



## Prediction of Hydropower Plant Electricity Production Dependence on Weather Conditions Using the SARIMAX Model

Dennis Hasnan Zulfialda<sup>1</sup>, Catur Arie Nugroho<sup>1</sup>, Hakim Luthfi Malasan<sup>2</sup>

<sup>1</sup>Program Studi S2, Sains Komputasi, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Bandung PT PLN (Persero) Jakarta, Indonesia

<sup>2</sup>KK Astronomi dan Program Studi Sains Komputasi, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Bandung Bandung, Indonesia

\*Corresponding Author: Dennis Hasnan Zulfialda

Email: : [dennis.zulfialda@pln.co.id](mailto:dennis.zulfialda@pln.co.id)



### Article Info

#### Article history:

Received 11 December 2024  
Received in revised form 17  
January 2025  
Accepted 30 January 2025

#### Keywords:

Hydropower  
SARIMAX  
El Niño  
Electricity Production  
Weather

### Abstract

Electricity production from Hydroelectric Power Plants (PLTA) that depends on the water capacity in the dam. The water capacity depends on uncertain weather conditions such as drought caused by the El Niño Storm, which has an impact on the lack of water supply that enters the hydropower turbine which is then converted into electrical energy. Accurate predictions are needed to be able to mitigate existing weather fluctuations. In this study, the SARIMAX model on electricity production data integrated with weather data for 4 years from January 2020 to December 2023. The SARIMAX model with optimal parameters ( $p=0$ ,  $d=1$ ,  $q=1$ ,  $P=1$ ,  $D=0$ ,  $Q=1$ ,  $s=12$ ) provides quite satisfactory prediction results for hydropower power production. SARIMAX obtained MSE values of 0.00101, MAE 0.0274, and RMSE 0.0318. The study also highlights the significance of accurate prediction of hydropower production, emphasizing the importance of external factors such as weather in particular El Niño. Understanding and predicting weather patterns is critical to the power generation system of hydropower in making decisions and optimizing the operation of the electricity system efficiently.

## Introduction

Electricity production from Hydroelectric Power Plants (PLTA) has a crucial role in providing sustainable and environmentally friendly energy. With the increasing demand for electricity and the importance of optimizing water resources, the ability to predict electricity production from hydropower has become very important. Accurate predictions can aid in operation planning, maintenance, and strategic decision-making to ensure stable availability of electricity (Ezeigweneme et al., 2024; Eti et al., 2007).

Weather data can have a significant role in predicting electricity production in hydropower plants which depends on the availability of water which is directly affected by weather conditions (Schaepli, 2015; Wasti et al., 2022; Monteiro et al., 2013). A weather factor that can significantly affect the water capacity is the El Niño phenomenon (Goddard & Gershunov, 2020). El Niño is a climate event characterized by abnormal warming in the central and eastern Pacific Oceans, which can result in changes in rainfall patterns in different parts of the world. In Indonesia, El Niño often causes longer dry seasons and lower rainfall, reducing river water discharge, which is the main source of hydropower (Sahu et al., 2012; Irwandi et al., 2021). Therefore, models such as SARIMAX, which takes into account seasonal data, can

be quite useful when used with weather data (Elshewey et al., 2022; Kumar & Richariya, 2024).

SARIMAX has been used in various fields such as prediction of electricity load (Wang et al., 2021), prediction of energy consumption (Manigandan et al., 2021), prediction of parking lot placement (Fokker et al., 2022), infectious diseases (He et al., 2022), arrival of international tourists (Balli et al., 2019). For the record, this study is compiled as follows, chapter 2 briefly explains the review of the literature used in this study. Chapter 3 describes the methods used. Chapter 4 provides information on the results of the research. Then for chapter 5 are conclusions and suggestions.

## SARIMAX

SARIMAX stands for *Seasonal Autoregressive Integrated Moving Average with exogenous variables*. SARIMAX is a type of time series model that combines *autoregressive*, *differencing* (D), and *moving average operators*, and takes into account seasonal factors (*seasonal ARIMA*) and exogenous variables (Sameh & Elshabrawy, 2023; Papaioannou et al., 2016). This model is useful for forecasting time series data influenced by seasonal factors and external variables, thereby improving prediction accuracy. It extends simple models such as ARIMA by adding exogenous variables and seasonal components for more accurate predictions in complex contexts.

SARIMAX Components: 1) *Seasonal (S)*: SARIMAX takes seasonality into account in the data. Seasonal patterns occur when there are regular fluctuations or patterns in data that repeat over a given period; 2) *Autoregressive (AR)*: This component captures the relationship between the variable and the previous values. It works with models based on *lagged observations*; 3) *Integrated (I)*: The "integrated" part of SARIMAX that handles differentiation. Stationary time series have statistical properties that are constant over time, making them easier to model; 4) *Moving Average (MA)*: This component describes the relationship between the variables and residual errors of the *moving average* model applied to lagging observations; 5) *Exogenous Regressors (X)*: SARIMAX allows the inclusion of exogenous variables that affect the time series being analyzed. These variables are not directly related to the time series but have an impact on it. By including exogenous regressors in the model, forecasting accuracy is improved, especially when there are external factors affecting the time series.

