MODEL DEVELOPMENT OF NIÑO 3.4 AND INDIAN OCEAN DIPOLE (IOD) ANOMALIES TELECONNECTION

Antonni Saputra[1], Eddy Hermawan[2], Denny Darmawan[1,*)]

[1]*Physics Study Programs at Yogyakarta State University*
[2]*Center of Atmospheric Science and Technology at National Institute of Aeronautics and Space*
*email korespondensi: darmawan@uny.ac.id

Abstract

The purpose of this research is to develop the teleconnection model of Niño 3.4 and IOD anomalies which can be used as reference to explain precipitation anomalies. El-Niño and IOD cycles are shown as the warming process of sea surface temperatures where for El-Niño is in the Pacific Ocean and IOD is in the Indian Ocean and each of them forms a cycle over a certain period of time. The method used to determine the dominant oscillation of the teleconnection of Niño 3.4 and IOD anomalies is Power Spectral Density (PSD), and to model the teleconnection of Niño 3.4 and IOD anomalies is ARIMA (Autoregressive Integrated Moving Average). The data used are Niño 3.4 index which is one type of index for El-Niño and IOD index. The results are Power Spectral Density (PSD) graphs for the teleconnection of Niño 3.4 and IOD anomalies which oscillates around 5 years. By the ARIMA method, the approximate model for the data of teleconnection of Niño 3.4 and IOD is ARIMA (1,1,2) with equation of $Z_t = 1.516 \; Z_{t-1} - 0.516 \; Z_{t-2} - 0.256 \; a_{t-1} + 0.021 \; a_{t-2}$.

Keywords: Anomalies Teleconnection, Indian Ocean Dipole, Niño 3.4

Introduction

El Niño–Southern Oscillation (ENSO) and Indian Ocean Dipole (IOD) are parts of a global phenomenon that also affects the condition of rainfall. ENSO is a climate phenomenon with area of origin is in the Pacific Ocean but has far-reaching consequences for weather around the world, and is particularly associated with droughts and floods [1]. There are several indices used to monitor the tropical Pacific region, all of them are based on anomalies of mean sea surface temperature in a given region. The Niño 3.4 index and the Oceanic Niño Index (ONI) are the most commonly used indices for determining El Niño and La Niña events [2]. Indian Ocean Dipole (IOD) is a pattern of surface and sub-surface temperatures 'basin' that greatly affect the annual climate anomaly of many countries around the Indian Ocean rim, as well as the global climate system [3].

ENSO and IOD are two different phenomena. IOD had peak power in the 1960s and 1990s, while ENSO was in the 1970s and 1990s. In addition, ENSO has a continuous broad band spectrum from the 1970s to the late 1980s or early 1990s, whereas IOD has an IOD spectrum showing two narrow bands that are elongated and distinctly separated [4]. At the time of strong El Niño occurrence, variations in the climate index are often analyzed in terms of their association with extreme climatic events, particularly long dry seasons or long wet seasons. The same goes to IOD [5]. The purpose of this research is to develop the teleconnection model of Niño 3.4 and IOD anomalies which can be used as reference to explain precipitation anomalies.

Data and Method

We used the monthly data of Niño 3.4 anomaly index from National Oceanic and Atmospheric Administration for period of January 1950 – November 2017 and IOD anomaly data from Japan Agency for Marine-Earth Science and Technology for period of January 1950 – November 2017. We examine the teleconnection patterns of Niño 3.4 anomaly index and IOD anomaly index using time series plot and Power Spectral Density plot. Model development method used in this research is Autoregressive Integrated Moving Average (ARIMA) method. This model is also known as the "Box-Jenkins" model. This model can estimate (forecast) data for the future based on past data that already exists.

Result

The characteristics of the teleconnection between Niño 3.4 and IOD

The teleconnection between Niño 3.4 and IOD anomaly sometimes they are mutually reinforcing and sometimes mutually debilitating depends on their patterns. If both are in the same phases at the same time, they will mutually be debilitating but if both are in different phases at the same time they will mutually debilitating. In attempt to identify the teleconnection of Niño 3.4 and IOD anomaly, we sum up both indexes as one. To describe the influences of Niño 3.4 and IOD we plot the Niño 3.4, IOD, and teleconnection between those data with time series plot and Power Spectral Density (PSD) plot. For time series we used the data from January 1981 until December 2016. The time series graph as shown at Figure 1 below.

