

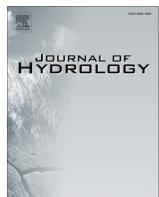


ELSEVIER

Contents lists available at ScienceDirect

# Journal of Hydrology

journal homepage: [www.elsevier.com/locate/jhydrol](http://www.elsevier.com/locate/jhydrol)



## Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over Mainland China at multiple spatiotemporal scales



Guoqiang Tang<sup>a</sup>, Yingzhao Ma<sup>a</sup>, Di Long<sup>a,\*</sup>, Lingzhi Zhong<sup>b</sup>, Yang Hong<sup>a,c,\*</sup>

<sup>a</sup>State Key Laboratory of Hydroscience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing, China

<sup>b</sup>Chinese Academy of Meteorological Sciences, Beijing, China

<sup>c</sup>Department of Civil Engineering and Environmental Science, University of Oklahoma, Norman, OK, United States

### ARTICLE INFO

#### Article history:

Received 7 September 2015

Received in revised form 26 November 2015

Accepted 5 December 2015

Available online 12 December 2015

This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Matthew Rodell, Associate Editor

#### Keywords:

GPM

IMERG

TRMM

Precipitation

### SUMMARY

The post-real time product of Day-1 Integrated Multi-satellite Retrievals for Global Precipitation Measurement (IMERG) is evaluated over Mainland China from April to December 2014 at the hourly timescale, against data from hourly ground-based observations. In addition, the IMERG product is compared with its predecessor—the Version-7 post-real-time 3B42 (3B42V7) product of Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA) at its original 3-hourly and then daily timescales for the same period. All the products are cross-evaluated at gridded, regional, and national scales. Results show that: (1) the Day-1 IMERG shows appreciably better performance than 3B42V7 at both sub-daily and daily timescales, and all the three spatial scales. The gap between the two products is more significant at the sub-daily resolution; (2) Out of the six sub-regions of China, IMERG especially performs better than 3B42V7 at the mid- and high-latitudes, as well as relatively dry climate regions; (3) IMERG can better reproduce the probability density function (PDF) in terms of precipitation intensity, particularly in the low ranges; and (4) although IMERG better captures the precipitation diurnal variability, both products have room to further improve their capability, particularly in the dry climate and high-latitude regions. This study is among the earliest evaluation and comparison of IMERG and 3B42V7 products, which could be valuable in providing reference for the development of IMERG algorithms, associated global products, and various applications as well.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

Precipitation is one of the most important components of global water and energy cycles, playing an important role in the interactions between hydrosphere, atmosphere, and biosphere (Kidd and Huffman, 2011). Accurate precipitation measurement or estimation is vital to water resource management, weather prediction, disaster monitoring, controlling, and so on (Hou et al., 2014). In addition, precipitation input has great impacts on the performance of a range of hydrological, climatic, and atmospheric models (Shen and Xiong, 2015). However, obtaining accurate precipitation has always been challenging for scientists. Currently, there are three mainstream methods to measure precipitation, i.e., gauge, weather

radar, and satellite-based sensors (Li et al., 2013). Gauges provide the most straightforward and accurate precipitation observations so far (Ma et al., 2015). However, gauge data are provided at specific sites, and various interpolation methods would result in potential errors when obtaining continuous spatial precipitation estimates. In addition, networks of rain gauges are sparse over most of continents, and few gauges are located over the ocean (Kidd and Huffman, 2011). In regard to the weather radar, it can provide the internal structure of storms, and real-time high-resolution monitoring over large areas (Doviak and Zrnic, 2006; Germann et al., 2006). However, the radar also suffers from various sources of errors, including mean-field systematic error, range-dependent systematic error, and random error (Dinku et al., 2002). The radar network is often not dense enough over most parts of the world. The only practical way to achieve comprehensive estimation of precipitation on a global basis relies on earth observation satellites (Hong et al., 2012; Hou et al., 2014; Villarini and Krajewski, 2008).

In recent years, a large number of quasi-global satellite precipitation products with various temporal and spatial resolutions

\* Corresponding authors at: Room 224b, New Hydraulic Engineering Building, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China (D. Long). Room A207, State Key Laboratory of Hydroscience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China (Y. Hong). Tel.: +86 010 62787394.

E-mail addresses: [dlong@tsinghua.edu.cn](mailto:dlong@tsinghua.edu.cn) (D. Long), [hongyang@tsinghua.edu.cn](mailto:hongyang@tsinghua.edu.cn) (Y. Hong).

have been developed and released to the public, such as TMPA (Huffman et al., 2007), Climate Prediction Center (CPC) MORPHing technique (CMORPH) (Joyce et al., 2004), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) (Sorooshian et al., 2000), and Global Satellite Mapping of Precipitation (GSMap) (Kubota et al., 2007). Those free and open access products have been widely studied and applied globally (Hong et al., 2007a, 2007b, 2007c; Long et al., 2015a; Yong et al., 2015), and regionally (Bitew et al., 2012; Kirstetter et al., 2013; Long et al., 2014; Long et al., 2015b; Xue et al., 2013), and could potentially bring substantial scientific and societal benefits (e.g., disaster forecast and monitoring, and water resource management).

The Global Precipitation Measurement (GPM) mission is an international constellation of satellites, including one Core Observatory satellite and approximately ten partner satellites. As the successor to the TRMM satellite, which was launched by the National Aeronautics and Space Administration (NASA) and National Space Development Agency (NASDA) on November 27, 1997, the GPM Core Observatory was deployed on February 28, 2014 by a joint effort of NASA and the Japan Aerospace Exploration Agency (JAXA), marking a transition from the TRMM era to the GPM era. The GPM Core Observatory carries a dual-frequency precipitation radar (DPR; the Ku-band at 13.6 GHz and Ka-band at 35.5 GHz) and a conical-scanning multichannel GPM Microwave Imager (GMI; frequencies range between 10 and 183 GHz). GPM extends the sensor package compared to TRMM instruments, which had a single-frequency precipitation radar (PR; the Ku-band at 13.8 GHz) and a multichannel TRMM Microwave Imager (TMI; frequencies range between 10 and 85.5 GHz). Therefore, the GPM sensors can detect light and solid precipitation more accurately than TRMM sensors (Hou et al., 2014). This study will focus on the Level-3 product, provided by the Day-1 Integrated Multi-satellE Retrievals for GPM (IMERG) algorithm, which is intended to intercalibrate, merge, and interpolate all microwave (MW) estimates of the GPM constellation, infrared (IR) estimates, gauge observations, and other data from potential sensors at  $0.1^\circ \times 0.1^\circ$  and half-hour temporal resolutions (Huffman et al., 2014). IMERG provides three kinds of products, including the near real time "Early" and "Late" run products, and the post real time "Final" run product. The "Final" run is the research level product used in this study.

As GPM's precursor, the TRMM satellite re-entered the Earth's atmosphere on June 15, 2015 after over 17 years of productive data gathering (<http://pmm.nasa.gov/gpm-news/trmm-spacecraft-re-enters-over-tropics>). The TMPA is intended to provide the "best" satellite precipitation estimate (Huffman et al., 2007), which has been studied and applied widely over the past years. Once the TRMM satellite has decommissioned, the TMPA products are intended to be generated using other calibrators instead of TRMM satellite instruments, until the GPM products can totally substitute TMPA products (Huffman et al., 2015). It is necessary and meaningful to evaluate the Day-1 IMERG product and compare it with TMPA products, which will shed light on research and application during the transition from TMPA to IMERG, and future improvement and retrospective construction of IMERG. There have been a multitude of statistical and hydrological studies comparing and evaluating various satellite precipitation products (Bitew et al., 2012; Li et al., 2015; Scheel et al., 2011; Su et al., 2008). However, the majority of these studies were conducted at relatively coarse temporal scales (e.g., daily or monthly scales), which may be far less than satisfactory from revealing the characteristics of those products comprehensively at 3 hourly or even finer temporal resolutions. The IMERG products are characterized by high temporal and spatial resolutions (half-hour and  $0.1^\circ \times 0.1^\circ$ ), and its capability to detect light and solid precipitation (Hou et al., 2014;

Huffman et al., 2015). Due to the lack of ground snowfall observations, this study only focuses on the evaluation and comparison of total precipitation (including liquid, solid, and mixed precipitation). Further studies are needed to evaluate the quality of IMERG solid precipitation products.

Therefore, the objectives of this study are twofold: (1) evaluating the quality of Day-1 IMERG "Final" run precipitation products over Mainland China at hourly,  $0.1^\circ \times 0.1^\circ$  resolutions against hourly ground-based observations from the China Meteorological Administration (CMA) over Mainland China, and (2) comparing IMERG and 3B42V7 products synchronously at 3-hourly and daily timescales, to explore the continuity and differences between the two products in both GPM and TRMM eras. This study will reveal characteristics of hourly, 3-hourly and daily errors of IMERG and TMPA products at gridded, regional, and national scales, and provide insight into subsequent studies and applications of IMERG precipitation products at high temporal and spatial resolutions. The remaining parts of the paper are organized as follows: Section 2 introduces the study area, precipitation datasets, and metrics. Section 3 evaluates the quality of IMERG products at  $0.1^\circ \times 0.1^\circ$  and hourly resolutions. Section 4 compares IMERG and TMPA 3B42 products at multiple spatiotemporal scales. Finally, discussion and summary are involved in Section 5.

