The Effect of Spatiotemporal Resolution Degradation on the Accuracy of IMERG Products over the Huai River Basin

JIANBIN SU

State 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

HAISHEN LÜ

State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, National Cooperative Innovation Center for Water Safety and Hydroscience, College of Hydrology and Water Resources, and Joint International Research Laboratory of Global Change and Water Cycle, Hohai University, Nanjing, China

WADE T. CROW

USDA–ARS Hydrology and Remote Sensing Laboratory, Beltsville, Maryland

YONGHUA ZHU

State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, National Cooperative Innovation Center for Water Safety and Hydroscience, College of Hydrology and Water Resources, and Joint International Research Laboratory of Global Change and Water Cycle, Hohai University, Nanjing, China

YIFAN CUI

Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China

(Manuscript received 22 July 2019, in final form 24 February 2020)

ABSTRACT

The rapid development of the Integrated Multisatellite Retrievals for Global Precipitation Measurement (IMERG) precipitation product provides new opportunities for a wide range of Earth system and natural hazard applications. Spatiotemporal averaging is a common method for IMERG users to acquire suitable resolutions specific to their research or application purpose and has a direct impact on the overall quality of IMERG precipitation estimates. Here, three different IMERG, version 06 (V06), latency run products (i.e., early, late, and final) are assessed against a ground-based benchmark along a continuous series of spatiotemporal resolutions over the Huai River basin (HuaiRB) between June 2014 and May 2017. In general, IMERG products better capture the spatial pattern of precipitation, and demonstrate better reliability, in the southern portion of the HuaiRB relative to its northern region. Furthermore, the degradation of spatiotemporal resolution is associated with better rain/no-rain determination and the consistent improvement of rainfall product performance metrics. This improvement is more pronounced for IMERG products at fine spatiotemporal resolution. However, due to the presence of autocorrelated errors, the performance improvement associated with the degradation of spatiotemporal resolution is less than theoretical expectations assuming purely uncorrelated errors. Component analysis indicates that while both temporal and spatial aggregation do not mitigate temporally autocorrelated errors, temporal averaging can remove spatially autocorrelated error. Hence, temporal averaging is found to be more effective than spatial averaging for improving the quality of IMERG products. These results will inform users of the reliability of IMERG products at different spatiotemporal scales and assist in unifying former disparate validation assessments applied at different scales within the literature.

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JHM-D-19-0158.s1.

Corresponding author: Haishen Lü, lvhaishen@hhu.edu.cn

DOI: 10.1175/JHM-D-19-0158.1

1. Introduction

Precipitation, as a fundamental component of the global water and energy cycles, provides a primary source of freshwater essential to life on Earth (Kidd et al. 2009). Since precipitation plays a crucial role in water resource management and agribusinesses, it also has important socioeconomic value (Kidd et al. 2009). Too much rainfall causes floods, endangering life and property, whereas too little may generate droughts and degrade agricultural productivity. Hence, comprehensive knowledge of the occurrence and distribution of precipitation is crucial for the forecasting, detection, and mitigation of environmental disasters (Kidd and Huffman 2011). However, obtaining accurate precipitation estimations has always been challenging for scientists, especially for remote and inaccessible areas (Tang et al. 2016b).

Currently, there are three mainstream approaches for measuring precipitation: rain gauges, ground-based weather radars, and satellite-based remote sensing. Rain gauges generally provide the most direct and reliable source and usually serve as the benchmark for assessing radar-based (Li et al. 2017; Wang et al. 2008; Xie et al. 2011) and satellite-based precipitation estimations (SPEs) (Su et al. 2017; Wei et al. 2018; Zhao et al. 2018). However, to obtain details on the spatial variability of precipitation, very dense rain gauge networks are required. This requirement is unrealistic for most developing countries due to high upfront costs and maintenance difficulties (Su et al. 2018). Likewise, ground-based weather radar systems are very valuable for providing real-time precipitation monitoring at high spatiotemporal resolution. However, they suffer from their own shortcomings including limited and irregular spatial coverage, variable reliability, and poor utility in cold weather and mountainous terrains (Dinku et al. 2002; Tang et al. 2016b). In contrast, the remote sensing retrieval from Earth observation satellites is the only practical way to achieve continuous global precipitation estimations.

Since the Tropical Rainfall Measuring Mission (TRMM) satellite was launched as the first dedicated meteorological precipitation satellite in 1997, several quasi-global, TRMM-based SPEs have been released publicly. The wide-scale application of these open-access SPEs in different fields (e.g., disaster monitoring and forecasting and water resources management) has yielded substantial scientific and societal benefits. After over 17 years of productive data gathering, the TRMM satellite reentered Earth's atmosphere on 15 June 2015. Nevertheless, the TRMM era achieved great success by not only providing vast volumes of quasi-global precipitation data but also accumulating adequate instrument and algorithm expertise for the follow-up Global Precipitation Measurement (GPM) mission. The GPM mission is an international constellation of satellites that provides next-generation global observations of rain and snow accumulation rates. Building upon the success of the TRMM, the GPM mission consists of one "core" satellite (*Core Observatory*) and approximately 10 partner satellites. The *Core Observatory* satellite is equipped with the first-ever spaceborne Dual-Frequency Precipitation Radar (DPR; the Ku band at 13.6 GHz and Ka band at 35.5 GHz) and a multichannel GPM Microwave Imager (GMI; frequency ranging from 10 to 183 GHz). The *Core Observatory* satellite is designed to measure precipitation from space and serves as a reference standard to unify precipitation measurements obtained from a constellation of other research and operational satellites.

As the level 3 product of the GPM mission, the Integrated Multisatellite Retrievals for GPM (IMERG), version 06 (V06), product provides global (90°S–90°N) precipitation estimations at fine spatial ($0.1° \times 0.1°$) and temporal (0.5 h) resolution via the unified U.S. algorithm (Huffman et al. 2019). The IMERG system runs three separate analyses for each observation time and provides three different estimates—each accommodating different user requirements for latency and accuracy. Although the IMERG is designed to be the best estimator of global precipitation, it still suffers from various sources of systematic biases and random errors (Guo et al. 2016; Su et al. 2018; Tang et al. 2016b). This uncertainty is more pronounced for subdaily products (Asong et al. 2017; Sharifi et al. 2018; Wu et al. 2019). Therefore, thorough validation, particularly at subdaily time scales, is necessary before IMERG products can be regionally applied with confidence.

To this end, a number of validation efforts have been conducted to statistically evaluate IMERG products using gauge-based observations acquired in India (Prakash et al. 2016), West and East Africa (Dezfuli et al. 2017), northeast Austria (Sharifi et al. 2018), Iran (Sharifi et al. 2016), the United States (Tan et al. 2017), Singapore (Tan and Duan 2017), and mainland China (Guo et al. 2016; Su et al. 2018; Tang et al. 2016b; Zhao et al. 2018). Meanwhile, other studies have focused on the reliability of IMERG products for detecting extreme precipitation (e.g., Huang et al. 2019; Omranian et al. 2018) and providing precipitation forcing data for hydrological models (see, e.g., He et al. 2017; Su et al. 2019b; Yuan et al. 2018; Wang et al. 2017). Most of these studies have found that IMERG products are generally superior to their predecessor SPEs and therefore

provide superior precipitation forcing for hydrological simulations.

However, these prior assessments have been performed at a specific set of temporal and spatial resolutions, and relatively little study has focused on how the quality of IMERG products varies across a range of spatiotemporal scales. One exception is Tang et al. (2016b), who assessed IMERG products at multiple time scales and their native spatial resolution (0.1° × 0.1°) and suggested that the statistical metrics of IMERG improved significantly as accumulation time scales are increased from hourly to 3 hourly and daily. This is typically the case as temporal aggregation tends to cancel out temporally uncorrelated retrieval errors. Analogously, spatial averaging is also expected to neutralize spatially random errors. However, the exact impact of spatiotemporal resolution degradation on the accuracy of IMERG products has not been extensively studied.

Besides, with the rapid development of satellite observation and retrieval techniques, the available spatial and temporal resolution of SPEs is being progressively refined (e.g., from 3 hourly and 0.25° × 0.25° resolution during the TRMM era to half hourly and 0.1° × 0.1° within the GPM era). These refinements have made SPE's more valuable for applications (e.g., hurricane monitoring, flood forecasting, and landslide prediction) requiring fine spatiotemporal resolutions. At the same time, many applications remain that can utilize coarse spatiotemporal data (e.g., water resource management and regional drought monitoring)—provided they are sufficiently accurate. Spatial and temporal averaging is a common method to acquire precipitation at a suitable resolution from high-resolution SPEs. However, selecting the appropriate spatial and temporal resolution for SPEs to accommodate both resolution and accuracy user requirements is a meaningful topic.

Therefore, the objectives of this study are twofold: 1) to evaluate the three runs of products of IMERG V06 against the ground-based observations and 2) to examine the effect of spatial and temporal degradation on accuracy measures. The Huai River basin (HuaiRB), a midlatitude basin that is frequently affected by floods and droughts, has been selected as the study area. This study will reveal the error features of IMERG products over the HuaiRB and provide useful guidelines for potential IMERG users. In addition, the sensitivity of various performance metrics to spatial and temporal resolution degradation is examined. These results are valuable for unifying assessment results of IMERG previously acquired at different spatiotemporal scales and providing end users with quantitative performance metrics at scale(s) suitable for their specific application.

Accordingly, this paper is structured as follows: Section 2 introduces the study area, precipitation datasets, and performance metrics utilized in the analysis. Sections 3 and 4 present the main results and associated discussion, and these results are summarized in section 5.

2. Study area, datasets, and methodology

a. Study area

The HuaiRB covers a drainage area of approximately 270 000 km^2 between the Yellow River and the Yangtze River basins in eastern China (Fig. 1a) and is bounded by the box 111.93°–121.38°E and 30.95°–36.31°N. The terrain of the HuaiRB is relatively flat. About two-thirds of it is a vast expanse of plains, and the rest is the low mountains, hills, and highlands along with the western, southern, and northern edges of the basin. Based on water resources and river classifications (http://www.resdc.cn), the HuaiRB can be divided into four subcatchments (i.e., the upstream, midstream, downstream and Yishusi basins; see Fig. 1). Here, we will use this classification to organize the subbasin evaluation of IMERG products. The number of native IMERG grids within each of these four subcatchments is 340, 1360, 352, and 865, respectively.

