pers/2022/02/konsumsi-listrik-naik-pln-cetak-pendapatan-rp-2513-triliun-di-januari-2022/

Evania, D. C., Wiyanti, W., & Author, C. (2024). Comparison of Holt's Linear Method and GPT-4 for Forecasting the Number of Diabetes Mellitus Death Cases in Indonesia. Journal of Mathematics, 7(1), 1–9.

Makridakis S, Wheelwright SC, H. R. (1997). 1 / the Forecasting Perspective. Forecasting Methods and Applications, 1–632.

Melantika, B. T. E., & Wiyanti, W. (2024). Comparison of the Box-Jenkins Method and ChatGPT in Predicting the Number of Tuberculosis Deaths in Indonesia. Journal of Mathematics, Computations and Statistics, 7(1), 10–19. https://doi.org/10.35580/jmathcos.v7i1.1948

Murdifi, A. (2021). PLN Electricity Consumption Rises to 146 TWh, Signaling Indonesia's Economy Starting to Recover. September 8. https://web.pln.co.id/cms/media/siaran-pers/2021/09/konsumsi-listrik-pln-naik-jadi-146-twh-sinyal-ekonomi-ri-mulai-bangkit/

Nugroho, D. A., Arifin, A., & Wiyanti, W. (2024). Effectiveness of Exponential Smoothing and Autoregressive Conditional Heteroscedasticity (ARCH) Methods in Predicting Stock Prices. Jurnal Sains Matematika Dan Statistika, 10(2), 242. https://doi.org/10.24014/jsms.v10i2.27141

Trianto, G. A. (2024). Throughout 2023, PLN Boosts Electricity Infrastructure, Now National Electricity Capacity Reaches 72,976.30 Megawatts. January 29. https://web.pln.co.id/media/siaran-pers/2024/01/sepanjang-2023-pln-genjot-infrastruktur-kelistrikan-kini-kapasitas-listrik-nasional-mencapai-72-97630-megawatt

Wahyuni, M. S., Zaki, A., Hidayat, S., & Pratama, M. I. (2024). Application of the ARIMA Method in Forecasting Electricity Demand in Makassar City. 7(2), 323–331.

Wiyanti, W., & Siregar, B. (2023). Comparison Predictions of the Demam Berdarah Dengue (DBD) using Model Exponential Smoothing: Pegel's Classification and ChatGPT. Jurnal Matematika, Statistika Dan Komputasi, 20(2), 351–359. https://doi.org/10.20956/j.v20i2.32122