https://doi.org/10.48048/tis.2024.7592

Bias Correction of IMERG Data in the Mountainous Areas of Sumatra Based on A Single Gauge Observation

Ravidho Ramadhan[1,2], Marzuki Marzuki[2], Wiwit Suryanto[1,*], Sholihun Sholihun[1], Helmi Yusnaini[2], Hiroyuki Hashiguchi[3] and Toyoshi Shimomai[4]

[1]Department of Physics, Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia
[2]Department of Physics, Universitas Andalas, Padang 25163, Indonesia
[3]Research Institute for Sustainable Humanosphere, Kyoto University, Kyoto 606-8501, Japan
[4]Department of Electronic and Control Systems Engineering, Shimane University, Shimane 690-8504, Japan

([*]Corresponding author's e-mail: ws@ugm.ac.id)

Received: 10 October 2023, Revised: 26 October 2023, Accepted: 27 October 2023, Published: 15 February 2024

Abstract

The performance of surface precipitation data from satellite precipitation products (SPPs) in mountainous areas has greater error and bias than in plain areas. In this study, linear scaling (LS), local intensity (LOCI), power transformation (PT), and cumulative distribution function (CDF) methods are used to correct the bias of Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (IMERG) data in the mountainous region of Sumatra based on long-term and high-resolution optical rain gauge (ORG) observations. The ORG is installed at Equatorial Atmospheric Observatory (EAO) in Kototabang, West Sumatra, Indonesia (100.32 °E, 0.20 °S, 865 m above sea level (ASL) with an observation period from 2002 to 2016. The impact of the bias correction method is tested based on accuracy and capability detection tests. The bias correction method is more effective at the daily resolution than the hourly resolution of the IMERG data in the mountainous region of Sumatra. The LS method exhibited the best improvement in accuracy with reduced root-mean-square error (RMSE) and relative bias (RB), although there was no significant increase in coefficient correlation (CC) values. However, the accuracy improvement was not–observed in the bias correction for hourly data. The lack of improvement in the accuracy of the hourly IMERG data is due to the high local variability of rainfall in the mountainous area of Sumatra. The high data variability causes large differences in the mean and variance of the IMERG calibration and evaluation data periods. On the other hand, the LOCI, PT, and CDF methods were successfully improved the rain detection capability of IMERG, as indicated by the better critical succession index (CSI) values compared to the original hourly and daily IMERG data. It increased the CSI value by reducing false alarms for rain with intensity below 2 mm/h. Furthermore, the CDF method can improve the analysis of extreme rainfall in the mountainous region of Sumatra by improving the estimation of the extreme rainfall index. Therefore, these methods can be applied to improve the accuracy and detectability of IMERG data in the mountainous region of Sumatra. However, the scale factor and transfer function constructed in this study need to be further evaluated on other rain gauge observation data in Sumatra's mountainous region to improve performance.

Keywords: IMERG, Mountainous, Bias correction, Optical rain gauge, Kototabang

Introduction

The development of surface rainfall estimation technology from satellite observations has grown rapidly over the past few decades. One of the best resolution satellite precipitation products (SPPs) data is SPPs developed by the National Aeronautics and Space Administration (NASA). NASA's best resolution SPPs data is currently surface rainfall data based on satellite observations of the Global Precipitation Measurement (GPM) constellation. NASA's multi-satellite rainfall data is known as Integrated Multi-Satellite Retrievals for GPM (IMERG) [1]. IMERG data has a resolution of 0.1 ° - 30 min with a quasi-global observing range (60 °S - 60 °N). The good resolution and wide coverage of IMERG have the potential to be utilized in various applications in hydrology, meteorology, and climatology [2-5].

Researchers worldwide have been evaluated the performance of IMERG due to its wide potential utilization [6], providing valuable information for NASA to improve algorithms. The validation results also serve as a reference for users to assess the accuracy of IMERG data in various applications. However, it is important to note that IMERG data still contains errors and uncertainties related to sensor limitations, incorrect observation assumptions, and algorithm usage [7,8]. These factors contribute to the need for ongoing development and improvement of IMERG data. Thus, it is expected that the accuracy of IMERG data will improve, further expanding its utilization and applicability in different fields.

To improve the accuracy and reliability of IMERG data, one approach is to employ bias correction methods [9]. Bias correction involves using observed rainfall data on the surface as a reference to correct the IMERG data. Previous studies have demonstrated the effectiveness of bias correction methods in reducing errors and uncertainties in IMERG data [10-12]. However, correcting IMERG data in areas with complex topography, like mountainous regions, remains a challenge [13-15]. The challenges of IMERG bias correction in complex topographic regions are due to the complex weather systems in these regions. The complex weather system in these areas is caused by complex local wind patterns and orographic lift due to topographic effects. Overall, bias correction methods have proven beneficial in enhancing the quality of IMERG data, but further advancements are needed to address the challenges in correcting IMERG data in complex terrain areas.

There are several studies that have conducted to improve the accuracy of IMERG data in mountainous areas by minimizing false detection of rainfall events [10,12,13]. However, bias correction in these areas remains unsatisfactory due to limited surface rainfall observations as reference points. The number and resolution of surface gauges used for bias correction affect the accuracy improvement and detection capabilities of IMERG data [16]. In developing countries with limited surface networks, it is challenging to apply bias correction methods to mountainous regions. Therefore, further research is needed to test bias correction methods of IMERG data based on single gauge observations in mountainous areas. It would contribute to better understanding and correction of rainfall data in the mountainous areas with very limited gauge observation.

This research focuses on correcting biases in the IMERG data for rainfall in the mountainous area of Sumatra, which is located in the western of Indonesia maritime continent (IMC). The IMC is known for its active cumulus activity, making accurate rainfall data crucial for monitoring and mitigating hydrometeorological disasters. In addition, the mountainous region of Sumatra located in the western part of the IMC plays an important role in local rainfall patterns due to land-sea interaction [17,18]. The localized rainfall pattern in this region means that surface rainfall data with high accuracy and resolution is needed to improve regional weather and climate forecasting models. Previous evaluations have shown that the IMERG data in the IMC still needs improvement, especially for daily and hourly resolution and in areas with complex topography [19-24]. The IMC region has a complex land surface and numerous mountains due to its location in the Ring of Fire. Therefore, the mountainous regions of the IMC largely lack an adequate surface observation network. Thus, correcting the biases in the IMERG data specifically in mountainous areas of the IMC would enhance the accuracy of surface rainfall data for daily and hourly resolution in these regions.

Materials and methods

Study area and optical rain gauge data

The research site is the Equatorial Atmospheric Observatory (EAO) of the National Research and Innovation Agency (BRIN) in Kototabang (100.32 °E, 0.20 °S, 865 m above sea level (ASL)), West Sumatra, Indonesia (**Figure 1**). EAO was established in 2001 by the Indonesian government in cooperation with Kyoto University, Japan. The main purpose of EAO is to observe the dynamics of the equatorial atmosphere. Its main instrument is the Equatorial Atmospheric Radar (EAR) [25]. In addition to EAR, there are several other observation instruments, including high-resolution observation of surface rainfall, namely Optical Rain Gauge (ORG).

The Optical Rain Gauge (ORG) in Kototabang is an output of the Optical Scientific inc (OSI) company with the ORG-815 series that measures rainfall based on scintillation technology. The basic principle of scintillation technology in measuring rainfall is based on variations in the intensity of the laser captured by the sensor when raindrops pass through it. In detail, the mechanism of ORG in measuring rainfall can be seen on the website of the development company (OSI website). The recording resolution of the ORG in Kototabang is set for every 1-min with rainfall observation accuracy in the range of 0.1 - 500 mm/h [26]. The ORG in Kototabang started operating in March 2002 and stopped recording in December 2016. The percentage of ORG data availability over the 15 years can be seen in **Figure 2(a)**.

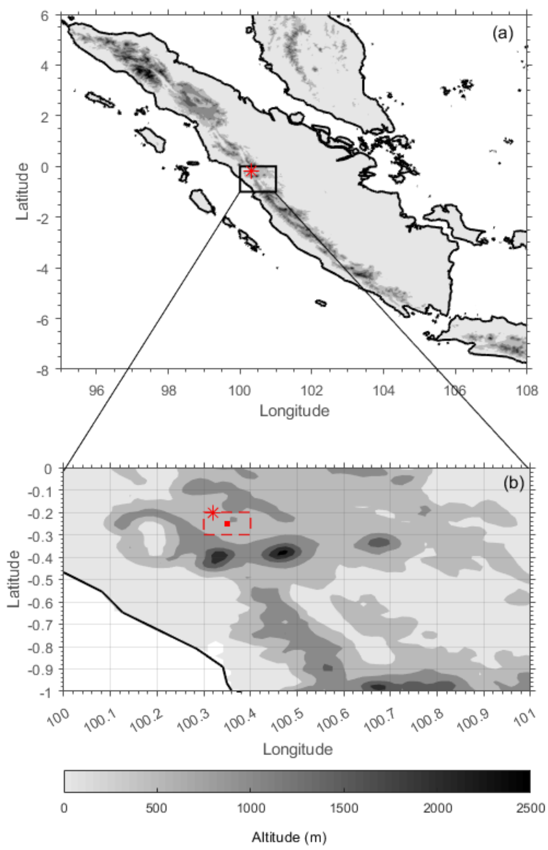

Figure 1 The location of Kototabang on Sumatra Island (a) and the topography of the ORG installation site in Kototabang (b). The red asterisk indicates the location of ORG, the red rectangle is the centroid of the IMERG grid closest to ORG, and the dashed red line is the IMERG data grid being compared.

