The ratio of Indonesia BMKG agency and TRMM satellite Rainfall Data in West Java Province by Using Statistical Parameter and Correlation Analysis

Rian Alfian[1], Ahmad Zakaria[1], Endro Prasetyo Wahono[1], Dyah Indriana Kusumastuti[1], Ahmad Herison[1]

[1]*Departement of Civil Engineering, Faculty of Engineering, Universitas Lampung, Lampung, Indonesia*

[*]*Email: rian19922017@gmail.com*

Article Information:

Received:
24 September 2021

Received in revised form:
14 November 2021

Accepted:
24 December 2021

Volume 3, Issue 1, December 2021
pp. 51 – 60

http://doi.org/10.23960/jesr.v3i2.82

Abstract

Weather observations can be done in two ways, namely weather observations based on weather stations and based on remote sensing such as satellites. One of these weather study data is rainfall measured from the Indonesia BMKG agency and TRMM satellite rain observations. To see the pattern of the distribution of rain that has occurred, the two weather observations can be connected as a reference for the distribution of rain. The purpose of this study was to analyze the correlation value of rainfall data between BMKG and TRMM by looking at the comparison graph and analyzing the comparison of statistical parameters. This research was conducted using daily rainfall data from 1998-2018 at four rain stations in West Java Province and taking descriptive decisions in the form of pictures and graphs in the form of daily, monthly and annual data. Based on the analysis results, the largest correlation value is in the annual cumulative with a value of 0,88-0,94, the smaller the number of days, the smaller the correlation value. The BMKG and TRMM rainfall data have relatively the same pattern, but the maximum data have differences so that it reduces the correlation value. In the Statistical Parameter analysis, it can be stated that the difference in the values of the statistical parameters is directly proportional to the daily cumulative, that the difference in the median and mean values in the BMKG and TRMM data is greater in the data with a larger cumulative. Meanwhile, for linear regression analysis, it was found that the greatest value was the cumulative 1-year value with a coefficient of determination from 0,78 - 0,89.

Keywords: *BMKG agency, TRMM satellite, rainfall, regression analysis.*

I. Introduction

Rainfall that occurs in the tropics, namely Indonesia, has influential elements such as meteorology compared to other factors. The variation of rainfall in the territory of Indonesia is very large both spatially and temporally [1]. In general, to know accurate rainfall and can be used in engineering planning, rainfall data is needed to support information in the form of temporal (time series) and spatial (spatial) [2]. The amount of rain as a result of measurements with a rain gauge for some time years can be used to determine the nature (characteristics) of rainfall in a place [3]. However, there is also the concept of probability where the characteristics of rain based on time are ignored and the calculation will only be profitable if the data processing is long enough [4].

Good rainfall data can be obtained from recordings that are maintained and always monitored. Whether recording or observations can be done in two ways, namely weather observations based on weather stations and weather observations based on remote sensing such as satellites. As well as rainfall data originating from the meteorological observation post owned by the Meteorology, Climatology and Geophysics Agency (BMKG) it can predict weather conditions in areas such as West Java Province with a maximum average temperature of 32,2ºC and a minimum temperature of 18.2ºC.

West Java Province is geographically located between 5º50' – 7º50' South Latitude and 104º48' –

108°48' East Longitude [5]. West Java Province is a land area that is distinguished by steep mountainous areas, gently sloping hillsides, broad plains, and watersheds. Rainfall measurements were carried out at rain observation posts in the West Java Province. The area consists of several rain observation points, but not all rain observation points are used in this study.

The distribution of rain is not only seen from the measured rainfall data from the BMKG rain observation post but also satellite observations are needed to see the pattern of rain distribution that has occurred, namely by looking at the data from the TRMM. NASA's Tropical Rainfall Measurement Mission (TRMM) was the first coordinated international effort to provide reliable rainfall measurements from space [6]. TRMM data is precipitation data (rain) obtained from the TRMM meteorological satellite with several sensors in it. Based on this, several studies have been obtained using TRMM satellite data which can verify rainfall data from the TRMM satellite and at the BMKG rain observation post in West Java Province.

Several studies have shown that monthly rainfall predictions in the western part of Java are still in areas with low accuracy, especially in the mountains [7]. This also affects the suitability of the rain pattern between the observed rainfall data and the estimated TRMM rainfall [8]. In this study, we will present the relationship or correlation between the rainfall data from the BMKG rain observation post and the TRMM data which is usually used as a reference for the distribution of rain. The rainfall for each rain observation post in the area is then summed and averaged to get the average rainfall value for the area. The purpose of this study is to analyze the correlation value of BMKG rainfall data and TRMM data and analyze the comparison of statistical parameters and the ability of the rainfall data.

II. Materials and Methods

The stages of this research consist of the initial preparation stage, the data processing stage, and the data analysis stage. The research flow chart is shown in Figure 1. This study uses secondary data from related parties such as Indonesia BMKG agency and TRMM satellite at West Java Province at four stations namely Jatiwangi Station, Citeko Station, Bandung Geophysics Station, and Bogor Station. The data used is rainfall data with a span of 20 years, the data range from 1998 - 2018 which is downloaded from the official website of BMKG and TRMM West Java Province.

