Prisma Sains: Jurnal Pengkajian Ilmu dan Pembelajaran Matematika dan IPA IKIP Mataram
https://e-journal.undikma.ac.id/index.php/prismasains/index
e-mail: prismasains.pkpsm@gmail.com

April 2023. Vol. 11, No. 2
p-ISSN: 2338-4530
e-ISSN: 2540-7899
_pp. 385-392_

Identification of Rainfall events on Climate Phenomena in Medan based on Machine Learning

[1,2]Deassy Eirene Diana Doloksaribu, [1,*]Kerista Tarigan, [3]Richard Mahendra Putra, [4]Yahya Darmawan

[1]Faculty of Mathematics and Natural Sciences, University of Sumatera Utara, Medan, Indonesia.
[2]Meteorological Station of Kualanamu, Deli Serdang, Indonesia
[3]Meteorological, Climatological, and Geophysical Agency (BMKG), Indonesia
[4]State College of Meteorology Climatology and Geophysics (STMKG), Jakarta, Indonesia

*Corresponding Author e-mail: kerista@usu.ac.id

Received: March 2022; Revised: March 2023; Published: April 2023

Abstract

Indonesia has diverse topographical conditions that result in Indonesia having a unique climate. One of the unique climate elements to be studied is rainfall, because rainfall has a different pattern in each region, this different rainfall pattern is caused by several climate phenomena factors that affect the rainfall pattern, including El-Nino Southern Oscillation (ENSO), Indian Ocean Dipole (IOD) and Madden Julian Oscillation (MJO). Medan City is the capital of North Sumatra province which is one of the areas in the flood-prone category in North Sumatra, where the factor of flooding is due to rainfall events in a long period of time, so the author wants to know which climatic phenomena factors can affect rainfall events in Medan city by using Machine Learning technology through the Matlab application, where in this study has a method by forming four combination models, namely the combination of the influence of IOD, SOI and MJO; second combination of IOD and SOI; third combination of SOI and MJO; and fourth combination of MJO and IOD, these four combinations will be the rainfall value of the four models. Furthermore, the rainfall value of the model is compared with the observed rainfall value and verification test using Mean Absolute Error (MAE) and correlation. Then the calculation of the comparison between the four rainfall models with the observed rainfall obtained the lowest MAE value during the SOI and MJO phenomenon of 15.0 mm and the highest correlation value during the IOD and SOI and SOI and MJO phenomena. So it is concluded that the combination of SOI and MJO has the best verification value. This shows that based on Machine Learning modeling, the model shown as the best predictor in Medan city is when the model combination consists of SOI and MJO.

Keywords: Rainfall Events, Climate Phenomena, Machine Learning

How to Cite: Doloksaribu, D., Tarigan, K., Putra, R., & Darmawan, Y. (2023). Identification of Rainfall events on Climate Phenomena in Medan based on Machine Learning. _Prisma Sains : Jurnal Pengkajian Ilmu dan Pembelajaran Matematika dan IPA IKIP Mataram_, _11_(2), 385-392. doi:https://doi.org/10.33394/j-ps.v11i2.7738

https://doi.org/10.33394/j-ps.v11i2.7738

INTRODUCTION

Indonesia has a variety of topographical conditions because Indonesia is located between the Asian and Australian continents, the Pacific Ocean and the Indian Ocean and is surrounded by vast oceans. This diverse topography makes the climate in Indonesia unique to study, where one of the elements of climate is Rainfall. Rainfall is the height of rainwater collected in rain gauges on a flat place, not absorbing, not seeping and not flowing. This element of rainfall is interesting to study because rainfall patterns in the Indonesian region are not all the same, but we can see the factors of global phenomena that can affect rainfall patterns in the Indonesian region including El-Nino Southern Oscillation (ENSO), Indian Ocean Dipole (IOD) and Madden Julian Oscillation (MJO).

ENSO is a form of climate aberration in the Pacific Ocean characterized by rising sea surface temperatures (SST) in the Central and Eastern equatorial regions, the phenomenon plays an important role in annual climate variations. ENSO has an influence in some parts of Indonesia characterized by the amount of rainfall. Likewise, IOD is also a climate aberration produced by the interaction of the ocean and atmosphere in the Indian Ocean around the Equator (Kailaku, 2009). However, in addition to ENSO and IOD, MJO, which is a global factor, is also a phenomenon of oscillation wave propagation that moves eastward with a recurrence period of 30 - 90 days. This oscillation has a very strong impact on the areas through which it passes, with the first occurrence occurring in the Indian Ocean and moving eastward (Yana, 2014). This MJO phenomenon can be used to predict climate, especially in the tropics during active MJO periods (Windayati *et al*., 2016).

