COMPARATIVE ANALYSIS OF RAINFALL DATA ESTIMATION METHODS: A CASE STUDY OF TANGGAMUS REGENCY

Syahlan Naufal Fridayanto[1], Ahmad Zakaria[1]*, Dyah Indriana Kusumastuti[1], Endro Prasetyo Wahono[1]
[1]Department of Civil Engineering, University of Lampung, Jl. Prof. Dr. Sumantri Brojonegoro No. 1, Bandar Lampung, 35145
***E-mail:** *ahmad.zakaria@eng.unila.ac.id*

Received: April 6, 2024 Reviewed: July 5, 2024 Accepted: August 1, 2025

ABSTRACT

Rainfall data is vital for water resource systems and natural disaster management. It is required to have complete data to know the needs and structural requirements. However, in reality, plenty of cases of missing or broken rainfall data happened due to specific circumstances. To estimate missing rainfall data, several methods can be used, such as the Arithmetic Mean Method, Normal Ratio Method, Inversed Square Distance Method, and Modified Method. This study analyzed and compared Pearson's Correlation Coefficient and Root Mean Square Error (RMSE) among the methods. The result indicates that based on Pearson's Correlation Coefficient, the Modified Inversed Square Distance Method tends to be the best method for estimating missing rainfall data in the Tanggamus Regency. Several factors contribute to Pearson's Correlation Coefficient, such as rainfall data consistency and the elevation of stations used. Otherwise, based on the Mean Square Error, the Modified Arithmetic Mean Method is the best to use. This research also indicated that Pearson's Correlation Coefficient is better for evaluating missing data than RMSE.

Keywords: Hydrological Analysis, Missing Rainfall Data, Modified Method, Pearson's Correlation Coefficient, Tanggamus Regency.

1. Introduction

Precipitation within the tropics has powerful components, such as meteorology, compared to other variables. The sum of rain as a result of estimations with a rain gauge for a few times a long time can be utilized to decide the nature (characteristics) of precipitation in a specific location [1]– [3]. Within the application of logical calculations, precipitation is decided in mm/time, which can be deciphered as the depth of the surface immersion caused by the rain itself inside hours, day by day, week after week, or yearly time units [4], [5].

Rainfall data is important for water resource systems and natural disaster management [6]. It is crucial to have consistent and continuous rainfall data for statistical analysis. Accurate and timely observations of regional and global precipitation are crucial for various research and applications. The accuracy of the rainfall data is crucial for numerous research and analysis [7]. Rainfall data in time series recording can provide trend information from the nature of rain in a place, whether it has increased or decreased. However, in practice, the consistency of rainfall data may be disturbed or missed due to several factors, such as changes in observational procedure and incomplete records (missing observations) [8], [9].

Before estimating missing rainfall data, it is highly required to test the consistency of the rainfall data. The accuracy of rainfall-runoff models depends upon the consistency and continual rainfall data. But, when rainfall data is interrupted or even missing, it affects the quality of the models. There are several methods that can be used for determining data consistency. One of the common methods to test the consistency of rainfall data is the Double-Mass Curve (DMC), which has been applied in various hydrological analyses, including rainfall, runoff, and sediment as well [10]–[15].

Missing rainfall data often needs to be estimated to complete rainfall records or assist in areal rainfall calculations [16]. The missing rainfall data can occur in several amounts whether to a specific short duration or the whole data. However, for the most part, missing rainfall data often occurs in longer periods, making it difficult to estimate the missing data. Methods such as Arithmetic Mean Method (AM) [17]–[21], Normal Ratio (NR) [17]–[20], and Inversed Square Distance (ISD) [17], [20]–[22] are commonly used for estimating missing rainfall data in an area. Recently, researchers have developed various methods to enhance the accuracy of missing rainfall data that has been estimated. Hence, copious modification methods have been created.

This research uses four methods, the Arithmetic Mean Method (AM), Normal Ratio Method (NR), Inversed Square Distance Method (ISD), and Modification Method (M). These particular methods are preferred to estimate missing rainfall data in short-term periods. Furthermore, these methods are highly recommended in case the neighbouring stations are not too far away and are highly correlated with the specific target station [23]. This research is primarily focused on analysing missing rainfall data and comparing each method. Besides, the rainfall data of each station used in this research are correlated and relatively close.

The main objective of this research is to estimate missing rainfall data, ranging between 2013 and 2022, and compare all the methods used afterward, such as AM, NR, ISD, and M. The data used in this research must be tested for consistency using the Double-Mass Curve Method. The data estimation then being compared by Pearson's Correlation Coefficients and Mean Square Errors. Later on, it implies the best method that can be used to estimate missing rainfall data in a Tanggamus Regency, especially at each station used in this research. Moreover, this research aims to determine whether the Modified Method is better than conventional or common methods.

2. Methods

This study estimates missing rainfall data from five stations in Tanggamus Regency, Lampung Province, Indonesia, such as R067 Air Naningan, R284 Gunung Megang, R011 Banjar Agung, PH020 Gisting Atas, and R040 Bulok, as shown in Figure 1. All the rainfall data used in this research were obtained from Balai Besar Wilayah Sungai Mesuji Sekampung (BBWS-MS), with a range of 10 years from 2013 to 2022. The stations used in this research have various elevations (shown in Table 1 and Table 2).

Table 1. Elevation of each station.