And the equation for SARIMAX (p,d,q)(P,D,Q,s) can be written as follows:

$$\Theta(L)^p \theta (L^s)^P \Delta^d \Delta_s^D y_t = \Phi(L)^q \phi (L^s)^Q \Delta^d \Delta_s^D \epsilon_t + \sum_{i=1}^n \beta_i x_t^i \quad (1)$$

Where,  $y_t$  is the given time series,  $p$  is the amount of time lag for the regression,  $\epsilon_t$  is the noise at the time  $t$ ,  $\beta$  is the constant,  $L$  is the regular lag operator,  $\Theta(L)^p$  is the  $p$ -order polynomial function of  $L$ ,  $q$  is the amount of time lag of the error for the regression,  $\phi$  Defined analogously with  $\Theta$ , is the integration operator where  $\Delta^d$  is the differencing order used, is the  $\Delta_s^D$  *seasonal differencing* of the time series where  $s$  is the sum of the time lag that forms a full period of the season and  $D$  is the *differencing* order applied to the seasonal lag. Then,  $n$  is the exogenous variable defined at each step of time  $t$ , denoted by for  $x_t^i \leq n$

The reasons for choosing SARIMAX are because: 1) The SARIMAX model allows for the inclusion of exogenous variables, such as weather features, which can have a significant impact on demand patterns. By incorporating these variables, the SARIMAX model can capture the influence of external factors on demand and improve prediction accuracy (Du et al., 2022); 2) The SARIMAX model is particularly suitable for capturing seasonal patterns and autocorrelation in demand data. The seasonal and autoregressive components of the SARIMAX model allow them to model and forecast demand accurately, especially when

there are recurring patterns and dependencies in the data (Alharbi & Csala, 2022); 3) SARIMAX models offer flexibility in terms of model specifications and customization. They can be tailored to specific demand patterns and adapted to include a variety of exogenous variables, allowing for more precise and tailored demand predictions (Andrianajaina et al., 2022); 4) The SARIMAX model provides interpretable results, allowing for a better understanding of the underlying factors driving demand. The coefficients and parameters of the SARIMAX model can provide insights into the relationship between demand and exogenous variables, aiding in decision-making and strategy development (Benitez et al., 2023; Vivek et al., 2024; Serrano et al., 2024).

One potential drawback is the model's sensitivity to its parameter selection, which may require careful tuning and optimization for different types of data (Ma et al., 2023). In addition, SARIMAX may be less effective when dealing with highly volatile or irregular demand patterns, as this model relies on capturing and modeling fundamental time series components, which may be difficult in such cases (Tsui, 2017). After obtaining the optimal parameters, the performance of the algorithm can be measured with several measurements to understand which method is better than the others, such as: MSE, MAE, and RMSE. *Mean Squared Error* (MSE), calculates the average of the squared variance between our prediction and the actual value. This method measures the effectiveness of our prediction model. A smaller MSE value indicates that the prediction is more in line with the actual value, while a larger MSE value indicates a greater difference between the prediction and reality, which we can see in the following equation:

$$MSE = \frac{\sum_{i=1}^n (\hat{Y}_i - Y_i)^2}{n} \quad (2)$$

The *Mean Absolute Error* (MAE) is the average of the absolute difference between your prediction and the actual value. MAE provides a simple and easy-to-understand way to measure the accuracy of your predictions. Like MSE, a lower MAE value indicates that the prediction is closer to the actual value, and the MAE equation can be seen as follows:

$$MAE = \frac{\sum_{i=1}^n |\hat{Y}_i - Y_i|}{n} \quad (3)$$

The *Root Mean Square Error* (RMSE) is a way to calculate the average size of the error in our predictions, with the square root ensuring the error metric scale matches the original value scale. Like MSE and MAE, a lower RMSE value indicates that the prediction is closer to the actual value, signaling a more accurate prediction model. The equation can be seen as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n \|Y(i) - \hat{Y}(i)\|^2}{n}} \quad (4)$$

For MAE, MSE, and RMSE, the symbol can be read as follows: Where n indicates the number of data points, indicates the observed value, and indicates the predicted value.  $Y_i \hat{Y}_i$

## EL NIÑO

One of the exogenous variables of SARIMAX that will be used is SST 3.4 *Anomaly* (*Sea Surface Temperature*) and ONI *Index* (*Oceanic Niño Index*). These 2 attributes are indicators used by scientists and meteorologists to monitor and analyze the development of ENSO (*El Niño-Southern Oscillation*) and its impact on weather and climate patterns in various parts of the world.

ENSO is a natural phenomenon in the tropical Pacific Ocean that involves the interaction between sea surface temperature, atmosphere, and wind. ENSO consists of two main phases: El Niño and La Niña, as well as a neutral phase.

El Niño occurs when sea surface temperatures in the equatorial regions of the eastern and central Pacific Ocean become warmer than normal. This has resulted in significant changes in global weather patterns, including increased rainfall in some regions and drought elsewhere. So from this explanation, the impact of El Niño that will be predicted is the one that will have an impact due to the effect of drought. This can cause a decrease in electricity production at hydropower plants due to a decrease in water discharge that enters hydropower water turbines.