So that when these global phenomena factors occur in the Indonesian region, it causes differences of rainfall in each region of Indonesia, one of which is in the city of Medan in the province of North Sumatra which is geographically located in the western part of Indonesia, where the city of Medan is included in the 5 largest growth center cities in Indonesia with a population of around 2,210. 743 people (BPS, 2020) and is also one of the areas included in the flood-prone category in the North Sumatra region (Tampubolon, 2018) where one of the factors for flooding is rainfall in a long period of time (Anggraini *et al*., 2021) for this reason it is necessary to know what global phenomenon factors can affect rainfall events in Medan city, so that as forecasters we can predict the rainfall events. To find out these global phenomena, we need technology that can analyze and forecast rainfall in the city of Medan.

Machine learning (ML) is one of the technologies commonly used in analyzing and predicting. ML is a branch of Artificial Intelligence (AI) which means programming from computers that require human intelligence (Mehr *et al*., 2017). The capabilities of this AI include understanding and monitoring information visually/spatially, making predictions, hearing, interacting with humans and machines and continuously learning to improve themselves (Yang *et al*., 2017). So that by using ML we can reflect the given data pattern through a (mathematical) model. The concept in creating an ML model is to utilize input parameters then training is carried out to get the output results with a 3-layer division structure consisting of input layers, hidden layers and output layers. The results of the input layer calculation from the hidden layer are received as a new input layer and then recalculated to become the output layer (Putra *et al*., 2020).

So in this study, monthly rainfall data from January 2011 to December 2022 for 12 years is used as rainfall modeling data as monthly rainfall modeling data that utilizes global phenomenon factors such as ENSO, IOD and MJO data with each data in the same month as rainfall for 12 years based on ML to find out which global phenomenon factors affect monthly rainfall events in Medan city, so that the results of this study are expected for forecasters to be able to know which global phenomenon factors can cause rain in the Medan city area and can anticipate flooding when it rains (Wang *et al*., 2013).

Theory and Calculation

In each region has a different climate, because the character of the weather is determined based on the probability of the value of one or more specified climate elements such as rain, temperature and wind. With the differences in climate in each region, there are differences in rainfall patterns that occur based on the occurrence of climatic phenomena factors such as *El-Nino Southern Oscillation* (ENSO), *Indian Ocean Dipole* (IOD) and *Madden Julian Oscillation* (MJO).

ENSO (El-Nino Southern Oscillation)

It is a global climate phenomenon that consists of three phases: El Nino, La Nina and Neutral. EL Nino is known as a hot phase followed by a decrease in rainfall and La Nina is known as a cold phase that stimulates an increase in rainfall above normal. ENSO is characterized by measuring changes that occur in sea surface temperatures in the eastern Pacific and western Pacific (Lefkovitz, 2013), to determine the occurrence of El Nino and La Nina can be calculated from the Southern Oscillation Index (SOI) data, where this value is obtained from the value of air pressure in Tahiti and in Darwin (Maulidiya *et al*., 2012). If the air pressure in the western Pacific weakens, the eastern and central Pacific tends to be strong and SOI is positive, otherwise if the air pressure in the western Pacific strengthens, the eastern and central Pacific tends to weaken and SOI is negative (Haryanto, 1998). The calculation of the ENSO Index is calculated using the following formula:

$$\text{SOI} = \frac{(Pdiff - Pdiffav)}{SD} x\ 10$$

Information :
Pdiff = Difference between Tahiti and Darwin Sea Surface Temperature
Pdiffav= Long-term average of Pdiff in the month
SD = Standard Deviation of Pdiff
SOI = *Southern Oscillation Index*

Table 1. El-Nino, La-Nina and normal prediction guide for SOI values

SOI values	The Phenomenon That Will Happen
<-10 for 6 months	Strong El Nino
-5 s/d -10 for 6 months	Weak El Nino
-5 s/d +5 for 6 months	Neutral
+5 s/d +10 for 6 months	Weak to moderate La-Nina
>+10 for 6 months	Strong La-Nina

IOD (Indian Ocean Dipole)

