

Geophysical Research Letters

RESEARCH LETTER

10.1029/2021GL094092

Key Points:

- Four-component error decomposition (4CED) is proposed to quantify positive and negative error components for precipitation products
- The 4CED incorporates and outdoes the original error decomposition method
- The 4CED can reveal distinct spatial patterns and seasonal characteristics of individual error components

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. Ye, azye@bnu.edu.cn

Citation:

Zhang, Y., Ye, A., Nguyen, P., Analui, B., Sorooshian, S., & Hsu, K. (2021). New insights into error decomposition for precipitation products. *Geophysical Research Letters*, 48, e2021GL094092. https://doi.org/10.1029/2021GL094092

Received 27 APR 2021 Accepted 13 AUG 2021

© 2021. 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.

New Insights Into Error Decomposition for Precipitation Products

Yuhang Zhang¹, Aizhong Ye¹, Phu Nguyen², Bita Analui², Soroosh Sorooshian², and Kuolin Hsu²

¹State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of Geographical Science, Beijing Normal University, Beijing, China, ²Center for Hydrometeorology and Remote Sensing, Department of Civil and Environmental Engineering, University of California Irvine, Irvine, CA, USA

Abstract It is very important to quantify errors of precipitation estimation products. However, the existing methods do not describe all error components and are therefore not comprehensive enough. In this study, we propose a four-component error decomposition method (4CED) that decomposes the total errors of precipitation products into four independent parts: hit positive bias, hit negative bias, false bias, and missed bias. And we use it to evaluate the performance of the three latest satellite precipitation products in the eastern monsoon region of China. Our study reveals 4CED has apparent improvements compared with the previous method. Results also provide new insights for tracking error sources and quantifying the error magnitudes of precipitation products. Moreover, the proposed 4CED can be extended to different spatial and temporal scales. Our new method will not only contribute to product upgrades, but also provide guidance for potential applications.

Plain Language Summary Precipitation products still have biases compared to the precipitation that actually occurs. The possible biases are divided into precipitation overestimation and precipitation underestimation. And they can be further divided into hit precipitation overestimation, false alarm overestimation, hit precipitation underestimation and missed underestimation. Recognizing and understanding these biases can better leverage the full effectiveness of precipitation products. Here, we have quantified these biases using statistical metrics and found that we may have underestimated the magnitude of the biases when using the classical method, especially in some regions with complex topography where precipitation estimation remains a challenge.

1. Introduction

Precipitation is not only a key variable in climate studies, but also a crucial forcing factor of the hydrological cycle (Do et al., 2020; Jiang & Bauer-Gottwein, 2019; Yao et al., 2020). Currently, various precipitation products are developed through different technologies, such as gauge interpolation, radar observation, satellite retrieval, and assimilation reanalysis (Foufoula-Georgiou et al., 2020; Sun et al., 2018). However, despite the significant advances of these products, there still exists errors associated with them due to many factors (e.g., algorithms, data sources, bias correction). Therefore, it is important to identify and quantify these error sources before proceeding to applications, which is also beneficial to the iterative updates of products (Alijanian et al., 2019; Sorooshian et al., 2000).

Many efforts have been done to address this task, including the comparison of gauge observations, intercomparison, and hydrological validation (Beck et al., 2017; Chen et al., 2020; Liu et al., 2017; Massari et al., 2017; Qi et al., 2016; Shen et al., 2014). The statistical indicators used include those describing consistency (e.g., Pearson correlation coefficient, PCC), continuity (e.g., root mean square error, RMSE; mean error, ME) and probability (e.g., probability of detection, POD) (Tang et al., 2016). Moreover, some studies further decomposed these metrics to better identify error sources (Beck et al., 2019; Gupta et al., 2009). For example, Tian et al. (2009) separated the errors (ME) into three independent components (hit bias, missed precipitation and false precipitation). AghaKouchak et al. (2012) decomposed the errors (MSE) into systematic and random error components. A more recent study attempted to quantify the uncertainties of different precipitation products using a three-dimensional variance partitioning approach (Zhou et al., 2020). Among these methods, the ME error decomposition was well explained and has been widely used in many





(a) Contingency 2*2 classes

Figure 1. The contingency table (a), four components error decomposition (b) and an example of precipitation error decomposition (c).

studies (Chen et al., 2019; Tian et al., 2009; Su et al., 2018). An improved error decomposition scheme for satellite-based precipitation products was simultaneously validated and emphasized in India (Chaudhary & Dhanya, 2021).