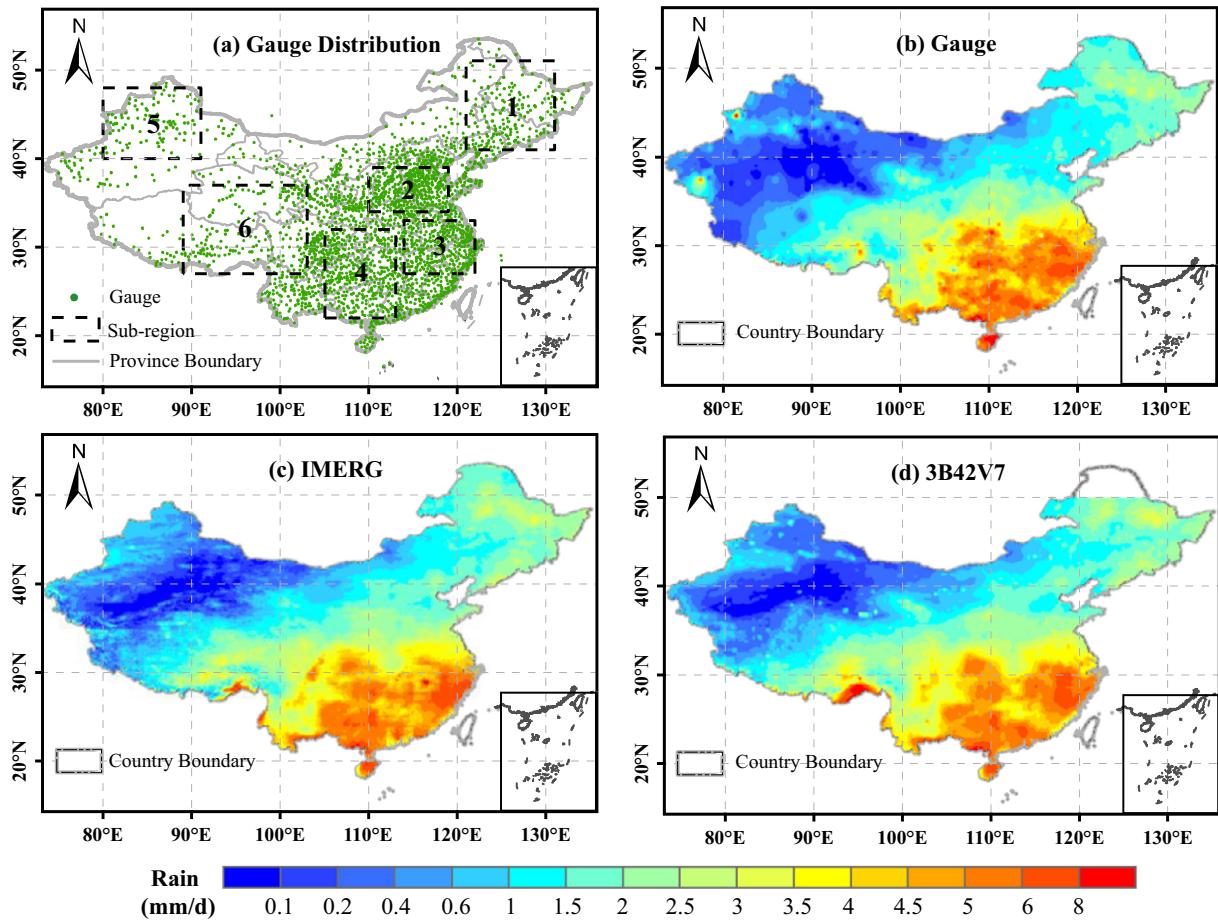
## 2. Study area, datasets and metrics

### 2.1. Study area

The study area is the whole Mainland China, located between  $73^\circ\text{--}135^\circ\text{E}$  and  $18^\circ\text{--}53^\circ\text{N}$  (Fig. 1(a)). Six sub-regions of Mainland China with relatively high density of gauges were chosen to conduct regional evaluation, i.e., Region 1 (Reg1) covers part of northeast Mainland China ( $121^\circ\text{--}131^\circ\text{E}$ ,  $41^\circ\text{--}50^\circ\text{N}$ ), Region 2 (Reg2) involves the lower reach of the Yellow River, part of the North China Plain ( $110^\circ\text{--}119^\circ\text{E}$ ,  $34^\circ\text{--}39^\circ\text{N}$ ), Region 3 (Reg3) encompasses the lower reach of the Yangtze River and Huaihe River ( $114^\circ\text{--}122^\circ\text{E}$ ,  $27^\circ\text{--}33^\circ\text{N}$ ), Region 4 (Reg4) includes the middle reach of the Yangtze River and most part of the Pearl River ( $105^\circ\text{--}113^\circ\text{E}$ ,  $22^\circ\text{--}32^\circ\text{N}$ ), Region 5 (Reg5) includes the northern part of Xinjiang Uygur Autonomous Region ( $80^\circ\text{--}91^\circ\text{E}$ ,  $40^\circ\text{--}48^\circ\text{N}$ ), and Region 6 (Reg6) covers the eastern part of the Tibetan Plateau ( $89^\circ\text{--}103^\circ\text{E}$ ,  $27^\circ\text{--}37^\circ\text{N}$ ) (Fig. 1(a)). Reg1 and Reg2 are primarily controlled by the monsoon climate of medium latitudes, characterized by hot and wet summer, and cold and dry winter (Meng et al., 2014). Reg3 and Reg4 are dominated by the subtropical monsoon climate, with a hot and wet summer, and mild and dry intensity in winter. Reg5 is dictated by the temperate continental climate with little precipitation over the whole year. Reg6 is dominated by the Plateau mountain climate, characterized by complex topography and high elevation, and showing great precipitation variability (Guo et al., 2007).

### 2.2. Gauge precipitation observation

Hourly gridded precipitation data sets from April through December 2014 with  $0.1^\circ \times 0.1^\circ$  resolution were provided by the CMA, which were produced using inverse distance weighting (IDW) interpolation method (Ahrens, 2006). The interpolated value is estimated by a weighted mean of the observations and the weights are proportional to the inverse square of the distance. The IDW interpolation method has been used to derive the areal precipitation distribution in some previous regional studies (Yong et al., 2010; Tong et al., 2014), but uncertainty still exists in the interpolated precipitation field particularly over West China with relatively sparse gauge networks. To reveal the probable



**Fig. 1.** (a) Number of gauges over Mainland China used as reference in the study at a typical hour (8 a.m., June 27, 2014). Six sub-regions are outlined for regional averaging; (b-d) spatial distributions of mean daily precipitation over Mainland China for the period April–December 2014, from gauge observations, IMERG, and 3B42V7 precipitation products.

uncertainty involved in the evaluation of satellite precipitation, the products based on the Biharmonic Spline Interpolation (BSI) method (Sandwell, 1987) which were also provided by CMA were used as the other ground reference in Section 4.1. All the gauge data have undergone strict quality control in three levels, which include (1) the extreme values' check, (2) internal consistency check, and (3) spatial consistency check (Shen et al., 2010). For grid boxes with gauges, the observed precipitation values are exactly the gauge observation or the averaged observation when more than one gauge locates in a grid, and used as the benchmark in grid-scale comparison and metric calculation. The number of gauges changes over time, ranging between 2200 and 2800 for each hour. The gauge distribution shown in Fig. 1(a) is for a typical hour (8 a.m., June 27, 2014 UTC), which represents the median gauge number for all the hours between April and December 2014. Most gauges are located over the eastern and southern parts of Mainland China, and relatively sparse gauge network are located across the northern and western parts, especially over the Tibetan Plateau. The limited number of gauges could be a source of error in evaluation of satellite precipitation products in such areas (Shen et al., 2014).

### 2.3. Satellite precipitation products

The latest Version-7 post-real-time 3B42 (hereafter referred as 3B42V7) product of TMPA at  $0.25^\circ \times 0.25^\circ$ , 3-hourly resolutions was used in this study. The 3B42V7 product provides precipitation

estimates between  $50^\circ\text{N}$ – $S$ , combining remote sensing data from various MW and IR sensors, and GPCC monthly gauge analysis (Huffman et al., 2007, 2010; Huffman and Bolvin, 2015). The TRMM Combined Instrument (TCI) estimate (TRMM 2B31 product) employed data from PR and TMI onboard the TRMM satellite to provide the best rain estimates for TRMM (Huffman and Bolvin, 2015). The 2B31 product was used as a satellite calibrator in the production of 3B42 and 3B43, until October 2014 when the routine production of TRMM PR precipitation estimates ended due to the ongoing descent of the TRMM satellite (Huffman and Bolvin, 2015), and thus the production of TCI was terminated. Climatological satellite calibrations used in the real-time TMPA (3B40RT, 3B41RT and 3B42RT) products were used to succeed TCI as the calibrator of 3B42V7 since October 2014, which could bring a discontinuity with previous data. The TMI data was still used as an input to the production of 3B42V7, until the instrument was turned off on 8 April, 2015, marking the full decommission of TRMM. The TMPA products, including 3B42V7, will continue until early 2017, when the retrospective processing of its IMERG will probably be carried out (Huffman et al., 2015). The 3B42V7 data used in this study were downloaded from the Precipitation Measurement Missions (PMM) website (<http://pmm.nasa.gov/data-access/downloads/trmm>).

IMERG is the Level 3 multi-satellite precipitation algorithm of GPM, which combines intermittent precipitation estimates from all constellation microwave sensors, IR-based observations from geosynchronous satellites, and monthly gauge precipitation data

**Table 1**List of the statistical metrics used in the evaluation and comparison<sup>a</sup>.

Statistic metrics	Equation	Perfect value
Correlation Coefficient (CC)	$CC = \frac{\frac{1}{N} \sum_{n=1}^N (S_n - \bar{S})(G_n - \bar{G})}{\sigma_S \sigma_G}$	1
Mean Error (ME)	$ME = \frac{1}{N} \sum_{n=1}^N (S_n - G_n)$	0
Relative Bias (BIAS)	$BIAS = \frac{\frac{1}{N} \sum_{n=1}^N (S_n - G_n)}{\sum_{n=1}^N G_n} \times 100\%$	0
Root Mean Squared Error (RMSE)	$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (S_n - G_n)^2}$	0
Probability of Detection (POD)	$POD = \frac{n_{11}}{n_{11} + n_{01}}$	1
False Alarm Ratio (FAR)	$FAR = \frac{n_{01}}{n_{11} + n_{01}}$	0
Critical Success Index (CSI)	$CSI = \frac{n_{11}}{n_{11} + n_{01} + n_{10}}$	1

<sup>a</sup> Notation:  $n$  represents number of samples;  $S_n$  represents satellite precipitation estimate;  $G_n$  represents gauge observed precipitation;  $\sigma_G$  represents standard deviations of gauge precipitation;  $\sigma_S$  represents standard deviations of satellite precipitation.  $n_{11}$  represents precipitation observed by the gauge and satellite simultaneously;  $n_{10}$  represents precipitation observed by the satellite but not observed by the gauge;  $n_{01}$  is contrary to  $n_{10}$ ;  $n_{00}$  represents precipitation observed neither by the gauge nor the satellite.

(Hou et al., 2014). Currently, IMERG is at its very early Day-1 stage. IMERG employs the 2014 version of the Goddard Profiling Algorithm (GPROF2014) to compute precipitation estimates from all passive microwave (PMW) sensors onboard GPM satellites, which is an improvement compared with TMPA (GPROF2010) (Huffman et al., 2014, 2015). IMERG “Final” run combines the GPCC Monitoring Product (currently Version 4) in the product, whose data source is limited to the Global Telecommunications System (GTS) with only about 7000 stations over the globe. The Full Data Reanalysis (currently Version 6) involves much more stations than the Monitoring Product, but only covers the period 1901–2010. The IMERG data were also downloaded from the PMM website (<http://pmm.nasa.gov/data-access/downloads/gpm>).

In terms of the independence of comparison between gauge- and satellite-based products, GPCC only use data from 194 China's International Exchange Stations over the whole country (Shen et al., 2013), which merely account for small parts of the total gauges (<9%) used in our study. Therefore, the evaluation and comparison in the study are based on more than 91% independent gauge stations. In addition, both TRMM and GPM research level products are combined with GPCC gauge data at the monthly scale, while the comparison in our study was conducted at sub-daily and daily scales.

#### 2.4. Metrics

To evaluate IMERG and 3B42V7 precipitation products comprehensively, seven metrics were selected, which could be generally divided into three categories (Yong et al., 2010). The first category includes the correlation coefficient (CC), describing the agreement between satellite estimates and gauge observations. The second category includes the mean error (ME), the relative bias (BIAS), and the root mean square error (RMSE), which are used to describe the error and bias of satellite estimates compared with gauge observations. And the third category includes the probability of detection (POD), the false alarm ratio (FAR), and the critical success index (CSI), which are used to describe the contingency of satellite precipitation estimates. POD gives the fraction of precipitation events that the satellite detects among all the actual precipitation events, and FAR gives the fraction of unreal events among all the events satellites detect (Ebert et al., 2007; Wilks, 2006). The CSI can actually be expressed as a function of POD and FAR (Gerapetritis and Pelissier, 2004), combining the characteristics of false alarms and missed events, and is therefore a more balanced score. Considering the rain gauge's detection resolution (0.1 mm/h), the rain/no rain threshold was set to 0.1 mm for hourly metrics, and 0.3 mm for 3-hourly metrics, and 2.4 mm for daily metrics. Formulas and perfect values of those metrics are listed in Table 1.

