Statistical and hydrological evaluation of the latest Integrated Multi-satelliteE Retrievals for GPM (IMERG) over a midlatitude humid basin in South China

Shanhui Jiang, Liliang Ren, Chong-Yu Xu, Bin Yong, Fei Yuan, Yi Liu, Xiaoli Yang, Xinmin Zeng

ARTICLE INFO

Keywords:
Satellite precipitation product
GPM IMERG
Statistical evaluation
Hydrological utility
Xinanjiang model

ABSTRACT

Recently, the Integrated Multi-satelliteE Retrievals for Global Precipitation Measurement (IMERG) products with high spatial (0.1° × 0.1°) and temporal (half-hourly) resolution have become operationally available. It is of crucial importance to comprehensively evaluate these new products before they are applied extensively. This study focuses on statistical and hydrological evaluations of the latest IMERG (Version 05) products: the near-real-time “Early” run and “Late” run IMERG products (IMERG-E and IMERG-L, respectively), and the post-real-time “Final” run IMERG product (IMERG-F) over the mid-latitude humid Mishui basin in South China during 2014–2015, in comparison with their predecessors, the Tropical Rainfall Measuring Mission Multi-satellite Precipitation Analysis (TMPA) products (3B42RT and 3B42V7). The post-real-time IMERG-F presents the best performance among the five satellite precipitation products (SPPs), with the highest daily correlation coefficient (CC) of 0.85 and the lowest daily root-mean square error of 5.58 mm at the basin scale. The near-real-time IMERG-E and IMERG-L demonstrate comparable performance with 3B42RT. In addition, the Taylor diagrams visually demonstrate that the IMERG products are better than the 3B42 products. For hydrological simulations under scenario I (model calibration based on rain gauge observations), the post-real-time IMERG-F performs obviously better than the 3B42V7 does, with a relatively high CC of 0.81, a good Nash-Sutcliffe coefficient of 0.63, and an acceptable relative bias of –3.98%. Both the IMERG-E and IMERG-L demonstrate a better performance than the 3B42RT does. For scenario II (model recalibration based on each satellite dataset), the hydrological performances of both the IMERG and 3B42 products are improved. The IMERG-E, IMERG-L, demonstrate streamflow simulation performance comparable to that of the 3B42V7, both for the whole simulation period and flood season, indicating the great potential of the latest near-real-time IMERG-E product for flood simulation and prediction. Overall, this systematic evaluation highlights that the latest IMERG products have desirable hydrological utility in the study region. This study will provide useful guidelines for hydrological applications of the new generation SPPs, as well as IMERG algorithm development.

1. Introduction

Precipitation is a fundamental component of the global water cycle, and it also represents crucial forcing data for various applications such as hydrology, water resources, weather, and climate (Ebert et al., 2007; Kidd and Huffman, 2011; Skofronick-Jackson et al., 2017). Due to the significant spatial and temporal variability of precipitation, accurate and rapid acquisition of surface precipitation information has always been a difficult problem, especially for complex terrain regions (Hong et al., 2007; Wu et al., 2014; Yong et al., 2015). The emergence of multi-satellite inversion technology provides a new method for the rapid capture of continuous precipitation information and is also the most effective means for the global scale precipitation observation (Huffman et al., 2007; Kucera et al., 2013; Sun et al., 2018). Currently, satellite inversion precipitation can be divided into three important eras, i.e., the Global Precipitation Climatology Project era (GPCP) (Huffman et al., 1997), the Tropical Rainfall Measuring Mission era (TRMM) (Huffman et al., 2007), and the Global Precipitation Measurement era (GPM) (Hou et al., 2014). These satellite precipitation products (SPPs) have provided quasi-global, high-temporal (higher than
implemented a real-time global temporal and spatial scales. Wu et al. (2014) developed and conducted hydrological simulation results. Duan et al. (2016) evaluated eight high-resolution precipitation products (i.e., PERSIANN-CDR and NCEP-CFSR), with optimal statistical parameters and the best hydrological simulation utility. Therefore, this study focuses mainly on two aspects. First, it aims to statistically evaluate the accuracy and performance of the new generation of satellite precipitation products at a spatial resolution of 0.1° × 0.1° and a temporal resolution of a half hour (Hou et al., 2014; Huffman et al., 2014). The GPM satellite constellation consists of one core observational satellite and approximately ten partner satellites. The GPM core observational satellite was launched on 27 February 2014, with the latest Dual-Frequency Precipitation Radar (DPR), the Ku-band at 13.6 GHz and Ka-band at 35.5 GHz and multichannel GPM Microwave Imager (GMI, frequency ranging from 10 to 183 GHz) (Huffman et al., 2014). Compared with TRMM, there are three critical improvements in GPM: 1) the orbital inclination has been increased from 35° to 65°, expanding the coverage area; 2) the radar has been upgraded to two frequencies, adding sensitivity to light precipitation; and 3) “high-frequency” channels have been added to the passive microwave (PMW) imager, facilitating the sensing of light and solid precipitation (Hou et al., 2014; Huffman et al., 2014). The enhanced measurement and sampling capabilities allow GPM to provide comparatively larger coverage, higher spatial-temporal resolution and more accurate and reliable rainfall and snowfall estimations. The GPM products are divided into four levels based on different algorithms, in which the Level 3 product is the Integrated Multi-satellite Retrievals for GPM (IMERG) estimate. IMERG provides three runs to accommodate different user requirements for latency and accuracy, including the near-real-time “Early” run (hereafter called IMERG-E) and “Late” run (hereafter called IMERG-L) products, and the post-real-time “Final” run (hereafter called IMERG-F) product (Huffman et al., 2014). IMERG products were firstly provided as Version 03 (V03) from March 2014, after that the newer IMERG Version 04 (V04) and Version 05 (V05) products were successively released on March and November 2017 (Huffman et al., 2018). The changes from V03 to V04 and from V04 to V05 are multifactorial, including the algorithm upgraded, new sources data absorbed, new fields and components added and so on (Huffman et al., 2017a and Huffman et al., 2017b).