The HuaiRB is situated in the climate transition zone between the humid southern region of China and the semihumid northern region and is chiefly characterized by monsoonal weather conditions (Pan et al. 2017). Note that, following Yong et al. (2016), the division between humid and semihumid climate is based on a threshold of 800 mm for mean annual precipitation (as shown in Fig. 1). Mean annual temperatures in the HuaiRB range from 11° to 16°C, with the highest monthly mean temperature (25°C) in July and the lowest (0°C) in January (M. Yang et al. 2016). Precipitation in the HuaiRB strongly varies both monthly and annually due to the unstable duration, intensity, and location of the subtropical high over the northern Pacific. The average annual precipitation of the HuaiRB is about 833 mm but ranges from less than 600 mm in the northwest part of the basin to more than 1400 mm in the southeast (M. Yang et al. 2016). Most annual precipitation (60%–70%) occurs between June and August (i.e., the flood season). Frequent rainstorms during the flood season often generate basin-scale flooding because of slow-moving surface water flow across flat terrain within the midstream and downstream portions of the basin. Therefore, within the HuaiRB, the timely and accurate monitoring of precipitation, as well as its spatiotemporal distribution, is critical for hydrological and meteorological prediction and natural hazard monitoring.

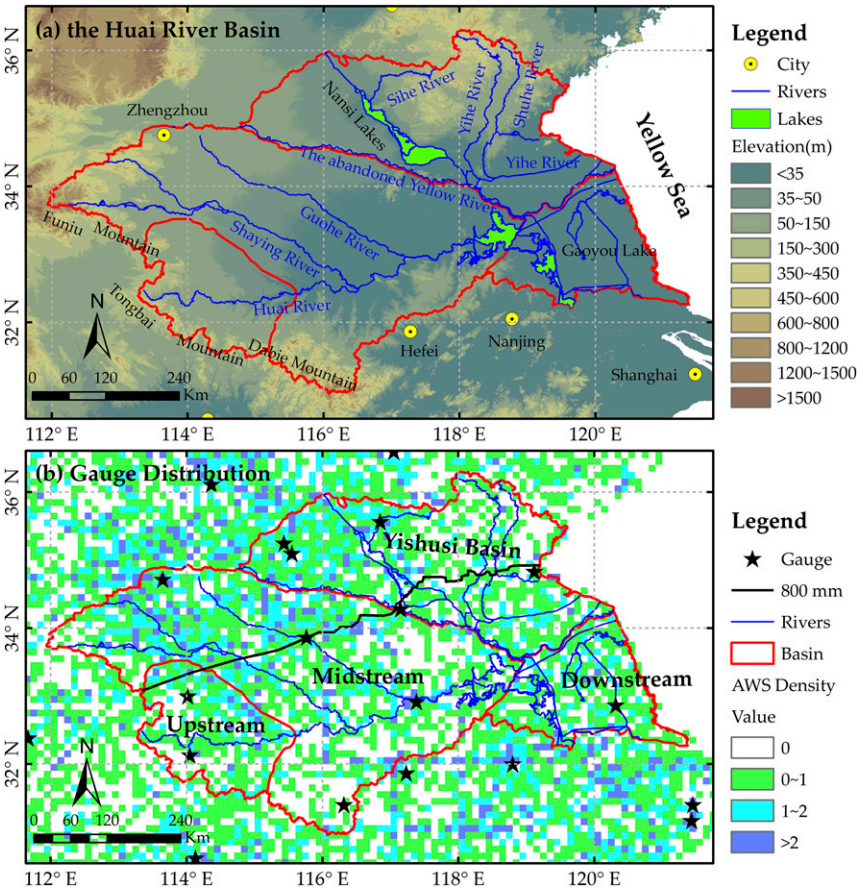

FIG. 1. (a) A close-up map of the Huai River basin showing major rivers and elevation differences and (b) the spatial distribution of GPCC rain gauges (stars) and the spatial density of AWS stations used to produce the CMPA.

b. *Ground reference data*

To validate the quality of IMERG products at different spatiotemporal resolutions, a high spatial (0.1° × 0.1°) and temporal (hourly) resolution gauge-satellite merge precipitation product, the China Merge Precipitation Analysis Hourly, version 1.0 (V1.0), product (CMPA; Shen et al. 2014), was selected as the benchmark. These data were developed by the National Meteorological Information Center and are publicly available on the official website of the China Metrological Administration (http://data.cma.cn). It was produced by merging ground data from over 30 000 automatic weather stations (AWSs) across China with precipitation estimates from the satellite-based Climate Prediction Center (CPC) morphing technique (CMORPH) product via an improved probability density function–optimal interpolation method (PDF-OI). Verification results demonstrated that the CMPA performs reasonably well in China and can capture time-varying features of hourly precipitation during storm events (Shen and Xiong 2016) owing to the high density of the AWSs within the basin (see Fig. 1b) and the rigorous quality control of these source data. An independence assessment based on approximately 2400 national weather stations showed that the monthly spatial distribution of precipitation accumulation provided by the CMPA is very similar to the China gauge-based daily precipitation analysis (Shen et al. 2014). Moreover, the AWSs employed in developing the CMPA are not used for the bias correction of IMERG products. Therefore, the CMPA can act as an objective, independent reference for evaluating IMERG products. Note that the CMPA has already been widely applied in past assessments of SPEs (Su et al. 2018; Tang et al. 2017; Wang et al. 2017).

c. *IMERG products*

To cope with the increased likelihood of extreme weather under climate change scenarios, the GPM mission was designed to provide improved understanding (and therefore forecasting) of extreme weather events

and systems. Based on the GPM constellation, IMERG products provide global precipitation estimations at relatively fine spatiotemporal resolution (0.1° and 0.5 h). In particular, the IMERG system was designed to intercalibrate, merge, and interpolate precipitation microwave estimates (obtained within the entire GPM constellation) with (microwave calibrated) infrared satellite estimates, rain gauge analyses, and all other available precipitation estimators (Huffman et al. 2019). To this end, the IMERG processor provides three different IMERG products—each accommodating a different set of user data latency and accuracy requirements.

The ''early'' run (IMERG-E) gives a quick estimate of precipitation with an approximate 4-h delay and is particularly suitable for real-time applications like short-term hazard prediction. The ''late'' run (IMERG-L) has an 18-h delay but provides improved estimates of precipitation intensity as more remote sensing data can be collected and applied by IMERG during its longer data latency period. The ''final'' run (IMERG-F) is released approximately 3.5 months after the observation time to create a research-level product by maximally leveraging the (relatively) slow collection of ground-based gauge rain data. Note that both IMERG-E and IMERG-L are near-real-time products that adopt climatological gauge adjustments while IMERG-F is a post-real-time product calibrated against the Global Precipitation Climatology Centre (GPCC) monthly gauge analyses. For a detailed description of all three runs, please see Huffman et al. (2019). Weather stations utilized by the GPCC project within and around the HuaiRB are displayed in Fig. 1b.

Currently, the IMERG products have been updated several times, that is, V03, V04, V05, and V06. Here, all three runs of the latest IMERG version (V06) are employed. These datasets are publicly available on the NASA Precipitation Processing System (https://pmm.nasa.gov/data-access/downloads/gpm). Note, like IMERG, the hourly CMPA product is provided in coordinated universal time (UTC). Hence, IMERG products are first averaged to an hourly scale, matching the temporal resolution of CMPA. Subsequently, the same upscaling operation is applied to both IMERG and CMPA products to isolate the impact of spatiotemporal resolution degradation on the accuracy of IMERG products.

d. Methodology

For a comprehensive quality assessment of the IMERG products, two categories of statistical metrics are selected. The first category contains widely applied basic statistical indices including Pearson correlation coefficient (CC), mean absolute error (MAE), relative bias (Rbias), and root-mean-square error (RMSE). CC describes the linear agreement between SPEs and ground references, while the MAE, Rbias, and RMSE metrics are used to describe the magnitude of error and bias (Yong et al. 2010). The second category consists of categorical statistical indices including the probability of detection (POD), the false alarm ratio (FAR), and the critical success index (CSI). The categorical statistical indices describe the consistency between SPEs and observed rain event occurrence (Omranian and Sharif 2018). According to Tang et al. (2016b), given the rain gauge's detection resolution, a small value (i.e., 0.1 mm h^{-1}) is selected as the rain/no-rain threshold to calculate the categorical statistical indices over the HuaiRB. Detailed formulas and perfect values for all statistical metrics are listed in Table 1.

Simple spatiotemporal averaging is employed to obtain precipitation estimates at relatively coarser temporal and spatial resolutions. The abovementioned statistical metrics are independently sampled for each upscaled IMERG product. In this way, the effects of depredating spatiotemporal resolution on the accuracy of IMERG products are examined. Note that we are working with precipitation rates and not precipitation accumulation. In other words, at all spatiotemporal resolutions, precipitation is measured in millimeters per hour. Hence, a constant rain/no-rain threshold of 0.1 mm h^{-1} is adopted for rain occurrences.

The presence of autocorrelated errors in IMERG products will tend to weaken the effect of spatiotemporal aggregation. Therefore, using RMSE as a measure of error level, an extreme end case of purely uncorrelated errors is considered. In this scenario, errors in IMERG products are assumed to be temporally and spatially uncorrelated. Based on the central limit theorem, it follows immediately that the standard deviation of a sample mean is the population standard deviation divided by $\sqrt{N}$, where N is the number of independent samples (Wilks 2011). In our case of assumed purely unrelated errors, the theoretical RMSE values of IMERG products at the resolution $0.l° \times 0.l°$ and t hours are thus calculated as

$$\mathrm{RMSE}(l,t) = \frac{\mathrm{RMSE}(0.1°, 1\,\mathrm{h})}{\sqrt{\frac{0.l°}{0.1°} \times \frac{0.l°}{0.1°} \times t}}. \quad (1)$$

The difference between the calculated and theoretical RMSE values given by (1) provides a measure of error autocorrelation.