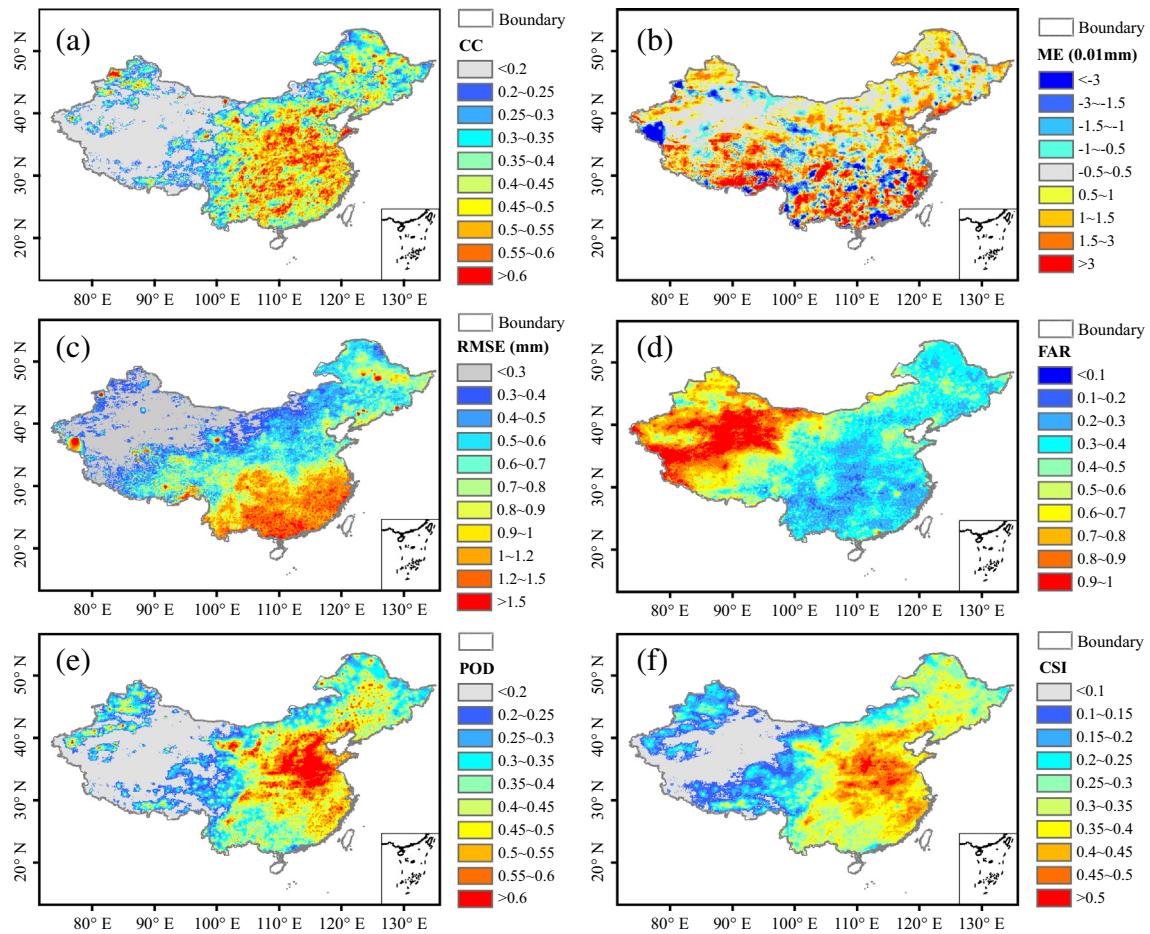
Please note that all the expression for precipitation intensity in this study is the accumulation of precipitation over the corresponding time period, e.g., for the 3-hourly and daily precipitation, the expression for intensity is the volume of precipitation during the 3 h and day, respectively. Therefore the unit is “mm”, instead of “mm/h” or “mm/day”. Another thing to be noted is that in the grid-scale comparison, only grid boxes with at least one gauge were taken to calculate the metrics to reduce error and bias arising from the interpolation. Such metrics were presented and analyzed in the paper context. However, every grid box over Mainland China was assigned a calculated metric value to enable the visualization of distribution (Figs. 2 and 5–7).

### 3. Evaluation of IMERG hourly estimates

#### 3.1. Grid-scale evaluation

Spatial distributions of metrics for IMERG hourly precipitation at  $0.1^\circ \times 0.1^\circ$  resolution over Mainland China are shown in Fig. 2. IMERG agreed well with the gauge data over East and South Mainland China, which are equipped with more gauges (Fig. 2(a)). However, the CC was a little lower over North China, and even under 0.2 over West China. Several factors could contribute to relatively low CC of IMERG products over such areas: (1) the topography and climate over West China are complex, posing a great challenge for accurate satellite precipitation estimation (Dinku et al., 2007); (2) few gauges are used in the production of GPCC monthly gauge analysis, and thus the quality of IMERG products is potentially degraded; (3) the interpolated precipitation could deviate far from the real values due to the sparse gauge networks, which would reduce reliability of metrics calculated against such ground reference (Shen and Xiong, 2015). The ME was generally small over most parts of Mainland China, and characterized by an alternate distribution of positive and negative values. Distributions of FAR, POD and CSI showed similar spatial characteristics to CC and RMSE, indicating that IMERG performed better over East and South China. All the metrics (except ME and BIAS) showed similar pattern with the precipitation intensity distribution (Fig. 1(b) and (c)). This phenomenon is reasonable because those regions have greater precipitation intensities and relatively flatter topography than other parts of Mainland China, which is beneficial to precipitation estimation.

Fig. 3 shows box plots of metrics over grid boxes with at least one gauge for IMERG and 3B42V7 precipitation products at hourly, 3-hourly, and daily scales. The upper and lower edges of the central box are the first and third quartiles (25% and 75%, respectively), and the band inside the box is the median. The ends of the outliers can represent several possible alternative values, which were



**Fig. 2.** Spatial distributions of metrics for IMERG hourly precipitation at  $0.1^\circ \times 0.1^\circ$  resolution over Mainland China: (a) CC, (b) ME, (c) RMSE, (d) FAR, (e) POD, and (f) CSI.

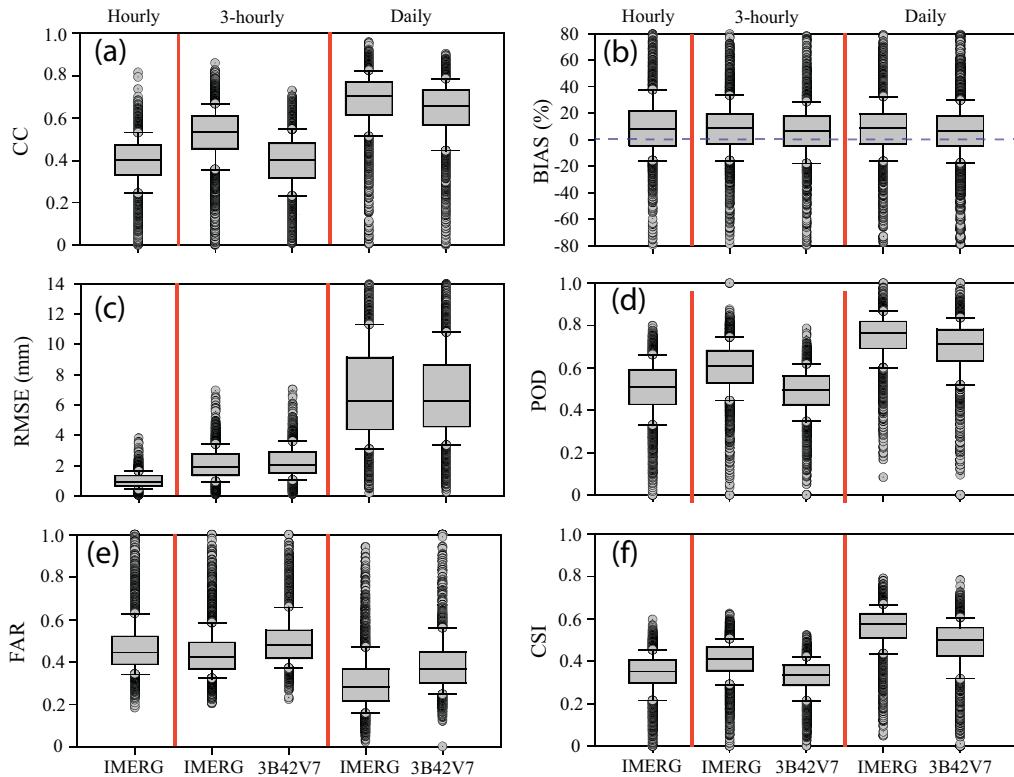
defined as the 10th percentile and the 90st percentile in this study. All the metrics are included in Fig. 3, except ME which varies greatly as the temporal resolution increases from hourly to daily, and was therefore not suitable to be presented in the boxplots. The distributions of the metrics for hourly IMERG precipitation showed strong symmetry, and the values of the first/third quartiles and upper/lower end of outliers were close to the median, indicating that the majority of these metrics were concentrated. Note that some outliers were far from the median, particularly for CC, BIAS, POD, and FAR, which suggested that IMERG precipitation deviated from gauge observations greatly over the corresponding grid boxes for the outliers (Fig. 3). The CC was low (0.40), and the RMSE was relatively high (1.06 mm) for the evaluation of IMERG at the grid scale. The ME and BIAS were 0.01 mm and 9.32%, respectively, which could be attributed to the offset of positive/negative biases for IMERG (Fig. 2(b)). The POD, FAR, and CSI were 0.50, 0.46 and 0.35, respectively. Based on the metrics above, the quality of IMERG precipitation at the hourly resolution is limited, but seems to be acceptable at the hourly scale over Mainland China.

### 3.2. Regional and national-scale evaluations

Fig. 4 shows scatterplots of precipitation comparisons for Mainland China at hourly, 3-hourly, and daily scales between IMERG, 3B42V7, and gauge. Precipitation values are calculated by averaging data from IMERG, 3B42V7, and gauge interpolation field over Mainland China at the three timescales, respectively. As shown in Fig. 4 (a), hourly IMERG precipitation agrees well with gauge observations

over Mainland China. A statistical summary of metrics for IMERG and 3B42V7 products at hourly (IMERG only), 3-hourly, and daily resolutions over Mainland China and six sub-regions is shown in Table 2. The indices were based on mean areal precipitation over Mainland China and sub-regions. Generally, IMERG overestimated hourly precipitation slightly. The low ME (0.10 mm) and BIAS (8.64%) were probably caused by the cancelation of positive/negative values. The national CC was 0.90 and the RMSE was 0.03 mm. The POD, FAR and CSI (0.86, 0.19 and 0.72, respectively) were also very satisfying, indicating that IMERG can capture most precipitation events over Mainland China.

Regional evaluation over six sub-regions was also conducted to reveal the error characteristics of IMERG over Mainland China comprehensively. In general, metrics of IMERG hourly precipitation in the six sub-regions were worse than those for Mainland China (Table 2). The IMERG product showed the best metrics in Regs2–4, corresponding to mid- and low-latitudes, and have relatively mild and wet climates. The CC in the above three sub-regions was greater than 0.85, and the CSI was greater than 0.7. Reg5 is the driest among the six sub-regions with the smallest RMSE and ME. Most metrics for Reg5 were the worst among the six sub-regions, including CC (0.50), POD (0.39), FAR (0.53), and CSI (0.27). Reg6 corresponds to the eastern part of the Tibetan Plateau with the highest altitudes among the six sub-regions, in which IMERG overestimated precipitation with the highest BIAS (20.69%) among the six sub-regions. Most gauges located in Reg6 are built at low-altitude sites and could probably lead to the underestimation of interpolated precipitation field over the high mountainous area.



**Fig. 3.** Box plots of metrics over grid boxes with at least one gauge for IMERG and 3B42V7 precipitation products at hourly, 3-hourly, and daily timescales: (a–f) represents CC, BIAS, RMSE, POD, FAR, and CSI, respectively.

But other metrics were generally good, indicating that IMERG can capture the general trend of precipitation well in Reg6. For Reg1, the metrics were a little worse than those in Reg6 except BIAS.

In conclusion, the IMERG product overestimated hourly precipitation slightly with small BIAS for grid-scale and Mainland China evaluations (9.32% and 8.14%, respectively), which was caused primarily by the cancellation of positive/negative values. But the great overestimation in high altitude area (Reg5) is notable. The performance of IMERG hourly precipitation estimates was not very satisfactory in terms of all the metrics in the grid-scale evaluation with small CC and CSI, but could still be acceptable considering the hourly temporal resolution. However, when it comes to the national comparison, all the metrics were improved greatly except ME and BIAS, indicating the stronger utility of IMERG to describe areal precipitation over large areas. As far as the regional evaluation over six sub-regions is concerned, IMERG showed better performance in Regs2–4 corresponding to mid- and low-latitudes, and a relatively mild and wet climate. Precipitation estimation over high latitudes (e.g., Reg1) and altitudes (e.g., Reg6), and dry area (e.g., Reg5) should be improved further in the future version of IMERG products, particularly for dry areas with relatively low precipitation intensity.

