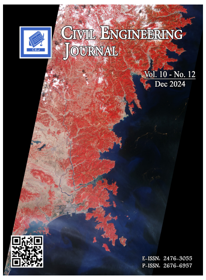
Available online at www.CivileJournal.org

Civil Engineering Journal

(E-ISSN: 2476-3055; ISSN: 2676-6957)

Vol. 10, No. 12, December, 2024



Evaluation of GPM IMERG Product Against Ground Station Rainfall Data in Semi-Arid Region

Denik Sri Krisnayanti [1*], Jusuf J. S. Pah [1], Ralno R. Klau [1], Alex Decaprio [1], Syamsumarlin [1]

[1] *Faculty of Science and Engineering, Nusa Cendana University, Kupang 85228, Indonesia.*

Received 18 June 2024; Revised 24 October 2024; Accepted 07 November 2024; Published 01 December 2024

Abstract

Benanain River is the longest and largest river on Timor Island, with a length of 132 km and an area of 6,460.12 km². In this region, a significant factor affecting the presence of surface water sources is rainfall. To compensate for the lack or unavailability of automatic Rainfall Data (RD) in the Benanain River Basin (BRB), Global Rainfall Measurement (GPM) data from 1998 to 2018 (20 years) were used. The accuracy of GPM rainfall analysis was obtained when parameter conformity and compatibility with data recorded at Rainfall Station (RS) were maintained. The difficulty of predicting rainfall values, spatially and temporally, in the field led to data gaps and unreliable data for analysis needs. Additionally, RD obtained from observation stations contributed to measuring rainfall because there was insufficient RD for analysis in a few regions. The challenge of accurately predicting rainfall values in the field led to differences in data, rendering it unreliable for analysis. To address this issue, satellite data was required as an alternative method to estimate rainfall. Among a total of 7 RS, only 2 passed rainfall characteristic tests. Following this discussion, Lahurus station showed a correlation coefficient of 0.7046, an RMSE of 25.89, and an NSE of 0.476. In addition, the rainfall characteristic test result for Haliwen Station was 1.66 (R_{100}/R_2). The second station that passed was Kaubele Station, signifying a correlation coefficient of 0.7907, RMSE of 25.28, and NSE of 0.604. Additionally, the rainfall characteristic test result for Haliwen Station was 3.04 (R_{100}/R_2) and the daily performance of the GPM product in the rainy season with low rainfall ($\leq$ 50 mm) was better compared to extreme rainfall ($\geq$ 100 mm). In this study, corrected GPM daily RD in the range >100 mm was underestimated. This analysis implied that the GPM IMERG Final Run product on daily and monthly rainfall timescales had strong detection capabilities and provided data support for long-time series investigations on Timor Island.

Keywords: Rainfall Characteristic; Benanain; Timor; Precipitation.

1. Introduction

A rainfall observation station is used to observe and record rainfall, manually or automatically. Data monitored by rainfall station (RS) includes daily, monthly, and annual rainfall data (RD). This data is useful for water resource planning, which affects agricultural and economic activities. RD is very important in the hydrological analysis process. Temporal and spatial variability in data is the primary source of many complex hydrological events. Therefore, accurate temporal-spatial and long-term RD is essential for climate change study/prediction, simulation studies, hydrological predictions, natural disaster studies such as floods, landslides, and droughts, as well as water resource management [1]. A major problem in this study is the lack of sufficient RD for analysis in a few regions, specifically daily data over a long period or at a specific time. Additionally, the difficulty in predicting rainfall values spatially and temporally in the field is caused by the lack of good direct rainfall measurements and the highly variable nature of parameters [2].

* Corresponding author: denik.krisnayanti@staf.undana.ac.id

http://dx.doi.org/10.28991/CEJ-2024-010-12-09

The lack of sufficient RD in unsurveyed river basins certainly poses a challenge for hydrological analysis. Due to the scarcity or unavailability of RD, several studies in recent years have explored the use of satellite-based RD as a complement to manually measured RD. Moreover, the significant advancement of satellite remote sensing technology and onboard equipment has led to the development of multi-satellite rainfall products. This product eliminates the limitations and drawbacks of ground-based rain instruments as well as radar systems and has become the primary technology for obtaining high-resolution RD at both spatial and temporal scales [3]. Following this discussion, IMERG (Integrated Multi-Satellite Retrievals) represents the latest generation of multi-satellite integrated RD designed for the Global Precipitation Measurement (GPM) mission [4]. Satellite-derived RD from GPM can be used in Indonesia because the method provides information about extreme rainfall in tropical regions [5, 6]. In water resources management, GPM IMERG is widely adopted across the world as a substitute for predicting ground station RD [7, 8]. Additionally, the mission of GPM is [9] a next-generation global rainfall product with half-hourly temporal and spatial resolutions of 0.1° initiated in 1992. This information is derived from the success of TRMM (Tropical Rainfall Measuring Mission) [2, 10].

The need for RD in hydrological analysis can be supplemented by using GPM IMERG data when automatic rainfall recording (ARR) is limited. The three categories of GPM IMERG products include IMERG-E (GPM IMERG Early Run), IMERG-L (GPM IMERG Late Run), as well as IMERG-F (GPM IMERG Final Run). This product has a grid resolution of 0.1º × 0.1º and 30 min temporal resolution. IMERG-E refers to products that deliver near real-time data with a latency of approximately 4 hours. In addition, IMERG-L is another product that offers near real-time data with a latency of about 14 hours. This GPM product is ideal for monitoring natural disasters such as floods and hurricanes. A study in southwestern Iran found that GPM IMERG-F performed well in capturing extreme rainfall events. However, IMERG-E and IMERG-L had issues with rainfall estimation in mountainous regions, indicating overestimated values [11]. An investigation into extreme rainfall in the North China Plain from 2000 to 2018 was conducted to examine the accuracy of IMERG-F and IMERG-L. Relating to the exploration, the findings showed that IMERG-F had better accuracy than IMERG-L [12]. As a result, IMERG-F was recommended for study due to its calibration [13].

The advantages of satellites in measuring rainfall intensity compared to surface rain observation stations generally include high spatial and temporal resolution, wide region coverage, near real-time and continuous recording of data, quick access, less impact from climate and terrain variability, as well as free economic benefits [14]. Many studies have evaluated the efficiency of IMERG to capture light rainfall. Additionally, the performance of IMERG is better than TRMM, specifically in regions with relatively dry climates as well as middle and high latitudes. This method is also found to have better performance in detecting light rainfall [15]. There is also an opinion showing that IMERG has limited capability in estimating light rainfall, such as in Mainland China. Following the discussion, IMERG-L shows the best performance among other rainfall products (IMERG-E, PERSIANN-CCS, GSMaP-NRT, GSMaP-MVK, and TMPA-RT). However, GSMaP and IMERG overestimate the proportions of light rainfall events. Both also have relatively larger errors in light rainfall (0.2-0.4 mm/h or 1-2 mm/day) [16]. This same issue occurs in the high tropical Andes region at an elevation range from 3,800 to 4,600 m, where IMERG is significantly less effective at detecting low-intensity rainfall events [17].