There have been many initial statistical performance evaluations of IMERG data with rain gauge observations or other prevailing SPPs in different climatic zones and on different spatial-temporal scales (Liu, 2016; Tang et al., 2016a; Tan and Duan, 2017; Sungmin et al., 2017; Xu et al., 2017; Prakash et al., 2018; Tan and Santo, 2018; Wei et al., 2018; Anjum et al., 2018; Sharifi et al., 2018; Gebregiorgis et al., 2018; Chiaravalloti et al., 2018; Wang et al., 2018b). For instance, Liu (2016) compared V03 IMERG-F with TMPA 3B43 on a global scale and showed that IMERG-F provides better performance over high precipitation land regions; furthermore, systematic differences over land are much smaller than those over oceans. Tang et al. (2016a) evaluated the V03 IMERG-F and TMPA products over mainland China at multiple spatiotemporal scales and found that the IMERG-F product outperforms TMPA products in most cases, especially at the mid- and high-latitudes. Anjum et al. (2018) evaluated the V04 IMERG-F with ground observations and TMPA products over the northern highlands of Pakistan and demonstrated that the performance of IMERG is better than that of TMPA products. Gebregiorgis et al. (2018) compared the V03 IMERG-L product with TMPA 3B42RT product across the conterminous United States and confirmed the advances of the new generation of satellite precipitation relative to that of its predecessor. Wang et al. (2018b) globally intercompared and regionally evaluated different version IMERG products and conducted that both V04 and V05 IMERG-F show significant differences and improvements from V03 at global level and V05 IMERG-F generally improves upon both V04 and V03 over Mainland China. Compared with the statistical evaluations, there are only a few investigations performed on the hydrological simulation utilities of IMERG products at the basin scale (Tang et al., 2016b; Li et al., 2017; Zubieta et al., 2017; Yuan et al., 2017; Wang et al., 2017; He et al., 2017; Yuan et al., 2018; Tan et al., 2018). Wang et al. (2017) evaluated the hydrological utility of V03 IMERG products over the Beijiang River Basin in China and summarized that the IMERG-F has high accuracy and good hydrological utility, while the IMERG-E and IMERG-L products have satisfactory hydrological utility during the flood season. Zubieta et al. (2017) analyzed the hydrological simulation utility of V03 IMERG over the Peruvian–Ecuadorian Amazon Basin and indicated that IMERG-F is as useful as TMPA V7 in southern regions, though they do not properly simulate streamflows in northern regions. Yuan et al. (2018) evaluated the hydrological utility of V05 IMERG-F over the Yellow River source region in China and highlighted that the IMERG-F demonstrates comparable streamflow simulation performance with the gauge-based result at daily and sub-daily time scales. In general, IMERG products present preliminary satisfactory accuracy and hydrological simulation utility in most cases, while different types of IMERG products might have distinct hydrological utility in different regions.

Studies focusing on a comprehensive comparison of the “Early”, “Late” and “Final” runs of the product are rare. The statistical precision characteristics and hydrological simulation utility of the latest V05 IMERG-E, IMERG-L and IMERG-F products, as well as their advantages compared those of with TMPA products over different river basins, still require further exploration. Therefore, this study focuses mainly on two aspects. First, it aims to statistically evaluate the accuracy and performance of the latest V05 IMERG-E, IMERG-L and IMERG-F products, compared with those of 3B42RT and 3B42V7, by using dense rain gauge network observations; then, it assesses the hydrological simulation utility of the five SPPs under two simulation scenarios (i.e., gauge-based calibration and each SSP-based recalibration, respectively) in Mishui Basin, a mid-latitude humid basin in South China. This study will provide useful guidelines for hydrological applications of the latest IMERG products as well as algorithm development. The remainder of this paper is organized as follows: Section 2 introduces the study area and the datasets used. Section 3 describes the detailed methodology. Section 4 discusses the simulation results of different simulation scenarios. Finally, Section 5 draws the conclusions.

S. Jiang et al.


419
2. Study area and data

2.1. Study area

The Mishui basin, a tributary of the Xiangjiang River, is located southeast of Hunan Province in South China and extends from longitudes 112.85°E to 114.20°E and latitudes 26.00°N to 27.20°N (Fig. 1). The drainage area above the Ganxi hydrologic station is 9972 km². The elevation within the basin ranges from 49 m to 2093 m above sea level, and the topography significantly descends from southeast to northwest. The land use in the basin is composed of forest and shrubs (54.4%), grassland (33.5%), cropland (11.8%), and urban and water (0.3%). The basin has a typical humid subtropical monsoon type, with long-term average annual temperature and precipitation of 18°C and 1561 mm, respectively. Influenced by atmospheric circulation, the distribution of precipitation is characterized by strong temporal and spatial variability. The rainy season spans from May to July, with generally more than 50% of the annual rainfall. The basin is a high incidence zone of flash flood disasters, with a basin flood disaster recorded every three years. It is meaningful to explore the potential of multi SPPs in streamflow and flood simulation. Jiang et al. (2012) evaluated the hydrological utility of TMPA 3B42V6, 3B42RT, and CMORPH over Mishui basin and concluded that the 3B42V6 performs the best streamflow simulation among the three products. Therefore, this quantitative streamflow and flood simulation utility evaluation of the latest IMERG products is an important continuation of research.

