**Evaluation and Hydrologic Validation of Three Satellite-Based Precipitation Products in the Upper Catchment of the Red River Basin, China**

Yueyuan Zhang 1, Yungang Li 1,*, Xuan Ji 1, Xian Luo 1 and Xue Li 2

1 Yunnan Key Laboratory of International Rivers and Transboundary Eco-Security, Yunnan University, Kunming 650091, China; zhangyueyuan1@163.com (Y.Z.); jixuan@ynu.edu.cn (X.J.); luoxian@ynu.edu.cn (X.L.)

2 School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China; lixue33333z@126.com (X.L.)

* Correspondence: ygli@ynu.edu.cn; Tel.: +86-871-65034577

Received: 15 October 2018; Accepted: 21 November 2018; Published: 24 November 2018

**Abstract:** Satellite-based precipitation products (SPPs) provide alternative precipitation estimates that are especially useful for sparsely gauged and ungauged basins. However, high climate variability and extreme topography pose a challenge. In such regions, rigorous validation is necessary when using SPPs for hydrological applications. We evaluated the accuracy of three recent SPPs over the upper catchment of the Red River Basin, which is a mountain gorge region of southwest China that experiences a subtropical monsoon climate. The SPPs included the Tropical Rainfall Measuring Mission (TRMM) 3B42 V7 product, the Climate Prediction Center (CPC) Morphing Algorithm (CMORPH), the Bias-corrected product (CMORPH_CRT), and the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) Climate Data Record (PERSIANN_CDR) products. SPPs were compared with gauge rainfall from 1998 to 2010 at multiple temporal (daily, monthly) and spatial scales (grid, basin). The TRMM 3B42 product showed the best consistency with gauge observations, followed by CMORPH_CRT, and then PERSIANN_CDR. All three SPPs performed poorly when detecting the frequency of non-rain and light rain events (<1 mm); furthermore, they tended to overestimate moderate rainfall (1–25 mm) and underestimate heavy and hard rainfall (>25 mm). GR (Génie Rural) hydrological models were used to evaluate the utility of the three SPPs for daily and monthly streamflow simulation. Under Scenario I (gauge-calibrated parameters), CMORPH_CRT presented the best consistency with observed daily (Nash–Sutcliffe efficiency coefficient, or NSE = 0.73) and monthly (NSE = 0.82) streamflow. Under Scenario II (individual-calibrated parameters), SPP-driven simulations yielded satisfactory performances (NSE >0.63 for daily, NSE >0.79 for monthly); among them, TRMM 3B42 and CMORPH_CRT performed better than PERSIANN_CDR. SPP-forced simulations underestimated high flow (18.1–28.0%) and overestimated low flow (18.9–49.4%). TRMM 3B42 and CMORPH_CRT show potential for use in hydrological applications over poorly gauged and inaccessible transboundary river basins of Southwest China, particularly for monthly time intervals suitable for water resource management.

**Keywords:** TRMM 3B42 V7; CMORPH_CRT; PERSIANN_CDR; GR models; hydrological simulation; Red River Basin

---

1. **Introduction**

Precipitation is one of the most important water balance components of the global water cycle, and has great variability across different spatial and temporal scales [1,2]. The accurate observation
or estimation of precipitation has important theoretical and practical significance for flood warnings, drought monitoring, and water resource management [3,4]. Gauge observations provide relatively accurate point-based measurements of precipitation [5]; however, owing to significant precipitation heterogeneity across a variety of spatiotemporal scales, rain gauge observations only represent local conditions, and can result in potential errors when interpolated to larger scales, especially in mountainous areas with complex terrain [6]. Moreover, the spatial distribution of rain gauges is extremely uneven, with sparse gauges in remote areas, less developed regions, or areas with complex terrain [7]. Therefore, in situ gauge data usually cannot meet the requirements of applications that depend on high spatial–temporal resolution precipitation data (e.g., hydrological simulations [5]).

Satellite-based precipitation estimates have the advantage of adequate temporal resolution and fine spatial resolution with wide coverage, enabling accurate precipitation estimates in data-scarce or ungauged regions [5,8–10]. A number of satellite-based precipitation products (SPPs) are currently available, including the Global Precipitation Climatology Project (GPCP) [11], the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) [12], the Climate Prediction Center (CPC) morphing algorithm (CMORPH) [13], the Tropical Rainfall Measuring Mission (TRMM) Multi-satellites Precipitation Analysis (TMPA) [14], the Global Precipitation Measurement (GPM) [3], and the global gridded precipitation dataset Multi-Source Weighted-Ensemble Precipitation (MSWEP) [15]. With wide coverage and high spatial–temporal resolution (mostly finer than 0.25° × 0.25° at spatial and three-hour temporal scales), SPPs have been extensively applied in many fields, including hydrological simulation [16–18], extreme event analysis [19–21], and water resource management [22,23].

