Earth and Space Science

RESEARCH ARTICLE
10.1029/2019EA000977

Key Points:
- Both near-real-time and post-real time TMPA and IMERG products were evaluated and compared over the Mainland China.
- IMERG-F exhibits slight overestimation but perform best in three different categories.
- In the basin-scale comparison, IMERG-F outperforms TMPA-V7, with IMERG-E and IMERG-L show performance close to or even better than TMPA-V7.

Correspondence to:
H. Lü, lvhaishen@hhu.edu.cn

Citation:
Su, J., Lü, H., Ryu, D., & Zhu, Y. (2019). The assessment and comparison of TMPA and IMERG products over the major basins of Mainland China. Earth and Space Science, 6, 2461–2479. https://doi.org/10.1029/2019EA000977

Received 30 OCT 2019
Accepted 24 NOV 2019
Accepted article online 6 DEC 2019
Published online 21 DEC 2019

©2019. The Authors. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

The Assessment and Comparison of TMPA and IMERG Products Over the Major Basins of Mainland China

Jianbin Su1, Haishen Lü1,2, Dongryeol Ryu1, and Yonghua Zhu1,2

1State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, National Cooperative Innovation Center for Water Safety and Hydroscience, College of Hydrology and Water Resources, Hohai University, Nanjing, China,
2Joint International Research Laboratory of Global Change and Water Cycle, Hohai University, Nanjing, China,
3Department of Infrastructure Engineering, University of Melbourne, Melbourne, Victoria, Australia

Abstract The Integrated Multi-satellite E Retrievals for Global Precipitation Measurement mission (IMERG) aims to deliver the “best” precipitation estimation from space and has attracted much attention. The Version 05 of IMERG products including the near-real-time “Early” and “Late” run products (IMERG-E and IMERG-L, respectively), and the post-real-time “Final” run IMERG product (IMERG-F) are assessed at both national and basin scales against gauge observations over Mainland China for a 4-year period (from April 2014 to March 2018). As control products for comparison, their predecessor Tropical Rainfall Measurement Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) products (i.e., TMPA-RT and TMPA-V7) are also employed. Components analysis confirms the best performance of IMERG-F among the five SPEs in three different categories. All five SPEs feature increasing bias and root mean square difference (RMSD) with increasing daily gauge total precipitation, and such issue is less pronounced for IMERG-F— as evidenced by the lowest bias and RMSD across all precipitation rates. Besides, compared to TMPA, IMERG products exhibit better accuracy in detecting real precipitation events, especially for light-to-medium rain (<60 mm/day), but they do not demonstrate significant improvement in the assessment of severe over/underestimation. In the basin-scale comparison, all five SPEs catch the key variation feature of basin-averaged precipitation time series (except TMPA-RT over Continental River Basin). Overall, IMERG-F demonstrates the best performance over all nine basins despite the slight overestimation, followed by TMPA-V7. IMERG-E and IMERG-L show performance close to or even better than TMPA-V7 in terms of the correlation coefficient and RMSD.

1. Introduction

Floods are among the most common and costly natural hazards worldwide (Hong et al., 2010). Over the past decades, flood-related disasters have caused economic losses of tens of billions of dollars (USD) and a large number of casualties each year (Hirabayashi et al., 2013). Most of these losses and casualties appeared in densely populated and underdeveloped countries (Wu et al., 2012), and the increasing climate variability may exacerbate the flood damages in the future (Hirabayashi et al., 2013). China, as a populous and developing country, has experienced several devastating flood events (Ma et al., 2018), such as the unprecedented great flood of Yangtze River in 1998, which affected 223 million people and resulted in direct economic loss of 166,600 million CNY (Chinese Yuan) (Zong & Chen, 2000). Frequently occurring flood events highlight the importance of flood forecasting system (FFS), which can provide sufficient lead time for risk response. In practice, the core component of an operational FFS involves an applicable hydrological model, which calibrated by retrospective hydrometeorological data to reproduce the observed hydrographs with acceptable errors, and real-time forcing to run the model and provide the up-to-date hydrological states (Zhang & Tang, 2015). However, appropriate past- or real-time precipitation data are often insufficient in most developing countries (e.g., China) and many remote parts of the world, which makes it difficult to build the operational FFS.

Precipitation is a crucial component of global water and energy cycles, and the reliability of FFS is strongly dependent on the quality of the rainfall inputs (Brunetti et al., 2018). Generally, rain gauges and ground-based weather radars are major sources of precipitation data, and they are considered as the most reliable sources (Tapiador et al., 2012). However, uneven distribution, insufficient density, limited spatial coverage hamper the wide application of these ground-based precipitation observations (GPOs), especially in the
basins with complex topography (Anjum et al., 2018; Li et al., 2017). In recent years, the rapid development of satellite observations and retrieval techniques has prompted the continuous improvement of satellite-based precipitation estimations (SPEs), providing new opportunities for regional, national, and global scale applications [Kim et al., 2018; J Su et al., 2017; Wu et al., 2012] to overcome the limitation of GPOs.

Since the Tropical Rainfall Measuring Mission (TRMM), operated by the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA), was launched in 1997 as the first dedicated meteorological precipitation satellite, several diverse quasi-global SPEs have been released publicly through different techniques, such as the TRMM Multi-satellite Precipitation Analysis (TMPA) [George J Huffman et al., 2007], Climate Prediction Center (CPC) MORPHing technique (CMORPH) (Joyce et al., 2004), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) (Sorooshian et al., 2000), and Global Satellite Mapping of Precipitation (GSMaP) (Ushio et al., 2009). Among them, the TMPA products, especially the latest Version 7 (V7), are regarded as the most reliable and extensively used products in the TRMM era over most parts of world (Dinku et al., 2008; Gao et al., 2017; Qin et al., 2014; F Su et al., 2008; J Su et al., 2017; H Yang et al., 2007; Yong et al., 2015). Over Mainland China, Yong et al. (2016) focused on evaluating both near-real-time and post-real-time TMPA by an improved error-component analysis procedure and suggested that the error components in TMPA exhibited rather strong regional and seasonal differences over diverse climate regimes. Shen et al. (2010) assessed six high-resolution SPEs over China and pointed out that all of six SPEs could capture the overall spatial distribution and temporal variations of precipitation reasonably well while CMORPH demonstrated a slightly better accuracy than TMPA. Sheng et al. (2013) systematically evaluated and compared two successive TMPA V6 and V7 over Mainland China and demonstrated that the latest TMPA V7 clearly improved upon its predecessor TMPA V6 over China. Benefited from the contributions of previous researchers, the spatial and temporal distribution, structure, and magnitude of errors contained in TRMM era SPEs have been preliminarily understood.

The TRMM satellite re-entered the earth’s atmosphere on 15 June 2015, ending its over 17 years of productive data gathering; however, it achieved great success for both providing huge volumes of quasi-global SPEs and accumulating adequate inversion experience and techniques [J Su et al., 2019]. Introduced as a successor to TRMM, the Global Precipitation Measurement (GPM) mission’s Core Observatory (CO) satellite was launched on 27 February 2014, marking a transition from the TRMM era to the GPM era. Equipped by a latest Dual-Frequency Precipitation Radar (DPR, the Ku-band at 13.6 GHz and Ka-band at 35.5 GHz) and a conical scanning, multichannel, GPM Microwave Imager (GMI, frequency ranging from 10 to 183 GHz), the CO satellite of GPM is designed to improve the accuracy in detecting light and solid precipitation and thus provide a better reference standard from a constellation of research and operational satellites (Hou et al., 2014; George J. Huffman et al., 2014; J Su et al., 2019). In addition, the GPM mission is also supported by approximately 10 partner satellites. The joint action of CO satellite and its partner satellites can effectively enhance the temporal and spatial coverage of sampling capabilities and further improve the spatiotemporal resolution.