2.2. Data

2.2.1. Satellite precipitation products

The SPPs used in this study include two post-real-time research products, i.e., IMERG-F and 3B42V7, and three near-real-time products, i.e., IMERG-E, IMERG-L and 3B42RT. The TMPA method was designed to combine precipitation estimates from different satellite systems, as well as rain gauges and was intended to provide the best satellite precipitation estimates in the TRMM-era (Huffman et al., 2007). The method retrieves real-time precipitation through three consecutive stages: 1) polar-orbiting MW precipitation estimates are calibrated by TRMM MW estimates and then combined together; 2) geostationary IR precipitation estimates are calibrated using the merged MW precipitation to fill in gaps of the MW coverage; and 3) MW and IR data are combined to form the real-time pure satellite precipitation product (i.e., 3B42RT). Additionally, the 3B42V7 estimate adjusts its bias based on monthly rain gauge observations (Huffman et al., 2007). The IMERG inter-calibrates, merges, and interpolates “all” satellite MW precipitation estimates, together with microwave-calibrated IR satellite estimates, precipitation gauge analyses, and potentially other precipitation estimators at fine temporal and spatial scales over the entire globe (Huffman et al., 2018). As noted above, the algorithm for the IMERG was recently upgraded to Version 05 on 1 December 2017. From the early validation results over the conterminous US, it appears that Version 05 IMERG-F is generally an improvement over Version 04 (Huffman et al., 2018). In this study, the latest near-real-time IMERG-E and IMERG-L daily products and post-real-time IMERG-F daily product from 1 April 2014 to 31 December 2015 were used. The IMERG products were aggregated to 0.25° × 0.25° datasets for grid-scale evaluation by using the standard bilinear interpolation method (Wang et al., 2017).

2.2.2. Gauged precipitation and discharge data

The observed daily precipitation data for 2008 to 2015 were derived from 35 rain gauge stations in the Mishui basin, using roughly two rain gauges within one 0.25° grid. For the same period, daily streamflow and potential evapotranspiration data were collected from the Ganxi hydrologic station and Wulipai evaporation station, respectively. The inverse distance weighting of the three nearest rain gauges was used to obtain the spatially distributed precipitation database of the Mishui basin (Bartier and Keller, 1996). The 30 arc-second global digital elevation model data were obtained from the U.S. Geological Survey, whereas the vegetation-type data were obtained from the International Geosphere-Biosphere Program.

3. Methodology

3.1. XAJ model

The Xinanjiang model is a well-known, conceptual hydrological model developed by Zhao in the 1970s (Zhao, 1992). Since its
development, the Xinjiang model has been successfully and widely used in the humid and semi-humid regions of China (Zhao, 1992; Ren et al., 2008; Jiang et al., 2018). In this study, a gridded-structured Xinjiang model (hereafter called XAJ) for streamflow simulation was constructed. The simulation was performed by computing the runoff and dividing the runoff types within each grid. The slope and river network convergence processes were then integrated to obtain the streamflow series of the hydrologic station. The model was operated daily, with a 0.25° × 0.25° spatial resolution from January 2008 to December 2015. The model has 16 parameters, and they were automatically calibrated using the SCE-UA algorithm, which is an effective and efficient global optimization algorithm (Duan et al., 1994). The calibration was processed automatically with the objective function of maximizing the Nash–Sutcliffe efficiency coefficient (NSE) value, and the model parameters were selected in the experimental numerical range (Jiang et al., 2012 and Jiang et al., 2018).

3.2. Evaluation statistics

The validation statistical indices of the correlation coefficient (CC), mean error (ME), root-mean square error (RMSE), and relative bias (BIAS) were employed to qualitatively evaluate the latest IMERG and TMPA products with rain gauge observations. CC shows the agreement between the satellite precipitation and rain gauge observations. ME demonstrates the average difference between the satellite precipitation and rain gauge observations. BIAS describes the systematic bias of the satellite precipitation. The standard deviation (SD) and centered root-mean-square error (CRMSE) were also calculated for the Taylor diagram, which can visually show and compare the “closeness” between the SPPs and the rain gauge observations (Taylor, 2001). In addition, two categorical statistical indices, including the probability of detection (POD), and false-alarm rate (FAR) were adopted to measure the correspondence between the SPPs and rain gauge observations. POD, also known as the hit rate, represents how often the rain occurrences are correctly detected by the satellite. FAR denotes the fraction of cases in which the satellite records precipitation when the rain gauges do not. For the evaluation of the hydrological simulation utility of the SPPs, the CC, NSE, and BIAS were employed. These three indices jointly measure the consistency of the simulated and observed streamflow series, both in terms of temporal distribution and amount. The formulas for the indices mentioned above are listed in Table 1.