Since SPPs are based on an indirect approach that utilizes sensors, the results inevitably contain uncertainties caused by measurement errors, sampling, retrieval algorithms, and bias correction processes [17,24,25]. Furthermore, the error characteristics change depending on the climate region, season, altitude, and other factors [10,26]. In general, quantitative statistical and hydrological modelling evaluations are effective tools that are used to evaluate the precision of SPPs [4,17]. The former focuses on the comparison and evaluation of SPPs against gauge data or ground-based radar estimates. By this principle, temporal characteristics and spatial distributions of SPPs are not only investigated, but can also be quantitatively analyzed; however, the scale discrepancy problem remains when using rain gauge data for validation. The latter evaluates SPPs based on their predictive ability of streamflow rate in a hydrological modeling framework; precipitation products are evaluated at the watershed scale with respect to a specific application [27].

Over the past decades, numerous studies have improved our understanding of SPP performance at global and regional scales [8,28–31]. For example, TMPA products were validated in various parts of the world [32–36], and those results showed that TMPA products perform reasonably well over most regions. Following the successful TMPA, the Integrated Multisatellite Retrievals for GPM (IMERG), which incorporates observations from several satellites, offers improvements over the TMPA in quality and spatiotemporal resolution of precipitation data [3]. A range of studies comparing the TMPA and GPM products for the United States [37], Brazil [38], Africa [39], Iran [40], India [41], Pakistan [42], Malaysia [43], Singapore [44], and China [45] indicated that GPM is superior to TMPA products. In China, evaluation and validation using hydrologic simulations have been explored over many basins, including the Ganzi River basin [46], the upper Yangtze River and upper Yellow River basins over the Tibetan Plateau [47], the Huaihe River basin of eastern China [18], the Beijiang and Dongjiang River basins of southern China [5], the Luanhe River basin of northern China [4], the Tiaoxi watershed, which is part of the southern catchment of Taihu Lake in southeastern China [48], the Lancang River basin of southwest China [49], and the Huifa River basin of northeast China [50]. These studies all highlighted the great potential of different SPPs in hydrologic simulations, although SPPs have variable accuracy and distinct hydrological performance in different basins.

The Red River is an important transboundary river in Southeast Asia. Precipitation distribution is significantly uneven across the basin due to the complex terrain and subtropical monsoon climate [51].
About 85% of the annual total precipitation falls during the summer season [52]. Consequently, the Red River has an irregular flow regime. The high variability of river discharge in space and time leads to substantial challenges related to flooding and water stress, particularly in the Red River delta, which is a densely populated area of great importance to Vietnam for its agricultural productivity and economic activity [35]. The upstream region in China has a mean annual flow of 48.3 billion m$^3$, which contributes 37% of total flow of the Red River (131.4 billion m$^3$) [53]. The transboundary water resource is virtual for agriculture irrigation, hydropower, and ecosystem services. However, the rain gauge network has a low density (around 300 km$^2$ per rain gauge in Yunnan province, China), spatially uneven distribution, and is insufficient over mountainous areas. The scarcity and mismatch of the precipitation observations from upstream and downstream countries make it imperative to use SPPs in hydrological modeling, drought monitoring, and water resources management. Unfortunately, there is little work focusing on the evaluation of SPPs and their hydrological applicability over the Red River Basin in China.

This study aimed to assess the performance of three latest SPPs over the upper catchment of the Red River Basin for the time period 1998–2010. The SPPs included TRMM 3B42 V7, CMORPH_CRT (CMORPH Bias-corrected product), and PERSIANN_CDR (PERSIANN Climate Data Record). The main objectives were to: (1) statistically evaluate the quality of the three SPPs through comparison with rain gauge observations; and (2) comprehensively explore and compare the capability of these three SPPs in streamflow simulations using GR (Génie Rural) hydrological models at daily and monthly scales. This study will improve our understanding of the reliability of the three latest SPPs, and provide a reference for their applications in hydrological simulation and transboundary water resource management in the Red River Basin.

