JIT Journal of Innovation and Technology
e-ISSN: 2721-8562; p-ISSN: 2721-8570
Vol. 3, No. 2, October 2022. Page: 45 – 51

Time Lag Application on Correlation Analysis among the national and international agencies for the rainfall data of Lampung Province, Indonesia

Lintang Kurnia Aridini[1], Ahmad Zakaria[1*], Dyah Indriana Kusumastuti[1], Ahmad Herison [1], Endro Prasetyo Wahyono[1]

[1] *Departement of Civil Engineering, Faculty of Engineering, Universitas Lampung, Jl. Prof. Soemantro Brojonegoro no. 1, Bandar Lampung 35145, Indonesia*

*Corresponding author: ahmad.zakaria@eng.unila.ac.id

Article history
Received: 18.06.2022
Revised: 23.09.2022
Accepted: 25.10.2022

DOI:10.31629/jit.v3i2.5060

Abstract
The rainfall data in Lampung tends to be small and cannot represent a large location, so other sources of rainfall data are needed that can be used. One of them is TRMM on that basis it is deemed necessary to see the correlation of the TRMM and BMKG data. The purpose of this research is to analyze how big the correlation between TRMM data and BMKG data with Time Lag. And its Predictability, as well as improved correlation with Time Lag. The method used in this study is the Pearson correlation. The result of this research is that the biggest correlation is on H0 and H+1 for the highest daily cumulative 0.3376 - 0.4808, for cumulative 2 days at 0.417- 0.555, for cumulative 7 days 0.569 - 0.706, for cumulative 15 days at 0.612 - 0.785 and cumulative 30 daily 0.710 -0.811 the conclusion of this study is that TRMM and BMKG have a better correlation on H + 1 and (H0), the results of linear regression analysis with Time Lag concluded that the predictive ability of the data is better on Time Lag and the difference in correlation between H + 1 and H0 obtained an increase in the correlation value. declared to have increased 0.01 - 0.2).

Keywords: rainfall, correlation, time lag, rainfall station

1. Introduction

Indonesia is one of tropical country in the southeast Asia, Indonesia have 2 seasons, rainy dan summer and because of that Indonesia depends on rainfall data to do project about water resources management and natural disasters management. Rainfall data in Indonesia which issued by BMKG have many problems, such as blank data and immeasurable rainfall data. This problem can be solved by Using supporting rain data issued by TRMM, based on supporting data TRMM, ground data issued by BMKG can be predict by using analytical methods that can estimate the value of the measured Ground rainfall data. TRMM using Remote sensing technology and the advantages of remote sensing technology should be further utilized to study the characteristics of the weather and climate in an area for the benefit of water resource management and its utilization for the welfare of the community [1].

In Indonesia, there are various regions with different elevations and unique nature, because of thar rainfall data from satellite or TRMM are affected by topography. To reduce error from predicting BMKG rainfall data using TRMM data

DOI:10.31629/jit.v3i2.5060

there are study conducted [2] it was explained by using a time lag of 1-3 days can affect the correlation value of the BMKG and TRMM. Based on this research it's important to do research about Correlation using BMKG and TRMM.

Correlation is one method to see the similarity of data patterns and shapes. General correlation is used to mathematically predict things that happen in the field. Correlation is also useful as a parameter for creating modeling functions. In terms of rainfall data, correlation analysis is used to see how much similarity the pattern between BMKG and TRMM data is with a certain time lag. Based on the results of previous studies, it was found that the time lag can change the correlation value and also allows better prediction of modeling results [3].

The purpose of this research is to analyze correlation strength value of TRMM and BMKG with Time Lag, to analyze predict value of TRMM rainfall data can predict BMKG rainfall data and to analyze Differences the predictive value with and without Time Lag [4].

2. Materials and Methods

The method used in this Research are the Pearson correlation method and Linear Regression. The Pearson correlation is a method to measure strength value between two different variables using a quantitative data. If the data observations or measurements have many variables, then it must be analyzed which variables are strongly related and carried out to further analysis [5].