The Niño 3.4 region is located in the equatorial Pacific Ocean, between latitudes 5° N and 5° S and longitudes 120° E and 170° E. SST in this region is often used as an indicator to monitor sea surface temperature anomalies related to climate change and the ENSO phenomenon. Indonesia also uses SST data in the Niño 3.4 region to monitor and analyze El Niño and La Niña climate phenomena. Although the region is not directly near Indonesia, changes in sea surface temperatures in the Niño 3.4 region have a significant impact on weather and climate patterns in Indonesia and the surrounding region. Therefore, SST data from the Niño 3.4 region is very important for weather and climate monitoring and prediction in Indonesia.

Historically, scientists have classified El Niño intensity based on SST anomalies that exceed a pre-selected threshold in a given region of the equatorial Pacific Ocean. The most commonly used region is the Niño 3.4 region, and the most commonly used threshold is a positive SST excess of normal greater than or equal to +0.5°C. Because it covers the western half of the equatorial cold tongue region, it provides a good measure of important changes in the SST and SST gradients resulting in changes in the pattern of deep tropical convection and atmospheric circulation. The criteria often used to classify El Niño episodes are five SST anomalies running on average for 3 consecutive months exceeding that threshold (Boegehold et al., 2023).

The ONI Index is one of the main indicators to monitor and measure the severity of El Niño and La Niña phenomena. The ONI Index measures sea surface temperature anomalies (SST) in the Niño 3.4 region, which is the region in the equatorial Pacific Ocean between latitudes 5° N and 5° S and longitudes 120° E and 170° E. ONI is calculated as the average monthly sea surface temperature anomaly within the Niño 3.4 region over a period of three consecutive months. A positive ONI indicates El Niño conditions (sea surface temperatures are warmer than normal), while a negative ONI indicates La Niña conditions (sea surface temperatures are colder than normal). ONI is used by meteorological and climate agencies around the world to predict and manage the impacts of El Niño and La Niña on global weather and climate

Studies have shown that a necessary condition for the development and persistence of deep convection (increased cloudiness and rainfall) in the Tropics is that the local sea surface temperature should be 28°C or higher. After the deep convection pattern has changed due to the SST Anomaly, tropical and subtropical atmospheric circulation adapts to the new tropical warming pattern, resulting in anomalous rainfall and temperature patterns that extend far beyond the equatorial Pacific Ocean. An SST anomaly of +0.5°C in the Niño 3.4 region is enough to reach this threshold from late March to mid-June. During the rest of the year, larger SST anomalies, up to +1.5°C in November-December-January, are required to reach a threshold that favors persistent deep convection in the region[10].

## Methods

In this study, the data used was data on hydropower plants in one of the cities of South Sulawesi. The flowchart can be seen in Figure 1.

For the SARIMAX simulation process in this study, the python programming language was used in the Jupyter Notebook application.

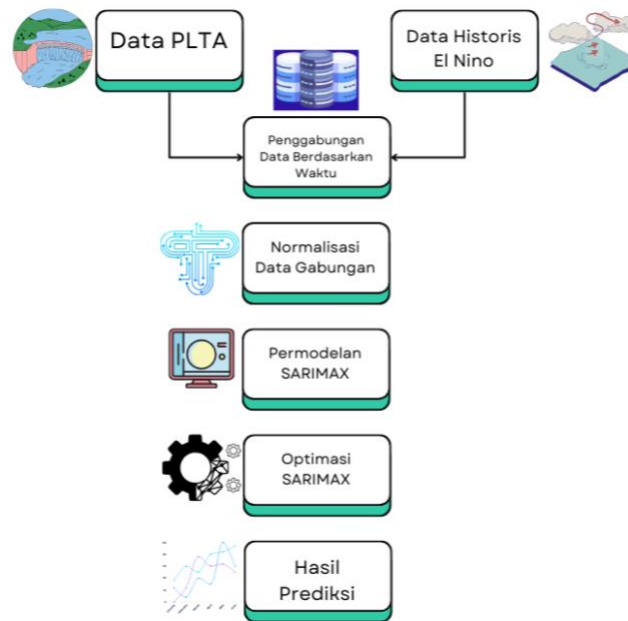


Figure 1. Flow of the research process

## Results and Discussion

From Figure 1, a detailed explanation of each test step can be explained as follows: 1) Preparing data from hydropower plants obtained from daily data which is then done on average to get monthly values. The daily data taken is data from January 2020 to December 2023. The variable data taken were electricity production in MWh, dam data in the form of elevation, discharge, runoff, and inflow. The data can be seen in Table 1; 2) Then, weather data related to SST 3.4 Anomaly and Oceanic Niño Index (ONI) values obtained from the National Center for Environmental Information (NCEI) website were prepared; 3) Then the hydropower data and weather data above are combined into one dataset which can be seen in Table 2.