Stations	Elevations (MASL)
R067 Air Naningan	557 m
R284 Gunung Megang	533 m
R011 Banjar Agung	199 m
PH020 Gisting Atas	545 m
R040 Bulok	411 m

This research mainly consists of three steps. The first one is to test the consistency of the rainfall data. After all of the data have been tested for consistency, it shows that the rainfall data is in good condition or very strong (Table 3). The next step is to estimate the missing rainfall data of each station using four methods. The last compares the observed and predicted data using Pearson's Correlation Coefficient and RMSE.

Double-Mass Curve (DMC). DMC is a method to check the data consistency. It is essential to some statistical research since data consistency is crucial for obtaining the best results. The Double-Mass Curve theory consists of a graph of the accumulation of one quantity and the accumulation of another quantity over the same period, which is plotted as a straight line as long as the data are proportional (*x*, *y* in mm). The consistency of the data was measured using the coefficient of determination, R^2 (see Figure 2). The slope represents the constant of proportionality between the quantities [10].

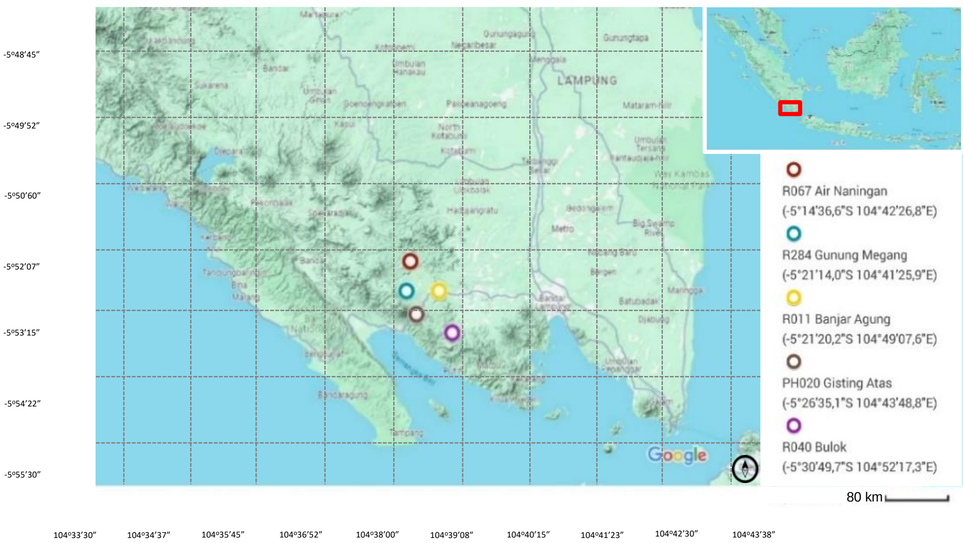

Figure 1. Geographic location of the study area

Table 2. Rainfall station Coordinates.

Stations	Coordinates		data length
	latitudes	longitudes	
R067	-5º14'36.6''	104º32'26.8''	3652 days
R284	-5º21'14.0''	104º41'25.9''	3652 days
R011	-5º21'20.2''	104º49'07.6''	3652 days
PH020	-5º26'35.1''	104º43'48.8''	3652 days
R040	-5º30'49.7''	104º52'17.3''	3652 days

Arithmetic Mean Method. This is simply the simplest method to estimate missing rainfall data in an area. The rainfall data from all surrounding stations simultaneously are added up and then divided by the number of surrounding stations. In this case, the stations used in this calculation method are located near each other. The formula used in this method is given in Eq. (1) [24].

$$Px = \frac{P1 + P2 + \cdots + Pn}{n} \quad (1)$$

Px is missing rainfall data station X; $P1$, $P2$, …, Pn is rainfall data in reference stations, and n is the total number of reference stations.

Normal Ratio Method. This method is used when the normal annual rainfall from the surrounding stations exceeds 10% of the annual rainfall in a station under consideration and vice versa. The formula used in this method is given in Eq. (2) [25].

$$Px = \frac{1}{n}\left(\frac{Nx \times P1}{N1} + \cdots + \frac{Nx \times Pn}{Nn}\right) \quad (2)$$

Px is missing rainfall data in station X; Nx is annual rainfall data in station X; $P1$, …, Pn is rainfall data in reference stations; $N1$, …, Nn is annual rainfall data in reference stations, and n is the total number of reference stations.

Inversed Square Distance Method. This method interprets that the weight of each station is assumed to be inversely proportional to the squared distance of the target station from the surrounding stations. The formula used in this method is given in Eq. (3) [26].

$$Px = \frac{\left(\frac{P1}{L1^2} + \frac{P2}{L2^2} + \cdots + \frac{Pn}{Ln^2}\right)}{\left(\frac{1}{L1^2} + \frac{1}{L2^2} + \cdots + \frac{1}{Ln^2}\right)} \quad (3)$$

Px is missing rainfall data from station X; $P1$, $P2$, …, Pn is rainfall data in reference stations; and $L1$, $L2$, …, Ln is a distance of reference stations from Station X.

Modified Method. This method is a modification form of three methods commonly known such as AM, NR, and ISD. Modifications are made by adding another factor to the calculation, that is, elevation of the rain gauge station [27]. The formulas used in this method are given in Eq. (4), Eq. (5), and Eq. (6).

$$Px = \frac{(P1 \times H1) + \cdots + (Pn \times Hn)}{\Sigma H} \quad (4)$$

$$Px = \frac{\left(\frac{Nx \times H1 \times P1}{N1} + \frac{Nx \times Hn \times Pn}{Nn}\right)}{\Sigma H} \quad (5)$$

$$Px = \frac{n}{\Sigma H}\left(\frac{\frac{P1 \times H1}{L1^2} + \cdots + \frac{Pn \times Hn}{Ln^2}}{\frac{1}{L1^2} + \cdots + \frac{1}{Ln^2}}\right) \quad (6)$$

$H1$, …, Hn is the elevation of reference stations and ∑H is a summation of the elevation of any of reference stations.