Previous studies have characterized the rainfall climatology of the Kototabang region using ORG observations and other instruments. Kototabang has high annual rainfall with a range of 2400 mm/year due to its location near the equator and in the mountainous region of Sumatra [20]. The Inter-tropical convergence zone (ITCZ) strongly influences the seasonal rainfall in this region, along with the southeast and northwest monsoons [27-29]. As a result, Kototabang has 2 peaks in the rainy season, occurring in March-April and November-December [20,22]. Additionally, the area exhibits a significant diurnal rainfall pattern, with the peak rainfall occurring in the afternoon between 1600 - 1700 Local Time (LT) [22,30,31]. This diurnal pattern is mainly influenced by rain migration from the west coast of Sumatra [17,18,32]. In addition, rainfall in Kototabang is also strongly influenced by the Madden-Julian Oscillation (MJO) [22,33]. Rainfall accumulation in Kototabang shows an increase when the MJO is strong and in phase 2 - 4, which is caused by an increase in the frequency of long-duration rainfall events. Research on characteristics and mechanism of rainfall in the Kototabang and mountainous regions of Sumatra is still ongoing, so the availability of accurate and high-resolution surface rainfall data is needed.

Methodology

We conducted bias correction of ORG-based IMERG data in Kototabang for IMERG version 06 (V06) data. The focus of the study was on bias correction of the final type IMERG data, which is considered to have the best accuracy over IMC due to the use of more complete data by the IMERG algorithm [24]. The final IMERG data was also calibrated with monthly data from the Deutscher Wetterdienst (DWD) Global Precipitation Climatology Center (GPCC) Full/Monitoring product [1]. By performing bias correction of the IMERG Final data using daily and ORG data at Kototabang, the aim was to improve accuracy with a more specific approach. The approach used for the IMERG and ORG data was a direct or point-to-pixel method, allowing for comparison with previous studies conducted at the same location [20-22].

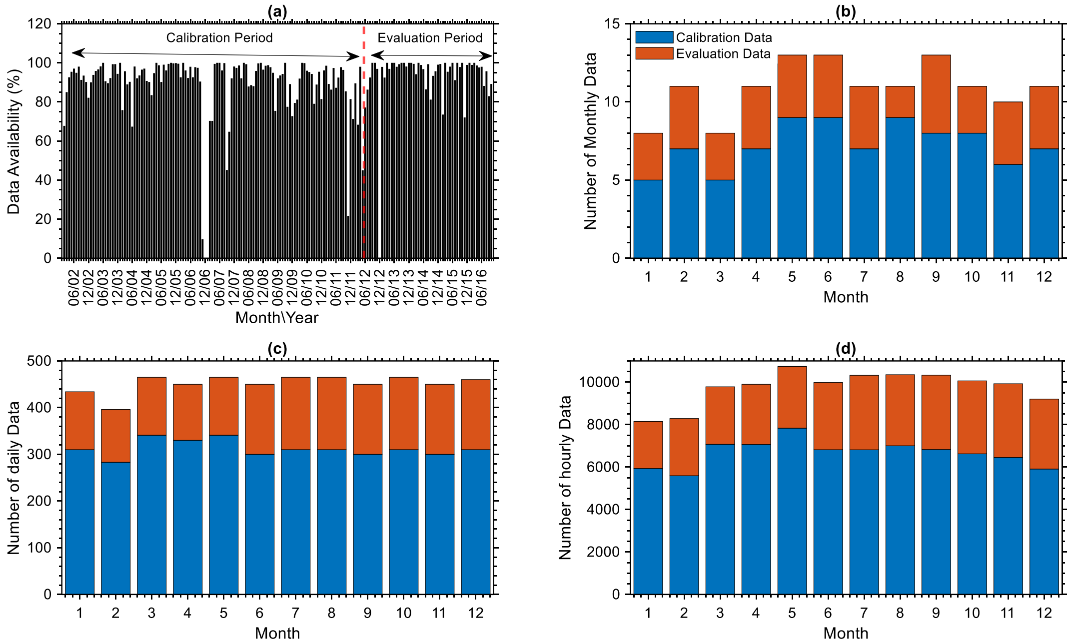

Figure 2 Monthly data availability of ORG data in Kototabang during the observation period (a), the number of complete monthly (b), daily (c), and hourly (d) ORG data.

In this bias correction method, the observation data from ORG is divided into 70 % calibration data and 30 % validation data. The calibration and validation periods can be seen in **Figure 2(a)**. The calibration and validation data include complete months, days, and hours, as shown in **Figures 2(b) - 2(d)**. Complete hourly and daily data are defined as ORG 1-min data without any gaps in recording, while complete monthly data is the sum of complete hourly data with > 90 % completeness in that month. The bias correction and validation are performed only on the complete hourly and daily data to avoid errors caused by missing recordings. The bias correction methods used in this study are linear scaling (LS), local intensity (LOCI), power transformation (PT), and cumulative distribution function (CDF). These methods are very useful for the correction of bias in precipitation data from model and satellite data processing [11,34,35].

Linear scaling (LS) method

The LS method is a bias correction method based on the ratio of the mean precipitation value of the reference data to the estimated data [36]. In this case, the reference data used are ORG data, while the estimated data are IMERG data. The calibrated rainfall data is calculated by multiplying the rainfall data by a scaling factor calculated from the average monthly ORG and IMERG rainfall ratios. The calculation of the precipitation with the LS-Method can be described as follows:

$$P^*_{m,i} = P_{m,i} \,.\, s_m \tag{1}$$

$$\mathrm{s_m} = \frac{\mu_m(G_{m,i})}{\mu_m(P_{m,i})} \tag{2}$$

where $P^*_{m,i}$ is the bias corrected data of the *i*-th day or era in month m, $P_{m,i}$ dan $G_{m,i}$ are the *i*-th day or era in month m from IMERG and ORG data, s_m is the scale factor in month *m*, and μ_m is the monthly average.

Local intensity (LOCI) method

The bias correction in the LOCI method follows a similar principle as the LS method: The IMERG data are multiplied by a scaling factor. The difference is that the LS method first sets a threshold event value (P_{th}) that is defined as rain or not [37]. In this study, the P_{th} value is set based on the rain detection capability of the IMERG data before correction. The data transformation can be written mathematically:

$$P^*_{m,i} = \begin{cases} 0, & if\ P_i < P_{th} \\ P_{m,i} \,.\, c, & otherwise \end{cases} \tag{3}$$

where c is the scaling factor in the LOCI method calculated by the following equation:

$$c = \frac{\mu_m(G_{m,i}|G_{m,i} \geq 0.1)}{\mu_m(P_{m,i}|P_{m,i} \geq P_{th})} \quad (4)$$

c values are calculated for each month and used for bias correction of daily rainfall data and the same ages as the LS method.

Power transformation (PT) method

PT method is a bias correction technique that aims to adjust the variance of the estimated data with the reference data [38]. The correction is done in a non-linear way in the form of an exponential which is written as follows:

$$P^*_{m,i} = aP^{b_m}_{m,i} \quad (5)$$

where b_m is power parameter that calculated based on the minimum value of the difference in coefficient of variation (CV) calculated iteratively based on daily and monthly rainfall data. The correction CV was calculated as follows:

$$f(b_m) = 0 = \frac{\sigma_m(G_{m,i})}{\mu_m(G_{m,i})} - \frac{\sigma_m(P^{b_m}_{m,i})}{\mu_m(P^{b_m}_{m,i})} \quad (6)$$

where σ_m is the monthly standard deviation. In this study, iteration values with b_m of 0.01 were used. After obtaining the appropriate b_m value, the value of the scaling factor *a* was calculated based on the comparison of the mean value of $P^{b_m}_{m,i}$ and $G_{m,i}$.

Cumulative distribution function (CDF) method

Unlike the 3 previous methods, which are based on the scaling factor of the mean value, the CDF method corrects the bias of the IMERG data based on the probabilistic distribution of the IMERG and ORG data [39]. The step in the CDF method is to match the rainfall data values from IMERG and ORG that have the same cumulative probability, and then equalize them with a transformation function. The transformation function of the CDF method is based on the following equation:

$$C_1(P_i) = \int f_1(P_i)\, dP_i \quad (7)$$

$$C_1(G_i) = \int f_1(G_i)\, dG_i \quad (8)$$

$$f_3(P_i) = C_2^{-1}(C_1(P_i)) \quad (9)$$

$$P_i^* = f_3(P_i) \quad (10)$$

where f_1 and f_2 are the probability functions of P_i and G_i, f_3 is a transformation function that describes the relationship between P_i and Gi.