(b) Four components error decomposition

In this study, an alternative four-component error decomposition method (4CED) is proposed to better describe the error components of precipitation products. The proposed scheme of 4CED decomposes the total errors into four independent parts: hit positive bias, hit negative bias, missed bias, and false bias. Moreover, we apply it to three state-of-the-art satellite precipitation products (SPPs) to illustrate the feasibility of the method. This study has two objectives: (1) To illustrate the enhancement of the new method; (2) To provide new insights for identifying and tracking errors in precipitation products.

2. Methods

Traditionally, the statistical indicator (mean error, ME) is commonly used to quantify the bias of satellite precipitation. This indicator is obtained by summing the difference between SPPs and observations. Such an indicator is usually valid when comparing different SPPs, but there may be some cases where misleading results occur. For example, a dataset with a large daily bias (both positive and negative) may add up to a small bias when averaged over a long time. Conversely, a product with a small daily bias (only positive) may add up to a larger bias. More importantly, the fact that we only know a total bias (even with some other metrics, see Text S2 for more discussion) prevents us from tracking the sources of the error in depth and getting better feedback to improve the retrieval algorithms. Tian et al. (2009) decomposed the error of SPPs into three parts (hit bias, missed precipitation, and false precipitation) to better track the error sources associated with the satellite retrieval processes. However, hit bias, as one of the three error components in their method, is still a combined value obtained after offsetting positive and negative bias, and therefore does not accurately describe the magnitude of all individual errors. In this study, we extend it to a more comprehensive version called four-component error decomposition (4CED) to further decompose hit bias into a positive part and a negative part. Illustrations and an example can be found in Figure 1.

Firstly, given a precipitation field $R_d(\vec{x},t)(d = 1)$, reference data; d = 2, satellite data), we can derive a binary-valued precipitation event mask $P_d(\vec{x},t)$ based on a rain/no rain threshold (1 mm/d in this study).



$$P_d(\vec{x},t) = \begin{cases} 1, & \text{if } R_d(\vec{x},t) \ge \text{threshold} \\ 0, & \text{if } R_d(\vec{x},t) < \text{threshold} \end{cases}$$
(1)

Then we can define four different masks by product the $P_1(\vec{x},t)$ and $P_2(\vec{x},t)$ as follows:

hit mask :
$$P_{12} = P_1 \times P_2$$
 (2)

miss mask :
$$P_{1\overline{2}} = P_1 \times P_{\overline{2}}$$
 (3)

false mask :
$$P_{\overline{12}} = P_{\overline{1}} \times P_2$$
 (4)

correct negative mask :
$$P_{\overline{12}} = P_{\overline{1}} \times P_{\overline{2}}$$
 (5)

where $P_{\overline{n}}$ denotes the Boolean complement of a binary mask P_n .

If we define the difference between $R_1(\vec{x},t)$ and $R_2(\vec{x},t)$ as $R(\vec{x},t)$, then the hit precipitation field can be expressed by:

$$H(\vec{x},t) = R(\vec{x},t) \times P_{12}$$
(6)

$$R\left(\vec{x},t\right) = R_2\left(\vec{x},t\right) - R_1\left(\vec{x},t\right)$$
(7)

and we can further derive a binary-valued hit precipitation event mask as:

$$P_{H}(\vec{x},t) = \begin{cases} 1, & \text{if } H_{1}(\vec{x},t) \ge H_{2}(\vec{x},t) \\ 0, & \text{if } H_{1}(\vec{x},t) < H_{2}(\vec{x},t) \end{cases}$$
(8)

where $P_H(\vec{x},t) = 1$ denotes reference precipitation greater than satellite precipitation; $P_H(\vec{x},t) = 0$ denotes reference precipitation smaller than satellite precipitation.

Then we define another two masks:

hit positive mask :
$$P_{H_1H_2} = P_{H_1} \times P_{H_2}$$
 (9)

hit negative mask : $P_{\overline{H_1}H_2} = P_{\overline{H_1}} \times P_{H_2}$ (10)

where $P_{\overline{n}}$ denotes the Boolean complement of a binary mask P_n .

Finally, we can decompose total error into four different parts as follows:

hit positive bias : HPB =
$$R(\vec{x}, t) \times P_{H_1H_2}$$
 (11)

hit negative bias : HNB =
$$R(\vec{x}, t) \times P_{H_1H_2}$$
 (12)

false bias : FB =
$$R(\vec{x}, t) \times P_{12}$$
 (13)

missed bias : MB =
$$R(\vec{x},t) \times P_{1\bar{2}}$$
 (14)

Among four error components, HPB and FB are always positive, and HNB and MB are always negative. The relationships between them with mean error (ME) and mean absolute error (MAE) are:

$$ME = \underbrace{FB + HPB}_{overestimation} + \underbrace{HNB + MB}_{underestimation}$$
(15)



$$MAE = \underbrace{FB + HPB}_{overestimation} - \underbrace{(HNB + MB)}_{underestimation}$$
(16)

The MAE $[0, +\infty)$ accumulates the absolute difference between the SPPs and reference data for each day compared to the ME $(-\infty, +\infty)$ which directly accumulates the difference. The closer MAE (or ME) is to 0, the smaller the total errors.