Pearson's Correlation Coefficient. The Pearson's Correlation Coefficient, denoted by the letter R, approaches 1 when there is a strong correlation between observed and predicted values, and approaches 0 when the correlation is weak. The formula used in this method is given in Eq. (7) [28].

$$R = \frac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2} \times \sqrt{n\Sigma y^2 - (\Sigma y)^2}} \quad (7)$$

With x is observed rainfall data, and y is predicted rainfall data. The classification of R, as explained in [28], is shown in Table 3.

Table 3. Conventional approach to interpreting a correlation coefficient

Correlation Coefficient (R)	Interpretation
0.00 – 0.10	Negligible
0.10 – 0.40	Weak
0.40 – 0.50	Moderate
0.50 – 0.80	Strong
0.80 – 1.00	Very Strong

Root Mean Square Error. In statistics, Root Mean Square Error (RMSE) is a function that measures the root mean squared differences between the model's predicted values x and the actual values y of the observed data. The lower RMSE value indicated a better calculation data. The formula used in this method is given in Eq. (8) [29].

$$RMSE = \sqrt{\frac{\sum[(x - y)^2]}{n}} \quad (8)$$

With x is observed rainfall data; y is predicted rainfall data, and n is the amount of data used.

3. Result and Discussion

Cumulative Rainfall Consistency Test. Before estimating missing rainfall data, it is required to verify the data consistency. This study uses the DMC method to check the consistency of the rainfall data used. The R^2 of this data consistency is shown in Figure 2. As shown below, the consistency of rainfall data used in this research is pretty varied. It shows the R^2 of each station, such as R067 Air Naningan with a value of 0.9974; R284 Gunung Megang with a value of 0.9980; R011 Banjar Agung with a value of 0.9934; PH020 Gisting Atas with a value of 0.9986; and R040.

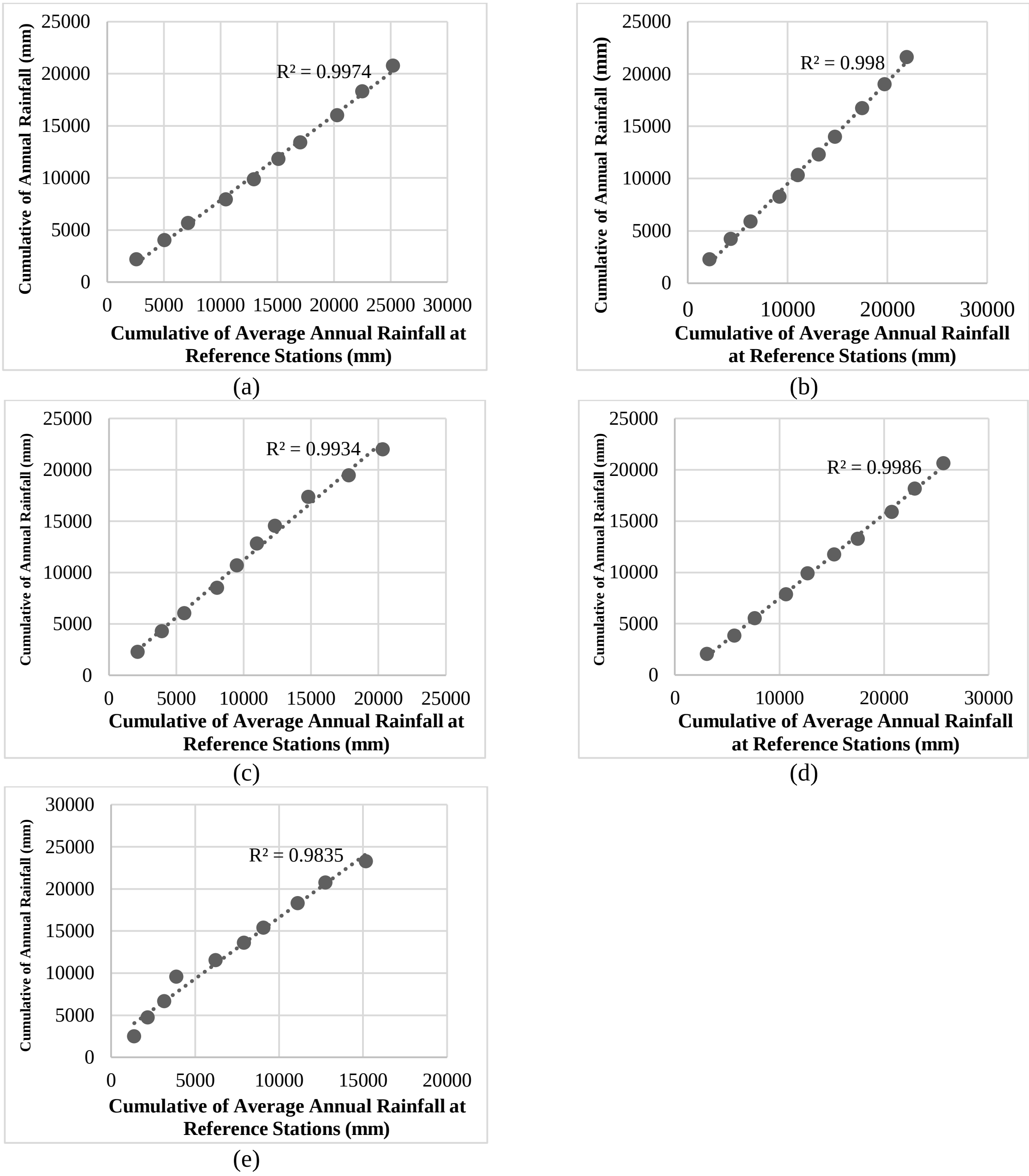

Figure 2. Double-Mass Curve at: a) R067 Air Naningan, b) R284 Gunung Megang, c) R011 Banjar Agung, d) PH020 Gisting Atas, dan e) R040 Bulok.