2. Materials and Methods

2.1. Study Area

The Red River originates in a mountainous area of Yunnan Province, China; it flows 1200 km to the southeast, and ends in the Gulf of Tonkin, in the South China Sea [54,55]. The Red River Basin drains an area of 156,451 km$^2$, of which 50.3% is in Vietnam, 48.8% is in China, and 0.9% is in Laos [52]. The upper catchment of the Red River Basin (URRB) refers to the catchment north of the China–Vietnam border (Figure 1). The catchment covers an approximate area of 33,614 km$^2$. The elevation of the catchment ranges from 76 m to 3123 m above sea level, and decreases from the northwest to the southeast [51]. It is characterized by a subtropical monsoon climate [52], with annual average temperatures of 14.8–23.8°C. The annual average precipitation over the URRB from 1998 to 2010 was about 1044 mm, ranging from 772 mm to 1276 mm; approximately 85% of the annual precipitation is concentrated in the rainy season (May to October) [52]. The climate confers the typical hydrologic regime characterized by large runoff during the summer and low runoff during the winter [56]. The annual average discharge at Manhao Station for the period 1998–2010 was 282 m$^3$/s.

2.2. Datasets

Daily observed discharge data for the period 1998–2010 at the Manhao hydrological station was obtained from the Hydrological Year Book and the Hydrological Bureau of Yunnan Province (HBYP). Daily precipitation data from 25 rain gauges were provided by the Meteorological Agency of Yunnan Province (MAYP) (Figure 1). It is noteworthy that these rain gauges are independent from the Global Precipitation Climatology Centre (GPCC) gridded gauge-analysis precipitation dataset. Daily meteorological observations at 21 stations for the same period were collected from the China Meteorological Administration (CMA), including mean air temperature, mean relative humidity, mean wind speed at 10-m height, and hours of bright sunshine. The quality of the above datasets has been checked by the HBYP, MAYP, and CMA. We also performed routine quality assessment including statistical tests, visual data plots, and histograms, to ensure that there were no missing or erroneous records. The descriptive statistics of precipitation and discharge during 1998–2010 are provided in the
supplementary material (Table S1 and Table S2). The daily potential evapotranspiration for each station was estimated using the Penman–Monteith equation, as recommended by the Food and Agriculture Organization of the United Nations (FAO) [57]. The FAO Penman–Monteith method is provided in the supplementary material (S1). Areal average rainfall and potential evapotranspiration over the catchment were produced by using the Thiessen polygon approach [58].

![Map of the Red River Basin](image)

**Figure 1.** Location of the upper catchment of the Red River Basin (URRB) and distribution of meteorological and hydrological stations.

Three SPPs (TRMM 3B42 V7, CMORPH_CRT, and PERSIANN_CDR) were considered. The TRMM 3B42 V7 product is one of the TMPA products, which were designed based on a wide variety of satellite datasets and are supplied by the National Aeronautics and Space Administration (NASA) [14]. This product provides precipitation at a spatial resolution of 0.25° and a three-hour temporal resolution; it has quasi-global coverage of 50°N–50°S from 1998 to the present, combining information from calibrated passive microwave (PMW) and infrared (IR) data. The 3B42 V7 product was adjusted using monthly rain gauge precipitation data from the GPCP [59]. Here, the 3B42 V7 with a daily temporal resolution and a 0.25° spatial resolution from 1998 to 2010 was employed. Daily precipitation was obtained by the accumulation of three-hour precipitation data.

NOAA's (National Oceanic and Atmospheric Administration) CPC CMORPH contains global satellite-based precipitation generated by integrated PMW and IR data [13]. The latest CMORPH V1.0 product includes a raw satellite-only precipitation product (CMORPH_RAW), CMORPH_CRT, and a satellite-gauge blended product (CMORPH_BLD), covering 60°S–60°N and 180°W–180°E. The CMORPH_CRT product is generated by adjusting the CMORPH_RAW product against the CPC unified daily gauge analysis over land, and the pentad GPCP over ocean using the probability density function (PDF) matching bias correction method [60]. Three combinations of spatial–temporal resolutions are available: eight km and 30 min, 0.25° and 3 h, and 0.25° and daily. Here, the CMORPH_CRT product with 0.25° and daily spatial–temporal resolution for the period 1998–2010 was used.

The original PERSIANN is one of the popular global precipitation estimations for estimating historical precipitation from March 2000 to present; it was developed by combining PMW observations and IR data. The latest PERSIANN_CDR product used the archive of the GridSat-B1 IR data [61] as input to the trained PERSIANN model; then, the biases in the PERSIANN estimated precipitation were adjusted using GPCP 2.5° monthly data version 2.2 [12,27]. Since no PMW is used in the PERSIANN_CDR product, the PERSIANN model parameters were pretrained using National Centers for Environmental Prediction (NCEP) stage-IV hourly precipitation data. Currently, this version of PERSIANN_CDR provides daily precipitation estimates at a spatial resolution of 0.25° for quasi-global
coverage (60°N–60°S) from 1983 to the present. In this study, a subset of data for the period 1998–2010 was employed.