Statistical evaluation matrices

To evaluate the accuracy of the IMERG data bias correction method in the mountainous region of Sumatra, we used 6 statistical evaluation matrices that are also commonly used worldwide [6]. The 6 statistical evaluation matrices consist of 3 accuration tests and 3 capability of detection tests. The 3 accuration test matrices include coeffecient correlation (CC), root-mean square error (RMSE), and relative bias (RB). The 3 matrices are calculated based on the following equation [40]:

$$CC = \frac{\sum_{i=1}^{n}(P_i^* - \overline{P_i^*})(G_i - \overline{G_i})}{\sqrt{\sum_{i=1}^{n}(P_i^* - \overline{P_i^*})^2 \sum_{i=1}^{n}(G_i - \overline{G_i})^2}} \quad (11)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i^* - G_i)^2} \quad (12)$$

$$RB = \frac{\sum_{i=1}^{n}(P_i^* - G_i)}{\sum_{i=1}^{n} G_i} x100\% \quad (13)$$

where P_i^* is the corrected IMERG data, Gi is the ORG data for evaluation, $\overline{P_i^*}$ and $\overline{G_i}$ is mean values of $P_i^*$and Gi. The CC value ranges from 0 to 1, indicating no correlation (0) and perfect correlation (1). The best RMSE value is 0, and the larger the RMSE, the larger the average error. The best RB value is 0, while a negative (positive) RB value indicates an underestimation (overestimation) of the IMERG calibrated data with respect to the ORG observations.

Furthermore, the 3 matrices of the capability of detection test used are probability of deteciton (POD), false alarm ratio (FAR), and critical succession index (CSI). The POD, FAR, and CSI values are calculated based on the following equations [41]:

$$POD = \frac{H}{H+M} \quad (14)$$

$$FAR = \frac{F}{H+F} \quad (15)$$

$$CSI = \frac{H}{H+F+M} \quad (16)$$

where H (hit) is the condition that calibrated IMERG data and ORG both identify a rain event (P_i^* ≥ threshold & G_i ≥ threshold), M (miss) is the condition that calibrated IMERG data does not identify rain while ORG identifies rain (P_i^* < threshold & G_i ≥ threshold), and F (false) is the condition that calibrated IMERG identifies rain while ORG does not (P_i^* ≥ threshold & G_i < threshold). The values of the 3 detection capability tests range from 0 to 1. The best value for POD and CSI is 1, while the best value for FAR is 0.

Results and discussion

Scaling factor parameterization

The first step in scaling-based bias correction is to determine the scaling factor parameter values of the IMERG calibration data and reference data. The scaling factors include the S_m parameter for LS, the *a* parameter for LOCI, and the *c* parameter for PT. In the LS method, the S_m value can be calculated directly as described in the methodology section, but in the LOCI and PT methods there are several steps. This sub-section describes each step in determining the scaling factor values of these methods.

The LOCI method determines the threshold value based on the detection capability of IMERG data compared to ORG data during the calibration period. The POD, FAR, and CSI parameter values of IMERG data are analyzed with different thresholds (**Figures 3(a) - 3(b)**). In hourly scale, increasing the threshold of observed rainfall intensity decreases FAR but also decreases the POD, resulting in a trade-off between accurate detection and information loss. The best threshold value is determined when the decrease in FAR is greater than the decrease in POD, indicated by the optimum value of the CSI parameter. In this study, the optimum threshold value for rainfall in Kototabang is found to be 0.5 mm/h. This finding is consistent with the previous study that the probability density function (PDF) value of the IMERG data at Kototabang overestimates (underestimates) rainfall with a threshold of > 0.1 mm (≤ 0.1 mm) [21]. Therefore, this threshold value of 0.5 mm/h will be used as the P_{th} value in the LOCI method for the IMERG data in the Kototabang.

Unlike the time data, the CSI value in the daily data tends to decrease when the threshold value of IMERG data rainfall intensity is increased (**Figure 3(b)**). This decrease is primarily driven by a decrease in the Probability of Detection (POD) value and an increase in the False Alarm Ratio (FAR) value. The study uses a threshold of 0.5 mm/day as the optimum value to identify epochal rainfall events in the daily data. It notes that a threshold of 1 mm/day is commonly used to identify wet days and extreme weather events globally [42,43]. However, in the mountainous region of Sumatra, using a threshold of 1 mm/day results in a slight bias due to the loss of wet days data. To avoid this bias, the study recommends using a threshold of 0.5 mm/day for bias correction of IMERG data with the LOCI method in Kototabang.

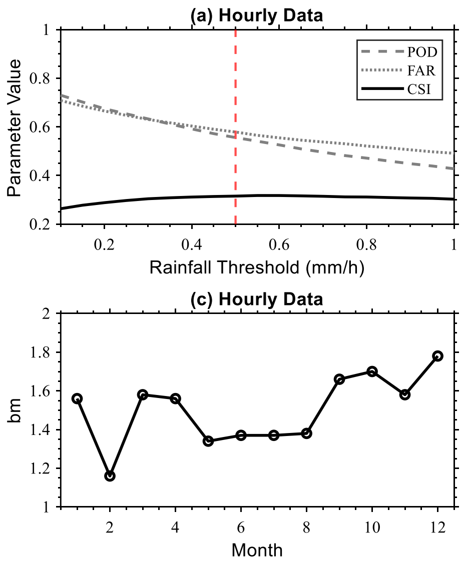

Figure 3 Monthly values of the capability of the detection parameters value and the b_m parameters of the hourly (a and c) and daily (b and d) data for the IMERG and ORG data during the calibration period. The dashed vertical red line is the optimum value of CSI.

Table 1 Values of the scaling parameters of the LS, LOCI, and PT methods for the IMERG data at Kototabang.

	Month											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Hourly												
S_m	0.557	0.667	0.693	0.752	0.894	0.951	0.680	0.837	0.745	0.749	0.869	0.846
c	0.933	1.140	1.038	1.143	1.239	1.506	1.097	1.174	1.116	1.065	1.128	1.112
a	0.271	0.536	0.335	0.327	0.608	0.671	0.471	0.529	0.301	0.289	0.442	0.291
Daily												
S_m	0.570	0.751	0.696	0.715	0.906	0.931	0.704	0.789	0.737	0.746	0.845	0.795
c	0.967	1.152	1.068	0.974	1.561	1.550	1.273	1.218	1.087	1.022	1.077	1.087
a	0.107	0.475	0.238	0.517	0.617	0.236	0.138	0.388	0.194	0.259	0.449	0.395

The PT method of IMERG data with ORG data at Kototabang has a b_m parameter value ranging from 1 to 2 for both hourly and daily data. This leads to an increase in rainfall values over 1 mm and a decrease in rainfall values below 1 mm. While the daily and period data b_m values are within the same range, the monthly b_m values exhibit different patterns. This indicates that the difference in coefficient of variation (CV) between IMERG and ORG data at Kototabang varies on a monthly basis. The disparity can be attributed to the varying extreme values in hourly and daily scales. Extreme precipitation plays a significant role in the difference of mean and standard deviation values [44], which in turn affects the CV values. Hence, it is necessary to separately determine the exponential parameters for hourly and daily data in mountainous regions.

After defining the values of P_{th} in the LOCI method and b_m in the PT method, the values of the scaling parameters can be calculated as shown in **Table 1**. In general, the values of the scaling parameters S_m and c show the same tendency, but the value of the scaling parameter a shows a different pattern. The similarity of S_m and c scaling parameter values is related to LOCI scaling methods, which only reduce low-intensity rainfall, so the average ratio generally does not change. In contrast, the PT method first reduced the variance of the IMERG data before calculating the scaling factor. The reduction of the variance value can be seen from the value of the scaling parameter *a*, which is smaller than the value of S_m. However, the value of the scaling factor *a* that is less than 1 shows that the IMERG data are still overestimated compared to ORG after correcting the variance value with increasing values by the value of b_m that is greater than 1. In contrast, the value of the scaling parameter *c* tends to exceed 1, showing that the IMERG value is smaller than ORG. This tendency is because the number of false events detected by IMERG is still high despite correcting the rain definition with LOCI.

Transfer function of CDF matching

In bias correction with the CDF method, several orders of polynomial fitting were performed on IMERG hourly and daily data with ORG data at Kototabang. Several polynomial fitting orders were tried to get the best order for bias correction of IMERG data in the mountainous region of Sumatra. Several previous studies have used various polynomial orders in the correction of SPPs data [45-47]. Therefore, it is necessary to determine the optimal polynomial order before performing the CDF fitting.