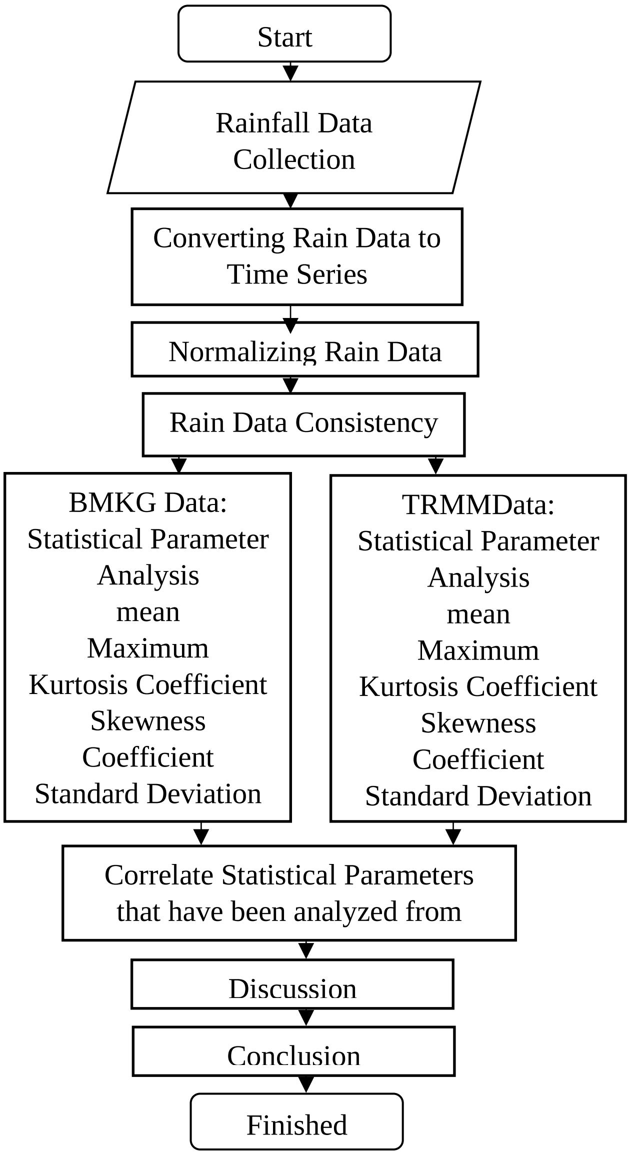

Figure 1. Research Flowchart

A. Procedure

Processing data in this study has several stages, including changing the BMKG station rainfall data and TRMM data into time series data, Normalizing rainfall data from the Meteorology, Climatology, and Geophysics Agency, conducting a consistency test for rain data that has been done searching for missing data. To find out whether this data is suitable for use as research data, Classify daily cumulative rainfall data, 7 daily, 15 daily, monthly, annual average, annual maximum, Analyzing rainfall data with statistical parameters namely Mean, Variance, Standard Deviation, Coefficient Skewness, and Kurtosis Coefficient, Correlate rainfall data from two different sources from different classifications, Correlate statistical parameters of rainfall data from two different rainfall data, Perform rainfall data modeling with regression analysis for BMKG agency and TRMM satellite data.

B. Analysis

The analysis of the research results is in the form of correlation values and also descriptive decision-making according to the results of the analysis. The results are pictures and graphs from each rain station and also from different methods in West Java Province. After obtaining the results of this study, a discussion was carried out to draw conclusions based on the research rules

III. Results and Discussions

A. Converting Rain Data into the Time Series

Precipitation data from the BMKG before being used in the study was still in the form of a time series with many empty days. In transforming rain data that are experiencing shortages, a program is used so that the process of filling in days that are not recorded and empty data can be processed much faster. The program used in this research is a program created using Force. This program is used to transform BMKG rainfall data that can be correlated well and by TRMM rainfall data.

B. Consistency Test Data

This test is used as one of the parameters that the rain data used is good data by looking at the possibilities of the data that is owned has a value that is far from its value in general. The consistency Test is carried out by adding up each existing daily data for a year or annual cumulative. Consistency tests were also carried out on the TRMM rainfall data, this was done to ensure that the TRMM rainfall data values did not have significant errors. The analyzed annual cumulative is used in determining the consistency test. This test is carried out by doing the cumulative last year compared to the cumulative of other stations. This annual cumulative value is the value used in linear regression analysis.

C. Correlation Analysis of BMKG Rainfall Data and Daily of TRMM

Correlation analysis of daily rainfall data was carried out using daily cumulative data. This comparison is used to see how much the pattern similarity value between BMKG data and TRMM data. This analysis was compared with 3 stages, namely Pearson correlation analysis, statistical data comparison, Curve Fitting, and linear regression analysis. The results of Pearson's correlation analysis show that from 4 rain stations in West Java, the correlation value is from 0,08 – 0,22. This states that the similarity of rainfall patterns between BMKG data and TRMM data on the same day has a low-scale correlation. This was also stated by the researcher where the results of the analysis obtained if the rainfall data measured by TRMM had the same temporal distribution pattern of rainfall as that measured by BMKG [9].

Analyzing the correlation, Curve Fitting was also carried out to see the comparison of the BMKG and TRMM rainfall data, for this reason, a comparison of the rain data was carried out in the graphical form. The results of the graph comparisons carried out show that the data graphs have similar patterns, but the maximum value of each data has a significant difference so that these values have a small correlation value.

In addition, for statistical parameter analysis, it was found that the mean value or average value of the BMKG and TRMM rainfall data has a difference with a scale of 0,35 - 1,44, for the median value of BMKG and TRMM data has a difference with a scale of 0. .05 to 1.91 states that the average value and the daily mean value of the BMKG and TRMM results have a small difference. the results of statistical parameter analysis can also be seen that the characteristics of the data have the same tendency, namely that the distribution of the data when viewed through the value of the skewness coefficient that all data has a positive distribution graph. As for the results of the kurtosis coefficient, all rain data has a Leptokurtic curve steepness level. The value of variance and standard deviation states that the data has a level of dispersion away from the zero value. It states that data values have dynamic variations in the amount of data. So that the analyzed rain data tends to move away from the average value of rainfall.

The results of the regression analysis show that the coefficient of determination or the ability of the regression function to model the BMKG rainfall data value from the TRMM rainfall data has a very low scale, it can be concluded that the TRMM daily rainfall data value can only predict the BMKG value with an accuracy rate of 1% - 5%. The results can be seen in the following figure 1 to 4.