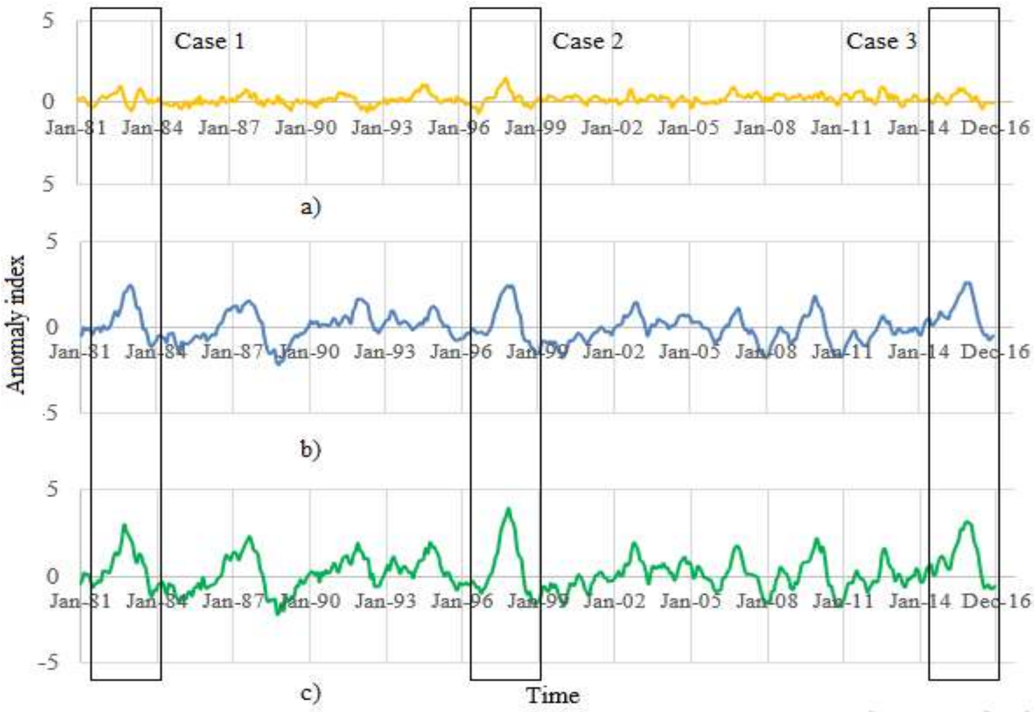


Figure 1. Time Series of (a) IOD, (b) Niño 3.4, and (c) Niño 3.4 and IOD teleconnection for period of January 1981 – December 2016.

As we can see above, IOD (Figure 1a) gives lesser influences compared with Niño 3.4 (Figure 1b) to the teleconnection between them (Figure 1c). But this does not mean IOD is not important. When IOD and Niño 3.4 have the same phase, the teleconnection between them becomes strong, as shown at Case 1, 2, and 3 from Figure 1. Next is to find the oscillation period of Niño 3.4, IOD, and teleconnection between them. At this step we used Power Spectral Density (PSD). Here, we used data from January 1950 until December 2016. The longer period of time was used to get more accurate result. The result as shown at Figure 2 below.