As the Level 3 products of GPM mission, the Integrated Multisatellite Retrievals for GPM (IMERG) precipitation estimations are produced by merging all available satellite microwave estimations from the low earth orbit satellites and microwave-calibrated infrared estimations from the geosynchronous orbit (GEO) satellites (George J. Huffman et al., 2017). To accommodate the different user requirements, the IMERG is generated to supply three types of products including: the near-real-time “Early” run (IMERG-E), the “Late” run (IMERG-L), and the post-real-time “Final” run (IMERG-F) products (Hou et al., 2014; George J. Huffman et al., 2014). Among them, the IMERG-E and IMERG-L are satellite-only SPEs releasing approximately 4 and 12 hr after the observation time, respectively. The IMERG-F is a satellite–gauge merged SPE, which is calibrated by employing the Global Precipitation Climatology Center (GPCC) monthly precipitation data via the gauge calibration algorithm (GCA). As the IMERG algorithm potentially delivers the “best” precipitation estimations (George J. Huffman et al., 2017), assessment and benchmarking of IMERG products have been in focus. For instance, C Wang et al. (2018) globally intercompared and regionally assessed different versions of the IMERG products (i.e., V03, V04, and V05). Their results indicate that both IMERG-F V04 and V05 show significant differences and improvements from V03, particularly in the ocean, and the IMERG-F V05 improves upon both V04 and V03 over Mainland China. An initial comparison of monthly TMPA and IMERG products operated by Liu (2015) at a global scale reported that large systematic
differences of them exist and vary with the surface type and precipitation rate. Simultaneously, several initial statistical performance evaluations of IMERG products have been conducted in different regions, such as in India (Ganesh et al., 2019; Prakash et al., 2016), Iran (Sharifi et al., 2016), South America (Palomino-Ángel et al., 2019), West and East Africa (Dezfuli et al., 2017), East Asia (Lee et al., 2019), Singapore (M L Tan & Duan, 2017), and USA (J Tan et al., 2016). Most of them found that IMERG products generally outperform the earlier TMPA products and better capture both the spatial and temporal variation of precipitation.

Mainland China has a complex topography and spatially heterogeneous precipitation, which makes it a good test-bed for evaluating the SPEs. More importantly, the natural pattern of northern droughts and southern floods in Mainland China has generated a large number of potential users of SPEs (J Su et al., 2018). Hence, the assessment of IMERG products over Mainland China is valuable for both algorithm developers and data users. Thus far, there have been several preliminary assessments focusing on evaluating IMERG products over Mainland China. For instance, Chen and Li (2016) monthly evaluated IMERG and compared it with TMPA 3B43 products while Ning et al. (2016) directly compared IMERG and GSMaP against gauge observations over a period from April 2014 to November 2015. Moreover, IMERG and TMPA are compared in the works performed by L Jiang and Bauer-Gottwein (2019) and Wei et al. (2018). These researches all used a sparse gauge network (less than 840 weather stations over entire China) as the benchmark. However, the sparse gauge network may not provide sufficient spatial coverage and details of precipitation information and thus leads to high uncertainty of the assessments. Additionally, IMERG and TMPA were evaluated in Tang, Ma, et al. (2016) and Guo et al. (2016) by employing a relatively dense network with more than 2,400 gauges. Nevertheless, the assessments and comparisons of both were only performed over a short time frame (9 months for Tang, Ma, et al. (2016) and 1 year for Guo et al., 2016, respectively). Similarly, works of Zhao, Yang, Yang, et al. (2018) and Zhao, Yang, You, et al. (2018) intercompared the performance of IMERG and GSMaP in eight geographical subregions based on gridded Chinese daily Precipitation Analysis Product for a short-term period of March 2014 to December 2015. Due to the great interannual variation of precipitation in Mainland China, it is hard to obtain reliable statistical performance metrics of these SPEs in such a short duration. Meanwhile, J Su et al. (2018) attempted to reveal the 3-hr error features of multiple IMERG products and GSMaP via the error-component analysis, but the reference data used in it are an hourly gauge-satellite merged product produced by merging the ground data with the CMORPH. Since the reference data are not truly independent reference, it may bias the assessment results. In addition, most of the previous studies have mainly focused on the earlier versions of IMERG products (i.e., V03 and V04); thus, the assessment of IMERG V05 is still urgently needed as it has undergone many improvements upon its previous versions [C Wang et al., 2018]. Also, note that almost all assessments related to IMERG focused on the “Final” run products, and comprehensive evaluation of “Early” and “Late” runs of the IMERG products are still rare. Considering the high practical value of near-real-time SPE in actual forecasting, it is necessary to evaluate the accuracy of near-real-time IMERG products.

Therefore, the objectives of this study are twofold: (1) evaluating the quality of near-real-time and post-real-time IMERG V05 and comparing them with the corresponding TMPA against gauge observation from the China Meteorological Administration (CMA) over Mainland China, and (2) comparing IMERG and TMPA products synchronously over major basins of Mainland China to provide basic accuracy information for the potential users in flood forecasting and water resources management. The remainder of this paper is organized as follows: Section 2 introduces the study area, precipitation datasets, and metrics; Sections 3 and 4 present the main results and the discussion, separately; and a summary of the work is given in the last section.

2. Study Area, Data Sets, and Methodology

2.1. Study Area

The study area is the entire Mainland China, which spans between 73°E–135°E and 18°N–53°N and covers a land area of about 9.6 million km². As shown in Figure 1a, the terrain of China is high in west but low in east, which determines the direction of major rivers and plays an important role in the formation of China’s climate. According to the water resources and river classification, the entire Mainland China can be divided into nine major hydrological zones (http://www.resdc.cn), namely, Continental River Basin (CRB), Yellow River Basin (YeRB), Haihe River Basin (HaiRB), Song-Liao River Basin (SLRB), Southwest River
Basin (SwRB), Yangtze River Basin (YaRB), Huai River Basin (HuaiRB), Pearl River Basin (PeRB), and Southeast Basin (SeRB). The basic characteristics of these basins are presented in Table 1 (Fang et al., 2016). The CRB is located in northwest China, which is a typical arid area with severe water shortage. The YeRB, HaiRB, and YaRB distribute in north and northeast of China with semi-arid and semi-humid climate. The SwRB, YaRB, and HuaiRB distribute at the mid-latitudes in China, while PeRB and SeRB locate at the low-latitudes. Most parts of these five hydrological zones (SwRB, YaRB, HuaiRB, PeRB, and SeRB) are in the humid zone of China, whereas PeRB and SeRB receive significantly more precipitation than other three hydrological zones. In addition, Mainland China is located in the typical Asian monsoon region where the monsoonal circulation significantly affects the spatial and temporal distribution of precipitation. Coupled

Table 1

| Hydrological zone                  | Area \((10^4 \text{ km}^2)\) | Annual average rainfall (mm) | Number of stations |
|------------------------------------|-------------------------------|-----------------------------|--------------------|
| Continental River Basin (CRB)     | 335.4                         | 172.8                       | 179                |
| Yellow River Basin (YeRB)          | 79.4                          | 410.1                       | 321                |
| Haihe River Basin (HaiRB)          | 31.8                          | 504.4                       | 255                |
| Song-Liao River Basin (SLRB)       | 123.8                         | 518.4                       | 247                |
| Southwest River Basin (SwRB)       | 85.3                          | 850.4                       | 97                 |
| Yangtze River Basin (YaRB)         | 180.9                         | 993.4                       | 719                |
| Huaihe River Basin (HuaiRB)        | 26.9                          | 841.8                       | 218                |
| Pearl River Basin (PeRB)           | 45.4                          | 1518.6                      | 241                |
| Southeast Basin (SeRB)             | 24.0                          | 1620.9                      | 128                |
with the diverse topography, the Mainland China has been a good test-bed for assessing the performance of SPEs, as well as the effects of such factors on the accuracy of the related retrieval algorithms.

2.2. Ground Reference Data

Ground-based observations are normally employed as reference to assess the performance of SPEs. In this study, the daily precipitation observations from the Chinese national ground weather network, which include more than 2,400 weather stations, are employed as the reference in April 2014–March 2018. All the observations from such weather network are subjected to rigorous quality control, including analysis of extremes, contiguous values, and spatial consistency checks. The spatial distribution of these stations is shown in Figure 1b. Generally, the weather stations are distributed unevenly over Mainland China with densely populated stations in eastern China (i.e., HaiRB, HuaiRB, east part of YaRB, SeRB, and PeRB) and relatively sparse stations in western China, especially in CRB. In addition, it should be noted that, over the whole country, only 194 weather stations of such weather network have been employed by GPCC to produce and calibrate TMPA and IMERG products. Given that the dependent stations merely account for a small fraction of the total gauges (<9%), the evaluation and comparison in this research are based on more than 91% independent weather stations. Besides, both post-real-time TMPA and IMERG products are calibrated by the GPCC at monthly scale, whereas the assessment in this study is operated on daily scale. Therefore, we believe that the observations from Chinese national ground weather network can act as a receivable benchmark in evaluating and comparing TMPA and IMERG products.