2.3. Methods

2.3.1. Evaluation Indices

Comparisons between the three SPPs and rain gauge data were performed both on grid and basin scales. For the grid scale, the SPPs precipitation at the grid boxes with rain gauge stations are extracted and compared with the corresponding rain gauge precipitation. For the basin scale, spatially averaged pixel values of the SPPs precipitation were compared with the areal-averaged precipitation from rain gauge stations using the Thiessen polygon approach [58].

Several widely used statistical indices were adopted to quantify the performance of the three SPPs against rain gauge observations, including Spearman’s Rank correlation coefficient (CC) [62], root mean squared error (RMSE), mean absolute error (MAE), and relative bias (Bias). In addition, the probability of detection (POD), frequency of hit (FOH), false alarm ratio (FAR), critical success index (CSI), and the Heidke skill score (HSS) indices were calculated to evaluate the precipitation detection ability of the three SPPs [63]. POD provides the fraction of precipitation events that the satellite products detect among all the actual precipitation events; FOH measures how often the satellite products detect rainfall when there is actually rainfall; FAR measures the fraction of unreal events among all the events that the satellite products detected; CSI represents the overall fraction of precipitation events correctly detected by the satellite products; and HSS measures the accuracy of the estimates accounting for matches due to random chance. Furthermore, CC, Bias, and the Nash–Sutcliffe efficiency coefficient (NSE) [64] were employed to evaluate the performance of the hydrological model in streamflow simulations. NSE describes the prediction skill of the simulated streamflow as compared to the observed. The formulas to calculate the indices mentioned above are listed in Table 1.

Table 1. Indices used to evaluate the performance of satellite precipitation estimates. CC: correlation coefficient, RMSE: root mean squared error, MAE: mean absolute error, Bias: relative bias, POD: probability of detection, FOH: frequency of hit, FAR: false alarm ratio, CSI: critical success index, HSS: Heidke skill score, NSE: Nash–Sutcliffe efficiency coefficient.

<table>
<thead>
<tr>
<th>Statistical Metric</th>
<th>Unit</th>
<th>Equation</th>
<th>Range</th>
<th>Perfect Value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>-</td>
<td>[ CC = 1 - \frac{\sum_{i=1}^{n}(X(i) - Y(i))^2}{\sum_{i=1}^{n}(X(i) + Y(i))^2} ]</td>
<td>-1 to 1</td>
<td>1</td>
<td>[62]</td>
</tr>
<tr>
<td>RMSE</td>
<td>mm</td>
<td>[ RMSE = \sqrt{\frac{\sum_{i=1}^{n}(S(i) - G(i))^2}{\sum_{i=1}^{n}G(i)^2}} ]</td>
<td>0 to ( \infty )</td>
<td>0</td>
<td>[16]</td>
</tr>
<tr>
<td>MAE</td>
<td>mm</td>
<td>[ MAE = \frac{\sum_{i=1}^{n}</td>
<td>S(i) - G(i)</td>
<td>}{n} ]</td>
<td>0 to ( \infty )</td>
</tr>
<tr>
<td>Bias</td>
<td>%</td>
<td>[ Bias = \frac{\sum_{i=1}^{n}</td>
<td>S(i) - G(i)</td>
<td>}{\sum_{i=1}^{n}G(i)} \times 100% ]</td>
<td>-( \infty ) to ( \infty )</td>
</tr>
<tr>
<td>POD</td>
<td>-</td>
<td>[ POD = \frac{N_{11}}{N_{11} + N_{00}} ]</td>
<td>0 to 1</td>
<td>1</td>
<td>[63]</td>
</tr>
<tr>
<td>FOH</td>
<td>-</td>
<td>[ FOH = \frac{N_{11}}{N_{11} + N_{01}} ]</td>
<td>0 to 1</td>
<td>1</td>
<td>[63]</td>
</tr>
<tr>
<td>FAR</td>
<td>-</td>
<td>[ FAR = \frac{N_{10}}{N_{11} + N_{10}} ]</td>
<td>0 to 1</td>
<td>0</td>
<td>[63]</td>
</tr>
<tr>
<td>CSI</td>
<td>-</td>
<td>[ CSI = \frac{N_{11}}{N_{11} + N_{01} + N_{10}} ]</td>
<td>0 to 1</td>
<td>1</td>
<td>[63]</td>
</tr>
<tr>
<td>HSS</td>
<td>-</td>
<td>[ HSS = \frac{N_{11}}{2N_{11} + N_{10} + N_{01} + N_{00} - N_{10} - N_{01}} ]</td>
<td>0 to 1</td>
<td>1</td>
<td>[63]</td>
</tr>
<tr>
<td>NSE</td>
<td>-</td>
<td>[ NSE = 1 - \frac{\sum_{i=1}^{n}(Q_{oi} - Q_{si})^2}{\sum_{i=1}^{n}(Q_{oi} - Q_{g})^2} ]</td>
<td>-( \infty ) to 1</td>
<td>1</td>
<td>[64]</td>
</tr>
</tbody>
</table>