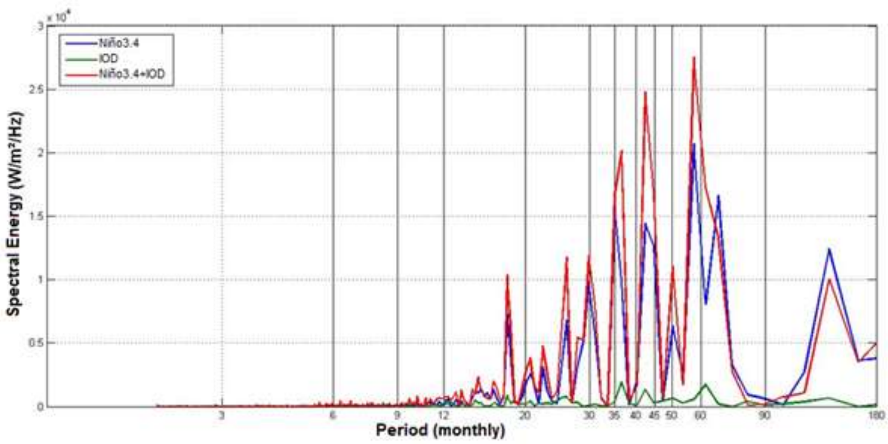
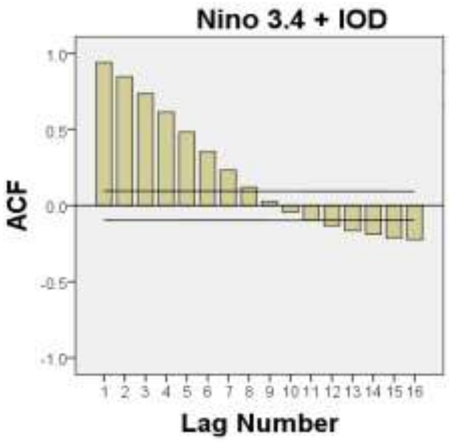


Figure 2. Power Spectral Density (PSD) of Niño 3.4, IOD, and the teleconnection between them for period of January 1950 – December 2016.

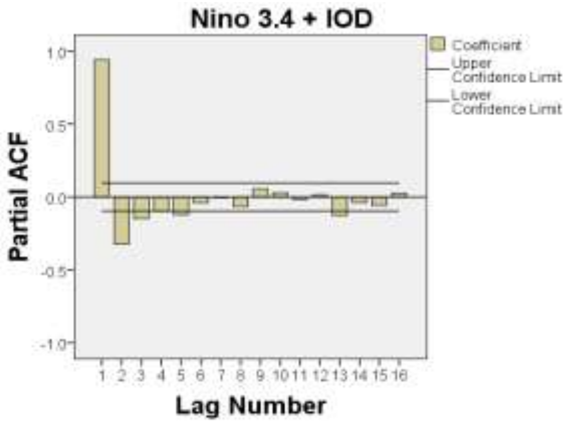
From Figure 2, we can see that the maximum oscillation period of Niño 3.4 is ~60 month (5 year), IOD is ~36 month (~3 year) and the teleconnection between them is ~60 month (5 year). Once again, from this PSD graph, Niño 3.4 gives more influences to the teleconnection of Niño 3.4 and IOD since the maximum oscillation of Niño 3.4 and the teleconnection of Niño 3.4 and IOD are almost the same. But it does not mean that IOD is not important, as the spectral energy for teleconnection of Niño 3.4 and IOD is higher than Niño 3.4 and this is because of the influences of IOD.

Stationary test for teleconnection data of Niño 3.4 and IOD

Before we build the model, it is important to make sure the data is stationary. In ARIMA method, there are few parameters that used to build the model, those are p, d, and q. PACF (Partial Autocorrelation function) used to get the p parameter, also known as AR (Autoregressive). PACF is a function that shows the magnitude of partial correlation between observation at the time of t with previous observations [6]. ACF (Autocorrelation Function) is used to get the q parameter, also known as MA (Moving Average). To make sure the data is stationary, this is where the d parameter comes from. The d parameter can be found from how many times the differencing is needed [7]. Sometimes differencing the data multiple times is needed until the data is stationary. The differencing is done by subtracting the data at the time of t with the previous data. Hence, the first step is to make sure the data is stationary or not. Using ACF and PACF, we can analyze whether the data is stationary or not. From here, we used the data from January 1981 until December 2016. The ACF and PACF result for the teleconnection of Niño 3.4 and IOD is shown at Figure 3.



(a)



(b)

Figure 3. (a) ACF (Autocorrelation Function) and (b) PACF (Partial Autocorrelation Function) of teleconnection between Niño 3.4 and IOD.

Figure 3 displayed that the lag graph of ACF forms a sine curve and the PACF shown that on lag-2 the curve decreased significantly although the lag-1 passed the upper confidence limit which means the data is not stationary yet. To make the data stationary, differencing the data is needed. Table 1 below shows the mean and variance of teleconnection between Niño 3.4 and IOD until second differencing.