Bulok with a value of 0.9835. As mentioned before, the correlation coefficient in Table 3 implies that the consistency of rainfall data used in this research is very strong, therefore, the rainfall data are good to use.

In addition, although data consistency at the five stations used in this research is firm, there is no specific research regarding the feasibility of rainfall data consistency. As a consequence, although these rainfall data are in good shape according to Schober, et.al [28], it does not indicate that the data consistency of this study will not affect the Pearson's Correlation Coefficient. In other words, there is still a chance of acquiring a small correlation coefficient at some stations with smaller data consistency.

Comparison of Pearson's Correlation. Following the data consistency test, it can be analyzed to estimate missing rainfall data at each station. These methods were applied with three various reference stations, such as four reference stations, three reference stations, and two reference stations. All the results obtained are being analyzed to get the Pearson's Correlation Coefficient and Mean Square Error value.

Based on the analysis results to estimate missing rainfall data using all of the methods, the Pearson's

Correlation Coefficient and Mean Square Error value are obtained as presented in Table 4 and Table 5. The results indicate the correlation coefficient and RMSE pattern of the data estimation. This shows, as previously mentioned, data consistency contributes to the correlation coefficient of each station.

The Pearson's Correlation Coefficient between the model's predicted data and the actual observation at each station using AM, NR, ISD, and M is generally moderate, with an average value above 0.4. However, certain stations with lower data consistency, such as R011 and R040, the average coefficient value drops to below 0.4. It suggests that data consistency may influence the Pearson's Correlation Coefficient

Table 4. Comparison of Pearson's Correlation Coefficient at each station.

Stations	Methods	4 Reference Stations	3 Reference Stations	2 Reference Stations
R067	AM	0.4496	0.4223	0.3792
	NR	0.4305	0.4215	0.3681
	ISD	0.4718	0.4388	0.3870
	MAM	0.4582	0.4317	0.3859
	MNR	0.4443	0.4189	0.3784
	MISD	**0.4722**	**0.4451**	**0.3942**
R284	AM	0.5344	0.5000	0.4455
	NR	0.5101	0.4796	0.4323
	ISD	0.5514	0.5119	0.4466
	MAM	0.5492	**0.5126**	**0.4526**
	MNR	0.5325	0.4990	0.4454
	MISD	**0.5530**	0.5120	0.4497
R011	AM	0.2873	0.2732	0.2512
	NR	0.2872	0.2730	0.2513
	ISD	0.2843	0.2711	0.2500
	MAM	0.2861	0.2721	0.2501
	MNR	**0.2877**	**0.2735**	**0.2515**
	MISD	0.2828	0.2696	0.2485
PH020	AM	0.5108	0.4780	0.4292
	NR	0.4990	0.4689	0.4209
	ISD	0.5161	0.4768	0.4222
	MAM	0.5179	0.4858	0.4330
	MNR	0.5131	0.4806	0.4302
	MISD	**0.5214**	**0.4887**	**0.4350**
R040	AM	0.3427	0.3251	0.2955
	NR	0.3446	**0.3263**	**0.2971**
	ISD	**0.3490**	0.3261	0.2916
	MAM	0.3292	0.3133	0.2868
	MNR	0.3315	0.3159	0.2892
	MISD	0.3381	0.3222	0.2951

Table 5. Comparison of Root Mean Square Error (RMSE) at each station (in mm)

Stations	Methods	4 Reference Stations	3 Reference Stations	2 Reference Stations
R067	AM	12.64	13.04	**13.80**
	NR	13.40	14.00	15.15
	ISD	**12.51**	**12.98**	13.84
	MAM	12.67	13.11	13.97
	MNR	13.40	14.05	15.34
	MISD	12.93	13.27	14.07
R284	AM	10.94	**11.44**	**12.37**
	NR	11.30	11.86	12.89
	ISD	11.01	11.60	12.71
	MAM	**10.94**	11.52	12.60
	MNR	11.14	11.76	12.98
	MISD	11.62	12.36	13.41
R011	AM	11.97	12.41	13.23
	NR	11.72	12.15	12.98
	ISD	12.20	12.59	13.34
	MAM	12.08	12.49	13.27
	MNR	**11.61**	**12.01**	**12.79**
	MISD	12.42	12.81	13.60
PH020	AM	12.85	13.23	**13.96**
	NR	13.30	13.94	15.15
	ISD	12.83	13.29	14.08
	MAM	**12.78**	**13.22**	14.08
	MNR	13.21	13.91	15.29
	MISD	12.83	13.29	14.02
R040	AM	13.64	13.99	14.66
	NR	**12.50**	**12.66**	13.15
	ISD	13.65	14.12	14.88
	MAM	14.08	14.43	15.09
	MNR	12.61	12.76	**13.07**
	MISD	13.77	14.14	14.79

Table 4 also indicates the correlation pattern between each number of reference stations used. For example, Pearson's Correlation Coefficient at R067 Air Naningan using the Arithmetic Mean Method with four reference stations is 0.4496, which means it is slightly better than using three reference stations with a value of 0.4223 and two reference stations with a value of 0.3792. For instance, at station R011, the Pearson's Correlation Coefficient obtained using the Arithmetic Mean Method was 0.2873 with four reference stations, 0.2732 with three reference stations, and 0.2512 with two reference stations. These results suggest that using more reference stations to estimate missing rainfall data in an area generally leads to a higher Pearson's Correlation Coefficient, whereas fewer reference stations result in lower values.