A study was conducted in the river basins of Northern Morocco using four satellite-based rainfall products (TRMM, GPM, PERISIANN-CDR, and CHIRPS). These four satellite-based rainfall products (SPPs) estimated rainfall better in humid and semi-humid regions than in semi-arid ones. However, TRMM and Climate Hazards Group InfraRed Rainfall with Station data (CHIRPS) outperformed the GPM product for semi-arid regions. SPP tends to overestimate rainfall, except in humid climates where TRMM and CHIRPS underestimate. Moreover, lower elevations (<500 m) provided more accurate results for SPP compared to higher altitudes (>500 m) [18].

A study in the semi-arid watershed of Morocco (Rheraya) during 2013-2018 was conducted using ERA5 reanalysis and GPM IMERG V06 data (GPM-E, GPM-L, and GPM-F). The results indicated that the four products often overestimated the observed rainfall. In addition, evaluation of IMERG showed satisfactory performance of the IMERG-E and IMER-L products on an hourly time scale, as the most efficient for simulating flood events [19]. This study indicated that the different rainfall products tested had satisfactory performance for flood hydrological modeling. Satellite-based rainfall could be an alternative for flood simulation in river basins that were not measured or not well surveyed.

Some studies often use traditional methods to evaluate satellite-derived RD, including comparison with ground-based observations and considering environmental factors such as topography and rainfall intensity [20]. The total bias is calculated by subtracting ground observations from satellite rainfall estimates, without considering the interdependence between various indicators [21]. Therefore, the use of traditional evaluation methods that compare satellite-based RD with ground station observations will be investigated in the semi-arid region of Timor Island, Indonesia. The limitations of existing RD and the mountainous topography, such as in the Benanain River Basin (BRB) on Timor Island, are among the reasons for using GPM IMERG (Figure 1).

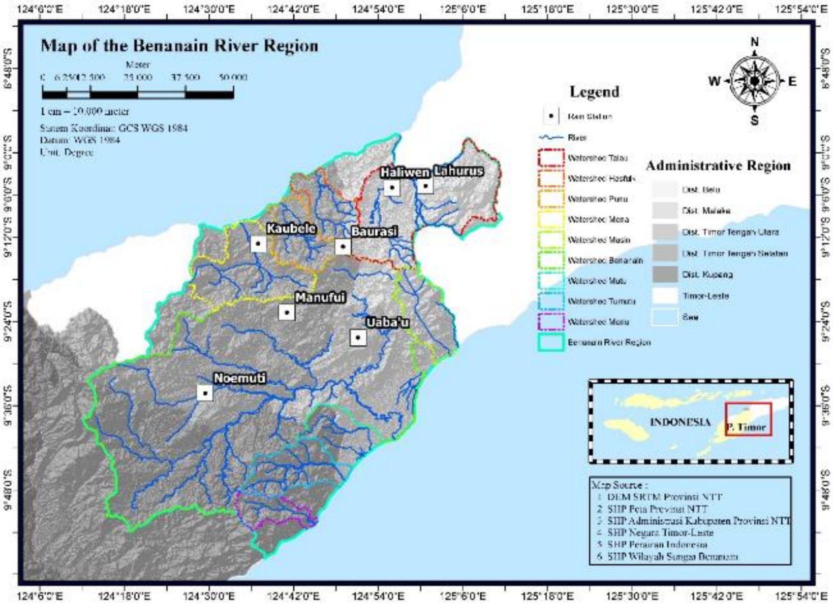


Figure 1. BRB in Timor Island

The use and evaluation of satellite remote sensing data have also been conducted on ground station RD in BRB [22]. Following this discussion, the study compares RD from ground stations with data from TRMM. The results include root mean square error (RMSE) and correlation coefficient (r) values for rainfall alone, which range from 0.1143 to 0.8760. The error values based on annual maximum daily rainfall have a range of 0.0610 - 1.3647, while values based on daily rainfall have a range of 0.0029 - 0.0463. No study has been conducted in the Benanain River region using GPM IMERG product data. Therefore, the exploration will be analyzed and reprocessed with consistent correction procedures over a very long period. This study aims to assess the performance of the post-real-time GPM IMERG-F product at three temporal resolutions (i.e., half-hourly, daily, and monthly) in the BRB region of Indonesia. The investigation is conducted to determine correction factor values using GPM IMERG product data in the semi-arid region of Timor Island, Indonesia. In addition, product data of this method is used to determine reasonable correction factor values for applying to ground station data.

2. Methods

2.1. Study Region and Data

East Nusa Tenggara (NTT) is an archipelagic province in eastern Indonesia with an area of 47,931.54 km² and 1,192 islands, of which 624 are inhabited [23]. The province consists of 3 large islands, namely Flores, Sumba, and Timor Island. In addition, Timor Island is located in the southeastern part of Indonesia and consists of 1 city and 5 regencies, including Kupang City, North Central Timor Regency, South Central Timor Regency, Kupang Regency, Malaka Regency, and Belu Regency. Benanain River is the longest and largest river in the territory of Timor Island, Indonesia. This river basin is a large and long river in the Province of NTT, which has an area of ± 3181.521 km², a length of about ± 128 km, and has ± 92 tributaries. There are 13 large rivers and 79 small rivers where the Benanain River flows across four regencies, namely North Central Timor (TTU), South Central Timor (TTS), Belu, and Malaka. The upstream of the river is in TTU, TTS, and Belu Regencies, while the downstream is in Malaka Regency into the Timor Sea [20]. BRB is geographically located between 124°11'45.64" E - 125°07'31.22" E and 8°56'33.21" S - 9°58'34.60" S.

The study region had a limited number of ground-based RS, and the uneven distribution led to gaps in the data. This limited number of stations made the investigation more difficult to analyze rainfall patterns based on ground observations. To address this issue, the daily RD accumulated in BRB from the GPM-IMERG satellite was used for the period between 2002 and 2021. The satellite-based rainfall estimates were then compared with ground-based observations from January 2002 to December 2021.

2.2. Study Framework

Several stages were included in this study, starting with identifying the watershed region to be modeled, which was essential for retrieving GPM data. The watershed region had adequate ground station data availability (Figure 2). Moreover, RD was tabulated from ground stations in the watershed supplied by the Indonesia Agency for Meteorology, Climatology, and Geophysics (BMKG) and the Nusa Tenggara II River Basin Agency (BWS NT II). The coordinates of GPM, which are close to the coordinates of the watershed being studied, were identified. Additionally, GPM grid creation was generated using AutoCAD and ArcGIS 10.2 applications. The length of RD from the ground station should range from the year 2000 to the desired time since satellite data has been available nationally since 2000.

