

Article

Comparison and Bias Correction of TMPA Precipitation Products over the Lower Part of Red–Thai Binh River Basin of Vietnam

Hung Manh Le ^{1,2,*} , Jessica R. P. Sutton ³, Duong Du Bui ², John D. Bolten ⁴ and Venkataraman Lakshmi ¹

¹ School of Earth, Ocean and the Environment, University of South Carolina, Columbia, SC 29208, USA; vlakshmi@geol.sc.edu

² National Center for Water Resources Planning and Investigation, Ministry of Natural Resources and Environmental, Hanoi 100000, Vietnam; duongdubui@gmail.com

³ Berry College, Mount Berry, GA 30149, USA; jsutton@berry.edu

⁴ Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA; john.bolten@nasa.gov

* Correspondence: hmle@email.sc.edu; Tel.: +1-702-861-4265

Received: 5 September 2018; Accepted: 24 September 2018; Published: 1 October 2018



Abstract: As the limitation of rainfall collection by ground measurement has been widely recognized, satellite-based rainfall estimate is a promising high-resolution alternative in both time and space. This study is aimed at exploring the capacity of the satellite-based rainfall product Tropical Rainfall Measurement Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA), including 3B42V7 research data and its real-time 3B42RT data, by comparing them against data from 29 ground observation stations over the lower part of the Red–Thai Binh River Basin from March 2000 to December 2016. Various statistical metrics were applied to evaluate the TMPA products. The results showed that both 3B42V7 and 3B42RT had weak relationships with daily observations, but 3B42V7 data had strong agreement on the monthly scale compared to 3B42RT. Seasonal analysis showed that 3B42V7 and 3B42RT underestimated rainfall during the dry season and overestimated rainfall during the wet season, with high bias observed for 3B42RT. In addition, detection metrics demonstrated that TMPA products could detect rainfall events in the wet season much better than in the dry season. When rainfall intensity was analyzed, both 3B42V7 and 3B42RT overestimated the no rainfall event during the dry season but underestimated these events during the wet season. Finally, based on the moderate correlation between climatology–topography characteristics and correction factors of linear-scaling (LS) approach, a set of multiple linear models was developed to reduce the error between TMPA products and the observations. The results showed that climatology–topography-based linear-scaling approach (CTLS) significantly reduced the percentage bias (PBIAS) score and moderately improved the Nash–Sutcliffe efficiency (NSE) score. The finding of this paper gives an overview of the capacity of TMPA products in the lower part of the Red–Thai Binh River Basin regarding water resource applications and provides a simple bias correction that can be used to improve the correctness of TMPA products.

Keywords: Red–Thai Binh River Basin; TMPA 3B42V7; TMPA 3B42RT; rainfall; bias correction; linear-scaling approach; climatology; topography

1. Introduction

Precipitation is the most crucial input variable enforced in water prediction models. Reliable precipitation is required for model calibration, forecast, and simulation [1–3]. Gauge observation is

the primary collection approach to obtain precipitation information [4]. However, gauge network is often sparse and nonexistent in many parts of the globe [5,6]. Moreover, it is often challenging to obtain gauge data, especially in developing countries and transboundary rivers, due to technical and administrative reasons [7–9]. In addition, gauge observations only provide point measurements of precipitation and cannot capture the full spatial variability. Space-based precipitation estimations, therefore, have great potential application to enhance the capacity of measuring this vital water cycle component [10,11].

Several satellite-derived datasets have been used in previous studies, such as the Tropical Rainfall Measurement Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) [12], the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) [13], the Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) [14], and National Oceanic and Atmospheric Administration/Climate Prediction Centre (NOAA/CPC) morphing technique (CMORPH) [15] products. Among them, TPMA—the first space-borne product of the Earth Science Mission aimed at studying tropical and subtropical rainfall—has performed well in a wide range of applications, such as hydrological modeling [16–18], drought monitoring [19,20], and agronomy [21,22]. TMPA products have also been evaluated as having better performance than other satellite-based rainfall products. For example, the TMPA 3B42V7 data is generally a better input in a distributed hydrological model compared to CMORPH and TMPA 3B42RT (real time) for multiple hydrological purposes, including annual water budgeting, monthly and daily streamflow simulation, and extreme flood modeling [23]. Similarly, Tong et al. [24] showed that 3B42V7 was a better driving force of hydrological model for both monthly and daily streamflow simulation over the Tibetan Plateau compared to CMORPH, PERSIANN, and 3B42RT. Moazami et al. used six statistical indices and contingency table to evaluate 3B42V7, concluding it was a better estimation of daily precipitation than PERSIAN and 3B42RT over Iran [25]. Simons et al. [26] identified that monthly TMPA 3B43 rainfall product was the most suitable satellite dataset compared to CHIRPS and CMORPH over the Red River Basin of Vietnam.

Differences between TMPA products and rain gauge observation analysis have been a cause of concern recently. Zad et al. [27] pointed out that 3B42V7 tended to overestimate rainfall measurement by approximately 26.95% at Pahang River Basin of Malaysia and that 3B42V7 was likely to have a high accuracy of detecting rainfall events at high-altitude and mid-altitude areas compared to low-altitude regions. Kneis et al. [28] analyzed that 3B42V7 and 3B42RT datasets were moderately correlated with their gauged-based counterpart at sub-basin level (4000 to 16,000 km²) at the lower Mahanadi River Basin of India but that the 3B42V7 and 3B42RT data often do not reflect gauge observation at high-intensity level (>80 mm/day). The TMPA product is also likely to perform better on a monthly scale when compared to the ground data. Curtarelli et al. [29] found that monthly 3B43 dataset had a great consistency (correlation coefficient >0.97) with ground observation data over the Itumbiara Reservoir drainage area in Central Brazil but that 3B43 tended to overestimate rainfall by 1.24%. Comparing monthly 3B43 dataset with 56 observations in Yangtze River Delta, Cao et al. [30] also showed an inclination of 3B43 to overestimate monthly rainfall, with the bias ranging between –10% and 10% most of the study area; its correlation coefficient with observation was found to peak in March (0.96) and reach bottom in August (0.79). Although the TRMM satellite has not been operated since 2014, TMPA products are still being generated regardless [31].

Following the highly successful TMPA, the Global Precipitation Measurement (GPM) mission was developed to continuously increase precipitation estimation over most of the globe [32]. A range of studies in many regions have demonstrated that GPM outperforms TMPA by having a better spatial resolution, coverage area, and lower systematic bias error [33–35]. However, GPM has only been available for a short time (since 2014), while TMPA products date back to January 1998. In addition, GPM is just a slight improvement over TMPA products [36]. Huffman et al. [32] aim to extend the GPM data to the same length as the longest TMPA data. Therefore, assessments on TMPA products

are of paramount importance to gain insights into their performance at various regions so that their algorithms can be improved and the next generation GPMs can be developed.

While there is a clear advantage of having a high temporal and spatial resolution using TMPA products, extra work is required because bias correction needs to be performed prior to application of any TMPA products in environmental, water resources, and ecological studies [27]. Climatology and topography are likely factors to induce errors in remote sensing retrievals [37]. Consequently, their effects on the quality of TMPA products are inevitable. Based on the moderate inverse linear relationship between the monthly 3B43 bias and elevation, Hashemi et al. [38] developed a linear model between 3B43 bias and elevation, especially for stations that have elevations above 1500 m above mean sea level in the U.S. The corrected monthly 3B43 product showed a significant improvement in the high elevation area. Thus, the empirical bias correction model using climatology and topography seems to be a potential investigation direction, although relatively little research has been conducted so far.

In Vietnam, ground observations provide poor spatial and temporal measurement of rainfall due to the lack of a dense network for rain gauge measurement. The average rain gauge network in Vietnam is around 400 km² per rain gauge, which is below the World Meteorological Organization standard (area per rainfall station of 100–250 km² for mountainous areas; area per rainfall station of 600–900 km² for lowland areas) [39]. Moreover, the rain gauge distribution in Vietnam is uneven, with insufficient gauged stations at high elevation areas. According to the Vietnam Meteorological and Hydrological Administration, most rain gauge stations (75%) are concentrated at low elevation areas (<200 m), which only cover half of Vietnam's land [40]. With these perspectives, satellite-based precipitation is an indispensable alternative source of precipitation data for Vietnam. Preliminary studies on satellite-based precipitation products in the country have been conducted recently. However, these studies either focused on monthly rainfall [26,41] or used directly satellite-derived rainfall without bias correction analysis [42]. Therefore, further research on satellite-based precipitation products is still of fundamental importance for the country.

This study selected the Red–Thai Binh River Basin—one of the largest river systems in Vietnam—as a case study. Although it plays an essential role in Vietnam's economic and social development, many parts of this basin do not have rainfall monitoring from ground, causing difficulties for basin rainfall estimation and water resources management. The first objective of this study was to compare the TMPA products 3B42V7 and 3B42RT with ground observation data over Red–Thai Binh River Basin in various aspects, such as calculating error statistics on a daily scale, monthly scale, dry and wet seasons, detecting rainfall events ability, and evaluating rainfall intensity. The second objective was to develop a linear-scaling bias correction model using climate–topography indices for both 3B42V7 and 3B42RT datasets. The results of the assessment and bias correction of TMPA precipitation products could help in supporting its potential application in hydrological modeling and drought monitoring in the studied region.

2. Materials

2.1. Study Area

The Red–Thai Binh River Basin is a transboundary river that flows through three countries—Vietnam, China, and Laos—with a total area of 169,000 km² (Figure 1). The area of this in Vietnam is 88,680 km², which makes up 51.3% of the total area. In this study, due to the lack of observation data, description of water resource characteristics and evaluation results of TMPA 3B42V7 and TMPA 3B42RT data only focused on the Vietnamese part of the basin. There are two primary river systems in the Red–Thai Binh River. The Red River system originates in China and flows into Vietnam through three main tributaries—Da, Lo, and the Thao River—while the Thai Binh River system is entirely located in Vietnam. The Red–Thai Binh River belongs to a tropical climate with two distinct seasons: the wet season and the dry season. The total annual rainfall is approximately 1700 mm, with high rainfall amounts (>2000 mm) observed in the mountainous areas between the

Vietnam and China border. The annual total flow of the Red–Thai Binh River is 131.4 billion m³—the Chinese territory part generates 48.3 billion m³, while the rest 83.1 billion m³ is generated in the Vietnamese side [43]. As the second largest river system in Vietnam, the Red–Thai Binh River is home to 29.1 million Vietnamese (2015 figure), making up for 22.6% of Vietnam’s GDP (2010 figure) (General Statistics of Vietnam) [44].

