Evaluation and Hydrological Utility of the Latest GPM IMERG V5 and GSMaP V7 Precipitation Products over the Tibetan Plateau

Dekai Lu, Bin Yong

1 State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China; sdludekai@hhu.edu.cn
2 School of Earth Sciences and Engineering, Hohai University, Nanjing 211100, China
* Correspondence: yongbin@hhu.edu.cn; Tel.: +86-25-8378-7485

Received: 15 November 2018; Accepted: 9 December 2018; Published: 12 December 2018

Abstract: Satellite precipitation products provide alternative precipitation data in mountain areas. This study aimed to assess the performance of the latest Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG) version 5 (IMERG V5) and Global Satellite Mapping of Precipitation version 7 (GSMaP V7) products and their hydrological utilities over the Tibetan Plateau (TP). Here, two IMERG Final Run products (uncalibrated IMERG (IMERG-UC) and gauge-calibrated IMERG (IMEEG-C)) and two GSMaP products (GSMaP Moving Vector with Kalman Filter (GSMaP-MVK) and gauge-adjusted GSMaP (GSMaP-Gauge)) were evaluated from April 2014 to March 2017. Results show that all four satellite precipitation products could generally capture the spatial patterns of precipitation over the TP. The two gauge-adjusted products were more consistent with the ground measurements than the satellite-only products in terms of statistical assessment. For hydrological simulation, IMERG-UC and GSMaP-MVK showed unsatisfactory performance for hydrological utility, while GSMaP-Gauge demonstrated comparable performance with gauge reference data, suggesting that GSMaP-Gauge can be selected for hydrological application in the TP. Our study also indicates that accurately measuring light rainfall and winter snow is still a challenging task for the current satellite precipitation retrievals.

Keywords: satellite precipitation; Tibetan Plateau; GPM; IMERG; GSMaP

1. Introduction

Precipitation is a key variable of the global water cycle and the driving force for land surface hydrological processes [1–3]. Reliable precipitation input is crucial for hydrological modeling and prediction [4]. Traditionally, rainfall observations are operated through rain gauges and ground-based weather radars. However, measurements from these in situ observational networks are restricted for the remote areas and mountain regions [5]. For instance, the Tibetan Plateau (TP) with an average elevation of over 4000 m, is the highest plateau in the world [6]. To date, precipitation observation networks are sparse or nonexistent in many parts of the TP due to the complex terrain, harsh climate, and high cost [7]. Satellite precipitation products are practical means of detecting rainfall distribution over vast regions that are poorly gauged, and provide a potential complementary source beyond ground in situ data [8,9].

Satellite precipitation products combine the advantage of the frequent sampling of infrared (IR) and more accurate precipitation estimates of passive microwave (PMW) observations and can provide quasi-global precipitation maps [10]. Since the Tropical Rainfall Measuring Mission (TRMM) started in 1997, various satellite precipitation products were made available to the public with different temporal...
and spatial resolutions [11–15]. This greatly facilitates hydrological prediction and water resource management [16–18]. However, errors and uncertainties exist in satellite precipitation products due to the indirectness of distant radiance measurements to surface precipitation, especially in high latitudes and altitudes [3,19,20]. For example, Tian and Peters-Lidard [21] assessed six different TRMM-era datasets over the globe, and found high relative uncertainties over complex terrain areas, including the TP, the Rockies, and the Andes. Yong et al. [22] argued that the highly accurate IR and PMW observations still have difficulties on the TP with snow cover and complex terrain.

Recently, as a successor to the TRMM, the Global Precipitation Measurement (GPM) Core Observatory (GPM CO), carrying advanced precipitation sensors with the Ka/Ku-band (35.5 and 13.6 GHz, respectively) dual-frequency precipitation radar (DPR) and the multifrequency (10–183 GHz) GPM microwave imager (GMI), was launched on 28 February 2014 [23]. The GPM CO and another 10 or so satellite partners together provide the best possible sampling and coverage of precipitation observations [24,25], and can improve the capabilities of measuring light rain and snowfall. With such a constellation of varying precipitation sensors, the Integrated Multi-satellite Retrievals for GPM (IMERG) algorithm was developed to provide a new generation of precipitation estimates around the globe with high temporal (half-hour) and spatial (0.1° × 0.1°) resolutions [23,26,27]. Based on the difference of data latency and accuracy, IMERG offers three distinct products: Early, Late, and Final Runs. These IMERG products take advantage of several existing precipitation retrieval algorithms, including TRMM Multi-satellite Precipitation Analysis (TMPA), Climate Prediction Center Morphing with Kalman filter (CMORPH-KF), and Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks and a Cloud Classification System (PERSIANN-CCS), and are expected to have less error and better performance than TRMM-era satellite precipitation products.

As the GPM era is now upon us, several recent studies evaluated the performance of IMERG data by comparing them against observations from gauges, radars, and other prior satellite precipitation products [10,28–32]. Huffman et al. [33] did a preliminary comparison between the IMERG and TMPA 3B42 over coterminous United States (US) in the months April through July 2014. They reported that IMERG has good agreement with 3B42 data (correlation of 0.88). Furthermore, Tang et al. [29] demonstrated that IMERG has better performance than 3B42V7 in almost every sub-region of mainland China, especially in relatively dry climate regions. Gebregiorgis et al. [32] investigated the degree of improvement of the IMERG by comparing the uncalibrated version 3 Late Run IMERG and version 7 TMPA-RT products. They concluded that IMERG shows substantial improvement with miss-rain and false-rain bias reduction and hit rate. Similar assessments of the IMERG were conducted in different regions and climatic conditions (e.g., Africa [34], the Central Amazon [35], and Iran [36], among many others). In summary, the earlier studies found that the IMERG is a good successor of TMPA, and is promising for hydrological and meteorological applications. Meanwhile, authors also highlighted that the continuous researching of IMERG for complex topography is needed as more IMERG data become available in the future [37–39].

The algorithm for IMERG underwent several upgrades since it was released with initial version 3. Recently, IMERG version 5 products were released in November 2017 and contain data from March 2014 to present. According to Huffman et al. [40], version 5 supersedes all prior IMERG versions, and users are recommended to shift to the new datasets as soon as practical. On the other hand, the Japanese counterpart of IMERG, the Global Satellite Mapping of Precipitation (GSMaP), operated by the Japan Aerospace Exploration Agency (JAXA), distributes global precipitation maps with different products: the near-real-time GSMaP_NRT, the standard GSMaP_MVK, the gauge-adjusted standard GSMaP_Gauge, and others [41]. All GSMaP products are available at 0.1° × 0.1° resolution every hour and can be found at http://sharaku.eorc.jaxa.jp/GSMaP/. The latest GSMaP algorithm is version 7, which was updated on January 2017, and corresponding products are available after March 2014. Thus, it is necessary to evaluate the latest versions of IMERG and GSMaP products and offer some insight for product developers and users.
As the “Asian Water Tower” and “Third Pole” [6,42], the TP plays an important role in water supply of Asia’s main river basins and global climate change. Knowing the spatial and temporal patterns of precipitation is vital for meteorology research and water management of the TP. With the improvement in measuring light rain and falling snow, the GPM-era precipitation products seem to have vast application prospects in the TP area. Therefore, there is an urgent need to validate the newly available version of GPM-era satellite precipitation products over the TP. In this study, the performances of IMERG version 5 (IMERG V5) and GSMaP version 7 (GSMaP V7) products were evaluated in the TP through extensive statistical and hydrological validation. We expect that the results could provide useful information for algorithm developers of the newest precipitation versions, which are scheduled to be reprocessed to TRMM-era datasets in the near future. The remaining parts of this paper are organized as follows: Sections 2 and 3 describe the study area, the satellite and gauge data, and the evaluation method. The results are presented and discussed in Section 4. Finally, Section 5 summarizes major conclusions of this study.