4. Results and discussion

4.1. Statistical evaluation and comparison of IMERG and 3B42 products

The IMERG and 3B42 products were evaluated against the dense rain gauge observations at both grid scale and basin scale, during the period from 1 April 2014 to 31 December 2015. For the grid scale evaluation, the IMERG and 3B42 products were compared against the interpolated rain gauge observations at 17 selected 0.25° × 0.25° grid boxes, with at least one rain gauge in each grid box (Fig. 1). For the basin scale evaluation, the interpolated rain gauge data, and both the IMERG and the 3B42 products were spatially accumulated into the basin average precipitation. Daily CC, ME, RMSE, BIAS, POD and FAR were calculated for the grid-based evaluation and basin-wide evaluation. Table 2 summarizes the daily statistical indices of the 17 selected grid boxes. For the grid scale evaluation, the IMERG-F presents the best performance among the five SPPs (with the highest daily CC of 0.77, and the lowest daily RMSE of 7.65 mm). By comparison, the near-real-time IMERG-E and IMERG-L demonstrate comparable performance with that of 3B42V7, in terms of slightly better CC and RMSE values. The near-real-time 3B42RT shows the worst performance among the five SPPs (with the lowest daily CC of 0.66, and the largest daily RMSE of 9.94 mm). In regard to the systematic error, all the five SPPs systematically underestimated the precipitation, in which the near-real-time IMERG-E and IMERG-L products have poor performance, with BIAS values of −8.55% and −10.74%, respectively, while the near-real-time 3B42RT has good performance with a BIAS of −0.83%. It should be noted that the BIAS of 3B43RT is very small, and the RMSE of 3B42RT is relatively large, which can be attributed to the fact that the positive/negative biases are offset for 3B42RT (Fig. 2). For the two post-real-time products, the IMERG-F presents lower ME (−0.29 mm) and BIAS of −5.36%. From the visual representation of CC, and RMSE in Fig. 2, it is noted that the near-real-time and post-real-time IMERG products exhibit relatively better performance than that of the near-real-time and post-real-time 3B42 products.

For the basin scale evaluation, the spatial scale enlarges, the performance of both the IMERG products and 3B42 products significantly improves compared to the grid scale evaluation. The CC values increase from 0.66–0.77 for the grid scale to 0.77–0.85 for the basin scale, while the RMSE values decrease from 7.65–9.94 mm for the grid scale to 5.58–6.91 mm for the basin scale. This finding is consistent with some previous research in different regions, which also confirmed the positive effect of area scale on satellite precipitation error (Yong et al., 2010; Wang et al., 2017; Tan and Santo, 2018). Similar to the former grid scale evaluation, the IMERG-F presents the best performance among the five SPPs (with the highest daily CC of 0.85 and the lowest daily RMSE of 5.58 mm). In general, it is noted that the near-real-time and post-real-time IMERG products exhibit significantly better performance than those of the near-real-time and post-real-time 3B42 products.

Table 1 Statistical evaluation indices for evaluating the SPPs and their hydrological utility.

<table>
<thead>
<tr>
<th>Evaluation Indices</th>
<th>Unit</th>
<th>Formulas</th>
<th>Perfect Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation coefficient (CC)</td>
<td></td>
<td>[ CC = \frac{\sum_{i=1}^{n} (Q_{\text{sim}} - \bar{Q}<em>{\text{obs}}) (Q</em>{\text{obs}} - \bar{Q}<em>{\text{obs}})}{\sqrt{\sum</em>{i=1}^{n} (Q_{\text{sim}} - \bar{Q}<em>{\text{sim}})^2} \sqrt{\sum</em>{i=1}^{n} (Q_{\text{obs}} - \bar{Q}_{\text{obs}})^2}} ]</td>
<td>1</td>
</tr>
<tr>
<td>Mean error (ME)</td>
<td>mm</td>
<td>ME = \frac{\sum_{i=1}^{n} (Q_{\text{sim}} - Q_{\text{obs}})}{n}</td>
<td>0</td>
</tr>
<tr>
<td>Root-mean-square error (RMSE)</td>
<td>mm</td>
<td>RMSE = \sqrt{\frac{\sum_{i=1}^{n} (Q_{\text{sim}} - Q_{\text{obs}})^2}{n}}</td>
<td>0</td>
</tr>
<tr>
<td>Relative bias (BIAS)</td>
<td>%</td>
<td>BIAS = \frac{\sum_{i=1}^{n} (Q_{\text{sim}} - Q_{\text{obs}})}{\sum_{i=1}^{n} Q_{\text{obs}}} \times 100</td>
<td>0</td>
</tr>
<tr>
<td>Standard deviation (SD)</td>
<td>mm</td>
<td>SD = \sqrt{\frac{\sum_{i=1}^{n} (Q_{\text{sim}} - Q_{\text{obs}})^2}{n}}</td>
<td>0</td>
</tr>
<tr>
<td>Centered root-mean-square error (CRMSE)</td>
<td>mm</td>
<td>CRMSE = \sqrt{\frac{\sum_{i=1}^{n} (Q_{\text{sim}} - \bar{Q}_{\text{sim}})^2}{n}}</td>
<td>0</td>
</tr>
<tr>
<td>Probability of detection (POD)</td>
<td></td>
<td>POD = \frac{t_{H}}{t_{H} + t_{F}}</td>
<td>1</td>
</tr>
<tr>
<td>False-alarm rate (FAR)</td>
<td></td>
<td>FAR = \frac{t_{F}}{t_{H} + t_{F}}</td>
<td>0</td>
</tr>
<tr>
<td>Nash–Sutcliffe efficiency (NSE)</td>
<td></td>
<td>NSE = 1 - \frac{\sum_{i=1}^{n} (Q_{\text{sim}} - Q_{\text{obs}})^2}{\sum_{i=1}^{n} (Q_{\text{obs}} - \bar{Q}_{\text{obs}})^2}</td>
<td>1</td>
</tr>
</tbody>
</table>