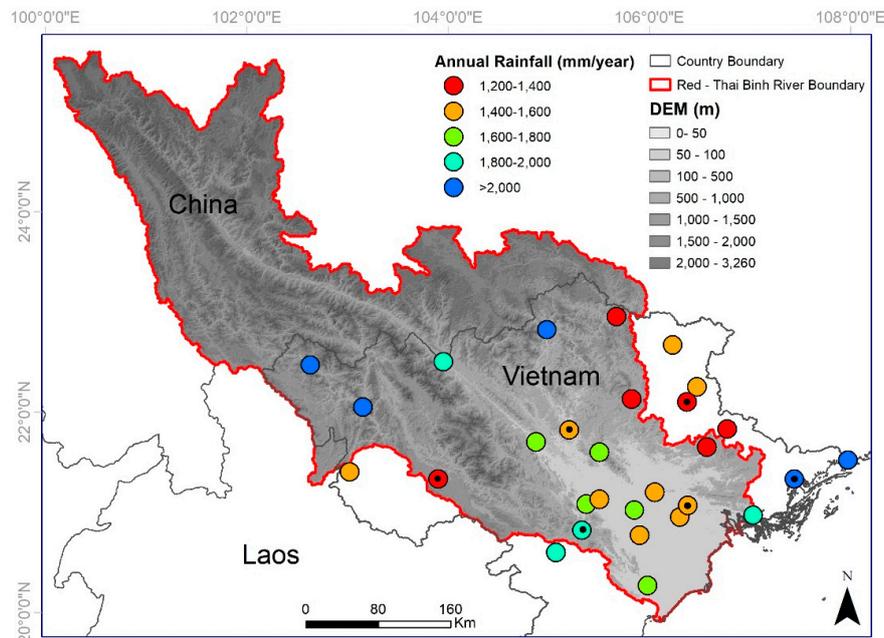


Figure 1. Overview of Red–Thai Binh River Basin. The stations with black dots at the middle were used for calibration climatology–topography-based linear-scaling approach.

2.2. Data

2.2.1. Observation Data

Rainfall measurements from a total of 29 daily rainfall stations (March 2000 to December 2016) within or neighboring the basin were collected from the Vietnam Meteorological and Hydrological Administration. The distribution of rainfall stations is presented in Figure 1, and their characteristics can be found in Table 1. The stations were selected due to their reliable data and low missing values (5–10%).

Table 1. Rainfall station descriptions for the ground observation stations over Red–Thai Binh River Basin (March 2000–December 2016).

No.	Station Name	Long. (°)	Lat. (°)	Elev. (m)	Annual Rainfall (AR) (mm/year)	Standard Deviation of Rainfall (SDR) (mm/day)	No. of Rain Days (NRD) (day)
1	Baccan	105.82	22.13	241	1389	11.29	250
2	Bacninh	106.05	21.20	8	1537	13.27	248
3	Baolac	105.67	22.95	348	1201	9.91	263
4	Caobang	106.23	22.67	244	1417	11.5	251
5	Dienbien	103.02	21.40	487	1535	11.75	248
6	Hagiang	104.98	22.82	117	2333	17.15	222
7	Bavi	105.37	21.08	535	1791	14.61	234
8	Lang	105.85	21.02	17	1686	14.5	246
9	Phuxuyen	105.90	20.77	9	1516	13.16	272
10	Sontay	105.50	21.13	14	1600	13.27	244
11	Chilinh	106.38	21.07	1	1489	12.51	250
12	Haiduong	106.30	20.95	3	1530	13.6	249

Table 1. Cont.

No.	Station Name	Long. (°)	Lat. (°)	Elev. (m)	Annual Rainfall (AR) (mm/year)	Standard Deviation of Rainfall (SDR) (mm/day)	No. of Rain Days (NRD) (day)
13	Hoabinh	105.33	20.82	48	1861	14.69	239
14	Maichau	105.07	20.60	579	1859	18.79	251
15	Muongte	102.63	22.47	354	2433	17.1	229
16	Tamduong	103.15	22.05	303	2333	14.49	216
17	Chilang	106.57	21.65	124	1324	11.9	267
18	Langson	106.77	21.83	263	1315	11.55	253
19	Thatkhe	106.47	22.25	157	1484	12.33	243
20	Vanmich	106.37	22.10	238	1341	11.33	240
21	Laocai	103.95	22.50	152	1810	14.11	229
22	Ninhbinh	105.98	20.27	3	1725	15.16	242
23	Baichay	107.03	20.97	59	1898	17.93	246
24	Mongcai	107.97	21.52	7	2735	24.5	230
25	Tienyen	107.44	21.33	16	2139	19.04	231
26	Sonla	103.90	21.33	709	1364	10.96	252
27	Thainguyen	105.50	21.60	784	1760	15.02	238
28	Tuyenquang	105.20	21.82	29	1575	14.02	242
29	Yenbai	104.87	21.70	41	1796	14.97	222

In Vietnam, daily ground rainfall data is often collected twice per day at 7.00 a.m. UTC + 7 and 7.00 p.m. UTC + 7, and the daily accumulation is calculated as accumulated rainfall from 7.00 a.m. UTC + 7 to the same time next day [45]. Figure 2 shows monthly rainfall distribution over Red–Thai Binh River Basin from gauge observation data. Wet season (May–October) has a high amount of rainfall, accounting for 85–90% of total annual rainfall. Very high amounts of rainfall are often observed during June, July, and August. During these periods, tropical storms often occur, with the accumulated rainfall reaching 200–600 mm within several days [44]. During the dry season (November–April), the total amount of rainfall only accounts for 10–15% of total annual rainfall.

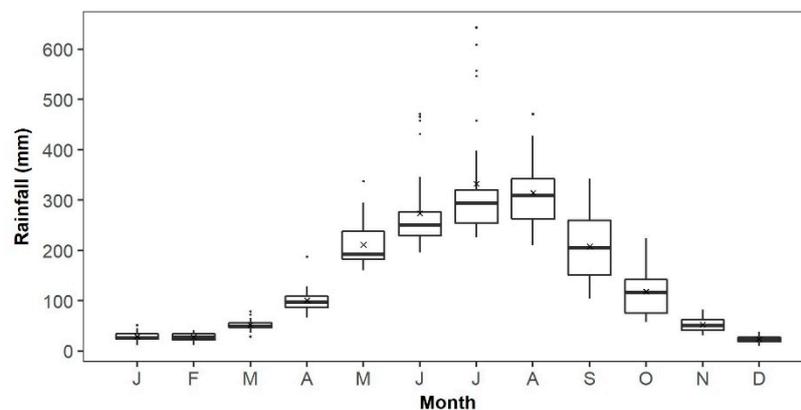


Figure 2. Monthly rainfall distribution over Red–Thai Binh River Basin (March 2000–December 2016). Cross symbol indicates average monthly rainfall.

2.2.2. TMPA Products

The TRMM is a low Earth orbits (LEO) satellite with sensors used to analyze and understand the characteristics of precipitation. The satellite is equipped with various instruments, such as Precipitation Radar (PR), TRMM Microwave Imager (TMI), Visible and Infrared Scanner (VIRS), and Lightning Imaging Sensor (LIS) [12]. The spatial coverage of TRMM is mainly in tropical and subtropical zones (50°S to 50°N) from an altitude of 400 km. The TMPA products used in this study were TMPA 3B42V7 and its real-time version TMPA 3B42RT at 0.25° spatial resolution. Detailed description of 3B42V7 can be found in Reference [12] and that of 3B42RT can be found in Reference [46]. The 3B42V7 dataset

ranges from January 1998 to present, while the 3B42RT product ranges from March 2000 to present. However, for comparison purpose, a consistent data length was required and data was therefore collected from March 2000 to December 2016 for both TMPA 3B42V7 and TMPA 3B42RT. Both products were downloaded through NASA Goddard Space Flight Center (<https://pmm.nasa.gov/data-access/downloads/trmm/>). In order to match the satellite rainfall products with the daily precipitation gauge data, the 3-hourly 3B42 products were accumulated to daily values at 0.00 UTC (equivalent to 7.00 a.m. UTC + 7).

3. Method

The comparison of TMPA 3B42V7 and TMPA 3B42RT precipitation against the ground observation data involved the extraction of data time series of TMPA products at the corresponding locations of the 29 meteorological stations. As one TMPA pixel contained one rainfall station, a total of 29 TMPA pixels were extracted to form the time series corresponding to the ground observation data.

3.1. Error Metric Assessment

To compare rainfall values between TMPA products and ground observation data, widely accepted error metrics—correlation coefficient (CC), Nash–Sutcliffe efficiency (NSE), root mean square error (RMSE), and percent bias (PBIAS)—were used [47,48]. The formulas for the statistical metrics are presented as follows:

$$CC = \frac{\sum_{i=1}^N (OBS_i - \overline{OBS})(TMPA_i - \overline{TMPA})}{\sqrt{\sum_{i=1}^N (OBS_i - \overline{OBS})^2 \sum_{i=1}^N (TMPA_i - \overline{TMPA})^2}} \quad (1)$$

$$NSE = 1 - \frac{\sum_{i=1}^N (TMPA_i - OBS_i)^2}{\sum_{i=1}^N (OBS_i - \overline{OBS})^2} \quad (2)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (TMPA_i - OBS_i)^2} \quad (3)$$

$$PBIAS = 100 \frac{\sum_{i=1}^N (TMPA_i - OBS_i)}{\sum_{i=1}^N OBS_i} \quad (4)$$

where N is the total of samples, OBS_i and $TMPA_i$ represent the rainfall values for the ground observation data and the TMPA data, respectively, and \overline{OBS} and \overline{TMPA} represent the mean of the corresponding variables. CC ranges from -1 to 1 , with strong positive correlation when the CC value is closer to 1 and strong negative correlation when the CC value is closer to -1 . NSE varies between $-\infty$ to 1 , indicating how well the plot of satellite product values and ground values fit the 1:1 line. A NSE value closer to 1 indicates a more perfect match between satellite product and ground data. RMSE is unit-based and would shed further light on the accuracy of the TMPA products. PBIAS measures the average tendency of the satellite values to be larger or smaller than the corresponding ground observations.

3.2. Detection Metric Assessment

The probability of detection (POD), false alarm ratio (FAR), the probability of false detection (POFD), and critical success index (CSI) were used to compare the occurrence and nonoccurrence of rainfall events between TMPA products and ground data [27,33,49]. The POD was the ratio of the total number of rainfall events correctly detected by the TMPA products to the total number of actual rainfall events. The FAR evaluated the ratio of the number of rainfall falsely detected by the TMPA products to the total rainfall events estimated by the TMPA products. The POFD was a fraction of false events detected by the TMPA products versus the correct observations of no rainfall events by the

TMPA products. The CSI, which is a function of POD and FAR, was the most accurate detection metric. The rainfall day threshold was set as 0.6 mm/day, which was defined as a threshold between no rainfall event and low rainfall event within 24 h based on long-term rainfall analysis over Vietnam [50]. These detection metrics can be computed as follows:

$$POD = \frac{\text{Hits}}{\text{Hits} + \text{Misses}} \quad (5)$$

$$FAR = \frac{\text{False Alarms}}{\text{Hits} + \text{False Alarms}} \quad (6)$$

$$POFD = \frac{\text{False Alarms}}{\text{False Alarms} + \text{Correct Rejections}} \quad (7)$$

$$CSI = \frac{\text{Hits}}{\text{Hits} + \text{False Alarms} + \text{Misses}} \quad (8)$$

The Hit, Miss, False Alarm, and Correct Rejection are presented in a contingency table in Table 2. The perfect scores of the POD and CSI are 1, while the perfect scores of the POFD and FAR are 0.

Table 2. Contingency table to measure the correspondence between ground observation data and Tropical Rainfall Measurement Mission Multi-satellite Precipitation Analysis (TMPA) product concerning the threshold intensity of 0.6 mm/day of a point-to-point event [51].