Notation: \( n \), number of samples; \( S \), satellite precipitation; \( G \), gauged observation; \( X(i) = \text{rank}[S(i)]; Y(i) = \text{rank}[G(i)]; Q_{si}, \) simulated streamflow; \( Q_{oi}, \) observed streamflow; \( Q_{g}, \) observed mean streamflow. \( N_{11} \), Satellite is > 0 and gauge is > 0; \( N_{00} \), Satellite is > 0 and gauge equals 0; \( N_{01} \), Satellite equals 0 and gauge is > 0; \( N_{10} \), Satellite equals 0 and gauge equals 0.
2.3.2. GR Hydrological Models

The GR hydrological models were developed by Irstea, which is a national applied research institute of France; they are lumped rain–runoff models that can be applied at various time steps, ranging from hourly to annual [65, 66]. In this study, only the daily (GR4J) and monthly (GR2M) models were employed.

(1) GR4J model

The GR4J model was originally developed and tested on 429 different catchments in France, the United States of America (USA), Brazil, and the Côte d’Ivoire [67]. GR4J simulates runoff via two functions. First, a production function accounts for precipitation (P) and potential evapotranspiration (PET), and determines the effective precipitation that contributes to flow and supplies the production reservoir. Second, a routing function calculates runoff at the catchment outlet. The quantity of water feeding the routing part of the model comprises the percolation that is added to the remaining fraction of water. This flow is divided into two components according to a fixed split: 90% is routed by a unit hydrograph, UH1, and then a non-linear routing store, and the remaining 10% are routed by a single unit hydrograph, UH2. The purpose of the unit hydrographs is to account for differences in runoff delays between the two conceptual reservoirs. GR4J requires the calibration of four free parameters (Table 2). The median values and approximate 80% confidence intervals for the four parameters are provided in Table 2, which were obtained using a large variety of catchments [68].

Table 2. Median values and approximate 80% confidence intervals of four model parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Median Value</th>
<th>80% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>x1: maximum capacity of the production store (mm)</td>
<td>350</td>
<td>100–1200</td>
</tr>
<tr>
<td>x2: groundwater exchange coefficient (mm)</td>
<td>0</td>
<td>−5 to 3</td>
</tr>
<tr>
<td>x3: maximum capacity of the routing store (mm)</td>
<td>90</td>
<td>20–300</td>
</tr>
<tr>
<td>x4: time base of unit hydrograph UH1 (days)</td>
<td>1.7</td>
<td>1.1–2.9</td>
</tr>
</tbody>
</table>

(2) GR2M Model

The GR2M has two parameters and has been shown to have the best performance among several similar models when using a benchmark test consisting of 410 basins around the world [69, 70]. The production function of the GR2M model has strong similarities with the daily version, but uses a simplified routing scheme [66]. This model is characterized by two functions: (1) a function of production that revolves around a reservoir ground of a maximum capacity (x1), which is the first parameter to be wedged (transferring a percolation of reservoir ground is ensured by the dependent feature of the stock status ‘S’), and (2) a transfer function represented by a quadratic draining reservoir with a capacity fixed at 60 mm. This reservoir is modified by an underground exchange, whose coefficient (x2) is the second parameter to optimize [71].

(3) Hydrological Simulation Scenarios

In this study, both models were warmed up for one year (1998), and then split into calibration (1999–2004) and validation (2005–2010) periods. Model calibration was achieved with the aid of the default algorithm provided in the airGR package developed at Irstea [66]. Two parameterization scenarios were conducted to evaluate the effect of precipitation uncertainty on runoff simulation results. In Scenario I, model parameters were first calibrated using gauged rainfall data in the calibration period, and then the model was rerun using gauged rainfall data and the three SPPs in both the calibration and validation periods. Scenario I was mainly used to evaluate the streamflow simulation utility of the different SPPs using gauge-calibrated parameters [72]. In Scenario II, model parameters were recalibrated with individual satellite rainfall data during the calibration period; then, streamflow was simulated for both the calibration and validation periods using the three individual satellite-based parameters. Scenario II was adopted to determine whether the evaluated SPPs have the potential to be
alternative data sources for hydrological simulations in data-poor or ungauged basins [9,10]. Table 3 shows calibrated parameter values in the GR4J model for different precipitation data inputs.