Table 1. Mean and variance for teleconnection of Niño 3.4 and IOD

	Mean	Variance
Niño 3.4 + IOD	0.1977	1.126
Diff1_ Niño 3.4 + IOD	-0.0003	0.132
Diff2_ Niño 3.4 + IOD	-0.0012	0.339

As shown in Table 1 above, the mean and variance after the first differencing are smaller than before differencing, the mean is from 0.1977 to -0.0003 meanwhile the variance is from 1.126 to 0.132. It means the data getting more stationary since both the mean and variance are closer to 0. But, after the second differencing, the mean becomes smaller and variance becomes bigger compared with the first differencing. The mean is from -0.0003 to -0.0012 and the variance is from 0.132 to 0.339. This means the data only need one differencing since the first differencing already get more stationary than second differencing. With this, we got the value of d parameter, that is 1, since the data only needs one differencing to become stationary.

Creating the model for the teleconnection of Niño 3.4 and IOD

After the data is stationary, the next step is repeating the ACF and PACF using the data after the first differencing to get the possibility value of p and q parameter. The result is shown in Figure 4.

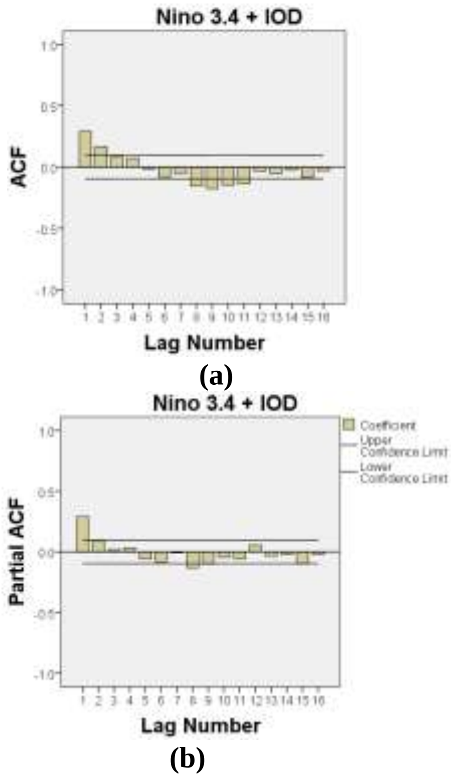


Figure 4. (a) ACF (Autocorrelation Function) and (b) PACF (Partial Autocorrelation Function) for the teleconnection of Niño 3.4 and IOD after first differencing.

As displayed in Figure 4, both ACF and PACF prove that the data are already stationary since both are not form a sine curve and decreasing exponentially. For ACF the lag-1 and 2 passed the upper confidence limit, which means the possible lags for q parameter are 1 and 2, and for PACF only lag-1 passed the upper confidence limit means the possible lag for p parameter is 1. To find the best parameter with the most less error, Root Mean Square Error (RMSE) and mean Absolute Percentage Error (MAPE) can be used to find out the best parameter for ARIMA model [7]. The result is shown at Table 2 below.

Table 2. Root Mean Square Error (RMSE) and mean Absolute Percentage Error (MAPE) for the teleconnection of Niño 3.4 and IOD after first differencing.

	MAPE	RMSE
ARIMA (1,1,1)	134.391	0.347
ARIMA (1,1,2)	134.089	0.347

In Table 2 we can see that ARIMA (1,1,2) has smaller MAPE value compared with ARIMA (1,1,1). This means the most suitable and the less error model is ARIMA (1,1,2) although the RMSE for both model are the same. The notation for ARIMA parameter is (p,d,q), from the previous steps we got that the p parameter is 1, d is 1 and q is 2, so we write it as ARIMA (1,1,2). The estimation values for each lags that will be used to find the ARIMA equation are shown in Table 3.

Table 3. Estimation values for lags from ARIMA (1,1,2).