		Ground Observation	
		Yes	No
TMPA Product	Yes	Hit	False Alarm
	No	Miss	Correct Rejection

3.3. Rainfall Intensity Evaluation

To evaluate the rainfall intensity, we used probability density function (PDF) to classify the daily rainfall amounts into six categories based on Vietnam's regulation on rainfall classification [50]: (1) 0 to 0.6 mm; (2) 0.6 to 6 mm; (3) 6 to 16 mm; (4) 16 to 50 mm; (5) 50 to 100 mm; (6) >100 mm. The PDF analysis has been previously applied for comparing satellite rainfall products and ground data in several studies [33,36].

3.4. Climate–Topography–Based Linear-Scaling (CTLS) Bias Correction Approach

The linear-scaling (LS) approach [52,53] was based on monthly correction factor, which was the ratio between long-term monthly mean data for ground observation and TMPA.

$$CF_m = \frac{\overline{OBS}_m}{\overline{TMPA}_m} \quad (9)$$

$$TMPA_{i,m}^{corrected} = TMPA_{i,m} \times CF_m \quad (10)$$

where CF_m is the monthly mean change factor at month m , \overline{OBS}_m and \overline{TMPA}_m represent the mean of ground observation and TMPA data at month m , respectively. $TMPA_{i,m}^{corrected}$ and $TMPA_{i,m}$ are the corrected TMPA data and original TMPA data at day i of month m , respectively. In this study, we developed a set of multiple linear models that predicted correction factors CF_m from climatology–topography characteristics. We acquired station information as longitude (LONG), latitude (LAT), elevation (ELEV), annual rainfall (AR), standard deviation of rainfall (SDR), and the number of rainfall day (NRD). The CF_m can be computed as follows:

$$CF_m = \alpha_{0m} + \alpha_{1m}LONG + \alpha_{2m}LAT + \alpha_{3m}ELEV + \alpha_{4m}AR + \alpha_{5m}SDR + \alpha_{6m}NRD \quad (11)$$

where $\alpha_{0m}, \alpha_{1m}, \alpha_{2m}, \alpha_{3m}, \alpha_{4m}, \alpha_{5m}, \alpha_{6m}$ are regression coefficients corresponding to correction factor at month m . In other words, we developed a set of 12 multiple linear models to estimate correction factors from climatology–topography data. In order to select the most suitable candidates for each multiple linear model, we analyzed the relationship between the correction factor and climatology–topography for a single month and selected the significant correlation candidates. We used 23 meteorological stations (80%) to develop the abovementioned multiple linear models and six meteorological stations (20%) to verify the models.

4. Results and Discussion

4.1. Comparison between TMPA Products and Ground Observation Data

4.1.1. Daily and Monthly Scale Assessment

Table 3 presents the TMPA 3B42V7 and TMPA 3B42RT data in daily scale and monthly scale performance over the Red–Thai Binh River compared to the ground observation stations for 17 years (March 2000–December 2016). The results showed that daily rainfalls from both 3B42V7 and 3B42RT had very weak correlations with the ground observation data; the average of the CC and the average of NSE were 0.387 and -0.152 for 3B42V7 data and 0.304 and -0.521 for 3B42RT data, respectively. The negative NSE values demonstrated that TMPA values were less accurate than the mean of observed data and were therefore very poor estimations.

Table 3. Descriptive statistics for observation rain gauge and TMPA data in daily and monthly scale.

	n	TMPA 3B42V7						TMPA 3B42RT					
		Daily Scale			Monthly Scale			Daily Scale			Monthly Scale		
		Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean
CC	29	0.510	0.320	0.387	0.959	0.833	0.896	0.395	0.216	0.304	0.900	0.731	0.842
NSE	29	0.207	-0.507	-0.152	0.884	0.593	0.765	0.002	-0.968	-0.521	0.792	0.131	0.480
RMSE	29	21.7	11.4	15.1	111.6	36.2	66.5	24.5	13.7	17.3	143.6	76.4	96.0
PBIAS	29	33.2	-21.5	3.2	33.2	-21.5	3.2	38.5	-18.1	14.8	38.5	-18.1	14.8

Note: n is total number of stations. RMSE unit on a daily scale is mm/day. RMSE unit on a monthly scale is mm/month.

The statistics metric for monthly scale showed a significant improvement for both 3B42V7 and 3B42RT compared to ground data (Table 3). However, the PBIAS did not change from a daily to monthly scale. Monthly 3B42V7 and monthly 3B42RT had similar CC, with an average value of 0.896 and 0.842, respectively. However, monthly 3B42V7 data greatly outperformed monthly 3B42RT data regarding NSE, RMSE, and PBIAS. Average NSE of monthly 3B42V7 was 0.765 and no single station had a value smaller than 0.5, while average NSE of monthly 3B42RT was only 0.480. The monthly CC and NSE scores of 3B42V7 compared to ground data in this case study were very similar to the results of monthly 3B43 data compared to observations in the same basin [26]. Average RMSE of monthly 3B42V7 was 66.5 mm/month, equivalent to 30% less than the figure of monthly 3B42RT. Average PBIAS of monthly 3B42V7 was approximately 5 times less than that of monthly 3B42RT, with values of 3.2% and 14.8% respectively. The positive of PBIAS also indicated that both TMPA products overestimated compared to ground observation data. This finding was consistent with the study at the Black Volta Basin of West African countries [17] or Pahang River Basin of Malaysia [27], but it was contrary to the study in Iran [54]. It should be mentioned that although belonging to the same South East Asia region, the 3B42V7 data over the Red–Thai Binh River Basin performed better than that for Malaysia’s basin as the PBIAS of 3B42V7 for Malaysia’s basin was up to 26.95% on average [27].

We calculated various error metrics—CC, NSE, RMSE, and PBIAS. However, for TMPA’s spatial performance purpose, we only showed the spatial PBIAS score distribution. There were two reasons for this: (1) PBIAS is recommended in water resources planning projects because the overall difference

between observed and estimated values is a criteria of paramount importance [55]; (2) PBIAS is precisely aimed at defining a poor model performance and has immense variation between seasons [56].

Looking at the PBIAS distribution, the PBIAS of 3B42V7 data mostly ranged within $\pm 10\%$, while the PBIAS of 3B42RT data mostly fell in the range of 10–40% (Figure 3). The poor performance of 3B42RT data was observed at the center of the Red–Thai Binh River Basin. Moderate performances for both TMPA products were found at the northwestern mountainous area between Vietnam and Chinese border as well as the northeast coastal area.

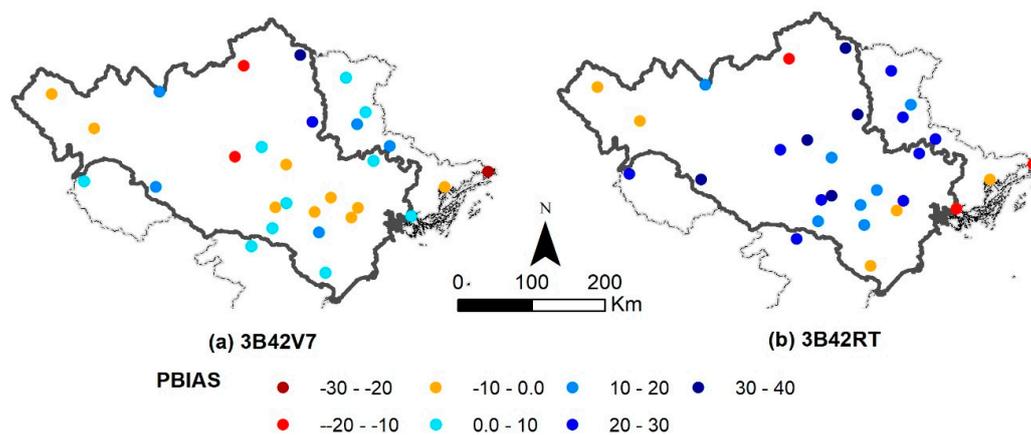


Figure 3. Percentage bias (PBIAS) score’s spatial performance of TMPA products (a) 3B42V7 and (b) 3B42RT against observation data on both daily and monthly scales from March 2000 to December 2016 over Red–Thai Binh River Basin. The grey line is the Red–Thai Binh River Basin boundary within the Vietnam territory.

4.1.2. Dry and Wet Season Assessment

Table 4 presents the performances of both daily and monthly 3B42V7 and 3B42RT during the dry season (November–April) and wet season (May–October) over the Red–Thai Binh River Basin. Generally, 3B42V7 data was better than 3B42RT data in all statistical metrics compared to ground stations, especially NSE, RMSE, and PBIAS. For example, monthly 3B42V7 had moderate NSE metric compared to ground observation, with averages of 0.586 and 0.566 in the dry season and wet season, respectively. In contrast, the figures of monthly 3B42RT were quite low, with 0.199 and 0.009, respectively. Interestingly, although RMSE of daily 3B42V7 during both dry and wet seasons were quite similar to those of daily 3B42RT, aggregation daily 3B42V7 to monthly was significantly less than monthly 3B42RT during both dry and wet seasons, with a reduction of approximately 30% for each. Regarding PBIAS, 3B42V7 and 3B42RT had almost the same bias during the dry season; however, in the wet season, 3B42V7 had significantly low PBIAS, with a value of 6.1% compared to 20.7% PBIAS of 3B42RT. In regard to the dry and wet seasons, although CC and NSE were slightly higher during the dry season than during the wet season, it was not clearly evident. On the other hand, RMSE during the dry season was observed to be much smaller than during the wet season. This can be explained as dry season receives a small amount of rainfall (10–15% of total annual rainfall), and its rainfall variation is not high as the fluctuation during the wet season.

Both TMPA products showed overall negative PBIAS values during the dry season and overall positive PBIAS values during the wet season, indicating overall underestimations during the dry season and overall overestimations during the wet season. 3B42V7 was observed to underestimate ground observation at 20 out of 29 stations, and this figure was 24 out of 29 stations for 3B42RT (Figure 4). When we used scatter plot to compare monthly dry season of TMPA products and that of ground observation (data not shown), we found that TMPA products reported zero values in many months. The wrong no-rainfall reported by TMPA data was also found at Chindwin River Basin of Myanmar [57]. The underestimation of TMPA rainfall during the dry season was consistent with

previous studies in Southwestern of China [58]. On the other hand, during the wet season, 22 out of 29 stations experienced overestimations for 3B42V7 data and 24 out of 29 stations experienced overestimation for 3B42RT data. The northwest mountain region and the northeast coastal area were the only places where both TMPA products underestimated ground observation data during two seasons. The overestimation of rainfall during the wet season agreed with a case study in Malaysia [27] but was contrary to a study involving the southwestern region of China [58].