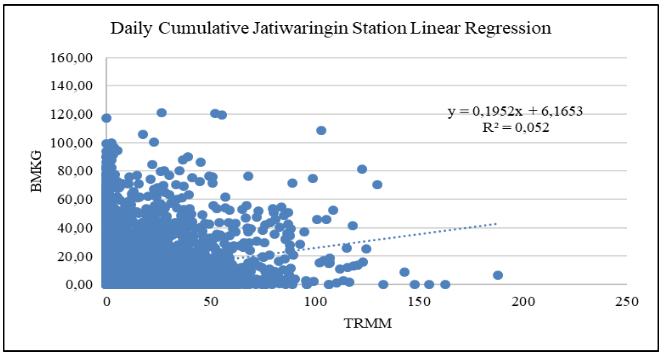

Figure 1. Daily Cumulative data at Jatiwaringin Station with Linear Regression Chart.

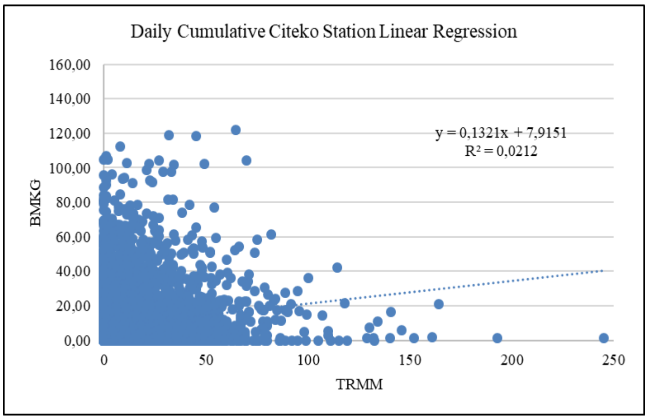

Figure 2. Daily Cumulative data at Citeko Station using Linear Regression Chart.

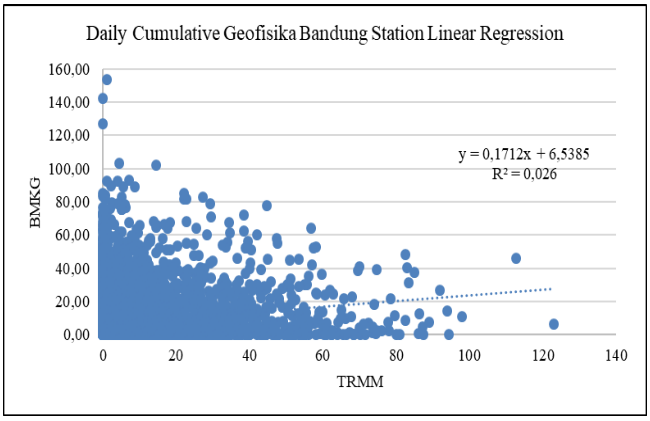

Figure 3. Daily Cumulative data from Station of Geofisika Bandung with Linear Regression Chart.

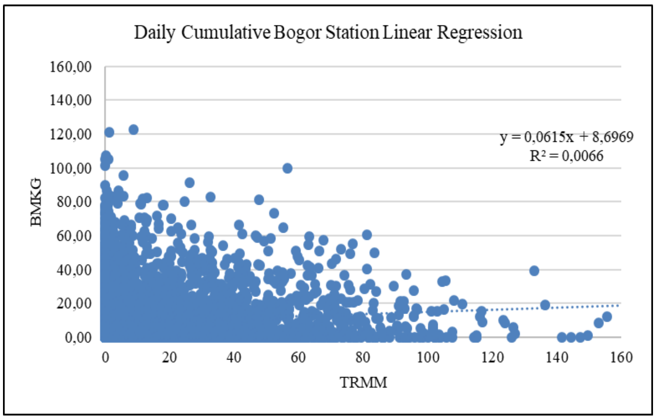

Figure 4. Daily Cumulative Bogor Station Linear Regression Chart.

D. Correlation Analysis of BMKG and TRMM 2 Daily Rainfall Data

Correlation analysis with a cumulative 2 days to see how much change occurs if the data used is data with a cumulative 2 days. The results of the correlation analysis stated that the correlation value produced was much better than the correlation value from the daily cumulative results. The results of the 2-day cumulative correlation analysis have a low scale with a value of 0,27 – 0,39. So it is stated that the 2 daily results are much better than the daily cumulative results. The graphical comparison between the two-rainfall data, namely BMKG and TRMM, it is stated that the values of the BMKG and TRMM rainfall data have the same pattern, but the maximum value of the BMKG rainfall data has a higher tendency than the TRMM data.

Statistical parameters obtained that the average value (Mean) and the median value (Median) of the BMKG and TRMM data have a greater difference than the daily cumulative data, namely 0,7 – 2,88 for the average value and 2,64 - 3,23 for the middle value. The variance and standard deviation values still have the same tendency, which is away from the zero value, so it can be concluded that the data has a diverse and dynamic distribution. The value of the Skewness Coefficient also states that the value has a positive distribution so it can be concluded that the graph tends to the right. From the results of the Kurtosis Coefficient, it is also shown that the value has the characteristic level of the steepness of the Leptokurtic curve. For the results of the regression analysis that has been produced, it is stated that the coefficient of determination or the model's ability to predict BMKG rainfall data is 0,07 – 0,15 with a very low scale. The results can be seen in the following figure 5 - 8.