#### 4. Comparison between IMERG and 3B42V7 products

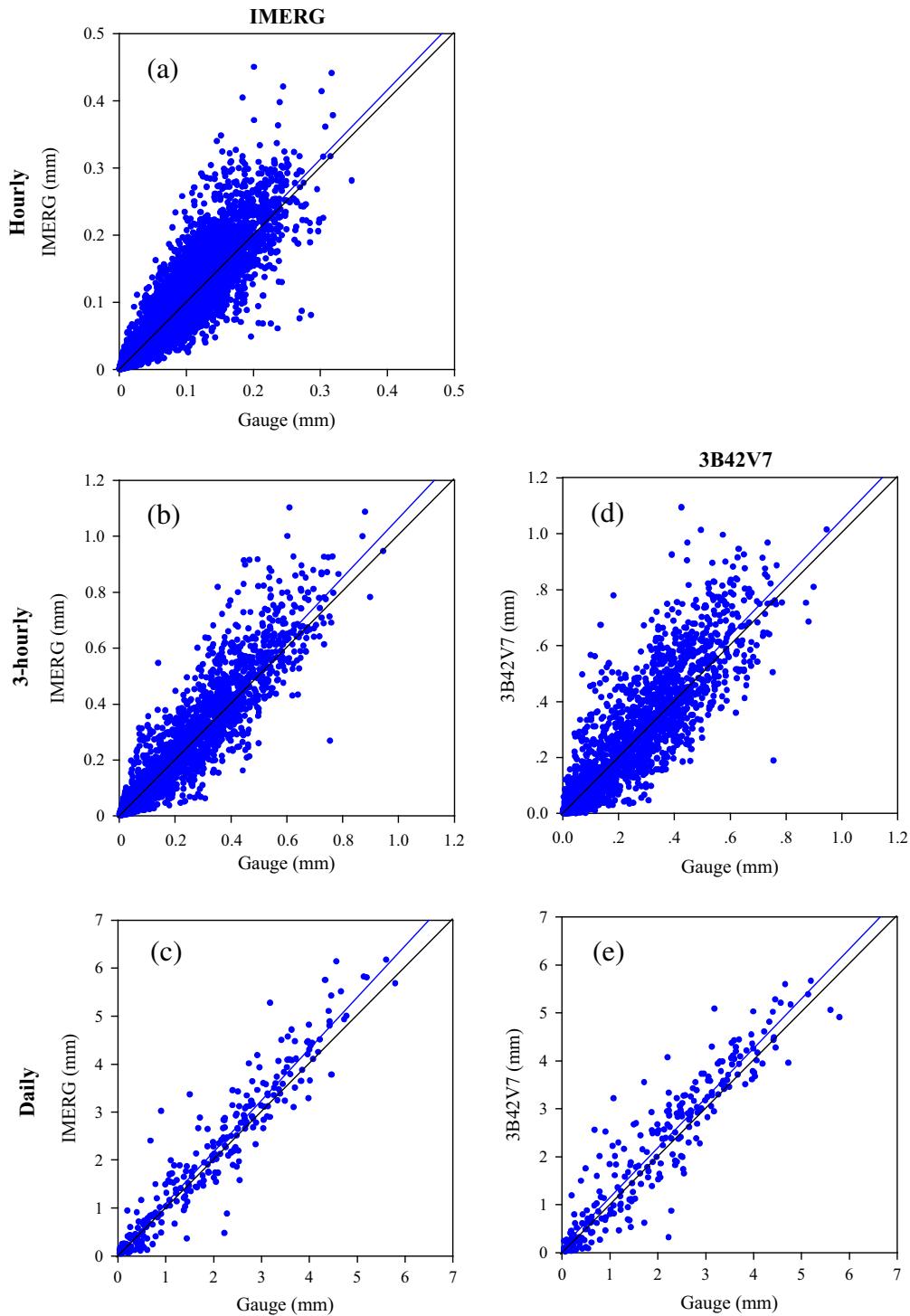
To explore the continuity and difference between the IMERG and 3B42V7 precipitation estimates, the two products were compared and evaluated at gridded, regional, and national scales. In addition, probability density functions (PDFs) of IMERG and 3B42V7 products with different intensities were compared to the PDF of the gauge observations over Mainland China. All the comparisons were conducted from April to December 2014 at

3-hourly and daily temporal resolutions, and  $0.1^\circ \times 0.1^\circ$  spatial resolutions. The IMERG and the interpolated gauge precipitation were aggregated to 3-hourly datasets, with the same time zone as 3B42V7, representing a nominal  $\pm 90$ -min span around the nominal hour (Huffman and Bolvin, 2015). In addition, the diurnal cycle of the three types of precipitation was calculated over the six sub-regions at the hourly (gauge and IMERG) and 3-hourly (3B42V7) resolutions from June to August 2014, corresponding to the warm season in most of Mainland China with relatively strong precipitation intensity and diurnal cycle. The 3B42V7 products were downscaled to the  $0.1^\circ \times 0.1^\circ$  resolution, the same as IMERG and the interpolated gauge precipitation, using the standard bilinear interpolation method suited for gridded datasets.

As shown in Fig. 1(c) and (d), both IMERG and 3B42V7 products capture the spatial distribution and magnitude of daily precipitation compared to the interpolated gauge precipitation, showing a decreasing trend from south to north, and from east to west. The 3B42V7 product is restricted to  $50^\circ\text{N}$ – $S$ , so the IMERG product and gauge observation beyond  $50^\circ\text{N}$  were precluded from this comparison.

#### 4.1. Grid-scale Comparison

Fig. 5 shows spatial distributions of CC, BIAS, and CSI computed from IMERG and 3B42V7 products against IDW interpolated results at 3-hourly and  $0.1^\circ \times 0.1^\circ$  resolutions over Mainland China. Theoretically, the BIAS of IMERG has the same distribution map with the same magnitude at hourly, 3-hourly, and daily scales, because the accumulation of precipitation would not change the relative bias. The ME of IMERG shows the same distribution, but the magnitude increases as temporal resolutions changing. Therefore, Fig. 7 is the same as Fig. 5 but for the daily metrics, except that the BIAS is replaced by ME, which could avoid redundant maps. The



**Fig. 4.** Scatterplots of precipitation comparisons for Mainland China at hourly (the first panel), 3-hourly (the second panel), and daily scale (the third panel) between IMERG and gauge (the first column), and 3B42V7 and gauge (the second column). The diagonal line is black, and the best fit line (using the least squares method) is blue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

distribution of CC seemed better for 3B42V7 than IMERG at the 3-hourly resolution (Fig. 5(a) and (b)). But at the daily resolution, IMERG performing better than 3B42V7 in North China. The CC values were improved greatly for daily precipitation (Fig. 7 (a) and (b)). The most striking improvement occurred in the sparse gauge region, which could be explained by the better representatively of the interpolated rain data with a sparse network over a longer time period. If the precipitation is given longer time to move

around, it has a better chance of being captured by one or more gauges, and so has a better chance of showing up in the interpolation. ME values were generally small in West and North China due to relatively small precipitation intensity (Fig. 7(c) and (d)). BIAS shows a different spatial distribution with most small values distributed over East and South China. The unusual large BIAS, as well as low CC and CSI over regions with sparse gauge networks, could be attributed to the poor interpolation quality.

**Table 2**

Summary of evaluation metrics for IMERG and 3B42V7 products at hourly (IMERG only), 3-hourly, and daily resolutions over Mainland China and six sub-regions. The metrics were calculated based on mean areal precipitation.

	Product	CC	RMSE (mm)	ME (0.1 mm)	BIAS (%)	POD	FAR	CSI
<i>China</i>								
Hourly	IMERG	0.90	0.03	0.10	8.14	0.86	0.19	0.72
3-hourly	IMERG	0.91	0.09	0.20	9.58	0.88	0.18	0.74
	3B42V7	0.87	0.11	0.30	10.56	0.86	0.21	0.70
Daily	IMERG	0.96	0.50	1.80	9.17	0.91	0.16	0.77
	3B42V7	0.94	0.56	2.00	10.14	0.91	0.23	0.72
<i>Reg1</i>								
Hourly	IMERG	0.74	0.12	0.05	6.25	0.73	0.24	0.59
3-hourly	IMERG	0.77	0.31	0.26	11.39	0.72	0.28	0.56
	3B42V7	0.67	0.38	0.22	9.68	0.71	0.30	0.54
Daily	IMERG	0.86	1.64	2.08	11.39	0.81	0.24	0.64
	3B42V7	0.78	1.83	1.76	9.61	0.79	0.28	0.61
<i>Reg2</i>								
Hourly	IMERG	0.88	0.10	0.09	11.81	0.85	0.20	0.70
3-hourly	IMERG	0.80	0.37	0.20	8.72	0.77	0.26	0.60
	3B42V7	0.78	0.38	0.23	10.06	0.75	0.33	0.55
Daily	IMERG	0.87	1.94	1.63	8.72	0.81	0.19	0.68
	3B42V7	0.85	1.94	1.88	10.06	0.82	0.24	0.65
<i>Reg3</i>								
Hourly	IMERG	0.85	0.19	0.09	4.93	0.84	0.11	0.76
3-hourly	IMERG	0.81	0.65	-0.19	-3.12	0.77	0.13	0.69
	3B42V7	0.80	0.67	-0.05	-0.81	0.76	0.14	0.68
Daily	IMERG	0.86	3.52	-1.55	-3.12	0.81	0.08	0.75
	3B42V7	0.85	3.66	-0.40	-0.81	0.84	0.10	0.77
<i>Reg4</i>								
Hourly	IMERG	0.88	0.15	0.10	5.31	0.82	0.10	0.75
3-hourly	IMERG	0.86	0.48	0.02	0.33	0.77	0.10	0.71
	3B42V7	0.82	0.49	-0.09	-1.53	0.77	0.11	0.70
Daily	IMERG	0.90	2.73	0.16	0.33	0.81	0.06	0.77
	3B42V7	0.88	2.54	-0.73	-1.53	0.81	0.07	0.77
<i>Reg5</i>								
Hourly	IMERG	0.50	0.05	0.02	7.25	0.39	0.53	0.27
3-hourly	IMERG	0.49	0.14	0.06	9.12	0.34	0.57	0.24
	3B42V7	0.45	0.14	-0.07	-11.31	0.29	0.55	0.22
Daily	IMERG	0.68	0.62	0.47	9.12	0.27	0.63	0.19
	3B42V7	0.61	0.64	-0.58	-11.31	0.09	0.67	0.08
<i>Reg6</i>								
Hourly	IMERG	0.78	0.07	0.16	20.69	0.82	0.27	0.63
3-hourly	IMERG	0.79	0.20	0.52	23.86	0.82	0.28	0.62
	3B42V7	0.73	0.26	0.62	28.40	0.77	0.28	0.59
Daily	IMERG	0.89	1.05	4.18	23.86	0.89	0.19	0.74
	3B42V7	0.85	1.31	4.98	28.40	0.93	0.20	0.75