Tabel 1. PLTA Data

Tahun	Bulan	Produksi (MWh)	Elevasi (m)	Discharge (m <sup>3</sup> /s)	Limpasan (m <sup>3</sup> /s)	Inflow (m <sup>3</sup> /s)
2020	2020-01-01	79540.90	615.45	39.31	26.61	66.69
2020	2020-02-01	82403.30	615.01	42.95	20.59	64.11
2020	2020-03-01	87843.40	615.44	42.85	23.59	66.38
2020	2020-04-01	87437.50	615.46	44.03	23.13	91.51
2020	2020-05-01	90510.60	615.48	44.03	46.46	92.39
...	...	...	...	...	...	...
2023	2023-08-01	57053.00	615.09	28.54	0.00	28.23
2023	2023-09-01	45130.00	615.16	23.08	0.00	25.02
2023	2023-10-01	44693.00	615.17	22.54	0.00	25.13
2023	2023-11-01	59668.00	615.21	31.19	0.00	51.10
2023	2023-12-01	80109.00	615.41	39.16	12.10	50.12

Table 2. Hydropower and weather dataset

Tahun	Bulan	Produksi (MWh)	Elevasi (m)	Discharge (m <sup>3</sup> /s)	Limpasan (m <sup>3</sup> /s)	Inflow (m <sup>3</sup> /s)	SST 3.4 Anomaly	ONI
2020	2020-01-01	79540.90	615.45	39.31	26.61	66.69	0.52	0.50
2020	2020-02-01	82403.30	615.01	42.95	20.59	64.11	0.50	0.50
2020	2020-03-01	87843.40	615.44	42.85	23.59	66.38	0.41	0.40

2020	2020-04-01	87437.50	615.46	44.03	46.13	91.51	0.21	0.20
2020	2020-05-01	90510.60	615.48	44.03	46.46	92.39	0.00	0.00
...	...	...	...	...	...	...	...	...
2023	2023-08-01	57053.00	615.09	28.54	0.00	28.23	1.30	1.30
2023	2023-09-01	45130.00	615.16	23.08	0.00	25.30	1.53	1.50
2023	2023-10-01	44693.00	615.17	22.54	0.00	25.13	1.90	1.90
2023	2023-11-01	59668.00	615.21	31.19	0.00	51.10	1.90	1.90
2023	2023-12-01	80109.00	615.41	39.16	12.10	52.18	1.99	2.00