3. Results

a. Assessment of IMERG products

In this section, a systematic assessment of the three IMERG V06 products is structured over the HuaiRB

TABLE 1. Statistical performance metrics used for the evaluation of satellite-based precipitation estimations. Notation: n represents the number of samples; $R_{s,i}$ and $R_{o,i}$ are IMERG and gauged observations, respectively; $\overline{R_o}$ and $\overline{R_s}$ are the mean values of the corresponding elements; H is the number of observed rainfall events that are correctly detected by IMERG; M is the observed rainfall events that are not detected by IMERG; and F is the number of rainfall events detected by IMERG but not occurring in the observation.

Statistic metrics	Unit	Equation	Perfect value
Correlation coefficient (CC)	—	$\mathrm{CC} = \dfrac{\sum_{i=1}^{n}(R_{o,i}-\overline{R_o})(R_{s,i}-\overline{R_s})}{\sqrt{\sum_{i=1}^{n}(R_{o,i}-\overline{R_o})^2}\sqrt{\sum_{i=1}^{n}(R_{s,i}-\overline{R_s})^2}}$	1
Mean absolute error (MAE)	mm	$\mathrm{MAE} = \dfrac{1}{n}\sum_{i=1}^{n}\lvert R_{s,i}-R_{o,i}\rvert$	0
Relative bias (Rbias)	%	$\mathrm{Rbias} = \dfrac{\sum_{i=1}^{n}(R_{s,i}-R_{o,i})}{\sum_{i=1}^{n}R_{o,i}} \times 100\%$	0
Root-mean-square error (RMSE)	mm	$\mathrm{RMSE} = \sqrt{\dfrac{1}{n}\sum_{i=1}^{n}(R_{s,i}-R_{o,i})^2}$	0
Probability of detection (POD)	—	$\mathrm{POD} = \dfrac{H}{H+M}$	1
False alarm ratio (FAR)	—	$\mathrm{FAR} = \dfrac{F}{H+F}$	0
Critical success index (CSI)	—	$\mathrm{CSI} = \dfrac{H}{H+M+F}$	1

against an (assumed) perfect benchmark (i.e., the CMPA) from June 2014 to May 2017. As discussed above, our assessment framework is based on measures of spatial pattern consistency, basic statistical performance, and the capacity to accurately identify rain occurrences.

Figure 2 shows the spatial distributions of daily average precipitation rates (mm day^{-1}) sampled from each of the three IMERG products and the CMPA benchmark. Note that daily average precipitation is examined here, instead of hourly, to reduce visual clutter and average across diurnal variability. Results from the reference CMPA (Fig. 2d) show a clearly decreasing geographic trend from south and southeast to north and northwest with the highest daily precipitation in the Dabie Mountains (>4 mm day^{-1}) in the southwest portion of the basin (see Fig. 1a) and the lowest daily precipitation in the northern plains area (approximately 1 mm day^{-1}). The three IMERG products all capture this spatial tendency well; however, visual differences are still apparent. For instance, although all three IMERG products tend to overestimate mean precipitation rates, IMERG-E and IMERG-L generally exhibit a more severe overestimation than IMERG-F, particularly over regions with high precipitation (i.e., the south and southeast portions of HuaiRB). In addition, there are some patches of positive bias in the IMERG-E and IMERG-L products within the Yishusi basin and along the northern part of the midstream basin. Both of these biases are effectively corrected in the post-real-time IMERG-F product.

Interestingly, precipitation rates in the IMERG-F product demonstrate a gradual spatial descent from south and southeast to north and northwest that is much spatially smoother than in the other two IMERG products. The most likely explanation for this is that the sparse gauge network (Fig. 1b) used to derive the GPCC monthly gauge analyses does not adequately sample real rainfall spatial variability—which subsequently smooths out observed spatial heterogeneity in the IMERG-F product during the calibration process.

To quantify spatial differences between the three IMERG products, long-term daily average precipitation values sampled from all grids within the corresponding subbasin are compared against the CMPA benchmark in Fig. 3. Relevant statistical metrics, including CC, Rbias, MAE, and RMSE, are listed for each comparison. It notable that all three IMERG products significantly overestimate precipitation rates in each of the four subbasins. Summarized in terms of Rbias, the three IMERG products present the best metrics over the downstream basin (Rbias values of 31.33%, 32.88%, and 21.05% for the three IMERG products, respectively) and the poorest metrics over the Yishusi basin (Rbias values of 44.97%, 45.22%, and 27.59% for the three IMERG products, respectively). Even so, according to Fig. 3, a nearly constant systematic bias across all rain rates is seen in the four subbasins. The highest CC metrics occur within the midstream basin area (CC > 0.90), indicating that IMERG products best reproduce the spatial

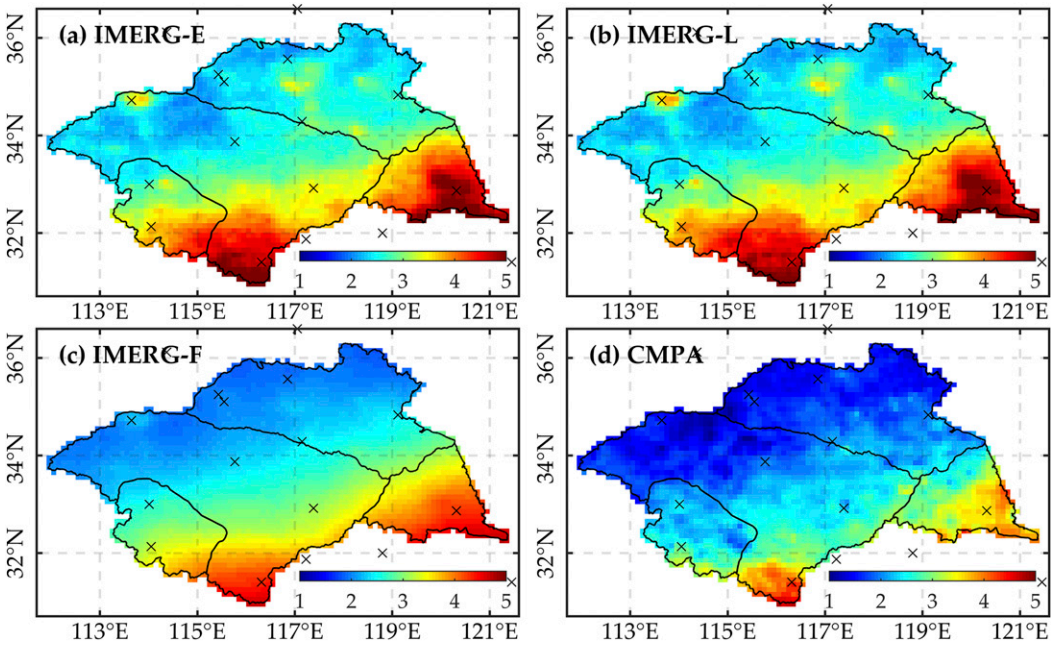

FIG. 2. The spatial distribution of daily average precipitation (mm day^{-1}) sampled from June 2014 to May 2017 at 0.1° × 0.1° resolution over Huaihe River basin for: (a) IMERG-E, (b) IMERG-L, (c) IMERG-F, and (d) the CMPA benchmark.

variations of CMPA there. In the upstream subbasin area, the consistency of the precipitation distribution between the IMERG products and the CMPA benchmark is the lowest—as evidenced by the relatively small CC values sampled there.

Among the three IMERG products, IMERG-F generally exhibits the best performance in all subbasins. As expected, it appears to benefit from GPCC monthly calibration and the use of additional passive microwave (PW) data relative to the near-real-time products. In terms of these products, IMERG-L has slightly better accuracy than IMERG-E in most subbasins. In summary, as evidenced by their high CC metrics (0.93, 0.92, and 0.96 for the three IMERG products over the entire HuaiRB, respectively), all three IMERG products capture the spatial variation of daily average precipitation relatively well.

To analyze spatial variations in their temporal accuracy, hourly temporal performance metrics for all IMERG products were sampled separately for all spatial grids within the HuaiRB. The spatial distributions of these temporal metrics are plotted in Figs. 4 and 5, respectively, with associated boxplots shown in Fig. 6. In terms of the CC metric, the post-real-time IMERG-F shares an almost identical spatial distribution with the near-real-time IMERG-E and IMERG-L products, although it has a slightly higher grid-averaged value (0.46 vs 0.45 and 0.45 for three IMERG products, respectively). Note that regions of low temporal CC (CC < 0.35) are mainly located in the northwestern, northeastern, and central parts of the HuaiRB and are spatially consistent with the demarcation line between semihumid and humid climates (see Fig. 1b). This suggests that these low CC metrics are mainly attributed to the highly variable weather in the transition area between climate zones, which complicates the remote retrieval of precipitation. Likewise, both MAE and RMSE display a spatial trend that mirrors that of daily average precipitation (Fig. 2)—indicating their close relationship with total precipitation accumulation. Given the uneven spatial distribution of precipitation over the HuaiRB, the meaningful intracomparison of the spatial distribution of MAE or RMSE results is difficult. However, an intercomparison between the three IMERG products indicates that the magnitude of error in IMERG-F is slightly lower than that of IMERG-E and IMERG-L (Figs. 6b,d).