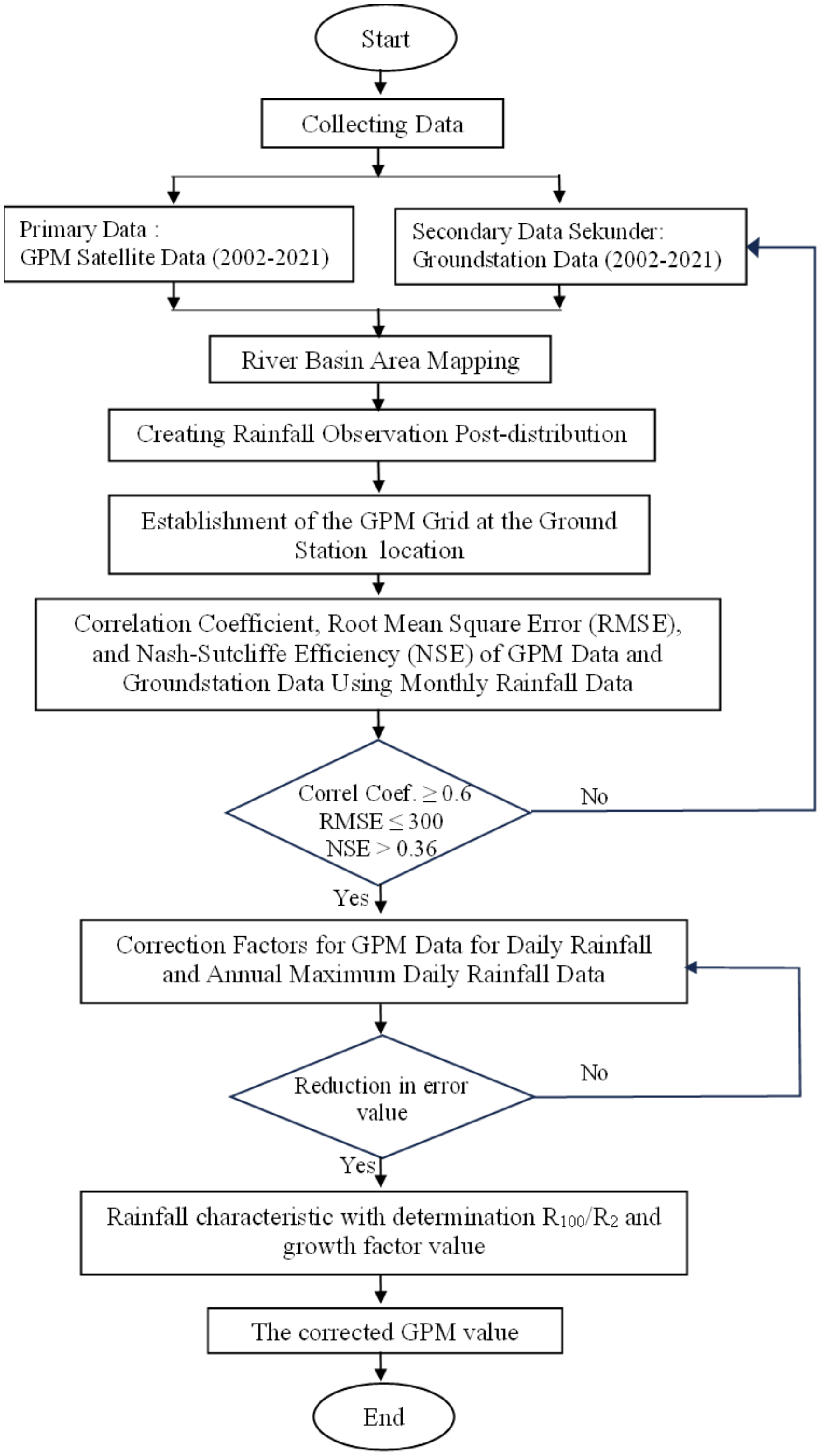


Figure 2. Flow Chart of Study

The tabulated data from the ground station was correlated with the existing GPM data of the same length of time. Ground station data was considered good when the correlation coefficient between GPM products was 0.60 or higher. Moreover, correlation was a term used to determine the strength of the relationship between variables [14, 24]. The correlation coefficient was an index or number used to determine the tightness (strong, weak, or none) of the relationship between variables [25]. In this study, the coefficient calculation of monthly GPM data and ground stations was performed. The examination of monthly RD against ground stations aimed to assess the total quality of monthly data. The best correlation between the two datasets was calculated using statistical metrics, including correlation coefficient, RMSE, and NSE (Nash Sutcliffe Error), as shown in Table 1. In the context of this study, "r" described the agreement between IMERG data and measurement data [26].

RMSE was computed to assess the error variance of the GPM-IMERG rainfall product and the ground station from the rain gauge. In the context of this study, RMSE values ranged from 0 to ∞. A value of zero showed no errors, while variables greater than zero reflected the errors between the estimated rainfall of GPM-IMERG and the measurements from the rain gauge. NSE was applied to measure the performance of the estimated rainfall by GPM-IMERG with the rain instrument observed in the catchment region of Benanain. Additionally, NSE was calculated using the empirical equation in Table 1. The value of NSE ranged between ∞ and 1.0, where the value of NSE = 1 showed a perfect match between GPM-IMERG rainfall estimates and the observed rainfall [27, 28].

Table 1. Metric of statical analysis

Name	Formula	Value	Interpretation
Correlation coefficient (r)	$\frac{\sum_1^n (O_i - \bar{O})(O_i - \bar{P})}{\sqrt{\sum_1^n (O_i - \bar{O})^2}\ \sqrt{\sum_1^n (P_i - \bar{P})^2}}$	0 – 0.19	Very weak
		0.20 – 0.39	Weak
		0.40 – 0.59	Moderate
		0.60 – 0.79	Strong
		0.81 – 1.00	Very strong
Root Mean Square Error (RMSE)	$\sqrt{\frac{1}{n}\sum_1^n (O_i - P_i)^2}$	equal to zero	No errors
Nash Sutcliffe Error (NSE)	$1 - \frac{\sum_1^n (O_i - P_i)^2}{\sum_1^n (O_i - \bar{O})}$	NSE > 0.75	Very good
		0.36 < NSE < 0.75	Good
		NSE < 0.36	Poor

The correlation coefficient ranged from -1 to +1, as strong correlations had coefficients approaching +1 or -1, while weak correlations moved toward 0. When the connection between two variables correlated to -1 or +1, it implied that the values had a perfect relationship. Consequently, when the relationship between two variables correlated to 0, the process implied that there was no connection between the values. When the correlation coefficient was positive, the variables had a direct relationship, but when the correlation coefficient was negative, the values had an inverse relationship.

NTT was included in the monsoon rainfall pattern region, and the correlation coefficient ranged from 0.60 to 0.80. When the correlation coefficient was < 0.60, then RS could not be used for flood discharge calculations at the dam [29]. To establish the accuracy of GPM IMERG satellite data compared to ground stations, GPM conducted satellite data correction. Moreover, consistency in the curve pattern or the trend of the two compared datasets served as one identification in concluding the quality of GPM data. Daily RD inspection was performed when the ground station had a minimum RD length of 3 years. Additionally, RS data to be corrected was still in the range of years covered by the available GPM data (2000-2021) [30].