Through the intermediate steps in the above equations, we cannot only deduce the classical ME error decomposition, but also get many relevant statistical indicators, such as POD, FAR (see Text S1). Moreover, these equations can also be applied to other temporal scales (e.g., hourly).

3. Data

We apply our error composition scheme to three state-of-the-art satellite precipitation estimates products (SPPs) in this study, namely the Precipitation Estimate from Remotely Sensed Information using Artificial Neural Networks-Dynamic Infrared Rain Rate near-real-time (PDIR-Now, hereafter, PDIR) (Nguyen, Ombadi, et al., 2020, Nguyen, Shearer, et al., 2020), Global Precipitation Measurement Mission Integrated Multi-satellite Final Run V6B (hereafter, IMERG) (Huffman et al., 2015, 2019) and Gauge-calibrated Global Satellite Mapping of Near Real-time Precipitation product version 6 (hereafter, GSMaP) (Kubota et al., 2020). Their main differences lie in algorithm design, input data, parameter estimation, presence of bias correction. More detailed information can be found in Table S1 as well as in some earlier studies (Hsu et al., 1997; Hong et al., 2004; Nguyen et al., 2018, 2019).

The China Meteorological Administration (CMA) observed gridded precipitation dataset with a 0.5° spatial and daily temporal resolution was used in this study as the reference. This dataset was developed by interpolating the high-quality precipitation observations from more than 2400 weather stations over China to $0.5^{\circ} \times 0.5^{\circ}$ grid using the Global 30 Arc Second Elevation Data Set (GTOPO30) and the thin plate smooth spline (TPS) method. To date, this is probably the most accurate gridded precipitation product for China and has been also confirmed to best describing the spatiotemporal distribution of precipitation (Zhao et al., 2014).

The study area is set as the eastern monsoon region of China because of the relatively high density of gauge distribution (see Figure S1; see also in figure 1b in Su et al., 2019). Seventeen years of precipitation data (January 1, 2003 to December 31, 2019) were used for the analysis. And the difference between SPPs and CMA was termed as precipitation bias.

4. Results

The results of our study are divided into three parts. Firstly, we randomly choose one summer (June, July, and August 2012, hereafter JJA12) as an example to highlight the improvement of our new method relative to the old one. In the second part, we show the spatially averaged multi-year average intra-annual variability of different error components for their frequency, magnitude and percentage, as well as a comparison with the classical method. And in the last part, we illustrate the uniqueness of the new method, on the basis of which the characteristics of the spatial (both magnitude and frequency) and seasonal (winter, summer and all seasons) dominant factors of the error components are detected.

4.1. The Enhancement of 4CED

As we mentioned in Section 2, in some cases, the two individual components (positive and negative bias) may cancel each other, resulting in a smaller total bias (ME). The three-component error decomposition method (Tian et al., 2009) was proposed to better understand the error characteristics hidden in the expression of the total error. However, the same situations will occur with the hit bias in the three-component error decomposition. The hit bias was not completely decomposed, because of the fact hit bias is calculated similarly to ME, with both positive and negative cases. Here, we selected JJA12 as a case to test these hypotheses and to illustrate the rationality of the four-component error composition.





Phase 1: Mean error (ME); Mean absolute error (MAE);

Phase 2: Underestimation (UE); Overestimation (OE);

Phase 3: Missed bias (MB); Hit negative bias (HNB); Hit positive bias (HPB); False bias (FB); and Hit bias (HB, Tian et al., 2009)

Figure 2. Accumulative (a) mean error (ME) and mean absolute error (MAE), (b) overestimation (OE) and underestimation (UE), (c) false alarm bias (FB), hit positive bias (HPB), hit negative bias (HNB) and missed bias (MB), and (d) hit bias (HB) of the three datasets for the summer of 2012 (JJA12). HB and HPB, HNB with yellow shade are the main difference between our method and Tian's method.