Table 4. Descriptive statistics for daily and monthly observation rain gauge and those of TMPA data during the dry and wet seasons.

n	TMPA 3B42V7						TMPA 3B42RT						
	Dry Season			Wet Season			Dry Season			Wet Season			
	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	
Daily													
CC	29	0.487	0.289	0.407	0.494	0.264	0.344	0.423	0.196	0.317	0.364	0.141	0.255
NSE	29	-0.048	-0.884	-0.325	0.185	-0.601	-0.201	-0.076	-1.237	-0.612	-0.031	-1.107	-0.598
RMSE	29	8.8	5.8	7.0	29.1	14.7	20.0	10.2	6.4	7.7	32.8	17.9	22.9
PBIAS	29	18.9	-38.6	-10.4	37.2	-19.4	6.1	31.7	-47.3	-14.1	43.2	-13.3	20.7
Monthly													
CC	29	0.957	0.657	0.827	0.924	0.588	0.796	0.873	0.551	0.718	0.817	0.358	0.691
NSE	29	0.881	0.191	0.586	0.788	0.276	0.566	0.672	-0.381	0.199	0.618	-0.845	0.009
RMSE	29	44.9	16.8	29.0	152.8	48.0	88.7	57.1	26.7	41.0	198.3	100.6	128.6
PBIAS	29	18.9	-38.6	-10.4	37.2	-19.4	6.1	31.7	-47.3	-14.1	43.2	-13.3	20.7

Note: n is the total number of stations. RMSE unit on a daily scale is mm/day. RMSE unit on a monthly scale is mm/month.

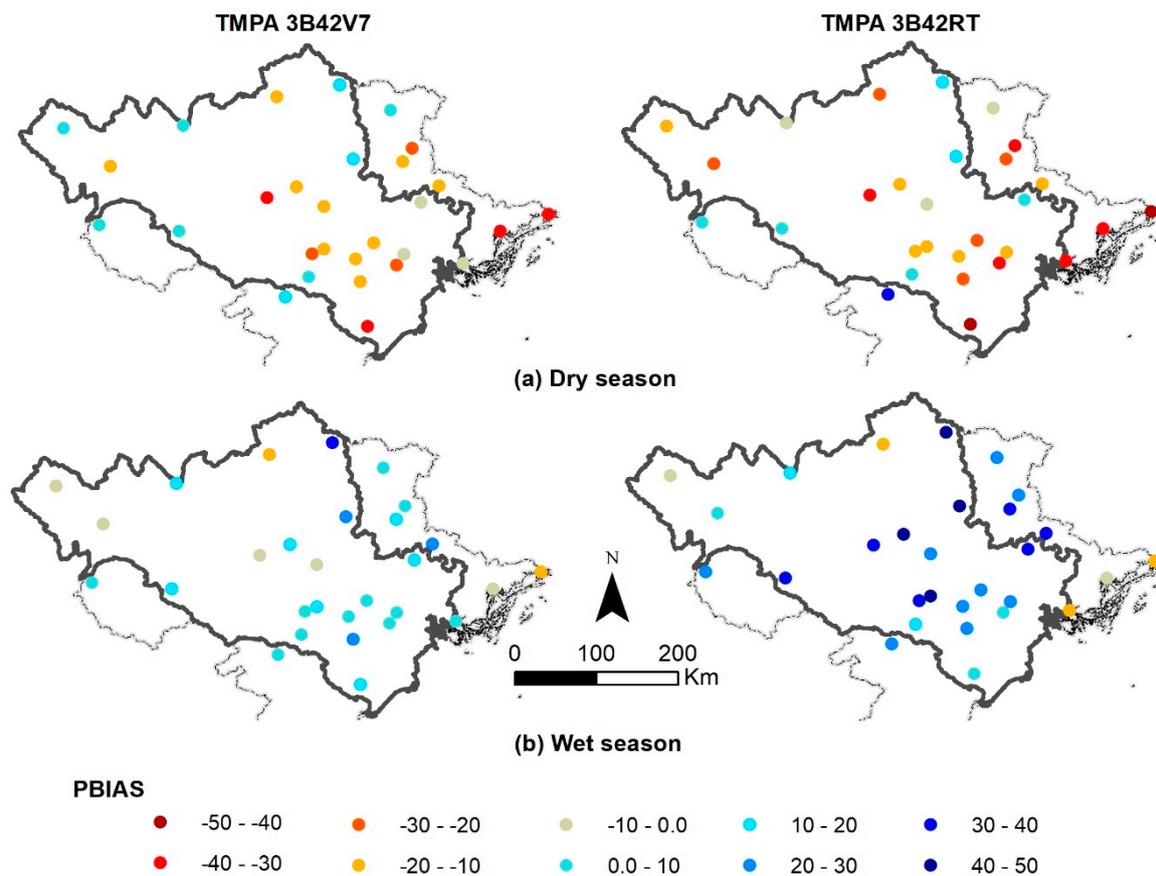


Figure 4. PBIAS score’s spatial performance of TMPA rainfall data against observation data during (a) the dry and (b) the wet season from March 2000 to December 2016 over the Red–Thai Binh River Basin. The grey line is the Red–Thai Binh River Basin boundary within Vietnam territory.

Although they had a generally positive PBIAS score, TMPA products seemed to underestimate large rainfall amounts. One possible explanation for this could be their spatial resolution. With a rather low 0.25° spatial resolution (approximately 25 km), rainfall observed in a grid was averaged over about 625 km^2 . However, rainfall can vary dramatically even with a few kilometers, and the resolution of TMPA products are often unable to pick up these differences. If we consider the complexity of terrain, this variation can be harder to estimate. Additionally, many convective storms can have a rapid evolution that a satellite will often not be able to observe accurately [59].

4.1.3. Rainfall Detection Assessment

The capacity of 3B42V7 and 3B42RT data regarding rainfall detection over the Red–Thai Binh River Basin from March 2000 to December 2016 is presented in Figure 5. Generally, the detection capacity of daily TMPA products during the wet season was much better than during the dry season, and 3B42V7 data had slightly better score than 3B42RT data. This may be associated with the temporal resolution of TMPA data as short-duration rainfall events are a typical characteristic of the dry season. Indeed, with 3-hourly products, it is easy for TMPA to miss events lasting less than this figure. On the other hand, TRMM is meant to capture and estimate convective precipitation rather than other types because of its on-board sensors. In Vietnam, precipitation is generally associated with heavier storms and cloud coverage during the wet season [60], meaning the precipitation is more likely to be detected. In contrast, in the dry season, there will be much lighter rainfall with less cloud coverage and convection, meaning that it will be more difficult to detect [59].

The POD for the whole daily TMPA data was stable over the years, with the average values of about 0.61 and 0.58 for 3B42V7 and 3B42RT, respectively. The POD scores for the daily time series in the wet season period were higher, with average values of 0.71 and 0.69 for 3B42V7 and 3B42RT, respectively. The POD scores of 3B42V7 and 3B42RT for the daily time series in the dry season period were typically low, with average values of 0.32 and 0.30, respectively. In the year 2012, the POD scores during the dry season were the lowest at about 0.2. The FAR of the daily time series and the daily values in wet season were moderate, with average values of 0.37 and 0.36 corresponding to 3B42V7 and 0.40 and 0.38 corresponding to 3B42RT. However, the FAR of the time series in the dry season was high, with an average of 43% of 3B42V7 rainfall prediction being wrong (FAR = 0.43). The wrong prediction of 3B42RT was even more than that of 3B42V7, with average FAR being 0.50. Interestingly, FAR scores of 3B42V7 and 3B42RT had great fluctuation over the years, reaching 0.6 and 0.62 in the year 2000, respectively, but the wrong prediction was reduced to only 0.30 for 3B42V7 and 0.42 for 3B42RT in the year 2014. The POFD scores were moderate for both TMPA products, with average values of 0.15 and 0.16 for 3B42V7 and 3B42RT, respectively. The POFD scores during the dry season were relatively low, with all years reporting values less than 0.1 for both TMPA products. During the wet season, POFD scores were higher than those of the dry season, with a range of 0.2–0.3. The CSI scores showed that there was no single year during the study time where the CSI scores of both 3B42V7 and 3B42RT were over 0.5. During the wet season, the average CSI values were around 0.52 and 0.50 for 3B42V7 and 3B42RT, respectively. Regarding the dry season, the CSI were only about 0.24 and 0.21 for each TMPA product, and the lowest CSI scores in the dry season were observed in 2006 and 2012.

As CSI combines different aspects of POD and FAR to give an overall assessment of TMPA performance, we used this metric to investigate the detection metric of TMPA products on the spatial scale (Figure 6). The lowland central part of the basin experienced the worst CSI score, while the northwestern mountainous part of the basin had moderate CSI score (>0.5). The better detection capacity at high elevation region than the lower land area was consistent with the study in Malaysia's basin [27].

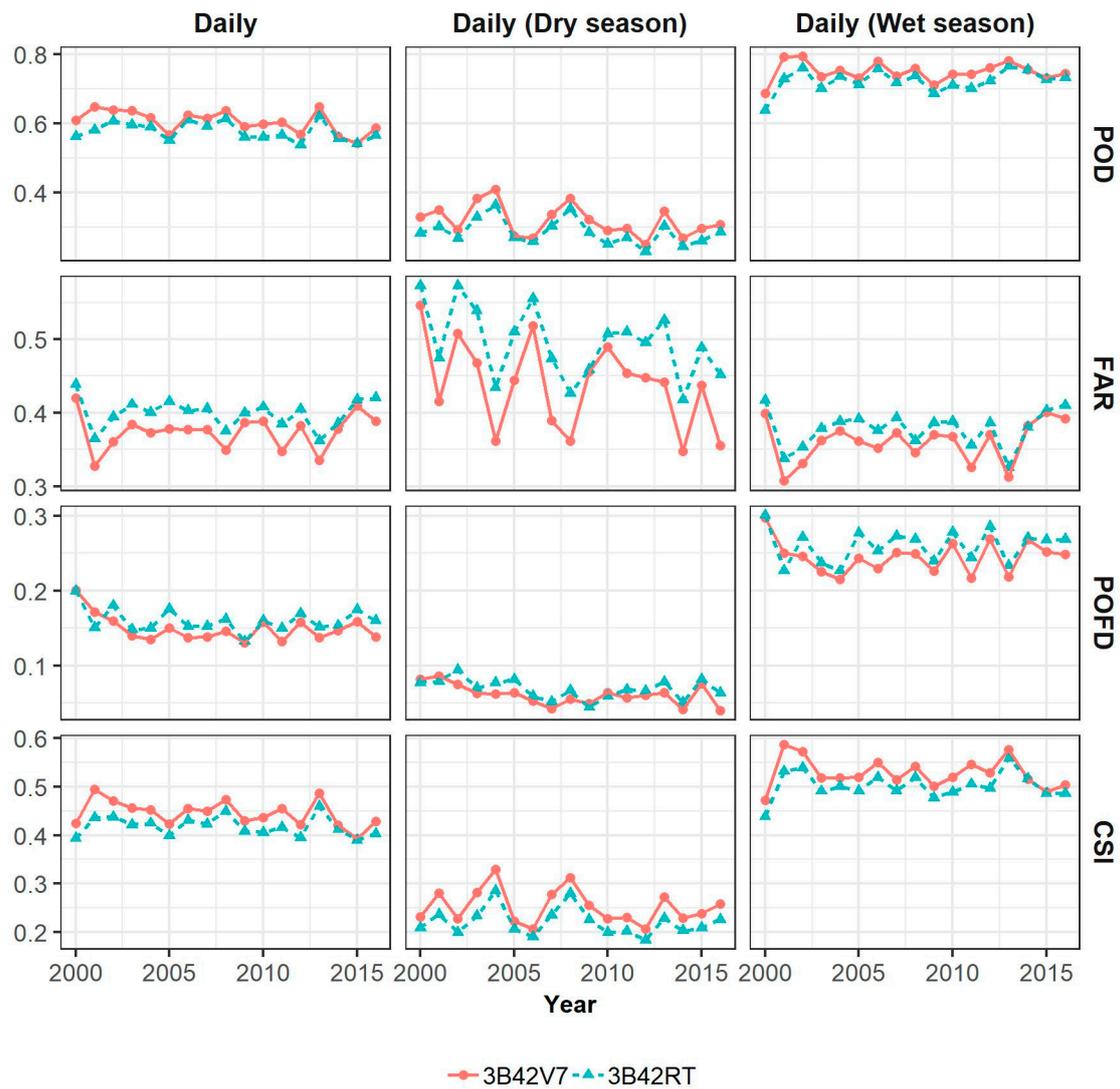


Figure 5. Average rainfall detection measurement of TMPA 3B42V7 and TMPA 3B42RT over the Red–Thai Binh River Basin from March 2000 to December 2016.