The datasets (GPM and ground stations) of the same period were associated in two adjacent columns. The probability value was then calculated for specific rainfall events that were smaller in each rainfall group with an interval of 10-20 mm. As an initial guide in entering correction values, a formula was used where GPM rainfall less than a certain value (ranging from 0 - 10 mm) was considered as 0. In addition, rainfall less than a certain value (ranging from 50-100 mm) was multiplied by a moving constant between 0.8 - 1.0; and RD greater than 100 mm was multiplied by a constant between 1.0 - 1.3 [29].

The results were evaluated based on rainfall characteristics, which included comparing the rainfall ratio for a 100-year return period (RP) to a 2-year time frame and calculating the growth factor (GF). The RP ratio was the comparison between two different periods. In this study, 2-year RP was the lowest period, often approximating the average rainfall value using the IDF method, while 100-year RP produced a non-exceedance probability close to 1 and represented the maximum rainfall value. This ratio was then used in further filtering to obtain better model performance, as the R_{100}/R_2 value ranged from 1.5 to 3.4 [29]. Following the discussion, reclassification based on the R_{100}/R_2 ratio produced the GF value. This GF was multiplied by the extreme rainfall index to estimate the T-year design rainfall [31]. Through the availability of the GF value and extreme rainfall index, rainfall for a specific RP at RS was estimated.

3. Results

The initial step was to select RD from the ground station to be used, which included about 7 RS/ground stations in the region of BRB. The determination of GPM grid depended on the available ground station, and data was mapped based on the coordinates of each RS using ArcGIS 10.5 software. After identifying GPM data, a grid was created to support the coordinates of RS. Each grid covered a region of 0.1° × 0.1° or about 11.1 × 11.1 km, ensuring that every grid contained a minimum of one RS. The distribution map of GPM Grid was shown in Figure 3.

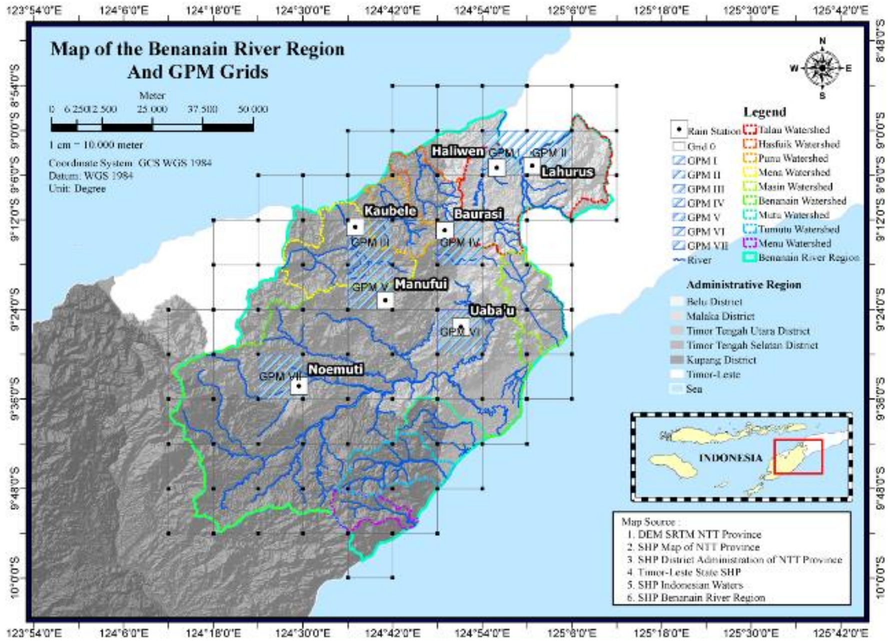


Figure 3. GPM Grid on Ground Station in BRB

Before using RD from IMERG satellite in hydrological modeling, data needed to be verified for accuracy and consistency by comparing the data with ground station data from rain instruments. In regions with ARR, GPM output was generally more accurate compared to manual rainfall records. After conducting RS and GPM grid distribution, the next step was to perform statistical analysis, correction factors, error analysis, and rainfall characteristic tests. In this study, the analysis sampling for the explanation used GPM Grid II as an example because the grid passed all the statistical tests. When any section of the statistical analysis did not meet the criteria, further calculations were stopped, and the analysis was considered unsuccessful. In addition, the method was also applied to subsequent calculations and studies. Table 2 showed the recapitulation of the distribution of RS and GPM grids for each ground station in BRB.

Table 2. The distribution of ground station in BRB

No.	Ground station	District	GPM Grids Coordinate	Number of GPM Grid
1	Haliwen	Belu	124.904, -9.063,125.004,-9.163	I
2	Lahurus	Belu	125.004,-9.063,125.104,-9.163	II
3	Kaubele	TTU	124.604,-9.163,124.704,-9.263	III
4	Baurasi	TTU	124.804,-9.163,124.904,-9.263	IV
5	Manufui	TTU	124.604,-9.363,124.704,-9.463	V
6	Uaba'u	Malaka	124.804,-9.363,124.904,-9.463	VI
7	Noemuti	TTU	124.404,-9.563,124.504,-9.663	VII

3.1. GPM Grid II (Lahurus RS)

GPM Grid II was located in the Belu District, with Lahurus RS positioned in this grid. In the context of this study, the analysis was conducted solely at Lahurus RS, as there was just one RS in the grid. The map of the GPM Grid II region was shown in Figure 4.

The initial analysis included comparing ground station and GPM data to specify the largest correlation that reflected the relationship between the two datasets applied to monthly data. To identify the strongest correlation in this investigation, a coefficient correlation equation was used. Moreover, RD in grid II was classified based on monthly rainfall, and then an analysis was conducted to obtain the correlation coefficient of RMSE and NSE values (Figure 5). Table 3 showed RD used in rainfall pattern suitability analysis of the study.