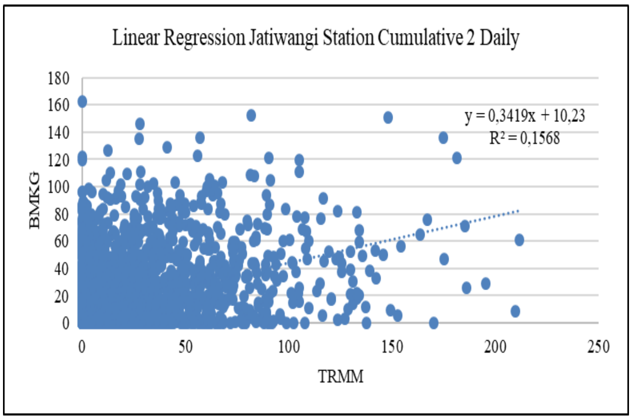

Figure 5. Linear Regression Graph Jatiwangi Station Cumulative 2 Daily.

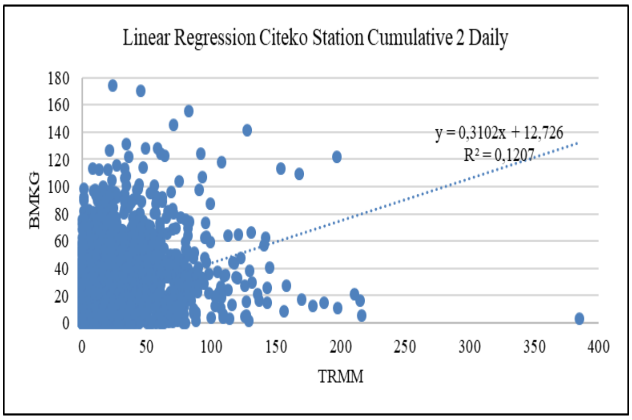

Figure 6. Linear Regression Graph Citeko Station Cumulative 2 Daily.

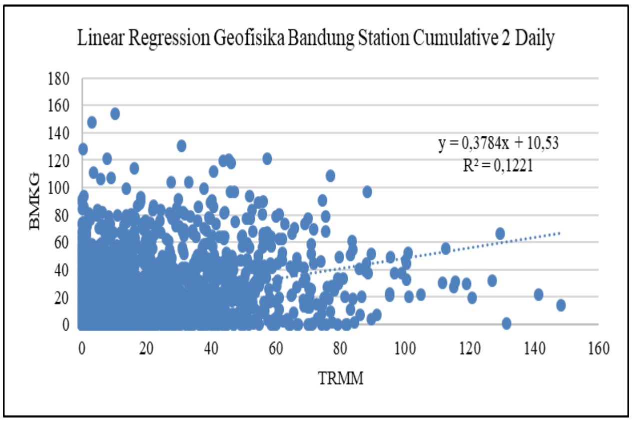

Figure 7. Linear Regression Graph Geofisika Bandung Station Cumulative 2 Daily.

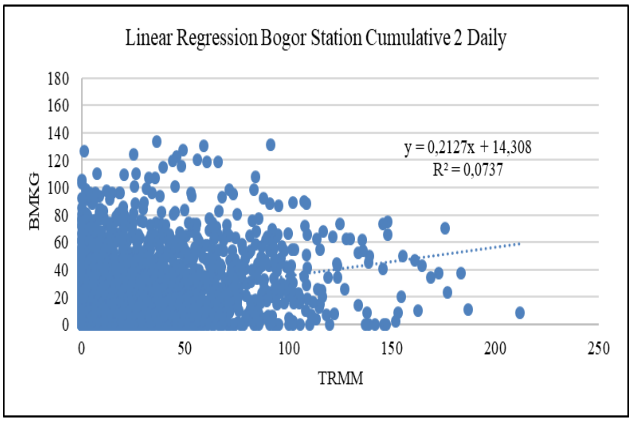

Figure 8. Linear Regression Graph Bogor Station Cumulative 2 Daily.

E. Correlation Analysis of BMKG and TRMM 7 Daily Rainfall Data

Analysis of 7-day rainfall is carried out 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 of statistical parameters. The correlation value shows that the similarity of rainfall data patterns in BMKG data to TRMM data has a value that tends to increase compared to the 2-day data. From the value, it is also stated that the correlation value is on a medium scale with the smallest correlation value being Bogor Station and the largest is Jatiwangi Station. Comparisons with graphs have been compared and it can be concluded that the values have a similar pattern, but it can also be seen that the maximum value of the rain data has a value that differs greatly between BMKG and TRMM so this maximum value may cause the correlation value to be on a medium scale.

Comparison of statistical parameter values from BMKG and TRMM data, it was found that the difference between BMKG and TRMM stations for the mean value was 2,4 – 10,12 and 1,53 – 11,6 for the median value. This concludes that the greater the cumulative days used, the greater the difference in value. From the results of the analysis of standard deviation and variance, it is stated that they still have the same tendency, which is away from the zero value, so it can be concluded that the data has a diverse and dynamic distribution.

The value of the Skewness Coefficient also states that the value has a positive distribution so it can be concluded that the graph tends to the right. From the results of the Kurtosis Coefficient, it is also shown that the value has the characteristic level of the steepness of the Leptokurtic curve. As for the results of the regression analysis, it was found that the coefficient of determination or the model's ability to predict BMKG rainfall data was 0,19 – 0,33 with a very low to low scale. The results can be seen in the following figure 9 - 12.

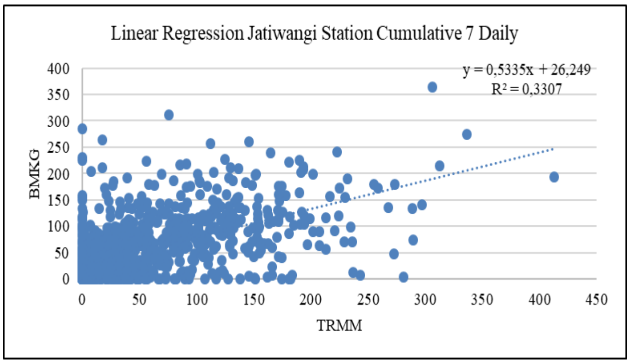

Figure 9. Linear Regression Graph Jatiwangi Station Cumulative 7 Daily.