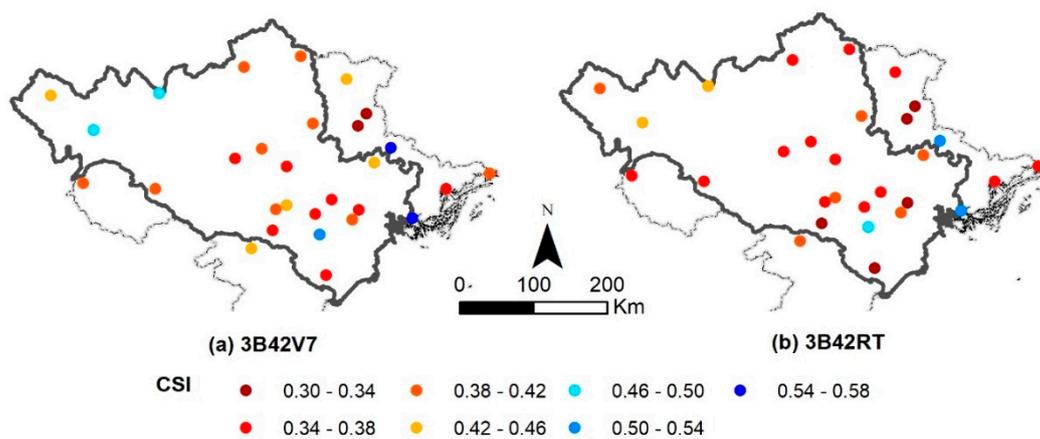


Figure 6. Critical success index (CSI) score’s spatial performance of TMPA rainfall data against observation data from March 2000 to December 2016 over the Red–Thai Binh River basin. The grey line is the Red–Thai Binh River Basin boundary within Vietnam territory.

4.1.4. Rainfall Intensity Analysis

The rainfall frequency distributions of ground observations, 3B42V7, and 3B42RT over the Red–Thai Binh River Basin are presented in Figure 7. Generally, rainfall intensity of both TMPA products followed the intensity of ground observations for the whole time series. Based on ground observation data, no rainfall (≤ 0.6 mm/day) accounted for 68.8% of total rainfall events, and 3B42V7 and 3B42RT data had similar figures. During the dry season, low rainfall intensity (0.6–6 mm/day) detected by TMPA datasets were relatively low (4.8% and 5.3% corresponding to 3B42V7 and 3B42RT) compared to the figure from ground measurement (13.1%). However, the no rainfall (≤ 0.6 mm/day) detected in the dry season was a different situation. The 3B42V7 estimated 86.4% of daily rainfall events during this season as no rainfall. Similarly, 88% of rainfall events during the dry season were considered as no rainfall events by 3B42RT. In contrast, the observations data only reported a figure of 79%. During the wet season, the no rainfall events by 3B42V7 and 3B42RT were relatively low (approximately 52% for both products), while the figure for observation data was nearly 60%. Regarding high rainfall events (50–100 mm/day) and heavy rainfall events (>100 mm/day), TMPA products had a high accuracy of detecting these, with the PDFs of both TMPA products almost the same as those of observation. The slight underestimation of low rainfall event (0.6–6 mm/day) was contrary to the overestimation conclusion of this rainfall intensity in a case study in Singapore [36].

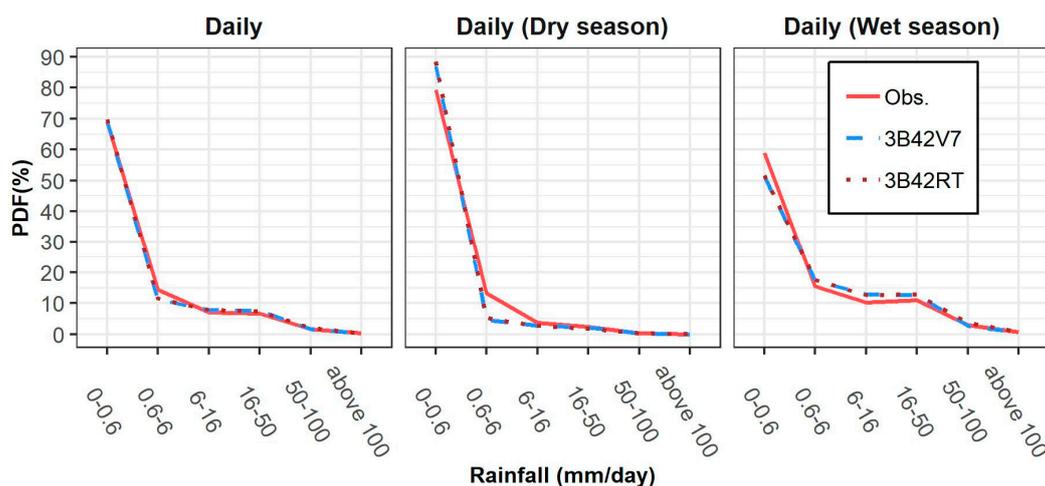


Figure 7. Average probability density function (PDF) of ground observation, TMPA 3B42V7, and TMPA 3B42RT for rainfall in daily, daily (dry season), and daily (wet season) over the Red–Thai Binh River Basin from March 2000 to December 2016.

As no rainfall and low rainfall intensity during the dry season and wet season experience significant differences between ground observations and TMPA data, we exploited the differences by analyzing seasonal spatial low rainfall and light rainfall's intensity of TMPA products. PDF differences between each TMPA data and ground observation were calculated and are presented in Figure 8. The 3B42V7 and 3B42RT data had similar characteristics, which overestimated no rainfall during the dry season (10–15%) and low rainfall intensity during the wet season (0–5%). On the other hand, the TMPA products underestimated no rainfall during the wet season (10–13%) and low rainfall intensity during the dry season (10–15%). It was noticed that the above characteristics occurred similarly for areas throughout the basin and were not specific to a typical region.

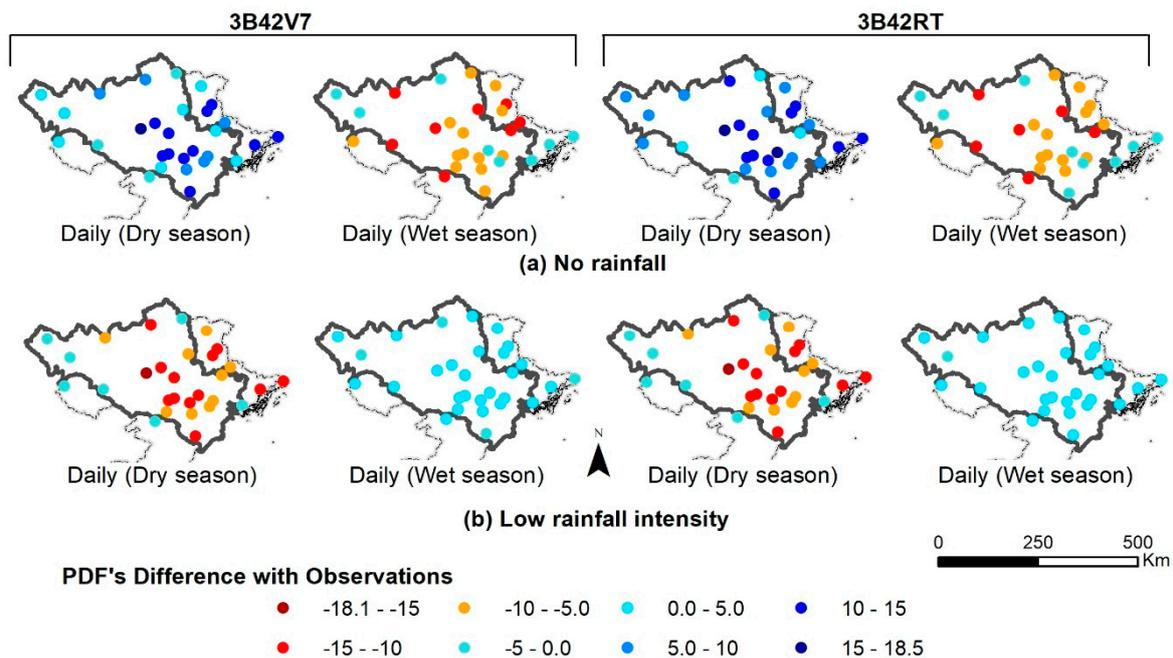


Figure 8. Percentage difference of PDF between TMPA 3B42V7, TMPA 3B42RT, and observation at (a) no rainfall intensity (0–0.6 mm/day) and (b) low rainfall intensity (0.6–6 mm/day) over the Red–Thai Binh River Basin from March 2000 to December 2016.

4.2. Development of Bias Correction Model Using Climatology–Topography Characteristics-Based Linear-Scaling (LS) Approach

4.2.1. Correlation Analysis between Climatology–Topography Characteristics and Correction Factors of LS Approach

In the LS approach, the correction factor is an important key to adjust satellite data closely to observation. Correction factors between TMPA products and observations were calculated for each month. In total, we had 12 group correction factors for 3B42V7 and 3B42RT data. Tables 5 and 6 present the relationship between correction factors in each month with climatology–topography characteristics. Based on the significant levels of the correlation coefficient, we found that topographical characteristics (LONG, LAT, and ELEV) were often associated with correction factors during dry months (except for April), while climate characteristics (AR, SDR, and NRD) were often linked with correction factors during wet months. A larger correction factor indicates larger error between satellite data and observations. ELEV (elevation) had a significant inverse relationship with the correction factor, meaning satellite data at higher elevation areas probably had a smaller error with observations compared to lower areas. This result agreed with an observation in Iran that compared 3B43V7 with rain gauge over this country [25]. Similarly, LAT (latitude) also had significant negative relationship with the correction factor. This meant that the higher the latitude area, the smaller was the satellite error. The frequency occurrence of clouds can affect the accuracy of satellite rainfall estimation [61], and NRD (a number of the rainy days) is a variable that reflects this frequency. As the number of rainy days had significant correlations with the correction factors with negative values, it seemed that the higher the number of rainy day stations, the more error of satellite-based rainfall there were. In addition, from Tables 5 and 6, AR (annual rainfall rate) and SDR (standard deviation of rainfall) had significant positive correlations with the correction factors. This means the higher the rainfall rate area, the higher was the correction factor, implying a more substantial error. This feature was the same as previous literature [62]. As a result, the correction factor for each month could be estimated from significant climatology–topography candidates.