Table 3. Calibrated parameter values in the GR4J model for different precipitation data inputs.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Rain Gauge</th>
<th>TRMM 3B42</th>
<th>CMORPH_CRT</th>
<th>PERSIANN_CDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>x1</td>
<td>1200</td>
<td>1737</td>
<td>1667</td>
<td>1998</td>
</tr>
<tr>
<td>x2</td>
<td>0.77</td>
<td>0.35</td>
<td>0.92</td>
<td>1.09</td>
</tr>
<tr>
<td>x3</td>
<td>25</td>
<td>37</td>
<td>23</td>
<td>35</td>
</tr>
<tr>
<td>x4</td>
<td>2.25</td>
<td>2.45</td>
<td>3.12</td>
<td>3.42</td>
</tr>
</tbody>
</table>

3. Results

3.1. Comparison of Rain Gauge and Satellite-Based Precipitation Data

3.1.1. Evaluation at the Grid and Basin Scales

Figures 2 and 3 present scatterplots of daily and monthly precipitation, respectively, from TRMM 3B42, CMORPH_CRT, and PERSIANN_CDR against rain gauge data at the grid scale. Intuitively, Figure 2 shows that the TRMM 3B42 product performed the best among the three SPPs for all metrics at the daily scale. Throughout the year, TRMM 3B42 had a much larger CC, but a smaller RMSE, MAE, and Bias than that of the other two products, with the CC = 0.62, RMSE = 7.7, MAE = 2.9, and Bias = 1.6%. Meanwhile, PERSIANN_CDR performed the worst, with the exception of the Bias (Bias = −5.4%), for which it was marginally better than that of CMORPH_CRT (Bias = −8.0%). The performances of the three products were also compared seasonally. According to the climate of the Red River Basin, May to October was considered to be the wet season, and November to April was considered to be the dry season [73]. For both the dry and wet seasons, TRMM 3B42 still showed a better performance than the other two products. All three SPPs achieved better estimations during the wet season than during the dry season.

On a monthly scale (Figure 3), the TRMM 3B42 product showed the best performance, regardless of season. Moreover, all three SPPs performed better for the wet season than for the dry season. In terms of Bias, TRMM 3B42, CMORPH_CRT, and PERSIANN_CDR products all showed underestimations during the dry season; such an underestimation in SPPs during the dry season has been reported in many studies [7,74,75]. In addition, the monthly results show much higher CC values with the observations than do the daily comparisons. This is because errors in the daily values are offset to some extent when added to the monthly values [4].

At the basin scale (Table 4), the daily and monthly CC values of the three SPPs improved, and the MAEs significantly decreased, respectively. Meanwhile, except for TRMM 3B42, the Bias values also showed a corresponding decreasing trend. From Table 4, the relative performances of the SPPs are similar to those at the grid scale. A distinct exception is that the TRMM 3B42 showed the largest overestimation in precipitation. Overall, all three SPPs had comparable and good performances for the monthly precipitation estimations.
Figure 2. Scatterplots of daily precipitation from TRMM 3B42, CMORPH_CRT, and PERSIANN_CDR against ground observations at the grid scale: the three panels show the results from the whole year (upper panel), dry season (mid panel), and wet season (lower panel). The red line indicates a 1:1 correspondence.

Figure 3. Scatterplots of monthly precipitation from TRMM 3B42, CMORPH_CRT, and PERSIANN_CDR against ground observations at the grid scale: the three panels show the results from the whole year (upper panel), dry season (mid panel), and wet season (lower panel). The red line indicates a 1:1 correspondence.
Table 4. Statistical indices between three satellite-based precipitation products (SPPs) and rain gauge data at the basin scale.