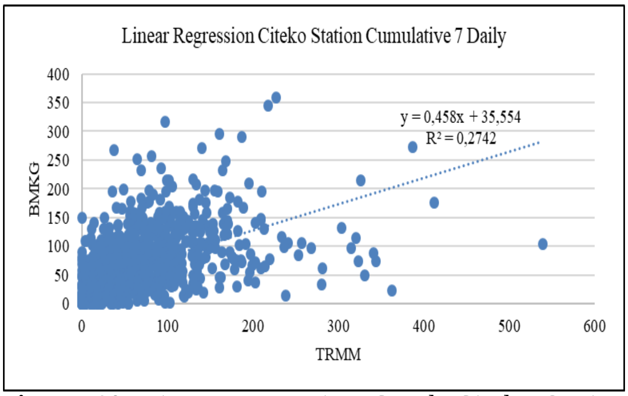

Figure 10. Linear Regression Graph Citeko Station Cumulative 7 Daily.

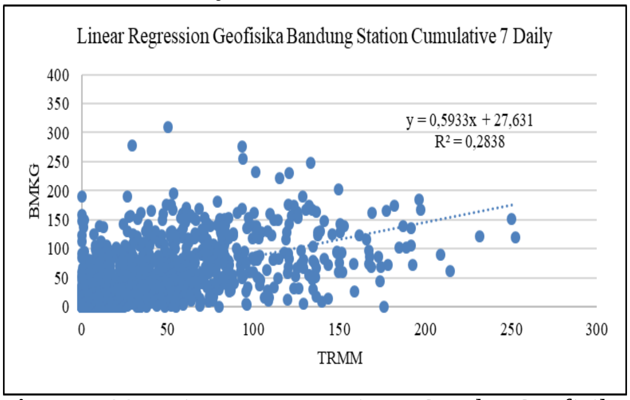

Figure 11. Linear Regression Graph Geofisika Bandung Station Cumulative 7 Daily.

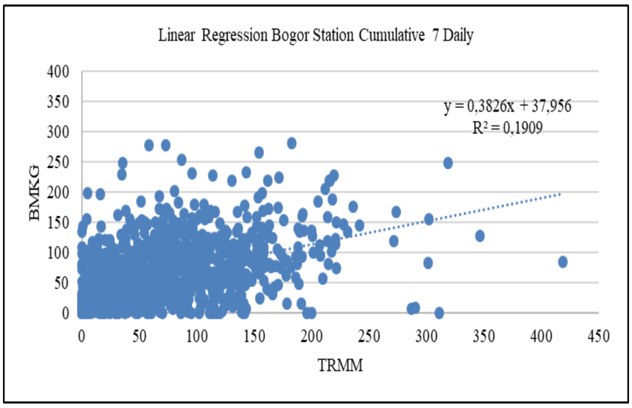

Figure 12. Linear Regression Graph Bogor Station Cumulative 7 Daily.

F. Correlation Analysis of BMKG and TRMM 15 Daily Rainfall Data

Analysis with a cumulative total of 15 days was analyzed as a means of comparison in analyzing rainfall data. The correlation value shows that the similarity of rainfall data patterns in BMKG data to TRMM data has a value that tends to increase compared to 7 daily data. From the value, it is also stated that the correlation value is on a high scale except for Bogor station data on a medium scale with the smallest correlation value being Bogor Station and the largest being Jatiwangi station. The comparison of the graphs can be concluded that the values have a similar pattern but it can also be seen that the maximum value of the rain data has a much different value between BMKG and TRMM so this maximum value may cause the correlation value to be on a medium to high scale.

Analysis of statistical parameter values found that the difference between the BMKG and TRMM stations for the mean value was 5,4 – 21 and 8,2 – 43,13 for the median value. This concludes that the greater the cumulative days used, the greater the difference in value. From the results of the analysis of standard deviation and variance, it is stated that they still have the same tendency, which is away from the zero value, so it can be concluded that the data has a diverse and dynamic distribution.

The value of the Skewness Coefficient also states that the value has a positive distribution so it can be concluded that the graph tends to the right. The results of the Kurtosis Coefficient also show that each station has a characteristic level of the steepness of the Leptokurtic curve except for the Jatiwangi TRMM station with Platykurtic characteristics. For regression analysis, it is stated that the coefficient of determination or the model's ability to predict BMKG rainfall data is 0,22 – 0,46 with a low and medium scale. The results can be seen in the following figure 13 - 16.

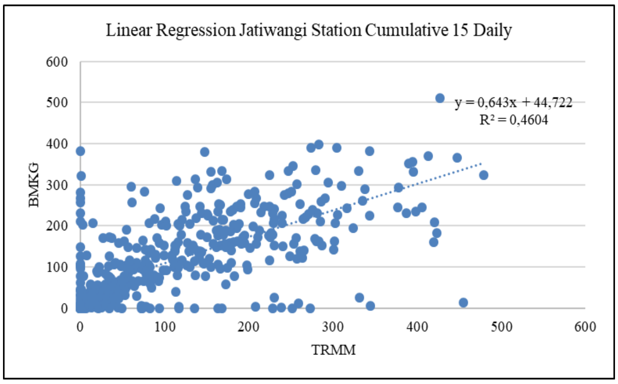

Figure 13. Linear Regression Graph Jatiwangi Station Cumulative 15 Daily.

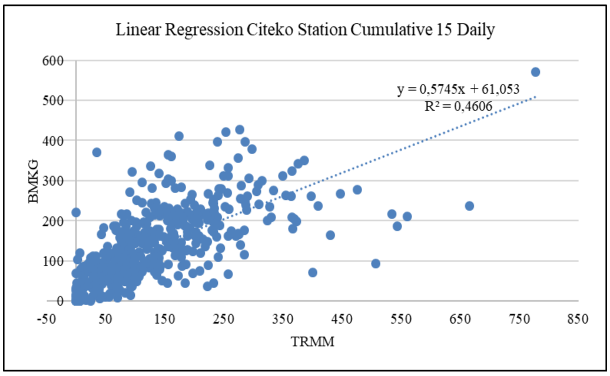

Figure 14. Linear Regression Graph Citeko Station Cumulative 15 Daily.