Figure 2 depicts the different cumulative biases for each product in the JJA12, which are ME bias and MAE bias (Figure 2a), overestimation (OE) and underestimation (UE) (Figure 2b), false bias (FB), hit positive bias (HPB), hit negative bias (HNB), and missed bias (MB) (Figure 2c), and hit bias (HB) (Figure 2d), respective-ly. For the magnitudes of the two forms of bias, the ME for all SSPs ranges from -200 to 200 mm. However, MAE can reach 1000 mm or even higher in some regions. Three datasets share similar spatial characteristics of ME, that is, underestimation in the source of the Yangtze River basin and overestimation in Most of the Pearl River basin. The MAE, on the other hand, exhibits different spatial patterns. The Hengduan Mountains are a hot spot for all SPPs. These differences can be further explained in Figure 2b, where we observe that all products share considerable similarities in their spatial distribution of bias (both underestimation and overestimation) in the Hengduan Mountains. Therefore, this confirms that some large biases may be canceled out resulting in a relatively small total bias. The Hengduan Mountains with high MAE for all SPPs is due to the synergistic effect of the overestimation and underestimation in the same region, but it is not

captured or underestimated by ME. These underestimations are prone to occur for regions with many factors influencing precipitation processes (e.g., coastal areas and high mountains) (Grose et al., 2019; Masson & Frei, 2014; Yao et al., 2016).

Similarly, by further decomposing the biases, especially the hit bias into positive and negative parts using our four-component approach, the sources of errors in SPPs and their magnitudes can be identified and quantified in greater depth. Comparing the FB with HPB, we can see that the FB dominated the precipitation overestimation in the summer of 2012, and this feature is present in all products. This corresponds to a relatively small magnitude of HPB, which only aggravates the total amounts of overestimation. On the other hand, MB and HNB show similar spatial patterns, with MB dominating the underestimation of precipitation in the summer of 2012. However, if we use the old method (Tian et al., 2009), we are only able to obtain HB in Figure 2d and MB and FB in Figure 2c, and cannot further compare and analyze the magnitudes and spatial characteristics of HNB and HPB. On the contrary, our method yields HB as well as its components HPB and HNB, which enables us to obtain a more comprehensive understanding of all error components of precipitation products. Therefore, the 4CED incorporates and outdoes the original error decomposition method (Tian et al., 2009).

In summary, Figure 2 reveals that only using the ME bias and three-component error decomposition may underestimate the magnitudes of errors and may miss some critical error components (HPB and HNB) and regions with relatively poor precipitation forecast skills, such as the Hengduan Mountains. In contrast, the use of the newly proposed method not only yields the results of ME and MAE, but also provides a more comprehensive description of the spatial distribution and magnitudes of the individual error components.

4.2. Temporal Characteristics of Error Components

To explore the variation of errors over time, especially the differences between seasons, we calculated the spatially averaged and multi-year average (2003–2019) intra-annual variation of the error components for three SPPs (Figure 3). A 31-day moving average was used for each time series to smooth and reduce visual cluttering.

For different event frequencies (Figure 3a), all SPPs were able to detect no rain events (CN) more accurately. At the same time, most of the products were able to achieve a high probability of rain/no rain recognition without considering hit bias. Especially in the winter season, the probability can reach 70% or even higher. However, when we analyze the error component in addition to correct negative events (CN), the old method (Tian et al., 2009) can only give the total number of hit events, which is the sum of two events (hit positive and hit negative events in 4CED) and certainly accounts for a higher percentage. But in fact, among the four error events, FB is the one that occurs most frequently. Meanwhile, we are also able to split the proportion of hit positive and negative events, respectively. In addition to correct negative events (CN), the most frequent events were false alarm precipitation. And there was also a proportion of missed precipitation, which occurred more frequently in summer.

Figure 3b depicts the seasonal variation in the magnitudes of ME, HB, HPB, HNB, FB, and MB. With the 4CED, we can capture more error components (HPB and HNB) and recognize their interactions. This has been ignored in previous studies. It is easy to obtain a relatively small ME bias when the magnitudes of positive and negative values in the error components are about at the same amplitude, which may mislead us to think that a sufficiently accurate precipitation estimate is obtained, especially after averaging during a long-term period. But, especially in daily precipitation estimates, the bias of SPPs may still be very large, which makes it very unfavorable when we engage in daily planning-related applications. Similarly, a very small hit bias is obtained when the HPB and HNB are canceled out by each other. Using the three-components error decomposition, the hit bias is one of the final error components, so the HPB and HNB are missed and their magnitudes are unknown to us.