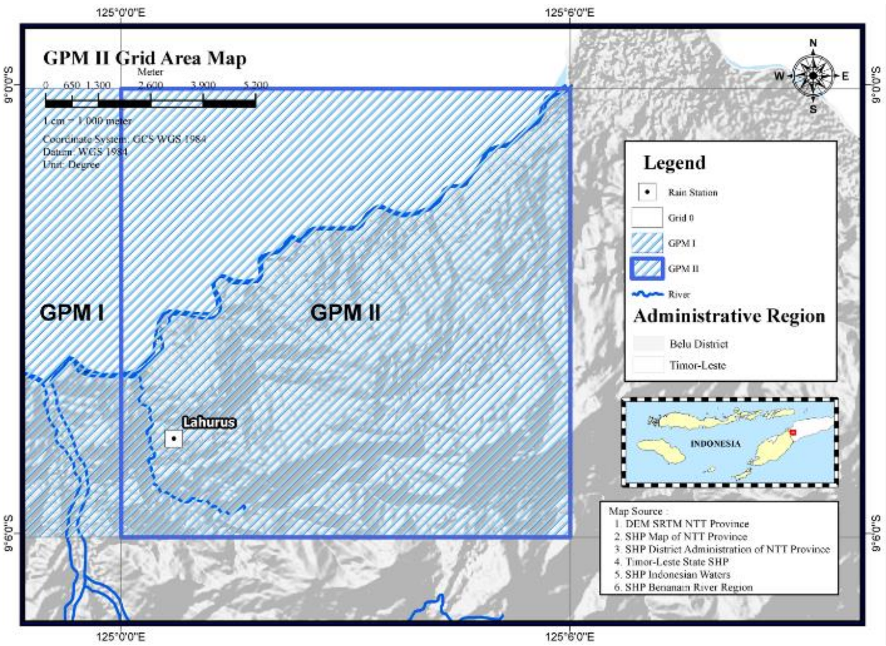


Figure 4. Map of GPM II Region

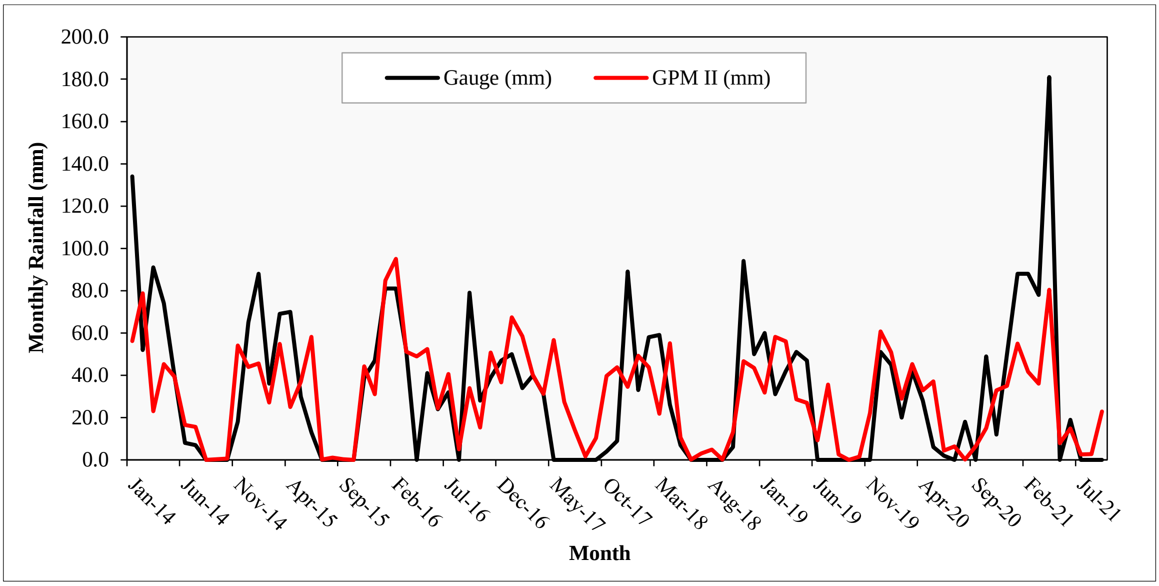


Figure 5. The correlation of Lahurus gauge and GPM II (monthly rainfall)

Table 3. Monthly maximum rainfall at Lahurus RS and GPM grid II

Month	Lahurus (mm)	GPM II (mm)	Month	Lahurus (mm)	GPM II (mm)
Jan-14	134.0	56.1	Jan-15	88.0	45.6
Feb-14	52.0	78.7	Feb-15	36.0	27.2
Mar-14	91.0	23.0	Mar-15	69.0	54.9
Apr-14	74.0	45.3	Apr-15	70.0	25.0
May-14	40.0	39.4	May-15	29.9	37.0
Jun-14	8.0	16.5	Jun-15	13.0	58.2
Jul-14	7.0	15.7	Jul-15	0.0	0.2
Aug-14	0.0	0.1	Aug-15	0.0	1.1
Sep-14	0.0	0.3	Sep-15	0.0	0.3
Oct-14	0.0	0.6	Oct-15	0.0	0.0
Nov-14	18.0	54.1	Nov-15	39.0	44.2
Dec-14	65.0	44.0	Dec-15	47.0	31.1

The calculated correlation coefficient value of 0.6454 shown at the monthly stages was quite strong, with a value above 0.5. The value showed IMERG GPM was able to reflect the value of the ground station, which was categorized with moderate tightness. In this study, the RMSE value was 26.48 (less than 300), hence, dropping to the class of very small. The value implied that the error from both RS data was very minimal. NSE value was 0.415, which was greater than 0.36 and was categorized as meeting the interpretative criteria. Based on the correlation coefficient, RMSE, and NSE values, Lahurus RS data was suitable for further analysis, particularly in determining correction factors.

3.2. Correction Factor and Error Value

The objective was to assess the extent of the difference between GPM data and ground stations in determining the correction factor for GPM satellite data as shown in Table 4. Consistency in the curvature pattern or the trend of rise and fall between the two compared datasets served as an identification factor in concluding the quality of GPM data. Moreover, data used in determining the correction factor included daily and annual maximum daily RD.

Table 4. Correction factor for GPM product

Daily Rainfall (mm)	Daily Correction Factor	Annual Maximum Daily Rainfall (mm)	Annual Correction Factor
10	0.8	10	0.8
30	0.9	30	0.9
50	1.1	50	1.1
110	1.2	110	1.2
>110	1.3	>110	1.3

These datasets were used for daily data correction and annual maximum daily RD analysis. In this study, the error calculation results for the probability of each rainfall class according to the corrected and GPM data were shown in Table 5.

Table 5. The error value for daily rainfall and annual rainfall

The average error for daily rainfall	GPM II (mm)	The average error for annual rainfall	GPM II (mm)
Before	0.0114	Before	0.1480
After	0.0109	After	0.1086
Deviation	4%	Deviation	27%

The Table 5 showed a decrease in error from GPM RD to corrected RD. The graphical representation of daily RD from Observation Post, GPM II, and Corrected GPM II was shown in Figure 6.

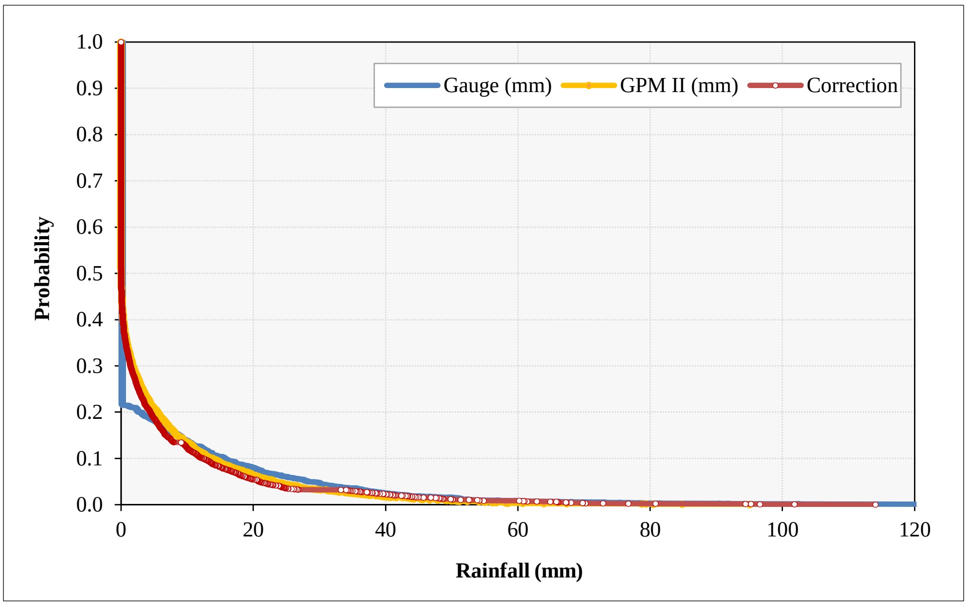


Figure 6. Daily rainfall correction with a probability curve

The annual maximum daily rainfall correction at the Lahurus rain gauge station was conducted using the same method that was shown in Figure 7. Based on the annual maximum daily rainfall, GPM data showed a pattern similar to that of the rain gauge station, though rainfall magnitudes differed at Lahurus station. Following this discussion, the error reduction achieved was 27%, as shown in Table 5.