The IMERG-F product also has relatively better Rbias results—with a spatial average of 27.14% in comparison to 41.04% and 41.34% for IMERG-E and IMERG-L, respectively. According to the spatial distributions of Rbias shown in Fig. 4, regions of high Rbias for IMERG-E and IMERG-L are mainly distributed in the northern parts of HuaiRB—where positive Rbias is generally higher than 45%. This Rbias trend is probably caused by the fact that a smaller precipitation value is

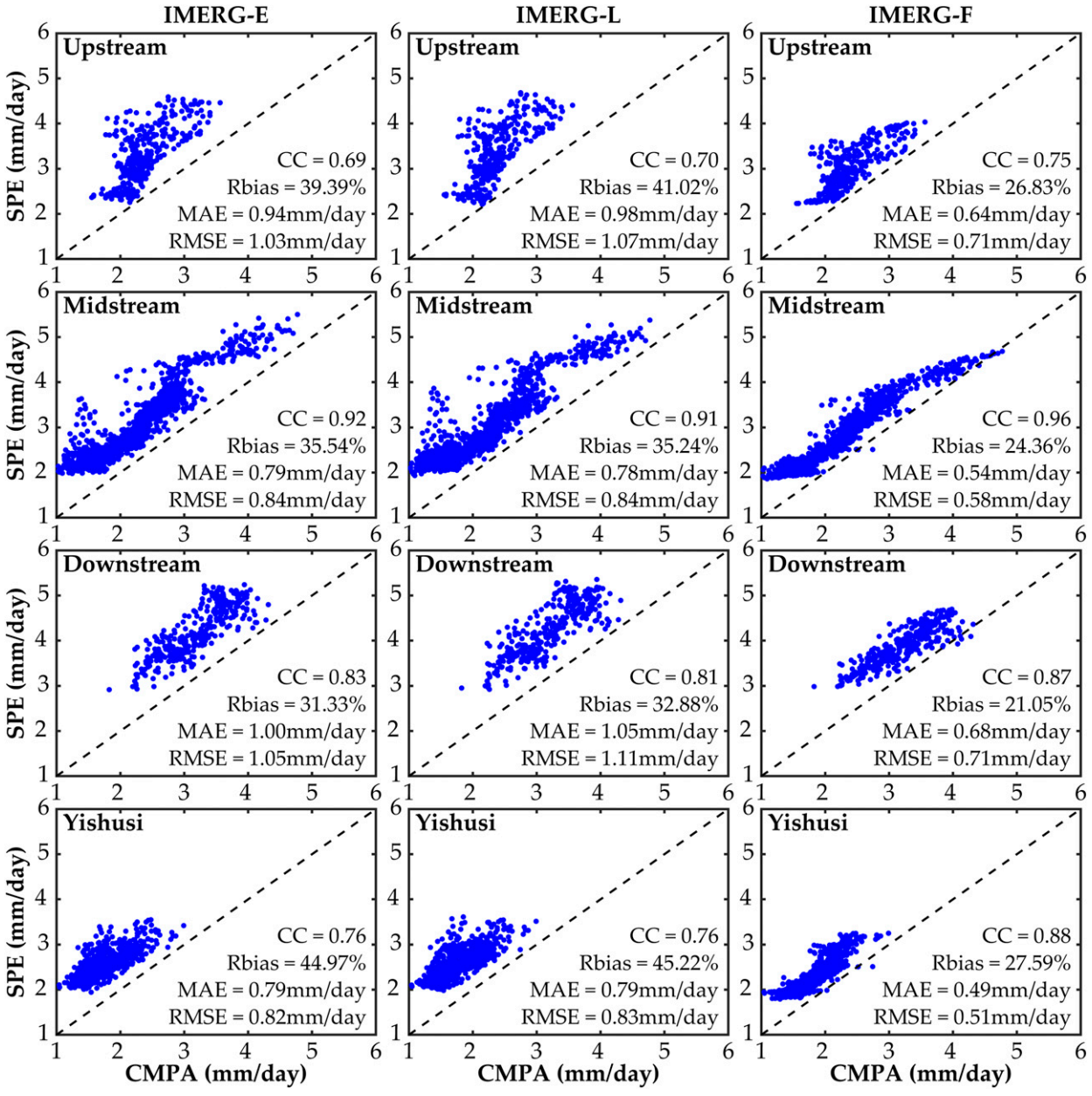

FIG. 3. Scatterplots of daily average precipitation (mm day^{-1}) for all three IMERG products against the CMPA over the four subbasins of the HuaiRB during the period June 2014 and May 2017. Each point represents a single spatial grid box.

used in the normalization over northern parts of HuaiRB. In contrast, although IMERG-F also overestimates the precipitation in most areas of HuaiRB, the severity of the overestimation is more modest than that seen in near-real-time IMERG products. In general, the three IMERG products display similar spatial patterns of CC, MAE, Rbias, and RMSE with relatively higher accuracy and lower bias in the southern HuaiRB and the opposite scenario in northern HuaiRB. On the whole, the near-real-time IMERG-E and IMERG-L products demonstrate significantly poorer performance than IMERG-F, and IMERG-L performs slightly better than IMERG-E.

For the POD, FAR, and CSI metrics shown in Figs. 6e–g, the hourly IMERG products demonstrate high POD metrics (grid-averaged values of 0.56, 0.58, and 0.59, respectively), indicating high accuracy in detecting actual precipitation events. At the same time, they also have high FAR metrics (grid-averaged values of 0.62, 0.60, and 0.59, respectively). Given that a relatively small threshold (0.1 mm h^{-1}) is employed to calculate

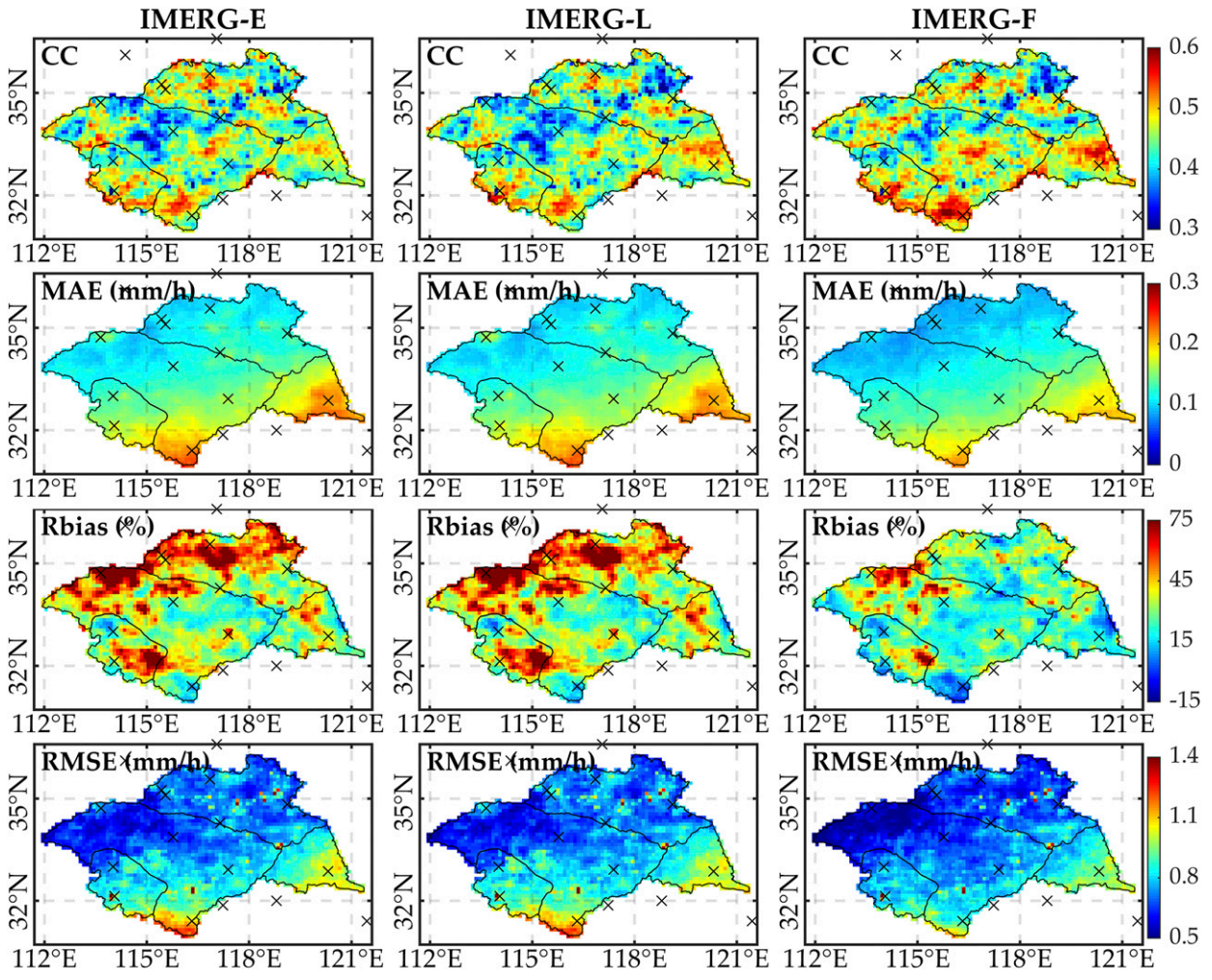

FIG. 4. The spatial distribution of basic statistical indices at an hourly time scale for the three IMERG products vs the CMPA.

the categorical indices, some low-rainfall ''drizzle'' events—detected by IMERG products but not recorded by AWSs—likely contribute to the high FAR metric values. As reported by Su et al. (2019a), the enhanced sensitivity of sensors used in producing the IMERG products and the increased sampling frequency of GPM mission effectively improve IMERG's ability in detect light rain but also increase the risk of false alarms. Given the high FAR values sampled in the HuaiRB, reducing the frequency of false precipitation alarms remains an important future task for IMERG product developers. It is also interesting to note that IMERG-L and IMERG-F products present higher POD metrics than IMERG-E but still have slightly lower FAR metrics. Since IMERG-E is based only on the forward propagation of the PW data (unlike both IMERG-L and IMERG-F), the use of the forward and backward propagation of the PW data appears to improve the precipitation detection capability of IMERG products.