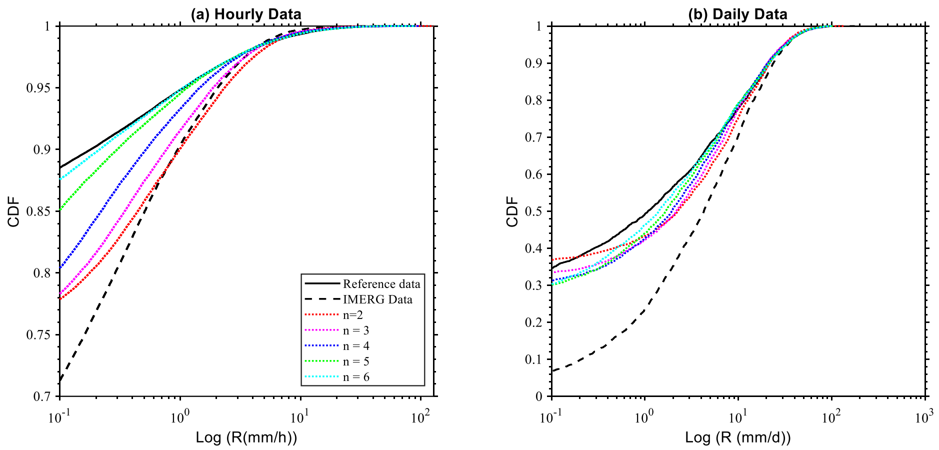

Figure 4 Cumulative distribution function (CDF) of ORG, IMERG, and the results of the fitting of both data in several polynomial orders for the hourly scale (a) and the daily scale (b).

Table 2 Transfer function from the IMERG data to the ORG data for polynomials of order 3 to 6.

Degree of polynomial	**Transfer function**
Hourly data	
$n = 2$	$5.6\ 10^{-3}\ x^2 + 0.2157\ x - 0.1740$
$n = 3$	$6.146\ 10^{-4}\ x^3 + 0.0484\ x^2 - 0.1228\ x - 0.1112$
$n = 4$	$3.208\ 10^{-5}\ x^4 - 0.0041\ x^3 + 0.1336\ x^2 - 0.5184\ x - 0.0577$
$n = 5$	$-1.376\ 10^{-6}\ x^5 + 2.296\ 10^{-3}\ x^4 - 0.0126\ x^3\ 0.2523\ x^2 - 0.8961\ x - 0.0184$
$n = 6$	$2.936\ 10^{-8}\ x^6 - 6.636\ 10^{-4}\ x^5 + 5.521\ 10^{-3}\ x^4 - 0.0206\ x^3 + 0.3240\ x^2 - 1.0611\ x - 0.0052$
Daily data	
$n = 2$	$-1.5\ 10^{-3}\ x^2 + 0.0381\ x - 2.1674$
$n = 3$	$-4.350\ 10^{-5}\ x^3 + 6.193\ 10^{-2}\ x^2 - 0.1995\ x - 1.3662$
$n = 4$	$6.906\ 10^{-7}\ x^4 - 2.280\ 10^{-3}\ x^3 + 0.0187\ x^2 - 0.4203\ x - 0.8674$
$n = 5$	$-1.612\ 10^{-8}\ x^5 + 5.971\ 10^{-6}\ x^4 - 7.606\ 10^{-3}\ x^3 + 0.0381\ x^2 - 0.6367\ x - 0.5237$
$n = 6$	$3.755\ 10^{-10}\ x^6 - 1.633\ 10^{-7}\ x^5 + 2.598\ 10^{-5}\ x^4 - 1.921\ 10^{-2}\ x^3 + 0.0660\ x^2 - 0.8605\ x - 0.2547$

The results show that the polynomial order has a significant impact on the CDF method in the hourly data compared to the daily data (**Figure 4**). For hourly data, higher polynomial orders bring the CDF values of the IMERG data closer to the observed data. This finding is particularly important for bias correction in mountainous areas, while previous studies have mainly focused on daily IMERG data [45-47]. For the daily data, different polynomial orders show varying CDF values depending on the intensity of rainfall. Lower polynomial orders (2 and 3) perform better at lower daily rainfall intensities, while higher orders fit better at higher intensities. Therefore, this study chooses to use the 6th order polynomial for both time series and daily data, as accurate representation of high-intensity rainfall plays a larger role in reducing bias [48]. The transfer function of each polynomial order in the CDF fitting is shown in **Table 2**. These transfer functions are then used for the bias correction of the IMERG data in the evaluation period.

Evaluation of bias correction method

Bias correction is applied to the IMERG data during the evaluation period of the ORG data, then the corrected data are statistically evaluated based on the evaluation matrix. The statistical evaluation results of the corrected IMERG data are shown in **Table 3**.

Table 3 Statistics evaluation from different bias correction method against ORG observation during evaluation period.

	CC	RMSE	RB (%)	POD	FAR	CSI
Hourly data						
IMERG	0.265	2.002	14.4	0.750	0.696	0.276
IMERG_LS	0.265	1.949	−11.6	0.735	0.681	0.286
IMERG_LOCI	0.263	2.053	16.3	0.591	0.563	0.335
IMERG_PT	0.174	2.305	−17.7	0.616	0.594	0.324
IMERG_CDF	0.215	2.180	−20.4	0.530	0.519	0.337
Daily data						
IMERG	0.485	12.357	17.1	0.989	0.288	0.706
IMERG_LS	0.485	11.610	−9.8	0.986	0.283	0.709
IMERG_LOCI	0.482	12.256	13.8	0.954	0.244	0.729
IMERG_PT	0.435	12.981	−11.1	0.953	0.246	0.727
IMERG_CDF	0.468	12.444	−12.8	0.883	0.179	0.740

According to a study by Yusnaini *et al.* [20], the statistical test value of IMERG data without bias correction during the evaluation period shows better results compared to the entire observation range. This improvement is observed in the increased of CC, POD, and CSI values, both on an hourly and daily basis, compared to previous studies. In addition, the decreased of the RMSE, RB, and FAR also indicate better accuracy of IMERG data during the evaluation period. The reason for this improved accuracy is attributed to the use of different satellite constellations in the Tropical Rainfall Measuring Mission (TRMM)-Era and GPM-Era [49]. The IMERG data utilizes TRMM observations before the launch of the GPM satellite in 2014 to intercalibrate passive microwave observations from geosynchronous satellites. The advancements in sensors and algorithms on the GPM satellite have enhanced the accuracy of the IMERG data in the GPM-Era compared to the TRMM-Era [50]. Consequently, the evaluation period data, which mainly covers the GPM-Era, exhibits more accurate results.

In the hourly resolution, the bias correction did not show any significant improvement from the 3 accuracy tests conducted on all evaluation period data (**Table 3**). It shows that the LS and LOCI methods had no significant improvement in the CC and RMSE values compared to the original IMERG data. The PT and CDF methods, however, slightly decreased the CC value and slightly increased the RMSE value. The change in CC values in the PT and CDF methods is because both methods make adjustments that are not only multiplied by the scale factor. The PT method impacted the distribution of data below and above 1, while the CDF method altered the rainfall value using a transfer function. Furthermore, the RB value showed that the LOCI method overestimated the data, while the other methods underestimated it. This is due to the scaling factor value greater than 1 in the LOCI method. In terms of the CDF method, the transfer function of order 6 tends to increase the percentage of light rainfall data and decrease the number of higher-intensity rainfall events (**Figure 4**). Overall, the LS method was the most successful in reducing the RMSE and RB values compared to the original IMERG data. Therefore, the LS method is recommended for improving the accuracy of hourly IMERG data in Kototabang. In general, however, bias correction of hourly IMERG data has not resulted in a satisfactory improvement in accuracy. The difficulty of bias correction of IMERG data in hourly resolution is related to the high spatial variability of precipitation in the mountains at short time resolutions. Rainfall of such short duration is often localized and cannot be detected by satellites. Therefore, other bias correction methods involving other weather parameters besides rainfall data are needed.

Contrary to accuracy test, the detection test of hourly IMERG data shows an increase in CSI values after calibration, indicating improved accuracy in detecting rainfall events. This improvement is supported by a decrease in the FAR value, although the POD value shows a slight decrease as well, indicating that some rain events may have been missed after correction. Among the methods used, the CDF and LOCI methods show the largest increase in CSI value and decrease in FAR value, suggesting that these methods are particularly effective in detecting low intensity rain errors in mountainous regions like Sumatra. Similar light rain detection errors have been found in other mountainous regions as well [51,52]. Therefore, using the CDF and LOCI methods can potentially enhance the detection capability of hourly IMERG data in mountainous areas, which is crucial for studying the diurnal pattern of precipitation frequency in Indonesian mountainous regions [18,22,53].