Method which used to measure degree of association or degree of relationship between variables are known as correlation analysis. Correlation and Linear Regression are two stages to analyze relation between variable and predicted the result. because if the observed variable has correlation strong value with other variables, then we can predict the value of the variable using Linear Regression

Pearson Correlation function used in this research is as follow:

$$R = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{\{n\Sigma x^2 - (\Sigma x)^2\}\{n\Sigma y2 - (\Sigma y)2\}}} \quad (1)$$

In this study, several research procedures were used which became the stages of the research carried out. The research procedures consist of three stages That are Literature Study, Data Collection and Data Processing [6].

Literature study is the stage carried out by collecting literature and previous research studies. These are collected and then used as research references so that basic hypotheses and research frameworks can be produced.

This Research used secondary data obtained from several sources, such as BMKG rain data and TRMM satellite rain data for a period of 20 years. The data used was checked first and then normalized and validated [7].

The next stage after data collection is data processing carried out by several methods, the stages of data processing are as follows:

a. Transform rainfall data into Time Series
b. Normalize rainfall data and replacing the unmeasured rainfall data with value of 0
c. Correlate rainfall data between TRMM and BMKG with time lag from D-3 to D+3
d. Comparing correlation strength value between time lag and take greatest variable for further analysis
e. Analyze Linear Regression using variable which has greatest correlation strength value
f. Analyze value of the Determination coefficient and see how much improvement the model's ability to predict rainfall data is analyzed using Time Lag.
g. Calculate error from linear regression mathematical function.

3. Result

This Research use Pearson Correlation method as a data processing method. The BMKG rainfall data has several weaknesses, including missing and unmeasured data, Because of that data weaknesses can affect this research. This research was conducted to see the time lag influence rainfall correlation strength value. Within D-3 to D+3 of the rainfall being compared, with rainfall data from two different sources

3.1 Transform Rainfall Data to Time Series

Before using Precipitation data from the BMKG as analyzing data, BMKG still have many unmeasured. To normalize BMKG data required replace blank and unmeasured data with value of 0 [8]. In this research is a software created using the Force 2.0.

DOI:10.31629/jit.v3i2.5060

3.2 Data Consistency Test

Consistency Test is carried out by adding up each existing daily data for a year or annual cumulative. And performed linear regression with rainfall data from other stations. And the results of the data consistency test, it can be stated that the rainfall data is good because it has an R2 value with a value above 0.9 with a very Strong value [9].

3.3 BMKG and TRMM Daily Cumulative Rainfall Data Correlation with the Timelag

Correlation analysis consist of two daily cumulative rainfall data. However, the correlation analysis was carried out using Time Lag from D-3 to D+3. The results of the correlation analysis can be seen in Table 1.

Table 1. Statistics of research participant

Station	Time Lag Correlation						
	D-3	**D-2**	**D-1**	**D0**	**D+1**	**D+2**	**D+3**
Kotabumi	0,11	0,11	0,12	0,36	0,28	0,11	0,09
Maritim	0,03	0,09	0,11	0,33	0,17	0,08	0,05
Masgar	0,07	0,1	0,09	0,18	0,48	0,16	0,1
Raden Inten II	0,09	0,08	0,09	0,27	0,36	0,16	0,08

From the results of the correlation analysis using Time Lag, the strongest correlation value is obtained, namely correlation in H0 for Kotabumi and Maritim rainfall Stations while H+1 for Masgar and Radin Inten II Stations, and after correlation have been done and strongest correlation value determined so that linear regression analysis can be carried out to find out how much the data is capable of [10]. Predicting BMKG data using TRMM rainfall data. The results of the linear regression analysis are as follows:

Table 2. Linear Regression Daily Cumulative Result

Station	Time Lag	Function	R^2
Kotabumi	H0	y = 0,369x + 6,0663	0,1297
Maritim	H0	y = 0,3446x + 4,5664	0,114
Masgar	H+1	y = 0,4819x + 4,4786	0,2312
Radin Inten II	H+1	y = 0,3549x + 5,1147	0,13

From the Linear Regression results Determined with the strongest value is located at Masgar station with a value of 0.2312, Radin Inten II with 0.13, Kotabumi 0.1297 and Maritim 0.1140. Thus, it is found that the influential Masgar and Radin Inten II Time Lag Stations are H+1 [11]. Meanwhile, Kota Bumi and Maritime Stations have no effect on Time Lag for daily cumulative analysis.

3.4 BMKG and TRMM Cumulative 2 Days Rainfall data Correlation with the Timelag

This correlation analysis was carried out the same as the previous daily correlation analysis, but the data used in this analysis was cumulative 2 days. Correlation analysis with a cumulative 2 days to see how much change occurs if the data used is data with a cumulative 2 days. In this analysis, correlation analysis was carried out, from several time lags from D-3 to D+3 with a cumulative 2 days [12]. From the results of the 2 daily correlation analysis, it was found that the correlation value increased compared to the daily cumulative correlation value with a medium scale for each station. However, the highest correlation displacement also occurred at the Radin Inten II rain station (table 3) where for the daily cumulative there was a Time Lag H+1 while the cumulative 2 days occurred on the same day (H0).

DOI:10.31629/jit.v3i2.5060

The correlation analysis for BMKG rainfall data against TRMM rainfall data with cumulative 2 day data is as follows:

Table 3. Cumulative 2 Days Correlation with Timelag

Station	Time Lag Correlation						
	D-3	**D-2**	**D-1**	**D0**	**D+1**	**D+2**	**D+3**
Kotabumi	0,20	0,21	0,31	**0,49**	0,45	0,28	0,19
Maritim	0,11	0,16	0,27	**0,42**	0,36	0,17	0,09
Masgar	0,15	0,17	0,20	0,40	**0,56**	0,39	0,20
Radin Inten II	0,15	0,15	0,23	**0,45**	0,21	0,31	0,18

And based strongest value on correlation analysis the results of the linear regression analysis are as follows:

Table 4. Linear Regression Cumulative 2 Days Result

Station	*Time Lag*	Function	R^2
Kotabumi	H0	y = 0,5171x + 10,21	0,237
Maritim	H0	y = 0,4399x + 8,2064	0,174
Masgar	H+1	y = 0,5769x + 7,9635	0,308
Radin Inten II	H0	y = 0,4578x + 9,1388	0,203

From the results, obtained the predictive ability value on a low and very low scale for a cumulative 2 days. With the greatest capability, Masgar Station with a value of 0.308, Radin Inten II with 0.203, Kotabumi 0.237 and Maritim 0.174 [13]. For the influential Time Lag, it is also found that the Masgar Station with the influential Time Lag is H+1.

3.5 BMKG and TRMM Cumulative 7 Days Rainfall data Correlation with the Time lag

The analysis comparing rainfall data with cumulative 7 days. This is done by testing whether the more the number of cumulative days of rain, the greater the correlation value. The result of correlation analysis is as follows:

Table 5. Cumulative 7 Days Correlation with Time Lag

Station	Time Lag Correlation						
	D-3	**D-2**	**D-1**	**D0**	**D+1**	**D+2**	**D+3**
Kotabumi	0,54	0,59	0,63	**0,66**	0,64	0,60	0,54
Maritim	0,44	0,48	0,54	**0,56**	0,55	0,47	0,40
Masgar	0,47	0,54	0,59	0,64	**0,70**	0,67	0,61
Radin Inten II	0,47	0,53	0,59	0,63	**0,66**	0,62	0,55

From the results of the correlation analysis the largest correlation value is obtained, at H0 for Kotabumi, Maritim and Radin Inten II stations, while for Masgar Stations at H+1 [14], so that linear regression analysis can be carried out to find out how much capability rainfall data TRMM estimates BMKG rainfall data. The results of the linear regression analysis are as follows:

DOI:10.31629/jit.v3i2.5060

Table 6. Linear Regression Cumulative 7 Days Result

Station	*Time Lag*	Function	R^2
Kotabumi	H0	y = 0,7269x + 26,160	0,437
Maritim	H0	y = 0,6375x + 22,000	0,323
Masgar	H+1	y = 0,7677x + 20,862	0,498
Radin Inten II	H+1	y = 0,6932x + 23,203	0,439

From the results of the regression analysis, the value of the Determination Coefficient at Masgar Station was obtained with a value of 0.4978 with a medium scale so that the ability to know how much the BMKG value was on the Medium scale for cumulative 7 daily [15]. The results of the regression analysis for each station are as follows:

3.6 BMKG and TRMM Cumulative 15 Days Rainfall data Correlation with Timelag

The analysis was carried out by comparing rainfall data with a cumulative 15 days. The cumulative number of 15 days was analyzed as a means of comparison in analyzing irrigation data. For the analysis carried out in the correlation of 15-day cumulative rainfall data is the Pearson correlation analysis with Time Lag from 3 days before the comparison day to 3 days after the comparison day [16]. For the results of the correlation analysis, the following values are generated is as follow.

Table 7. Cumulative 15 Days Correlation with Time Lag

Station	Time Lag Correlation						
	D-3	D-2	D-1	D0	D+1	D+2	D+3
Kotabumi	0,69	0,73	0,74	**0,75**	**0,75**	0,72	0,71
Maritim	0,59	0,61	0,61	**0,61**	0,60	0,58	0,55
Masgar	0,69	0,70	0,75	0,77	**0,79**	0,77	0,76
Radin Inten II	0,67	0,69	0,70	0,71	**0,72**	0,70	0,68

From the results of the regression analysis, the value of the Determination Coefficient at Masgar Station was obtained with a value of 0.616 with a high scale so that the ability to find out how much the BMKG value was on a medium scale for a cumulative 15 days [17]. The results of the regression analysis for each station are as follows:

Tabel 8. Linear Regression Cumulative 15 Days

Station	*Time Lag*	Function	R^2
Kotabumi	H0	y = 0,8377x + 45,354	0,558
Maritim	H0	y = 0,6959x + 42,885	0,375
Masgar	H+1	y = 0,8539x + 37,946	0,616
Radin Inten II	H+1	y = 0,7278x + 46,973	0,523

From the results, it is found that the prediction value is on a low scale for Maritime Stations, while for Kotabumi and Radin Inten II Stations it is on a medium scale, and for Masgar Station it is on a strong scale for a cumulative 15 days [18]. With the greatest capabilities located at Masgar Station with a value of 0.616, Radin Inten II with 0.523, Kotabumi 0.558 and Maritim 0.375.

DOI:10.31629/jit.v3i2.5060

3.7 BMKG and TRMM Cumulative 30 Days Rainfall data Correlation with Timelag

The analysis was carried out by comparing rainfall data with a cumulative 30 days. This is done by testing whether the more cumulative days of rain the greater the correlation value, this is also done to prove how much the daily cumulative affects the value in Time Lag [19], this correlation analysis is carried out with Time Lag D-3 to D+3 from the day that becomes comparison in correlation values. The correlation result is as follow:

Table 9. Cumulative 30 Days Correlation with Timelag

Station	Time Lag Correlation						
	D-3	**D-2**	**D-1**	**D0**	**D+1**	**D+2**	**D+3**
Kotabumi	0,78	0,80	0,80	0,80	**0,81**	0,80	0,79
Maritim	0,70	0,71	0,71	**0,71**	0,70	0,69	0,67
Masgar	0,77	0,77	0,79	0,80	**0,81**	0,81	0,81
Radin Inten II	0,76	0,76	0,76	**0,77**	0,77	0,76	0,77