2. Study Area and Data

2.1. Study Area

The TP, which covers an area of about $2.5 \times 10^6$ km$^2$, is located in southwestern China (Figure 1). Our study focuses on the area within China ($73^\circ–105^\circ$E, $25^\circ–40^\circ$N), which encompasses six provinces, namely the entire Tibet Autonomous Region and the Qinghai Province, and parts of the Xin Jiang Uygur Autonomous Region, Gansu, Sichuan, and Yunnan Provinces. In the TP, topography is extremely complicated, but it has an overall descending tendency from west to east. Specially, the highest mountain in the world, Mount Everest, stands on the southern fringe of the TP. Due to its unique geographical location and high elevation, the TP has great influence on regional and even global energy and water cycles through thermal and dynamical forcings [43]. Climate on the TP is influenced by multiple climatic systems, which are characterized by the Indian monsoon and East Asian monsoon in the summer and Westerlies in the winter. The monsoon-dominated rainfall varies greatly in space with an obvious southeast to northwest gradient, and annual precipitation ranges from over 1500 mm to less than 100 mm [44]. In the summer, the southeast monsoon produces heavy precipitation, specifically in southeastern Tibet. In the winter, the snow brought by westerly winds is the main precipitation form in western regions [45]. The TP has an overall low annual mean temperature below 0 $^\circ$C, with the warmest month in July. Overall, the annual mean temperature decreases from east to west across the TP. Correspondingly, vegetation types also exhibit a spatial transition from forests in the southeast, to temperate shrubland or meadow in the middle, to desert land in the northwest regions [46]. Many of Asia’s major rivers, such as the Indus, Yangtze, Yellow, Brahmaputra, Mekong, and Salween rivers, all originate from the TP. In this study, the headwater region of the Yellow River, with the outlet of the Tangnaihai hydrological station, was selected for hydrological evaluation of the satellite precipitation products.

2.2. Satellite Data

IMERG is the level 3 multi-satellite precipitation algorithm of GPM, which combined all available constellation microwave precipitation estimates, infrared (IR) satellite estimates, and monthly precipitation gauge data, with the intention of creating a new generation of global precipitation products [23]. The IMERG system is run several times, first to produce an “early” multi-satellite product (~4 h after observation time) for giving a preliminary estimate, and successively providing “late” multi-satellite products (~12 h after observation time) as more data arrive. When the monthly gauge data is received, the “final” satellite-gauge product (~2.5 months after the observation month) is created as research [47]. For these three IMERG products, in addition to the difference in sensor data latency, the early and late estimates use climatological gauge data, while the final product is adjusted by monthly Global Precipitation Climatology Centre (GPCC) gauge data [48]. There
are two precipitation data field variables embedded in each IMERG product: precipitationCal and precipitationUncal. PrecipitationCal means a multi-satellite precipitation estimate with gauge calibration, and precipitationUncal presents the original multi-satellite precipitation estimate. For a more detailed description of IMERG algorithm, readers can refer to Huffman et al. [26,40,47]. In this study, we used both calibrated precipitation estimate and uncalibrated precipitation estimates from Final Run IMERG version 5 (hereafter referred to as “IMERG-C” and “IMERG-UC”, respectively).

**Figure 1.** Map and topography of the Tibetan Plateau (TP), meteorological stations, and Tangnaihai hydrological station.

GSMaP is a satellite-based precipitation map algorithm that combines various available PMW and IR sensors aiming to develop high-precision precipitation products. The GSMaP products are produced in several steps. Firstly, the instantaneous precipitation rate is retrieved based on the PMW radiometers from different satellite platforms, including GMI, advanced microwave scanning radiometer 2 (AMSR2), TRMM Microwave Imager (TMI), special sensor microwave imager/sounder (SSMIS), advanced microwave sounding unit-A (AMSU-A), and microwave humidity sounder (MHS) [49]. Then, the gaps between PMW-based estimates are propagated using the cloud motion vectors computed from geo-IR images, and a Kalman filter approach is applied to refine the precipitation rate [50,51]. Next, the forward and backward propagated precipitation estimates are weighted and combined to generate the GSMaP-MVK product. Unlike GSMaP-MVK, GSMaP-NRT is a simplified algorithm that only considers temporary forward cloud movement to keep operability and data latency in near real time. The GSMaP-Gauge is a gauge-calibrated product that adjusts the GSMaP-MVK estimate with National Oceanic and Atmospheric Administration (NOAA)/Climate Prediction Center (CPC) gauge-based analysis of global daily precipitation [52]. In the current study, GSMaP-MVK and GSMaP-Gauge with the latest version 7 were used.

Overall, four satellite-based precipitation datasets (IMERG-UC, IMERG-C, GSMaP-MVK, and GSMaP-Gauge), from April 2014 to March 2017, were selected for analysis in this study. Basic information of the four datasets is provided in Table 1. IMERG-UC and GSMaP-MVK belong to the “satellite-only” class, while IMERG-C and GSMaP-Gauge are “gauge-adjusted” products. In order to keep consistent with the ground reference dataset, both IMERG and GSMaP products were aggregated into daily readings, with 0.25° × 0.25° resolution. Generally, satellite precipitation products perform better when scaled up to a larger spatial average and longer time periods [48,53]. Readers should keep in mind that the resampling might introduce additional uncertainties.
Table 1. Coverage and spatiotemporal resolutions of satellite precipitation products used in this study.