Table 4 presented that the Modified Inverse Square Distance is better than the Modified Normal Ratio (MNR) and Modified Arithmetic Mean Method (MAM). But all of the modified methods are better than non modified methods (AM, NR, and ISD). The same results were also shown by Gultom, et.al [27], who conducted research using rainfall data from West Lampung Regency. However, the research only used Pearson's Correlation Coefficient to fit the rainfall data with the predicted rainfall data [27]. In this research, we use Pearson Correlation and RMSE, but in reference number 27 only using a Pearson Correlation coefficient. In this research also, we study a power of 2 for the distance of reference station (L). I guess, this is also another novelty of this research.

Comparison of RMSE. The results in Table 5 imply that the Root Mean Square Error (RMSE) tends to be smaller or better when the values are estimated using the Modified Arithmetic Mean Method. The result contrasts with the Pearson's Correlation Coefficient, where the Modified Inversed Square Distance Method proved to be more effective in estimating missing rainfall data. The difference between these two measures is largely a result of their different concepts: Pearson's Correlation Coefficient measures the correlation pattern between predicted and observed data while RMSE measures the difference of each value. It is possible for certain data can have both the highest Pearson's Correlation Coefficient and

Root Mean Square Error at the same time. The other difference is that we used a power of 2 for the distance of the reference station (L).

The Modification Method used in this research has better value than the unmodified method. Although it does not have a strong Pearson's Correlation Coefficient according to Schober, et. al. [28], it remains the largest one compared to the others. It indicates that the elevation of stations used has become one of the factors contributing to the data correlation. The other factor contributing highly to the results is the consistency of the data used.

Comparison of daily rainfall data. By using four reference stations, predicted or missing rainfall data series using the Modified Inversed Square Distance Method and observed rainfall data series for a year (2022) can be presented as follows.

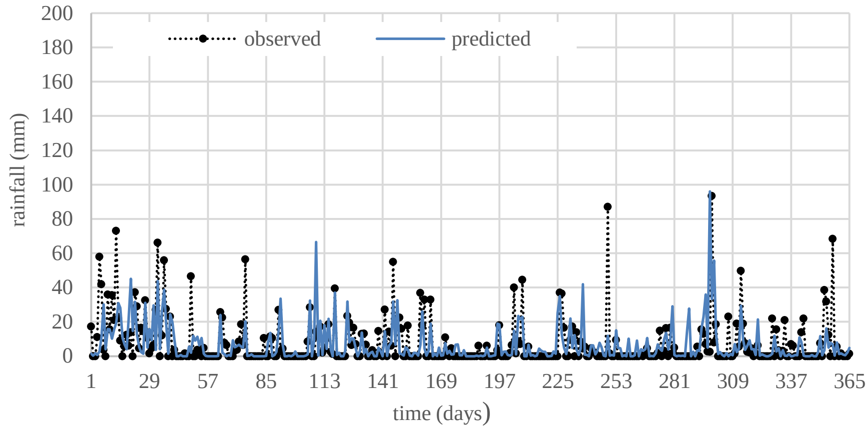

Figure 3. Predicted and observed daily rainfall series from Air Naningan station (year 2022).

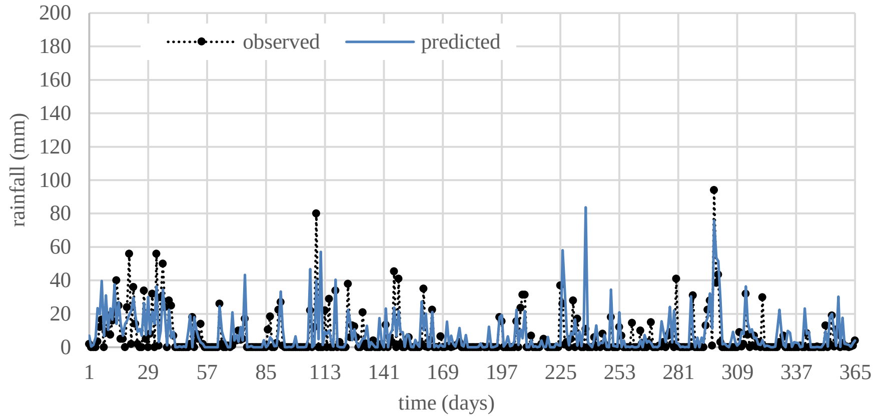

Figure 4. Predicted and observed daily rainfall series from Gunung Megang station (year 2022).