<table>
<thead>
<tr>
<th>Basin Scale</th>
<th>Daily CC</th>
<th>Daily RMSE</th>
<th>Daily MAE</th>
<th>Monthly CC</th>
<th>Monthly RMSE</th>
<th>Monthly MAE</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole year</td>
<td>0.82</td>
<td>11.3</td>
<td>1.7</td>
<td>0.99</td>
<td>17.1</td>
<td>12.0</td>
<td>9.6%</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>23.4</td>
<td>2.3</td>
<td>0.98</td>
<td>14.2</td>
<td>10.2</td>
<td>−0.9%</td>
</tr>
<tr>
<td></td>
<td>0.66</td>
<td>34.1</td>
<td>2.9</td>
<td>0.97</td>
<td>21.1</td>
<td>15.1</td>
<td>2.2%</td>
</tr>
<tr>
<td></td>
<td>0.57</td>
<td>2.7</td>
<td>0.7</td>
<td>0.96</td>
<td>6.2</td>
<td>4.6</td>
<td>4.8%</td>
</tr>
<tr>
<td>Dry season</td>
<td>0.49</td>
<td>4.7</td>
<td>0.8</td>
<td>0.90</td>
<td>10.3</td>
<td>7.0</td>
<td>−16.6%</td>
</tr>
<tr>
<td></td>
<td>0.42</td>
<td>6.8</td>
<td>1.1</td>
<td>0.83</td>
<td>10.6</td>
<td>8.0</td>
<td>−6.2%</td>
</tr>
<tr>
<td></td>
<td>0.81</td>
<td>19.8</td>
<td>2.7</td>
<td>0.96</td>
<td>23.4</td>
<td>19.3</td>
<td>10.4%</td>
</tr>
<tr>
<td>Wet season</td>
<td>0.66</td>
<td>41.8</td>
<td>3.8</td>
<td>0.97</td>
<td>17.3</td>
<td>13.4</td>
<td>2.0%</td>
</tr>
<tr>
<td></td>
<td>0.52</td>
<td>60.9</td>
<td>4.6</td>
<td>0.93</td>
<td>27.9</td>
<td>22.2</td>
<td>3.8%</td>
</tr>
</tbody>
</table>

3.1.2. Evaluation of Contingency

Figure 4 shows box plots of rainfall-detecting skill scores, including POD, FOH, FAR, CSI, and HSS, which were used to measure the contingency of the three SPPs. Among the three SPPs, TRMM 3B42 performed the best, having the highest FOH, CSI, and HSS values, and the lowest FAR value; PERSIANN_CDR performed the worst in terms of all five rainfall-detecting skill scores. These results illustrate that TRMM 3B42 yields the highest frequency of successful hits when rainfall really occurs, and the lowest erroneous detection rate when there is actually no rainfall. For the POD, TRMM 3B42 had a median value of 0.65, while that of CMORPH_CRT was 0.72; this implies that CMORPH_CRT is more likely to detect a larger fraction of precipitation events among all of the actual precipitation events. Note that the POD, FOH, FAR, CSI, and HSS of CMORPH_CRT tended to have slightly larger variations than they did for the other two products, indicating the considerable data instability of CMORPH_CRT. This instability can be partially attributed to the morphing processes in the CMORPH_RAW data, which determines the precipitation values as a weighted mean of PMW estimates from multiple sensors [74].

![Figure 4](image-url)
3.1.3. Evaluation of Rainfall Intensity Distribution

Figure 5 displays the occurrence frequency distribution of daily precipitation for different rain intensity classes and their relative contributions to the amount of total rainfall. Non-rainy days have the largest occurrence frequency, accounting for 58–67% of the total days. TRMM 3B42 identified significantly more non-rainy days (67%) than did the gauged data (63%); this is because the TRMM data have less skill in correctly specifying moderate and light rain rates on short time intervals [48,50,76] identified less non-rainy days (58%), but more light rain (0–1 mm) days. PERSIANN_CDR identified a similar number of no-rain events (64%) as the observation; however, it deviated the most for the light rain event. In addition, all three SPPs overestimated the moderate rain intensity class (1–25 mm) and underestimated heavy rainfall (25–50 mm) and hard rainfall events (>50 mm), with the exception of the TRMM 3B42 product, which slightly overestimated the heavy rainfall event.

![Figure 5](image-url)  
**Figure 5.** Occurrence frequency distribution (bars) of daily precipitation for different rain intensity classes and their relative contributions (lines) to total rainfall amount during 1998–2010.

For the distribution of relative contribution, PERSIANN_CDR showed the largest discrepancy, while the other two SPPs were very similar to the gauged data. It was found that although the frequency of light rainfall event was 4–15% of the total days, its contribution to the total rainfall amount was only 1.3% on average. On the contrary, although high rainfall occurred on only a small percentage of the total days, they had a significant contribution to the total rainfall amount. For example, the hard rainfall event had a frequency of just about 0.5%, but its contribution to the total rainfall amount was 10.7% on average. For the rain intensity of 10–25 mm, there was a similarly small percentage (6.5%) of total days for all four datasets, but this class contributed the most (35.2%) to the total rainfall amount. The different volume contribution performances greatly impacted the hydrologic simulations, since most of the hydrological processes within the models are sensitive to the total precipitation amount and rainfall intensity distribution [18,77].