<table>
<thead>
<tr>
<th>Product</th>
<th>Temporal Resolution</th>
<th>Spatial Resolution</th>
<th>Start Time</th>
<th>Latency</th>
<th>Coverage</th>
<th>Corrected by Gauges</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMERG-UC</td>
<td>0.5 h</td>
<td>0.1°</td>
<td>12 March 2014</td>
<td>3 months</td>
<td>90°N–90°S</td>
<td>No</td>
</tr>
<tr>
<td>IMERG-C</td>
<td>0.5 h</td>
<td>0.1°</td>
<td>12 March 2014</td>
<td>3 months</td>
<td>90°N–90°S</td>
<td>Yes (GPCC monthly)</td>
</tr>
<tr>
<td>GSMaP-MVK</td>
<td>1 h</td>
<td>0.1°</td>
<td>1 March 2014</td>
<td>3 days</td>
<td>60°N–60°S</td>
<td>No</td>
</tr>
<tr>
<td>GSMaP-Gauge</td>
<td>1 h</td>
<td>0.1°</td>
<td>1 March 2014</td>
<td>3 days</td>
<td>60°N–60°S</td>
<td>Yes (CPC daily)</td>
</tr>
</tbody>
</table>

2.3. Ground Gauge Data

The China Gauge-based Daily Precipitation Analysis (CGDPA), operated by the National Meteorological Information Center (NMIC) of the China Meteorological Administration (CMA), was used as the reference dataset in this study. All the gauge data used in CGDPA are manually recorded by bucket rain gauges and underwent a strict quality control procedure [54]. Using the optimal interpolation (OI) method, daily precipitation data from about 2400 gauge stations in mainland China were interpolated to obtain the gridded CGDPA product (0.25° × 0.25°). The CGDPA is considered as a high-quality ground precipitation dataset and was previously adopted to evaluate satellite precipitation products over China [55–57].

For the hydrological evaluation, the daily meteorological observations from 2009 to 2015 at the upper Yellow River basin were obtained from CMA, including daily maximum and minimum temperature, and daily average wind speed. The daily streamflow data at the Tangnaihai hydrological station from 2009 to 2015 were primarily collected from the Chinese Hydrology Almanac, which is published by the Hydrological Bureau of the Ministry of Water Resources in the People’s Republic of China.

2.4. Geographical Data

Geographical data, such as underlying surface information required by the Variable Infiltration Capacity (VIC) model, including soil texture, topography, and vegetation types, were also used in this study. The global Digital Elevation Model (DEM) data (GTOPO30) with a resolution of thirty arcseconds were downloaded from the website of the US Geological Survey (USGS; https://lta.cr.usgs.gov/GTOPO30). The soil texture information was obtained from the Food and Agriculture Organization dataset [58]. The vegetation data were taken from the global vegetation classifications database developed by the University of Maryland, which has 14 types of land use/cover with a spatial resolution of 1 km × 1 km [59].

3. Methodology

3.1. Verification Metrics

To quantitatively compare GPM-era satellite precipitation products, several widely used statistical metrics were applied in this study. The formulas and perfect values of these metrics are listed in Table 2. The Pearson correlation coefficient (CC) describes the agreement between the satellite precipitation and gauge observations. Mean error (ME) simply scales the difference between the satellite estimates and the reference. The root-mean-squared error (RMSE) corresponds to the square root of the average of the squared differences between the estimates and the observed rainfall, and was used to measure the average error magnitude, while relative bias (RB) and relative root-mean-squared error (RRMSE) were used to calculate the systematic and random components of the error in the satellite precipitation products. For the contingency table metrics, following Ebert et al. [19], we adopted four skill scores, namely probability of detection (POD), false alarm rate (FAR), frequency bias index (FBI), and equitable threat score (ETS), to examine the detection capabilities in rain events of satellite estimates. The ETS is commonly used as an overall skill measure by the numerical weather prediction community, and the POD, FAR, and FBI present complementary information about misses, false alarms, and bias.
The distinguishing characteristic of the VIC is that it incorporates land-cover vegetation 
were more consistent with the ground measurements than the satellite-only products. After gauge 
validation (Table 2).

### 3.2. Hydrological Model

The hydrological model used in this study was the Variable Infiltration Capacity (VIC) 
model [60,61], which is a grid-based distributed hydrological model maintained by the University of 
Washington. The distinguishing characteristic of the VIC is that it incorporates land-cover vegetation 
heterogeneity, multiple soil layers with spatially varying infiltration capacity, and base flow as a 
non-linear recession curve [62]. The model also considers dynamic changes of both water and energy 
balances over a grid mesh. For a more detailed description of the VIC model, the reader is referred 
to Nijssen et al. [63] and Xie et al. [64]. Presently, the VIC model is widely applied in hydrological 
simulations and evaluations of satellite precipitations over different basins [65–67]. Here, the VIC 
model was implemented at a spatial resolution of 0.1° for the upper Yellow River basin. The daily 
streamflow observations were used to calibrate and validate the VIC model. Three statistical indices 
(Nash–Sutcliffe efficiency (NSE), RB and RMSE) were used to assess the result of hydrologic simulation 
validation (Table 2).

### 4. Results and Discussion

#### 4.1. Rainfall Characteristics of the TP

Figure 2 shows the spatial distributions of three-year daily mean precipitation for CGDPA and 
four GPM-era precipitation products over the TP. Generally, all precipitation sets shared similar spatial 
distribution: the mean precipitation decreased from the southeast to the northwest. One can see that a 
large amount of precipitation (more than 5 mm/day) concentrates over the southern region of the TP. 
The reason for this is the effect of the Himalayas, which intercept the moisture from the Indian Ocean 
monsoon and induce rainfall. In contrast, less precipitation is observed in most of the west and north, 
where the Westerlies do not prevail and the Indian monsoon is relatively weaker [68]. Compared with 
the CGDPA, a pronounced difference can be found between the IMERG and GSMaP precipitation 
products. IMERG-UC significantly underestimated the TP’s precipitation and the GSMaP-MVK clearly 
overestimated the precipitation. However, the precipitation estimates of both gauge-adjusted products 
were more consistent with the ground measurements than the satellite-only products. After gauge 
correction, the under- and overestimations of satellite-only products were mitigated, especially in

### Table 2. List of the validation statistical metrics for evaluating satellite precipitation products.

<table>
<thead>
<tr>
<th>Statistic Index</th>
<th>Formula</th>
<th>Unit</th>
<th>Perfect Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation coefficient (CC)</td>
<td>$CC = \frac{\sum_{i=1}^{n}(G_i - \overline{G})(S_i - \overline{S})}{\sqrt{\sum_{i=1}^{n}(G_i - \overline{G})^2} \cdot \sqrt{\sum_{i=1}^{n}(S_i - \overline{S})^2}}$</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Root-mean-squared error (RMSE)</td>
<td>$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (S_i - G_i)^2}$</td>
<td>mm</td>
<td>0</td>
</tr>
<tr>
<td>Mean error (ME)</td>
<td>$ME = \frac{1}{n} \sum_{i=1}^{n} (S_i - G_i)$</td>
<td>mm</td>
<td>0</td>
</tr>
<tr>
<td>Relative root-mean-squared error (RRMSE)</td>
<td>$RRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^{n} (S_i - G_i)^2}}{\frac{1}{n} \sum_{i=1}^{n} G_i}$</td>
<td>%</td>
<td>0</td>
</tr>
<tr>
<td>Relative bias (RB)</td>
<td>$RB = \frac{\sum_{i=1}^{n} (S_i - G_i)}{\frac{1}{n} \sum_{i=1}^{n} G_i}$</td>
<td>%</td>
<td>0</td>
</tr>
<tr>
<td>Probability of detection (POD)</td>
<td>$POD = \frac{G_i}{G_i + M_i}$</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>False alarm ratio (FAR)</td>
<td>$FAR = \frac{F_i}{F_i + M_i}$</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Frequency bias index (FBI)</td>
<td>$FBI = \frac{F_i}{F_i + M_i}$</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Equitable Threat Score (ETS)</td>
<td>$ETS = \frac{H_i - H_e}{(H_i + M_i)(H_i + F_i)}$</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Nash–Sutcliffe coefficient of efficiency (NSE)</td>
<td>$NSE = 1 - \frac{\sum_{i=1}^{n} (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^{n} (Q_{obs,i} - Q_{mean})^2}$</td>
<td>-</td>
<td>1</td>
</tr>
</tbody>
</table>