Figure 3c depicts the seasonal variations in the contributions of different error components, from which the dominant factors can be discerned. For example, for all SPPs, the bias in summer is dominated by the HPB, with a contribution greater than 25%, leading to positive ME and HB in Figure 3b. On the contrary, for winter, the dominant factor is the MB, with a contribution close to 50%, resulting in negative ME bias. The relative contribution of the four components and their seasonal variations can yield new insights into





Figure 3. Spatially averaged and multi-year average (2003–2019) intra-annual variation (a) frequency of different events, (b) magnitude and (c) percentage of the error components (EC) for eight SPPs. FB: false alarm bias; HPB: hit positive bias; HNB: hit negative bias; MB: missed bias; CN: correct negative events; HB: hit bias; ME: mean error.

sources of error in precipitation products, especially the positive and negative parts of the hit events. This has not been mentioned in previous studies (Tian et al., 2009). And a combined value (HB with positive and negative) cannot be used with MB and FB to calculate relative contributions. Recognizing these new error components helps algorithm developers and data producers to further correct and reduce errors of precipitation products.

4.3. Spatial Characteristics of Error Components

As we have emphasized above, hit bias (HB), like ME, is a combined value, and it could be positive or negative. Therefore, even if a three-component error decomposition method is used, it is not appropriate to compare HB with MB and FB to identify the dominant factor of the error source. With the 4CED, however, each component is thoroughly decomposed and their positive and negative signs are constant, so we can take their absolute values and compare them with each other to determine the true dominant factor of the error source. For example, to compare the spatial characteristics of error components for three SPPs during the whole study period (2003–2019) and differences between seasons, we calculated the cumulative values





Figure 4. Spatial distribution of dominant factor of error components considering their (a) amounts and (b) frequency for all seasons, summer and winter during the whole study period (2003–2019). FB: false alarm bias; HPB: hit positive bias; HNB: hit negative bias; MB: missed bias.

(both magnitude and frequency) of all error components for each day to three different time scales (all seasons, summer and winter). Finally, both the total amounts and frequency of different error components were ranked for each grid, and the highest-ranked error component is seen as the dominant factor for the specific grid. The spatial distributions of dominant factors are shown in Figure 4.

For the winter season, all products exhibit similar spatial characteristics in both quantity and frequency. The FB and MB are the two dominant factors for all products (except for IMERG in southeastern China). And the MB is the dominant error source in South China for all SPPs. On other hand, there exist large differences in North China between the three datasets. Among them, GSMaP is dominated by FB in the Yellow River basin and source of the Yangtze River basin, while in these regions, PDIR and IMERG are dominated by MB. The spatial characteristics of the dominant factor for IMERG are similar to those of PDIR, but it

appears that the magnitude of HPB for IMERG in the southeast coastal area is much larger than MB. For the summer season, the dominant factors are more diverse. Four different error components-dominated regions exist for all products, but they have different spatial distributions and coverage. Moreover, there exist some regions with low frequency but high magnitude. For example, for IMERG and GSMaP, the HPB dominated most regions for the frequency of error components, where are dominated by MB for the magnitudes of error components (e.g., SongLiao River basin). Consequently, the dominant factors for the frequency of error components and their spatial distribution in all seasons are largely consistent with the winter season, while that for the magnitude of error components are closer to the spatial patterns of the summer season. This is possibly caused by the higher proportion of heavy rainfall events in summer. The interaction of these error components presents an enhancement of the total bias in some cases, while in other cases they cancel each other out, resulting in a total bias that is, much smaller than individual components.

Once again we would like to express that, compared to the three-component decomposition method (Tian et al., 2009), our method is able to directly compare all the independent error components of the full decomposition and determine their dominance as well as the temporal and spatial characteristics of the occurrence.

5. Discussion and Conclusions

Identifying and quantifying error sources of precipitation products can not only provide references for future algorithm research, but also serve users to better understand the advantages and limitations of these datasets. In this study, we propose a new framework for error decomposition of precipitation products, which can cover not only all components obtained from a previous study (Tian et al., 2009), but also provides new insights into the error sources, especially the magnitude of the hit bias (both positive and negative). The new framework named Four-component error decomposition (4CED). The three state-of-the-art products we selected include the latest PDIR, IMERG, and GSMaP. The evaluation of these products also confirms the validity of our proposed framework and how it complements the previous method in tracking and quantifying error sources. Furthermore, although three satellite precipitation estimates products were selected in this study for a daily-scale case study in the eastern monsoon region of China, we believe that the 4CED could be applied to more precipitation products (including reanalysis and merge products), as well as to different spatial (from pixel to watershed and even global) and temporal (e.g., hourly) scales.