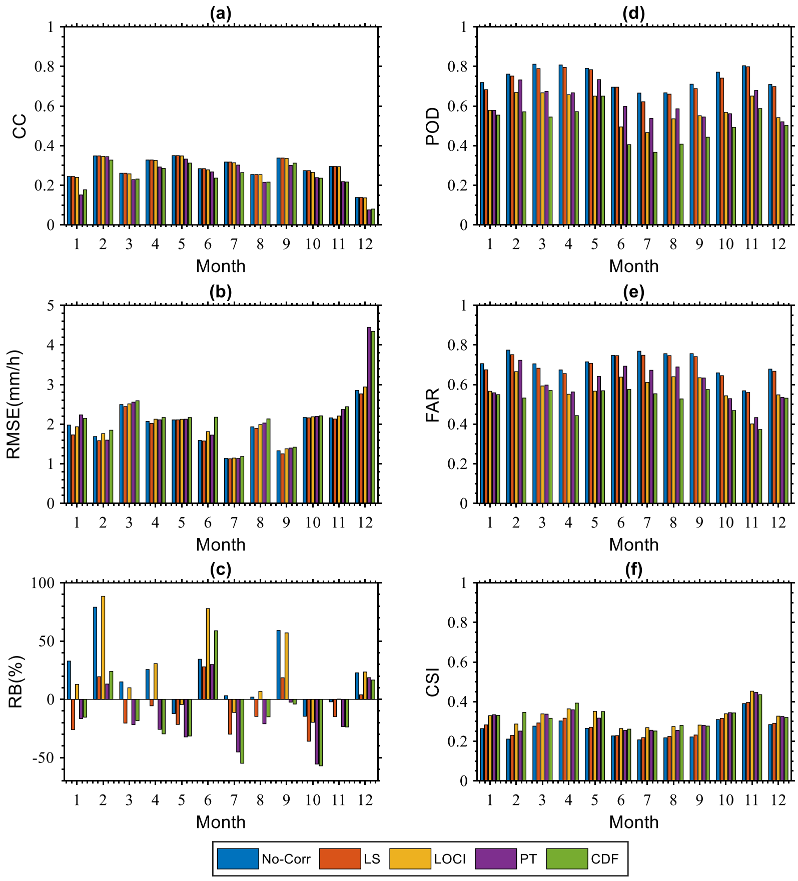

Figure 5 Monthly values of CC (a), RMSE (b), RB (c), POD (d), FAR (e), and CSI (f) from hourly IMERG data against ORG observations during the evaluation period.

The statistical test results of the corrected of hourly IMERG data on a monthly basis show a pattern similar to the overall data (**Figure 5**). The LS method performs best in improving accuracy, while the LOCI and CDF methods show the best performance in detecting capability improvement each month. However, an interesting observation is the significant increase in RMSE value in PT and CDF methods in December. The dominant increase in RMSE value in December is likely the main cause of the overall increase in RMSE of the PT and CDF methods (**Table 3**). This increase is likely due to the peak rainfall and occurrence of extreme rain in Kototabang [26,54], which leads to a higher variance in the data. The weak ability of

IMERG data to observe rain in the rainy season has also been observed in several previous studies in Indonesia [19,24]. Additionally, the CC value in December is the lowest, indicating reduced accuracy of IMERG data during the wet season. This decline in accuracy during the wet season supports previous research findings [20,24]. Consequently, it can be inferred that the PT and CDF methods tend to overestimate rainfall during the peak of the rainy season in mountainous regions.

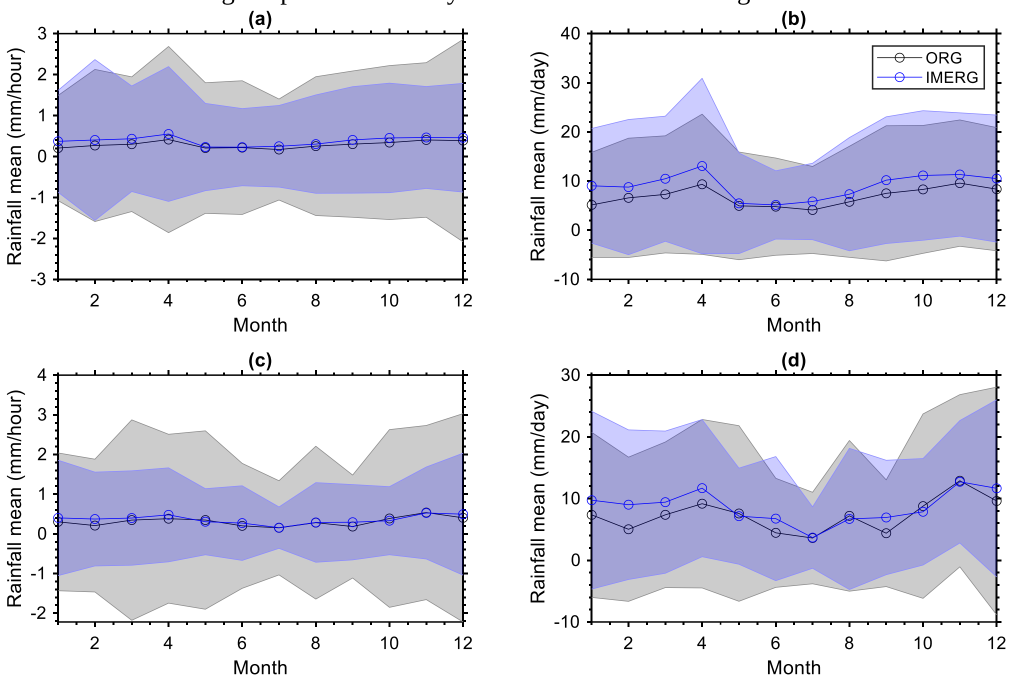

Figure 6 Monthly mean with shaded error bar values of hourly and daily rainfall from IMERG data against ORG observations during the calibration period (a,b) and evaluation period (c,d).

The monthly statistical test values reveal significant variations in the RB values of calibrated data and original IMERG data each month (**Figure 5(c)**). This is evident from the RB value of the original data, which fluctuates between positive and negative values. However, the RB value of the corrected data also exhibits irregular variations, indicating that the bias correction mechanism used during the calibration period has failed to address the bias in the evaluation data. Moreover, the average rainfall values and standard deviation of data during different calibration and evaluation periods highlight the difference in bias (**Figures 6(a) - 6(c)**). The calibration data consistently displays higher average rainfall values for IMERG data compared to ORG data, resulting in a positive RB value. However, during the evaluation data range, there are instances where the average rainfall value is lower than the ORG observation in a specific month. Additionally, the difference in standard deviation between the calibration and evaluation periods indicates distinct patterns in the data distribution. The inability of bias correction to mitigate the unsystematic bias in the highly variable monthly rainfall data in Sumatra's mountainous region can be attributed to the complex interactions of local factors, global oscillations, and regional oscillations [22,26,28,30].

Consistent with the time resolution data, the LS method shows the most satisfactory results in improving the accuracy test, while the LOCI and CDF methods show the most satisfactory results in improving the capability of detection test of the daily IMERG data (**Table 3**). However, all bias correction methods show success in reducing the RB value, indicating that the bias in the daily data is more systematic than in the hourly data. This more systematic bias is indicated by the more uniform variance value of ORG and IMERG daily data as shown by the distribution of standard deviation values of daily data (**Figures 6(b) - 6(d)**). However, the RB values of the monthly bias of the daily data also show diverse variations (**Figure 7**). The diversity of RB values that are not always positive is consistent with the distribution of RB in the hourly data. The non-uniformity of daily rainfall in Kototabang led to less satisfactory results in bias correction compared to other regions [55-57]. Nonetheless, the constructed bias correction method can be evaluated with other rain gauge observations in the Sumatran mountains, where data is incomplete and observational duration is short [58]. Evaluation in other observation areas will further enhance the effectiveness of the developed method in correcting IMERG data bias in mountainous regions.