In terms of spatial variations in the categorical indices show in Fig. 5, they also reveal the superior performance of IMERG products in the southern HuaiRB relative to the northern HuaiRB, demonstrating agreement with basic statistical indices shown earlier in Fig. 4. Besides, in the comparison to spatial distribution for FAR and Rbias, we can even infer that the severe positive Rbias in IMERG products is caused mainly by relatively high FAR metrics. Notably, the three IMERG products all show poor detection capacity over and near lakes in the HuaiRB (e.g., the Nansi Lakes and Hongze Lake) with low POD values and high FAR values. This may be an artifact of limitations in our ground-based CMPA reference—as there are fewer AWSs located in and near these lakes. Besides, Tian and Peters-Lidard (2007) confirmed that CMORPH products have much higher FAR over water bodies than over land pixels, whereas this issue is negligible for IMERG products due to their use of an unified and updated microwave algorithm

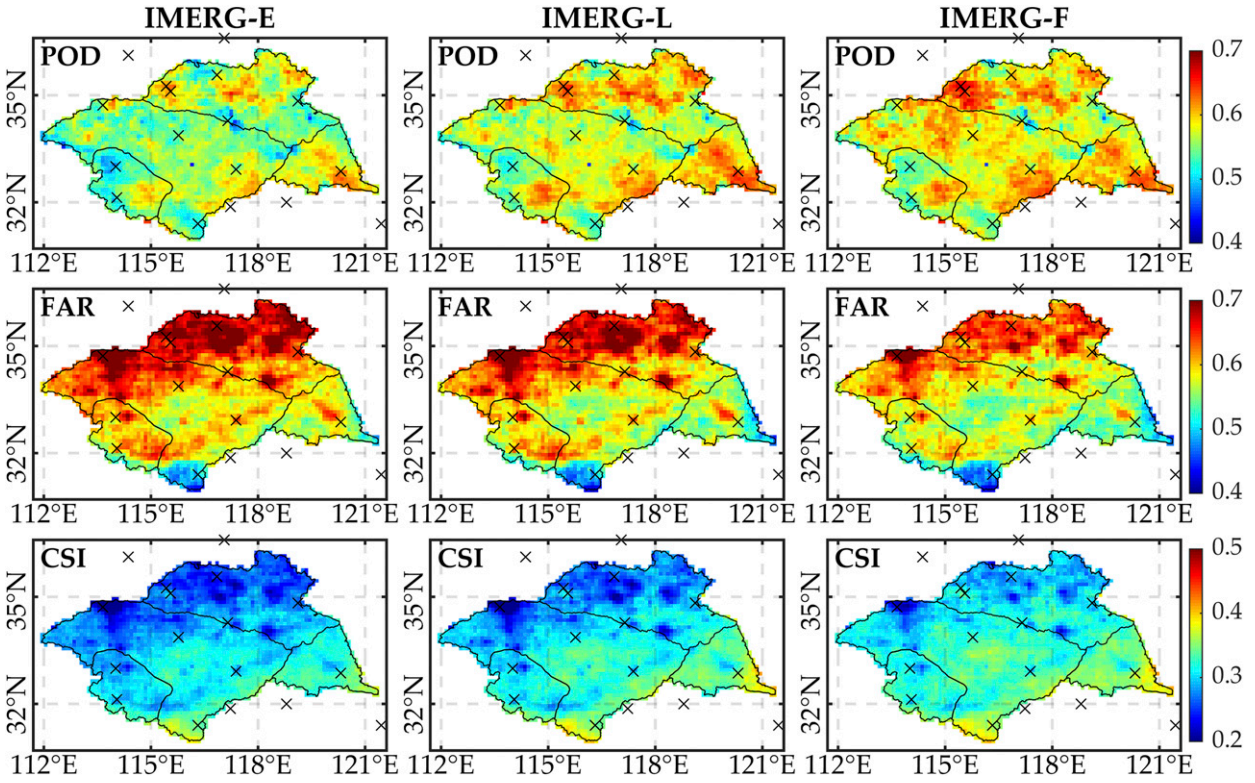

FIG. 5. As in Fig. 4, but for categorical error metrics.

(Tang et al. 2016a). Considering that CMORPH has been integrated into the CMPA, the high FAR of CMORPH potentially reduces the quality of CMPA over the lake areas and thus biases our assessment results.

b. *The effects of spatial and temporal resolution enlargement*

Generally, coarsening their spatial and temporal resolution will improve the reliability of SPEs (Tan et al. 2017). However, the exact variation in key performance metrics as a function of spatiotemporal resolution has not been well studied. Hence, taking the IMERG V06 products as an example, changes in IMERG performance metrics over a range of continuous spatiotemporal resolutions are examined in this section. Since the spatiotemporal averaging of precipitation does not impact Rbias (Tang et al. 2016b), it is not considered.

Based on the grid-averaged statistical metrics, Fig. 7 demonstrates how the performance of IMERG-F changes with the modification of spatiotemporal resolution. Similar analyses for IMERG-E and IMERG-L are provided in the supplemental material (Figs. S1 and S2 in the online supplemental material, respectively). For consistency, the units of MAE and RMSE at different spatiotemporal resolutions are all expressed as millimeters per hour. From the distributions of performance metrics shown in related figures, all three IMERG products share almost identical distributions of variation tendencies—with a diagonally zonal increase accuracy from the bottom left (i.e., fine spatiotemporal resolutions) to the top right (i.e., coarse spatiotemporal resolutions) of the figures. This means that, as expected, all IMERG products perform better at coarser spatiotemporal resolutions than at finer ones.

The accurate estimation of precipitation occurrence, amount, and distribution are three vital components of robust remote sensing precipitation retrievals. Degrading spatial resolution contributes little to the improvement of precipitation retrieval algorithm, but it reduces the sensitivity to the exact spatial location of precipitation and thus benefits statistical metrics. Analogously, the degradation of temporal resolution increases the ability of SPEs to identify rain occurrence by obscuring the exact timing of precipitation events. Hence, according to the POD, FAR, and CSI metrics, the detection capability of IMERG products improves rapidly with increased spatiotemporal resolution. However, as suggested by Tan et al. (2017), categorical rain detection results at different spatiotemporal scales are sensitive to the choice of a rain/no-rain threshold. Considering that a constant threshold of 0.1 mm h^{-1} is employed here, a scale-consistent set of thresholds should be examined in future research.

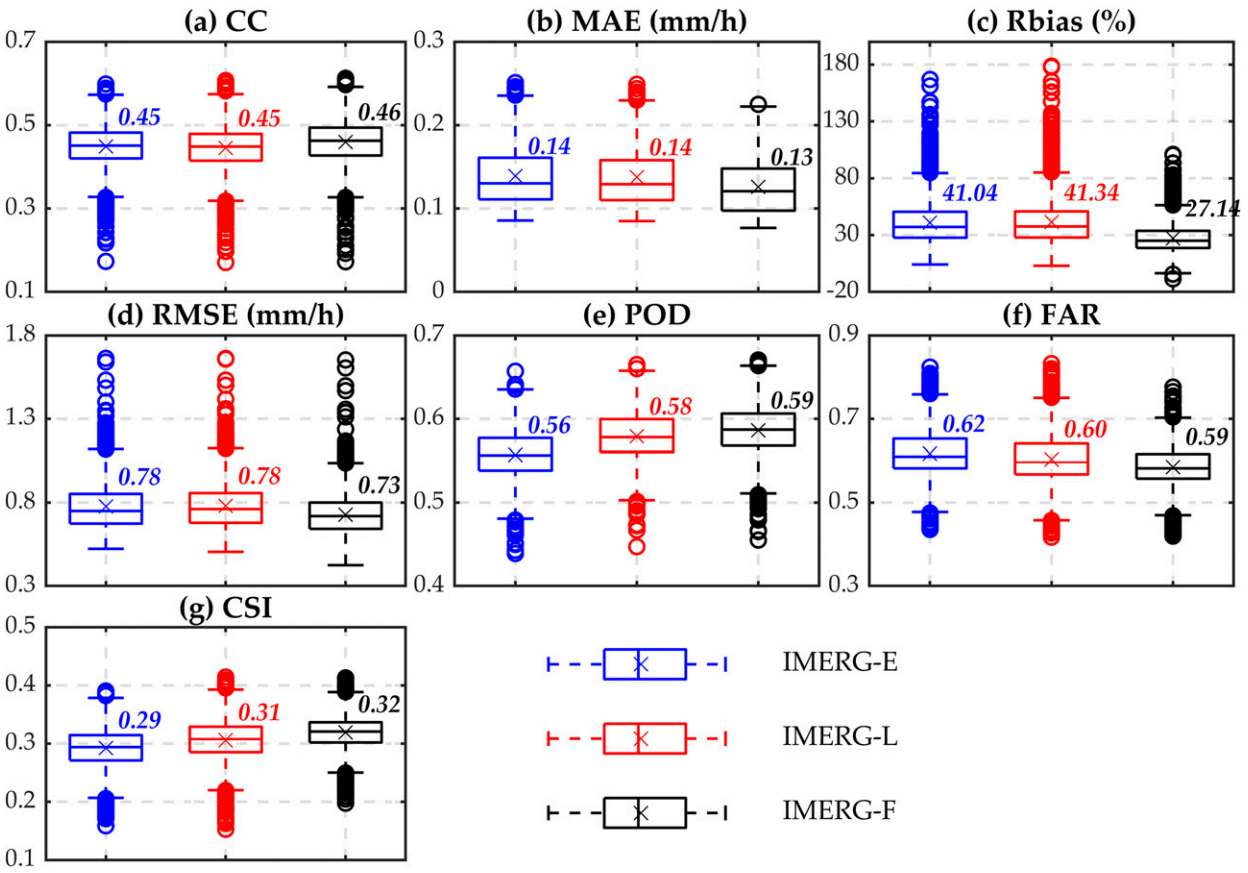

FIG. 6. The boxplot analysis of the statistical metrics over all grids within the HuaiRB. The five plotted horizontal lines (from bottom to top) in each box represent: minimum, 25th percentile, 50th percentile, 75th percentile, and maximum values, respectively. Plotted points mark the outliers.