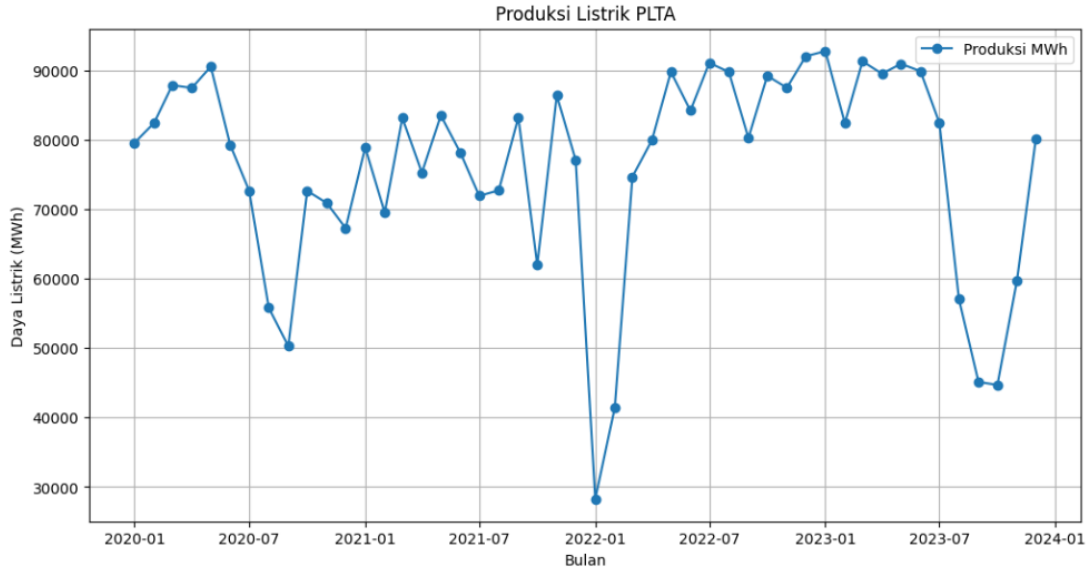


Figure 2. Electricity Production (MWh) data plot

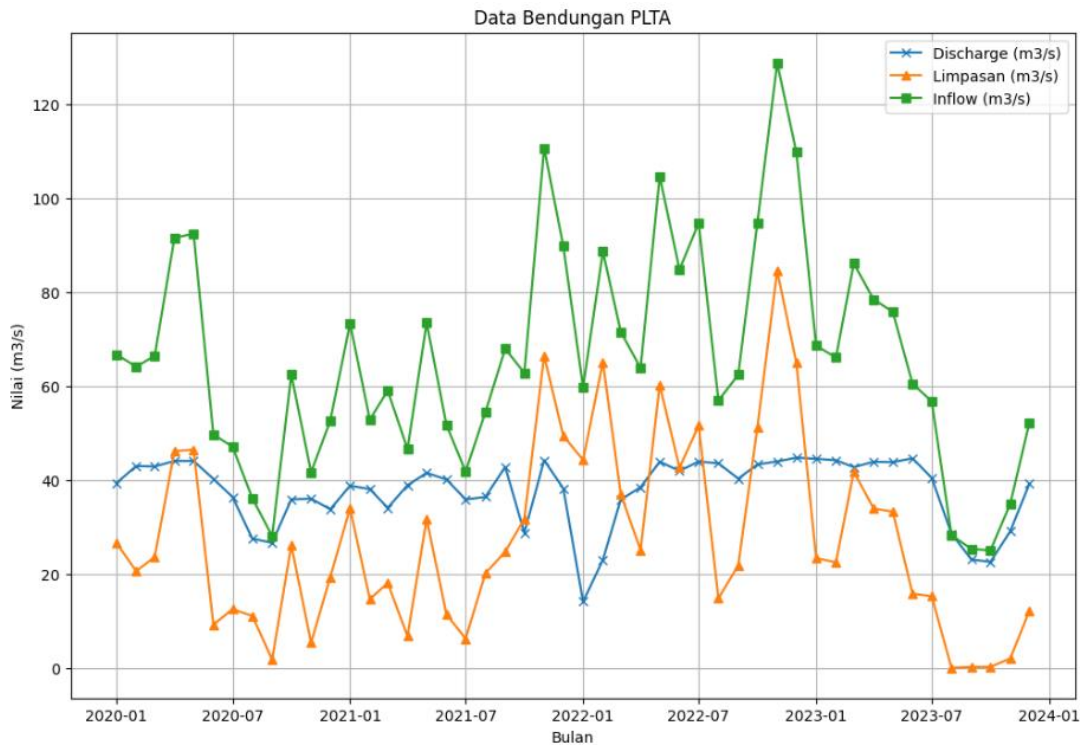


Figure 3. Runoff, discharge, and inflow data plots (m3/s)

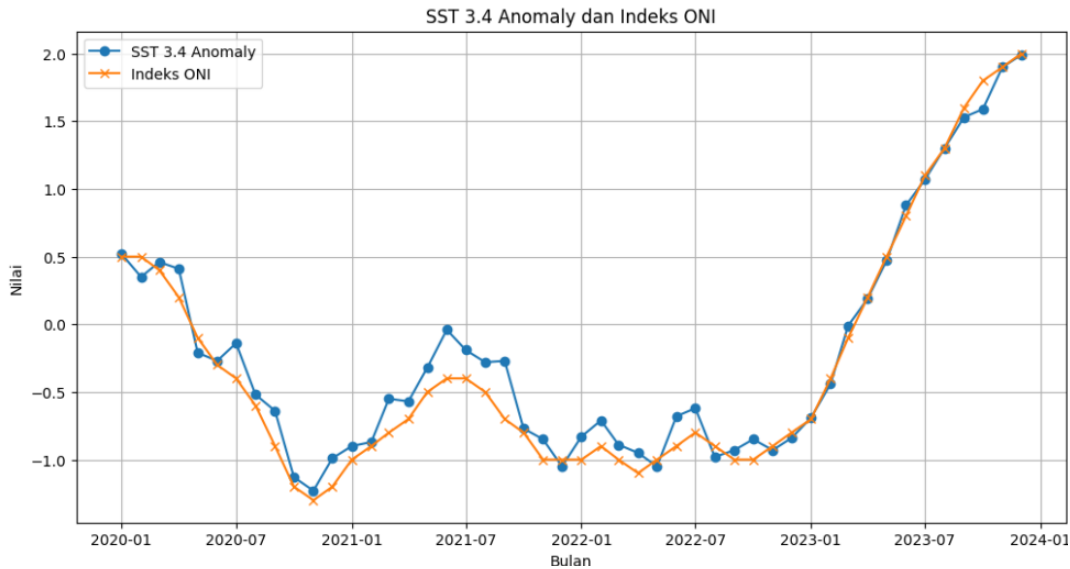


Figure 4. SST Anomaly 3.4 data plot and ONI Index

Table 3. SARIMAX parameter values

Parameter	Definition	Mark
ppp	The number of lag observations included in the model (autoregressive order)	0
ddd	Degree differencing to make data stationary	1
qqq	Ukuran rentang pengamatan rata-rata ( <i>order moving average</i> )	1
PPP	The number of lag observations included in the seasonal component (seasonal autoregressive order)	1
DDD	The number of times seasonal observations are differenced to achieve seasonal stationarity.	0
QQQ	Seasonal moving average observation size (seasonal moving average order)	1
sss	The number of time steps in one seasonal period (annual data)	12

For datasets, it is converted in the form of a time graph. The results of the graph plot can be seen in Figure 2-4. From Figure 2, it can be seen that electricity production is in Megawatt hours. From Figure 3, it can be seen that there are fluctuations in elevation, runoff, discharge, and inflow data as a result of weather effects. From Figure 4, it can be seen for SST Anomaly 3.4 data and the ONI Index. After the dataset is obtained, a correlation matrix calculation between variables is carried out using the '`corr()`' function from the '`pandas`' library that has the python programming language in the jupyter notebook application.