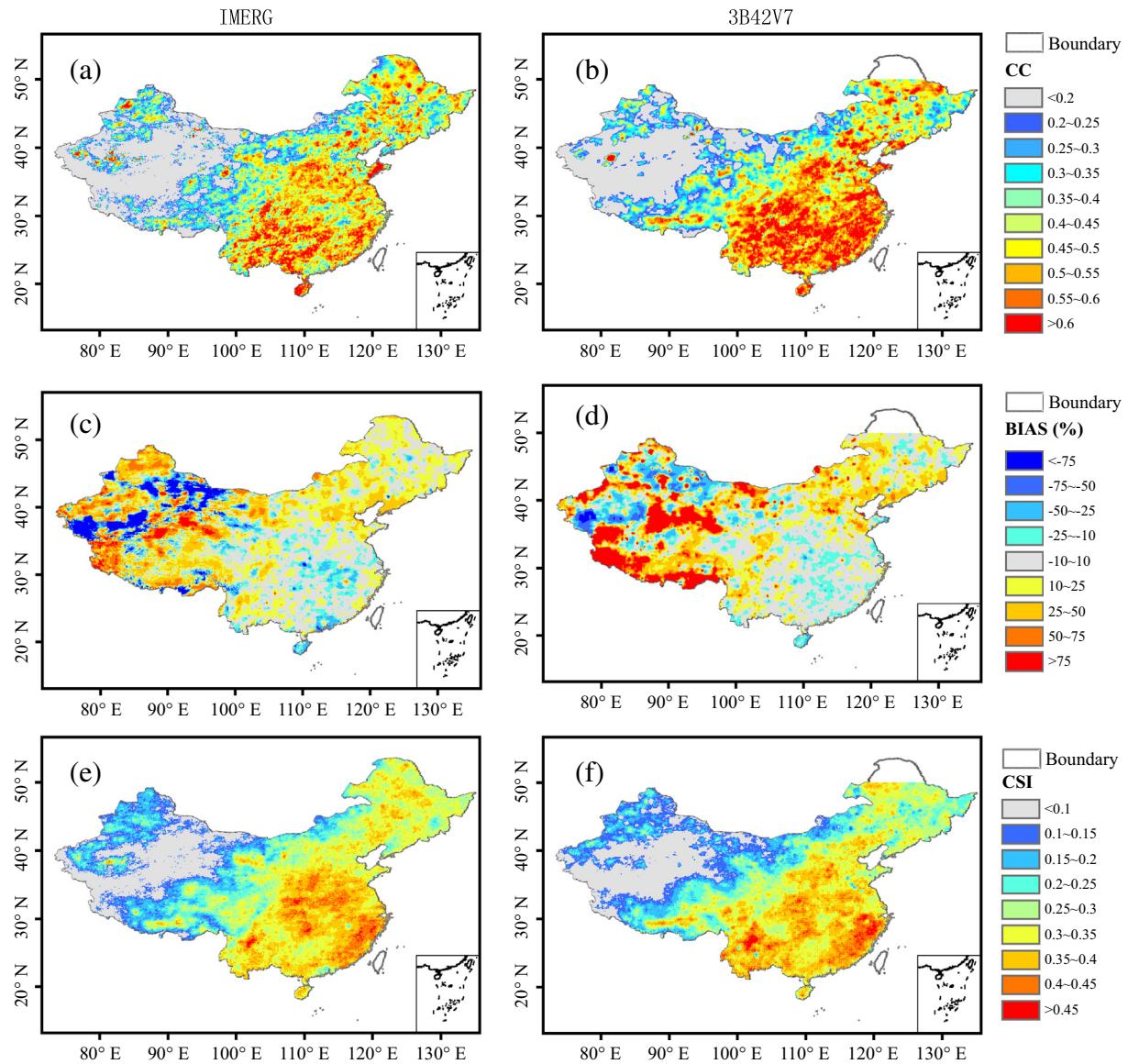
**Fig. 6** shows spatial distributions of CC, ME, BIAS, and CSI computed from IMERG and 3B42V7 products against BSI results at the 3-hourly scale. The spatial pattern of CC and CSI was similar with that in **Fig. 5**. However, the ME and BIAS of IMERG and 3B42V7 (**Fig. 6(c)–(f)**) shows greater underestimation rather than overestimation which was different with **Fig. 5(c)–(f)** due to the larger interpolated precipitation. Comparing **Fig. 6(c)** and **(d)** to the gauge distribution (**Fig. 1(a)**), it is notable that severe underestimation usually occurs over areas without gauges in West China, particularly for the Tibetan Plateau. In addition, previous studies showed that TMPA 3B42 products tended to overestimate precipitation over high-latitude areas ([Yong et al., 2010, 2012, 2014](#)), while **Fig. 6(d)** shows underestimation dominated over North China. The characteristic of BSI to find the smoothest surface could be a possible explanation for its poor performance, because precipitation varies greatly over space but the precipitation field is actually not smooth. Therefore, we considered that the BSI method is not suited for regions with sparse gauge networks.

The CC of IMERG was 0.53 and 0.71 at 3-hourly and daily resolutions over Mainland China, respectively, higher than 0.42 and 0.68 for 3B42V7 (**Fig. 3**), indicating higher ability of IMERG to capture temporal variations with reference to in situ data. Both IMERG and 3B42V7 products slightly overestimated precipitation with

positive ME and BIAS at 3-hourly and daily resolutions, and the mean value of BIAS for 3B42V7 is a little smaller than that for IMERG. The slightly worse performance IMERG than 3B42V7 in terms of BIAS and RMSE could be attributed to error accumulation from the original half-hourly resolution to 3-hourly. All the contingency metrics of IMERG were better than those of 3B42V7 at 3-hourly and daily resolutions, indicating that the IMERG product could capture precipitation events better than the 3B42V7 product.

#### 4.2. Regional and national-scale comparisons

The IMERG and 3B42V7 precipitation products were also compared over the whole country and six sub-regions at 3-hourly and daily resolutions (**Fig. 4(b)–(e)**), with the metrics calculated using mean values of the precipitation estimates from satellite products and interpolated gauge observations over Mainland China. The metrics for Mainland China were improved greatly compared with the grid-scale comparison for either product. The CC of IMERG was 0.91 and 0.96 at 3-hourly and daily resolutions, respectively, better than that of 3B42V7 with values of 0.87 and 0.94 (**Table 2**). Higher CC values indicated that both IMERG and 3B42V7 products could capture temporal variations better at the national scale than gridded scale. The RMSE was reduced to



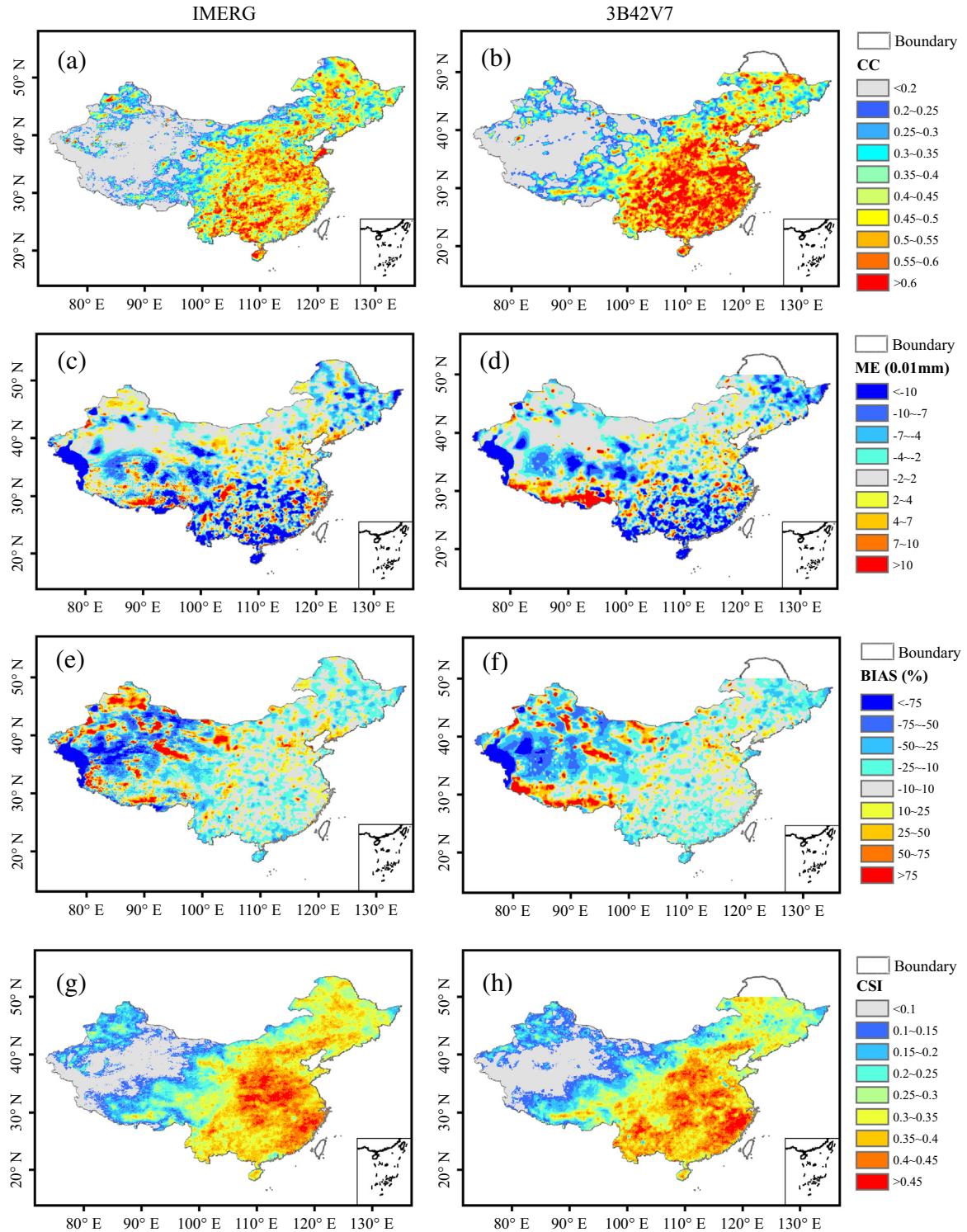
**Fig. 5.** Spatial distributions of metrics for IMERG (the left panel) and 3B42V7 (the right panel) precipitation at 3-hourly and  $0.1^\circ \times 0.1^\circ$  resolution over Mainland China: (a and b) CC, (c and d) BIAS, and (e and f) CSI.

0.01 mm for the 3-hourly comparison and 0.50 mm for the daily comparison, compared with 2.39 mm and 7.82 mm for the grid-scale comparison at the corresponding temporal resolutions. POD ( $\sim 0.9$ ), FAR ( $\sim 0.2$ ), and CSI ( $>0.7$ ) were also improved greatly, indicating that both IMERG and 3B42V7 could capture most precipitation events correctly over Mainland China. The ME and BIAS remained similar, because the two metrics were influenced greatly by the cancellation of negative and positive values (Figs. 2(b), 5(c) and (d) and 6(c) and (d)).

Regional performance of IMERG and 3B42V7 products based on mean areal precipitation varied over the six sub-regions. In general, both products showed better performance in Regs2–4 and worse performance in other three sub-regions at 3-hourly and daily resolutions. The characteristics of metrics were similar with the analysis in Section 3.2 for hourly IMERG data, and IMERG was always better than 3B42V7 concerning most metrics (Table 2). The metrics of 3B42V7 at the 3-hourly resolution are even worse than those of IMERG at the hourly resolution, but the gap between the two products is narrow. The POD, FAR and CSI for 3B42V7 at the daily resolution were anomalous, and even worse than those

at the 3-hourly resolution, because the precipitation events beyond the threshold (2.4 mm/d) were quite rare in Reg5, and only one event was captured by 3B42V7, resulting in the unreasonable contingency metric values. Similar problems existed in consistency metrics of Reg5 at the daily resolution with only around 7% precipitation events above 2.4 mm/d. For other sub-regions which are not as dry as Reg5, the effective samples were generally enough for metric calculation, and therefore the consistency metrics is effective. Both IMERG and 3B42V7 showed great overestimation in Reg6 with the BIAS up to 20%, but the CC and contingency metrics were quite well ranking only second to Regs2–4.

As in Section 3.2, we recommended that for Reg1 and Regs5–6 metrics calculated over grid boxes with at least one gauge would be more informative and convincing. Reg5 still showed the lowest metrics at both 3-hourly and daily resolutions. 3B42V7 showed the largest BIAS (20.24% at the 3-hourly resolution and 15.31% at the daily resolution) in Reg6 among six sub-regions. In contrast, IMERG showed much smaller BIAS in Reg6 (reduced from  $\sim 24\%$  in regional comparison to  $\sim 7\%$  in grid-scale comparison). For dry Reg5, IMERG showed slight overestimation, while 3B42V7 showed