CC metric results agree well with that of categorical statistical indices—indicating that the increased linear agreement of IMERG products with real precipitation can also be linked to the degradation of spatiotemporal resolution. Besides, given that gridded SPEs provide the mean precipitation for the corresponding grid cell, averaging over larger areas and longer periods is expected to filter random errors. Hence, as results show, the retrieved errors of IMERG products (in terms of MAE and RMSE) demonstrate a decreasing trend from fine resolutions to rough ones. However, if the errors are highly autocorrelated in space or time, temporal–spatial averaging will degrade spatiotemporal resolution without actually reducing the RMSE of IMERG products.

To quantify the effect of autocorrelated errors on the temporal aggregation of IMERG-F precipitation estimates, actual RMSE metrics and associated theoretical RMSE values are compared in Fig. 8. Note that plotted "theoretical" RMSE values are calculated based on an assumption of purely uncorrelated errors and (1). For convenience, the difference between these two set of RMSE values—providing a measure of error autocorrelation—is also plotted in Fig. 8. The curves of actual RMSE at different spatial resolutions all demonstrate a clear trend of decreasing RMSE values with coarsening temporal scales, although impacted by diminishing returns at coarser temporal resolutions. However, in comparison to the corresponding theoretical values, relatively less decay is observed. One plausible explanation is that while time-aggregation reduces error in IMERG-F, this reduction is limited by the presence of temporally autocorrelated errors. This explanation is supported by the RMSE differences, which remain roughly constant during temporal aggregation (see different subpanels in Fig. 8). Of course, given the nature of precipitation, the presence of continuous no-rainfall hours in the hourly precipitation time series also potentially weakens the positive effect of time aggregation and reduces the decline of actual RMSE. With regards to the degradation of spatial resolution, hourly RMSE metrics at coarse spatial resolutions are

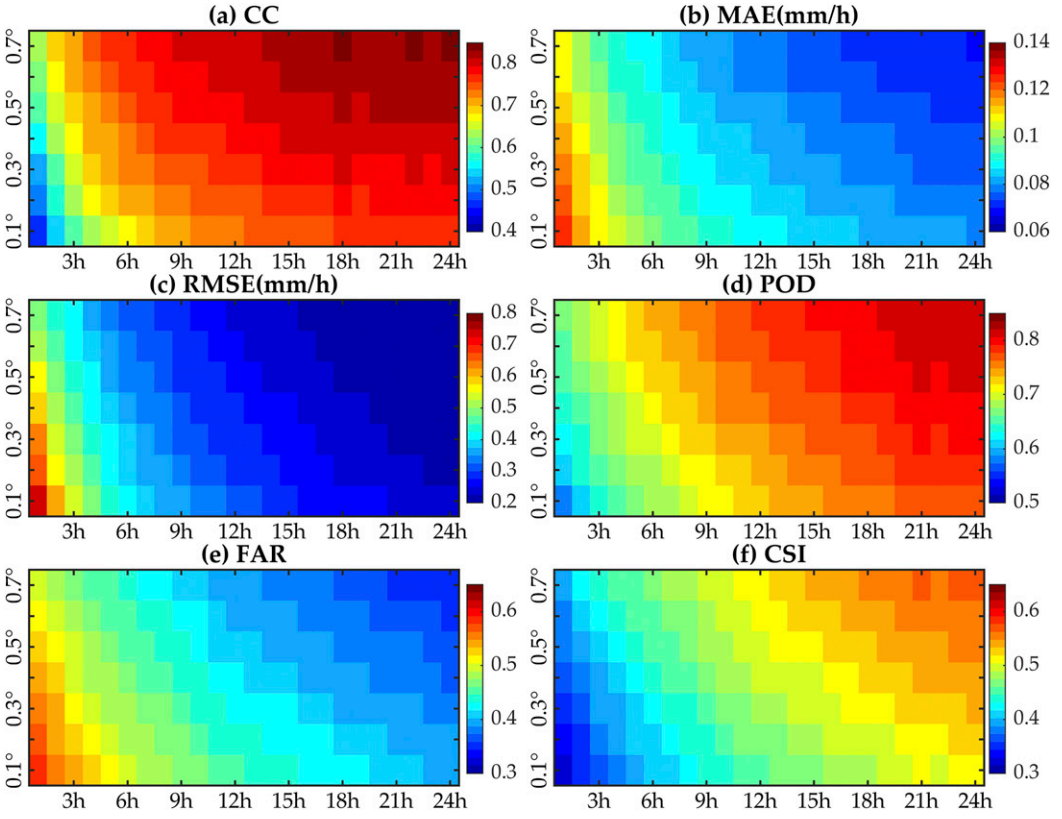

FIG. 7. The distribution of grid-averaged performance metrics for IMERG-F calculated at different spatial and temporal resolutions for the entire study period.

lower than that at fine spatial resolutions; however, the decay of RMSE during temporal aggregation is essentially the same at different spatial resolutions.

As for the effect of autocorrelated errors on spatial aggregation of IMERG-F, the comparison of actual RMSE metrics and associated theoretical values at multiple temporal resolutions (e.g., 1, 3, 6, 12, 18, and 24 h) are shown in Fig. 9. The theoretical RMSE values are also calculated for the case of purely uncorrelated errors. It can be observed that the actual RMSE curves decrease linearly with an increase of spatial resolution. This indicates that spatial averaging has a positive effect and reduces error in IMERG-F. Meanwhile, note that the reduction rate of actual RMSE at fine temporal resolutions, particularly at the hourly scale, is significantly faster than that at coarser scales. Given the relatively high time-dependence error at fine temporal resolutions, this is likely due to the presence of time-dependent error in IMERG-F that can be partly filtered by spatial aggregation (e.g., biases in adjacent grid cells may be different and can be averaged out through the degradation of spatial resolution). More importantly, as shown in Fig. 9, the theoretical RMSE curves decrease much faster than actual RMSE curves at all temporal resolutions, which suggests that spatial averaging can cancel out spatially uncorrelated error but fails to remove the spatially autocorrelated error. Since precipitation events generally cover a large area (i.e., multiple grid cells), IMERG products often contain sizeable spatial autocorrelation, which in turn, leads to strong spatial autocorrelation of errors. Therefore, increasing the spatial resolution of IMERG-F does not significantly reduce its errors. Interestingly, when the spatial resolution is finer than 0.5°, the difference between the two RMSE curves increases slightly with the coarsening of the spatial resolutions, indicating the increased suppression of RMSE reduction by spatial autocorrelated errors.

In summary, both temporal and spatial aggregation of IMERG products can improve their accuracy with regards to both precipitation amount and restoring precipitation events—particularly for precipitation estimates obtained at fine spatiotemporal resolutions. However, this improvement is somewhat compromised by the presence of autocorrelated error, and temporal aggregation is generally more efficient than spatial aggregation for filtering IMERG error.

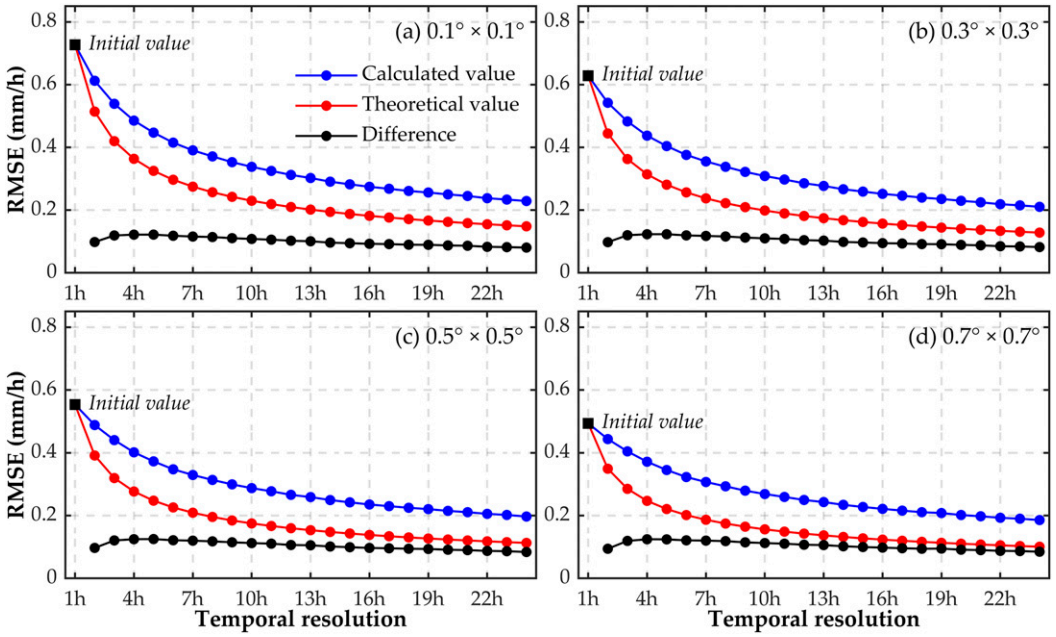

FIG. 8. Impact of temporal resolution on RMSE for four separate spatial resolutions.

4. Discussion

Due to the rapid development of satellite-based remote sensing and the establishment of Earth observation systems, the range of available remotely sensed geophysical datasets has increased rapidly. In particular, the development of SPEs had experienced three eras spanning across several decades, namely, 1) the Global Precipitation Climatology Project (GPCP) era from 1986 to 2000 (Huffman et al. 1997), 2) the TRMM era (Huffman et al. 2007) from 1998 to 2014, and 3) the GPM era since 2014 (Huffman et al. 2014). The application of these open-access SPEs in various fields (e.g., hydrology, water resources, weather, and climate) has yielded substantial scientific and societal benefits (W. Yang et al. 2016). However, as an indirect estimator of precipitation, the SPEs also suffer from relatively high uncertainties, such as inaccurate delineation rain area, imprecise precipitation occurrence time, and uncertainties in the relationship between cloud/precipitating particles and microwave radiances. The bias can be reduced in regions with gauge measurements via gauge calibration. However, sparse rain gauge networks have their own inherent error and uncertainties. Meanwhile, the misclassification of rain/no-rain pixels and imprecise precipitation event timing is hard to correct (Su et al. 2018; Xu et al. 2016). These biases and random errors directly affect the simulation performance of hydrometeorological models using SPE as input.