There are many statistical metrics in the existing literature to describe and assess the error of precipitation products. In this study, we illustrate the differences and connections between 4CED and two traditional metrics (ME and MAE), as well as the enhancement of 4CED compared to traditional error decomposition method. We also summarized a series of commonly used metrics in Table S2. These individual indicators can be used to describe only one or several aspects of the precipitation products. For example, correlation coefficient (R) can only describe the correlation between two series and cannot quantify the error between them. Similarly, probability of detection (POD) is a probabilistic indicator and is only used for the assessment of rain (or no-rain). It cannot quantitatively describe the error of the products. Mean error (ME) is a combination of positive and negative values that calculates only the average bias between two series. It describes the systematic bias between the two series, which leads to an underestimation of the magnitude of the precipitation error, especially on a specific day. Mean absolute error (MAE) measures the average magnitude of the errors without considering their direction. Root mean squared error (RMSE) is a quadratic scoring rule that also measures the average magnitude of the error. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors. Metrics like Nash-Sutcliffe efficiency (NSE) and Kling-Gupta efficiency (KGE) are integrated scoring rules that allow us to better compare the relative goodness of different products to each other, but they simultaneously lose the ability to track and analyze sources of error. Therefore, depending on the needs and the purpose of the application, end-users use multiple metrics in one case to more fully describe and characterize the different features of a product. However, the above metrics are usually only used to analyze the relative goodness of a product and require great care when interpreting the results (more discussion and examples see Text S2). In contrast, the error decomposition method proposed by Tian et al. (2009) or AghaKouchak et al. (2012) does not only evaluate the product, but provides a better analysis of the error sources for precipitation products. Therefore, it is important to highlight that the 4CED is a new member of this family of scoring rules, and we aim to use it

to better analyze the error sources and improve the quality of precipitation products based on different error sources. The improved error decomposition scheme (Chaudhary & Dhanya, 2021) is a very similar with 4CED, which was simultaneously validated and emphasized in India. This study chose some different precipitation products and study area, which also proves the generality of the methodology. This study closely compared our method with previous method (Tian et al., 2009). More importantly, This study elaborated the differences and connections between traditional scoring metrics and 4CED. 4CED is applicable to the error analysis of all types of precipitation products, which allows the method to be disseminated more widely.

The reader should be reminded of the strategy used in our analysis of the different error components over time in Section 3.2. Because our study period covers a relatively long period, we averaged the multi-year data to obtain an intra-annual variation. However, there may be some year-to-year or interannual variation in the accuracy and error features of SPPs over time. And these features may reflect improvements in terms of updates to the algorithms, or changes in available data sources. In future studies, we can further partition the study period and compare the characteristics of the errors over different study periods (Shen et al., 2020; Tang et al., 2020). Spatial averaging poses similar problems, but our approach is generalized to different spatial and temporal scales. Another issue is that all comparisons in this study were performed at a daily and $0.5^{\circ} \times 0.5^{\circ}$ grid due to the fixed resolution of the reference data. Therefore, the results presented may not fully demonstrate the advantages of some of the products with their original intentions. For example, PDIR is a product with the highest spatial resolution and the shortest lag time among all SPPs. And such a product may help us to focus on more regionalized precipitation features and near real-time applications, especially in regions with sparse stations. The same goes for the highest temporal resolution of the IMERG dataset. All these issues need to be further investigated in the future.

Data Availability Statement

The CMA reference dataset is publicly available at National Tibetan Plateau Third Pole Environment Data Center of China via (http://data.tpdc.ac.cn/en/data/653191c3-df99-479b-95bd-a18bee30f2bd/) provided by National Meteorological Information Center of China Meteorological Administration. PDIR dataset can be download from CHRS Data Portal (http://chrsdata.eng.uci.edu/); to access data used here, go to PDIR-Now data. GSMaP dataset is publicly available at (https://sharaku.eorc.jaxa.jp/GSMaP/index.htm); to access data used here, go to Global Rainfall Watch, GSMaP_Gauge_NRT, Gauge-calibrated Near-Realtime, v6, real-time_ver/v6/daily_G data. GPM IMERG Final Precipitation (IMERG Final Run) L3 1 day 0.1 degree × 0.1 degree V06 dataset is publicly available at (https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGDF_06/summary?keywords=%22IMERG%20final%22).