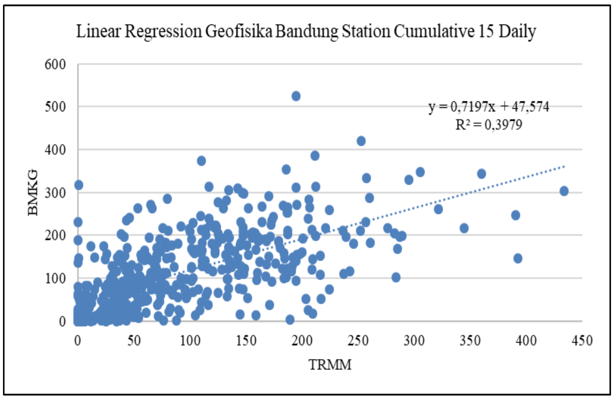

Figure 15. Linear Regression Graph Geofisika Bandung Station Cumulative 15 Daily.

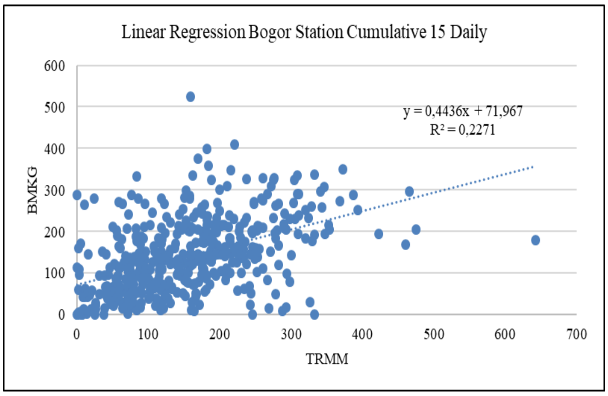

Figure 16. Linear Regression Graph Bogor Station Cumulative 15 Daily.

G. Correlation Analysis of BMKG and TRMM 30 Daily Rainfall Data

This analysis is carried out 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 of statistical parameters. The correlation value shows that the similarity of rainfall data patterns in BMKG data to TRMM data has a value that tends to increase compared to 15 daily data.

From the value, it is also stated that the correlation value is on a high scale except for Bogor station data on a medium scale with the smallest correlation value being Bogor Station and the largest being Citeko Station. The comparison of the results of the graph of the 30-day data pattern shows that the values have a similar pattern, but it can also be seen that the maximum value of the rain data has a much different value between BMKG and TRMM so this maximum value may cause the correlation value to be on a medium to high scale.

Statistical parameter analysis showed that the difference between the BMKG and TRMM stations for the mean value was 10,99 – 43,36 and 8,04 – 43,78 for the median value. This concludes that the greater the cumulative days used, the greater the difference in value. From the results of the analysis of the standard deviation and variance, it is stated that it still has the same tendency, namely away from the zero value, so it can be concluded that the data has a diverse and dynamic distribution.

The value of the Skewness Coefficient also states that the value has a positive distribution, so it can be concluded that the graph tends to the right. From the results of the Kurtosis Coefficient, it is also shown that each station has a characteristic steepness level of the Leptokurtic curve except for the Jatiwangi station with Platykurtic characteristics for Indonesia BMKG agency data, while for TRMM satellite for rainfall data all stations have a different sharpness level from BMKG agency, namely with Platykurtic characteristics.

While the regression analysis resulted that the coefficient of determination or the model's ability to predict BMKG rainfall data was with a value of 0,27 – 0,57 with a low to medium scale. The results can be seen in the following figure 17 - 20.

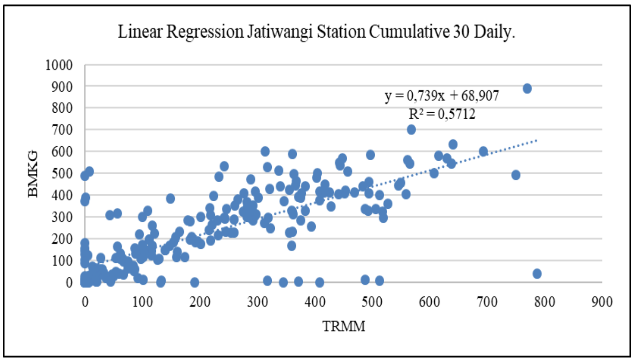

Figure 17. Linear Regression Graph Jatiwangi Station Cumulative 30 Daily.

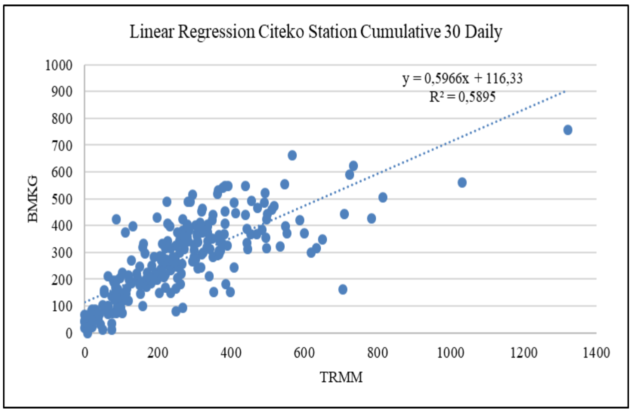

Figure 18. Linear Regression Graph Citeko Station Cumulative 30 Daily.