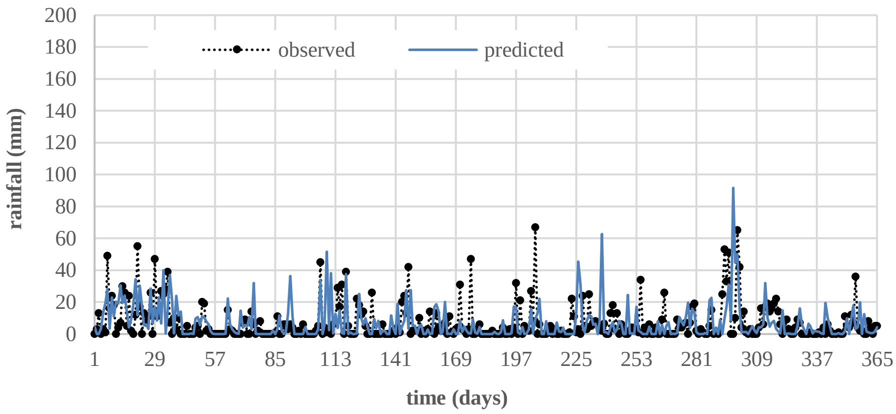

Figure 5. Predicted and observed daily rainfall series from Banjar Agung station (year 2022).

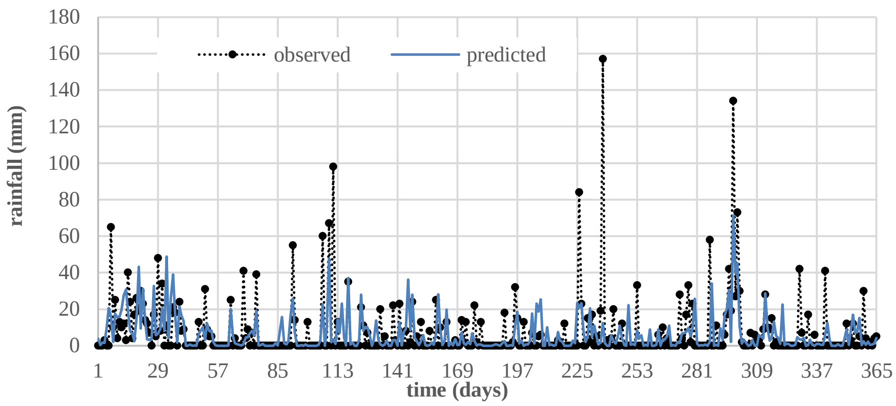

Figure 6. Predicted and observed daily rainfall series from Gisting Atas station (year 2022).

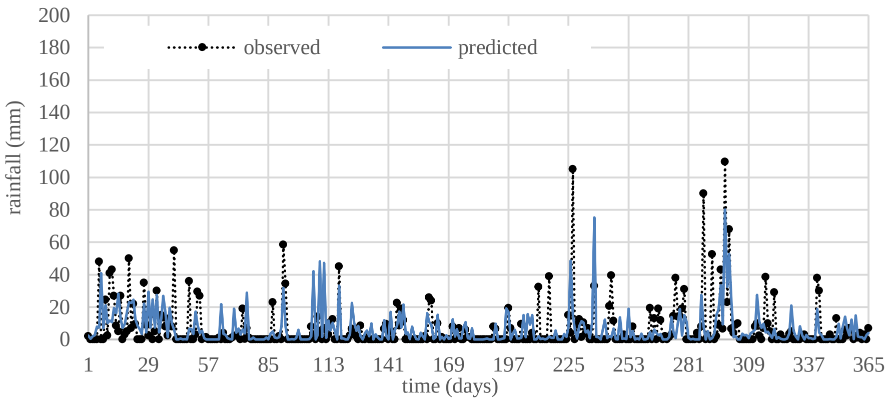

Figure 7. Predicted and observed daily rainfall series from Bulok station (year 2022).

Figure 3, 4, 5, 6, and 7 present observed daily rainfall series and predicted daily rainfall series generated by using the Modified Inversed Squared Distance Method for Air Naningan, Gunung Megang, Banjar Agung, Gisting Atas, and Bulok stations.

Pearson's Correlation Coefficient *R* for observed and predicted rainfall series at each station is 0.4409, 0.6081, 0.2908, 0.4908, and 0.4686. Root Mean Square Error (RMSE) for observed and predicted rainfall series in each station is 13.4899 mm, 10.8770 mm, 13.2582 mm, 14.9080 mm, and 13.1061 mm. This result presented that in the year 2022, observed and predicted rainfall data series from Gunung Megang station had the best results compared to the other stations for Pearson's Correlation Coefficient and RMSE values. Pearson's correlation coefficient value is 0.6081, and RMSE value is 10.8770 mm.

Optimise Power of 2. Inverse Square Distance (ISD) and Modified Inverse Square Distance (MISD) were used to produce Table 4 and Table 5 by using Equation (3) and Equation (6). In the equations, the power of 2 of a distance of reference station (*L*) was applied to predict missing rainfall data. Suppose it is assumed that the 2 is x variable. Using the four reference stations, the *x* was optimized to produce a better correlation coefficient and RMSE values. Using the solver tool in Excel, Table 6 and Table 7 were generated as follows.

Table 6. Rainfall correlation for different power.

Stations	Methods	Power	*R*	%
R067	ISD	2.00	0.4718	0.190
		1.60	0.4727	
	MISD	2.00	0.4722	0.740
		1.28	0.4757	
R284	ISD	2.00	0.5514	0.020
		1.79	0.5515	
	MISD	2.00	0.5530	0.540
		0.95	0.5560	
R011	ISD	2.00	0.2843	1.060
		0.00	0.2873	
	MISD	2.00	0.2828	1.150
		0.00	0.2861	
PH020	ISD	2.00	0.5161	0.330
		1.27	0.5178	
	MISD	2.00	0.5214	1.090
		1.05	0.5271	
R040	ISD	2.00	0.3490	0.490
		1.27	0.3507	
	MISD	2.00	0.3381	0.380
		1.40	0.3394	
			Average	**0.6**

Table 7. Rainfall RMSE for different power.