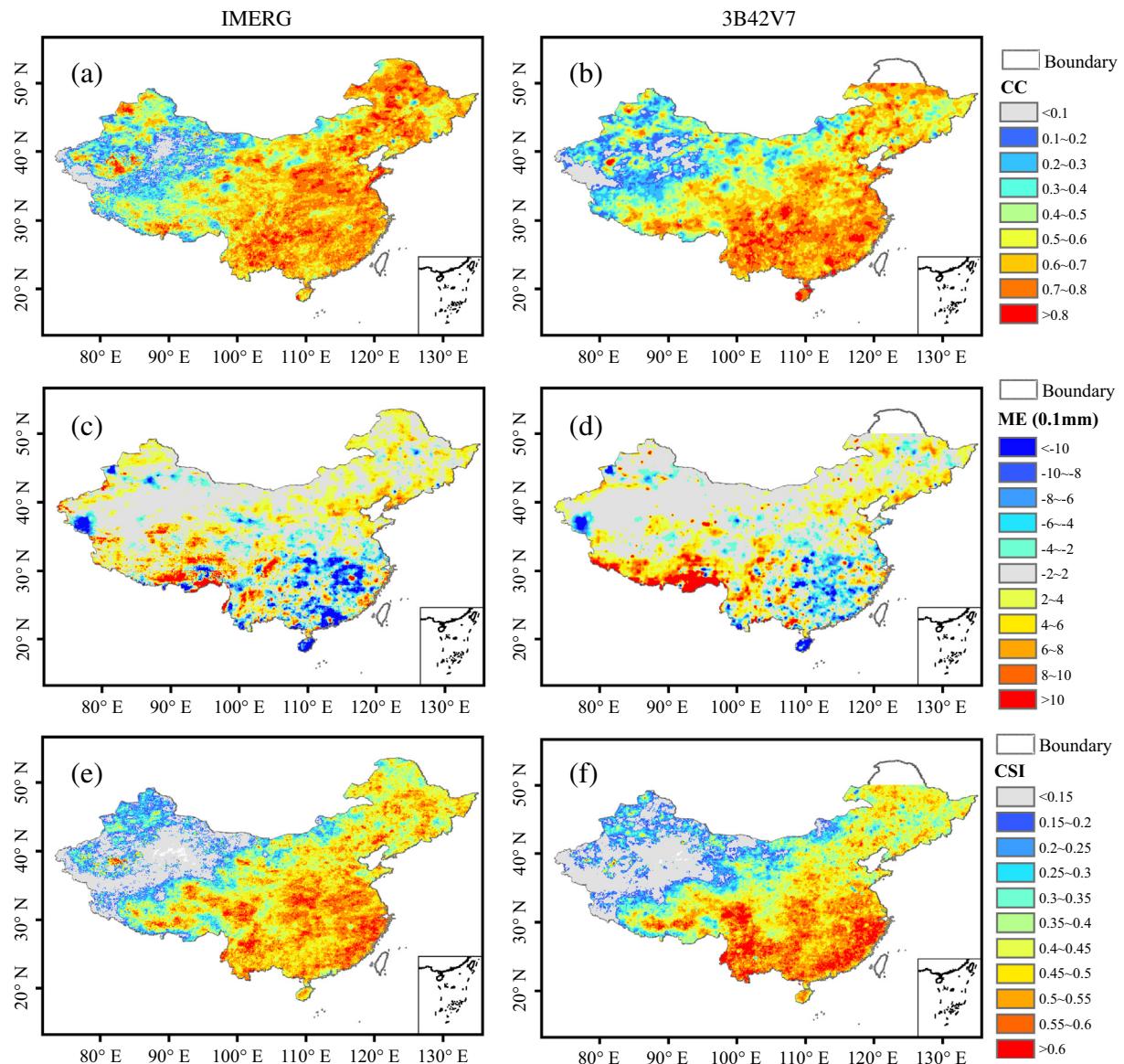


**Fig. 6.** Spatial distributions of metrics for IMERG (the left panel) and 3B42V7 (the right panel) precipitation against biharmonic spline interpolated results at 3-hourly and  $0.1^\circ \times 0.1^\circ$  resolution over Mainland China using : (a and b) CC, (c and d) ME, (e and f) BIAS, and (g and h) CSI.

underestimation, with the absolute value of BIAS being smaller for IMERG. The phenomenon indicated that IMERG was indeed improved compared with 3B42V7 in regions with smaller precipitation intensity and high altitudes. Generally, IMERG outperformed 3B42V7 as far as most metrics for each sub-region were concerned, and the improvement at the 3-hourly resolution was more significant than at the daily resolution.

All the metrics (except ME and BIAS) were improved as the temporal resolution increased from 3-hourly to daily (hourly to daily

for IMERG) at gridded, regional and national scales. For example, CC of IMERG increased by 1.1% (104%) from hourly to 3-hourly scales, and 3.3% (34%) from 3-hourly to daily scales in national (gridded) comparison over Mainland China. At the national scale, the increasing percentage was smaller since CC has already been high ( $>0.9$ ). Similar patterns could also be found for IMERG and 3B42V7 in other comparisons. In addition, as the comparison changed from the gridded scale to the national scale, the improvement attributed to spatial upscaling was significant. For instance, the CSI



**Fig. 7.** Spatial distributions of metrics for IMERG (the left panel) and 3B42V7 (the right panel) precipitation at 3-hourly and  $0.1^\circ \times 0.1^\circ$  resolution over Mainland China: (a and b) CC, (c and d) ME, and (e and f) CSI.

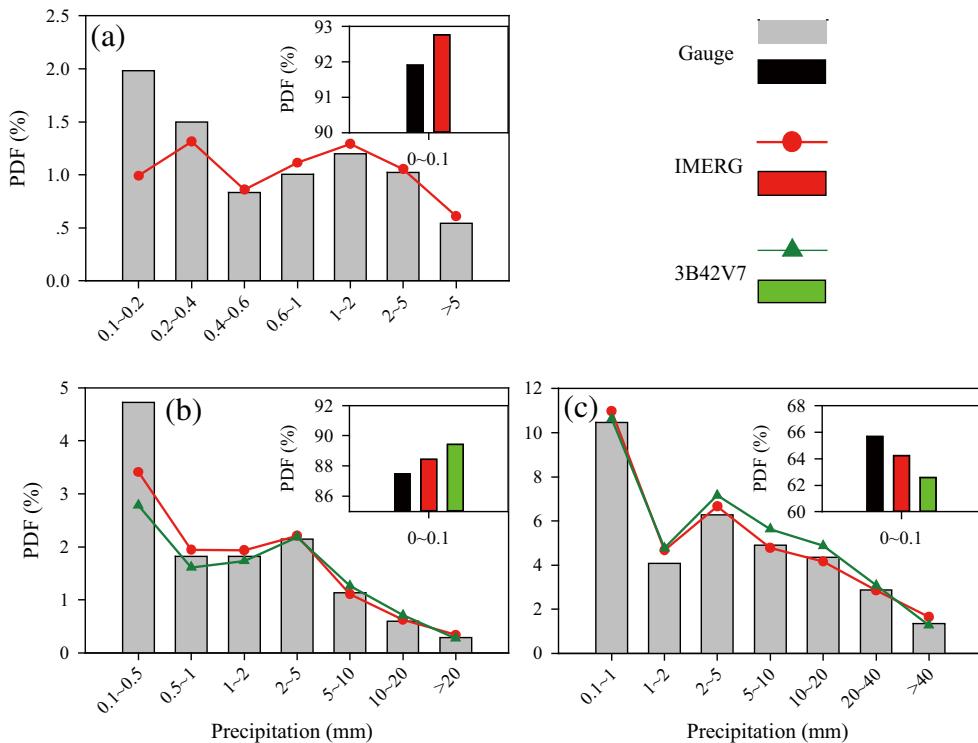
of IMERG increased by 140% from gridded (0.30) to national (0.72) scales over Mainland China.

#### 4.3. Probability density function of precipitation intensity

PDF represents the probability of rain-rate occurrence, and is used in many studies to evaluate the quality of satellite precipitation products (Kirstetter et al., 2013; Prat and Nelson, 2013; Shen et al., 2010; Wolff and Fisher, 2009). PDF tends to emphasize lighter rainfall rates (Kirstetter et al., 2013), which is suitable for examining the improved ability of IMERG to detect light precipitation. The PDFs of hourly, 3-hourly, and daily precipitation with different intensities over  $0.1^\circ \times 0.1^\circ$  grid boxes with at least one gauge were calculated from April to December 2014 (Fig. 8). The first bin of hourly, 3-hourly, and daily PDFs is displayed separately in the top right corner of Fig. 8(a)–(c), respectively, because most precipitation falls in the first bin representing no rain or drizzle.

For the hourly PDF shown in Fig. 8(a), only IMERG and gauge precipitation were included. Over 90% hourly precipitation was under 0.1 mm, and IMERG showed a little overestimation in this

bin. For precipitation with magnitudes ranging between 0.1 and 0.4 mm, the IMERG product showed obvious underestimation. Generally, for precipitation ranging between 0 and 0.2 mm, IMERG agreed well with gauge observations in terms of the frequency, but tended to underestimate the intensity. Further studies are needed to investigate the quality of IMERG estimates for light precipitation using more precise gauge data, which can detect precipitation smaller than 0.1 mm/h. For precipitation larger than 0.4 mm, IMERG could generally match the PDF of gauge with slight overestimation. As shown in Fig. 8(b), ~88% precipitation from gauge observation falls between 0 and 0.1 mm for the 3-hourly precipitation PDF. Both IMERG and 3B42V7 showed overestimation in the first bin, whereas IMERG was closer to the PDF of gauges, indicating the improved capability of IMERG in detecting no rain or drizzle events compared to 3B42V7, which could be attributed to the DPR carried by the GPM Core Observatory with higher sensitivity at light rain rates than PR onboard the TRMM satellite (Hou et al., 2014). In terms of precipitation ranging between 0.1 and 0.5 mm, both IMERG and 3B42V7 showed slight underestimation (1.3% for IMERG and 2% for 3B42V7), and IMERG is still closer to



**Fig. 8.** Probability density function (PDF) of (a) 1-hourly, (b) 3-hourly, and (c) daily precipitation with different intensities, as derived from IMERG and 3B42V7 products, and the gauge observations from April to December 2014.

the PDF of gauges. The 3-hourly IMERG PDF was similar to the hourly PDF for light precipitation with correct frequency but underestimated intensity. Both products agreed well with gauge observations with respect to the PDF of various precipitation intensities beyond 0.5 mm.

For the PDF of daily precipitation, both IMERG and 3B42V7 showed underestimation (~1.5% and ~3%, respectively) for precipitation between 0 and 0.1 mm, while IMERG was closer to the PDF of gauge (Fig. 8(c)). The IMERG product slightly overestimated light precipitation (0.1–5 mm), but generally captured the PDF for precipitation beyond 5 mm. The 3B42V7 product overestimated precipitation between 1 and 20 mm. Both products showed similar PDF structures for light (0.1–1 mm) and heavy (>20 mm) precipitation.

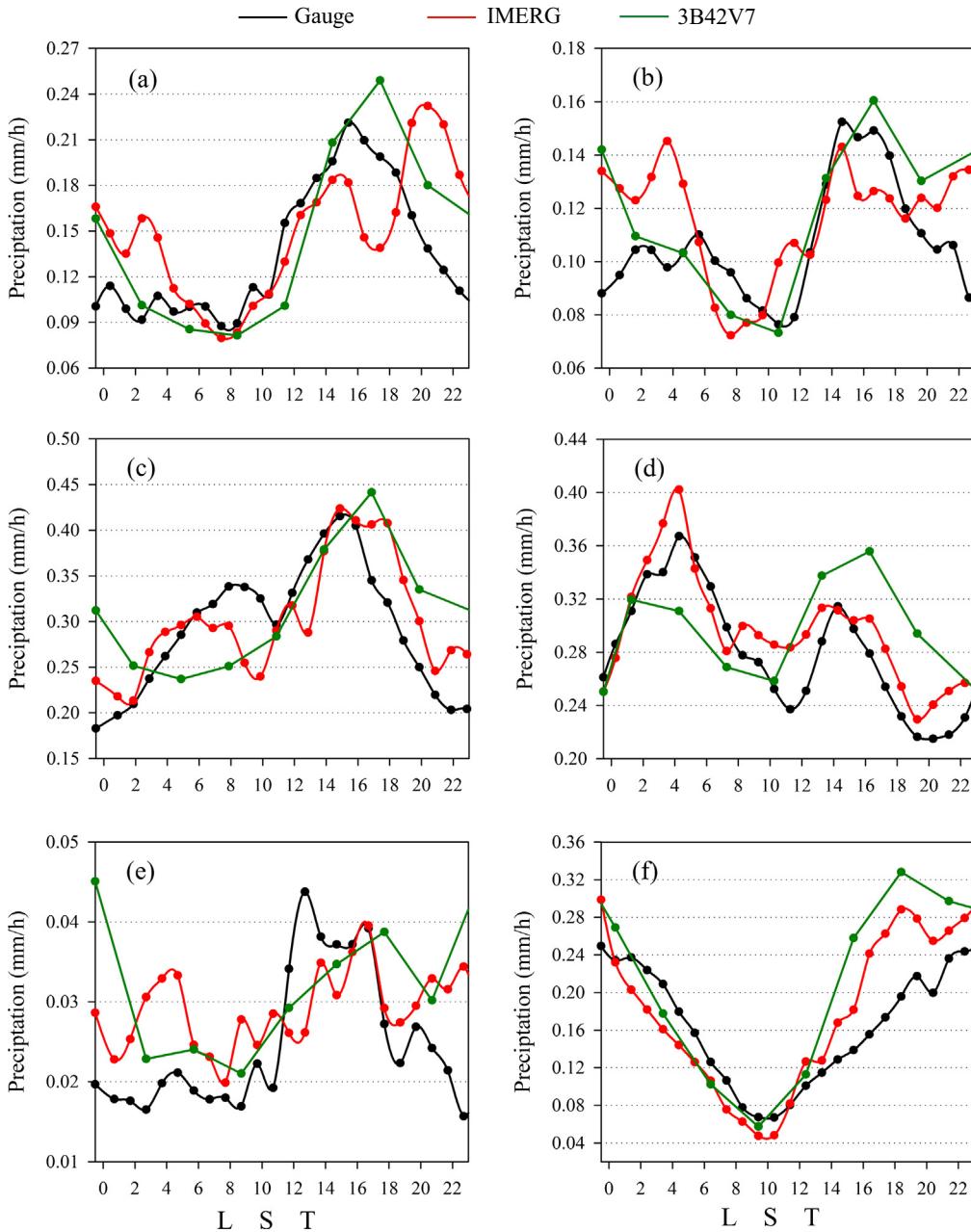
In conclusion, IMERG can reproduce the PDF structures of various precipitation intensities at hourly, 3-hourly, and daily scales, particularly for heavy precipitation, and performs better than 3B42V7 at daily and sub-daily scales. For light precipitation defined in our study, IMERG can reproduce the total frequency but underestimate the intensity, which should be noted by the algorithm developers.