It is the difference between sea surface temperature (SST) anomalies in the western and eastern parts of the Indian Ocean. The process of Ocean and atmosphere interaction by IOD affects the pattern of rainfall anomalies in the tropics, when the SPL anomaly in the western part of the Indian Ocean is greater than the eastern part of the Indian Ocean, the air pressure in the Indian Ocean atmosphere on the west coast of Sumatra will be low, so convection will be high and give an increase in rainfall from normal in the Indian Ocean west coast and give a decrease in rainfall in the eastern part of the Indian Ocean characterized by positive IOD values that affect the North Sumatra region, especially the city of Medan. Conversely, if the eastern SPL anomaly is greater than the western part of the Indian Ocean, the air pressure in the east coast Indian Ocean will be low and experience high convection which gives an increase in rainfall in the eastern Indian Ocean characterized by negative IOD values that affect the North Sumatra region, especially the city of Medan (Ashok *et al*., 2001).

The IOD climate variability data used is the DMI value, which is a 5-year monthly index of the difference between sea surface temperature anomalies in the western and eastern Indian Ocean. The calculation of the IOD Index is calculated using the following formula:

Table 2. Classification of Dipole Mode Index (Ramadhanty *et al*.,)

No	DMI values	Annotation
1	DMI > 0.48 ºC	DMI +
2	-0.48 ºC ≤ DMI ≤ 0.48 ºC	Netral
3	DMI < 0.48 ºC	DMI -

MJO (Madden Julian Oscillation)

An oscillation in the equatorial region, the MJO phenomenon has a period of 30-60 days, which is characterized by the appearance of Cumulonimbus (Cb) clouds over the Indian Ocean that move eastward along the equator. MJO, which is also climate variability, has an influence on rainfall in the area it passes through, when MJO is in a strong phase it causes high rainfall and when it is in a weak phase it causes low rainfall (Madden *et al*., 1971).

METHOD

The method used in this research is experimental and quantitative. This research will investigate which combination of monthly global climate phenomena is the most influential and strongly correlated with monthly rainfall in Medan city. In this study, four ML models of global phenomena input combinations were formed in influencing rainfall. The four combination models are the first combination of the influence of IOD, SOI, and MJO; the second is the combination of the influence of IOD and SOI; the third is the combination of the influence of SOI and MJO; and the fourth is the influence of the combination of MJO and IOD. So that from the results of this climate model data, which climate phenomena cause rainfall using ML techniques (Shrestha *et al.*,2018).

The ML model was built using MATLAB application. This ML model consists of two or three input variables (based on the four model combinations), 2 hidden layers with 50 neurons in the first hidden layer and 10 neurons in the second hidden layer, and one output variable, namely monthly rainfall. An illustration of the ML model structure can be seen in Figure 1.

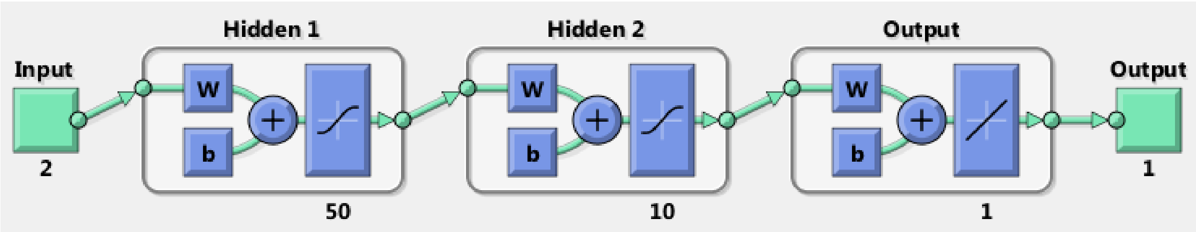

Figure 1. Illustration of the structure of ML model

After passing the modeling process using ML, the rainfall values of the four combination models will be obtained. Furthermore, the model rainfall values that have been obtained will be compared with the observed rainfall values and verification tests will be carried out using the mean absolute error (MAE) and correlation.

The MAE calculation is carried out by measuring the average absolute deviation of the model values from the observations. MAE shows the average of the magnitude of the error without the effect of eliminating each other between the sum of positive and negative values. Good model performance has an MAE value close to 0. The MAE calculation is expressed by the following equation (Adhikari *et al*., 2013).

$$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n} |e_t|$$

Information:

e_t = the value of the difference between modeled rainfall and observed rainfall.
n = Amount of data.