2.3. TMPA and IMERG Products

In this study, the TMPA and IMERG products are selected as the typical representatives of two overlapped eras in quantitative SPEs. For the primary objectives mentioned above, both near-real-time and post-real-time products of them are evaluated and compared at national and basin scales.

The TMPA were produced by combining various precipitation estimations form different satellite systems as well as gauge observations to provide the “best” SPEs in TRMM-era (George J Huffman et al., 2007). The Version 7 of TMPA including near-real-time TMPA-RT and post-real-time TMPA-V7 are available in the NASA website (https://pmm.nasa.gov/data-access/downloads/trmm) at a spatial resolution of 0.25° × 0.25° over the 50°N–50°S global latitude band. The main difference of them is that the TMPA-RT is a satellite-only product directly derived from the combination of the calibrated passive microwave (PMW) data and the PMW calibrated infrared (IR) data, whereas TMPA-V7 is a gauge calibrated product operated by employing GPCC monthly gauge precipitation analysis.

The IMERG algorithm was designed by integrating the TMPA, the PERSIANN–Cloud Classification System (PERSIANN-CSS), and the CMORPH–Kalman Filter (CMORPH-KF) techniques into the unified United States (US)-developed algorithm. In this algorithm, multiple satellite microwave precipitation estimates were intercalibrated, merged, and interpolated to provide quantitative SPEs at fine spatiotemporal resolution (0.1° × 0.1° and half hourly). As with TMPA, IMERG algorithm also provides three kinds of products to accommodate the different user requirements, among which IMERG-E and IMERG-L are produced in near real time and calibrated with climatological coefficients while IMERG-F is produced post-real-time and calibrated by GPCC. Since their first release in 2015, the IMERG products have been updated for several times, i.e., V03, V04, V05, and V06. Since the near-real-time products of IMERG V06 are not yet available, the V05 products acquired from https://pmm.nasa.gov/data-access/downloads/gpm are employed in this study.

In addition, to structure the daily evaluation framework, TMPA and IMERG products are temporally aggregated to produce daily precipitation accumulation commensurate with the gauge observations (Yuan et al., 2017). Then, for facilitating direct comparisons, only the grids containing at least one weather station are selected from TMPA and IMERG products, respectively. If there are two or more weather station in a same grid, the mean value of them is employed as the reference.

2.4. Methodology

This research focuses on evaluating and comparing the five SPEs over the major basins of Mainland China. To build the assessment frame, a series of widely used evaluation metrics including the Pearson correlation coefficient (CC), the Bias, and the root mean square difference (RMSD) are utilized. Among them, CC
describes the linear agreement between SPEs and gauge observations; Bias, assessing the system bias of SPEs (i.e., overestimation or underestimation), is the mean value of the difference between SPEs and gauge observations essentially; and RMSD shows the average magnitude of the error and gives greater weights to the larger errors. Besides, to evaluate the precipitation detection capability of SPEs, two categorical indices are employed: the probability of detection (POD) and the false alarm ratio (FAR). POD quantifies the fraction of true rainfall occurrences that are correctly detected by SPE while FAR denotes the fraction of zero rainfall events detected by SPE as non-zero events. Considering the detection resolution of weather stations, a small value (0.1 mm/day) is used as the rain/no-rain threshold in this study. Detailed formulas and perfect values of these metrics mentioned above are listed in the Table 2. For detailed information, please refer to Wilks (2006) and Ebert et al. (2007).

### 3. Results and Discussions

#### 3.1. Grid-scale Errors in IMERG and TMPA Products

To visually assess the accuracy of SPEs in capturing precipitation accumulation from April 2014 to March 2018, the spatial distributions of 4-year daily average precipitation derived from the five SPEs and ground weather stations are shown in Figure 2. Since the southeastern and eastern parts of Mainland China are located in the East Asian monsoon region, the spatial distribution of precipitation accumulation exhibits the influence of the monsoon over the corresponding regions. From Figure 2g, it is obvious that the precipitation pattern derived by gauge observations gradually decreases from the southeast (>4 mm/day) to the northwest (<1 mm/day) of China. This is mainly due to the fact that the southeast monsoon from the Pacific Ocean and the southwest monsoon from the Bay of Bengal (as shown in Figure 1a) bring abundant moisture to southeastern China, generating a large amount of rain. The moisture carried by the monsoon declines as it moves north and the precipitation accumulation gradually decreases accordingly. Since the moisture can hardly reach the northwest China, precipitation in such areas is scarce. Besides, Mainland China also suffers from the influence of the midlatitude westerlies. Since the midlatitude westerlies is much weaker than both southeast and southwest monsoons, vast northwest China always suffers from drought and shortage of water. Due to the sparse weather stations in northwest China, the impact of the midlatitude westerlies is not captured with enough details.

All five SPEs well capture this coarse-scale spatial distribution pattern of precipitation, whereas some differences at finer scales still exist among them. In order to quantify such difference, the daily average precipitation sampled from all grids with at least one weather station are compared in the scatter plots of Figure 3. The related error metrics including CC, Bias, and RMSD are also provided in Figure 3. The near-real-time TMPA-RT presents the poorest performance in terms of all three measures of error (CC, Bias, and RMSD). The near-

### Table 2

| Statistic metrics          | Unit | Equation                                                                 | Perfect value |
|----------------------------|------|--------------------------------------------------------------------------|---------------|
| Correlation coefficient (CC)| /    | \( CC = \frac{\sum_{i=1}^{n} (R_{o,i} - R_{s,i}) (R_{o,i} - \bar{R}_o)}{\sqrt{\sum_{i=1}^{n} (R_{o,i} - \bar{R}_o)^2} \sqrt{\sum_{i=1}^{n} (R_{s,i} - \bar{R}_s)^2}} \) | 1             |
| Relative bias (Rbias)      | %    | \( Rbias = \frac{\sum_{i=1}^{n} (R_{o,i} - R_{s,i})}{\sum_{i=1}^{n} R_{o,i}} \times 100\% \) | 0             |
| Root mean square difference (RMSD) | mm/day | \( \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (R_{o,i} - R_{s,i})^2} \) | 0             |
| Probability of detection (POD) | /    | \( \text{POD} = \frac{H}{H + M} \) | 1             |
| False alarm ratio (FAR)   | /    | \( \text{FAR} = \frac{F}{H + M} \) | 0             |

Note. \( n \) represents the number of samples; \( R_{o,i} \) and \( R_{s,i} \) are SPEs and gauge observations, respectively; \( \bar{R}_o \) and \( \bar{R}_s \) are the mean values of the corresponding elements; \( H \) is the number of observed rainfall events that are correctly detected by SPE; \( M \) is the observed rainfall events which are not detected by SPE; \( F \) is the number of rainfall events detected by SPE but not occurred in the observation.
real-time IMERG-E and IMERG-L show superior performance to TMPA-RT. Interestingly, the performance of IMERG-L is slightly lower than that of IMERG-E on Bias and RMSD although IMERG-L employs more sensor information and has a longer data latency. As for the post-real-time products, TMPA-V7 demonstrates a noticeable improvements over its near-real-time product, whereas the bias of IMERG-F is higher than that of TMPA-V7 and near-real-time IMERGs (0.20 mm/day for IMERG-F versus 0.10 mm/day for TMPA-V7, and 0.02 mm/day and −0.04 mm/day for IMERG-E and IMERG-L, respectively). Considering the high RMSD values and the loose distribution of points shown in IMERG-E and IMERG-L, IMERG-F features the largest bias but smallest spread whereas IMERG-E/L have smaller bias but wider spread (than both TMPA-V7 and IMERG-F). TMPA-V7 has smaller bias than IMERG-F but similar spread (their RMSD are similar). The systematic error, bias, is easier to correct than the random error, scatters around the mean trend. For further comparison and tracking the error source, there is a need to subdivide the error of SPEs into different components (i.e., error components in Hits, Misses, and False categories).