From the results of the correlation analysis, it was found that the largest correlation value was at H+1 for Kotabumi Station, H0 Maritim, H0 Radin Inten II while H+1 for Masgar Station, so that linear regression analysis can be carried out as follows:

Table 10. Linear Regression Cumulative 30 Days Result

Station	*Time Lag*	Function	R^2
Kotabumi	H + 1	y = 0,9025x + 77,993	0,649
Maritim	H0	y = 0,8683x + 60,588	0,511
Masgar	H + 1	y = 0,8802x + 71,762	0,657
Radin Inten II	H0	y = 0,7535x + 89,864	0,593

From the results, it was found that the predictive ability value was on a medium scale for Maritime Station and Radin Inten II [20], a strong scale for Kotabumi Station and Masgar Station with a cumulative 30 days. With the greatest capabilities located at Masgar Station with a value of 0.657, Radin Inten II with 0.593, Kotabumi 0.649 and Maritim 0.657.

4. Conclusion

From the results of the research and discussion, to overcome the missing rainfall data, it can be done by utilizing the TRMM (Tropical Rainfall Measuring Mission) satellite technology. In a study conducted in Lampung Province, it was concluded that the rainfall TRMM and BMKG at each station had a better correlation and regression at the time lag H+1 and without time delay (H0) with a cumulative value of 0.3602 – 0.4808 for the cumulative daily, 0.417 – 0.555 for cumulative 2 days, 0.569 – 0.706 for cumulative 7 days, 0.612 – 0.785 for cumulative 15 days and 0.715 – 0.811 and for cumulative 30 days.From the results of linear regression analysis of TRMM and BMKG rainfall data with Time Lag that the predictive ability of BMKG rainfall data is at H0 and H+1 which have the best correlation. So that there is an increase in predictive ability on the Time Lag.

References

[1]. Aditya, F., Sampurno, J., & Ihwan, A. Modifikasi Estimasi Curah Hujan Satelit TRMM Dengan Metode Jaringan Syaraf Tiruan Propagasi Balik Studi Kasus Stasiun Klimatologi Siantan. POSITRON, 4(2).

[2]. Akbar, P S, dan Gunawan, D. 2004. Prediksi Curah Hujan Bulanan Stasiun BMKG Di Provinsi Bengkulu Menggunakan Teknik Downscalling Statistik. Jurnal Meteorologi Klimatologi dan Geofisika 1(1):1–12.

[3]. Wilhelmus, B., Denik S., Krisnayanti, dan Yacobus A. Klau,2013. Analisis Hubungan Tebal Hujan dan Durasi Hujan pada Stasiun Klimatologi Lasiana

Kota Kupang. Jurnal Teknik Sipil Vol.II No. 2. Universitas Nusa Cendana. Kupang.

[4]. Cao, Y., Zhang, W., Wang, W. 2017. Evaluation of TRMM 3B43 data over the Yangtze River Delta of China. Scientific Reports.

[5]. Chapra, S. C, Canale, R. P. 1988. Numerical Methods for Engineers 2nd Edition.McCraw- Hill.

[6]. Fitria, Welly, dan Guntur Prayogo, 2019. Pendugaan Data Kosong Curah Hujan Observasi Dengan Metode Regresi Piecewise. Departemen Statistika, FMIPA, Institut Pertanian Bogor.

[7]. Gunawan, Dodo. 2008. Perbandingan Curah Hujan Bulanan dari Data Pengamatan Permukaan , Satelit TRMM dan Model Permukaan NOAH. Jurnal Meteorologi Dan Geofisika 9(1):1–10.

