Spectral Analysis and SARIMA Model for Forecasting Indian Ocean Dipole (IOD) and Rainfall in West Aceh Regency

Nuwairy El Furqany*, Miftahuddin, Ichsan Setiawan
Departement of Statistics, Faculty of Mathematics and Sciences, Syiah Kuala University, Banda Aceh, Indonesia

*nuwairy@gmail.com

Abstract. Indian Ocean Dipole (IOD) in the Indian Ocean is one of the climatic phenomena that affects climate conditions in western Indonesia. West of Aceh Regency is one of the districts in Aceh Province which borders and faces the Indian Ocean directly, so it has a significant impact due to climate events caused by the IOD phenomenon. Analysis for examines the occurrence of climate phenomena is needed to minimize the adverse effects due to the IOD phenomenon. Spectral analysis by calculating the periodogram and the SARIMA model for forecasting historical data. The data used are monthly data on the Indian Ocean IOD index and rainfall in West of Aceh in the period January 2010 to December 2019. The purpose of this study was to determine the periodicity of the IOD index data and rainfall, get a model and obtain the results of forecasting the IOD index and rainfall, and see the relationship of IOD to rainfall in West of Aceh. The results obtained in the spectral analysis are the characteristics of changes in the IOD index tend to increase or decrease in the IOD index every 17 months and the characteristics of changes in rainfall in West of Aceh tend to increase or decrease in rainfall every 6 months. The best model for predicting IOD index data is the SARIMA model $(2,1,3)(0,1,1)^{17}$. Forecasting for 2020 and 2021 tend to show a downward trend in 2020 and an upward trend in 2021. While the best model for predicting West of Aceh rainfall data is the SARIMA model $(2,0,2)(0,1,1)^{6}$. Forecasting for 2020 and 2021 tend to experience high rainfall in April and October for each year. The relationship obtained by the IOD index to rainfall in West of Aceh is when the IOD index is in a positive phase, rainfall in the West of Aceh region tends to decrease. Conversely, when the IOD index is in the negative phase, rainfall in the West of Aceh region tends to increase.

1. Introduction
The climate phenomenon in the western region of Sumatra, Indonesia and the western part of the Indian Ocean is often referred to as the Indian Ocean Dipole (IOD). At the time of the positive IOD phenomenon, the rainfall intensity in the western part of Indonesia was relatively low, while when the IOD was negative the rainfall intensity in the western part of Indonesia was relatively high \[^{[1]}\]. The phenomenon of drought on one side of the Indian Ocean and heavy rains on the other is a direct impact felt by this IOD phenomenon. When the IOD phase is positive, not only Indonesia will be affected by the drought, but other countries directly adjacent to the eastern part of the Indian Ocean, such as Australia will also be affected \[^{[2]}\]. West Aceh is one of the districts in Aceh Province which borders and faces the Indian Ocean so it often gets extreme climate impacts such as drought and floods.
Research methods of spectral analysis have been used in various fields of science, such as meteorology and climatology. An example is research on applying spectral analysis to determine the periodicity of the data by looking for periodograms and forecasting using the Seasonal Autoregressive Integrated Moving Average (SARIMA) model \[^{[10]}\]. Analysis for examines the occurrence of climate phenomena is needed to minimize the adverse effects due to the IOD phenomenon. In this study, we have objective to obtain the forecasting rainfall and IOD phenomena in the Indian Ocean by using method spectral analysis and SARIMA model.

2. Material and Methods
The data used in this study are secondary data, namely monthly IOD index data (°C) for the period January 2010 to December 2019 sourced from the Japan Agency for Marine-Earth Science Technology (JAMSTEC) \[^{[4]}\] and monthly rainfall data (mm / month) in the West Aceh Regency region for the period January 2010 to December 2019 sourced from the Meteorology, Climatology and Geophysics Agency (BMKG) \[^{[5]}\]. The variables and coordinates of the area in this study are shown in Table 1:

| Variable    | Area              | Coordinates                                      | Unit    |
|-------------|-------------------|--------------------------------------------------|---------|
| Dipole Mode | West Box          | (50°E – 70°E, 10°S – 10°N)                        | °C      |
| Index       | East Box          | (90°E – 110°E, 10°S)                             |         |
| Rainfall    | West Aceh Regency | (04° 06’ - 04° 47’ LU dan 95° 52’ - 96° 30’ BT)  | mm/month|

The procedures for this research are as follows:
1. Presenting IOD index data in the Indian Ocean and rainfall in West Aceh Regency in the form of a time series plot.
2. Perform spectral analysis to determine the periodicity of the IOD index time series data and rainfall obtained by calculating the periodogram. The main purpose of periodogram analysis is to find hidden periodicity of time series data.
3. Estimating a tentative (provisional) model through the ACF and PACF plots on the IOD index data and stationary rainfall relative to the mean and variance. Then proceed with the test of significance and model diagnostics \[^{[7]}\].
4. Selection of the IOD index model and the best rainfall using AIC, RMSE, and MAPE values with decision making based on minimum values \[^{[8]}\].
5. Conduct forecasts for the IOD index and rainfall for the period January 2020 to December 2021.