#### 4.4. Diurnal cycle

The diurnal cycle is one of the most important characteristics of precipitation, atmosphere, and climate with large spatial and seasonal variations (Shen et al., 2010; Yuan et al., 2012), and has been investigated in many studies (Dai et al., 2007; Yu et al., 2007; Yuan et al., 2012; Zhou et al., 2008). This study focuses on the diurnal cycle from June to August 2014 (92 days in total), because most precipitation occurs during this period in most parts of Mainland China (Tao, 1987; Zhou et al., 2008), which is the summer time in China and shows the strongest diurnal cycle (Dai et al., 2007; Yu et al., 2007; Zhou et al., 2008).

The diurnal cycles of the six sub-regions were calculated using hourly (gauge and IMERG) and 3-hourly (3B42V7) precipitation data over  $0.1^\circ \times 0.1^\circ$  grid boxes with at least one gauge (Fig. 9). Precipitation values are the mean of 92 days at specific hours, and converted from UTC to the Local Standard Time (LST) at the center latitude for each sub-region, respectively. The time zone conversion is a bulk correction to each region rather than correction to each gauge, which could result in some deviation. We deduced that the deviation is acceptable due to the limited longitudinal extent of sub-regions, and all the conversion procedure is the same for 3B42V7, IMERG and gauge data.

The gauge data showed a large precipitation peak around 1500 LST at Reg1, Reg3, and Reg5. For Reg1, IMERG and 3B42V7 data generally captured the trend, but both products showed severe lag in terms of the peak time, with the peak of 3B42V7 data around 1700 LST, and the peak of IMERG data around 2000 LST. IMERG almost got a valley when gauge and 3B42V7 had peak values around 1700 LST, which could be caused by the poor performance of hourly IMERG over Reg1 with CC only being 0.10 in the grid-scale comparison (Table 3). For Reg3, IMERG data showed close peak to gauge data, while 3B42V7 still showed peak around 1700 LST. In addition, both products overestimated precipitation around the mid-night and early morning in Reg1 and Reg3. Reg5 is the driest sub-region in this study, with the mean hourly precipitation from June to August lower than 0.15 mm. Both IMERG and 3B42V7 failed to reproduce the diurnal cycle structure in this region, showing underestimation for the peak from 11:00 to 18:00 LST, and overestimation for the remaining time in a day. According to the gauge data, two precipitation peaks around 6:00 and 16:00 LST existed in Reg2, but only the last one with higher intensity was captured by IMERG and 3B42V7. A similar problem existed in Reg2 as in Reg1 and Reg3, the overestimation of precipitation around mid-night and early morning. In Reg4,



**Fig. 9.** Warm season (June–August) diurnal cycle of precipitation (mm/h) over six sub-regions (Fig. 1(a)): (a–f) represents Reg1–Reg6, respectively. Results are derived from hourly (Gauge and IMERG) and 3 hourly (3B42V7) precipitation estimates, and converted to LST in the figures.

IMERG agreed well with the diurnal trend of gauge quite in both time and amplitude, while 3B42V7 showed great overestimation from the afternoon to mid-night. Reg6 is the eastern part of the Tibetan Plateau, at which the peak occurred in the mid-night and the valley occurred at noon. Both IMERG and 3B42V7 captured the time of peak and valley, but show great overestimation from the late afternoon to mid-night.

In general, the IMERG product shows improved capability in reproducing the diurnal cycle compared to the 3B42V7 product, but the improvement is not significant. Both products have a long way to go in reconstruction of the diurnal cycle, particularly for regions with dry climates (e.g., Reg5) and high latitudes (e.g., Reg1). Furthermore, the peak of the diurnal cycle predicted by satellite products lags that of rain gauge measurements over North and East Mainland China (e.g., Reg1, Reg2, and Reg3), with IMERG showing the greater lag than 3B42V7 in Reg1.

## 5. Discussion and summary

In this study, we evaluated the quality of the Day-1 IMERG precipitation product at unprecedented hourly and  $0.1^\circ \times 0.1^\circ$  resolutions, using hourly precipitation observation data from over 2200 gauges across Mainland China as reference. Then IMERG and 3B42V7 products were compared at 3-hourly and daily temporal resolutions to explore the continuity and difference between IMERG and 3B42V7. In addition, the PDFs of 3-hourly and daily precipitation with different intensities were used to examine the capability of IMERG and 3B42V7 to detect precipitation frequency of occurrence. Finally, the diurnal cycles were investigated to demonstrate both products' capability of reproducing precipitation temporal evolution in a given day. Further studies are needed to evaluate the performance of IMERG against other satellite and reanalysis products, e.g., CMORPH.

**Table 3**

Summary of evaluation metrics for IMERG and 3B42V7 products at hourly (IMERG only), 3-hourly, and daily resolutions over Mainland China and six sub-regions. The metrics were calculated based on grids with at least one gauge.

	Product	CC	RMSE (mm)	ME (0.1 mm)	BIAS (%)	POD	FAR	CSI
<i>China</i>								
Hourly	IMERG	0.26	1.40	0.06	4.59	0.44	0.51	0.30
3-hourly	IMERG	0.53	2.39	0.23	5.93	0.58	0.43	0.40
	3B42V7	0.42	2.59	0.18	4.68	0.49	0.48	0.34
Daily	IMERG	0.71	7.82	1.72	5.66	0.70	0.30	0.54
	3B42V7	0.68	7.64	1.30	4.27	0.70	0.36	0.50
<i>Reg1</i>								
Hourly	IMERG	0.10	1.99	-0.02	-2.07	0.40	0.56	0.27
3-hourly	IMERG	0.41	1.97	0.15	6.37	0.59	0.44	0.40
	3B42V7	0.26	2.21	0.07	3.11	0.41	0.53	0.28
Daily	IMERG	0.63	5.48	1.29	7.22	0.67	0.30	0.52
	3B42V7	0.55	5.71	0.56	3.11	0.61	0.39	0.44
<i>Reg2</i>								
Hourly	IMERG	0.28	0.83	0.12	15.76	0.58	0.51	0.36
3-hourly	IMERG	0.53	1.55	0.35	15.05	0.73	0.45	0.46
	3B42V7	0.38	1.80	0.32	13.73	0.57	0.51	0.36
Daily	IMERG	0.71	5.03	2.66	14.57	0.81	0.36	0.56
	3B42V7	0.65	5.34	2.43	13.30	0.78	0.44	0.48
<i>Reg3</i>								
Hourly	IMERG	0.30	1.53	0.09	4.69	0.48	0.47	0.34
3-hourly	IMERG	0.55	2.87	0.23	3.98	0.64	0.39	0.45
	3B42V7	0.43	3.28	0.29	5.04	0.53	0.42	0.38
Daily	IMERG	0.72	9.54	1.79	3.95	0.76	0.28	0.59
	3B42V7	0.71	9.46	2.31	5.09	0.76	0.31	0.57
<i>Reg4</i>								
Hourly	IMERG	0.33	1.55	0.10	5.07	0.41	0.48	0.30
3-hourly	IMERG	0.55	3.03	0.27	4.80	0.55	0.41	0.40
	3B42V7	0.45	3.22	0.17	2.96	0.47	0.44	0.34
Daily	IMERG	0.70	9.96	2.13	4.79	0.67	0.27	0.54
	3B42V7	0.69	9.60	1.33	3.00	0.68	0.33	0.51
<i>Reg5</i>								
Hourly	IMERG	0.20	0.32	0.00	0.46	0.26	0.72	0.16
3-hourly	IMERG	0.41	0.61	0.01	0.90	0.42	0.61	0.25
	3B42V7	0.22	0.64	-0.06	-8.49	0.28	0.74	0.15
Daily	IMERG	0.50	2.15	0.08	1.51	0.45	0.53	0.30
	3B42V7	0.38	2.19	-0.48	-8.97	0.36	0.63	0.22
<i>Reg6</i>								
Hourly	IMERG	0.16	0.94	0.05	5.66	0.26	0.56	0.19
3-hourly	IMERG	0.37	1.73	0.15	5.91	0.43	0.47	0.31
	3B42V7	0.34	1.73	0.50	20.24	0.45	0.54	0.30
Daily	IMERG	0.62	4.97	0.11	0.52	0.59	0.31	0.47
	3B42V7	0.62	4.79	3.10	15.31	0.63	0.39	0.45

The main conclusions are given as follows:

- (1) The IMERG hourly precipitation product shows slight overestimation over Mainland China, with small BIAS for grid-scale (9.3%) and national-scale (8.1%) evaluations, respectively. The performance of IMERG is not very satisfactory in the grid-scale hourly evaluation with small CC (0.40) and CSI (0.35) respectively, but could still be acceptable given such relatively fine temporal resolution. The quality of IMERG is improved significantly when it comes to the national evaluation, with CC and CSI increasing to 0.90 and 0.72, respectively, indicating the higher capability of the IMERG product over larger areas.
- (2) Compared with 3B42V7, IMERG shows better performance at 3-hourly and daily resolutions in grid-scale and national-scale evaluations. The differences between IMERG and 3B42V7 products are most notable for 3-hourly comparison at the gridded scale, with CC being 0.52 for IMERG and 0.39 for 3B42V7. Metrics are improved as the temporal resolution increases from 3-hourly to daily, but this effect is less remarkable as the comparison changes from the gridded scale to the national scale.
- (3) For the regional comparison, both IMERG and 3B42V7 products show better performance over sub-regions with mid- and low-latitudes, as well as relatively mild and wet climates. IMERG outperforms 3B42V7 in every sub-region in regard to almost all the metrics. However, both products should improve the quality of precipitation estimates in high latitudes (Reg1) and altitudes (Reg6), and dry areas (Reg5), particularly for the dry area with relatively low precipitation intensity.
- (4) IMERG can generally reproduce the PDF shape of various precipitation intensities at both 3-hourly and daily scales, and shows better performance than 3B42V7. The differences between values of PDF for IMERG products and gauge observations are generally within 1%, but 3B42V7 shows larger deviations for most precipitation intensities, especially the underestimation for the 3-hourly heavy precipitation.

(5) Regarding the reconstruction of the precipitation diurnal cycle, IMERG performs slightly better than 3B42V7 in most sub-regions. However, both products show lags of diurnal peaks of intensity over North and East China (Reg1, Reg2, and Reg3). The capability of IMERG and 3B42V7 in reproducing the diurnal cycle should be improved in the future, particularly for regions with dry climates (Reg5) and high latitudes (Reg1, Reg2, and Reg5).

The quality of IMERG needs to be improved further to provide more accurate hourly and half-hourly precipitation estimates, particularly for regions with high latitudes and altitudes, as well as dry climates. Future studies are advocated to explore the characteristics of error and uncertainty in IMERG products further, to prompt the development and application of IMERG algorithms and products in the GPM era.