Notation: $n$, number of samples; hit ($H$, observed rain correctly detected); miss ($M$, observed rain not detected); false ($F$, rain detected but not observed); null ($N$, no rain observed nor detected); $S_i$, satellite precipitation; $G_i$, gauged observation; $Q_{obs,i}$, observed streamflow; $Q_{sim,i}$, simulated streamflow; $Q_{mean}$, mean value of observed streamflow.
the eastern TP with more rain gauges. Therefore, the strategy of gauge adjustments using in situ measurements greatly improved the product accuracy.

![Spatial distributions of mean annual precipitation for (a) CGDPA, (b) IMERG-UC, (c) IMERG-C, (d) GSMaP-MVK, and (e) GSMaP-Gauge over the TP during the period of April 2014–March 2017.](image)

Figure 2. Spatial distributions of mean annual precipitation for (a) CGDPA, (b) IMERG-UC, (c) IMERG-C, (d) GSMaP-MVK, and (e) GSMaP-Gauge over the TP during the period of April 2014–March 2017.

Next, we displayed the rain rate distribution of daily precipitation amount and the daily number of precipitation events over the TP. As shown in Figure 3, for the intensity distribution from the CGDPA, the distribution of the precipitating amount over the TP presents a single-peak pattern (approximately at a rain rate of 14 mm/d), and the most precipitating days occurred in the precipitation intensity across the range of 2–8 mm·day$^{-1}$. The satellite precipitation products exhibit similar distribution patterns with CGDPA in terms of precipitation rates for both the occurrence and the volume of precipitation, but there are some differences in the intensity distribution curves. For example, at the light rain range (0.25–2 mm·day$^{-1}$), all the satellite datasets detected more precipitation events than the reference datasets (Figure 3b). Consequently, they contributed to the precipitation volumes, with the satellite-based estimates having more precipitation than the ground measurements at the light rain range.
with an RB value of 26.11%. After the gauge calibration, the RB had a downward trend and the IMERG-UC and IMERG-C products showed middle performance, with CC values of 0.67
precipitation products over the period from April 2014 to March 2017. In order to ensure a more accurate comparison, only grid pixels with at least one gauge (132 grids) were taken to calculate the statistical metrics.

Figure 3. Intensity distribution of (a) daily precipitation amount (mm day\(^{-1}\)) and (b) precipitation events (count per day) from the four satellite precipitation estimates at the TP. The logarithmic scale was used to bin the precipitation rates.

4.2. Statistical Performance of Satellite Precipitation Estimates

In this section, we evaluated the satellite precipitation products against gridded gauge-based precipitation products over the period from April 2014 to March 2017. In order to ensure a more accurate comparison, only grid pixels with at least one gauge (132 grids) were taken to calculate the statistical metrics.

Figure 4 shows the scatterplots of daily IMERG-UC, IMERG-C, GSMaP-MVK, and GSMaP-Gauge data versus CGDPA for the selected grids. The evaluation metrics are also given in the figure. For the contingency table statistics (i.e., POD, FAR, FB, and ETS), a common threshold of 1.0 mm day\(^{-1}\) was used to determine the rain/no rain event, as suggested by many previous studies [7,69–71]. Generally, among the four satellite precipitation products, GSMaP-Gauge exhibited the best performance with the highest CC of 0.77, while GSMaP-MVK had the worst performance with a poor CC of 0.52. The IMERG-UC and IMERG-C products showed middle performance, with CC values of 0.67 and 0.70, respectively. For the satellite-only products, we can see that IMERG-UC significantly underestimated the precipitation by about −39.32%, and GSMaP-MVK overestimated the precipitation with an RB value of 26.11%. After the gauge calibration, the RB had a downward trend and the scatter points were clustered more closely to the 1:1 line than those of satellite-only estimates. Consequently, both gauge-adjusted products only showed slight underestimation relative to reference observations (Figure 4b,d). In terms of the contingency table statistics, when compared to satellite-only products, the gauge-adjusted products also had better performance (with higher POD and ETS values). The gauge-adjusted products had more detected events compared to satellite-only products. However, it is worth noting that the FAR increased from IMERG-UC to IMERG-C. We argue that the calibration scheme of IMERG resulted in an increase in the number of false events over the TP. Thus, it contributed to the observed increases in FAR values. Correspondingly, the RMSE and RRMSE of IMERG-C did not improve over that of IMERG-UC.
As shown in Figure 5c, IMERG-UC underestimated reference precipitation over almost all of TP, while other satellite precipitation products. For GSMaP-MVK, the magnitude of the total RB was reduced. However, focusing on the spatial distribution of RMSE, we can see that GSMaP-MVK showed overestimation in the east and underestimation in the west (Figure 5i). Thus, for GSMaP-MVK, the reason is that the positive and negative biases could cancel each other out. As shown in Figure 5c, IMERG-UC underestimated reference precipitation over almost all of TP, while GSMaP-MVK showed overestimation in the east and underestimation in the west (Figure 5i). Thus, for GSMaP-MVK, the magnitude of the total RB was reduced. However, focusing on the spatial distribution of RMSE, we can see that GSMaP-MVK still demonstrated the largest error among the four satellite precipitation products.

To investigate the spatial distributions of error metrics, CC, RMSE, and RB were computed from the four satellite precipitation products, as shown in Figure 5. In general, the CC values of all products were good over most regions of TP. Spatially, higher CC values are observed in the east TP compared to the west. This pattern of CC is attributed to the limitations of retrieval of satellite precipitation in mountainous and high-elevation regions [19,66]. We also note that the GSMaP-Gauge showed best correspondence with gauge measurements, with larger correlation and smaller error (Figure 5j–l), which is consistent with the above statistical results. Interestingly, the spatial distributions of error metrics could explain the phenomenon of the RMSE of IMERG-UC being lower than that of GSMaP-MVK, but the bias performance of IMERG-UC was worse than that of GSMaP-MVK. The reason is that the positive and negative biases could cancel each other out. As shown in Figure 5c, IMERG-UC underestimated reference precipitation over almost all of TP, while GSMaP-MVK showed overestimation in the east and underestimation in the west (Figure 5i). Thus, for GSMaP-MVK, the magnitude of the total RB was reduced. However, focusing on the spatial distribution of RMSE, we can see that GSMaP-MVK still demonstrated the largest error among the four satellite precipitation products.