2.1. Spectral Analysis
The application of spectral analysis in determining the time series data forecasting model is by knowing the periodicity of the data using the periodogram plot. The period obtained is then used to determine the period of a time series data model \[^{[6]}\]. The periodogram equation can be written as follows:

\[
I(\omega) = \begin{cases} 
    n_{a_0^2}, & k = 0 \\
    \frac{n}{2}(a_k^2 + b_k^2), & k = 1, \ldots, \left[\frac{n-1}{2}\right] \\
    n_{a_{n/2}^2}, & k = 0
\end{cases}
\]

(1)
where $a_k$ and $b_k$ is the Fourier coefficient which is written as follows:

$$
a_k = \begin{cases} 
\frac{1}{n} \sum_{t=1}^{n} Z_t \cos \omega_k t, & k = 0 \text{ and } k = \frac{n}{2} \text{ if } n \text{ is even}, \\
\frac{2}{n} \sum_{t=1}^{n} Z_t \cos \omega_k t, & k = 1, 2, ..., \left\lfloor \frac{n-1}{2} \right\rfloor \text{ if } n \text{ is odd},
\end{cases}
$$

(2)

$$
b_k = \frac{2}{n} \sum_{t=1}^{n} Z_t \sin \omega_k t, \quad k = 1, 2, ..., \left\lfloor \frac{n-1}{2} \right\rfloor,
$$

(3)

where

$$
\omega_k = \frac{2\pi k}{n}, \quad k = 0, 1, 2, ..., \left\lfloor \frac{n}{2} \right\rfloor
$$

(4)

2.2. SARIMA Model

Autoregressive Integrated Moving Average (ARIMA) is a method first introduced by Box and Jenkins (1976) on forecasting univariate time series data. The ARIMA model is derived from the AR (Autoregressive) model, the MA (Moving Average) model, and the ARMA (Autoregressive Moving Average) model a combination of the AR and MA models. Then Box and Jenkins developed the ARIMA model on data that has a seasonal component, namely the Seasonal Autoregressive Integrated Moving Average (SARIMA) model [9]. In general, the SARIMA model is denoted as follows:

$$ARIMA(p, d, q)(P, D, Q)_s
$$

(5)

3. Results and Discussion

3.1. Plot Time Series

The data time series plot can be seen in Figure 1:

![Figure 1. Time series plot of IOD index and rainfall](image)

Figure 1 shows that the time series plot of the IOD index data for the period January 2010 to December 2019 is not stable every month. In the IOD index plot there is a trend pattern, namely a downward trend from around 2013 to 2017 and an upward trend from around 2017 to 2019. It can also be seen from the plot that the lowest IOD index is around the end of 2016 and the highest IOD index is around the end in 2019. Then in the time series plot the rainfall data has an unstable pattern every month and experiences the rainy season at the end of each year.

3.2. Spectral Analysis

The periodogram graph using spectral analysis can be seen in Figure 2:
Figure 2 on the IOD index periodogram chart for the period January 2010 to December 2019 shows that, the highest frequency peak is at the 13th frequency $\omega_7 = \frac{2\pi(7)}{60} = 0.3665$ and periodogram values $I(\omega_7) = 3.7059$. The resulting frequency is related to the period $P = \frac{2\pi}{\omega_7} = 17.1429$ or can be rounded up to 17 cycles per month. This shows that the characteristics of the changes in the IOD index tend to experience an increase or decrease in the IOD index every 17 months.

Meanwhile, on the graph of the periodogram of rainfall for the period January 2010 to December 2019 shows that, the highest frequency peak is at the 20th frequency $\omega_{20} = \frac{2\pi(20)}{60} = 1.0472$ and periodogram values $I(\omega_{20}) = 166052.0553$. The resulting frequency relates to the period $P = \frac{2\pi}{\omega_{20}} = 6$. This shows that the characteristics of changes in rainfall in West Aceh Regency tend to experience an increase or decrease in rainfall every 6 months.

3.3. SARIMA Model
The initial stage in determining the time series model is by examining the stationarity of the IOD index data and the rainfall for the period January 2010 to December 2019 both towards the mean and the variance. Then proceed with the identification of models and forecasting.