Table 5. Correlation coefficient between correction factors of TMPA 3B42V7 against climatology–topography characteristics.

	LONG	LAT	ELEV	AR	SDR	NRD
CF ₁	−0.46 **	−0.32	−0.35	0.18	0.27	−0.21
CF ₂	0.47 **	−0.61 **	−0.48 **	0.09	0.29	−0.02
CF ₃	0.43 *	−0.58 **	−0.42 *	−0.03	0.11	0.07
CF ₄	0.01	0.00	−0.11	0.52 **	0.47 *	−0.57 **
CF ₅	0.01	0.07	−0.02	0.66 **	0.56 **	−0.65 **
CF ₆	−0.04	0.02	−0.09	0.63 **	0.45 *	−0.46 *
CF ₇	0.03	0.12	−0.08	0.69 **	0.60 **	−0.44 *
CF ₈	0.06	−0.16	−0.34	0.53 **	0.54 **	−0.50 **
CF ₉	−0.29	0.17	0.25	0.52 **	0.50 **	−0.58 **
CF ₁₀	0.26	−0.42 *	−0.40 *	0.58 **	0.68 **	−0.47 *
CF ₁₁	0.29	−0.05	−0.47 **	0.31	0.33	−0.32
CF ₁₂	0.39 *	−0.39 *	−0.41 *	0.12	0.23	−0.17

Note: * 0.05 significant level; ** 0.01 significant level

Table 6. Correlation coefficient between correction factors of TMPA 3B42RT against climatology–topography characteristics.

	LONG	LAT	ELEV	AR	SDR	NRD
CF ₁	−0.02	−0.50 **	−0.10	0.29	0.28	−0.30
CF ₂	0.48 **	−0.69 **	−0.56 **	−0.03	0.16	0.23
CF ₃	0.25	−0.55 **	−0.42 *	−0.06	0.01	0.05
CF ₄	−0.15	0.22	−0.16	0.61 **	0.43 *	−0.57 **
CF ₅	−0.22	0.37	−0.03	0.69 **	0.47 *	−0.59 **
CF ₆	−0.28	0.44 *	0.02	0.68 **	0.39 *	−0.53 **
CF ₇	0.10	−0.13	−0.30	0.61 **	0.56 **	−0.18
CF ₈	0.48 **	−0.30	−0.52 **	0.58 **	0.74 **	−0.23
CF ₉	30	0.04	−0.03	0.54 **	0.69 **	−0.34
CF ₁₀	0.19	−0.42 *	−0.34	0.57 **	0.66 **	−0.35
CF ₁₁	0.55 **	−0.22	−0.51 **	0.46 *	0.62 **	−0.24
CF ₁₂	0.64 **	−0.37	−0.45 *	0.23	0.43 *	−0.16

Note: * 0.05 significant level; ** 0.01 significant level.

4.2.2. Multiple Linear Model Development to Estimate Correction Factors

As climatology–topography characteristics have various units, before building the multiple linear regression models for correction factors, we made it dimensionless for all input climatology–topography data by scaling them to a range [0.1, 0.9]. The multiple linear models for the correction factors of 3B42V7 and 3B42RT are presented in Tables 7 and 8. All *p*-values were smaller than 0.5, indicating that sets of linear models using climatology–topography characteristics could well predict correction factors.

Table 7. Multiple linear models to predict correction factors of TMPA 3B42V7 data.

Formulas	CC	<i>p</i> -Value
CF ₁ = 1.004 * LONG + 0.947	0.446	0.045
CF ₂ = 0.736 * LONG − 2.331 * LAT − 1.347 * ELEV + 3.256	0.779	<0.001
CF ₃ = 0.242 * LONG − 1.298 * LAT − 0.735 * ELEV + 2.200	0.768	<0.001
CF ₄ = 0.103 * AR + 0.134 * SDR − 0.262 * NRD + 0.921	0.604	0.003
CF ₅ = 0.046 * AR + 0.218 * SDR − 0.404 * NRD + 1.068	0.748	0.001
CF ₆ = 1.019 * AR − 0.528 * SDR + 0.116 * NRD + 0.662	0.712	0.003
CF ₇ = 0.937 * AR − 0.170 * SDR + 0.140 * NRD + 0.615	0.733	0.002
CF ₈ = −0.351 * AR + 0.544 * SDR − 0.393 * NRD + 1.093	0.694	0.006
CF ₉ = −0.491 * AR + 0.636 * SDR − 0.563 * NRD + 1.182	0.687	0.006
CF ₁₀ = −0.286 * LAT − 0.291 * ELEV − 0.665 * AR + 0.965 * SDR − 0.678 * NRD + 1.488	0.840	<0.001
CF ₁₁ = −0.755 * ELEV + 1.373	0.600	0.003
CF ₁₂ = 0.437 * LONG − 1.422 * LAT − 1.202 * ELEV + 2.671	0.564	0.038

Note: *p*-value shows significant level between predicted correction factors using multiple linear models and calculated correction factors.

Table 8. Multiple linear models to predict correction factors of TMPA 3B42RT data.

Formulas	CC	p-Value
$CF_1 = -1.537 * LAT + 2.582$	0.501	0.041
$CF_2 = 1.102 * LONG - 3.343 * LAT - 2.204 * ELEV + 4.038$	0.875	<0.001
$CF_3 = -2.871 * LAT - 2.045 * ELEV + 4.432$	0.748	0.001
$CF_4 = 1.041 * AR - 0.515 * SDR - 0.003 * NRD + 0.672$	0.720	0.003
$CF_5 = 1.197 * AR - 0.687 * SDR - 0.082 * NRD + 0.545$	0.761	<0.001
$CF_6 = 0.215 * LAT + 1.524 * AR - 0.909 * SDR + 0.386 * NRD + 0.147$	0.875	<0.001
$CF_7 = 0.605 * AR + 0.112 * SDR + 0.696$	0.608	0.002
$CF_8 = 0.196 * LONG - 0.241 * ELEV - 0.011 * AR + 0.580 * SDR + 0.676$	0.877	<0.001
$CF_9 = -0.311 * AR + 1.037 * SDR + 0.681$	0.712	0.003
$CF_{10} = -0.664 * AR + 0.429 * SDR + 0.587 * SD + 0.948$	0.674	0.007
$CF_{11} = 1.233 * LONG - 0.638 * ELEV + 0.720 * AR + 0.616 * SDR + 0.250$	0.838	<0.001
$CF_{12} = 2.233 * LONG - 0.978 * ELEV + 0.936 * SDR + 0.842$	0.729	0.002

Note: p-value shows significant level between predicted correction factors using multiple linear models and calculated correction factors.

4.2.3. Calibration and Validation of the CTLS Bias Correction Approach

Table 9 compares the TMPA products before and after using the LS and CTLS approaches against the observations on a daily scale. Both calibration and validation data showed that LS and CTLS performed very well in reduction PBIAS scores but had moderate performances regarding NSE scores, slight improvements in RMSE scores, and almost no change in CC scores. Moreover, the linear-scaling model seemed to reduce errors better for 3B42RT data compared to that for 3B42V7 data. The reason for this may be that 3B42V7 data had already passed through the correction stage before the online public, meaning other bias correction approaches did not improve this product's quality significantly. The good performances during calibration and validation stations of the CTLS approach indicated that empirical correction factors calculated by climatology and topography characteristics could be applied for the satellite-based data bias correction process throughout the Red–Thai Binh River Basin.

Table 9. The average performance of calibration and validation for climatology–topography-based linear-scaling approach (CTLS) with TMPA 3B42V7 and TMPA 3B42RT on a daily scale.

	Before Bias Correction				LS				CTLS			
	CC	NSE	RMSE	PBIAS	CC	NSE	RMSE	PBIAS	CC	NSE	RMSE	PBIAS
3B42V7												
Calibration	0.389	−0.147	15.2	3.0	0.389	−0.130	15.2	1.1	0.389	−0.119	15.2	0.8
Validation	0.378	−0.175	14.8	4.0	0.375	−0.153	14.7	1.5	0.372	−0.165	14.8	1.7
3B42RT												
Calibration	0.303	−0.509	17.3	13.5	0.309	−0.299	16.3	−0.2	0.306	−0.342	16.6	2.2
Validation	0.307	−0.565	17.1	19.4	0.308	−0.299	15.8	0.8	0.300	−0.409	16.4	7.4

Regarding bias correction models on a monthly scale, similar results were observed as the daily scale, with a significant reduction in PBIAS scores after bias correction (Table 10). Moreover, the NSE scores of corrected monthly 3B42RT improved profoundly compared to those before bias correction. Before applying bias correction, the average monthly NSE for calibration and validation stations for 3B42RT data were 0.488 and 0.447, respectively. After using the LS approach, these figures improved to 0.734 and 0.713, respectively. Also, the empirical CTLS approach had considerable monthly NSE improvement, with values of 0.677 and 0.642 corresponding to calibration and validation stations.

Table 11 presents the performance of TMPA products regarding the PBIAS score before and after bias correction using LS and CTLS during the dry and wet seasons. The wet season seemed to benefit from bias correction more than the dry season. Using the LS approach, PBIAS scores for both 3B42V7 and 3B42RT were equal to 0, while the figures for the dry season were up to 10%. With the CTLS approach, PBIAS scores during the wet season also observed a significant decrease, ranging from 0.07% to 4.55%. During the dry season, highly positive PBIAS scores (up to 24%) were observed, indicating a high overestimation of dry season after bias correction.

Table 10. The average performance of calibration and validation for CTLS with TMPA 34B42V7 and TMPA 3B42RT on a monthly scale.

	Before Bias Correction				LS				CTLS			
	CC	NSE	RMSE	PBIAS	CC	NSE	RMSE	PBIAS	CC	NSE	RMSE	PBIAS
3B42V7												
Calibration	0.899	0.767	66.7	3.0	0.904	0.816	59.5	1.1	0.899	0.799	62.1	0.8
Validation	0.881	0.755	65.8	4.0	0.883	0.778	63.2	1.5	0.873	0.755	65.9	1.7
3B42RT												
Calibration	0.843	0.488	96.1	13.5	0.866	0.734	71.8	−0.2	0.850	0.677	78.8	2.2
Validation	0.838	0.447	95.9	19.4	0.854	0.713	71.3	0.8	0.831	0.642	79.3	7.4

Table 11. The average PBIAS score’s performance of calibration and validation for CTLS with TMPA 34B42V7 and TMPA 3B42RT during the dry and wet seasons.