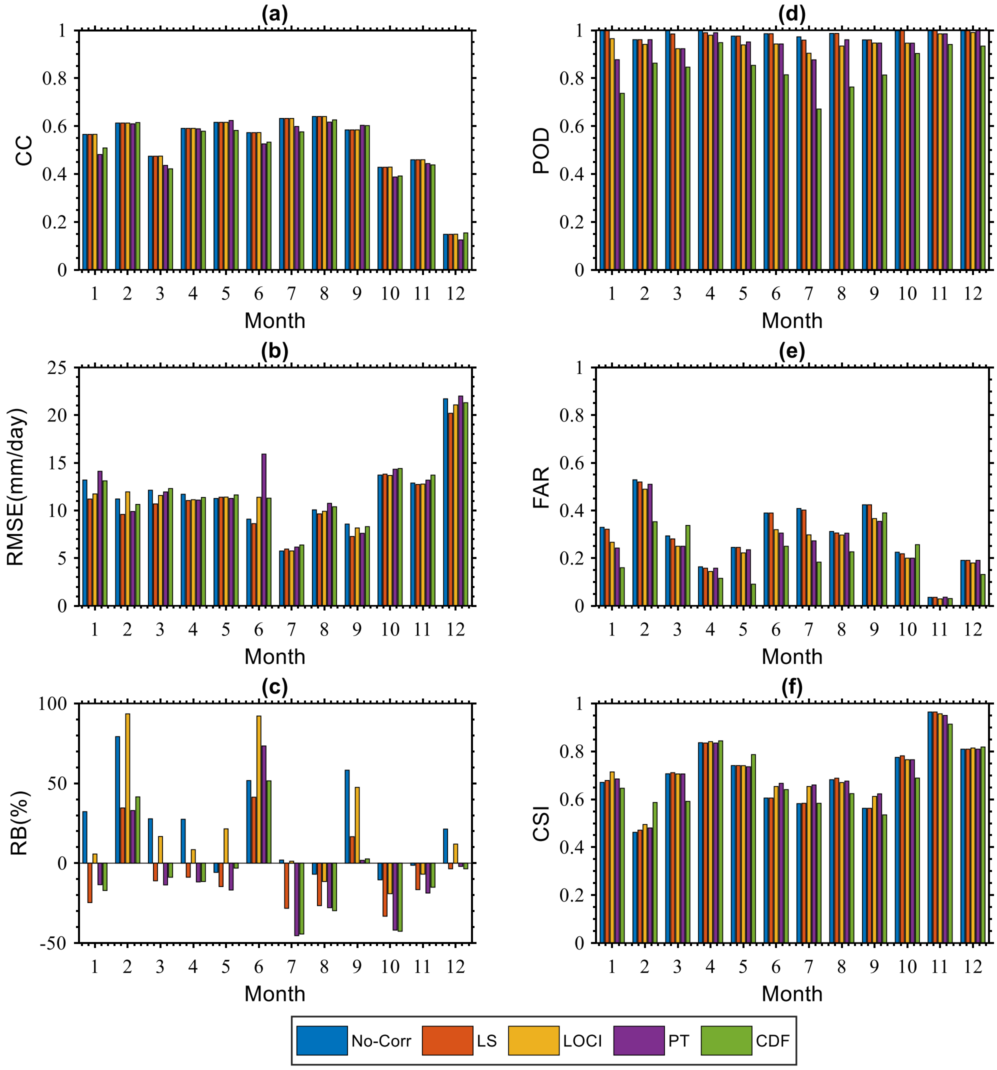

Figure 7 Monthly values of CC (a), RMSE (b), RB (c), POD (d), FAR (e), and CSI (f) from daily IMERG data against ORG observations during the evaluation period.

The bias correction capability was also evaluated in relation to the ability of the corrected IMERG data to calculate the frequency of daily rainfall occurrences with multiple rainfall thresholds. The determination of the number of occurrences of rainy days is very important for the determination of the extreme rainfall index [42]. The detection capability of corrected and original IMERG data on a daily scale with multiple rainfall thresholds is shown in **Figure 8**. In general, the POD and CSI values of all data, both corrected and original IMERG, show a decrease with increasing FAR. This indicates that the determination of the number of daily rainfall events decreases as the rainfall intensity threshold increases. The best CSI value of the IMERG daily rainfall data at Kototabang (CSI ≥ 0.5) is in the range of 1 - 5 mm/day. This is consistent with that found in southern China based on extreme rainfall with a recurrence interval of 60 years [59]. In addition, the relatively large RMSE value of the daily data at Kototabang (~12 mm/day) makes it difficult to determine the specific daily rainfall intensity. This leads to inaccuracies of IMERG data in the determination of some extreme precipitation indices involving high precipitation intensity, such as number of days with rainfall ≥ 50 mm/day (R50mm) and daily maximum rainfall (RX1day) indices [21,23,60]. However, previous research shows that the IMERG data at Kototabang can still capture the probability density function (PDF) of rainfall well [21]. In addition, the tendency of IMERG to underestimate high-intensity rainfall may also be the reason for the decrease in CSI values at such high rainfall intensities [61-63].

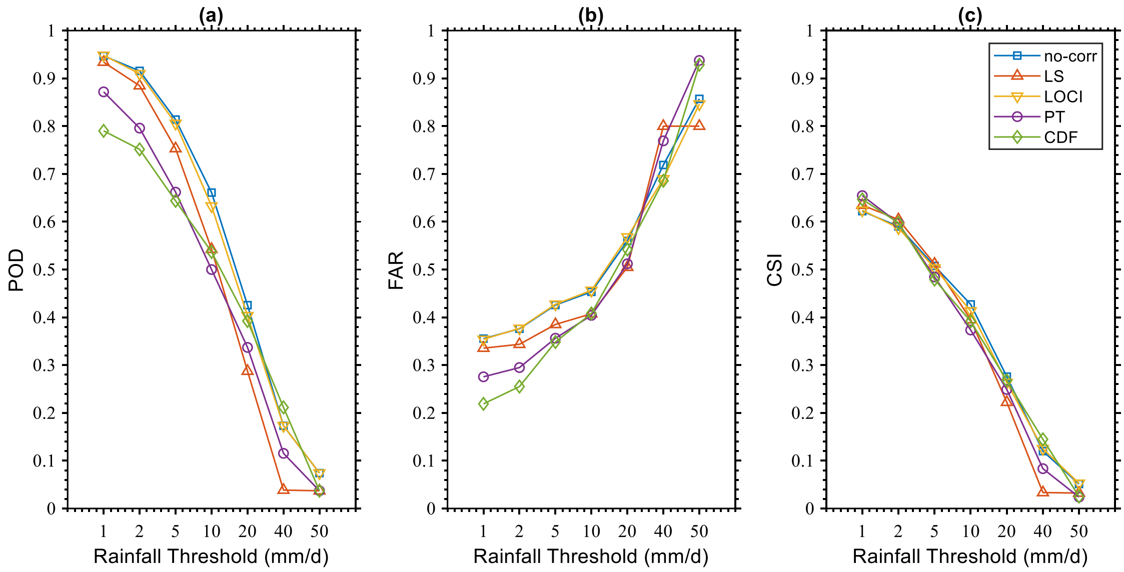

Figure 8 The values of POD (a), FAR (b), and CSI (c) from daily non-corrected and corrected IMERG data against ORG observations during the evaluation period.

Although all corrected and original IMERG data show the same detection trend, each method shows different POD, FAR and CSI values for each threshold (**Figure 8**). All methods show better CSI values at a threshold of 1 mm/day compared to the original IMERG data. This indicates that all methods used can improve the accuracy of IMERG data in observing the number of rainy days (R1mm), consecutive wet days (CWD), and consecutive dry days (CDD) indices. These 3 indices are very important in observing extreme rainfall indices, which are closely related to hydrometeorological disasters in Indonesia [2,64,65]. Furthermore, the LOCI method shows a trend that does not change the value of the 3 detection tests compared to the original IMERG data from a threshold of 1 - 50 mm/h. This indicates that the high CSI value in the LOCI method is only caused by the reduction of FAR at intensities below 0.5 mm. On the other hand, the CDF method shows a lower FAR value compared to the original IMERG < 50 mm/day. However, the low FAR value of the CDF method is followed by a lower POD value compared to the original IMERG data, so the CSI value of each threshold is not very satisfactory. Nevertheless, the CSI values of the LOCI and CDF methods show the best value at the threshold of 40 mm/day, which is often the threshold of extreme rainfall in Indonesia [66,67]. Thus, the LOCI and CDF methods remain a recommendation for improving the accuracy of extreme rainfall index observations because they can improve the accuracy of identifying rainy days (R ≥ 1 mm/day) and extremes (R ≥ 40 mm/day).

Conclusions

The bias-corrected IMERG data at Kototabang based on long-term observation of the optical rain gauge (ORG) shows improved accuracy and detection capability. Improvements depend on the temporal resolution of the data and the method used. In general, daily data is better corrected than hourly data. The best accuracy improvement is observed in the Linear Scaling (LS) method, as indicated by a decrease in the Root Mean Square Error (RMSE) and the Relative Bias (RB) compared to the original IMERG data. However, the accuracy improvement is not clearly observed in the bias correction of IMERG hourly data. On the other hand, the local intensity (LOCI), power transformation (PT), and cumulative distribution function (CDF) methods show a good ability to improve the rain detection capability of IMERG. The improved detection capability of these 3 methods is observed from the better critical succession index (CSI) values compared to the original IMERG data. The success of the 3 methods in improving the rain detection capability is due to the reduction of the false alarm ratio (FAR) values for rain with intensity < 2 mm. The improvement in rain detection of the IMERG data is also relatively satisfactory for the hourly data at Kototabang, especially for the LOCI and CDF methods. In addition, the CDF method also shows more satisfactory results for rainfall detection with thresholds ≥ 1 and ≥ 40 mm/day, so it is highly recommended to be applied to IMERG data before performing extreme rainfall analysis. Therefore, combining the LS, LOCI, and CDF methods on IMERG hourly and daily data will improve weather and climate analysis in the mountainous region of Sumatra. Although this study was conducted for one observation in Kototabang only, this correction method can also be applied in other mountainous regions. However, future research is needed to test the accuracy of the bias correction method on IMERG data in mountainous areas with multiple rain gauge observations in one IMERG data grid.