	Lag	Estimation value
AR ($\varnothing$)	1	0.516
MA (θ)	1	0.256
	2	-0.021

The general equation for ARIMA (1,1,2) is shown in Equation 1.

$$Z_t = (1+\varnothing_1)\, Z_{t-1} + (-\varnothing_1+\varnothing_2)\, Z_{t-2} + \dots + (-\varnothing_{m-1}+\varnothing_m)\, Z_{t-m} - \theta_1\, a_{t-1} - \theta_2\, a_{t-2} - \dots - \theta_n\, a_{t-n} \quad (1)$$

For ARIMA (1,1,2), Equation 1 will become

$$Z_t = (1+\varnothing_1)\, Z_{t-1} + (-\varnothing_1)\, Z_{t-2} - \theta_1\, a_{t-1} - \theta_2\, a_{t-2} \quad (2)$$

Substitute the estimation value from Table 2 to equation 2, then the ARIMA (1,1,2) equation for the teleconnection of Niño 3.4 and IOD is shown in Equation 3.

$$Z_t = 1.516\, Z_{t-1} - 0.516\, Z_{t-2} - 0.256\, a_{t-1} + 0.021\, a_{t-2} \quad (3)$$

Model learning process

At this step we compare the model that has been created using equation (3) with actual data, and analyze how good the model is to follow the actual data. We use time series plot along with error histogram and coefficient of determination. The result is shown in Figure 5.

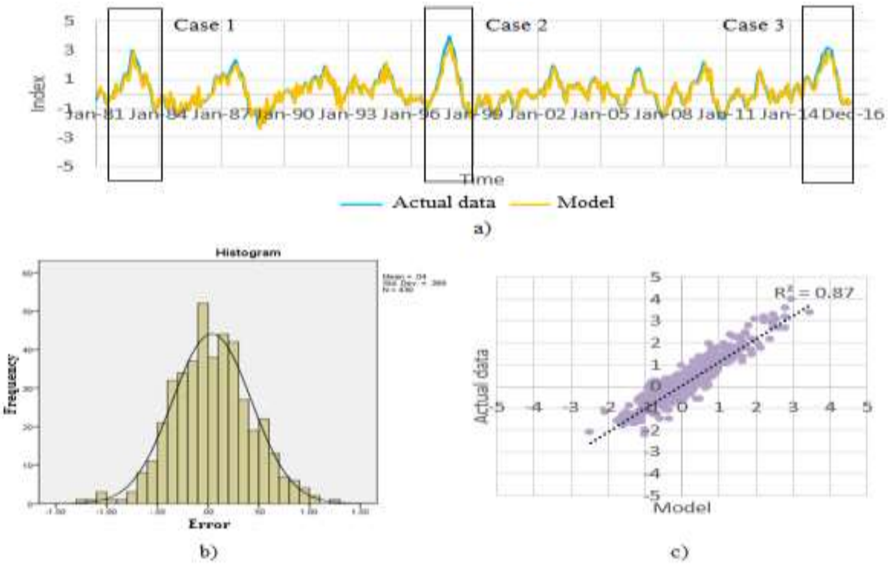


Figure 5. (a) Time Series, (b) error histogram, (c) Coefficient determination of the actual data and model of teleconnection between Niño 3.4 and IOD.

From Figure 5a, model seems can follow the actual data even when the actual data is significantly high as shown in Case 1, 2 and 3. The error histogram as shown in Figure 5b also follow the Gauss distribution which means data is reliable enough. How reliable is shown in Figure 5c by the coefficient of determination (R2) which is 0.87.

Model validation

To validate the model we use the data from January 2017 until November 2017 which is not used to build the model before to see if the model can actually follow the data after that and how reliable it is. Time series and coefficient of determination are used to analyze this. The result is shown in Figure 6.