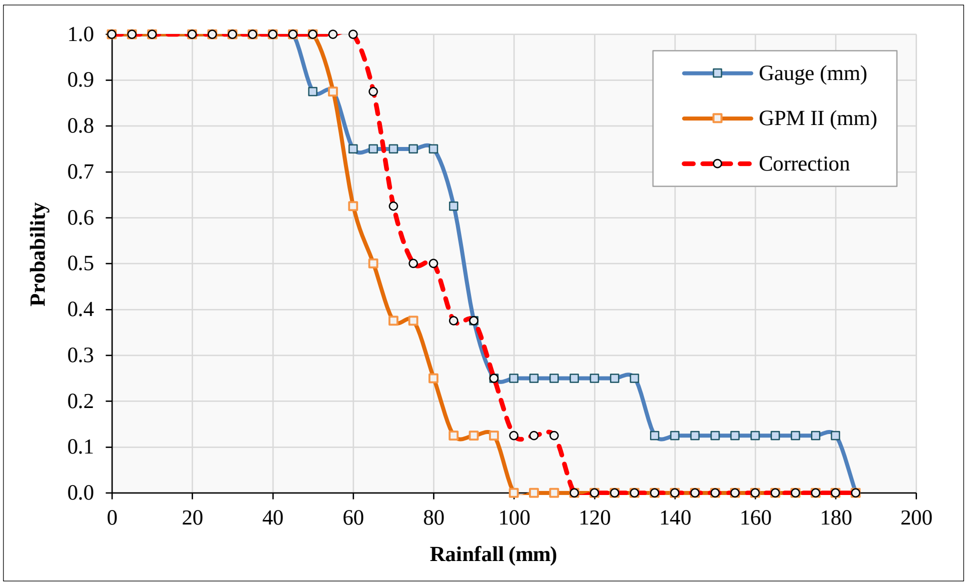


Figure 7. The value of the annual maximum daily rainfall correction

Figure 7 was a graph based on the curve models (annual maximum daily rainfall), in which the blue line was the curve of measurement data, the orange line represented the curve model of GPM RD, and the red line was the curve model of rainfall validation data. After calibration, the curve was more consistent with the CDF curve of the measurement data, showing the effectiveness of the modified model. Following this discussion, the precision of the curve model and correlation of the model were slightly higher compared to the correlation of the GPM original rainfall product. However, the RMSE of the curve was less than that of the original data, and the precision increased.

The results of rainfall distribution analysis for return periods of 2, 5, 10, 25, 50, 100, 200, 500, and 1000 years were obtained using the selected rainfall distribution. The selection of the distribution was based on the highest percentage of significance level in the Kolmogorov-Smirnov test. Subsequently, an analysis was conducted on rainfall values for return periods of 100 years compared to 2 years, GF, and the determination of the family curve. In this study, the test was performed using a Generalized Extreme Value (GEV) distribution.

Table 6. Return Period and GF for GPM II

Return Period	Probability	Rainfall		GF	
		Gauge Lahurus (mm)	Corrected of GPM II	Gauge Lahurus (mm)	Corrected of GPM II
1000	0.999	389.30	152.46	4.88	1.74
500	0.998	341.60	146.84	4.28	1.68
200	0.995	285.65	138.92	3.58	1.59
100	0.99	248.08	132.49	3.11	1.51
50	0.98	214.12	125.58	2.68	1.43
25	0.96	183.33	118.07	2.30	1.35
10	0.9	146.64	106.84	1.84	1.22
5	0.8	121.08	96.84	1.52	1.11
2	0.5	87.54	79.60	1.10	0.91
R_{100}/R_2		**2.83**	**1.66**		
Median		**79.85**	**87.60**		

Table 6 showed that R_{100}/R_2 was still in the range of 1.3 - 3.5. This result implied that the corrected GPM data was used in hydrological studies. GPM data calculation, along with correction factor values for RD calibration at five rain gauge stations in BRB, was shown in Table 7.

Table 7. Recapitulation of correction factor value for each GPM in BRB

Daily Rainfall (mm)	GPM I	GPM II	GPM III	GPM V	GPM VI
	Haliwen	Lahurus	Kaubele	Manufui	Ua'bau
	Correction Factor	Correction Factor	Correction Factor	Correction Factor	Correction Factor
10	0.8	0.8	0.8	0.8	0.8
30	0.9	0.9	0.9	0.9	0.9
50	1.1	1.1	1.1	1.1	1.1
110	1.2	1.2	1.2	1.2	1.2
>110	1.3	1.3	1.3	1.3	1.3
Annual Maximum Daily Rainfall (mm)	**Correction Factor**	**Correction Factor**	**Correction Factor**	**Correction Factor**	**Correction Factor**
10	0.8	0.8	0.8	0.8	0.8
30	0.9	0.9	0.9	0.9	0.9
50	0.95	0.95	0.95	0.95	0.95
110	1.2	1.2	1.1	1.1	1.1
>110	1.3	1.3	1.2	1.1	1.2

Based on the statistical analysis result, the approved RD was shown in Table 8.

Table 8. The result of the statistical analysis for the correlation of RD

No.	Ground station Name	Grid GPM	GPM vs Ground station			Description
			Correlation Coefficient	RMSE	NSE	
1	Haliwen	I	0.7046	25.89	0.476	Passes
2	Lahurus	II	0.6454	26.48	0.415	Passes
3	Kaubele	III	0.7907	25.28	0.604	Passes
4	Baurasi	IV	0.6275	32.46	0.306	Not Passes
5	Manufui	V	0.6568	35.23	0.422	Passes
6	Ua'bau	VI	0.6672	30.25	0.381	Passes
7	Noemuti	VII	0.4808	86.84	-0.29	Not Passes

Table 8 showed a strong correlation of the data, with correlation coefficient for GPM IV reaching 0.6275, implying a high level of correlation exception for GPM VII. In this study, RMSE values were 32.46 and 86.84, both less than 300, and were categorized as very small, signifying minimal error between the two datasets. However, the NSE value of GPM IV and VII was categorized to be unsatisfactory, where the values were less than 0.36 as shown in Table 1. These results implied that Lahurus and Noemuti ground station data were not reliable for further analysis. Based on the calculations from the previous point, the station was tested for rainfall characteristics and obtained the following results as shown in Table 9.