	Before Bias Correction		LS		CTLS	
	Dry Season	Wet Season	Dry Season	Wet Season	Dry Season	Wet Season
3B42V7						
Calibration	−10.32	5.75	6.97	0.00	3.37	0.51
Validation	−10.92	19.29	9.63	0.00	12.67	0.07
3B42RT						
Calibration	−14.48	7.32	−0.58	0.00	5.59	1.71
Validation	−12.75	26.23	5.59	0.00	24.32	4.55

Table 12 shows the CSI scores for TMPA products against observations before and after bias correction using LS and CTLS during daily, daily (dry season), and daily (wet season). Generally, there was no significant change in CSI scores after bias correction compared to before bias correction for both products and for both bias correction approaches. By analyzing the intensity metric before and after doing bias correction (data not shown), we also obtained the same results as that of detection metric, i.e., after doing bias correction, there was no significant change in the intense rainfall for TMPA products.

Table 12. Average CSI score’s performance of calibration and validation for CTLS with TMPA 34B42V7 and TMPA 3B42RT for daily, daily (dry season), and daily (wet season).

	Before Bias Correction			LS		CTLS			
	Daily	Daily (Dry Season)	Daily (Wet Season)	Daily	Daily (Dry Season)	Daily (Wet Season)	Daily	Daily (Dry Season)	Daily (Wet Season)
3B42V7									
Calibration	0.450	0.258	0.531	0.449	0.259	0.529	0.448	0.259	0.529
Validation	0.429	0.225	0.519	0.428	0.225	0.516	0.428	0.226	0.518
3B42RT									
Calibration	0.422	0.226	0.505	0.422	0.226	0.505	0.422	0.226	0.505
Validation	0.402	0.202	0.492	0.402	0.205	0.492	0.402	0.206	0.492

5. Conclusions

TMPA products are recommended for wide use over the tropical and subtropical regions due to their high temporal–spatial resolution. Therefore, this study carried out a comparison and bias correction between TMPA 3B42V7 and TMPA 3B42RT data and 29 ground observations over the lower part of the Red–Thai Binh River Basin from March 2000 to December 2016.

Based on various error metrics—CC, NSE, RMSE, and PBIAS—we compared 3B42V7 and 3B42RT against observations at different scales, including daily and monthly scales and dry and wet seasons. Our analysis showed that both products had relatively weak relationships with observations on a daily scale but this significantly improved on a monthly scale. Except for the CC score, 3B42V7 data was considerably better than 3B42RT data in the rest of the error metrics on a monthly scale. In addition, 3B42V7 data showed better performance than 3B42RT data during both dry and wet seasons, especially regarding NSE and PBIAS measurements. Both products showed overall underestimations during the

dry season and overestimations during the wet season. Spatial analysis of the PBIAS score indicated significant bias of TMPA products at the lowland area of the Red–Thai Binh River Basin, while the northwestern mountainous area and the northeast coastal area had low PBIAS for both products.

The comparison between 3B42V7 and 3B42RT was also viewed from a different angle using detection metrics—POD, FAR, POFD, and CSI—against observations on daily time series, daily time series in the dry season, and daily time series in the wet season. In this case, the 3B42V7 showed a slightly better performance compared to 3B42RT for the metrics mentioned. Both products had better detection metrics in the wet season compared to the dry season. Spatial CSI score distribution showed that the lowland area of the central basin had the lowest score compared to other parts.

From the perspective of the assessment on rainfall intensity on daily time series for the dry and wet seasons, it was found out that 3B42RT performed the same as 3B42V7 data. Both products overestimated no rainfall (≤ 0.6 mm/day) during the dry season and underestimated rainfall intensity during the wet season. The overestimation and underestimation compromised the daily time series for the dry and wet seasons, meaning the frequency distributions of no rainfall events were almost the same for TMPA products and ground observations. On the other hand, TMPA products underestimated low rainfall intensity (0.6–6 mm/day) during the dry season and overestimated rainfall intensity during the wet season. The underestimation of low rainfall was more significant than the overestimation, resulting in a slightly lower rainfall estimation by TMPA products at the daily time series compared to observations.

In addition, we used the LS approach to do bias correction for 3B42V7 and 3B42RT products. In this approach, the correction factor is an important key to adjust satellite rainfall data closely to observations. We found that the correction factors of the LS approach were associated with climatology–topography characteristics. Therefore, a set of multiple linear regression models was developed to predict correction factors from climatology–topography characteristics for 3B42V7 and 3B42RT. After bias correction using LS and CTLS approaches, corrected TMPA products showed significant improvement compared to the results before bias correction, especially for the 3B42RT dataset with PBIAS and NSE scores. However, we found that both bias correction approaches did not improve the TMPA products significantly on other measurement scores.

In conclusion, 3B42V7 and 3B42RT data should be a good alternative source for a wide range of hydrological purposes on a monthly scale. The 3B42V7 data is also a good source for typical analysis of dry and wet seasons, although these datasets should be used with caution for daily scale purposes. The post-TMPA products after using climatology–topography characteristics are promising sources, especially for total water resource estimation.

The biggest advantage of the LS approach was to reduce PBIAS score; however, other error scores remained almost the same. Future studies may merge satellite-based and ground-based rainfall product to further improve rainfall product quality [63]. The finding of this paper gives an overview of the capacity of TMPA products in the lower part of the Red–Thai Binh River Basin regarding water resource application and provides a simple bias correction that can be used to improve the correctness of TMPA products. Additionally, the study is beneficial for regions, such as Vietnam, that are seeking alternative rainfall sources. The reason for this is that approximately 60% of Vietnam’s water resources come from abroad, and hydro-climatology acquisition from upstream countries faces many challenges due to limited administration interaction [64].

Author Contributions: Conceptualization, H.M.L., and V.L.; Methodology, H.M.L.; Validation, J.R.P.S., D.D.B.; Resources, H.M.L., and J.R.P.S.; Writing—Original Draft Preparation, H.M.L.; Writing—Review & Editing, H.M.L. and J.R.P.S.; Supervision, V.K. and D.D.B.; Project Administration, V.L. and J.D.B.; Funding Acquisition, V.L. and J.D.B.

Acknowledgments: The authors wish to thank the NASA SERVIR project (Program Manager Nancy Searby) for the funding. The authors thank the hydrological modeling group of the National Central for Water Resources Planning and Investigation (NAWAPI) for providing essential support for data collection. The authors also thank anonymous reviewers for their valuable comments to improve the manuscript’s quality.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kumar, B.; Lakshmi, V. Accessing the capability of TRMM 3B42 V7 to simulate streamflow during extreme rain events: Case study for a Himalayan River Basin. *J. Earth Syst. Sci.* **2018**, *127*, 27. [[CrossRef](#)]
2. Brutsaert, W. *Hydrology: An Introduction*; Cambridge University Press: Cambridge, UK, 2005.
3. Yilmaz, K.K.; Hogue, T.S.; Hsu, K.-L.; Sorooshian, S.; Gupta, H.V.; Wagener, T. Intercomparison of rain gauge, radar, and satellite-based precipitation estimates with emphasis on hydrologic forecasting. *J. Hydrometeorol.* **2005**, *6*, 497–517. [[CrossRef](#)]
4. Kidd, C. Satellite rainfall climatology: A review. *Int. J. Climatol.* **2001**, *21*, 1041–1066. [[CrossRef](#)]
5. Rana, S.; McGregor, J.; Renwick, J. Precipitation seasonality over the Indian subcontinent: An evaluation of gauge, reanalyses, and satellite retrievals. *J. Hydrometeorol.* **2015**, *16*, 631–651. [[CrossRef](#)]
6. Xie, P.; Arkin, P.A. Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *J. Clim.* **1996**, *9*, 840–858. [[CrossRef](#)]
7. Plengsaeng, B.; Wehn, U.; van der Zaag, P. Data-sharing bottlenecks in transboundary integrated water resources management: A case study of the Mekong River Commission's procedures for data sharing in the Thai context. *Water Int.* **2014**, *39*, 933–951. [[CrossRef](#)]
8. Gerlak, A.K.; Lautze, J.; Giordano, M. Water resources data and information exchange in transboundary water treaties. *Int. Environ. Agreem. Polit. Law Econ.* **2011**, *2*, 179–199. [[CrossRef](#)]
9. Viglione, A.; Borga, M.; Balabanis, P.; Blöschl, G. Barriers to the exchange of hydrometeorological data in Europe: Results from a survey and implications for data policy. *J. Hydrol.* **2010**, *394*, 63–77. [[CrossRef](#)]
10. García, L.; Rodríguez, D.; Wijnen, M.; Pakulski, I. *Earth Observation for Water Resources Management: Current Use and Future Opportunities for the Water Sector*; World Bank Publications: Washington, DC, USA, 2016.
11. Sun, Q.; Miao, C.; Duan, Q.; Ashouri, H.; Sorooshian, S.; Hsu, K.L. A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Rev. Geophys.* **2018**. [[CrossRef](#)]
12. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J.; Wolff, D.B.; Adler, R.F.; Gu, G.; Hong, Y.; Bowman, K.P.; Stocker, E.F. The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* **2007**, *8*, 38–55. [[CrossRef](#)]
13. Sorooshian, S.; Hsu, K.-L.; Gao, X.; Gupta, H.V.; Imam, B.; Braithwaite, D. Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Am. Meteorol. Soc.* **2000**, *81*, 2035–2046. [[CrossRef](#)]
14. Funk, C.C.; Peterson, P.J.; Landsfeld, M.F.; Pedreros, D.H.; Verdin, J.P.; Rowland, J.D.; Romero, B.E.; Husak, G.J.; Michaelsen, J.C.; Verdin, A.P. *A Quasi-Global Precipitation Time Series for Drought Monitoring*; US Geological Survey Data Series 832; U.S. Geological Survey: Sioux Falls, SD, USA, 2014.
15. Joyce, R.J.; Janowiak, J.E.; Arkin, P.A.; Xie, P. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.* **2004**, *5*, 487–503. [[CrossRef](#)]
16. Ha, L.T.; Bastiaanssen, W.G.; Griensven, A.V.; Van Dijk, A.I.; Senay, G.B. Calibration of Spatially Distributed Hydrological Processes and Model Parameters in SWAT Using Remote Sensing Data and an Auto-Calibration Procedure: A Case Study in a Vietnamese River Basin. *Water* **2018**, *10*, 212. [[CrossRef](#)]
17. Adjei, K.A.; Ren, L.; Appiah-Adjei, E.K.; Odai, S.N. Application of satellite-derived rainfall for hydrological modelling in the data-scarce Black Volta trans-boundary basin. *Hydrol. Res.* **2015**, *46*, 777–791. [[CrossRef](#)]
18. Xue, X.; Hong, Y.; Limaye, A.S.; Gourley, J.J.; Huffman, G.J.; Khan, S.I.; Dorji, C.; Chen, S. Statistical and hydrological evaluation of TRMM-based Multi-satellite Precipitation Analysis over the Wangchu Basin of Bhutan: Are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins? *J. Hydrol.* **2013**, *499*, 91–99. [[CrossRef](#)]
19. Sahoo, A.K.; Sheffield, J.; Pan, M.; Wood, E.F. Evaluation of the tropical rainfall measuring mission multi-satellite precipitation analysis (TMPA) for assessment of large-scale meteorological drought. *Remote Sens. Environ.* **2015**, *159*, 181–193. [[CrossRef](#)]
20. Zhang, A.; Jia, G. Monitoring meteorological drought in semiarid regions using multi-sensor microwave remote sensing data. *Remote Sens. Environ.* **2013**, *134*, 12–23. [[CrossRef](#)]
21. Arvor, D.; Dubreuil, V.; Ronchail, J.; Simões, M.; Funatsu, B.M. Spatial patterns of rainfall regimes related to levels of double cropping agriculture systems in Mato Grosso (Brazil). *Int. J. Climatol.* **2014**, *34*, 2622–2633. [[CrossRef](#)]