n is the total number of samples; \( S_i \) and \( G_i \) are the ith values of the evaluated data and reference data, respectively; \( S \) and \( G \) are the mean values of \( S_i \) and \( G_i \), respectively. \( H, M, F \) are different cases: \( H \) is observed rain, correctly detected; \( M \) is observed rain, not detected; and \( F \) is rain, detected but not observed. \( t_{H}, t_{M}, t_{F} \) and \( t_{F} \) are the times of occurrence of the corresponding case. The SD formula listed in the table is for the SSPs, which is also fit for the rain gauge observations with the \( S \) replaced by \( G \).
the grid scale and basin scale, respectively), while the FARs of the IMERG products (ranging from 0.18–0.19 and 0.12–0.14 for the grid scale and basin scale, respectively) are slightly higher than those of the 3B42 products (ranging from 0.09–0.10 and 0.09–0.09 for the grid scale and basin scale, respectively). Similar to the findings of Wang et al. (2017) and Tan and Santo (2018), the IMERG products with higher PODs than those of 3B42 products might suggest that the extension of the GPM sensor package effectively improves the ability to detect light precipitation compared to that of TRMM instruments, but this improved sensitivity may sometimes increase the FARs. There are also some different findings than those in Wang et al. (2017); the POD and FAR values of IMERG-F perform slightly better than those of IMERG-E and IMERG-L products, which may be because there is a clear improvement from Version 4 to Version 5 of IMERG in POD (Huffman et al., 2018).

Furthermore, we employed Taylor diagrams to assess the performances of the five SPPs at both grid and basin scale (Fig. 3). As shown in the Taylor diagrams, the post-real-time IMERG-F and 3B42V7 products are closer to the gauge observations than the near-real-time IMERG-E, IMERG-F, and 3B42RT products are, and the aggregated basin scale evaluations performed much better than did the grid-scale results for all five SPPs. Similar to Tang et al. (2016b), Wang et al. (2017) and Anjum et al. (2018), the post-real-time IMERG-F demonstrates lower CRMSE values than those of 3B42V7 both at the grid and basin scales. For the near-real-time products, IMERG-E and IMERG-L demonstrate similar performances, and they both show lower CRMSE values than those of 3B42RT. Overall, the Taylor diagram assessment reveals that the IMERG products are comprehensively better than the 3B42 products are.

The probability density function (PDF) represents the probability of the occurrence of different rain intensities and is often used to evaluate the quality of the SPPs (Jiang et al., 2012; Tang et al., 2016a; Anjum et al., 2018). Fig. 4 shows the PDF of IMERG and 3B42 products from April 2014 to 2015 for the Mishui basin. All of the five SPPs slightly overestimated the occurrence frequency of the precipitation between 0 and 0.2 mm/day, with an overestimation frequency of approximately 3% and 6% for IMERG and 3B42 products respectively, and the IMERG

| Table 2 | Daily and monthly statistical measures of the IMERG and 3B42 products at grid and basin scales. |
|-----------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
|                | QPEs              | Grid Scale        | Basin Scale       |                  |                  | Grid Scale        | Basin Scale       |                  |                  | Grid Scale        | Basin Scale       |                  |                  |
|                |                   | CC               | ME (mm)           | RMSE (mm)        | BIAS (%)         | POD               | FAR               | CC               | ME (mm)           | RMSE (mm)        | BIAS (%)          | POD               | FAR               |
| Daily time     | IMERG-E           | 0.71             | −0.44             | 8.45             | −8.55            | 0.71              | 0.19              | 0.80             | −0.42             | 6.38             | −8.46             | 0.77              | 0.14              |
| series         | IMERG-L           | 0.73             | −0.55             | 8.12             | −10.74           | 0.72              | 0.18              | 0.80             | −0.54             | 6.20             | −10.78            | 0.77              | 0.12              |
|                | 3B42RT            | 0.66             | −0.04             | 9.94             | −0.83            | 0.52              | 0.09              | 0.77             | −0.06             | 6.91             | −1.15             | 0.70              | 0.09              |
|                | IMERG-F           | 0.77             | −0.20             | 7.65             | −3.65            | 0.73              | 0.18              | 0.85             | −0.20             | 5.58             | −3.93             | 0.78              | 0.13              |
|                | 3B42V7            | 0.70             | −0.29             | 8.95             | −5.63            | 0.53              | 0.10              | 0.82             | −0.29             | 5.96             | −5.89             | 0.72              | 0.09              |
| Monthly time    | IMERG-E           | 0.81             | −13.60            | 56.60            | −8.55            | –                 | –                 | 0.86             | −12.81            | 46.68            | −8.46             | –                 | –                 |
| series         | IMERG-L           | 0.79             | −16.90            | 60.08            | −10.74           | –                 | –                 | 0.83             | −16.31            | 51.41            | −10.78            | –                 | –                 |
|                | 3B42RT            | 0.80             | −1.24             | 62.89            | −0.83            | –                 | –                 | 0.85             | −1.74             | 51.32            | −1.15             | –                 | –                 |
|                | IMERG-F           | 0.93             | −6.22             | 36.89            | −3.65            | –                 | –                 | 0.97             | −5.95             | 24.46            | −3.93             | –                 | –                 |
|                | 3B42V7            | 0.90             | −8.88             | 41.80            | −5.63            | –                 | –                 | 0.94             | −8.92             | 31.04            | −5.89             | –                 | –                 |