From Figure 5, it can be seen that the weather data represented by SST 3.4 Anomaly and the ONI Index have weak negative correlation matrix values of -0.15 and -0.14 with electricity production in hydropower data. This indicates that when 1 variable increases, the other variables tend to decrease slightly and vice versa. However, the correlation matrix between SST 3.4 Anomaly and the ONI Index shows a strong positive value of around 0.9 which means that both variables have a positive tendency where when 1 variable increases, the other variables tend to increase with the same proportion. Then for the correlation matrix values between electricity production and runoff, inflow and discharge respectively are 0.3, 0.5, and 0.9. So from this, it can be concluded that the variables of electricity production, discharge, runoff, inflow, SST 3.4 Anomaly and ONI Index are features. From this, electricity production is the target, and other features are exogenous data. 1) From the existing datasets, preprocessing is carried out first using the '`Min-Max Scaler`' library because the range of datasets is quite far. This aims to ensure that the features are in the same range so that it can

improve the performance of the SARIMAX algorithm which is sensitive to feature scale; 2) Then the training data consists of 36 months of the first data, namely data from January 2020 to December 2022. Then for the remaining testing data, 12 months, namely data from January 2023 to December 2023; 3) SARIMAX optimization was carried out to obtain optimal parameter values for (p,d,q) and (P,D,Q) using the 'itertools' library with the smallest RMSE value referring to the SARIMAX parameter dataset, the results are seen in Table 3. The best parameters were obtained namely  $p=0$ ,  $d=1$ ,  $q=1$ ,  $P=1$ ,  $D=0$ ,  $Q=1$  with an RMSE value of 0.03182.

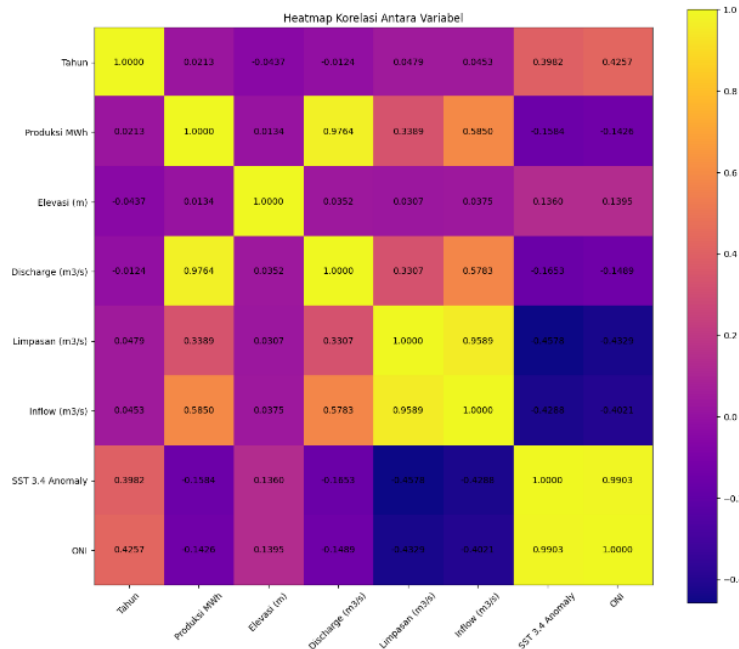


Figure 5. Dataset correlation matrix

From the predetermined SARIMAX parameters, it can be seen that the results of SARIMAX predictions made with the python programming language can be seen in Figure 5 below.