The correlation test is a method used to test the presence or absence of a relationship and the direction of the relationship (positive or negative) of two or more variables. Where the correlation value of 0.8 - 1.0 has an interpretation of a very strong relationship level, the correlation value of 0.6 - 0.79 has a strong relationship level, the correlation value of 0.4 - 0.59 has a moderate relationship level, the correlation value of 0.2 - 0.39 has a low relationship level, and the correlation value of 0.0 - 0.19 has a low relationship level. The calculation of the correlation value is as follows (Sugiyono, 2015).

$$r = \frac{n\sum XY - \sum X \sum Y}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum Y^2 - (\sum Y)^2]}}$$

Information:
X = observation rainfall data
Y = rainfall data model
n = amount of data

After obtaining the MAE and correlation values of each rainfall model against the observed rainfall, in the end, the model with the smallest MAE value and the largest correlation value will be obtained, which is the combination of global phenomena most influential on monthly rainfall values in Medan City based on the ML model-based approach.

RESULTS AND DISCUSSION

Figure 2 shows the modeled and observed rainfall values in time series. There is a total of 144 data (12 years multiplied by 12 months) with each color indicator listed in the upper right corner. Visually, the ML model has been able to follow the pattern of monthly rainfall changes that occur. However, it can be seen that the ML model occasionally overestimates the value of monthly rainfall.

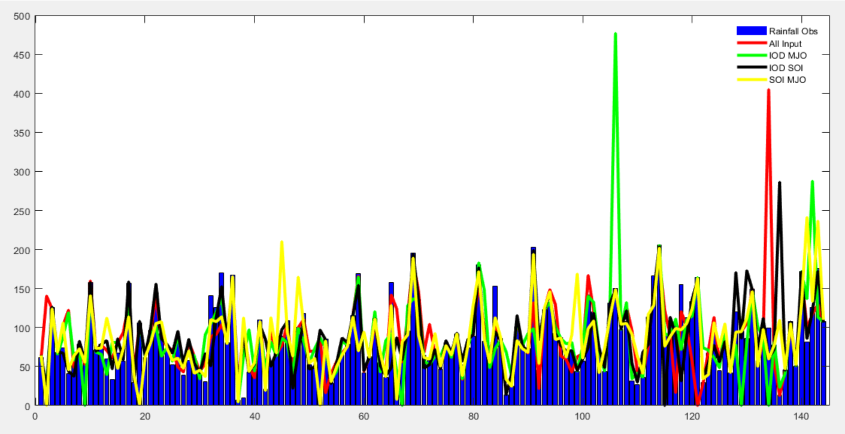

Figure 2. Comparison of ML modeled monthly rainfall to observed rainfall

From the value of MAE in Figure 3, it can be seen that the SOI and MJO combination model has the smallest absolute error value (MAE) of rainfall with 15.0 millimeters, followed by the MJO and IOD combination model which is worth 19.81 millimeters, the all input combination model which is worth 19.84 millimeters, and the last is the IOD and SOI combination model with an MAE of 20.76 millimeters.

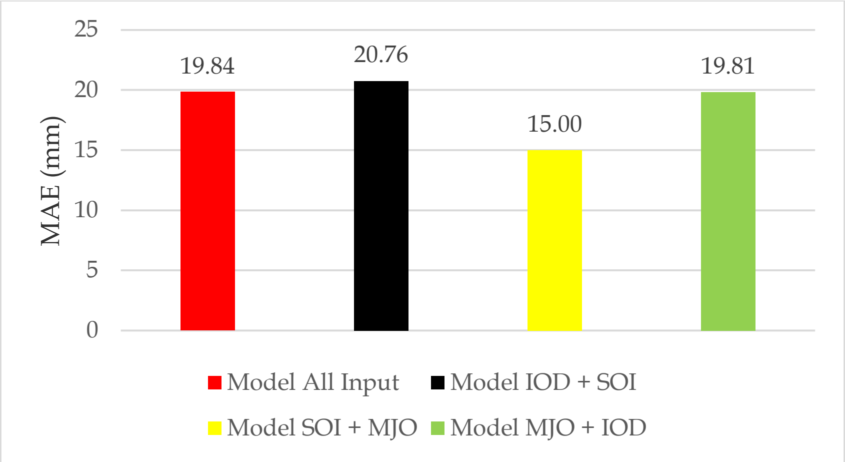

Figure 3. MAE Value of Output Model

Considering the model correlation values in Figure 4, there are two models with the highest correlation values, namely the IOD and SOI combination model, and the SOI and MJO combination model, with a correlation value of 0.70. While the MJO and IOD combination model obtained a correlation value of 0.64, and the All Input model obtained the lowest correlation value of 0.62.