Fig. 2. Box plots of the statistical accuracy indices of IMERG and 3B42 SPPs on a daily time scale versus the gauge observations for the selected grids. The upper and lower edges of the box mark the upper and lower quartiles (75% and 25%, respectively); the solid line in the box marks the median value; the upper and lower horizontal lines out of the box mark the 90% and 10% quartiles, respectively; and the points mark the maximum and minimum values.
products were closer to the PDF of the gauge than were the 3B42 products. For the light rainfall event between 0.2 and 1 mm/day, the IMERG products show a much better fitting with the gauge observation than that with the 3B42 products, which highlights the improved ability of IMERG over 3B42 in detecting light precipitation. While for the heavy rainfall event more than 50 mm/day, IMERG products overestimated the occurrence frequency, especially for the IMERG-F product, with occurrence frequency of 1.7% (the occurrence frequency for gauge observations is 0.6%), which should be noted by the algorithm developers. In general, the PDF analysis shows that there is a distinct improvement from 3B42 to IMERG in capturing the probability of the occurrence of different rain intensities in the Mishui basin.

To further analyze the monthly and seasonal precisions and variations, the monthly statistic indices of the five SPPs are summarized in Table 2 for grid and basin scales. The monthly basin average precipitation time series and the ME values of the five SPPs are shown in Fig. 5. At the monthly time scale, the IMERG-F has the best fitting with the rain gauge observations with the highest CC values of 0.93 and 0.97 and the lowest RMSE values of 36.89 mm and 24.46 mm for the grid and basin scales, respectively. The near-real-time IMERG-E and IMERG-L products present comparable performance with the 3B42RT estimate, and they significantly improve the overestimation of the 3B42RT for the rainy season in summer. However, the IMERG-E and IMERG-L demonstrate some overestimations of the rainfall for the dry season in winter.

4.2. Hydrological utility evaluation of IMERG and 3B42 products

The streamflow prediction utilities of the IMERG and 3B42 products were evaluated using the XAJ model through two different parameter calibration scenarios. In scenario I, model parameters were calibrated and validated based on the rain gauge observations from January 2008 to December 2013 and January 2014 to December 2015, respectively. Then, the model was run using IMERG and 3B42 products from January 2014 to December 2015 (April 2014 to December 2015 for IMERG products) with rain gauge-calibrated model parameters. In scenario II, model parameters were recalibrated based on each satellite precipitation dataset to explore how the hydrological model can tolerate and adjust the error of the SSP (Jiang et al., 2012; Maggioni and Massari, 2018). Both the calibrations were processed automatically by using the SCE-UA algorithm (Duan et al., 1994; Jiang et al., 2018).

Fig. 6 shows the daily comparisons of the observed streamflow with the simulated hydrograph, based on rain gauge observations for both the calibration and validation periods, respectively. In general, there is good agreement between the simulated daily streamflow and the observed time series both during calibration and validation periods, with a daily CC of 0.93 and 0.91, NSE of 0.86 and 0.82, and BIAS of 6.72% and 1.52% for the two periods, respectively. The evaluation results indicate that the XAJ model captured key features of the observed hydrograph, and thus, it is suitable to evaluate the hydrological utility of the SPPs. Then, the IMERG and 3B42 SPPs were used as forcing data to conduct the streamflow simulation for the period from January 2014.
to December 2015 (April 2014 to December 2015 for IMERG products), respectively. The streamflow simulation results and the statistical measures of each SPP for scenario I are shown in Fig. 7 and are summarized in Table 3. For the whole simulation period, the post-real-time IMERG-F presents the best performance in the streamflow simulation among the five SPPs, with relatively high CC of 0.81, good NSE of 0.63, and acceptable BIAS of −3.98%; the 3B42V7 product takes second place, with a relatively high CC of 0.77, good NSE of 0.55, and acceptable BIAS of −7.89%. For the near-real-time SPPs, the IMERG-E presents the best performance, with a relatively high CC of 0.73, good NSE of 0.50, and high but acceptable BIAS of −11.36%; the IMERG-L product takes second place and the 3B42RT demonstrates the worst performance. For the flood season (May to October), the model presents a similar performance as for the whole simulation period. The IMERG-F shows a desirable performance for the flood simulation, with a CC of 0.83, NSE of 0.66, and BIAS of −6.53%; the 3B42V7 product takes second place. The IMERG-E and IMERG-L exhibit good CC and NSE values, while they have relatively high BIAS values. The 3B42RT reveals the worst performance, with a relatively low NSE of 0.24. Overall, the IMERG products demonstrate comprehensively better streamflow simulation utilities than those of the corresponding 3B42 products.