Currently, SPEs are available at relatively high temporal and spatial resolution (e.g., 0.1° × 0.1° and half hourly for IMERG products). However, such high spatiotemporal resolutions are associated with a higher risk of error. Therefore, it is potentially advantageous to resample SPEs in time or space prior to their application. Choosing a suitable spatiotemporal resolution for the specific application based on the accuracy of SPEs and/or various application considerations may, in certain circumstances, yield better simulations. Temporal and spatial averaging is a common resampling method to acquire suitable (coarser) resolutions. However, such aggregation conforms to theoretical expectations only for the unrealistic case of purely random and uncorrelated errors. Therefore, exploring the variation of basic statistical performance and detection capability with the spatiotemporal resolution, as well as the potential effect of error autocorrelation error, is a meaningful topic.

Here, using the latest IMERG V06 products as examples, we examine the impact of spatiotemporal degradation on the accuracy of IMERG products. As illustrated in Fig. 7 (Figs. S1 and S2), degrading the spatiotemporal resolution of IMERG products improves their key statistical metrics and detection capacity. However, when compared to the theoretical curves, less RMSE decay is seen for both temporal and spatial aggregation (see Figs. 8 and 9). The difference between these two RMSE curves (i.e., actual vs theoretical) reveals the presence of autocorrelated error and/or bias. Figure 8 verifies that temporal autocorrelated error

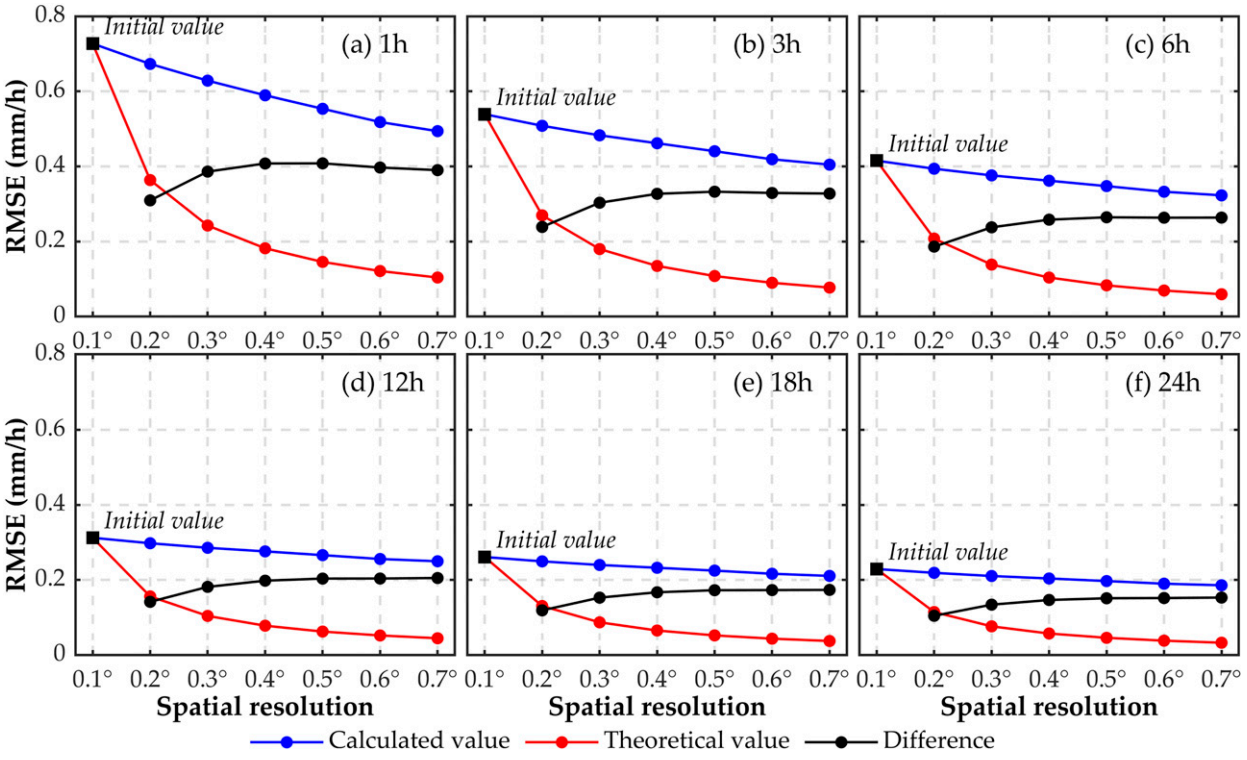

FIG. 9. The analyses of error variations (in terms of RMSE) under the degeneration of spatial resolution.

contained in hourly IMERG-F cannot be canceled out via temporal and spatial averaging. In contrast, Fig. 9 indicates that suppression of RMSE reduction by the spatially autocorrelated error in IMERG-F may slightly increase with the spatial aggregation when the spatial resolution is finer than 0.5°. It should be stressed that, some of the autocorrelated error responsible for inflating coarse-scale RMSE values may reasonably be defined as a stable bias. However, given significantly ambiguity in differentiating between highly autocorrelated errors and true bias, we will not attempt to distinguish between the two concepts here.

Generalizations about the autocorrelation structure of SPE's presented here are based on a single midlatitude basin (i.e., the HuaiRB). Since Wang et al. (2017) suggested that errors in SPEs are highly sensitive to latitude, climate, and elevation factors, the effect of degrading temporal and spatial resolution should be explored over a wider range of basins with different climates in different latitude zones.

Finally, although degrading spatiotemporal resolution improves the statistical accuracy of IMERG products, it also obscures the exact occurrence time and location of precipitation events. In other words, simple averaging of IMERG products is not able to enhance the actual event detection capability of the IMERG algorithm. Therefore, to promote the continued development of IMERG products, more effort should be made to develop efficient retrieval algorithms that provide more reliable precipitation estimations. Beyond that, for many disaster-related risk assessments such as flash flood predicting and landslide forecasting, more details about both when and where heavy precipitation occurs and the exact precipitation amount need to be provided to the associated application integration systems. Downscaling of SPEs is an effective way to provide these required precipitation details. However, due to the lack of an adequate benchmark datasets with spatial resolution is finer than 0.1°, such are approaches are difficult to validate. Even so, future research focusing on downscaling SPEs is strongly recommended to add value to IMERG products for disaster and natural hazard applications.

5. Conclusions

As TRMM's successor, the GPM mission has been operating for five years and continues to produce a vast quantity of quasi-global precipitation data at fine space–time resolutions (0.5 h and 0.1°). As one of the most important SPEs in the GPM era, the latest IMERG V06 products, including the near-real-time IMERG-E and IMERG-L products and the post-real-time IMERG-F, are here assessed against a ground-based CMPA benchmark

from June 2014 to May 2017 over the HuaiRB. Subsequently, the effects of spatial and temporal resolution degradation on performance metrics are examined. Our main conclusions are summarized as follows:

1) Compared to the ground-based CMPA benchmark, all three IMERG products accurately capture the spatial variation of mean precipitation despite the presence of a positive bias. The post-real-time IMERG-F demonstrates the best performance at both the subbasin-scale and grid-scale comparisons, followed by the near-real-time IMERG-L and IMERG-E products, respectively. IMERG products are more reliable in the southern HuaiRB than in the northern HuaiRB.
2) Degrading the spatiotemporal resolution improves the accuracy of IMERG products and improve a wide variety of SPE performance metrics (including detection-based metrics). However, due to the presence of autocorrelated error and/or bias, improvements associated with upscale averaging are significantly lower than theoretical expectation assuming purely uncorrelated error.
3) Component analysis indicates that the temporally autocorrelated error cannot be averaged out by both temporal and spatial aggregation. Hence, at multiple spatial resolutions, the RMSE differences share almost identical curves with respect to temporal aggregation. In terms of spatial autocorrelated error, it increases slightly in spatial averaging and reaches the asymptotic value at the spatial resolution 0.5. Results also indicate that temporal averaging can cancel out some spatially autocorrelated errors. Hence, temporal aggregation is somewhat more effective than spatial aggregation for reducing error in IMERG products.

The objective of this study is to examine the performance variation of SPEs as a function of spatiotemporal resolution and guide IMERG users on the performance specific to their purpose. Hence, the performance improvement of IMERG products accompanying upscale averaging is examined over the HuaiRB. However, spatiotemporal averaging is not suitable for disaster-related risk assessment, which requires detailed precipitation information as the forcing of the hydrological model. Therefore, future research focused on the downscaling of SPEs is recommended.

Acknowledgments. The authors declare no conflicts of interest. This research is supported by the National Key Research and Development Program (2018YFA0605400), NNSF (41830752 and 41971042), and the Open Funding of the Laboratory (2018B44114).

REFERENCES

Asong, Z. E., S. Razavi, H. S. Wheater, and J. S. Wong, 2017: Evaluation of Integrated Multisatellite Retrievals for GPM (IMERG) over southern Canada against ground precipitation observations: A preliminary assessment. *J. Hydrometeor.*, **18**, 1033–1050, https://doi.org/10.1175/JHM-D-16-0187.1.

Dezfuli, A. K., C. M. Ichoku, K. I. Mohr, and G. J. Huffman, 2017: Precipitation characteristics in West and East Africa from satellite and in situ observations. *J. Hydrometeor.*, **18**, 1799–1805, https://doi.org/10.1175/JHM-D-17-0068.1.

Dinku, T., E. N. Anagnostou, and M. Borga, 2002: Improving radar-based estimation of rainfall over complex terrain. *J. Appl. Meteor.*, **41**, 1163–1178, https://doi.org/10.1175/1520-0450(2002)041<1163:IRBEOR>2.0.CO;2.

Guo, H., S. Chen, A. Bao, A. Behrangi, Y. Hong, F. Ndayisaba, J. Hu, and P. M. Stepanian, 2016: Early assessment of Integrated Multi-Satellite Retrievals for Global Precipitation Measurement over China. *Atmos. Res.*, **176–177**, 121–133, https://doi.org/10.1016/j.atmosres.2016.02.020.