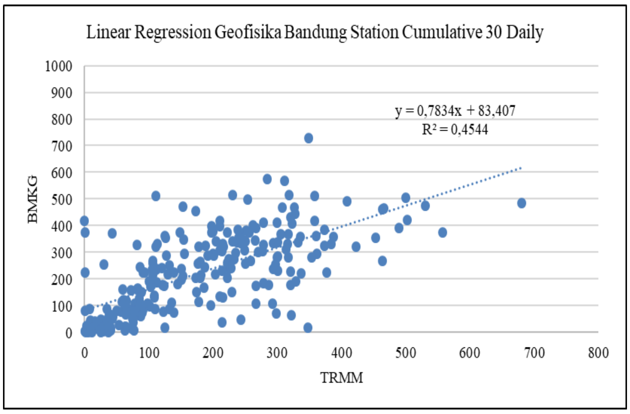

Figure 19. Linear Regression Graph Geofisika Bandung Station Cumulative 30 Daily.

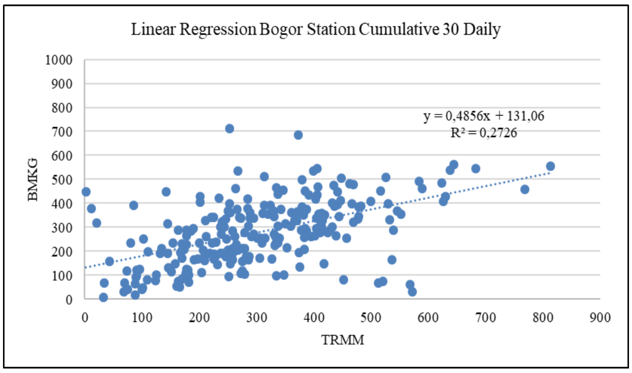

Figure 20. Linear Regression Graph Bogor Station Cumulative 30 Daily.

H. Correlation Analysis of BMKG and TRMM 1 Year Rainfall Data

This analysis is carried out 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 of statistical parameters. The correlation value shows that the similarity of rainfall data patterns in BMKG data to TRMM data has a value that tends to increase compared to 1-year data. From the value, it is also stated that the correlation value is on the High scale except for the Citeko station data on a very high scale with the smallest correlation value being Bogor Station and the largest being Bogor Station. Comparison of the graphs shows that the values have a similar pattern, but it can also be seen that the maximum value of the rain data has a much different value between BMKG and TRMM so this maximum value may cause the correlation value to be on a medium to high scale.

Statistical parameter analysis stated that the difference between BMKG and TRMM stations for the mean value was 76,82 – 484,96 and 213,28 – 592,84 for the median value. This concludes that the greater the cumulative days used, the greater the difference in value. From the results of the analysis of the standard deviation and variance, it is stated that they still have the same tendency, namely away from the zero value, so it can be concluded that the data has a diverse and dynamic distribution.

The value of the Skewness Coefficient also states that the value has a positive distribution, so it can be concluded that the graph tends to the right. From the results of the Kurtosis Coefficient, it is also shown that each station has a characteristic level of sharpness of the Platikurtic curve. For regression analysis, it is found that the coefficient of determination or the model's ability to predict BMKG rainfall data is 0,78 – 0,89 with a very high scale. The results can be seen in the following figure 21 - 24.

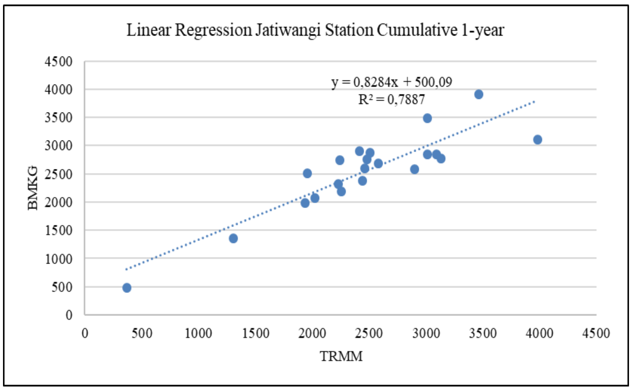

Figure 21. Linear Regression Graph Jatiwangi Station Cumulative 1-year.

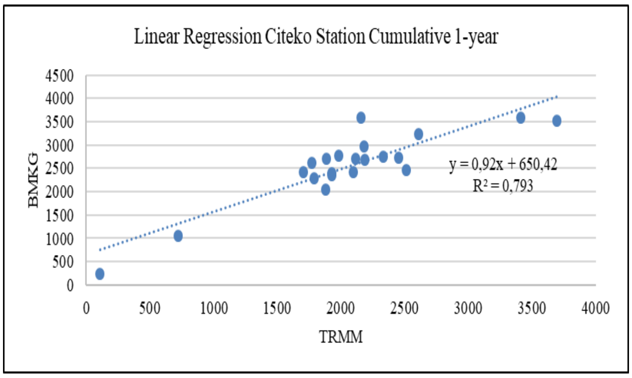

Figure 22. Linear Regression Graph Citeko Station Cumulative 1-year.

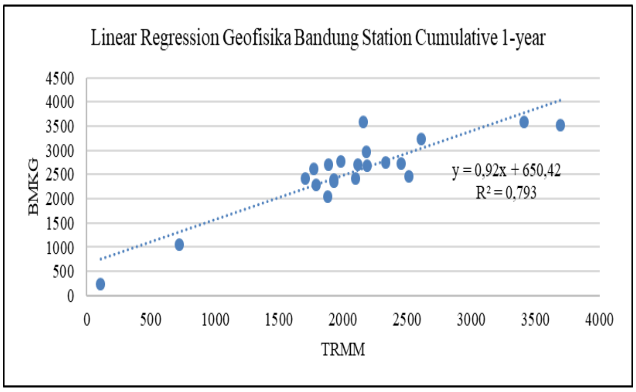

Figure 23. Linear Regression Graph Geofisika Bandung Station Cumulative 1-year.