Stations	Methods	Power	RMSE	%
R067	ISD	2.00	12.5076	0.439
		1.60	12.4528	
	MISD	2.00	12.9294	2.485
		1.28	12.6121	
R284	ISD	2.00	11.0055	0.306
		1.79	10.9719	
	MISD	2.00	11.6233	4.117
		0.95	11.1544	
R011	ISD	2.00	12.2004	1.875
		0.00	11.9738	
	MISD	2.00	12.4210	2.816
		0.00	12.0760	
PH020	ISD	2.00	12.8304	0.271
		1.27	12.7957	
	MISD	2.00	12.8328	2.816
		1.05	12.0760	
R040	ISD	2.00	13.6492	0.500
		1.27	13.5812	
	MISD	2.00	13.7670	0.003
		1.40	13.7666	
			Average	**1.6**

Table 6 and 7 indicate that to improve Pearson's correlation coefficient (*R*) and RMSE values, we can use power of less than 2. The Tables present that using a power of less than 2, we can improve Pearson's correlation coefficient and RMSE by about 0.6% and 1.6% better than using a power of 2.

4. Conclusion

In addition to the four methods used such as AM, NR, ISD, and M, this research also used three approaches as well such as using four, three, and two reference stations to obtain better results. Based on the results, it can be concluded that each station has its own best method. However, the best method to estimate missing rainfall data in Tanggamus Regency as indicated by Pearson's Coefficient Correlation is the Modified Inversed Square Distance Method while Root Mean Square Error indicates that the Modified Arithmetic Mean Method is the best method. It suggests that the Modification Method used in this research tends to be more effective than the other three common unmodified methods. Furthermore, Pearson's Coefficient Correlation proved to be more reliable than the Root Mean Square Error in evaluating model fitting. The last calculation presented that for Inversed Square Distance and Modified Inverse Square Distance methods, we still can improve the results if using the power of less than 2 for the distance of reference station (*L*).

Suggestion

This research uses rainfall data from five stations. Estimating missing rainfall data in an area requires rainfall data from the nearest surrounding stations; the more reference stations used, the better the data obtained. In addition, reliable estimation also depends on the quality of the data, particularly its consistency.

Acknowledgment

We acknowledge Mesuji Sekampung River Basin Development Agency (BBWS-MS) for providing complete rainfall data used in this research.

References

[1] D. Gunawan, “Perbandingan Curah Hujan Bulanan Dari Data Pengamatan Permukaan, Satelit Trmm Dan Model Permukaan Noah,” *J. Meteorol. dan Geofis.*, vol. 9, no. 1, pp. 1–10, 2014, doi: 10.31172/jmg.v9i1.17.

[2] R. Alfian, A. Zakaria, E. P. Wahono, D. I. Kusumastuti, and A. Herison, “The Ratio of BMKG and TRMM Rainfall Data in West Java Province by Using Statistical Parameter and Correlation Analysis,” *J. Eng. Sci. Res.*, vol. 3, no. 2, 2021, doi: 10.23960/jesr.v3i2.82.

[3] A. I. Hadi and S. Dan Herliana, “Analisis Karakteristik Intensitas Curah Hujan di Kota Bengkulu,” *J. Fis. FLUX*, vol. 7, no. 2, pp. 119–129, 2010.

[4] C. D. Soemarto, *Hidrologi Teknik*. Surabaya: Usaha Nasional, 1985.

[5] B. Setiyawan, A. Zakaria, and E. P. Wahono, “Correlation Between the NASA ’ s Tropical Rainfall Measuring Mission Data and Indonesia Meteorology and Geophysics Agency of Kalimantan Island,” *Int. J. Aviat. Sci. Eng.*, vol. 4, no. 1, pp. 52–58, 2022, doi: 10.47355/AVIA.V4I1.68.

[6] L. K. Aridini, A. Zakaria, D. I. Kusumastuti, A. Herison, and E. P. Wahono, “Time Lag Application on Correlation Analysis among the national and international agencies for the rainfall data of Lampung Province, Indonesia,” *J. Innov. Technol.*, vol. 3, no. 2, pp. 45–51, 2022, doi: 10.31629/jit.v3i2.5060.

[7] N. Yang *et al.*, “Evaluation of the TRMM Multisatellite Precipitation Analysis and Its Applicability in Supporting Reservoir Operation and Water Resources Management in Hanjiang Basin, China,” *J. Hydrol.*, vol. 549, pp. 313–325, 2017, doi: 10.1016/j.jhydrol.2017.04.006

[8] L. Suryaningtyas, S. Ery, and Rispiningtati, “Hydrological Analysis of TRMM (Tropical Rainfall Measuring Mission) Data in Lesti Sub Watershed),” *Civ. Environ. Sci. J.*, vol. III, no. 01, pp. 18–30, 2020, doi: 10.21776/ub.civense.2020.00301.3.

[9] R. P. De Silva and N. D. K. Dayawansa, “A Comparison of Methods Used in Estimating

Missing Rainfall Data," *J. Agric. Sci. – Sri Lanka*, vol. 3, p. 2, 2007, doi: 10.4038/jas.v3i2.8107.

[10] J. K. Searcy and C. H. Hardison, *Double-Mass Curves*. Washington DC: U.S. Government Printing Office, 1960.