Table 9. The result of GPM data passed the characteristics test

GPM grid	Statistic Passes	Characteristics Test
I	√	-
II	√	√
III	√	√
V	√	-
VI	√	-

The performance of GPM was not shown to be influenced by the elevation, amount, and nature of rainfall (convective or orographic). The method performed well in all zones of the country, except for the semi-arid region of Timor Island in Indonesia. However, the correlation of GPM with station rainfall was not statistically significant in the region.

4. Discussion

The results of this study further strengthened previous investigations showing that monthly GPM data has a good correlation coefficient compared to daily data. GPM monthly products had strong applicability and accurately reflected rainfall events in the BRB region. The comparison of monthly average rainfall showed that the 12-month rainfall distribution from satellite data closely matched the data from ground meteorological stations, effectively representing rainfall for each month. For daily data, the deviation between GPM data and rain gauge data was quite significant. Moreover, Table 6 showed that the correction multiplier for GPM data in terms of daily rainfall ranged from 0.8 to 1.30. As the intensity of rainfall became greater, the multiplier value became higher. This process implied that during extreme rainfall events exceeding 100 mm, GPM data tends to miscalculate the actual daily rainfall. Consequently, for rainfall intensities ≤ 50 mm, GPM values were closer to the conditions observed at ground stations.

The performance of GPM products in the rainy season with rainfall ≤ 50 mm was better compared to extreme rainfall ≥ 100 mm. However, the precision of GPM satellites in undulating regions was poor, which could easily lead to serious errors. The curve models were successfully used to modify GPM daily as well as the GPM annual maximum daily rainfall product and improve the accuracy. Moreover, corrected GPM daily RD in the range >100 mm was underestimated. The results obtained from the study were almost the same as those of a previous study conducted in Northeast China [28]. The only difference was that rainfall values ≥ 100 mm in the study conducted in Beijing tended to be overestimated, while in Timor of Indonesia, the values were underestimated. The GPM IMERG-F product at daily and monthly timescales in this study had good detection ability and could provide data support for long-time series investigations in the semi-arid region of Indonesia. Therefore, daily GPM data and annual maximum daily rainfall needed to be corrected before proceeding with the hydrological analysis.

5. Conclusion

In conclusion, this study examined the potential use of RD obtained from GPM IMERG satellite measurements for hydrological analysis in semi-arid regions. The analysis was conducted by comparing monthly, daily, and annual maximum daily RD from GPM with data obtained from ground stations. In addition, data validation was performed to obtain the correlation coefficient for monthly data and the correction factor for daily data. The correlation coefficient values for GPM IMERG ranged from 0.6275 to 0.7907, except for GPM VII, which had values less than 0.600. Moreover, RMSE error values for GPM ranged from 25.28 to 86.84, and NSE error values for GPM ranged from (-0.29) to 0.604. In BRB, 7 ground stations were available, but only 5 met the statistical test requirements. Additionally, the ground stations and GPM grids that passed the statistical test were 5 RS, namely Haliwen RS (GPM I), Lahurus RS (GPM II), Kaubele RS (GPM III), Manufui RS (GPM V), and Ua'bau RS (GPM VI).

Daily data had 2% GPM I, 4% GPM II, 9% GPM III, 13% GPM V, and 6% GPM VI error reduction percentages, respectively. Consequently, annual data had 16% GPM I, 27% GPM II, 18% GPM III, 4% GPM V, and 23% GPM VI error reduction. The performance of GPM products on the daily scale was better during the transition from summer to the rainy season (rainfall ≤ 50 mm) compared to the rainy season (rainfall ≥ 100 mm). The daily RD and annual maximum daily rainfall from the GPM product IMERG-F are necessary to be corrected before proceeding with the hydrological analysis.

GPM grids that passed the rainfall characteristic test (R_{100}/R_2 in the range of 1.5 - 3.4) consisted of 2 RS, including Lahurus Station (GPM II) with a 100-year divided by a 2-year return period of 1.66 and Kaubele Station (GPM III) with a 100-year divided by a 2-year return period of 3.04. For daily GPM data, the HHMT correction factor obtained ranged from 0.029 to 0.252, and for daily RD, the correction factor obtained ranged from 0.006 to 0.013.

6. Declarations

6.1. Author Contributions

Conceptualization, D.S.K. and A.D.; methodology, D.S.K. and A.D.; software, R.R.K. and J.J.S.P.; validation, D.S.K., A.D., and J.J.S.P.; formal analysis, D.S.K. and A.D.; investigation, D.S.K., J.J.S.P., and A.D.; resources, A.D. and S.; data curation, D.S.K. and S.; writing—original draft preparation, D.S.K. and A.D.; writing—review and editing, D.S.K., R.R.K., and S.; visualization, R.R.K.; supervision, D.S.K. and J.J.S.P.; project administration, D.S.K. and S. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Acknowledgments

The authors are grateful to the Civil Engineering Study Program at Nusa Cendana University, Balai Wilayah Sungai NT II, and several other parties that supported this study.

6.5. Conflicts of Interest

The authors declare no conflict of interest.

7. References

[1] Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., Huang, M., Yao, Y., Bassu, S., Ciais, P., Durand, J. L., Elliott, J., Ewert, F., Janssens, I. A., Li, T., Lin, E., Liu, Q., Martre, P., Müller, C., … Asseng, S. (2017). Temperature increase reduces global yields of major crops in four independent estimates. Proceedings of the National Academy of Sciences of the United States of America, 114(35), 9326–9331. doi:10.1073/pnas.1701762114.

[2] Kummerow, C., Barnes, W., Kozu, T., Shiue, J., & Simpson, J. (1998). The Tropical Rainfall Measuring Mission (TRMM) sensor package. Journal of Atmospheric and Oceanic Technology, 15(3), 809–817. doi:10.1175/1520-0426(1998)015<0809:TTRMMT>2.0.CO;2.

[3] Huffman, G. J., Adler, R. F., Morrissey, M. M., Bolvin, D. T., Curtis, S., Joyce, R., McGavock, B., & Susskind, J. (2001). Global Precipitation at One-Degree Daily Resolution from Multisatellite Observations. Journal of Hydrometeorology, 2(1), 36–50. doi:10.1175/1525-7541(2001)002<0036:gpaodd>2.0.co;2.

[4] Wang, H., Yuan, Y., Zeng, S., Li, W., & Tang, X. (2021). Evaluation of satellite-based precipitation products from GPM IMERG and GSMaP over the three-river headwaters region, China. Hydrology Research, 52(6), 1328–1343. doi:10.2166/NH.2021.029.

[5] Ramadhan, R., Muharsyah, R., Marzuki, Yusnaini, H., Vonnisa, M., Hashiguchi, H., Suryanto, W., & Sholihun. (2022). Evaluation of GPM IMERG Products for Extreme Precipitation over Indonesia. Journal of Physics: Conference Series, 2309(1), 1–8. doi:10.1088/1742-6596/2309/1/012008.