The comparison of the five SPEs versus gauge observations in three different detection categories (i.e., Hits, Misses, and False categories) is summarized in Table 3. The IMERG products show a remarkable improvement in detecting real precipitation events—with the POD values increasing from less than 0.6 for TMPA products to more than 0.75 for IMERG products. Two major factors could contribute to such improvement: (1) the extension of the GPM sensors, which effectively improves the ability in detecting light rain [J Su et al., 2019]; (2) the finer temporal resolution of IMERG products, providing higher chance to capture short-term
precipitation events. However, the enhanced sensitivity of sensors and increased sampling frequency also increase the detection of false precipitation events—as evidenced by the increased FAR values from approximately 0.31 for TMPA products to approximately 0.39 for IMERG products. In terms of Bias and RMSD, since IMERG products have much lower Bias and RMSD values in all three categories, they exhibit superior performance to TMPA. In the intercomparison of IMERG products, IMERG-F exhibits the lowest negative bias in Misses category (Bias is −1.81 mm/day for IMERG-F versus −2.04 mm/day and −1.86 mm/day for IMERG-E and IMERG-L, respectively), but the positive bias of it is slightly higher than IMERG-L in False category. Besides, in Hits category, two near-real-time IMERG products both trend to underestimate precipitation, whereas IMERG-F exhibits a slight overestimation. The relatively high precipitation accumulation of IMERG-F in Hits and False categories is mainly caused by the monthly gauge calibration, which attempts to alleviate the underestimation of near-real-time IMERG products over the southeast China but in turn results in slight overestimation for IMERG-F detections over such regions (Xu et al., 2016). Besides, note that

**Figure 3.** Scatter plots of SPEs versus the corresponding gauge observations for (a) TMPA-RT, (b) TMPA-V7, (c) IMERG-E, (d) IMERG-L, and (e) IMERG-F in April 2014–March 2018 over all grids containing at least one weather station over entire Mainland China. Each point represents the 4-year daily average precipitation of a single grid and the color represents the density of the points.

| Sample count | Error metrics | Bias (mm/day) | RMSD (mm/day) |
|--------------|---------------|---------------|---------------|
|              | POD | FAR  | Hits | Misses | False | Hits | Misses | False | Hits | Misses | False |
| TMPA-RT      | 639,514 | 470,051 | 295,344 | 0.58  | 0.32  | 2.50 | −3.04 | 4.88  | 18.64 | 5.90  | 8.89  |
| TMPA-V7      | 630,474 | 479,091 | 264,864 | 0.57  | 0.30  | 1.30 | −3.00 | 3.57  | 16.37 | 5.83  | 6.23  |
| IMERG-E      | 839,218 | 270,347 | 590,488 | 0.76  | 0.41  | −0.42 | −2.04 | 1.66  | 14.37 | 4.13  | 3.31  |
| IMERG-L      | 853,872 | 255,693 | 545,925 | 0.77  | 0.39  | −0.51 | −1.86 | 1.36  | 13.81 | 3.78  | 2.81  |
| IMERG-F      | 857,453 | 252,112 | 525,945 | 0.77  | 0.38  | 0.45  | −1.81 | 1.44  | 12.88 | 3.70  | 2.81  |
IMERG-F shows the lowest RMSD metrics in all three categories, which highlight the excellent performance of IMERG-F.

Severe over/underestimation of precipitation can cause serious issues in the disaster-related risk assessment (e.g., flood, landslide, and hurricanes) and thus reduce the actual usability of SPEs. Hence, the fraction of severe over/underestimation events for the five SPEs are compared in Figure 4. Here, the severe overestimation is defined as the precipitation detected by SPE exceeds gauge observation by or over 100% (including part of Hits category and the whole False category), and the severe underestimation is defined as precipitation detected by SPE is less than half of gauge observation. From Figure 4, it is observed a high proportion (approximately 60% of SPE’s detections) of severe overestimation for all five SPEs, but the components of severe overestimation are different. In comparison with TMPA products, IMERG products have a high proportion of false alarms, and a lower probability of severe overestimation in the Hits category. With regard to severe underestimation, for all five SPEs, the fraction of it is relatively low (approximately 17% of SPE’s detections). However, since the severe underestimation of rainstorms can cause catastrophic damage, such a drawback still need to be addressed. In addition, by the comparison between near-real-time and post-real-time SPE (i.e., TMPA-RT versus TMPA-V7, and IMERG-E/L versus IMERG-F), the gauge calibration contributes very little in reducing the fraction of severe overestimation/underestimation, which highlights the potential utility of the near-real-time SPEs in the analysis of extreme precipitation events.

### 3.2. Grid-scale Error Characteristics at Different Precipitation Rates

Table 4 summarizes the Bias and RMSD values of the five SPEs over all grids (with at least one gauge) at 11 different ranges of precipitation rates, and their associated difference from gauge observations (i.e., SPE minus gauge observation) is shown in Figures 5a–5e. Meanwhile, taking TMPA-V7 and IMERG-F as representatives of TMPA and IMERG products respectively, the variation of their POD metrics at different precipitation rates are demonstrated in Figure 5f. A notable feature of the error shown in Figures 5a–e is the large number of outliers, which indicates extreme overestimation or underestimation of the five SPEs. The bulk of the boxplot outliers across all rain rates except >100 mm/day highlights the positively skewed distributions of difference caused by the overestimation. Besides, with increasing daily precipitation rates, all five SPEs exhibit a trend of gradually increasing negative bias—from a slight positive bias of 1.02–1.82 mm/day for the light rain (<10 mm/day). In other words, the SPE usually overestimate the light rain but underestimate the medium-to-heavy rain (>20 mm/day) (Pipunic et al., 2015; Yong et al., 2010). Meanwhile, RMSD of SPEs increases with the precipitation rates, indicating a multiplicative nature of the error. In the comparison of

| Precipitation rate (mm/day) | TMPA-RT | TMPA-V7 | IMERG-E | IMERG-L | IMERG-F | TMPA-RT | TMPA-V7 | IMERG-E | IMERG-L | IMERG-F |
|-----------------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| <10                         | 1.82    | 1.21    | 1.15    | 1.02    | 1.33    | 9.24    | 7.59    | 6.47    | 6.03    | 6.19    |
| 10–20                       | -0.79   | -1.56   | -1.97   | -1.81   | -0.64   | 17.09   | 14.88   | 14.36   | 13.95   | 12.92   |
| 20–30                       | -3.76   | -4.50   | -5.51   | -5.13   | -3.10   | 22.05   | 19.50   | 19.50   | 19.15   | 17.59   |
| 30–40                       | -6.64   | -7.76   | -9.46   | -8.82   | -5.95   | 26.96   | 23.65   | 24.21   | 23.77   | 21.96   |
| 40–50                       | -10.50  | -11.97  | -13.85  | -13.05  | -9.80   | 31.13   | 27.66   | 29.68   | 29.16   | 26.06   |
| 50–60                       | -14.68  | -16.55  | -18.91  | -18.05  | -13.64  | 35.37   | 32.24   | 34.05   | 33.11   | 30.56   |
| 60–70                       | -18.41  | -20.95  | -23.86  | -22.79  | -17.59  | 39.98   | 36.86   | 38.95   | 37.77   | 34.73   |
| 70–80                       | -21.77  | -25.24  | -28.96  | -27.64  | -21.72  | 43.95   | 41.26   | 43.74   | 42.63   | 39.21   |
| 80–90                       | -28.18  | -31.66  | -34.54  | -33.08  | -26.26  | 48.49   | 46.18   | 49.83   | 48.54   | 44.50   |
| 90–100                      | -30.76  | -35.97  | -36.84  | -35.14  | -28.16  | 52.56   | 50.68   | 53.54   | 51.91   | 47.06   |
| >100                        | -58.17  | -62.62  | -61.59  | -58.96  | -50.00  | 81.92   | 80.09   | 82.13   | 79.94   | 71.66   |
multiple SPEs, near-real-time IMERG products generally have the highest negative bias (except for light rain), but post-real-time IMERG-F exhibits the lowest negative bias, which highlights the importance of GCA in producing IMERG-F. Besides, among the five SPEs, IMERG-F also has the lowest RMSD metrics. In Figure 5f, results show a clear trend of increasing POD with increasing rain rate—indicating that the higher daily total precipitation, the lower risk of missing. Even so, for the heavy rain events, the relative high negative Bias and RMSD metrics as well as the numerous outliers indicate that it still far away from excellent.

Besides, for light/moderate rain events (<90 mm/day), IMERG products has higher POD metrics than TMPA products.

3.3. Spatial Distribution of Errors in IMERG and TMPA Products

In order to analyze the spatial variations of SPEs, the temporal error metrics (including CC, Bias, and RMSD) of the five SPEs over all grids with at least one weather station are calculated and compared in this section. Note that the no-rain days are excluded from the analysis to reduce their impact on the error metrics, particularly over arid regions. Besides, since TMPA products cannot provide the precipitation information above the latitude of 50°N, the weather stations over such regions are not involved in the comparison.