Figure 4. Scatterplots of daily precipitation for (a) IMERG-UC; (b) IMERG-C; (c) GSMaP-MVK; (d) and GSMaP-Gauge against CGDPA from the 132 selected 0.25° grid boxes over the TP.
Figure 5. Spatial distribution of (a,d,g,j) correlation coefficient (CC); (b,e,h,k) root-mean-squared error (RMSE); and (c,f,i,l) relative bias (RB) between the four satellite precipitation estimates and CGDPA at the 132 grid central points.

Figure 6 shows the temporal variations of averaged spatial precipitation and statistics for the selected grid boxes. Table 3 lists the statistical summary of seasonal comparisons including spring (March–May), summer (June–August), autumn (September–November), and winter (December–February) by computing at the daily scale. Overall, the patterns of monthly mean precipitation for all satellite precipitation products exhibited similar fluctuations with gauge observations. Precipitation in the summer is the main water source over the TP, while precipitation in the winter only contributes a minor part of annual precipitation. The performance of satellite precipitation productions showed distinct seasonal variations. The statistical indices performed better in the summer than the other three seasons with high correlation, low relative error, and better detection for rain events (see Table 3). For instance, the CC value of the four satellite precipitations ranged from 0.59 to 0.76 during the summer, while a lower CC occurred in the winter. We also note that the RMSE was larger during the summer months than that of the winter season (Table 3 and Figure 6c). This is because RMSE could be affected by large precipitation bias. The RB results indicate that GSMaP-MVK overestimated the reference precipitation, except during the winter, and IMERG-UC, IMERG-C, and GSMaP-Gauge underestimated the gauge observations in all seasons. Similar to the results of Duan et al. [50] and Ning et al. [72], satellite precipitation products showed higher error and poor capability of rainfall detection in the winter months. During the winter season, although the DPR improved the skill of snowfall observations, satellite precipitation products still showed unsatisfactory
performance over the TP. This can be attributed to the limitation of passive microwave retrievals and IR information at cold or snow-covered background surfaces [22], suggesting that the current GPM-era estimates still have room for improvement in the winter.

Figure 6. (a) Average monthly precipitation time series and monthly variations of statistical indices: (b) CC, (c) RMSE, and (d) RB.

Table 3. Seasonal statistics of four satellite precipitation estimates (IMERG-UC, IMERG-C, GSMaP-MVK, and GSMaP-Gauge) against ground observations over the Tibetan Plateau (TP) during the study period of April 2014–March 2017.

<table>
<thead>
<tr>
<th>Season</th>
<th>Product</th>
<th>CC</th>
<th>RMSE (mm)</th>
<th>ME (mm)</th>
<th>RRMSE (%)</th>
<th>RB (%)</th>
<th>POD</th>
<th>FAR</th>
<th>FBI</th>
<th>ETS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring</td>
<td>IMERG-UC</td>
<td>0.56</td>
<td>3.26</td>
<td>-0.62</td>
<td>242.87</td>
<td>-46.38</td>
<td>0.51</td>
<td>0.30</td>
<td>0.74</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>IMERG-C</td>
<td>0.61</td>
<td>3.17</td>
<td>-0.18</td>
<td>236.01</td>
<td>-13.60</td>
<td>0.64</td>
<td>0.38</td>
<td>1.03</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>GSMaP-MVK</td>
<td>0.38</td>
<td>6.24</td>
<td>0.43</td>
<td>464.42</td>
<td>31.94</td>
<td>0.64</td>
<td>0.44</td>
<td>1.14</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>GSMaP-Gauge</td>
<td>0.71</td>
<td>2.71</td>
<td>-0.21</td>
<td>201.69</td>
<td>-15.53</td>
<td>0.82</td>
<td>0.30</td>
<td>1.17</td>
<td>0.52</td>
</tr>
<tr>
<td>Summer</td>
<td>IMERG-UC</td>
<td>0.67</td>
<td>5.15</td>
<td>-1.26</td>
<td>139.24</td>
<td>-34.02</td>
<td>0.72</td>
<td>0.22</td>
<td>0.93</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>IMERG-C</td>
<td>0.68</td>
<td>5.27</td>
<td>-0.10</td>
<td>142.52</td>
<td>-2.74</td>
<td>0.80</td>
<td>0.26</td>
<td>1.07</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>GSMaP-MVK</td>
<td>0.89</td>
<td>8.01</td>
<td>0.86</td>
<td>216.56</td>
<td>23.16</td>
<td>0.78</td>
<td>0.27</td>
<td>1.08</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>GSMaP-Gauge</td>
<td>0.76</td>
<td>4.36</td>
<td>-0.08</td>
<td>117.87</td>
<td>-2.03</td>
<td>0.91</td>
<td>0.24</td>
<td>1.20</td>
<td>0.50</td>
</tr>
<tr>
<td>Autumn</td>
<td>IMERG-UC</td>
<td>0.66</td>
<td>3.18</td>
<td>-0.57</td>
<td>218.98</td>
<td>-39.09</td>
<td>0.60</td>
<td>0.25</td>
<td>0.81</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>IMERG-C</td>
<td>0.71</td>
<td>3.12</td>
<td>-0.11</td>
<td>214.77</td>
<td>-7.44</td>
<td>0.69</td>
<td>0.31</td>
<td>1.01</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>GSMaP-MVK</td>
<td>0.44</td>
<td>6.96</td>
<td>0.52</td>
<td>478.63</td>
<td>35.58</td>
<td>0.70</td>
<td>0.37</td>
<td>1.12</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>GSMaP-Gauge</td>
<td>0.77</td>
<td>2.67</td>
<td>-0.06</td>
<td>183.5</td>
<td>-4.41</td>
<td>0.87</td>
<td>0.27</td>
<td>1.19</td>
<td>0.57</td>
</tr>
<tr>
<td>Winter</td>
<td>IMERG-UC</td>
<td>0.54</td>
<td>1.45</td>
<td>-0.21</td>
<td>522.96</td>
<td>-76.56</td>
<td>0.13</td>
<td>0.43</td>
<td>0.23</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>IMERG-C</td>
<td>0.58</td>
<td>1.39</td>
<td>-0.17</td>
<td>500.66</td>
<td>-62.17</td>
<td>0.19</td>
<td>0.54</td>
<td>0.42</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>GSMaP-MVK</td>
<td>0.30</td>
<td>1.95</td>
<td>-0.05</td>
<td>701.76</td>
<td>-18.30</td>
<td>0.29</td>
<td>0.66</td>
<td>0.85</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>GSMaP-Gauge</td>
<td>0.58</td>
<td>1.40</td>
<td>-0.13</td>
<td>501.71</td>
<td>-45.07</td>
<td>0.52</td>
<td>0.38</td>
<td>0.84</td>
<td>0.38</td>
</tr>
</tbody>
</table>
For analyzing the error characteristics of different precipitation event, following the error decomposition approach proposed by Tian et al. [69], the total bias can be decomposed into different parts: hit bias, bias due to rainfall misses, bias due to false detections, and bias with selected threshold. As shown in Figure 7, hit bias and total bias share considerable similarities in their spatial distributions, suggesting that hit bias is the dominant component of total bias. Considering that the negative bias with missed precipitation and positive bias with false precipitation have opposite signs, they can offset each other, resulting in a smaller total bias of satellite precipitation. The lighter precipitation (<1 mm/day) that we considered unreliable for either gauge data or satellite measurements contributed to only a small part of total bias and can be ignored (Table 4). In addition, the error components also showed seasonal dependence. Generally speaking, it is apparent that the values of total bias are lower in the summer than in the winter. Particularly, in the winter, higher miss bias was found, and missed precipitation was the dominant source of errors. This phenomenon indicates that the satellite estimates miss a lot of precipitation events in winter, and confirms our aforementioned speculation: the GPM-era satellite precipitation products still exhibit some deficiencies for detecting snowing events.