He, Z., L. Yang, F. Tian, G. Ni, A. Hou, and H. Lu, 2017: Intercomparisons of rainfall estimates from TRMM and GPM multisatellite products over the upper Mekong River basin. *J. Hydrometeor.*, **18**, 413–430, https://doi.org/10.1175/JHM-D-16-0198.1.

Huang, C., and Coauthors, 2019: How well can IMERG products capture typhoon extreme precipitation events over southern China? *Remote Sens.*, **11**, 70, https://doi.org/10.3390/rs11010070.

Huffman, G. J., and Coauthors, 1997: The Global Precipitation Climatology Project (GPCP) Combined Precipitation Dataset. *Bull. Amer. Meteor. Soc.*, **78**, 5–20, https://doi.org/10.1175/1520-0477(1997)078<0005:TGPCPG>2.0.CO;2.

——, and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.*, **8**, 38–55, https://doi.org/10.1175/JHM560.1.

——, D. T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, and P. Xie, 2014: NASA Global Precipitation Measurement (GPM) Integrated Multi-Satellite Retrievals for GPM (IMERG). Algorithm Theoretical Basis Doc., version 4.4, 30 pp., https://pps.gsfc.nasa.gov/Documents/IMERG_ATBD_V4.pdf.

——, and Coauthors, 2019: NASA Global Precipitation Measurement (GPM) Integrated Multi-Satellite Retrievals for GPM (IMERG). Algorithm Theoretical Basis Doc., version 06, 32 pp., https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V06.pdf.

Kidd, C., and G. Huffman, 2011: Global precipitation measurement. *Meteor. Appl.*, **18**, 334–353, https://doi.org/10.1002/met.284.

——, V. Levizzani, J. Turk, and R. Ferraro, 2009: Satellite precipitation measurements for water resource monitoring. *J. Amer. Water Resour. Assoc.*, **45**, 567–579, https://doi.org/10.1111/j.1752-1688.2009.00326.x.

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

Omranian, E., and H. O. Sharif, 2018: Evaluation of the Global Precipitation Measurement (GPM) satellite rainfall products over the lower Colorado River basin, Texas. *J. Amer. Water Resour. Assoc.*, **54**, 882–898, https://doi.org/10.1111/1752-1688.12610.

——, ——, and A. A. Tavakoly, 2018: How well can Global Precipitation Measurement (GPM) capture hurricanes? Case study: Hurricane Harvey. *Remote Sens.*, **10**, 1150, https://doi.org/10.3390/rs10071150.

Pan, Z., X. Ruan, M. Qian, J. Hua, N. Shan, and J. Xu, 2017: Spatiotemporal variability of streamflow in the Huaihe River basin, China: Climate variability or human activities? *Hydrol. Res.*, **49**, 177–193, https://doi.org/10.2166/nh.2017.155.

Prakash, S., A. K. Mitra, D. S. Pai, and A. AghaKouchak, 2016: From TRMM to GPM: How well can heavy rainfall be detected from space? *Adv. Water Resour.*, **88**, 1–7, https://doi.org/10.1016/j.advwatres.2015.11.008.

Sharifi, E., R. Steinacker, and B. Saghafian, 2016: Assessment of GPM-IMERG and other precipitation products against gauge data under different topographic and climatic conditions in Iran: Preliminary results. *Remote Sens.*, **8**, 135, https://doi.org/10.3390/rs8020135.

——, ——, and ——, 2018: Multi time-scale evaluation of high-resolution satellite-based precipitation products over northeast of Austria. *Atmos. Res.*, **206**, 46–63, https://doi.org/10.1016/j.atmosres.2018.02.020.

Shen, Y., and A. Xiong, 2016: Validation and comparison of a new gauge-based precipitation analysis over mainland China. *Int. J. Climatol.*, **36**, 252–265, https://doi.org/10.1002/joc.4341.

——, P. Zhao, Y. Pan, and J. Yu, 2014: A high spatiotemporal gauge-satellite merged precipitation analysis over China. *J. Geophys. Res. Atmos.*, **119**, 3063–3075, https://doi.org/10.1002/2013JD020686.

Su, J., H. Lü, J. Wang, A. M. Sadeghi, and Y. Zhu, 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 Sens.*, **9**, 1176, https://doi.org/10.3390/rs9111176.

——, ——, Y. Zhu, X. Wang, and G. Wei, 2018: Component analysis of errors in four GPM-based precipitation estimations over mainland China. *Remote Sens.*, **10**, 1420, https://doi.org/10.3390/rs10091420.

——, ——, D. Ryu, and Y. Zhu, 2019a: The assessment and comparison of TMPA and IMERG products over the major basins of Mainland China. *Earth Space Sci.*, **6**, 2461–2479, https://doi.org/10.1029/2019EA000977.

——, ——, Y. Zhu, Y. Cui, and X. Wang, 2019b: Evaluating the hydrological utility of latest IMERG products over the upper Huaihe River basin, China. *Atmos. Res.*, **225**, 17–29, https://doi.org/10.1016/j.atmosres.2019.03.025.

Tan, J., W. A. Petersen, P.-E. Kirstetter, and Y. Tian, 2017: Performance of IMERG as a function of spatiotemporal scale. *J. Hydrometeor.*, **18**, 307–319, https://doi.org/10.1175/JHM-D-16-0174.1.

Tan, M. L., and Z. Duan, 2017: Assessment of GPM and TRMM precipitation products over Singapore. *Remote Sens.*, **9**, 720, https://doi.org/10.3390/rs9070720.

Tang, G., D. Long, and Y. Hong, 2016a: Systematic anomalies over inland water bodies of high mountain asia in TRMM precipitation estimates: No longer a problem for the GPM era? *IEEE Geosci. Remote Sens. Lett.*, **13**, 1762–1766, https://doi.org/10.1109/LGRS.2016.2606769.

——, Y. Ma, D. Long, L. Zhong, and Y. Hong, 2016b: Evaluation of GPM Day-1 IMERG and TMPA version-7 legacy products over Mainland China at multiple spatiotemporal scales. *J. Hydrol.*, **533**, 152–167, https://doi.org/10.1016/j.jhydrol.2015.12.008.

——, Z. Zeng, M. Ma, R. Liu, Y. Wen, and Y. Hong, 2017: Can near-real-time satellite precipitation products capture rainstorms and guide flood warning for the 2016 summer in South China? *IEEE Geosci. Remote Sens. Lett.*, **14**, 1208–1212, https://doi.org/10.1109/LGRS.2017.2702137.

Tian, Y., and C. D. Peters-Lidard, 2007: Systematic anomalies over inland water bodies in satellite-based precipitation estimates. *Geophys. Res. Lett.*, **34**, L14403, https://doi.org/10.1029/2007GL030787.

Wang, X., H. Xie, H. Sharif, and J. Zeitler, 2008: Validating NEXRAD MPE and stage III precipitation products for uniform rainfall on the upper Guadalupe River basin of the Texas hill country. *J. Hydrol.*, **348**, 73–86, https://doi.org/10.1016/j.jhydrol.2007.09.057.

Wang, Z., R. Zhong, C. Lai, and J. Chen, 2017: Evaluation of the GPM IMERG satellite-based precipitation products and the hydrological utility. *Atmos. Res.*, **196**, 151–163, https://doi.org/10.1016/j.atmosres.2017.06.020.

Wei, G., H. Lü, W. T. Crow, Y. Zhu, J. Wang, and J. Su, 2018: Evaluation of satellite-based precipitation products from IMERG V04A and V03D, CMORPH and TMPA with gauged rainfall in three climatologic zones in China. *Remote Sens.*, **10**, 30, https://doi.org/10.3390/rs10010030.

Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. International Geophysics Series, Vol. 100, Academic Press, 704 pp.

Wu, Y., Z. Zhang, Y. Huang, Q. Jin, X. Chen, and J. Chang, 2019: Evaluation of the GPM IMERG v5 and TRMM 3B42 v7 precipitation products in the Yangtze River basin, China. *Water*, **11**, 1459, https://doi.org/10.3390/w11071459.

Xie, H., X. Zhang, B. Yu, and H. Sharif, 2011: Performance evaluation of interpolation methods for incorporating rain gauge measurements into NEXRAD precipitation data: A case study in the upper Guadalupe River basin. *Hydrol. Processes*, **25**, 3711–3720, https://doi.org/10.1002/hyp.8096.

Xu, S., Y. Shen, and Z. Du, 2016: Tracing the source of the errors in hourly IMERG using a decomposition evaluation scheme. *Atmosphere*, **7**, 161, https://doi.org/10.3390/atmos7120161.

Yang, M., X. Chen, and C. S. Cheng, 2016: Hydrological impacts of precipitation extremes in the Huaihe River basin, China. *SpringerPlus*, **5**, 1731, https://doi.org/10.1186/s40064-016-3429-1.

Yang, W., V. O. John, X. Zhao, L. Hui, and K. R. Knapp, 2016: Satellite climate data records: Development, applications, and societal benefits. *Remote Sens.*, **8**, 331, https://doi.org/10.3390/rs8040331.

Yong, B., L.-L. Ren, Y. Hong, J.-H. Wang, J. J. Gourley, S.-H. Jiang, X. Chen, and W. Wang, 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 Resour. Res.*, **46**, W07542, https://doi.org/10.1029/2009WR008965.

——, B. Chen, Y. Tian, Z. Yu, and Y. Hong, 2016: Error-component analysis of TRMM-based multi-satellite precipitation estimates over Mainland China. *Remote Sens.*, **8**, 440, https://doi.org/10.3390/rs8050440.

Yuan, F., and Coauthors, 2018: Evaluation of hydrological utility of IMERG final run V05 and TMPA 3B42V7 satellite precipitation products in the Yellow River source region, China. *J. Hydrol.*, **567**, 696–711, https://doi.org/10.1016/j.jhydrol.2018.06.045.

Zhao, H., S. Yang, S. You, Y. Huang, Q. Wang, and Q. Zhou, 2018: Comprehensive evaluation of two successive V3 and V4 IMERG final run precipitation products over Mainland China. *Remote Sens.*, **10**, 34, https://doi.org/10.3390/rs10010034.