Acknowledgements

This study was supported by Penelitian Disertasi Doktor (PDD) 2023 Grants from Ministry of Education and Culture Nr. 122/E5/PG.02.00.PL/2023 (Contract number: 3108/UN1/DITLIT/DitLit/PT.01.03/2023). The Optical Rain Gauge (ORG) was operated and maintained by Kyoto University in Japan, Shimane University in Japan, under collaboration with National Research and Innovation Agency (BRIN), Indonesia. We also thank National Aeronautics and Space Administration (NASA), United States and the Japan Aerospace Exploration Agency (JAXA) which has provided IMERG final run product version 06 (V06).

References

[1] GJ Huffman, DT Bolvin, D Braithwaite, K Hsu, R Joyce, C Kidd, EJ Nelkin and P Xie. *Algorithm theoretical basis document (ATBD) version 4.5*. NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). National Aeronautics and Space Administration, Washington DC, 2015.

[2] M Marzuki, R Ramadhan, H Yusnaini, M Vonnisa, R Safitri and E Yanfatriani. Changes in extreme rainfall in new Capital of Indonesia (IKN) based on 20 years of GPM-IMERG data. *Trends Sci.* 2023; **20**, 6935.

[3] Y Zhang, R Li and K Wang. Climatology and changes in internal intensity distributions of global precipitation systems over 2001-2020 based on IMERG. *J. Hydrol.* 2023; **620**, 129386.

[4] T Caloiero, GN Caroletti and R Coscarelli. IMERG-based meteorological drought analysis over Italy. *Climate* 2021; **9**, 65.

[5] Z Wang, R Zhong, C Lai and J Chen. Evaluation of the GPM IMERG satellite-based precipitation products and the hydrological utility. *Atmos. Res.* 2017; **196**, 151-63.

[6] RK Pradhan, Y Markonis, MRV Godoy, A Villalba-Pradas, KM Andreadis, EI Nikolopoulos, SM Papalexiou, A Rahim, FJ Tapiador and M Hanel. Review of GPM IMERG performance: A global perspective. *Rem. Sens. Environ.* 2022; **268**, 112754.

[7] C Massari and V Maggioni. *Error and uncertainty characterization*. *In*: V Levizzani, C Kidd, D Kirschbaum, C Kummerow, K Nakamura and F Turk (Eds.). Satellite Precipitation Measurement. Switzerland, 2020.

[8] V Maggioni, C Massari and C Kidd. Chapter 13 - Errors and uncertainties associated with quasiglobal satellite precipitation products. *Precipitation Sci.* 2022; **2022**, 377-90.

[9] L Zhou, T Koike, K Takeuchi M Rasmy, K Onuma, H Ito, H Selvarajah, L Liu, X Li and T Ao. A study on availability of ground observations and its impacts on bias correction of satellite precipitation products and hydrologic simulation efficiency. *J. Hydrol.* 2022; **610**, 127595.

[10] S Xiao, L Zou, J Xia, Z Yang and T Yao. Bias correction framework for satellite precipitation products using a rain/no rain discriminative model. *Sci. Total Environ.* 2022; **818**, 151679.

[11] L Wei, S Jiang, L Ren, L Zhang, M Wang, Y Liu and Z Duan. Bias correction of GPM IMERG early run daily precipitation product using near real-time CPC global measurements. *Atmos. Res.* 2022; **279**, 106403.

[12] C Lober, J Fayne, H Hashemi and LC Smith. Bias correction of 20 years of IMERG satellite precipitation data over Canada and Alaska. *J. Hydrol. Reg. Stud.* 2023; **47**, 101386.

[13] R Li, S Tang, Z Shi, J He, W Shi and X Li. Spatial patterns of errors in GPM IMERG summer precipitation estimates and their connections to geographical features in complex topographical Area. *Rem. Sens.* 2022; **14**, 4789.

[14] X Lu, G Tang, X Wang, Y Liu, L Jia, G Xie, S Li and Y Zhang. Correcting GPM IMERG precipitation data over the Tianshan Mountains in China. *J. Hydrol.* 2019; **575**, 1239-52.

[15] H Li, Y Zhang, H Lei and X Hao. Machine learning-based bias correction of precipitation measurements at high altitude. *Rem. Sens.* 2023; **15**, 2180.

[16] PDA Borges, J Franke, YMTD Anunciação, H Weiss and C Bernhofer. Comparison of spatial interpolation methods for the estimation of precipitation distribution in Distrito Federal, Brazil. *Theor. Appl. Climatol.* 2016; **123**, 335-48.

[17] S Mori, H Jun-Ichi, YI Tauhid, MD Yamanaka, N Okamoto, F Murata, N Sakurai, H Hashiguchi and

T Sribimawati. Diurnal Land–Sea rainfall peak migration over Sumatera Island, Indonesian maritime continent, observed by TRMM satellite and intensive Rawinsonde soundings. *Mon. Weather Rev.* 2004; **132**, 2021-39.

[18] M Marzuki, H Yusnaini, F Tangang, R Muharsyah, M Vonnisa and H Harmadi. Land - sea contrast of diurnal cycle characteristics and rain event propagations over Sumatra according to different rain duration and seasons. *Atmos. Res.* 2022; **270**, 106051.

[19] CY Liu, P Aryastana, GR Liu and WR Huang. Assessment of satellite precipitation product estimates over Bali Island. *Atmos. Res.* 2020; **244**, 105032.

[20] H Yusnaini, R Ramadhan, M Marzuki, AP Ningsih, H Hashiguchi, T Shimomai, M Vonnisa, H Harmadi, W Suryanto and S Sholihun. Statistical comparison of IMERG precipitation products with optical rain gauge observations over Kototabang, Indonesia. *Jurnal Ilmu Fisika* 2022; **14**, 10-20.

[21] R Ramadhan, M Marzuki, H Yusnaini, AP Ningsih, H Hashiguchi, T Shimomai, M Vonnisa, S Ulfah, W Suryanto and S Sholihun. Ground validation of GPM IMERG-F precipitation products with the point rain gauge records on the extreme rainfall over a mountainous Area of Sumatra Island. *Jurnal Penelitian Pendidikan IPA* 2022; **8**, 163-70.

[22] M Marzuki, H Yusnaini, R Ramadhan, F Tangang, AAB Amirudin, H Hashiguchi, T Shimomai and M Vonnisa. Characteristics of precipitation diurnal cycle over a Mountainous Area of Sumatra Island including MJO and seasonal signatures based on the 15-year optical rain gauge data, WRF model and IMERG. *Atmosphere* 2021; **13**, 63.

[23] R Ramadhan, M Marzuki, H Yusnaini, R Muharsyah, W Suryanto, S Sholihun, M Vonnisa, A Battaglia and H Hashiguchi. Capability of GPM IMERG products for extreme precipitation analysis over the Indonesian maritime continent. *Rem. Sens.* 2022; **14**, 412.

[24] R Ramadhan, H Yusnaini, M Marzuki, R Muharsyah, W Suryanto, S Sholihun, M Vonnisa, H Harmadi, AP Ningsih and A Battaglia. Evaluation of GPM IMERG performance using gauge data over Indonesian maritime continent at different time scales. *Rem. Sens.* 2022; **14**, 1172.

[25] S Fukao, H Hashiguchi, M Yamamoto, T Tsuda, T Nakamura, MK Yamamoto, T Sato, M Hagio and Y Yabugaki. Equatorial atmosphere radar (EAR): System description and first results. *Radio Sci.* 2003; **38**, 1053.

[26] HH Marzuki, T Shimomai and WL Randeu. Cumulative distributions of rainfall rate over Sumatra. *Electromagn. Waves* 2016; **49**, 1-8.

[27] R Ramadhan, Marzuki, M Vonnisa, Harmadi, H Hashiguhci and T Shimomai. Seasonal variation in the vertical profile of the raindrop size distribution for stratiform rain as inferred from micro rain radar observations at Kototabang. *AIP Conf. Proc.* 2020; **2221**, 090002.

[28] T Kozu, KK Reddy, S Mori, M Thurai, JT Ong, DN Rao and T Shimomai. Seasonal and diurnal variations of raindrop size distribution in Asian Monsoon Region. *J. Meteorol. Soc. Jpn.* 2006; **84A**, 195-209.

[29] E Aldrian and RD Susanto. Identification of three dominant rainfall regions within Indonesia and their relationship to sea surface temperature. *Int. J. Climatol.* 2003; **23**, 1435-52.

[30] P Wu, JI Hamada, S Mori, YI Tauhid, MD Yamanaka and F Kimura. Diurnal variation of precipitable water over a Mountainous Area of Sumatra Island. *J. Appl. Meteorol. Climatol.* 2003; **42**, 1107-15.

[31] R Ramadhan, Marzuki, M Vonnisa, Harmadi, H Hashiguchi and T Shimomai. Diurnal variation in the vertical profile of the raindrop size distribution for stratiform rain as inferred from micro rain radar Observations in Sumatra. *Adv. Atmos. Sci.* 2020; **37**, 832-46.