Figure 7. Spatial patterns of the error components of satellite precipitation estimates against CGDPA at the 132 grid central points (mm/day): total bias (first row), hit bias (second row), missed precipitation (third row), false precipitation (fourth row), and bias with selected threshold (fifth row).
worse in some cases. For example, IMERG-UC overestimated gauge observations at lower rain rates; works in precipitation estimates. It is seen that the IMERG-C elevated the precipitation estimates precipitation products. This gauge-calibration effectively reduced total bias while making things however, the bias calibration using GPCC gauge data elevated the precipitation estimates further in low rain rates. On the other hand, from the results of Figure 8, we can see how the calibration scheme evidencing that current satellite precipitation products need to continue improving the performance at higher rain rates [76,77]. In terms of RRMSE, satellite-based retrievals, as documented in previous studies [73–75]. This error feature of rain-rate light rain and underestimations for heavy rain (Figure 8a), which is a common error feature of obviously, all precipitation products showed a similar variation of error, with overestimations for high rain rates compared to moderate–high rain rates (Figure 8b), indicating that current satellite precipitation products need to continue improving the performance at low rain rates. On the other hand, from the results of Figure 8, we can see how the calibration scheme works in precipitation estimates. It is seen that the IMERG-C elevated the precipitation estimates and GSMaP-Gauge decreased the precipitation values compared to their corresponding uncalibrated precipitation products. This gauge-calibration effectively reduced total bias while making things worse in some cases. For example, IMERG-UC overestimated gauge observations at lower rain rates; however, the bias calibration using GPCC gauge data elevated the precipitation estimates further augmenting the overestimation at lower rain rates. Another case involved the GSMaP-Gauge showing large negative hit bias than GSMaP-MVK in the winter season (Table 4). Thus, it seems important to calibrate satellite-based precipitation estimates at difference rain rates or seasons in the future.

Table 4. Seasonally averaged error components as percentages of total observed precipitation.

<table>
<thead>
<tr>
<th>Season</th>
<th>Product</th>
<th>H (%)</th>
<th>−M (%)</th>
<th>F (%)</th>
<th>N (%)</th>
<th>E = H − M + F + N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring</td>
<td>IMERG-UC</td>
<td>−27.84</td>
<td>−28.62</td>
<td>8.07</td>
<td>2.04</td>
<td>−46.35</td>
</tr>
<tr>
<td></td>
<td>IMERG-C</td>
<td>−13.73</td>
<td>−19.75</td>
<td>16.66</td>
<td>3.24</td>
<td>−13.58</td>
</tr>
<tr>
<td></td>
<td>GSMaP-MVK</td>
<td>18.94</td>
<td>−22.63</td>
<td>33.22</td>
<td>2.41</td>
<td>31.94</td>
</tr>
<tr>
<td></td>
<td>GSMaP-Gauge</td>
<td>−21.52</td>
<td>−8.31</td>
<td>11.14</td>
<td>3.16</td>
<td>−15.53</td>
</tr>
<tr>
<td>Summer</td>
<td>IMERG-UC</td>
<td>−28.70</td>
<td>−12.31</td>
<td>6.36</td>
<td>0.62</td>
<td>−34.03</td>
</tr>
<tr>
<td></td>
<td>IMERG-C</td>
<td>−5.89</td>
<td>−8.20</td>
<td>10.59</td>
<td>0.74</td>
<td>−2.76</td>
</tr>
<tr>
<td></td>
<td>GSMaP-MVK</td>
<td>18.45</td>
<td>−10.12</td>
<td>13.99</td>
<td>0.85</td>
<td>23.17</td>
</tr>
<tr>
<td></td>
<td>GSMaP-Gauge</td>
<td>−9.47</td>
<td>−3.26</td>
<td>9.49</td>
<td>1.20</td>
<td>−2.04</td>
</tr>
<tr>
<td>Autumn</td>
<td>IMERG-UC</td>
<td>−26.67</td>
<td>−20.41</td>
<td>6.69</td>
<td>1.30</td>
<td>−39.09</td>
</tr>
<tr>
<td></td>
<td>IMERG-C</td>
<td>−6.72</td>
<td>−14.80</td>
<td>12.22</td>
<td>1.88</td>
<td>−7.42</td>
</tr>
<tr>
<td></td>
<td>GSMaP-MVK</td>
<td>19.75</td>
<td>−15.31</td>
<td>28.57</td>
<td>2.57</td>
<td>35.58</td>
</tr>
<tr>
<td></td>
<td>GSMaP-Gauge</td>
<td>−11.85</td>
<td>−5.85</td>
<td>11.28</td>
<td>2.00</td>
<td>−4.42</td>
</tr>
<tr>
<td>Winter</td>
<td>IMERG-UC</td>
<td>−12.94</td>
<td>−67.39</td>
<td>4.42</td>
<td>−0.65</td>
<td>−76.56</td>
</tr>
<tr>
<td></td>
<td>IMERG-C</td>
<td>−13.16</td>
<td>−60.48</td>
<td>9.31</td>
<td>2.19</td>
<td>−62.14</td>
</tr>
<tr>
<td></td>
<td>GSMaP-MVK</td>
<td>−6.04</td>
<td>−53.34</td>
<td>37.97</td>
<td>5.07</td>
<td>−18.34</td>
</tr>
<tr>
<td></td>
<td>GSMaP-Gauge</td>
<td>−32.33</td>
<td>−28.44</td>
<td>11.40</td>
<td>4.31</td>
<td>−43.06</td>
</tr>
<tr>
<td>All Season</td>
<td>IMERG-UC</td>
<td>−27.46</td>
<td>−19.59</td>
<td>6.71</td>
<td>1.01</td>
<td>−39.33</td>
</tr>
<tr>
<td></td>
<td>IMERG-C</td>
<td>−7.96</td>
<td>−14.06</td>
<td>12.14</td>
<td>1.56</td>
<td>−8.32</td>
</tr>
<tr>
<td></td>
<td>GSMaP-MVK</td>
<td>17.89</td>
<td>−15.59</td>
<td>22.09</td>
<td>1.72</td>
<td>26.11</td>
</tr>
<tr>
<td></td>
<td>GSMaP-Gauge</td>
<td>−13.35</td>
<td>−5.84</td>
<td>10.30</td>
<td>1.90</td>
<td>−6.99</td>
</tr>
</tbody>
</table>