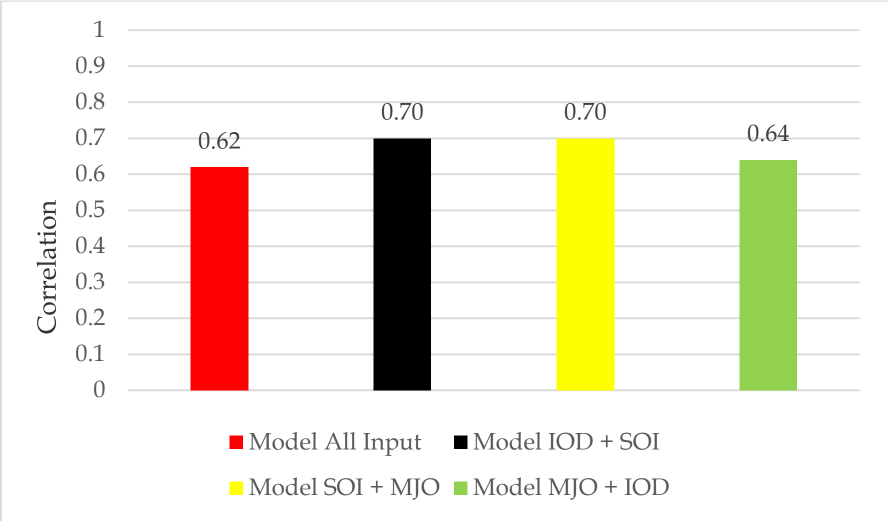

Figure 4. Correlation Value of Model Output

From the results of the comparison of rainfall values, MAE values and model correlation values above, it indicates that the four modeling of rainfall values using ML with a combination of input values of global climate phenomena as a whole produce strong correlation values with a range (0.62 - 0.70) and MAE values with a range of 15.0 - 20.76 mm per month.

The combination of climate phenomena that produces rainfall values closest to the observation value and that has the best correlation value is the combination of SOI and MJO phenomena with an MAE value of 15.0 mm and a correlation of 0.7. The combination of the IOD and SOI phenomena also shows the same value with good correlation results of 0.7, but there is a difference in the MAE value in the combination of the IOD and SOI phenomena, where the combination of the IOD and SOI phenomena has a greater MAE value of 20.67 mm per month which means the error value is greater than the combination of the SOI and MJO phenomena. For this concludes that from the overall combination of climate phenomena that cause rainfall in Medan city is when the combination of SOI and MJO climate phenomena.

CONCLUSION

The occurrence of rainfall in Medan city can be influenced by factors of climate phenomena seen through research by several combinations of global climate phenomena such as ENSO, IOD and MJO by using methods that form four ML models of the combination of climate phenomena inputs, namely the first combination of the influence of IOD, SOI, and MJO; the second is the combination of the influence of IOD and SOI; the third is the combination of the influence of SOI and MJO; and the fourth is the combination of the influence of MJO and IOD. The results of modeling using ML obtained rainfall values with the four combination models, then the model rainfall values will be compared with the observed rainfall values and verification tests will be carried out using the mean absolute error (MAE) and correlation.

Based on the model verification value using MAE and Correlation, it can be seen that the SOI + MJO model has the best verification value with an MAE value of 15.0 mm and a correlation of 0.7. This shows that based on ML modeling, the monthly rainfall value in

Medan City is most correlated and influenced by climate phenomena factors when there is a combination of SOI and MJO.

ACKNOWLEDGMENT

Thanks to Climatological Station of Deli Serdang for providing data so that we can conduct research and lecturers who take the time to guide until this research is completed, as well as colleagues who helped.

REFERENCES

Adhikari, R., dan Agrawal, R. K. (2013). An Introductory Study on Time Series Modeling and Forecasting, Lambert Academy Publishing, New Delhi.

Anggraini, N., Pangaribuan, B., Siregar, A.P., Sintampalam, G., Muhammad, A., Damanik, M.R.S., Rahmadi, M.T. (2021). Analisis Pemetaan Daerah Rawan Banjir Di Kota Medan Tahun 2020. Jurnal Kajian Ilmu dan Pendidikan Geografi, Vol. 4, No. 2.