Figures 6a–6e are spatial distributions of CC for the five SPEs, with associated boxplots in Figure 6f. In general, the spatial patterns of CC are similar to that of daily average precipitation shown in Figure 2, which shows higher values in the southeast but lower values in the northwest. For near-real-time TMPA-RT, the
Figure 6. Spatial distributions of daily CCs for the five SPEs over the weather stations of Mainland China from April 2014 to March 2018: (a) TMPA-RT, (b) TMPA-V7, (c) IMERG-E, (d) IMERG-L, (e) IMERG-F, and (f) is related boxplots of them. The edges of the box in (f) are 25th and 75th percentiles, the central measure is the median, and the ends of outliers are based on 1.5 times of interquartile range.

Figure 7. Scatterplots of daily precipitation time series sampled from the five SPEs against gauge observations over Tunxi and Ruoqiang weather stations. Note that the no-rain days (no precipitation is detected by all five SPEs and gauge) are removed from the time series for reducing visual clutter.
CC is acceptable (generally higher than 0.6) over the humid southeast China, but it decreases appreciably over the semi-arid and semi-humid areas (varies between 0.4 and 0.6). The CC of TMPA-RT is a little lower over northwest China, particularly around the northern boundary of the Tibetan Plateau where CC values are even lower than 0.2. In contrast, TMPA-V7 shares almost identical distributions of CC with TMPA-RT, but exhibits higher values over most parts of Mainland China. The near-real-time IMERG products (i.e., IMERG-E and IMERG-L) demonstrate close or even slightly better CC values to TMPA-V7, and the post-real-time IMERG-F displays best CC of the five selected SPEs—with CC values generally higher than 0.7 over most of central and southern China. Even so, IMERG products still suffer from the low CC over northwest China. Tang, Ma, et al. (2016) suggested that the varied topography and arid climate contribute most to the low CC metrics over such areas, and few rain gauges over these regions are adopted in GPCC system is also an important factor for the post-real-time SPEs (J Su et al., 2018). For visual comparison, the Tunxi and Ruoqiag weather stations (refer to Figure 6a) are chosen as representative of humid and arid areas, respectively. The precipitation time series of the five SPEs versus gauge observations over these two stations are compared in Figure 7. It is obvious that the high probability of false precipitation (FAR values are higher than 0.85 for TMPA products and approximately 0.8 for IMERG products) is the key shortcoming of all SPEs over arid regions. This is mainly due to that hydrometeors detected by spaceborne sensors are

Figure 8. Spatial distributions of daily Bias (mm/day) and RMSD (mm/day) for multiple SPEs over the weather stations of Mainland China from April 2014 to March 2018. The boxplots of Bias and RMSD are also exhibited in it.
partially or totally evaporate before reaching surface (J Su et al., 2018). Since rain areas delineation and rain/no-rain delineation cannot be corrected by the GCA (J Su et al., 2018; Xu et al., 2016), compared to corresponding near-real-time products, there is no significant improvement of post-real-time SPEs. Hence, for further improving the performance of SPEs (particularly for arid regions), more attention should be focused on developing more accurate rain areas delineation algorithm.

As with CC, the spatial distributions of Bias (mm/day) and RMSD (mm/day) for the five SPEs are shown in Figure 8, respectively. TMPA-RT exhibits severe overestimation over most parts of China, particularly for north China, Tibetan Plateau, and southeast China (R1 in Figure 8a) where Bias values of TMPA-RT are generally higher than 1.5 mm/day. IMERG-E and IMERG-L also overestimate the precipitation over most of north China and the North-China Plain (R3 in Figure 8a), but the magnitude of Bias is significantly lower than TMPA-RT. Besides, IMERG-E and IMERG-L display remarkable underestimation over south China and Sichuan Basin (R2 in Figure 8a), unlike the overestimation of TMPA-RT over such regions. In regard to the post-real-time products, given that the same ground-based benchmark (i.e., GPCC) and GCA are employed by TMPA-V7 and IMERG-F, they share almost identical distributions of Bias. However, in these two post-real-time products, the GCA takes effect in different ways. In TMPA-V7, the GCA attempts to alleviate the overestimation shown in TMPA-RT, whereas, in IMERG-F, it tends to suppress the overestimation.
over northwest China, serving as an important factor for the poor performance of SPEs over such region. In comparison to TMPA products, IMERG products exhibit relatively high POD values (generally higher than 0.8) and FAR values (generally higher than 0.45) over the North-China Plain (R3 in Figure 9a), which closely relates to the high frequency of light rain. As for the central and southeast China, TMPA products demonstrate low POD metrics, whereas IMERG products show significant improvement over these regions (except for Sichuan Basin). In summary, IMERG products can better catch the precipitation events but also have higher detection of false precipitation.

### 3.4. Basin-scale Assessments and Comparisons

Many studies used the SPEs in their works on water balance analyses, water resource management, hydrological simulation, and disaster forecasting (W Yang et al., 2016). However, basin-scale assessments of SPEs is the process necessary for their widely operational applications, as the inherent biases and indeterminate errors of them can be propagated or even amplified through the applied integration processes and thus result in poor results (J Su et al., 2017; J Su et al., 2019). Figure 10 is the summary of evaluation metrics for the daily basin mean precipitation time series sampled from the five SPEs against gauge observations—with associated monthly basin mean precipitation time series in Figure 11. Here, the monthly time series (instead of daily) is selected for reducing visual clutter.

Compared to the evaluation metrics demonstrated in Figures 6 and 8, Figure 10 is featured by the increased CC and the reduced RMSD, which primarily attributes to the ability of spatial aggregation to partly filter out the uncorrelated error and thus improve the performance in estimating real precipitation. Therefore, all the five SPEs in Figure 11 (except TMPA-RT over the CRB) can catch the key variation feature of basin-averaged precipitation, further indicating that the SPEs can provide accurate information on precipitation occurrence and amount in basin-averaged estimations. Meanwhile, among the five SPEs, two post-real-time SPEs best catch the monthly precipitation variation of all basins as expected. Furthermore, in terms of daily CC and RMSD shown in Figure 10, IMERG-F demonstrates the best performance over all nine basins—as
evidenced by the highest CC and lowest RMSD. However, in comparison to TMPA-V7, it still demonstrates a slight higher overestimation, particularly over YaRB, HuaiRB, and SeRB. The high probability of false precipitation (shown in Figure 9) and the overestimation in the Hits scenario (shown in Table 3) contribute mainly to such overestimation. In practical applications, if the Bias of IMERG-F is spatially and temporally homogeneous, it is not the key drawback as the hydrologic integration processes can tolerate some bias in a certain extent (J Su et al., 2019), and the calibration of model parameters can also suppress it (Li et al., 2017; Yuan et al., 2017). Hence, although the outliers of IMERG-F in Figure 8f indicate some severe overestimation/underestimation, the short interquartile range, which means low spatial heterogeneity also highlights its high potential hydrological utility over Mainland China.

The near-real-time spatial precipitation products with quantitative uncertainty are extremely valuable for shorter-term streamflow forecasting, which can provide sufficient lead time for flood-related risk responses. From this study, it observes that the performance of near-real-time TMPA and IMERG products generally exhibit greater spatial heterogeneity. TMPA-RT shows limited performance over the CRB as its high overestimation and poor agreement with gauge observations, whereas IMERG-E and IMERG-L demonstrate a significant improvement by evidence of both improved daily evaluation metrics and the better matching with gauge-based month precipitation. For other north basins at high latitude (e.g., YeRB, HaiRB, and SLRB), TMPA-RT significantly overestimate the precipitation, particularly in winter seasons. However, benefiting from the joint collaboration of DPR and GMI, which helps improve the accuracy in detecting solid precipitation, the performance of IMERG-E and IMERG-L is improved obviously in winter seasons. The SwRB situating across the southeast Tibetan Plateau and Yun-Gui Plateau is featured with complex topography and volatile climate. Over such basin, TMPA-RT demonstrates a remarkable overestimation, whereas IMERG-E and IMERG-L show notable underestimation especially in summer seasons. A similar situation also occurs in moist YaRB, PeRB, and SeRB. With regard to HuaiRB characterized by the transition zone between

Figure 11. The monthly basin mean precipitation time series of five SPEs and gauge observations during the study period from April 2014 to March 2018.
southern and northern climates, IMERG-E and IMERG-L show severe overestimation in winter seasons (Figure 11g), which is not observed in TMPA-RT. Such overestimation can principally ascribe to the super-sensitive of sensors in GPM’s CO satellite over such region. Generally, near-real-time TMPA-RT usually exhibits worst performance while IMERG-E and IMERG-L close or even better performance to that of post-real-time TMPA-V7 in spite of slight higher Bias.