[32] N Sakurai, F Murata, MD Yamanaka, S Mori, JI Hamada, H Hashiguchi, YI Tauhid, T Sribimawati and B Suhardi. Diurnal cycle of cloud system migration over Sumatera Island. *J. Meteorol. Soc. Jpn.* 2005; **83**, 835-50.

[33] Marzuki, H Hashiguchi, T Kozu, T Shimomai, Y Shibagaki and Y Takahashi. Precipitation microstructure in different Madden-Julian Oscillation phases over Sumatra. *Atmos. Res.* 2016; **168**, 121-38.

[34] C Teutschbein and J Seibert. Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *J. Hydrol.* 2012; **456-457**, 12-29.

[35] EZX Soo, WZW Jaafar, SH Lai, F Othman, A Elshafie, T Islam, P Srivastava and HSO Hadi. Precision of raw and bias-adjusted satellite precipitation estimations (TRMM, IMERG, CMORPH, and PERSIANN) over extreme flood events: Case study in Langat river basin, Malaysia. *J. Water Clim. Change* 2020; **11**, 322-42.

[36] G Lenderink, A Buishand and WV Deursen. Estimates of future discharges of the river Rhine using two scenario methodologies: Direct versus delta approach. *Hydrol. Earth Syst. Sci.* 2007; **11**, 1145-

59.
[37] J Schmidli, C Frei and PL Vidale. Downscaling from GCM precipitation: A benchmark for dynamical and statistical downscaling methods. *Int. J. Climatol.* 2006; **26**, 679-89.
[38] R Leander and TA Buishand. Resampling of regional climate model output for the simulation of extreme river flows. *J. Hydrol.* 2007; **332**, 487-96.
[39] AVM Ines and JW Hansen. Bias correction of daily GCM rainfall for crop simulation studies. *Agr. Forest Meteorol.* 2006; **138**, 44-53.
[40] DS Wilks. *Statistical methods in the atmospheric sciences*. Elsevier, Amsterdam, Netherlands, 2020.
[41] EE Ebert, JE Janowiak and C Kidd. Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Am. Meteorol. Soc.* 2007; **88**, 47-64.
[42] MG Donat, LV Alexander, H Yang, I Durre, R Vose, RJH Dunn, KM Willett, E Aguilar, M Brunet, J Caesar, B Hewitson, C Jack, AMGK Tank, AC Kruger, J Marengo, TC Peterson, M Renom, CO Rojas, M Rusticucci, J Salinger, ..., S Kitching. Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset. *J. Geophys. Res. Atmos.* 2013; **118**, 2085-454.
[43] X Zhang, L Alexander, GC Hegerl, P Jones, AK Tank, TC Peterson, B Trewin and FW Zwiers. Indices for monitoring changes in extremes based on daily temperature and precipitation data. *WIREs Clim. Change* 2011; **2**, 851-70.
[44] S Pfahl, PA O'Gorman and EM Fischer. Understanding the regional pattern of projected future changes in extreme precipitation. *Nat. Clim. Change* 2017; **7**, 423-7.
[45] M Gentilucci, M Barbieri and G Pambianchi. Reliability of the IMERG product through reference rain gauges in Central Italy. *Atmos. Res.* 2022; **278**, 106340.
[46] S Camici, L Ciabatta, C Massari and L Brocca. How reliable are satellite precipitation estimates for driving hydrological models: A verification study over the Mediterranean Area. *J. Hydrol.* 2018; **563**, 950-61.
[47] L Brocca, P Filippucci, S Hahn, L Ciabatta, C Massari, S Camici, L Schüller, B Bojkov and W Wagner. SM2RAIN–ASCAT (2007–2018): Global daily satellite rainfall data from ASCAT soil moisture observations. *Earth Syst. Sci. Data* 2019; **11**, 1583-601.
[48] PS Smitha, B Narasimhan, KP Sudheer and H Annamalai. An improved bias correction method of daily rainfall data using a sliding window technique for climate change impact assessment. *J. Hydrol.* 2018; **556**, 100-18.
[49] GJ Huffman. The transition in multi-satellite products from TRMM to GPM (TMPA to IMERG). TMPA-to-IMERG Transition. NASA, Washington DC, 2020.
[50] G Tang, Z Zeng, D Long, X Guo, B Yong, W Zhang and Y Hong. Statistical and hydrological comparisons between TRMM and GPM Level-3 products over a midlatitude basin: Is day-1 IMERG a good successor for TMPA 3B42V7? *J. Hydrometeorology* 2016; **17**, 121-37.
[51] Y Rojas, JR Minder, LS Campbell, A Massmann and R Garreaud. Assessment of GPM IMERG satellite precipitation estimation and its dependence on microphysical rain regimes over the mountains of South-Central Chile. *Atmos. Res.* 2021; **253**, 105454.
[52] H Wang, L Wang, J He, F Ge, Q Chen, S Tang and S Yao. Can the GPM IMERG hourly products replicate the variation in precipitation during the wet season over the Sichuan Basin, China? *Earth Space Sci.* 2020; **7**, e01090.
[53] R Ramadhan, H Yusnaini, M Marzuki, ZV Adha, M Vonnisa and R Muharsyah. *Diurnal rainfall pattern in Riau Islands as observed by Rain Gauge and IMERG data*. International Conference on Radioscience, Equatorial Atmospheric Science and Environment. Springer Nature Singapore, Gateway East, Singapore, 2022, p. 317-24.
[54] Marzuki, H Hashiguchi, M Vonnisa, Harmadi and Muzirwan. Long-term change in rainfall rate and melting layer height in Indonesia. *In*: Proceedings of the 2018 Progress in Electromagnetics Research Symposium (PIERS-Toyama), Toyama, Japan. 2018, p. 1154-8.
[55] W Sun, Y Sun, X Li, T Wang, Y Wang, Q Qiu and Z Deng. Evaluation and correction of GPM IMERG precipitation products over the capital circle in northeast china at multiple spatiotemporal scales. *Adv. Meteorol.* 2018; **2018**, 4714173.
[56] H Ouatiki, A Boudhar and A Chehbouni. Accuracy assessment and bias correction of remote sensing–based rainfall products over semiarid watersheds. *Theor. Appl. Climatol.* 2023, https://doi.org/10.1007/s00704-023-04586-y.
[57] F Yuan, L Zhang, K Win, L Ren, C Zhao, Y Zhu, S Jiang and Y Liu. Assessment of GPM and TRMM multi-satellite precipitation products in streamflow simulations in a data-sparse mountainous watershed in myanmar. *Rem. Sens.* 2017; **9**, 302.

[58] M Marzuki, K Suryanti, H Yusnaini, F Tangang, R Muharsyah, M Vonnisa and D Devianto. Diurnal variation of precipitation from the perspectives of precipitation amount, intensity and duration over sumatra from rain gauge observations. *Int. J. Climatol.* 2021; **41**, 4386-97.

[59] A Zhang, L Xiao, C Min, S Chen, M Kulie, C Huang and Z Liang. Evaluation of latest GPM-Era high-resolution satellite precipitation products during the May 2017 guangdong extreme rainfall event. *Atmos. Res.* 2019; **216**, 76-85.

[60] R Ramadhan, R Muharsyah, Marzuki, H Yusnaini, M Vonnisa, H Hashiguchi, W Suryanto and Sholihun. Evaluation of GPM IMERG products for extreme precipitation over Indonesia. *J. Phys. Conf. Series* 2022; **2309**, 012008.

[61] Y Xin, Y Yang, X Chen, X Yue, Y Liu and C Yin. Evaluation of IMERG and ERA5 precipitation products over the mongolian plateau. *Sci. Rep.* 2022; **12**, 21776.

[62] N Lu. Evaluation of IMERG precipitation products in the Southeast Costal Urban Region of China. *Rem. Sens.* 2022; **14**, 4947.

[63] Y Gao, X Zhong and F Gan. Inspection of IMERG precipitation estimates during typhoon cempaka using a new methodology for quantifying and evaluating bias. *J. Hydrol.* 2023; **620**, 129554.

[64] Supari, F Tangang, L Juneng and E Aldrian. Observed changes in extreme temperature and precipitation over Indonesia. *Int. J. Climatol.* 2017; **37**, 1979-97.

[65] R Ramadhan, M Marzuki, W Suryanto, S Sholihun, H Yusnaini, R Muharsyah and M Hanif. Trends in rainfall and hydrometeorological disasters in new capital city of Indonesia from long-term satellite-based precipitation products. *Rem. Sens. Appl. Soc. Environ.* 2022; **28**, 100827.

[66] BS Wiwoho, IS Astuti, IAG Alfarizi and HR Sucahyo. Validation of three daily satellite rainfall products in a humid tropic watershed, brantas, Indonesia: Implications to land characteristics and hydrological modelling. *Hydrology* 2021; **8**, 154.

[67] WK Cheong, B Timbal, N Golding, S Sirabaha, KF Kwan, TA Cinco, B Archevarahuprok, VH Vo, D Gunawan and S Han. Observed and modelled temperature and precipitation extremes over Southeast Asia from 1972 to 2010. *Int. J. Climatol.* 2018; **38**, 3013-27.