Figure 8 displays the error characteristics of satellite precipitation estimates with rain rate. Obviously, all precipitation products showed a similar variation of error, with overestimations for light rain and underestimations for heavy rain (Figure 8a), which is a common error feature of satellite-based retrievals, as documented in previous studies [73–75]. This error feature of rain-rate dependency is important for meteorological and hydrological applications, especially for typhoon monitoring and flood forecast, which are sensitive to higher rain rates [76,77]. In terms of RRMSE, higher values were found at low rain rates compared to moderate–high rain rates (Figure 8b), indicating that current satellite precipitation products need to continue improving the performance at low rain rates. On the other hand, from the results of Figure 8, we can see how the calibration scheme works in precipitation estimates. It is seen that the IMERG-C elevated the precipitation estimates and GSMaP-Gauge decreased the precipitation values compared to their corresponding uncalibrated precipitation products. This gauge-calibration effectively reduced total bias while making things worse in some cases. For example, IMERG-UC overestimated gauge observations at lower rain rates; however, the bias calibration using GPCC gauge data elevated the precipitation estimates further augmenting the overestimation at lower rain rates. Another case involved the GSMaP-Gauge showing large negative hit bias than GSMaP-MVK in the winter season (Table 4). Thus, it seems important to calibrate satellite-based precipitation estimates at difference rain rates or seasons in the future.

![Figure 8](image-url)  
**Figure 8.** The relationship between the validation indices and precipitation rate: (a) RB, and (b) RRMSE.
4.3. Hydrological Evaluation of Satellite Precipitation Estimates

In the previous section, we compared the GPM-era satellite precipitation products against the rain gauge observations; the next step was to evaluate the hydrological utility of these precipitation datasets. In this section, since the streamflow data after 2015 were not available at the Tangnaihai hydrological station, the hydrological evaluation of four satellite precipitation estimates was performed for the whole year of 2015. We also calculated statistical indices of precipitation estimates over the upper Yellow River basin in 2015 (Table 5). By analyzing this indices, we can conclude that the error characteristics of satellite precipitation during 2005 are consistent with prior comparison results. For example, the CC values of daily IMERG-UC, IMERG-C, GSMaP-MVK, and GSMaP-Gauge estimates were 0.57, 0.61, 0.52, and 0.75 over the upper Yellow River basin during 2005, and 0.67, 0.70, 0.52, and 0.77 over the TP during the periods of April 2014 to March 2017, respectively. For the basin-scale evaluation, statistical values with the basin-averaged data were better than those with the grid-scale evaluation. This is expected because random errors would decrease with spatial scale averaging. Thus, both grid-scale and basin-scale analysis confirm that the performance of satellite precipitation in the upper Yellow River basin was similar to that on the TP.

Table 5. Statistical summary of the precipitation products at grid and basin scales for the year of 2015 in the upper Yellow River basin.

<table>
<thead>
<tr>
<th>Product</th>
<th>CC</th>
<th>RMSE (mm)</th>
<th>RB (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid scale</td>
<td>IMERG-UC</td>
<td>0.57</td>
<td>3.10</td>
</tr>
<tr>
<td></td>
<td>IMERG-C</td>
<td>0.61</td>
<td>3.30</td>
</tr>
<tr>
<td></td>
<td>GSMaP-MVK</td>
<td>0.52</td>
<td>5.78</td>
</tr>
<tr>
<td></td>
<td>GSMaP-Gauge</td>
<td>0.75</td>
<td>2.41</td>
</tr>
<tr>
<td>Baseline</td>
<td>IMERG-UC</td>
<td>0.74</td>
<td>1.65</td>
</tr>
<tr>
<td></td>
<td>IMERG-C</td>
<td>0.79</td>
<td>1.60</td>
</tr>
<tr>
<td></td>
<td>GSMaP-MVK</td>
<td>0.74</td>
<td>2.70</td>
</tr>
<tr>
<td></td>
<td>GSMaP-Gauge</td>
<td>0.93</td>
<td>0.82</td>
</tr>
</tbody>
</table>

Next, the VIC model was calibrated and validated with observed precipitation and streamflow for the periods of 2009–2011 and 2012–2014 over the upper Yellow River basin. Figure 9 shows the CGDPA-simulated and observed streamflow at the daily scale. Comparing the observed and simulated streamflow, the values of NSE and RB were 0.61% and −2.56% during the calibration period, and NSE increased to 0.73 and RB of 0.96% in the validation period. It can be seen that the simulated streamflow generally agrees with observations very well, although overestimation and underestimation of peak floods existed in some cases. Improved results were obtained at the monthly scale for both the calibration and validation periods (NSE of 0.88 and 0.87, respectively).

After the model was benchmarked by the in situ data, the VIC model was then driven by gauge- and satellite-based precipitation datasets for the period from 1 January 2015 to 31 December 2015, without any further adjustment of parameters. The simulated and observed hydrographs are shown in Figure 10, and the statistical comparisons are summarized in Table 6. As shown, the CGDPA had a worse performance in 2015, with NSE of 0.41 and a runoff overestimation of 26.39%. The observed mean daily discharge from 2009 to 2014 was 711.61 m$^3$/s, whereas that in 2015 was 480.82 m$^3$/s. Differences in hydrological features during the two periods may potentially influence the simulation performance. Using the same parameters enabled us to compare the performance of simulated streamflow from different precipitation inputs. For the GPM-era satellite precipitation products, the GSMaP-Gauge showed the best performance in the streamflow simulation; the IMERG-C took second place; and the two purely satellite-derived estimates demonstrated poor performance due to the large precipitation bias, especially for GSMaP-MVK with 151.97% runoff overestimation at a daily scale. Interestingly, the simulated streamflow with GSMaP-Gauge inputs had lightly better performance than the CGDPA (e.g., 0.53 versus 0.41 for NSE). We considered that gauge corrections involved in the GSMaP-Gauge products remarkably improved the skill of streamflow simulation. For the monthly comparisons,
GSMaP-Gauge performed the best again, while the NSE value of IMERG-C reached 0.63. However, the satellite-only products still had unsatisfactory performance with negative NSEs, suggesting they have low hydrological utility for this region.

![Figure 9](image_url)  
**Figure 9.** Observed and Variable Infiltration Capacity (VIC) model simulated streamflow with the CGDPA precipitation for the calibration period (2009–2011), and validation period (2012–2014) over the upper Yellow River basin.

![Figure 10](image_url)  
**Figure 10.** Daily observed and simulated streamflow with gauge benchmarked parameters over the upper Yellow River basin.