3.5. Implications of the Results

The rapid development of satellite observation and retrieval algorithm has provided unprecedented opportunities for the improvement of SPEs. The IMERG products supported by the state-of-the-art GPM mission have been updated four times (i.e., V03, V04, V05, and V06) since it was firstly released in early 2015. Even so, IMERG is still in its infancy and the TMPA, as the predecessor of IMERG, will continue to be generated until the IMERG products can totally substitute TMPA products. This study demonstrates a systematic evaluation framework to explore the continuity and difference between IMERG and TMPA products across Mainland China, contributing toward a better understanding of errors in these SPEs. However, note that this evaluation is conducted at pixel-to-point scale, which means that the average precipitation of a large area (>100 km²) has been directly compared to the observations at special points. The spatial mismatch may bias the assessment results especially in regions with great spatial variability precipitation, such as mountain areas and monsoon regions. Given a fact that the spatial resolution of IMERG products and TMPA products are 0.1° and 0.25° respectively, such spatial mismatch is more pronounced for TMPA products. In this context, it may potentially make sense that IMERG products are better than TMPA products.

For promoting the potential hydrological application, a preliminary evaluation and comparison of these SPEs are performed over the major basins of Mainland China. Our results highlight their excellent performance over the humid southeast China (e.g., YaRB, PeRB, and SeRB) and confirm their poorer performance over the arid northwest China (e.g., CRB) and the high-latitude basins (e.g., YeRB, HaiRB, and SLRB) as reported by previous studies (J Su et al., 2018; Tang, Ma, et al., 2016; Wei et al., 2018), providing some guidance for potential hydrologic applications over these basins. However, according to Camici et al. (2018) and L Jiang and Bauer-Gottwein (2019), the relation between the quality of the precipitation product and the hydrological performance is not straightforward that means a better precipitation product does not necessarily guarantee a better discharge simulation. Qi et al. (2016) suggested that a good discharge simulation closely depends on a good coalition of hydrological model and precipitation product. So, although the satellite-based precipitation products are not as accurate as the gauge-based products, they could also have better performance in discharge simulations when appropriately combined with hydrological models (Qi et al., 2016). Therefore, hydrological assessment of SPEs over given catchments is also a meaningful topic.

Over Mainland China, previous studies have assessed the discharge simulation using SPEs as forcing input, such as Li et al. (2017) and Tang, Zeng, et al. (2016) over the Ganjiang River basin (81,258 km²), Z Wang et al. (2017) over the Beijiang River basin (38,672 km²), J Su et al. (2019) over the upper Huaihe River basin (16,000 km²), Yuan et al. (2018) and J Su et al. (2017) over the source region of Yellow River basin (122,000 km²), and S Jiang et al. (2018) over the Mishui basin (9,972 km²). Most of them confirmed that SPE forced hydrological simulations are hard to outperform or equal to gauge based hydrological simulations due to their seasonal and regional systematic biases and random errors [D Jiang & Wang, 2019] although their hydrologic utility may be acceptable. In addition, some studies discovered that the model recalibration by employing SPEs as rainfall forcing could increase the performance of the discharge simulation significantly (S Jiang et al., 2018; Z Wang et al., 2017; Yuan et al., 2018), while the model parameters may be unrealistic and thus limit the model’s performance at the sub-basin scale (D Jiang & Wang, 2019; Maggioni & Massari, 2018). In addition, the hydrological assessments of SPEs over Mainland China are usually conducted in a small number of catchments, and therefore, it is not clear to what extent the results can be generalized (Beck et al., 2017). Furthermore, only fairly large catchments (usually larger than 10,000 km²) and complex distributed hydrological models (e.g., the Grid-based Xinanjiang model (S Jiang et al., 2018; Yuan et al., 2017), the Variable Infiltration Capacity model (J Su et al., 2019; Z Wang et al., 2017) and the Coupled Routing and Excess Storage model (Li et al., 2017; Tang, Zeng, et al., 2016)) are employed in evaluating hydrological utility of SPEs, leading to combined rainfall and model uncertainty that is not easily interpreted (Beck et al., 2017). In this context, it is of great value to evaluate and compare the hydrological utility of IMERG and TMPA products over different small watershed with diverse landscapes.
topography, and climatic conditions using the traditional lumped hydrological model (i.e., simple rainfall-runoff model). Ji Jiang and Bauer-Gottwein (2019) explored the feasibility of using SPEs to force a lumped Hydrologiska Byrans Vattenbalansavdelning hydrological model over 300 headwater catchments of varying size and climate for a 2-year period (2016–2017). Results indicated that IMERG products provide comparable performances to gauge-based precipitation, whereas post-real-time TMPA-V7 performs relatively poor in terms of discharge simulation. However, the short study period, which cannot be representative for long-term catchment conditions, is a major limitation. Also, when a different rainfall-runoff model was used, the results may be different. Hence, more efforts are urgently needed to explore how do SPEs perform as hydrological model forcing, particularly over ungauged or poorly gauged basins.

4. Conclusion

In this study, we focus on evaluating and comparing IMERG and TMPA products over Mainland China against a relatively dense gauge network from April 2014 to March 2018. For our research objectives, both the near-real-time and post-real-time products of them are employed. Subsequently, their quality is compared in the major basins of China to promote the widely hydrological application. The main conclusions are summarized as follows:

1. Compared to the gauge observations, all five SPEs can capture the coarse-scale spatial distribution pattern of precipitation accumulation well. In the pixel-to-point comparison, post-real-time IMERG-F presents a comparable accuracy to TMPA-V7, in spite of a slightly higher overestimation (0.20 mm/day for IMERG-F versus 0.1 mm/day for TMPA-V7). The near-real-time IMERG-E and IMERG-L demonstrate a poorer accuracy than post-real-time products, but both of them do much better than TMPA-RT in terms of almost all evaluation metrics.

2. Due to the enhanced sensitivity of sensors and increased sampling frequency, IMERG products have better accuracy in detecting real precipitation events, but also contain more false precipitation. Component analysis in three different detection categories highlights the excellent performance of IMERG-F among the five SPEs. However, IMERG-F does not demonstrate significant improvement in severe overestimation/underestimation comparison.

3. All five SPEs present high accuracy in detecting heavier precipitation, in spite of the notable underestimation. As for light and moderate rain, IMERG products are more reliable by evidence of increased POD metrics. With lower Bias and RMSD, IMERG-F shows the best performance at almost all precipitation rates.

4. For the spatial variations, all five SPEs show better performance over south China but poor performance over northwest China. Attributing to the GCA, which alleviates the overestimation/underestimation effectively, the post-real-time SPEs usually exhibit low spatial heterogeneity over entire Mainland China.

5. All five SPEs can catch the key variation feature of basin-averaged precipitation time series (except TMPA-RT over CRB), indicating high potential hydrological utility. In terms of CC and RMSD, post-real-time IMERG-F demonstrates the best performance over all nine basins. Nevertheless, it also contains a slight higher overestimation than TMPA-V7. Although IMERG-E and IMERG-L exhibit poorer Bias, they demonstrate close or even better performance to that of TMPA-V7 in terms of daily CC and RMSD.

6. In summary, to reply to the original scientific question, we confirmed the significant improvement of IMERG, particularly for near-real-time products. However, care should be taken when employing them (especially for TMPA-RT) over the CRB and high-latitude catchments (e.g., YeRB, HaiRB, and SLRB). Hence, it is necessary to develop appropriate predictive error models, thus reducing their uncertainty. Besides, further studies are needed to evaluate and compare their hydrological utility over different catchments of varying size and climate.

References

Anjum, M. N., Ding, Y., Shangguan, D., Ahmad, I., Ijaz, M. W., Farid, H. U., et al. (2018). Performance evaluation of latest integrated multisatellite retrievals for Global Precipitation Measurement (IMERG) over the northern highlands of Pakistan. *Atmospheric Research*, 205, S0169809517311018.