Ashok, K., Z. Guan, and T. Yamagata. 2001. Impact of the Indian Ocean dipole on the relationship between the Indian monsoon rainfall and ENSO. Geophysical Research Letters 28.23: 4499-4502.

BPS. (2020). Jumlah Penduduk kota Medan menurut kecamatan dan jenis kelamin (jiwa), 2018-2020. Available online from: https://medankota.bps.go.id/indicator/12/31/1/jumlah-penduduk-kota-medan-menurut-kecamatan-dan-jenis-kelamin.html. [Accessed April 27, 2023]

Haryanto, U. (1998). Keterkaitan Fase SOI terhadap curah hujan di DAS Citarum. http://repository. ipb.ac.id/bitstream/handle/ 123456789/ 4016/1998 uha.pdf?sequence=. diakses tanggal 25 April 2023.

Kailaku, T.E. (2009). Pengaruh ENSO (El Nino-Southern Oscillation) DAN IOD (Indian Ocean Dipole) Terhadap Dinamika Waktu Tanam Padi di Wilayah Tipe Hujan Equatorial dan Monsunal (Studi Kasus Kabupaten Pesisir Selatan, Sumatera Barat dan Kabupaten Karawang, Jawa Barat). Skripsi. Department of Meteorology FMIPA IPB. Bogor.

Lefkovitz, Y. (2013). Climate and Commodity Prices: An Analysis of the Role of ENSO Forcasts in Agricultural Comodity Markets. New York: New York University.

Madden, R. A. dan Julian Paul. (1971). Detection of a 40±50 day oscillation in the zonal wind in thetropical Pacific. J Atmos Sci, 28, 702-708.

Maulidiya, H., Ihwan, A., Jumarang, M.I. (2012). Penentuan Kejadian El Nino dan La Nina berdasarkan nilai Southern Oscillation Indeks. Jurnal Positron, Vol.2, No. 2.

Mehr, H., Ash, H., and Fellow, D. (2017). Artificial Intelligence for Citizen Services and Government. Ash Cent. Democr. Gov. Innov. Harvard Kennedy Sch., no. August, p. pp 1-12.

Putra, R.M., Alfiandy, S., Haq, B.E.A. (2020). Identifikasi Pengaruh El Nino Southern Oscillation (ENSO), Indian Ocean Dipole (IOD) dan Madden Julian Oscilation (MJO) terhadap Intensitas Curah Hujan Bulanan di Indonesia Berbasis Machine Learning. Buletin Ngurah Rai, 6(2), pp. 1-8.

Ramadhanty, F.W., Muslim., Kunarso., Rochaddi, B., and Ismunarti, D.H. (2021) Pengaruh Fenomena IOD (*Indian Ocean Dipole*) Terhadap Sebaran Temperatur dan Salinitas di Perairan Barat Sumatera. *Indonesian Journal of Oceanography*, Vol.03, No. 01.

Shrestha, B., Rifaat, M.R. (2018) A Machine learning Approach to rainfall forecasting based on climate data. Journal of Computer Science and Information Security.

Sugiyono. (2015). Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif, dan R&D). Alfabeta, Bandung.

Tampubolon, K. (2018). Aplikasi Sistem Informasi Geografis (SIG) sebagai penentuan Kawasan Rawan Banjir di Kota Medan. Jurnal Pembangunan Perkotaan, Vol. 6, No. 2, 63-88.

Wang, H.W., Kuo, P.H. and Shiau, J.T., (2013) Assessment of climate change impacts on flooding vulnerability for lowland management in southwestern Taiwan. Natural Hazards, 68(2): 1001-1019.

Windayati, R., Surinati, D. (2016). Fenomena *Madden Julian Oscillation.* Journal of Oseana, Vol.XLI, No. 3.

Yana, S., Ihwan. A., Jumarang, M.I., Apriansyah. (2014) Analisis Pengaruh Madden Julian Oscillation, Annual Oscillation, ENSO dan Dipole Mode terhadap curah hujan di Kabupaten Kapuas Hulu. Journal of Prisma of Physics, Vol. 2, No.2.

Yang, J., Chen, Y., Huang, W., and Y. Li. (2017) Survey on Artificial Intelligence for Additive Manufacturing. 23rd Int. Conf. Autom. Comput., no. September, pp. 7–8.