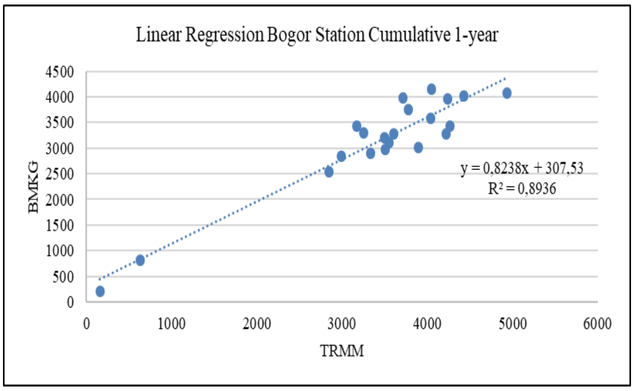

Figure 24. Linear Regression Graph Bogor Station Cumulative 1-year.

Based on the results of the regression analysis that has been carried out from daily to annual data, the equation results for each daily to annual rain station are presented in table 1.

Table 1. Results of Regression Analysis Equations at Each Rain Station.

Rain Station	Regression Analysis Equation					
	day	2 days	7 days	15 days	30 days	1-year
Jatiwangi	y=0,1952x + 6,1653 R^2 = 0,052	y=0,3419x + 10,23 R^2 = 0,1568	y=0,5335x + 26,249 R^2 = 0,3307	y=0,643x + 44,722 R^2 = 0,4604	y=0,739x + 68,907 R^2 = 0,5712	y=0,8284x + 500,09 R^2 = 0,7887
Citeko	y=0,1321x + 7,9151 R^2 = 0,0212	y=0,3102x + 12,726 R^2 = 0,1207	y=0,458x + 35,554 R^2 = 0,2742	y=0,5745x + 61,053 R^2 = 0,4606	y=0,5966x + 116,33 R^2 = 0,5895	y=0,8625x + 521,45 R^2 = 0,7858
Geofisika Bandung	y=0,1712x + 6,5385 R^2 = 0,026	y=0,3784x + 1053 R^2 = 0,1221	y=0,5933x + 27,631 R^2 = 0,2838	y=0,7197x + 45,574 R^2 = 0,3979	y=0,7834x + 83,407 R^2 = 0,4544	y=0,92x + 650,42 R^2 = 0,793
Bogor	y=0,0615x + 8,6969 R^2 = 0,0066	y=0,2127x + 14,308 R^2 = 0,0737	y=0,3826x + 37,956 R^2 = 0,1909	y=0,4436x + 71,967 R^2 = 0,2271	y=0,4856x + 131,06 R^2 = 0,2726	y=0,8238x + 307,53 R^2 = 0,8936

Source: Analysis Results, 2021

The results of research and discussion shows that from the results of the correlation analysis, it can be stated that the largest correlation value is in the annual cumulative with a correlation value of 0,88 – 0,94 while the lowest correlation value is in the daily cumulative between 0,08 – 0,22.

The correlation value obtained from the analysis is directly proportional to the cumulative amount, namely the smaller the number of days, the smaller the correlation value. The results of the graphical comparison analysis show that the pattern formed from the rainfall data from BMKG and TRMM has a relatively similar pattern, but the maximum data has differences so that it reduces the correlation value [9].

IV. Conclusions

The results of the Statistical Parameter analysis stated that the greater the difference between the median and mean values in the Indonesia BMKG agency and TRMM satellite data, the greater the cumulative value. The characteristics of the standard deviation, variance, and skewness coefficient have the same value, but the value of the standard deviation and variance increases as the number of days is accumulated. The kurtosis coefficient has a difference in value that is directly proportional to the number of days, but the characteristics of the kurtosis coefficient will be more different on a larger cumulative day. From the results of linear regression analysis, it was found that the greatest value was the cumulative 1-year value with a coefficient of determination from 0,78 - 0,89. The low coefficient of determination, the daily cumulative value is 0,01 – 0,05.

Acknowledgment

The authors of this research gratefully acknowledge various parties who have contributed to the completion of this research. This research was realized to complete the Master of Civil Engineering at the University of Lampung.

References

[1] Gunawan, D. 2014. Perbandingan Curah Hujan Dari Data Pengamatan Permukaan, Satelit TRMM dan Model Permukaan NOAH. Puslitbang BMKG.

[2] Syaifullah, M. D. 2014. Validasi data TRMM terhadap data hujan aktual di tiga DAS di Indonesia. Puslitbang BMKG.

[3] Hadi, A. I., Suwarsono., dan Heliana. 2006. Analisis Karakteristik Intensitas Curah Hujan Di Kota Bengkulu. Jurnal Fisika FLUX.

[4] Zakaria, A. 2010. Studi Pemodelan Stokastik Curah Hujan Harian Dari Data Curah Hujan Stasiun Purajaya. Journal Proc. Seminar Nasional Sain Mipa dan Aplikasinya page 8-9.

[5] Badan Pusat Statistik Provinsi Jawa Barat 2021, Provinsi jawa Barat Dalam Angka Tentang Geografi dan Iklim.

[6] 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.

[7] Apriyanal. Y. dan Lindawati. 2015, Aplikasi Model Prediksi Curah Hujan Pada Dua Sentra Produksi Padi Di Jawa Barat, Informatika Pertanian, Vol. 24, No.2, 149 – 156.

[8] Wiratri, M., 2012. Perbandingan Data curah hujan observasi dan estimasi curah hujan dari satelit TRMM (*Trofical Rainfall Measuring Mission*) untuk wilayah Banjarbaru dan Kotabaru. Laporan Kerja Praktik S-1 Fisika. Universitas Lambung Mangkurat. Banjarbaru.

[9] Pangestu, I. T. 2019. Analisis Korelasi Data Curah Hujan Bmkg Dengan Trmm (Studi Kasus Stasiun Bmkg di Sumatra Utara). Skripsi. Fakultas Teknik Universitas Lampung.