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I. J. M., Weedon, G. P., et al. (2017). Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences*, 21(12), 6201–6217. https://doi.org/10.5194/hess-21-6201-2017

Acknowledgments

This work was funded through the National Key Research and Development Program (grant nos. 2018YFA0605400 and 2016YFA0605404); NNSF (grant nos. 41830752, 41971042, and 41571015); and the Open Funding of the Laboratory (grant nos. OFSLRSS201806, HRM201704, and 2018B44114). The TMPA and IMERG data were provided by NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) at http://disc.sci.gsfc.nasa.gov/. The rain gauge data obtained from the China Meteorological Data Service Center at http://data.cma.cn.
Brunetti, M. T., Melillo, M., Peruccacci, S., Ciabatta, L., & Brocca, L. (2018). How far are we from the use of satellite rainfall products in landslide forecasting? Remote Sensing of Environment, 210, 65–75.

Camici, S., Ciabatta, L., Massari, C., & Brocca, L. (2018). How reliable are satellite precipitation estimates for driving hydrological models: A verification study over the Mediterranean area. Journal of Hydrology, 563, 950–961. https://doi.org/10.1016/j.jhydrol.2018.06.067

Chen, F., & Li, X. (2016). Evaluation of IMERG and TRMM 3B43 Monthly precipitation products over Mainland China. Remote Sensing, 8(6), 472.

D DezFál, A. K., Ichoku, C. M., Mohr, K. I., & Huffman, G. J. (2017). Precipitation characteristics in West and East Africa from satellite and in situ observations. Journal of Hydrometeorology, 18(6), 1799–1805. https://doi.org/10.1175/jhm-d-17-0068.1

Dinku, T., Chidzambwa, C., Ceccato, P., Connor, S. I., & Ropelewski, C. F. (2006). Validation of high-resolution satellite rainfall products over complex terrain. International Journal of Remote Sensing, 29(14), 4097–4110. https://doi.org/10.1080/01431160701772526

Elbert, E. E., Janowiak, J. E., & Kidd, C. (2007). Comparison of near-real-time precipitation estimates from satellite observations and numerical models. Bulletin of the American Meteorological Society, 88(1), 47.

Fang, G., Qi, H., Wen, X., & Zhou, L. (2016). Analysis of spatiotemporal evolution of extreme monthly precipitation in the nine major basins of China in 21st century under climate change (in Chinese). Journal of Natural Disasters, 25(2), 15–25.

Ganesh, S., Pradhan, A., & Indu, J. (2019). Evaluation of GPM sampling error over major basins in Indian subcontinent using bootstrap technique. Advances in Space Research, 68(10), 3289–3302. https://doi.org/10.1016/j.asr.2019.01.031

Gao, Z., Long, D., Tang, G., Zeng, C., Huang, J., & Hong, Y. (2017). Assessing the potential of satellite-based precipitation estimates for flood frequency analysis in ungauged or poorly gauged tributaries of China’s Yangtze River basin. Journal of Hydrology, 550, 478–496. https://doi.org/10.1016/j.jhydrol.2017.05.025

Guo, H., Chen, S., Bao, A., Behrangi, A., Hong, Y., Nadyisaba, F., et al. (2016). Early assessment of integrated multi-satellite retrievals for global precipitation Measurement over China. Atmospheric Research, 176–177, 121–133. https://doi.org/10.1016/j.atmosres.2016.02.020

Hirabayashi, Y., Mahendran, R., Koirala, S., Knoxohima, L., Dai, Y., Watanabe, S., et al. (2013). Global flood risk under climate change. Nature Climate Change, 3(9), 816–821.

Hong, Y., Adler, R. F., Huffman, G. J., & Pierce, H. (2010). Applications of TRMM-based multi-satellite precipitation estimation for global flood modeling: prototyping a global flood modeling system. Netherlands: Springer.

Hou, A. Y., Kakar, R. K., Neeki, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., et al. (2014). The global precipitation measurement mission. Bulletin of the American Meteorological Society, 95(5), 701–722.

Huffman, G. J., Bolvin, D., Nelkin, E., Wolff, D. B., Adler, R. F., Gu, G., et al. (2007). The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. Journal of Hydrometeorology, 8(1), 38–55.

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., & Xie, P. (2014). NASA Global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG), edited.

Huffman, G. J., Bolvin, D. T., & Nelkin, E. J. (2017). Integrated Multi-satellite Retrievals for GPM (IMERG) Technical Documentation, edited.

Jiang, D., & Wang, K. (2019). The role of satellite-based remote sensing in improving simulated streamflow: A review. Water, 11(8), 1615.

Jiang, L., & Bauer-Gottwein, P. (2019). How do IMERG precipitation estimation perform as hydrological model forcing? Evaluation for 300 catchments across Mainland China. Journal of Hydrology, 572, 468–500. https://doi.org/10.1016/j.jhydrol.2019.03.042

Jiang, S., Ren, L., Xu, C.-Y., Yong, B., Yuan, F., Liu, Y., et al. (2018). Statistical and hydrological evaluation of the latest Integrated Multi-satellite Retrievals for GPM (IMERG) over a midlatitude humid basin in South China. Atmospheric Research, 214, 416–429.

Joyce, R. I., Janowiak, J. E., Askin, P. A., & Xie, P. (2004). CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. Journal of Hydrometeorology, 5(3), 487–503.

Kim, S., Paik, K., Johnson, F. M., & Sharma, A. (2018). Building a flood-warning framework for ungauged locations using low resolution, open-access remotely sensed surface soil moisture, precipitation, soil, and topographic information. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 11(2), 375–387.

Lee, J., Lee, E.-H., & Seol, K.-H. (2019). Validation of Integrated multisatellite retrievals for GPM (IMERG) by using gauge-based analysis products of daily precipitation over East Asia. Theoretical and Applied Climatology, 137(3-4), 2401–2512. https://doi.org/10.1007/s00704-018-2749-1

Li, N., Tang, G., Zhao, P., Hong, Y., Gou, Y., & Yang, K. (2017). Statistical assessment and hydrological utility of the latest multi-satellite precipitation analysis IMERG in Ganjiang River basin. Atmospheric Research, 183, 212–223. https://doi.org/10.1016/j.atmosres.2016.07.020

Liu, Z. (2015). Comparison of integrated multisatellite retrievals for GPM (IMERG) and TRMM multisatellite precipitation analysis (TMPA) monthly precipitation products: Initial Results. Journal of Hydrometeorology, 17(3), 773–790. https://doi.org/10.1175/JHM-D-15-0068.1

Ma, M., He, B., Wan, J., Jia, P., Guo, X., Gao, L., et al. (2018). Characterizing the flash flooding risks from 2011 to 2016 over China. Water, 10, 704.

Maggioni, V., & Massari, C. (2018). On the performance of satellite precipitation products in riverine flood modeling: A review. Journal of Hydrology, 558, 214–224. https://doi.org/10.1016/j.jhydrol.2018.01.039

Ning, S., Wang, J., Jin, J., & Ishidaira, H. (2016). Assessment of the latest GPM-era high-resolution satellite precipitation products by comparison with observation gauge data over the chinese mainland. Water, 8(11). https://doi.org/10.3390/w8110481

Palomino-Ángel, S., Anaya-Acedo, J. A., & Botero, B. A. (2019). Evaluation of 3B42V7 and IMERG daily precipitation products for a very high-precision region in northwestern South America. Atmospheric Research, 217, 37–48. https://doi.org/10.1016/j.atmosres.2018.10.012

Pipunic, R. C., Ryu, D., Costelloe, J. F., & Su, C.-H. (2015). An evaluation and regional error modeling methodology for near-real-time satellite rainfall data over Australia. Journal of Geophysical Research: Hydrology, 120, 10,767–10,783. https://doi.org/10.1002/2015JD023512

Prakash, S., Misra, A. K., Pai, D. S., & AghaKouchak, A. (2016). From TRMM to GPM: How well can heavy rainfall be detected from space? Advances in Water Resources, 88, 1–7. https://doi.org/10.1016/j.adwatres.2015.11.008

Qi, W., Zhang, C., Fu, G., Sweetapple, C., & Zhou, H. (2016). Evaluation of global fine-resolution precipitation products and their uncertainty quantification in ensemble discharge simulations. Hydrology and Earth System Sciences, 20(2), 903–920. https://doi.org/10.5194/hess-20-903-2016

Qin, Y., Chen, Z., Shen, Y., Zhang, S., & Shi, R. (2014). Evaluation of satellite rainfall estimates over the Chinese Mainland. Remote Sensing, 6(11), 11649–11672.
短缺，对不同地表和气象条件下的评估进行了初步研究。Remote Sensing, 8(2), 135.