[6] Hutagaol, S. M., E. Ball, J., Suhartanto, E., & Wahyuni, S. (2023). The Evaluation of GPM IMERG v.06 Rainfall Product over the Lau Simeme Watershed in Indonesia. Civil and Environmental Science, 6(1), 33–42. doi:10.21776/ub.civense.2023.00601.5.

[7] Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., ... & Xie, P. (2020). NASA GPM Integrated Multi-SatellitE Retrievals for GPM (IMERG) Algorithm Theoretical Basis Document (ATBD) Version 06. National Aeronautics and Space Administration: Washington, United States.

[8] Sungmin, O., Foelsche, U., Kirchengast, G., Fuchsberger, J., Tan, J., & Petersen, W. A. (2017). Evaluation of GPM IMERG Early, Late, and Final rainfall estimates using WegenerNet gauge data in southeastern Austria. Hydrology and Earth System Sciences, 21(12), 6559–6572. doi:10.5194/hess-21-6559-2017.

[9] Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., & Iguchi, T. (2014). The global precipitation measurement mission. Bulletin of the American Meteorological Society, 95(5), 701–722. doi:10.1175/BAMS-D-13-00164.1.

[10] Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., & Wolff, D. B. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. Journal of Hydrometeorology, 8(1), 38–55. doi:10.1175/JHM560.1.

[11] Keikhosravi-Kiany, M. S., & Balling, R. C. (2024). Evaluation of GPM IMERG Early, Late, and Final Run in Representing Extreme Rainfall Indices in Southwestern Iran. Remote Sensing, 16(15), 2779. doi:10.3390/rs16152779.

[12] Zhang, D., Yang, M., Ma, M., Tang, G., Wang, T., Zhao, X., Ma, S., Wu, J., & Wang, W. (2022). Can GPM IMERG Capture Extreme Precipitation in North China Plain? Remote Sensing, 14(4), 928. doi:10.3390/rs14040928.

[13] Wahyuni, S., Hambali, R., & Fahriani, F. (2020). Evaluation of the Accuracy of GPM Satellite Rainfall Data on Surface Rainfall Data. Pertemuan Ilmiah Tahunan (PIT) HATHI-37. Universitas Bangka Belitung, 1083-1094. (In Indonesian).

[14] Mamenun, M., Pawitan, H., & Sopaheluwakan, A. (2014). Validasi Dan Koreksi Data Satelit Trmm Pada Tiga Pola Hujan Di Indonesia. Jurnal Meteorologi Dan Geofisika, 15(1). doi:10.31172/jmg.v15i1.169.

[15] Tang, G., Ma, Y., Long, D., Zhong, L., & Hong, Y. (2016). Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over Mainland China at multiple spatiotemporal scales. Journal of Hydrology, 533, 152–167. doi:10.1016/j.jhydrol.2015.12.008.

[16] Chen, H., Yong, B., Shen, Y., Liu, J., Hong, Y., & Zhang, J. (2020). Comparison analysis of six purely satellite-derived global precipitation estimates. Journal of Hydrology, 581. doi:10.1016/j.jhydrol.2019.124376.

[17] Bulovic, N., McIntyre, N., & Johnson, F. (2020). Evaluation of imerg v05b 30-min rainfall estimates over the high-elevation tropical andes mountains. Journal of Hydrometeorology, 21(12), 2875–2892. doi:10.1175/JHM-D-20-0114.1.

[18] Et-Takaouy, C., Aqnouy, M., Boukholla, A., & Stitou El Messari, J. E. (2024). Exploring the spatio-temporal variability of four satellite-based precipitation products (SPPs) in northern Morocco: a comparative study of complex climatic and topographic conditions. Mediterranean Geoscience Reviews, 6(2), 123–144. doi:10.1007/s42990-024-00119-5.

[19] Rachdane, M., El mehdi Saidi, M., & Tramblay, Y. (2022). Evaluation of GPM IMERG products and ERA5 reanalysis for flood modeling in a semi-arid watershed. Copernicus Meetings, No. IAHS2022-518. doi:10.5194/iahs2022-518.

[20] Tian, Y., Lv, X., Guo, H., Li, J., Meng, X., Guo, C., Zhu, L., & De Maeyer, P. (2024). Evaluation of GPM IMERG Product over the Yellow River Basin Using an Improved Error-Component Procedure. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17, 8918–8937. doi:10.1109/JSTARS.2024.3392601.

[21] Tian, Y., Nearing, G. S., Peters-Lidard, C. D., Harrison, K. W., & Tang, L. (2016). Performance metrics, error modeling, and uncertainty quantification. Monthly Weather Review, 144(2), 607–613. doi:10.1175/MWR-D-15-0087.1.

[22] Bunganaen, W., Krisnayanti, D. S., Nasjono, J. K., & Saudale, N. S. (2023). Pemanfaatan Data Tropical Rainfall Measuring Mission Pada Stasiun Hujan Di Wilayah Sungai Benanain. Jurnal Ilmiah Desain & Konstruksi, 22(2), 168–179. doi:10.35760/dk.2023.v22i2.8627.

[23] Badan Pusat Statistik NTT. (2021). Statistics of East Nusa Tenggara. Kupang, Indonesia. (In Indonesian).

[24] Asdak, C. (2023). Hydrology and river basin management. UGM Press, Yogyakarta, Indonesia. (In Indonesian).

[25] Iqbal Hasan, M. (2009). Main points of statistical material 1. Bumi Aksara, Jakarta, Indonesia. (In Indonesian).

[26] Jarwanti, D. P., Suhartanto, E., & Fidari, J. S. (2021). Validation of TRMM (Tropical Rainfall Measuring Mission) Satellite Rainfall Data with Rain Gauge Post Data in Grindulu Watershed, Pacitan Regency, East Java. Journal of Water Resources Technology and Engineering, 1(2), 772–785. doi:10.21776/ub.jtresda.2021.001.02.36.

[27] Liu, Y., Wu, Y., Feng, Z., Huang, X., & Wang, D. (2017). Evaluation of a Variety of Satellite Retrieved Precipitation Products Based on Extreme Rainfall in China. Tropical Geography, 37(3), 417–433. doi:10.13284/j.cnki.rddl.002949.

[28] Sun, W., Sun, Y., Li, X., Wang, T., Wang, Y., Qiu, Q., & Deng, Z. (2018). Evaluation and correction of GPM IMERG precipitation products over the capital circle in Northeast China at multiple spatiotemporal scales. Advances in Meteorology, 2018, 1–14. doi:10.1155/2018/4714173.

[29] Bendungan, B. (2019). Design Rainfall Analysis with TRMM Data. Bimbingan Teknis Balai Bendungan, Palembang, Indonesia.

[30] NASA. (2022). The Global Rainfall Measurement Mission. National Aeronautics and Space Administration (NASA), Washington, United States. Available online: https://gpm.nasa.gov/missions/GPM (accessed on November 2024).

[31] Murphy, C., Cunnane, C., Das, S., & Mandal, U. (2014). Flood Studies Update Technical Research Report Volume II Flood Frequency Estimation, Technical Research Reports, NNUI Galway, Galway, Ireland.