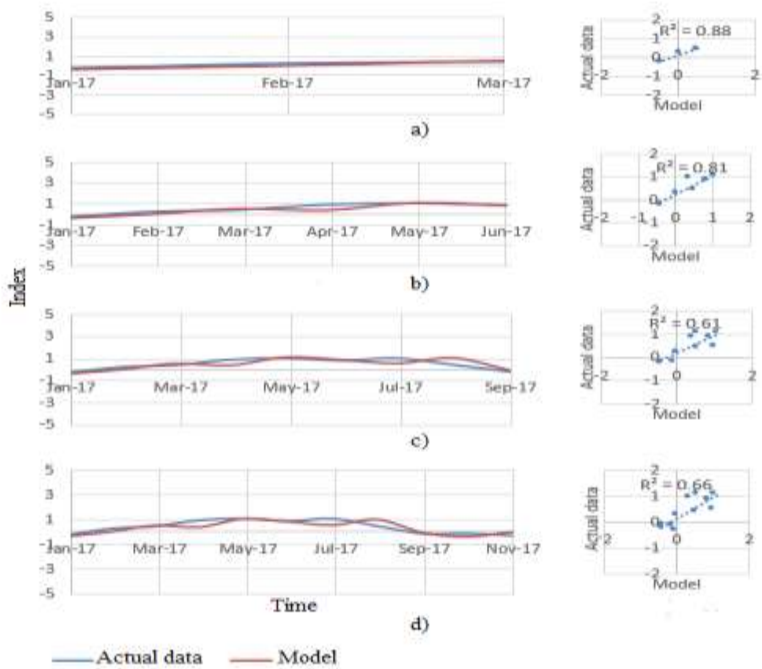


Figure 6. Time series and coefficient of determination for the model of teleconnection between Niño 3.4 and IOD for period of (a) January – March 2017, (b) January – June 2017, (c) January – September 2017, (d) January – November 2017.

As shown in Figure 6, the longer the time is used, the model becomes more deviate from the actual data. The coefficient of determination also becomes relatively lower as the time used is longer. When it is only 3 months (January until March) the coefficient of determination is 0.88, 6 month (January until June) is 0.81, 9 months (January until September) is 0.61, and 11 months (January until November) is 0.66. This means that the shorter the time is used, the more accurate is the model following the actual data.

Conclusion

We found that the maximum oscillation for the teleconnection is ~ 60 month (5 year). The best model that we found is using ARIMA (1,1,2) with the equation of $Z_t = 1.516 Z_{t-1} - 0.516 Z_{t-2} - 0.256 a_{t-1} + 0.021 a_{t-2}$. The model can follow the actual data with coefficient of determination is 0.87. Also, the model can be used to predict the data, but the longer the time is used, the model becomes more deviate from the actual data and the coefficient of determination becomes lower too. When 3 months (January until March) data is used, the coefficient of determination is 0.88, 6 months (January until June) is 0.81, 9 months (January until September) is 0.61, and 11 months (January until November) is 0.66.

References

[1] Kovats, R. S., Bouma, M. J., Hajat, S., Worrall, E., & Haines, A. (2003). El Niño and health. The *Lancet*, *362*(9394), 1481–1489.

[2] Trenberth, KE, dan National Center for Atmospheric Research Staff. (2016). The Climate Data Guide: Nino SST Indices (Nino 1+2, 3, 3.4, 4; ONI and TNI). From https://climatedataguide.ucar.edu/climate-data/nino-sst-indices-nino-12-3-34-4-oni-and-tni .

[3] Yamagata, T., Behera, S. K., Luo, J.-J., Masson, S., Jury, M. R., & Rao, S. A. (2013). Coupled Ocean-Atmosphere Variability in the Tropical Indian Ocean. *Geophysical Monograph Series*, 189–211.

[4] Saji, H. N., dan T. Yamagata. (2003). Structure of SST and Surface Wind Variability during Indian Ocean Dipole Mode Events: COADS Observations. *Journal of Climatology, 16*, 2735-2751.

[5] Hermawan, Eddy. (2012). *Model Interkoneksi Kejadian El-Nino dan Dipole Mode sebagai Indikasi Awal Datangnya Musim Kering/Basah Panjang di Kawasan Barat Indonesia*. Jakarta: Lembaga Penerbangan dan Antariksa Nasional.

[6] Lestari, Astuti Widya. (2012). *Pengembangan Model Prediksi Anomali Curah Hujan di Sentra Tanaman Pangan Kalimantan Timur Berbasis ARIMA*. Yogyakarta: Program Studi Fisika FMIPA UNY.

[7] Makridakis, S., Steven C. Wheelwright, dan Victor E. McGree. (1993). *Metode dan Aplikasi Peramalan*. Jakarta: Penerbit Erlangga.