Shen, Y., Xiong, A., Wang, Y., & Xie, P. (2010). Performance of high-resolution satellite precipitation products over China. Journal of Geophysical Research, 115, D02114. https://doi.org/10.1029/2009JD012097

Sheng, C., Yang, H., Cao, Q., Gourley, J. J., Kirstetter, P. E., Yong, B., et al. (2013). Similarity and difference of the two successive V6 and V7 TRMM multisatellite precipitation analysis performance over China. Journal of Geophysical Research: Atmospheres, 118, 13060–13074. https://doi.org/10.1002/2013JD018964

Sorooshian, S., Hsu, K. L., Gao, X., Gupta, H. V., Imam, B., & Dan, B. (2000). Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. Bulletin of the American Meteorological Society, 81(9), 2035–2046.

Su, F., Hong, Y., & Lettenmaier, D. P. (2008). Evaluation of TRMM multisatellite precipitation analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin. Journal of Hydrology, 394, 622–640. https://doi.org/10.11175/2007jhm944.1

Su, J., Li, H., Wang, J., Sadeghi, A., & Zhu, Y. (2017). Evaluating the applicability of four latest satellite–gauge combined precipitation estimates for extreme precipitation and streamflow predictions over the Upper Yellow River Basins in China. Remote Sensing, 9(11), 1176.

Su, J., Li, H., Zhu, Y., Cui, Y., & Wang, X. (2019). Evaluating the hydrological utility of latest IMERG products over the Upper Huaihe River Basin, China. Atmospheric Research, 225, 17–29. https://doi.org/10.1016/j.atmosres.2019.03.025

Su, J., Li, H., Zhu, Y., Wang, X., & Wei, G. (2018). Component analysis of errors in four GPM-based precipitation estimations over Mainland China. Remote Sensing, 10(9), 1420.

Tan, J., Petersen, W. A., Kirstetter, P.-E., & Tian, Y. (2016). Performance of IMERG as a function of spatiotemporal scale. Journal of Hydrology, 538, 307–319. https://doi.org/10.1017/jhdy.2016.107

Tan, M. L., & Duan, Z. (2017). Assessment of GPM and TRMM precipitation products over Singapore. Remote Sensing, 9(7), 720.

Tang, G., Ma, Y., Long, D., Zhong, L., & Hong, Y. (2016). Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over mainland China at multiple spatiotemporal scales. Journal of Hydrology, 533, 152–167.

Tang, G., Zeng, Z., Long, D., Guo, Y., Yong, B., Zhang, W., & Hong, Y. (2016). 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? Journal of Hydrology, 171(1), 121–137. https://doi.org/10.11175/jhm-d-15-0059.1

Tapiafero, F. J., Turk, F. J., Petersen, W., Hou, A. T., García-Ortega, E., Machado, L. A. T., et al. (2012). Global precipitation measurement: Methods, datasets and applications. Atmospheric Research, 104, 70–97. https://doi.org/10.1016/j.atmosres.2011.10.023

Tian, Y., Huffman, G. J., Adler, R. F., Tang, L., Sapiano, M., Maggioni, V., & Wu, H. (2013). Modeling errors in daily precipitation measurements: Additive or multiplicative? Geophysical Research Letters, 40, 2060–2065. https://doi.org/10.1002/gfl.50320

Ushio, T., Kubota, T., Shige, S., Okamoto, K., Aonashi, I., Inoue, T., et al. (2009). A Kalman filter approach to the global satellite mapping of precipitation (GSMaP) from combined passive microwave and infrared radiometric data. Journal of the Meteorological Society of Japan, 87A(3), 137–151. https://doi.org/10.2151/jmsj.87A.137

Wang, C., Tang, G., Han, Z., Guo, X., & Hong, Y. (2018). Global intercomparison and regional evaluation of GPM IMERG Version-03, Version-04 and its latest Version-05 precipitation products: Similarity, difference and improvements. Journal of Hydrology, 564, 342–356. https://doi.org/10.1016/j.jhydrol.2018.06.064

Wang, Z., Zhong, R., Lai, C., & Chen, J. (2016). Evaluation of the GPM IMERG satellite-based precipitation products and the hydrological utility. Atmospheric Research, 196, 151–163. https://doi.org/10.1016/j.atmosres.2017.06.020

Wei, G., Lü, H., Crow, W. T., Zhu, Y., Wang, J., & Su, J. (2018). Comprehensive evaluation of GPM-IMERG, CMORPH, and TMPA precipitation products with gauged rainfall over Mainland China. Advances in Meteorology, 2018, 18. https://doi.org/10.1155/2018/3024190

Wilks, D. S. (2006). Statistical methods in the atmospheric sciences. New York: Elsevier.

Wu, H., Adler, R. F., Hong, Y., Tian, Y., & Policelli, F. (2012). Evaluation of global flood detection using satellite-based rainfall and a hydrologic model. Journal of Hydrology, 413–414, 1268–1284.

Xu, S., Shen, Y., & Du, Z. (2016). Tracing the source of the errors in IMERG using a decomposition evaluation scheme. Atmosphere, 7(12), 161.

Yang, H., Adler, R. F., Hossain, F., Curtis, S., Huffman, G. J., & Hong, C. (2007). A first approach to global runoff simulation using satellite rainfall estimation. Water Resources Research, 43, W08502. https://doi.org/10.1029/2006WR005739

Yang, W., John, V. O., Zhao, X., Hui, L., & Knapp, K. R. (2016). Satellite climate data records: Development, applications, and societal benefits. Remote Sensing, 8(4), 331.

Yong, B., Chen, B., Tian, Y., Yu, Z., & Hong, Y. (2016). Error-component analysis of TRMM-based multi-satellite precipitation estimates over Mainland China. Remote Sensing, 8(5), 440.

Yong, B., Liu, D., Gourley, J. J., Tian, Y., Huffman, G. J., Ren, L., & Hong, Y. (2015). Global view of real-time trmm multisatellite precipitation analysis: Implications for its successor global precipitation measurement mission. Bulletin of the American Meteorological Society, 96(2), 293–296. https://doi.org/10.1175/bams-d-14-00017.1

Yong, B., Ren, J. L., Hong, Y., Wang, J. H., Gourley, J. J., Jiang, S. H., et al. (2010). Hydrologic evaluation of multisatellite precipitation analysis standard precipitation products in basins beyond its inclined latitude band: A case study in Laohahe basin, China. Water Resources Research, 46, W07542. https://doi.org/10.1029/2009WR008965

Yuan, F., Wang, B., Shi, C., Cui, W., Zhao, C., Liu, Y., et al. (2018). Evaluation of hydrological utility of IMERG Final Run V05 and TMPA 3B42V7 satellite precipitation products in the Yellow River source region, China. Journal of Hydrology, 567, 696–711. https://doi.org/10.1016/j.jhydrol.2018.06.045

Yuan, F., Zhang, L., Win, K., Ren, L., Zhao, C., Zhu, Y., et al. (2017). Assessment of GPM and TRMM multi-satellite precipitation products in streamflow simulations in a data-sparse mountainous watershed in Myanmar. Remote Sensing, 9(3), 302.

Zhang, X., & Tang, Q. (2015). Combining satellite precipitation and long-term ground observations for hydrological monitoring in China. Journal of Geophysical Research: Atmospheres, 120, 6426–6443. https://doi.org/10.1002/2015JD024400

Zhao, H., Yang, B., Yang, S., Huang, Y., Dong, G., Bai, J., & Wang, Z. (2018). Systematical estimation of GPM-based global satellite mapping of precipitation products over China. Atmospheric Research, 201, 206–217. https://doi.org/10.1016/j.atmosres.2017.11.005

Zhao, H., Yang, S., You, S., Huang, Y., Wang, Q., & Zhou, Q. (2018). Comprehensive evaluation of two successive V3 and V4 IMERG final run precipitation products over Mainland China. Remote Sensing, 10(1). https://doi.org/10.3390/rs10010034

Zong, Y., & Chen, X. (2000). The 1998 flood on the Yangtze, China. Natural Hazards, 22(2), 165–184. https://doi.org/10.1023/a:1008119805106