[8]. Ginting, Jody Martin, Joko Sujono, dan Rachmad Jayadi, 2019. Analisis Hubungan Data Hujan Satelit dengan Hujan Terukur ARR Kalibawang. Prosiding Konferensi Nasional Pascasarjana Teknik Sipil (KNPTS) x 2019. ISSN 2477-00-86

[9]. Hadi, Arif Ismul dan Suwarsono Herliana. 2006. Analisis Karakteristik Intensitas

[10]. Curah Hujan di Kota Bengkulu. Jurnal Fisika FLUX 7(0736):119–29. Jarwanti, Dieta Putri, Ery Suhartanto, dan Jadfan Sidqi Fidari, 2021. Validasi Data Curah Hujan Satelit TRMM (Tropical Rainfall Measuring Mission) dengan Data Pos Penakar Hujan di DAS Grindulu, Kabupaten Pacitan, Jawa Timur. Journal Teknologi dan Rekayasa Sumber Daya Air, Vol.1 No.2 (2021) p.772-785.

[11]. Kurniawan, Agusta. 2020. Evaluasi Pengukuran Curah Hujan Antara Hasil Pengukuran Permukaan (AWS, HELLMAN, OBS) dan Hasil Estimasi (Citra Satelit =GSMaP) Di Stasiun Klimatologi Mlati Tahun 2018. Jurnal Geografi, Edukasi dan Lingkungan (JGEL) Vol.4, No.1 Januari 2020: 1 - 7.

[12]. Marpaung, Sartono, Didi Satiadi, dan Teguh Harjana, 2012. Analisis Kejadian Curah Hujan Ekstrem di Pulau Sumatera Berbasis Data Satelit TRMM dan Observasi Permukaan. Jurnal Sains Dirgantara, Vol.9 No.2 Juni 2012: 127-138.

[13]. Mashudi, Imam, M.Anwar, dan Fengky F.Adji, 2020. Pemanfaatan Data Satelit Tropical Rainfall Measuring Mission (TRMM) Untuk Pemetaan Zona Agroklimat Neraca Air Lahan di Kalimantan Tengah. Journal of Environment and Management, E-ISSN 2722-6727.

[14]. Nachrowi, D.N, dan Hardius U. 2008. Pengunaan Teknik Ekonometrika. Jakarta: Raja Grafindo Persada.

[15]. Noor, Riza Arian, M. Ruslan, Gusti Rusmayadi, dan Badaruddin,2016. Pemanfaatan Data Satelit Tropical Rainfall Measuring Mission (TRMM) untuk Pemetaan Zona Agroklimat Oldeman di Kalimantan Selatan. EnviroScienteae Vol.12 No.3, November 2016. p-ISSN 1978-8096.

[16]. Orfa, Lourina Evanale, dan Abdul Samad, 2018. Kajian Penggunaan Data TRMM (Tropical Rainfall Measuring Missions) Untuk Analisis Hidrologi (Studi Kasus pada Bendungan Lolak Kab. Bolaang Mongondow, Sulawesi Utara). Seminar Nasional Teknologi dan Rekayasa (SENTRA) 2018.ISSN 2527-6042.

[17]. Pangestu, Ibnu Tuhu, 2019. Analisis Korelasi Data Curah Hujan BMKG dengan TRMM (Studi Kasus Stasiun BMKG diSumatera Utara). JRSDD, 8(1), 157-180.

[18]. Pratiwi, Destiana Wahyu, Joko Sujono, dan Adam Pamudji Rahardjo, 2017.Evaluasi Data Hujan Satelit Untuk Prediksi Data Hujan Pengamatan Menggunakan Cross Correlation. Jurnal.umj.ac.id/index.php/semnastek.

[19]. Renggono, F dan D. Syaifullah (2011). Kajian Meteorologis Bencana Banjir Bandang di Wasior, Papua Barat. Jurnal Meteorologi dan Geofisika, 12(1), 33-41

[20]. Wong, W. F. J., Chiu, L. S. 2008. Spatial and Temporal Analysis of Rain Gauge Data and TRMM Rainfall Retrievals in Hong Kong. Geographic Information Sciences, 14(2):105-112. World Meteorological Organization. 2008. Guide to Hydrological Pratices VolumeI. WMO -No.168.