[11] N. Sriwongsitanon *et al.*, "The Combined Power of Double Mass Curves and Bias Correction for the Maximisation of the Accuracy of an Ensemble Satellite-Based Precipitation Estimate Product," *Hydrology*, vol. 10, no. 7, 2023, doi: 10.3390/hydrology10070154.

[12] S. M. Hindall, "Temporal Trends in Fluvial-Sediment Discharge in Ohio, 1950-1987," *J. Soil Water Conserv.*, vol. 46, no. 4, pp. 311–313, 1991.

[13] P. Gao, V. Geissen, C. J. Ritsema, X.-M. Mu, and F. Wang, "Impact of Climate Change and Anthropogenic Activities on Stream Flow and Sediment Discharge in the Wei River Basin, China," *Hydrol. Earth Syst. Sci.*, vol. 17, no. 3, pp. 961–972, 2013, doi: 10.5194/hess-17-961-2013.

[14] W. Gumindoga, T. H. M. Rientjes, A. T. Haile, H. Makurira, and P. Reggiani, "Bias Correction Schemes for CMORPH Satellite Rainfall Estimates in the Zambezi River Basin," *Hydrol. Earth Syst. Sci. Discuss.*, vol. 2016, pp. 1–36, 2016, doi: 10.5194/hess-2016-33.

[15] A. Elgamal, P. Reggiani, and A. Jonoski, "Impact Analysis of Satellite Rainfall Products on Flow Simulations in the Magdalena River Basin, Colombia," *J. Hydrol. Reg. Stud.*, vol. 9, pp. 85–103, 2017, doi: 10.1016/j.ejrh.2016.09.001.

[16] N. D. Bennett, L. Newham, B. Croke, and A. J. Jakeman, "Patching and Disaccumulation of Rainfall Data for Hydrological Modelling," in *MODSIM07 - Land, Water and Environmental Management: Integrated Systems for Sustainability, Proceedings*, 2007.

[17] J. Das Agrawal, "ANN in Forecasting Missing Rainfall Data," *E3S Web Conf.*, vol. 04017, pp. 1–10, 2023, doi: 10.1051/e3sconf/202340504017.

[18] A. M. Armanuos, N. Al-Ansari, and Z. M. Yaseen, "Cross Assessment of Twenty-One Different Methods for Missing Precipitation Data Estimation," *Atmosphere (Basel).*, vol. 11, no. 4, 2020, doi: 10.3390/atmos11040389.

[19] C. Yozgatligil, S. Aslan, C. Iyigun, and I. Batmaz, "Comparison of Missing Value Imputation Methods in Time Series: The Case of Turkish Meteorological Data," *Theor. Appl. Climatol.*, vol. 112, no. 1, pp. 143–167, 2013, doi: 10.1007/s00704-012-0723-x.

[20] H. P. G. . Caldera, V. Piyathisse, and K. D. W. Nandalal, "A Comparison of Methods of Estimating Missing Daily Rainfall Data," *Eng. J. Inst. Eng. Sri Lanka*, vol. 49, p. 1, Nov. 2016, doi: 10.4038/engineer.v49i4.7232.

[21] E. Yogafanny and D. J. Legono, "A Comparative Study of Missing Rainfall Data Analysis Using the Methods of Inversed Square Distance and Arithmetic Mean," *ASEAN Eng. J.*, vol. 12, no. 2, pp. 69–74, 2022, doi: 10.11113/aej.V12.16974.

[22] N. F. Ahmad Radi, R. Zakaria, S. Satari, and M. A. Z. Azman, "Spatial Dependence of Extreme Rainfall," in *AIP Conference Proceedings*, 2017, p. 20003. doi: 10.1063/1.4982833.

[23] A. A. G. Nadiatul Adilah and H. Hannani, "Comparison of Methods to Estimate Missing Rainfall Data for Short Term Period at UMP Gambang," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 682, p. 12027, 2021, doi: 10.1088/1755-1315/682/1/012027.

[24] V. Te Chow, D. R. Maidment, and L. W. Mays, *Applied Hydrology*. New York SE - xiii, 572 pages : illustrations ; 26 cm.: McGraw-Hill New York, 1988.

[25] V. P. Singh, *Elementary Hydrology*. Englewood Cliffs, NJ: Prentice Hall Englewood Cliffs, NJ, 1992.

[26] F. Chen, F. Chen, and C. Liu, "Estimation of the Spatial Rainfall Distribution Using Inverse Distance Weighting (IDW) in the Middle of Taiwan," *J. Int. Soc. Paddy Water Environ. Eng.*, vol. 10, no. 2, pp. 209–222, 2012, doi: 10.1007/s10333-012-0319-1.

[27] W. M. R. Gultom, A. Zakaria, and D. I. Kusumastuti, "Perbandingan Analisis Data Curah Hujan yang Hilang di Daerah Lampung Barat dengan Menggunakan Metode Normal Ratio, Rata-Rata Aljabar, Inversed Square Distance, dan Modified Method," *J. Rekayasa Sipil dan Desain*, vol. 0, no. 3, pp. 503–512, 2021.

[28] P. Schober, C. Boer, and L. Schwarte, "Correlation Coefficients: Appropriate Use and Interpretation," *Anesth. Analg.*, vol. 126, no. 5, pp. 2763–2768, 2018, doi: 10.1213/ANE.0000000000002864.

[29] P. J. Bickel and K. A. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics*, 2nd ed. New York: Chapman and Hall/CRC, 2015. doi: 10.1201/b19822