RMSE: 0.031819361456859624  
 MAE: 0.02743405217278447  
 MSE: 0.0010124717635222838

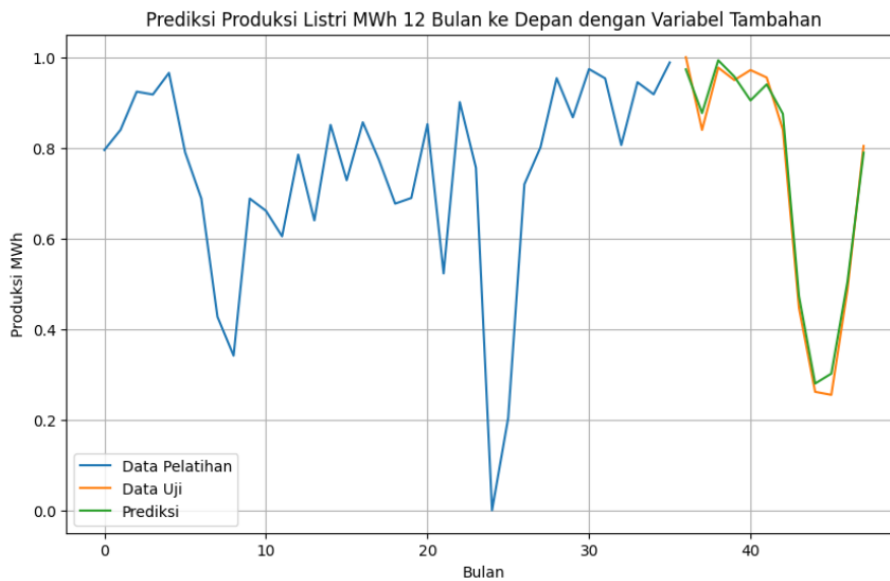


Figure 6. Plot Graph of SARIMAX Electricity Production Prediction

In the results of the simulation plot of prediction of hydropower production with exogenous power generation that has been determined predetermined, it can be seen that the training data and test data in the last 12 months show a visualization that is quite close. This is supported by the results of the model evaluation with an MSE value of 0.00101, MAE of 0.0274 and RMSE of 0.03182. From these several things, it can be concluded that the SARIMAX model that has been made is suitable for predicting hydropower power production that takes into account the condition of the hydropower dam and the weather, especially the El Niño phenomenon.

## Conclusion

Hydropower power production, hydropower dam conditions and weather conditions, especially the monthly El Niño phenomenon, can be used together with SARIMAX ( $p=0$ ,  $d=1$ ,  $q=1$ ,  $P=1$ ,  $D=0$ ,  $Q=1$ ,  $s=12$ ) for electricity production prediction with MSE of 0.00101, MAE of 0.0274 and RMSE of 0.03182. As has been analyzed, the correlation between weather features, namely SST 3.4 Anomaly and the ONI Index has a low correlation matrix value, but with growing data, weather data can be used as an exogenous feature in the SARIMAX algorithm. In the next research, other features can be added to get more optimal results such as air or land temperature index, wind speed and so on. It can also be added to the dataset range so that the modeling made is more up-to-date. We sincerely express our gratitude to PT PLN (Persero) for funding the Master's course process at the Bandung Institute of Technology through the Distance Learning Program. This is related to the evaluation of the operating system of one of the hydropower plants in South Sulawesi which is always struggling to mitigate every El Niño phenomenon that occurs.

## References

- Alharbi, F. R., & Csala, D. (2022). A seasonal autoregressive integrated moving average with exogenous factors (SARIMAX) forecasting model-based time series approach. *Inventions*, 7(4), 94. <https://doi.org/10.3390/inventions7040094>
- Andrianajaina, T., Razafimahefa, D. T., Rakotoarijaina, R., & Haba, C. G. (2022). Grid search for SARIMAX parameters for photovoltaic time series modeling. *Global Journal of Energy Technology Research Updates*, 9, 87–96. <https://doi.org/10.15377/2409-5818.2022.09.7>
- Balli, H. O., Tsui, W. H. K., & Balli, F. (2019). Modelling the volatility of international visitor arrivals to New Zealand. *Journal of Air Transport Management*, 75, 204–214. <https://doi.org/10.1016/j.jairtraman.2018.10.002>
- Benitez, I. B., Ibañez, J. A., Lumabad, C. I. I. D., Cañete, J. M., & Principe, J. A. (2023). Day-Ahead Hourly Solar Photovoltaic Output Forecasting Using SARIMAX, Long Short-Term Memory, and Extreme Gradient Boosting: Case of the Philippines. *Energies*, 16(23), 7823. <https://doi.org/10.3390/en16237823>
- Boegehold, A. G., Burtner, A. M., Camilleri, A. C., Carter, G., DenUyl, P., Fanslow, D., Fyffe Semenyuk, D., Godwin, C. M., Gossiaux, D., & Johengen, T. H. (2023). Routine monitoring of western Lake Erie to track water quality changes associated with cyanobacterial harmful algal blooms. *Earth System Science Data Discussions*, 2023, 1–39. <https://doi.org/10.5194/essd-15-3853-2023>
- Du, M., Zhu, H., Yin, X., Ke, T., Gu, Y., Li, S., Li, Y., & Zheng, G. (2022). Exploration of influenza incidence prediction model based on meteorological factors in Lanzhou, China, 2014–2017. *Plos One*, 17(12), e0277045. <https://doi.org/10.1371/journal.pone.0277045>