References

- AghaKouchak, A., Mehran, A., Norouzi, H., & Behrangi, A. (2012). Systematic and random error components in satellite precipitation data sets. *Geophysical Research Letters*, 39(9), L09406. https://doi.org/10.1029/2012GL051592
- Alijanian, M., Rakhshandehroo, G. R., Mishra, A., & Dehghani, M. (2019). Evaluation of remotely sensed precipitation estimates using PERSIANN-CDR and MSWEP for spatio-temporal drought assessment over Iran. *Journal of Hydrology*, 579, 124189. https://doi. org/10.1016/j.jhydrol.2019.124189
- Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., van Dijk, A. I. J. M., et al. (2019). Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. *Hydrology and Earth System Sciences*, 23(1), 207–224. https://doi.org/10.5194/ hess-23-207-2019
- 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
- Chaudhary, S., & Dhanya, C. T. (2021). An improved error decomposition scheme for satellite-based precipitation products. *Journal of Hydrology*, 598, 126434. https://doi.org/10.1016/j.jhydrol.2021.126434
- Chen, H., Yong, B., Gourley, J. J., Liu, J., Ren, L., Wang, W., et al. (2019). Impact of the crucial geographic and climatic factors on the input source errors of GPM-based global satellite precipitation estimates. *Journal of Hydrology*, 575, 1–16. https://doi.org/10.1016/j. jhydrol.2019.05.020
- Chen, S., Liu, B., Tan, X., & Wu, Y. (2020). Inter-comparison of spatiotemporal features of precipitation extremes within six daily precipitation products. *Climate Dynamics*, 54(1–2), 1057–1076. https://doi.org/10.1007/s00382-019-05045-z
- Do, H. X., Mei, Y., & Gronewold, A. D. (2020). To what extent are changes in flood magnitude related to changes in precipitation extremes? *Geophysical Research Letters*, 47(18), e2020GL088684. https://doi.org/10.1029/2020GL088684

Acknowledgments

This work is jointly supported by the National Key Research and Development Program of China (No. 2018YFE196000), U.S. Department of Energy (DOE Prime Award DE-IA0000018), the Natural Science Foundation of China (No. 51879009), the Second Tibetan Plateau Scientific Expedition and Research Program (No. 2019QZKK0405).



- Foufoula-Georgiou, E., Guilloteau, C., Nguyen, P., Aghakouchak, A., Hsu, K., Busalacchi, A., et al. (2020). Advancing precipitation estimation, prediction, and impact studies. *Bulletin of the American Meteorological Society*, 101(9), E1584–E1592. https://doi.org/10.1175/ BAMS-D-20-0014.1
- Grose, M. R., Syktus, J., Thatcher, M., Evans, J. P., Ji, F., Rafter, T., & Remenyi, T. (2019). The role of topography on projected rainfall change in mid-latitude mountain regions. *Climate Dynamics*, 53(5), 3675–3690. https://doi.org/10.1007/s00382-019-04736-x
- Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of hydrology*, 377(1), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003
- Hong, Y., Hsu, K., Sorooshian, S., & Gao, X. (2004). Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *Journal of Applied Meteorology* (1988), 43(12), 1834–1853. https://doi.org/10.1175/jam2173.1
- Hsu, K., Gao, X., Sorooshian, S., & Gupta, H. V. (1997). Precipitation estimation from remotely sensed information using artificial neural networks. *Journal of Applied Meteorology*, 36(9), 1176–1190. https://doi.org/10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2
 Huffman, G. J., Bolvin, D. T., & Nelkin, E. J. (2015). *Integrated multi-satellitE retrievals for GPM (IMERG) Technical Documentation*. Re-
- trieved from http://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc.pdf Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., & Tan, J. (2019). *GPM IMERG final precipitation L3 1 day 0.1 degree* × 0.1 *degree*
- V06 (In A. Savtchenko (Eds.)). Goddard Earth Sciences Data and Information Services Center (GES DISC). https://doi.org/10.5067/ GPM/IMERGDF/DAY/06
- Jiang, L., & Bauer-Gottwein, P. (2019). How do GPM IMERG precipitation estimates perform as hydrological model forcing? Evaluation for 300 catchments across Mainland China. *Journal of Hydrology*, 572, 486–500. https://doi.org/10.1016/j.jhydrol.2019.03.042
- Kubota, T., Aonashi, K., Ushio, T., Shige, S., Takayabu, Y. N., Kachi, M., et al. (2020). Global satellite mapping of precipitation (GSMaP) products in the GPM era (pp. 355–373). Springer. https://doi.org/10.1007/978-3-030-24568-9_20
- Liu, X., Yang, T., Hsu, K., Liu, C., & Sorooshian, S., & Univ. Of California, I. C. U. S. (2017). Evaluating the streamflow simulation capability of PERSIANN-CDR daily rainfall products in two river basins on the Tibetan Plateau. *Hydrology and Earth System Sciences*, 21(1), 169–181. https://doi.org/10.5194/hess-21-169-2017
- Massari, C., Crow, W., & Brocca, L. (2017). An assessment of the performance of global rainfall estimates without ground-based observations. *Hydrology and Earth System Sciences*, 21(9), 4347–4361. https://doi.org/10.5194/hess-21-4347-2017
- Masson, D., & Frei, C. (2014). Spatial analysis of precipitation in a high-mountain region: Exploring methods with multi-scale topographic predictors and circulation types. *Hydrology and Earth System Sciences*, 18(11), 4543–4563. https://doi.org/10.5194/hess-18-4543-2014
- Nguyen, P., Ombadi, M., Gorooh, V. A., Shearer, E. J., Sadeghi, M., Sorooshian, S., et al. (2020). PERSIANN dynamic infrared-rain rate (PDIR-Now): A near-real-time, quasi-global satellite precipitation dataset. *Journal of Hydrometeorology*, 21(12), 2893–2906. https://doi.org/10.1175/JHM-D-20-0177.1
- Nguyen, P., Ombadi, M., Sorooshian, S., Hsu, K., AghaKouchak, A., Braithwaite, D., et al. (2018). The PERSIANN family of global satellite precipitation data: A review and evaluation of products. *Hydrology and Earth System Sciences*, 22(11), 5801–5816. https://doi. org/10.5194/hess-22-5801-2018
- Nguyen, P., Shearer, E. J., Ombadi, M., Gorooh, V. A., Hsu, K., Sorooshian, S., et al. (2020). PERSIANN dynamic infrared-rain rate model (PDIR) for high-resolution, real-time satellite precipitation estimation. *Bulletin of the American Meteorological Society*, *101*(3), E286–E302. https://doi.org/10.1175/BAMS-D-19-0118.1
- Nguyen, P., Shearer, E. J., Tran, H., Ombadi, M., Hayatbini, N., Palacios, T., et al. (2019). The CHRS data portal, an easily accessible public repository for PERSIANN global satellite precipitation data. *Scientific Data*, *6*(1), 180296. https://doi.org/10.1038/sdata.2018.296
- 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
- Shen, Y., Zhao, P., Pan, Y., & Yu, J. (2014). A high spatiotemporal gauge-satellite merged precipitation analysis over China. Journal of Geophysical Research: Atmospheres, 119(6), 3063–3075. https://doi.org/10.1002/2013JD020686
- Shen, Z., Yong, B., Gourley, J. J., Qi, W., Lu, D., Liu, J., et al. (2020). Recent global performance of the climate hazards group infrared precipitation (CHIRP) with stations (CHIRPS). *Journal of Hydrology*, 591, 125284. https://doi.org/10.1016/j.jhydrol.2020.125284
- Sorooshian, S., Hsu, K., Gao, X., Gupta, H. V., Imam, B., & Braithwaite, D. (2000). Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bulletin of the American Meteorological Society*, 81(9), 2035–2046. https://doi.org/10.1175/1520-0477(2000)08 1<2035:EOPSSE>2.3.CO;2
- 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(12), 2461–2479. https://doi.org/10.1029/2019EA000977
- Su, J., Lü, 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. https://doi.org/10.3390/rs10091420
- Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K. L. (2018). A Review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, 56(1), 79–107. https://doi.org/10.1002/2017RG000574
- Tang, G., Clark, M. P., Papalexiou, S. M., Ma, Z., & Hong, Y. (2020). Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets. *Remote Sensing of Environment*, 240, 111697. https://doi.org/10.1016/j.rse.2020.111697