3.3.1. Data Stationarity Test
The visualization of the Box-Cox test plot after the transformation of the IOD index and rainfall data can be seen in Figure 3 below:

![Box-Cox test plot](image)
Based on the Box-Cox test plot after the transformation of the IOD index data and rainfall in Figure 3, it can be seen that the $\lambda$ value obtained is 1. This indicates that the IOD index data is stationary to variance after data transformation using $\sqrt{X_t + 1.5}$ and the rainfall data is stationary to the variance after transforming the data using $\log(10(X_t))$. The results of the data stationary test against the mean based on the number of differencing (differences) performed are shown in Table 2 below:

| Measuring instrument | Differencing | IOD Index P-value | Rainfall P-value |
|-----------------------|--------------|-------------------|------------------|
| ADF test              | 0            | 0.1328            | 0.01             |
|                       | 1            | 0.01              | -                |

The results of the stationary test on the mean in Table 2 show that the IOD index data is stationary after one-time differencing ($d = 1$). This is because the $p$-value ($0.01 < \alpha (0.05)$, so it can reject $H_0$ and stationary data. Furthermore, the rainfall data is stationary without differencing. This is because the $p$-value ($0.01 < \alpha (0.05)$, so it can reject $H_0$ and stationary data.

3.3.2. Model Identification
The identification of the tentative model is the stage of estimating the order of model parameters that are formed based on ACF and PACF plots on the IOD index data and the rainfall that is stationary to the mean and variance. The visualization of non-seasonal ACF and PACF plots for IOD index data and rainfall is as follows:

![Figure 4. Plots of non-seasonal ACF and PACF data for differencing IOD index (d = 1) and rainfall without differencing](image-url)
Figure 4 shows that the ACF plot decreases exponentially (tails off) and PACF plot decreases exponentially (tails off). So that the IOD index model is an ARMA model with orders (p, q) after one time differencing (d = 1) and the rainfall model is an ARMA model with orders (p, q) without differencing. Then the ACF and PACF plots in Figure 4 have not been stationary to seasonal factors. This is because both ACF and PACF plots on the IOD index and rainfall data form a repeating pattern at certain lags. To overcome the seasonal factor in the data, a seasonal differencing process was carried out (D = 1). The visualization of ACF and PACF plots after seasonal differencing is as follows:

Figure 5. Plot of ACF and PACF for seasonal differencing (D = 1) data IOD index and rainfall

Figure 5 shows that the ACF plot decreases exponentially (tails off) and PACF plot decreases exponentially (tails off). So that the IOD index model is the SARMA model with orders (P, Q) and the rainfall model is the SARMA model with orders (P, Q). To determine the order parameter by manually testing each order (trial and error). The results of the parameter significance test, white noise test, and residual normality test are shown in Table 3:

Table 3. Results of Parameter Significance Test, White Noise Test, and Residual Normality Test

| Model   | SARIMA             | Parameter Significance Test | White Noise Test | Residual Normality Test       |
|---------|--------------------|-----------------------------|-----------------|-------------------------------|
| IOD Index | (1,1,0)(1,1,0)\(^{17}\) | Significant                  | White Noise     | Not Distributed Normally      |
|         | (2,1,3)(0,1,1)\(^{17}\) | Significant                  | White Noise     | Not Distributed Normally      |
|         | (2,1,3)(1,1,0)\(^{17}\) | Significant                  | White Noise     | Not Distributed Normally      |
| Rainfall | (0,0,0)(2,1,6)\(^{6}\) | Significant                  | White Noise     | Distributed Normally          |
|         | (1,0,1)(2,1,0)\(^{6}\) | Significant                  | White Noise     | Distributed Normally          |
|         | (2,0,2)(0,1,1)\(^{6}\) | Significant                  | White Noise     | Distributed Normally          |
The results of the parameter significance test in Table 3 show that all parameters for the IOD index model and rainfall are significant at an error rate of 5%. White noise test results on Table 3 show that all models SARIMA (1,1,1)(1,1,0)\(^1\), SARIMA (2,1,3)(0,1,1)\(^1\) dan SARIMA (2,1,3)(1,1,0)\(^1\) qualify for white noise. Furthermore, the rainfall model shows that models SARIMA (0,0,0)(2,1,0)\(^6\), SARIMA (1,0,1)(2,1,0)\(^6\), dan SARIMA (2,0,2)(0,1,1)\(^6\) qualify for white noise.