## Acknowledgements

This study was supported by the National Natural Science Foundation of China Major Research Program (Project No. 91437214) and the National Natural Science Foundation of China (Project No. 51579128). The authors' great gratitude is extended to the China Meteorological Administration for providing ground-based rainfall data. The IMERG data were provided by the NASA/Goddard Space Flight Center's Mesoscale Atmospheric Processes Laboratory and PPS, which develop and compute IMERG as a contribution to GPM, and archived at the NASA PPS.

## References

- Ahrens, B., 2006. Distance in spatial interpolation of daily rain gauge data. *Hydrol. Earth Syst. Sci.* 10, 197–208.
- Bitew, M.M., Gebremichael, M., Gebremichael, L.T., Bayissa, Y.A., 2012. Evaluation of high-resolution satellite rainfall products through streamflow simulation in a hydrological modeling of a small mountainous watershed in Ethiopia. *J. Hydrometeorol.* 13, 338–350.
- Dai, A., Lin, X., Hsu, K.-L., 2007. The frequency, intensity, and diurnal cycle of precipitation in surface and satellite observations over low- and mid-latitudes. *Clim. Dyn.* 29, 727–744.
- Dinku, T., Anagnostou, E.N., Borga, M., 2002. Improving radar-based estimation of rainfall over complex terrain. *J. Appl. Meteorol.* 41, 1163.
- Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S.J., Ropelewski, C.F., 2007. Validation of satellite rainfall products over East Africa's complex topography. *Int. J. Remote Sens.* 28, 1503–1526.
- Doviak, R.J., Zrnic, D.S., 2006. Doppler Radar & Weather Observations. Academic Press Inc.
- Ebert, E.E., Janowiak, J.E., Kidd, C., 2007. Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Am. Meteorol. Soc.* 88, 47–64.
- Gerapetritis, H., Pelissier, J.M., 2004. The critical success index and warning strategy. In: Preprints, 17th Conference on Probability and Statistics in Atmospheric Sciences, Seattle, WA. American Meteorological Society, 2.10 <[http://ams.confex.com/ams/84Annual/techprogram/paper\\_70691.htm](http://ams.confex.com/ams/84Annual/techprogram/paper_70691.htm)>.
- Germann, U., Galli, G., Bosacchi, M., Bolliger, M., 2006. Radar precipitation measurement in a mountainous region. *Quart. J. R. Meteorol. Soc.* 132, 1669–1692.
- Guo, Z., Liu, X., Xiao, W., Wang, J., Meng, C., 2007. Regionalization and integrated assessment of climate resource in China based on GIS (in Chinese). *Resour. Sci.* 29, 2–9.
- Hong, Y., Adler, R.F., Negri, A., Huffman, G.J., 2007a. Flood and landslide applications of near real-time satellite rainfall products. *Nat. Hazards* 43, 285–294.
- Hong, Y., Gochis, D., Cheng, J.T., Hsu, K.L., Sorooshian, S., 2007b. Evaluation of PERSIANN-CCS rainfall measurement using the NAME event rain gauge network. *J. Hydrometeorol.* 8 (3), 469–482.
- Hong, Y., Adler, R.F., Hossain, F., Curtis, S., Huffman, G.J., 2007c. A first approach to global runoff simulation using satellite rainfall estimation. *Water Resour. Res.* 43,
- Hong, Y., Chen, S., Xue, X., Hodges, G., 2012. Global Precipitation Estimation and Applications. Multiscale Hydrologic Remote Sensing: Perspectives and Applications. CRC Press, 371–386.
- Hou, A.Y. et al., 2014. The global precipitation measurement mission. *Bull. Am. Meteorol. Soc.* 95, 701–722.
- Huffman, G.J., Adler, R.F., Bolvin, D.T., Nelkin, E.J., 2010. The TRMM Multi-Satellite Precipitation Analysis (TMPA). Satellite Rainfall Applications for Surface Hydrology. Springer, pp. 3–22.
- Huffman, G.J. et al., 2007. The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* 8, 38–55.
- Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R., Xie, P., 2014. GPM Integrated Multi-Satellite Retrievals for GPM (IMERG) Algorithm Theoretical Basis Document (ATBD) Version 4.4. PPS, NASA/GSFC, 30 pp. <[http://pmm.nasa.gov/sites/default/files/document\\_files/IMERG\\_ATBD\\_V4.4.pdf](http://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V4.4.pdf)>.
- Huffman, G.J., Bolvin, D.T., Nelkin, E.J., 2015. Integrated Multi-satellite Retrievals for GPM (IMERG) Technical Documentation. NASA/GSFC Code 612, 47 pp. <[http://pmm.nasa.gov/sites/default/files/document\\_files/IMERG\\_doc.pdf](http://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc.pdf)>.
- Huffman, G.J., Bolvin, D.T., 2015. TRMM and Other Data Precipitation Data Set Documentation. Mesoscale Atmospheric Processes Laboratory, NASA Global Change Master Directory Doc., 44 pp. <[http://pmm.nasa.gov/sites/default/files/document\\_files/3B42\\_3B43\\_doc\\_V7.pdf](http://pmm.nasa.gov/sites/default/files/document_files/3B42_3B43_doc_V7.pdf)>.
- Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.* 5, 487–503.
- Kidd, C., Huffman, G., 2011. Global precipitation measurement. *Meteorol. Appl.* 18, 334–353.
- Kirstetter, P.-E., Hong, Y., Gourley, J.J., Schwaller, M., Petersen, W., Zhang, J., 2013. Comparison of TRMM 2A25 Products, Version 6 and Version 7, with NOAA/NSSL ground radar-based national mosaic QPE. *J. Hydrometeorol.* 14, 661–669.
- Kubota, T. et al., 2007. Global precipitation map using satellite-borne microwave radiometers by the GSMAp Project: production and validation. *IEEE Trans. Geosci. Remote Sens.* 45, 2259–2275.
- Li, Z., Yang, D., Hong, Y., 2013. Multi-scale evaluation of high-resolution multi-sensor blended global precipitation products over the Yangtze River. *J. Hydrol.* 500, 157–169.
- Li, Z., Yang, D., Gao, B., Jiao, Y., Hong, Y., Xu, T., 2015. Multiscale hydrologic applications of the latest satellite precipitation products in the Yangtze River Basin using a distributed hydrologic model. *J. Hydrometeorol.* 16, 407–426.
- Long, D. et al., 2014. Drought and flood monitoring for a large karst plateau in Southwest China using extended GRACE data. *Remote Sens. Environ.* 155, 145–160.
- Long, D., Longuevergne, L., Scanlon, B.R., 2015a. Global analysis of approaches for deriving total water storage changes from GRACE satellites. *Water Resour. Res.* 51 (4), 2574–2594.
- Long, D., Yang, Y.T., Wada, Y., Hong, Y., Liang, W., Chen, Y.N., Wei, J.F., Chen, L., 2015b. Deriving scaling factor using a global hydrological model to restore GRACE signals for drought monitoring over China's Yangtze River basin, *Remote Sens. Environ.* 168, 177–193.
- Ma, Y., Zhang, Y., Yang, D., Farhan, S.B., 2015. Precipitation bias variability versus various gauges under different climatic conditions over the Third Pole Environment (TPE) region. *Int. J. Climatol.* 35 (7), 1201–1211.
- Meng, L., Long, D., Quiring, S.M., Shen, Y.J., 2014. Statistical analysis of the relationship between spring soil moisture and summer precipitation in East China. *Int. J. Climatol.* 34, 1511–1523.
- Prat, O.P., Nelson, B.R., 2013. Precipitation contribution of tropical cyclones in the Southeastern United States from 1998 to 2009 using TRMM satellite data. *J. Climate* 26, 1047–1062.
- Sandwell, D.T., 1987. Biharmonic spline interpolation of GEOS-3 and SEASAT altimeter data. *Geophys. Res. Lett.*
- Scheel, M.L.M., Rohrer, M., Huggel, C., Santos Villar, D., Silvestre, E., Huffman, G.J., 2011. Evaluation of TRMM Multi-satellite precipitation analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution. *Hydrol. Earth Syst. Sci.* 15, 2649–2663.
- Shen, Y., Xiong, A., 2015. Validation and comparison of a new gauge-based precipitation analysis over mainland China. *Int. J. Climatol.* <http://dx.doi.org/10.1002/joc.4341> (Advance online publication).
- Shen, Y., Xiong, A., Wang, Y., Xie, P., 2010. Performance of high resolution satellite precipitation products over China. *J. Geophys. Res.: Atmos.* 115.
- Shen, Y., Pan, Y., Yu, J., Zhao, P., Zhou, Z., 2013. Quality assessment of hourly merged precipitation product over China. *Trans. Atmos. Sci.* 36 (1), 37–46 (in Chinese).
- Shen, Y., Xiong, A., Hong, Y., Yu, J., Pan, Y., Chen, Z., Saharia, M., 2014. Uncertainty analysis of five satellite-based precipitation products and evaluation of three optimally merged multi-algorithm products over the Tibetan Plateau. *Int. J. Remote Sens.* 35, 6843–6858.
- Sorooshian, S., Hsu, K.-L., Gao, X., Gupta, H.V., Imam, B., Braithwaite, D., 2000. Evaluation of PERSIANN System Satellite-Based Estimates of Tropical Rainfall. *Bull. Am. Meteorol. Soc.* 81, 2035–2046.
- Su, F., Hong, Y., Lettenmaier, D.P., 2008. Evaluation of TRMM multisatellite precipitation analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin. *J. Hydrometeorol.* 9, 622–640.
- Tao, S.-Y., 1987. A review of recent research on the East Asian summer monsoon in China. *Monsoon Meteorol.*
- Tong, K., Su, F., Yang, D., Zhang, L., Hao, Z., 2014. Tibetan Plateau precipitation as depicted by gauge observations, reanalyses and satellite retrievals. *Int. J. Climatol.* 34 (2), 265–285.
- Villarini, G., Krajewski, W.F., 2008. Empirically-based modeling of spatial sampling uncertainties associated with rainfall measurements by rain gauges. *Adv. Water Resour.* 31, 1015–1023.
- Wilks, D.S., 2006. Statistical Methods in the Atmospheric Sciences. Elsevier, New York.
- Wolff, D.B., Fisher, B.L., 2009. Assessing the relative performance of microwave-based satellite rain-rate retrievals using TRMM ground validation data. *J. Appl. Meteorol. Climatol.* 48, 1069–1099.

- Xue, X. et al., 2013. Statistical and hydrological evaluation of TRMM-based multisatellite precipitation analysis over the Wangchu Basin of Bhutan: are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins? *J. Hydrol.* 499, 91–99.
- Yong, B., Liu, D., Gourley, J.J., Tian, Y., Huffman, G.J., Ren, L., Hong, Y., 2015. Global view of real-time TRMM multisatellite precipitation analysis: implications for its successor global precipitation measurement mission. *Bull. Am. Meteorol. Soc.* 96, 283–296.
- Yong, B. et al., 2010. Hydrologic evaluation of multisatellite precipitation analysis standard precipitation products in basins beyond its inclined latitude band: a case study in Laohahe basin, China. *Water Resour. Res.* 46, W07542.
- Yong, B. et al., 2012. Assessment of evolving TRMM-based multisatellite real-time precipitation estimation methods and their impacts on hydrologic prediction in a high latitude basin. *J. Geophys. Res.: Atmos.* (1984–2012) 117 (D9).
- Yong, B. et al., 2014. Intercomparison of the Version-6 and Version-7 TMPA precipitation products over high and low latitudes basins with independent gauge networks: is the newer version better in both real-time and post-real-time analysis for water resources and hydrologic extremes? *J. Hydrol.* 508, 77–87.
- Yu, R., Zhou, T., Xiong, A., Zhu, Y., Li, J., 2007. Diurnal variations of summer precipitation over contiguous China. *Geophys. Res. Lett.* 34.
- Yuan, W., Yu, R., Zhang, M., Lin, W., Chen, H., Li, J., 2012. Regimes of diurnal variation of summer rainfall over subtropical East Asia. *J. Climate* 25, 3307–3320.
- Zhou, T., Yu, R., Chen, H., Dai, A., Pan, Y., 2008. Summer precipitation frequency, intensity, and diurnal cycle over China: a comparison of satellite data with rain gauge observations. *J. Climate* 21, 3997–4010.