The simulation accuracy could be improved if the hydrology model was calibrated with different precipitation inputs. Subsequently, we recalibrated the model parameters using each satellite precipitation dataset during 2015. This scenario is also an alternative strategy for hydrological applications in ungauged basins where only satellite precipitation estimates are available [78,79]. As shown in Figure 11, the simulation performances were effectively improved after the model was recalibrated. For example, the daily NSE of IMERG-C significantly increased from 0.18 to 0.63. Furthermore, the RMSE also significantly decreased for all satellite products. As summarized in Table 6, simulations of IMERG-C and GSMaP-Gauge products had good statistical agreement with observed streamflow at daily and monthly scale. However, the NSE values of both satellite-only products was still below zero at the daily scale, further confirming that the undesirable hydrological utility of these two satellite-only products is mainly due to the unreliable precipitation estimates. The errors existing in these two precipitation datasets was propagated to simulated streamflow, and could not be removed.
upon model-parameter recalibration. Generally, the recalibration of the model parameters effectively improved the hydrological potential of satellite precipitation, especially for precipitation products with small errors; however, this recalibration approach should be taken with a grain of salt because it may result in unrealistic parameter values in some cases [80,81].

Table 6. Comparison of daily and monthly observed and simulated streamflow when the Variable Infiltration Capacity (VIC) model was forced by the gauge- and satellite-based precipitation datasets in 2015.

<table>
<thead>
<tr>
<th>Time Scales</th>
<th>Precipitation Products</th>
<th>Benchmarking Calibration</th>
<th>Product-Specific Calibration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>NSE</td>
<td>RB (%)</td>
</tr>
<tr>
<td>Daily</td>
<td>CGDPA</td>
<td>0.41</td>
<td>26.39</td>
</tr>
<tr>
<td></td>
<td>IMERG-UC</td>
<td>−0.12</td>
<td>−46.00</td>
</tr>
<tr>
<td></td>
<td>IMERG-C</td>
<td>0.18</td>
<td>17.97</td>
</tr>
<tr>
<td></td>
<td>GSMaP-MVK</td>
<td>−9.85</td>
<td>151.97</td>
</tr>
<tr>
<td></td>
<td>GSMaP-Gauge</td>
<td>0.53</td>
<td>11.64</td>
</tr>
<tr>
<td>Monthly</td>
<td>CGDPA</td>
<td>0.53</td>
<td>26.48</td>
</tr>
<tr>
<td></td>
<td>IMERG-UC</td>
<td>−0.17</td>
<td>−45.96</td>
</tr>
<tr>
<td></td>
<td>IMERG-C</td>
<td>0.63</td>
<td>18.05</td>
</tr>
<tr>
<td></td>
<td>GSMaP-MVK</td>
<td>−6.88</td>
<td>152.16</td>
</tr>
<tr>
<td></td>
<td>GSMaP-Gauge</td>
<td>0.74</td>
<td>11.73</td>
</tr>
</tbody>
</table>

Figure 11. Comparison of VIC simulated streamflow with recalibrated parameters using product-specific inputs.

5. Conclusions and Recommendations

After four years of operation of the GPM, it was proven to be a good successor to the TRMM, offering new opportunities for meteorological studies and hydrological applications, especially in mountainous areas. In this study, we investigated the performance of the latest GPM IMERG V5 and GSMaP V7 satellite precipitation products over the TP. Firstly, the statistical assessment of the accuracy of the satellite products was performed against gridded gauge-based data. Then, their hydrological simulation utilities were evaluated using the VIC model in the upper Yellow River basin.

Our study shows that IMERG and GSMaP products can appropriately capture spatial patterns of precipitation across the TP. IMERG-UC significantly underestimated the reference precipitation (−39.32%), while obvious overestimation was found in GSMaP-MVK (26.11%). After bias adjustment, the under- and overestimations were clearly reduced, with slight underestimation for IMERG-C.
Among the four studied satellite estimates, GSMaP-Gauge had the best performance in nearly all statistical indices with higher correlation, lower bias, and better detection, while GSMaP-MVK had worst performance. This results suggest that the calibration procedure in GSMaP-gauge performs better than that in IMERG-C. Considering that IMERG-C ingests monthly GPCC gauge analyses and that GSMaP-Gauge is calibrated by CPC daily gauge analyses, it seems important to use daily gauge data for IMERG in the future. In fact, the GPCC started computing daily adjustments based on station meteorological data, and it will be used in a future version of IMERG [47].

The GPM sensors were designed to extend the capability for detecting light-intensity precipitation and falling snow, making IMERG perform better than TMPA, which was confirmed by previous studies [29,32,82]. However, in the present study, we found that all satellite precipitation products had a relatively lower accuracy in winter, and missed bias was the main problem during the winter season. On the other hand, higher error values were found at low rain rates compared to at moderate–high rain rates. Such results indicate that the current GPM-era satellite precipitation products still have significant room for improving the estimation of light rainfall and falling snow events.

The hydrological simulations presented here show that the gauge-adjusted products are more promising than their unadjusted counterparts. The satellite-only products have unsatisfactory performance with large runoff errors. Even when the VIC model was recalibrated using precipitation-specific inputs, IMERG-UC and GSMaP-MVK still showed little potential for streamflow simulation over the upper Yellow River basin. This result implies that the VIC hydrologic model can tolerate relatively small precipitation errors but not unphysical bias through the hydrologic integration process. Thus, if a satellite precipitation estimate has good intrinsic data quality, then precipitation-specific calibration is recommended. Clearly, the performances of satellite precipitation estimates rely on gauge corrections, but purely satellite-derived estimates are the only available source of precipitation for areas where ground observations are nonexistent. At the current stage, the climatological calibration is proposed to make the purely satellite-derived product as consistent as possible with gauge-corrected products. In the future, more PMW/IR sources and more effective precipitation algorithms are needed for improving the retrieval ability of satellite precipitation.

We hope this study here will be helpful as a reference for precipitation dataset selection in the TP, and also provide the satellite precipitation community with a better understanding of GPM-era precipitation products. Considering the finer spatiotemporal resolution and improved precision, more studies exploring the potential of IMERG and GSMaP precipitation products in sub-daily or hourly time scales will be conducted in the near future.

Author Contributions: D.L. and B.Y. designed the study and conducted the experiments; D.L. wrote the draft of the manuscript; B.Y. supervised the research and revised the manuscript.

Funding: This work was financially supported by the National Key Research and Development Program of China (2018YFA0605402) and the Major Research Plan of the National Natural Science Foundation of China (91547101). Also, this work was partially sponsored by the National Science Foundation of Jiangsu Province (BK20161502).

Acknowledgments: The authors are grateful to the GPM science team at NASA and JAXA for making satellite precipitation data available, and thankful to the CMA for providing observation gauge data.

Conflicts of Interest: The authors declare no conflicts of interest.

References


36. Sharifi, E.; Steinacker, R.; Saghafian, B. Assessment of GPM-IMERG and other precipitation products against gauge data under different topographic and climatic conditions in Iran: Preliminary results. Remote Sens. 2016, 8, 135. [CrossRef]


42. Immerzeel, W.W.; Van Beek, L.P.H.; Bierkens, M.F.P. Climate change will affect the Asian water towers. Science 2010, 328, 1382–1385. [CrossRef]


70. Li, Z.; Yang, D.; Hong, Y. Multi-scale evaluation of high-resolution multi-sensor blended global precipitation products over the Yangtze River. *J. Hydrol.* 2013, 500, 157–169. [CrossRef]