For further analysis on how the hydrological model can tolerate and adjust the error of the SPPs, we recalibrated the model parameters based on each satellite precipitation dataset for the period from January 2014 to December 2015 (April 2014 to December 2015 for IMERG products). The streamflow simulation results and the statistical measures of each SPP for scenario II are shown in Fig. 8 and are summarized in Table 3. By comparison, the recalibration of the model parameters significantly improved the streamflow simulation performance of both IMERG and 3B42 products, especially in terms of BIAS. From the table and figure, we can highlight that IMERG-F presents a much more desirable hydrological simulation utility with CCs of 0.83 and 0.84, NSE of 0.70 and BIASs of 0.84% and −1.98% for the whole simulation period and flood season, respectively. The IMERG-E demonstrates comparable performance (with CCs of 0.77 and 0.80, NSEs of 0.60 and 0.64, and BIASs of 2.61% and −4.81%) with the 3B42V7, for both the whole simulation period and the flood season. However, the IMERG-L performs slightly worse than that of IMERG-E. In general, similar to findings of Tang et al. (2016a, 2016b), Wang et al. (2017), and He et al. (2017), it is evident that the IMERG products have good hydrological

Fig. 5. Monthly basin averaged precipitation time series of the five SPPs and their ME versus gauge observations during 2014–2015. ME: mean error.

Fig. 6. Observed and simulated streamflows using rain gauge data as input: (a) daily calibration period, and (b) daily validation period.
Fig. 7. Observed and simulated streamflows from IMERG and 3B42 SPPs with gauge calibrated parameters in scenario I: (a) IMERG-E, (b) IMERG-L, (c) 3B42RT, (d) IMERG-F, and (e) 3B42V7.

Table 3
Daily and monthly statistical measures of the hydrological simulation results of the IMERG and 3B42 products, under the two scenarios.

<table>
<thead>
<tr>
<th>Time series</th>
<th>QPEs</th>
<th>Scenario I</th>
<th>Scenario I Flood season</th>
<th>Scenario II</th>
<th>Scenario II Flood season</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CC</td>
<td>NSE</td>
<td>BIAS (%)</td>
<td>CC</td>
</tr>
<tr>
<td>Daily time series</td>
<td>IMERG-E</td>
<td>0.73</td>
<td>0.50</td>
<td>−11.36</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>IMERG-L</td>
<td>0.71</td>
<td>0.46</td>
<td>−14.66</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>3B42RT</td>
<td>0.69</td>
<td>0.32</td>
<td>2.35</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>IMERG-F</td>
<td>0.81</td>
<td>0.63</td>
<td>−3.98</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>3B42V7</td>
<td>0.77</td>
<td>0.55</td>
<td>−7.98</td>
<td>0.83</td>
</tr>
<tr>
<td>Monthly time series</td>
<td>IMERG-E</td>
<td>0.85</td>
<td>0.68</td>
<td>−11.36</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>IMERG-L</td>
<td>0.81</td>
<td>0.59</td>
<td>−14.66</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>3B42RT</td>
<td>0.80</td>
<td>0.56</td>
<td>2.35</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>IMERG-F</td>
<td>0.93</td>
<td>0.86</td>
<td>−3.98</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>3B42V7</td>
<td>0.88</td>
<td>0.75</td>
<td>−7.97</td>
<td>-</td>
</tr>
</tbody>
</table>
utility. The recalibration of the model parameters in scenario II effectively improved the precision of SPP-forced streamflow simulations, but caution should be taken when recalibrating model parameters with SPPs, which may result in unrealistic parameter values (Maggioni and Massari, 2018). It also should be noted that in both scenario I and scenario II, the near-real-time IMERG-E and IMERG-L overestimated the streamflow in the non-flood season of November 2014 and January 2015 (Figs. 7 and 8).

The simulated daily streamflow time series were aggregated to monthly time scales, and the performance of the monthly streamflow simulations of the five SPPs was further evaluated. The monthly streamflow simulation results and the statistical measures of each SPP are shown in Fig. 9 and are summarized in Table 3. Compared to the daily results, the monthly streamflow simulations were more accurate in terms of CC and NSE values, although both simulations had the same magnitudes of BIAS. The IMERG-F presents a desirable streamflow simulation for both simulation scenarios, with CCs of 0.93 and 0.94, NSEs of 0.86 and 0.88, and BIASs of −3.98% and 0.78% for scenarios I and II, respectively. The near-real-time IMERG-E and IMERG-L present a certain utility for monthly streamflow simulation, especially for scenario II, with CCs of 0.88 and 0.86, NSEs of 0.78 and 0.72, and BIASs of −2.61% and 4.64% respectively. However, the near-real-time IMERG-E and IMERG-L both seriously overestimated the streamflow in the non-flood season of November 2014 and January 2015. Generally, the monthly streamflow simulation evaluation demonstrates that the IMERG-F product has great application potential for monthly water resources planning and management.