3.3.3. Best Model Selection

The forecast results can be seen in table 4:

| Model                   | AIC   | MAPE   | RMSE  |
|-------------------------|-------|--------|-------|
| IOD Index               |       |        |       |
| SARIMA (1,1,1)(1,1,0)\(^1\) | -41.28 | 11,532 | 0.171 |
| SARIMA (2,1,3)(0,1,1)\(^1\) | -57.47 | 9,221  | 0.129 |
| SARIMA (2,1,3)(1,1,0)\(^1\) | -41.00 | 11,197 | 0.163 |
| Rainfall                |       |        |       |
| SARIMA (0,0,0)(2,1,0)\(^6\) | 13.22  | 7,904891 | 0.2390234 |
| SARIMA (1,0,1)(2,1,0)\(^6\) | 17.01  | 7,906423 | 0.2386987 |
| SARIMA (2,0,2)(0,1,1)\(^6\) | 2.01   | 6,797608 | 0.2092061 |

Table 4 shows that the SARIMA IOD index model (2,1,3)(0,1,1)\(^1\) and the SARIMA rainfall model (2,0,2)(0,1,1)\(^6\) are the best models to do forecasting because they have smaller AIC, MAPE, and RMSE values than other models. Based on the MAPE criteria, the two models have very good forecasting accuracy for predicting the IOD index and rainfall for the period January 2020 to December 2021.

3.3.4. Forecasting Results

The results of forecasting the IOD index and rainfall for the period January 2020 to December 2021 is presented in Figure 6:

![Figure 6](https://example.com/figure6.png)

Figure 6. Plot of forecast results for the IOD index and rainfall for Jan 2020 to Dec 2021

Figure 6 shows the plot of the IOD index and rainfall forecast results for 2020 to 2021. The results of the forecast are shown in the blue line, IOD has experienced data fluctuations, with a tendency to
show a decreasing trend in 2020 and an increase in 2021. Meanwhile, rainfall in West Aceh District has fluctuated, indicating fluctuation in rainfall, with a tendency to experience high rainfall in April and October for each year.

4. Conclusion
The conclusions obtained in this study are as follows:
1. The results of the spectral analysis of IOD index data for the period January 2010 to December 2019 obtained the highest peak frequency at the 7th frequency with period $P = \frac{2\pi}{\omega_7} = 17,1429$ thus indicating that the characteristics of the changes in the IOD index tend to experience an increase or decrease in the IOD index every 17 months. Meanwhile, the data on rainfall in West Aceh obtained the highest peak frequency at the 20th frequency with the period $P = \frac{2\pi}{\omega_{20}} = 6$ thus indicating that the characteristics of changes in rainfall in Aceh Barat tend to experience an increase or decrease in rainfall every 6 months.
2. The best model for predicting IOD index data is the SARIMA model $(2,1,3)(0,1,1)^{17}$ it was found that IOD experienced data fluctuation, with a tendency to show a downward trend in 2020 and an increase in 2021. While the best model for predicting rainfall data in West Aceh is the SARIMA model $(2,0,2)(0,1,1)^{6}$ it was found that the rainfall data fluctuated, with a tendency to experience high rainfall in April and October for each year.

References
[1] Saji, N. H., Goswami, B. N., Vinayachandran, P. N., & Yamagata, T. (1999). A Dipole Mode In the Tropical Indian Ocean. Nature, 401(6751), 360–363.
[2] Sahu, N., Yamashiki, Y., & Takara, K. (2010). Impact Assessment of IOD/ENSO in the Asian Region. Annuals of Disas, Kyoto University, 7.
[3] Ananda, S., & Miftahuddin. (2019). Generalized additive models fitting with autocorrelation for sea surface temperature anomaly data. IOP Conference Series: Materials Science and Engineering, 523, 012002.
[4] JAMSTECT. (2019). Indian Ocean Dipole. Diambil 2 Desember 2019, dari Japan Agency for Marine-Earth Science Technology website: http://www.jamstec.go.jp/aplnfo/sintexf/iod
[5] BMKG. (2019). Analisis Dinamika Atmosfer Laut: Analisis & Prediksi Curah Hujan. Diambil 2 Desember 2019, dari Badan Meteorologi, Klimatologi, dan Geofisika. website: https://www.bmkg.go.id/iklim/dinamika-atmosfir.bmkg
[6] Wei, W. W. S. (2006). Time Series Analysis Univariate and Multivariate Methods (Second Edition). New York: Pearson Education, Inc.
[7] Shumway, R. H., & Stoffer, D. S. (2011). Time Series Analysis and Its Applications: With R Examples (3rd ed). New York: Springer.
[8] Burnham, K. P., & Anderson, D. R. (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach (Second Edition). New York: Springer.
[9] Suhartono, S. (2011). Time Series Forecasting by Using Seasonal Autoregressive Integrated Moving Average: Subset, Multiplicative or Additive Model. Journal of Mathematics and Statistics, 7(1), 20–27.
[10] Maghfiroh, N., Hartatiati, S. S., Wahyuningsih, N., & Hakim, J. A. R. (2012). Peramalan Jumlah Wisatawan di Agrowisata Kusuma Batu Menggunakan Metode Analisis Spektral. Jurnal Sains dan Seni ITS, 1(1), 5.