22. Cashion, J.; Lakshmi, V.; Bosch, D.; Jackson, T.J. Microwave remote sensing of soil moisture: Evaluation of the TRMM microwave imager (TMI) satellite for the Little River Watershed Tifton, Georgia. *J. Hydrol.* **2005**, *307*, 242–253. [[CrossRef](#)]
23. Li, Z.; Yang, D.; Gao, B.; Jiao, Y.; Hong, Y.; Xu, T. Multiscale hydrologic applications of the latest satellite precipitation products in the Yangtze River Basin using a distributed hydrologic model. *J. Hydrometeorol.* **2015**, *16*, 407–426. [[CrossRef](#)]
24. Tong, K.; Su, F.; Yang, D.; Hao, Z. Evaluation of satellite precipitation retrievals and their potential utilities in hydrologic modeling over the Tibetan Plateau. *J. Hydrol.* **2014**, *519*, 423–437. [[CrossRef](#)]
25. Moazami, S.; Golian, S.; Kavianpour, M.R.; Hong, Y. Comparison of PERSIANN and V7 TRMM Multi-satellite Precipitation Analysis (TMPA) products with rain gauge data over Iran. *Int. J. Remote Sens.* **2013**, *34*, 8156–8171. [[CrossRef](#)]
26. Simons, G.; Bastiaanssen, W.; Ngô, L.A.; Hain, C.R.; Anderson, M.; Senay, G. Integrating global satellite-derived data products as a pre-analysis for hydrological modelling studies: A case study for the Red River Basin. *Remote Sens.* **2016**, *8*, 279. [[CrossRef](#)]
27. Zad, M.; Najja, S.; Zulkafli, Z.; Muharram, F.M. Satellite Rainfall (TRMM 3B42-V7) Performance Assessment and Adjustment over Pahang River Basin, Malaysia. *Remote Sens.* **2018**, *10*, 388.
28. Kneis, D.; Chatterjee, C.; Singh, R. Evaluation of TRMM rainfall estimates over a large Indian river basin (Mahanadi). *Hydrol. Earth Syst. Sci.* **2014**, *18*, 2493–2502. [[CrossRef](#)]
29. Curtarelli, M.P.; Rennó, C.D.; Alcântara, E.H. Evaluation of the Tropical Rainfall Measuring Mission 3B43 product over an inland area in Brazil and the effects of satellite boost on rainfall estimates. *J. Appl. Remote Sens.* **2014**, *8*, 1–14. [[CrossRef](#)]
30. Cao, Y.; Zhang, W.; Wang, W. Evaluation of TRMM 3B43 data over the Yangtze River Delta of China. *Sci. Rep.* **2018**, *8*, 5290. [[CrossRef](#)] [[PubMed](#)]
31. Huffman, G.J. The Transition in Multi-Satellite Products from TRMM to GPM (TMPA to IMERG); NASA/GSFC Code. Available online: https://pmm.nasa.gov/sites/default/files/document_files/TMPA-to-IMERG_transition_161025.pdf (accessed on 8 June 2018).
32. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. *Integrated Multi-satellite Retrievals for GPM (IMERG) Technical Documentation*; NASA/GSFC Code; NASA's Goddard Space Flight Center: Greenbelt, MD, USA, 2015; Volume 612, p. 47.
33. Xu, R.; Tian, F.; Yang, L.; Hu, H.; Lu, H.; Hou, A. Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau based on a high-density rain gauge network. *J. Geophys. Res. Atmos.* **2017**, *122*, 910–924. [[CrossRef](#)]
34. Kim, K.; Park, J.; Baik, J.; Choi, M. Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over Far-East Asia. *Atmos. Res.* **2017**, *187*, 95–105. [[CrossRef](#)]
35. He, Z.; Yang, L.; Tian, F.; Ni, G.; Hou, A.; Lu, H. Intercomparisons of Rainfall Estimates from TRMM and GPM Multisatellite Products over the Upper Mekong River Basin. *J. Hydrometeorol.* **2017**, *18*, 413–430. [[CrossRef](#)]
36. Tan, M.L.; Duan, Z. Assessment of GPM and TRMM precipitation products over Singapore. *Remote Sens.* **2017**, *9*, 720. [[CrossRef](#)]
37. Khan, S.I.; Hong, Y.; Gourley, J.J.; Khattak, M.U.K.; Yong, B.; Vergara, H.J. Evaluation of three high-resolution satellite precipitation estimates: Potential for monsoon monitoring over Pakistan. *Adv. Space Res.* **2014**, *54*, 670–684. [[CrossRef](#)]
38. Hashemi, H.; Nordin, M.; Lakshmi, V.; Huffman, G.J.; Knight, R. Bias Correction of Long-Term Satellite Monthly Precipitation Product (TRMM 3B43) over the Conterminous United States. *J. Hydrometeorol.* **2017**, *18*, 2491–2509. [[CrossRef](#)]
39. WMO. *Guide to Hydrological Practices: Data Acquisition and Processing, Analysis, Forecasting and Other Applications*; World Meteorological Organization: Geneva, Switzerland, 1994.
40. Duc, D.D. Assessment current situation and possibility of exploiting satellite rainfall for flood forecasting- an application for Chu River Basin. *J. Clim. Chang. Sci.* **2017**, *2*, 98–104. (In Vietnamese)
41. Poortinga, A.; Bastiaanssen, W.; Simons, G.; Saah, D.; Senay, G.; Fenn, M.; Bean, B.; Kadyszewski, J. A Self-Calibrating Runoff and Streamflow Remote Sensing Model for Ungauged Basins Using Open-Access Earth Observation Data. *Remote Sens.* **2017**, *9*, 86. [[CrossRef](#)]

42. Nguyen, T.H.; Masih, I.; Mohamed, Y.A.; van der Zaag, P. Validating Rainfall-Runoff Modelling Using Satellite-Based and Reanalysis Precipitation Products in the Sre Pok Catchment, the Mekong River Basin. *Geosciences* **2018**, *8*, 164. [[CrossRef](#)]
43. NAWAPI. *Red-Thai Binh River Basin Water Resources Planning, Term of Reference*; Ministry of Natural Resources and Environment: Hanoi, Vietnam, 2017. (In Vietnamese)
44. NAWAPI. *Red-Thai Binh River Basin Water Resources Planning, Main Report*; Ministry of Natural Resources and Environment: Hanoi, Vietnam, 2017. (In Vietnamese)
45. MONRE. *Circular on National Technical Standard of Meteorological Monitoring*; 25/2012/TT-BTNMT; Ministry of Natural Resources and Environment: Hanoi, Vietnam, 2012.
46. Huffman, G.J.; Bolvin, D.T. *TRMM and Other Data Precipitation Data Set Documentation*; NASA: Greenbelt, MD, USA, 2013; Volume 28.
47. Legates, D.R.; McCabe, G.J. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **1999**, *35*, 233–241. [[CrossRef](#)]
48. Moriasi, D.N.; Arnold, J.G.; Van Liew, M.W.; Bingner, R.L.; Harmel, R.D.; Veith, T.L. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* **2007**, *50*, 885–900. [[CrossRef](#)]
49. Schaefer, J.T. The critical success index as an indicator of warning skill. *Weather Forecast.* **1990**, *5*, 570–575. [[CrossRef](#)]
50. NCHMF. *Rainfall Classification of Vietnam*; NCHMF: Hanoi, Vietnam, 2000.
51. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences (International Geophysics Series; V. 91)*; Academic Press: Cambridge, MA, USA, 2006.
52. Lenderink, G.; Buishand, A.; Deursen, W.V. Estimates of future discharges of the river Rhine using two scenario methodologies: Direct versus delta approach. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1145–1159. [[CrossRef](#)]
53. Teutschbein, C.; Seibert, J. Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *J. Hydrol.* **2012**, *456*, 12–29. [[CrossRef](#)]
54. Alijanian, M.; Rakhshandehroo, G.R.; Mishra, A.K.; Dehghani, M. Evaluation of satellite rainfall climatology using CMORPH, PERSIANN-CDR, PERSIANN, TRMM, MSWEP over Iran. *Int. J. Climatol.* **2017**, *37*, 4896–4914. [[CrossRef](#)]
55. NAWAPI. *Bang Giang—Ky Cung Water Resources Planning Project, Water Resources Assessment Report*; Ministry of Natural Resources and Environment: Hanoi, Vietnam, 2018. (In Vietnamese)
56. Gupta, H.V.; Sorooshian, S.; Yapo, P.O. Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *J. Hydrol. Eng.* **1999**, *4*, 135–143. [[CrossRef](#)]
57. Yuan, F.; Zhang, L.; Win, K.W.W.; Ren, L.; Zhao, C.; Zhu, Y.; Jiang, S.; Liu, Y. Assessment of GPM and TRMM Multi-Satellite Precipitation Products in Streamflow Simulations in a Data-Sparse Mountainous Watershed in Myanmar. *Remote Sens.* **2017**, *9*, 302. [[CrossRef](#)]
58. Hu, Q.; Yang, D.; Li, Z.; Mishra, A.K.; Wang, Y.; Yang, H. Multi-scale evaluation of six high-resolution satellite monthly rainfall estimates over a humid region in China with dense rain gauges. *Int. J. Remote Sens.* **2014**, *35*, 1272–1294. [[CrossRef](#)]
59. Ebert, E.E.; Janowiak, J.E.; Kidd, C. Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Am. Meteorol. Soc.* **2007**, *88*, 47–64. [[CrossRef](#)]
60. Nguyen, D.Q.; Renwick, J.; McGregor, J. Variations of surface temperature and rainfall in Vietnam from 1971 to 2010. *Int. J. Climatol.* **2014**, *34*, 249–264. [[CrossRef](#)]
61. Ochoa-Sánchez, A.; Pineda Ordonez, L.E.; Crespo, P.; Willems, P. Evaluation of TRMM 3B42 precipitation estimates and WRF retrospective precipitation simulation over the Pacific-Andean region of Ecuador and Peru. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 3179–3193. [[CrossRef](#)]
62. Almazroui, M. Calibration of TRMM rainfall climatology over Saudi Arabia during 1998–2009. *Atmos. Res.* **2011**, *99*, 400–414. [[CrossRef](#)]
63. Nerini, D.; Zulkafli, Z.; Wang, L.-P.; Onof, C.; Buytaert, W.; Lavado-Casimiro, W.; Guyot, J.-L. A comparative analysis of TRMM–rain gauge data merging techniques at the daily time scale for distributed rainfall–runoff modeling applications. *J. Hydrometeorol.* **2015**, *16*, 2153–2168. [[CrossRef](#)]
64. Xuan, T.T.; Tuyen, H.M.; Thai, T.T.; Dung, N.K. *Water Resources on Vietnam’s River System*; Science and Technology Publisher: Hanoi, Vietnam, 2012. (In Vietnamese)