4.3. Earlier studies using the IMERG products

To compare the hydrological utility of the IMERG products around the world, we summarized some recent hydrological utility evaluation studies on the IMERG and TMPA products in Table 4 (Zubieta et al., 2017; Yuan et al., 2017; Tang et al., 2016a, 2016b; Wang et al., 2017; He et al., 2017; Yuan et al., 2018; Tan et al., 2018). In general, the summary indicates that the IMERG has a good daily streamflow
simulation utility in most of the regions, except for the northern regions of the Peruvian–Ecuadorian Amazon basin. Three main conclusions can be drawn from the comparisons: 1) in many cases, the IMERG-F outperforms TMPA V7 in hydrological simulations due to its improvement in rainfall estimation; 2) the near-real-time products usually show worse hydrological simulation results than those of the post-real-time product, either in IMERG or TMPA products; and 3) the recalibration of the model parameters, based on each SPP, effectively improved the precision of the SPP-forced streamflow simulations. Our study focuses on the hydrological simulation utility of the latest V05 IMERG near-real-time and post-real-time products. Similar findings were reported that the different version IMERG products performed better than the 3B42 products did in streamflow simulations (Tang et al., 2016a, 2016b; Wang et al., 2017; He et al., 2017; Yuan et al., 2018; Tan et al., 2018). Our study highlighted that the V05 IMERG-E demonstrates comparable streamflow simulation performance to that of the 3B42V7, both for the whole simulation period and flood season, which indicates the great potential of the latest near-real-time IMERG-E product in flood simulation and prediction, and the similar findings were verified by Wang et al. (2017) in Beijiang River Basin in China and Tan et al. (2018) in Kelantan River Basin in Malaysia. In addition, Yuan et al. (2018) implemented a case study of flood-event simulations based on V05 IMERG-F product and demonstrated that the MERRF-F product can be adopted as reliable precipitation source for hydrological simulations at 3-hourly time scale. The successful application of V05 IMERG-F product on flood-event simulations strengthened our confidence that new generation IMERG products can be used at flood-event simulation and forecast. While so far there are no flood-event simulation utility research for the near-real-time IMERG products, thus more studies to explore the potential of near-real-time IMERG products in sub-daily or hourly time-scale flood simulations are necessary.

5. Conclusions

This study conducted a comprehensive evaluation of the latest V05 IMERG-E, IMERG-L and IMERG-F products for the period of April 2014 to December 2015 in the Mishui Basin, a midlatitude humid basin in South China. First, the statistical evaluation of the accuracy of the IMERG products was compared with the 3B42RT and 3B42V7 by using dense rain gauge network observations. Then, the hydrological simulation utilities of the IMERG products were evaluated under two simulation scenarios (i.e., gauge-based calibration and each SSP-based recalibration). The main conclusions were drawn as follows:

1. The post-real-time IMERG-F presents the best performance among the five SPPs, with the highest daily CC of 0.85 and the lowest daily RMSE of 5.58 mm at the basin scale. The near-real-time IMERG-E and IMERG-L demonstrate comparable performance with that of 3B42V7 in terms of slightly better CC and RMSE values. For the capability of rainfall event detection, the newer IMERG products present higher PODs than those of the older 3B42 products, but with slightly higher FARS.

2. The Taylor diagrams visually show that the post-real-time IMERG-F and 3B42V7 products are closer to the gauge observations than are the near-real-time IMERG-E, IMERG-F, and 3B42RT products, and the aggregated basin scale evaluations perform much better than do the grid-scale results. In general, the Taylor diagrams clearly demonstrate that the IMERG products are comprehensively better than the 3B42 products are.

3. For the hydrological simulations under scenario I, the post-realtime IMERG-F performs obviously better than the 3B42V7 does, with a relatively high CC of 0.81, good NSE of 0.63, and acceptable BIAS of −3.98%. For the near-real-time SPPs, both the IMERG-E and IMERG-L demonstrate a better performance than that of the 3B42RT, and the IMERG-E, with a CC of 0.73, NSE of 0.50 and BIAS of −11.36%, takes the top spot.

4. As evidenced by the recalibration of the model parameters under scenario II, the performances of both the IMERG and 3B42 products are clearly improved, with significantly increased NSEs and decreased BIAS values. The IMERG-E especially demonstrates a comparable streamflow simulation performance to that of the 3B42V7, both for the whole simulation period and the flood season, indicating the great potential of the latest near-real-time IMERG-E product in flood simulation and prediction.

Overall, both the near-real-time and post-real-time IMERG products exhibit an apparent improvement over their corresponding 3B42 products. The IMERG-F presents the best performance in the streamflow simulation among the five SPPs, and the IMERG-E shows a comparable streamflow simulation performance with that of 3B42V7. With a finer spatiotemporal resolution and improved precision, the latest generation IMERG products demonstrate great potential in hydrological modeling,
disaster monitoring, water resources planning and management. This study will provide useful guidelines for hydrological applications of the latest generation of SPPs, as well as for algorithm development. To take full advantage of the high spatiotemporal resolution of IMERG, more studies to explore the potential of near-real-time and post-real-time IMERG products in sub-daily or hourly time-scale flood simulations are necessary.

Acknowledgments

The current study was jointly supported by the National Key Research and Development Program approved by Ministry of Science and Technology, China (2016YFA0601504); the Programme of Introducing Talents of Discipline to Universities by the Ministry of Education and the State Administration of Foreign Experts Affairs, China (B08048); the National Natural Science Foundation of China (41501017, 51779070), the Natural Science Foundation of Jiangsu Province (BK20150815) and the Fundamental Research Funds for the Central Universities (2017B10214).

References


files/document_files/IMERG_ATBD_V5.2.pdf,
Liu, Z., 2016. Comparison of integrated multisatellite retrievals for GPM (IMERG) and TRMM multisatellite precipitation analysis (TMPA) monthly precipitation products: initial results. J. Hydrometeorol. 17, 777–790.
Tang, G., Zeng, Z., Long, D., Guo, X., Yong, B., Zhang, W., Hong, Y., 2016b. Statistical and hydrological comparisons between TRMM and GPM level-3 products over a mid-latitude basin: is day-1 IMERG a good successor for TRMM 3B42V7? J. Hydrometeorol. 17, 121–137.