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. https://doi.org/10.1016/j.jhydrol.2015.12.008

- Tian, Y., Peters-Lidard, C. D., Eylander, J. B., Joyce, R. J., Huffman, G. J., Adler, R. F., et al. (2009). Component analysis of errors in satellite-based precipitation estimates. *Journal of Geophysical Research*, 114(D24), D24101. https://doi.org/10.1029/2009JD011949
- Yao, J., Yang, Q., Mao, W., Zhao, Y., & Xu, X. (2016). Precipitation trend-elevation relationship in arid regions of the China. *Global and Planetary Change*, 143, 1–9. https://doi.org/10.1016/j.gloplacha.2016.05.007
- Yao, L., Libera, D. A., Kheimi, M., Sankarasubramanian, A., & Wang, D. (2020). The roles of climate forcing and its variability on streamflow at daily, monthly, annual, and long-term scales. *Water Resources Research*, 56(7), e2020WR027111. https://doi.org/10.1029/2020WR027111
 Zhao, Y., Zhu, J., & Xu, Y. (2014). Establishment and assessment of the grid precipitation datasets in China for recent 50 years. *Journal of*
- the Meteorological Sciences, 34(4), 414–420. (In Chinese). https://doi.org/10.3969/2013jms.0008 Zhou, X., Polcher, J., Yang, T., & Huang, C. (2020). A new uncertainty estimation approach with multiple datasets and implementation for various precipitation products. *Hydrology and Earth System Sciences*, 24(4), 2061–2081. https://doi.org/10.5194/hess-24-2061-2020