- Elshewey, A. M., Shams, M. Y., Elhady, A. M., Shohieb, S. M., Abdelhamid, A. A., Ibrahim, A., & Tarek, Z. (2022). A Novel WD-SARIMAX model for temperature forecasting using daily delhi climate dataset. *Sustainability*, *15*(1), 757. <http://dx.doi.org/10.3390/su15010757>
- Eti, M. C., Ogaji, S. O. T., & Probert, S. D. (2007). Integrating reliability, availability, maintainability and supportability with risk analysis for improved operation of the Afam thermal power-station. *Applied Energy*, *84*(2), 202-221. <https://doi.org/10.1016/j.apenergy.2006.05.001>
- Ezeigweneme, C. A., Nwasike, C. N., Adefemi, A., Adegbite, A. O., & Gidiagba, J. O. (2024). Smart grids in industrial paradigms: a review of progress, benefits, and maintenance implications: analyzing the role of smart grids in predictive maintenance and the integration of renewable energy sources, along with their overall impact on the industri. *Engineering Science & Technology Journal*, *5*(1), 1-20. <http://dx.doi.org/10.51594/estj.v5i1.719>
- Fokker, E. S., Koch, T., van Leeuwen, M., & Dugundji, E. R. (2022). Short-term forecasting of off-street parking occupancy. *Transportation Research Record*, *2676*(1), 637–654. <https://doi.org/10.1177/03611981211036373>
- Goddard, L., & Gershunov, A. (2020). Impact of El Niño on weather and climate extremes. *El Niño Southern Oscillation in a changing climate*, 361-375. <https://doi.org/10.1002/9781119548164.ch16>
- He, J., Wei, X., Yin, W., Wang, Y., Qian, Q., Sun, H., Xu, Y., Soares Magalhaes, R. J., Guo, Y., & Zhang, W. (2022). Forecasting scrub typhus cases in eight high-risk counties in China: evaluation of time-series model performance. *Frontiers in Environmental Science*, *9*, 783864. <https://doi.org/10.3389/fenvs.2021.783864>
- Irwandi, H., Rosid, M. S., & Mart, T. (2021). The effects of ENSO, climate change and human activities on the water level of Lake Toba, Indonesia: a critical literature review. *Geoscience Letters*, *8*(1), 21. <https://doi.org/10.1186/s40562-021-00191-x>
- Kumar, D., & Richariya, P. (2024). Leveraging SARIMAX for Accurate Rainfall Forecasting in India. *International Journal of Innovations in Science, Engineering And Management*, 37-47.
- Ma, J., Jia, Y., & Lv, B. (2023). A hybrid forecasting model for monthly electricity consumption combining prophet and SARIMAX models. *Third International Conference on Advanced Algorithms and Neural Networks (AANN 2023)*, *12791*, 626–631. <https://doi.org/10.1117/12.3005085>
- Manigandan, P., Alam, M. D. S., Alharthi, M., Khan, U., Alagirisamy, K., Pachiyappan, D., & Rehman, A. (2021). Forecasting natural gas production and consumption in United States-evidence from SARIMA and SARIMAX models. *Energies*, *14*(19), 6021. <https://doi.org/10.3390/en14196021>
- Monteiro, C., Ramirez-Rosado, I. J., & Fernandez-Jimenez, L. A. (2013). Short-term forecasting model for electric power production of small-hydro power plants. *Renewable Energy*, *50*, 387-394. <http://dx.doi.org/10.1016/j.renene.2012.06.061>
- Papaioannou, G. P., Dikaiakos, C., Dramountanis, A., & Papaioannou, P. G. (2016). Analysis and modeling for short-to medium-term load forecasting using a hybrid manifold learning principal component model and comparison with classical statistical models (SARIMAX, Exponential Smoothing) and artificial intelligence models (ANN, SVM): The case of Greek electricity market. *Energies*, *9*(8), 635. <https://doi.org/10.3390/en9080635>

- Sahu, N., Behera, S. K., Yamashiki, Y., Takara, K., & Yamagata, T. (2012). IOD and ENSO impacts on the extreme stream-flows of Citarum river in Indonesia. *Climate dynamics*, 39, 1673-1680.
- Sameh, B., & Elshabrawy, M. (2023). Seasonal Autoregressive Integrated Moving Average for Climate Change Time Series Forecasting. *American Journal of Business and Operations Research*, 8(2), 25-35. <http://dx.doi.org/10.54216/AJBOR.080203>
- Schaeffli, B. (2015). Projecting hydropower production under future climates: a guide for decision-makers and modelers to interpret and design climate change impact assessments. *Wiley Interdisciplinary Reviews: Water*, 2(4), 271-289. <https://doi.org/10.1002/wat2.1083>
- Serrano, A. L. M., Rodrigues, G. A. P., Martins, P. H. D. S., Saiki, G. M., Filho, G. P. R., Gonçalves, V. P., & Albuquerque, R. D. O. (2024). Statistical comparison of time series models for forecasting brazilian monthly energy demand using economic, industrial, and climatic exogenous variables. *Applied Sciences*, 14(13), 5846. <https://doi.org/10.3390/app14135846>
- Tsui, K. W. H. (2017). Does a low-cost carrier lead the domestic tourism demand and growth of New Zealand? *Tourism Management*, 60, 390–403. <https://doi.org/10.1016/j.tourman.2016.10.013>
- Vivek, K., Deepti, A. R., & Basith, F. (2024, June). Forecasting the Consumer Price Index using SARIMAX Modeling. In *2024 Second International Conference on Inventive Computing and Informatics (ICICI)* (pp. 477-483). IEEE. <https://doi.org/10.1109/ICICI62254.2024.00083>
- Wang, S., Song, A., & Qian, Y. (2021). *Electricity Demand Prediction of Smart Grid in Smart Cities Using K-means Algorithm*.
- Wasti, A., Ray, P., Wi, S., Folch, C., Ubierna, M., & Karki, P. (2022). Climate change and the hydropower sector: A global review. *Wiley Interdisciplinary Reviews: Climate Change*, 13(2), e757. <https://doi.org/10.1002/wcc.757>