3.2. Hydrologic Validation Using GR Models

3.2.1. Daily and Monthly Streamflow Simulations under Scenario I

Table 5 shows the model performance results simulated using different precipitation inputs with gauge-calibrated parameters under Scenario I. Generally, the streamflow simulated by rain gauge data
agreed well with the observed hydrographs on both daily ($NSE = 0.82$, $CC = 0.92$, $Bias = 0.6\%$) and monthly scales ($NSE = 0.87$, $CC = 0.96$, $Bias = 1.8\%$). The GR4J and GR2M models were found to be robust and provided a sound basis for testing the applicability of the three SPPs.

Table 5. Comparison of daily and monthly observed and simulated streamflow under Scenario I.

<table>
<thead>
<tr>
<th>Precipitation Product</th>
<th>Daily</th>
<th>Monthly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NSE</td>
<td>CC</td>
</tr>
<tr>
<td>Rain gauge</td>
<td>0.82</td>
<td>0.92</td>
</tr>
<tr>
<td>TRMM 3B42</td>
<td>0.62</td>
<td>0.93</td>
</tr>
<tr>
<td>CMORPH_CRT</td>
<td>0.73</td>
<td>0.92</td>
</tr>
<tr>
<td>PERSIANN_CDR</td>
<td>0.53</td>
<td>0.88</td>
</tr>
</tbody>
</table>

On a daily scale (Table 5 and Figure 6), CMORPH_CRT showed the best performance for streamflow simulations over the entire period, with relatively desirable $NSE$ (0.73) and $CC$ (0.92), and a high but acceptable $Bias$ of $−7.5\%$; PERSIANN_CDR had the lowest $NSE$ (0.53) and $CC$ (0.88), and a $Bias$ of $−2.9\%$. The TRMM 3B42 had better $NSE$ (0.62) and $CC$ (0.93) than did PERSIANN_CDR; however, a significant overestimation (24.2%) was found for the TRMM forced simulation. It is possible that the $Bias$ (9.6%) of the daily TRMM 3B42 precipitation was magnified by the hydrological model, causing this large overestimation [17]. Moreover, all the three SPP-based streamflow simulations underestimated some high peak flows (e.g., 2001 and 2007) of the rainy season. This phenomenon can mainly be attributed to the precipitation estimate uncertainty of SPPs at the daily scale during heavy and hard rainfall events [72].

Figure 6. Simulated (red) and observed (black) daily streamflow under Scenario I from: (a) rain gauge data; (b) TRMM 3B42; (c) CMORPH_CRT; and (d) PERSIANN_CDR. Blue lines show precipitation data.

Figure 7 illustrates flow duration curves (FDCs) for the four simulations on the daily scale from 1999 to 2010 under Scenario I. The rain gauge, CMORPH_CRT, and PERSIANN_CDR simulations were all consistent with the observations. However, the FDC produced by the TRMM 3B42 simulation was apparently higher than that of the observations, which is consistent with the large $Bias$ (24.2%; Table 5).
Remote Sens. 2018, 10, x FOR PEER REVIEW 12 of 22

TRMM 3B42 performed the worst (NSE = 0.72, Bias = 24.2%) on the monthly streamflow simulations under Scenario I, which is consistent with former studies [18]. The CMORPH_CRT simulation had the best performance, with the highest NSE values (0.77, 0.92) and the lowest Bias (−3.1, 0.8%). PERSIANN_CDR performed the second best (NSE = 0.76, Bias = 5.5%), while TRMM 3B42 performed the worst (NSE = 0.72, Bias = 24.2%) on the monthly scale. All the three SPPs’ forced simulations had good agreements with the observed data, with CC values above 0.94.

The monthly streamflow simulations were generally found to be more accurate than the daily simulations in terms of NSE and CC values (Table 5 and Figure 8). Similar to the daily results, the CMORPH_CRT forced simulation had the best performance, with the highest NSE (0.82) and the lowest Bias (−0.9%). PERSIANN_CDR performed the second best (NSE = 0.76, Bias = 5.5%), while TRMM 3B42 performed the worst (NSE = 0.72, Bias = 24.2%) on the monthly scale. All the three SPPs’ forced simulations had good agreements with the observed data, with CC values above 0.94. As shown in Figure 8, the monthly streamflow hydrographs from the three SPPs reasonably well matched the observed streamflow, predicting peak flow and low flow conditions perfectly. These results indicated that all three SPPs are suitable for monthly streamflow simulation purposes in the study area. In general, considering the results for the daily and monthly scales, the CMORPH_CRT product is most suitable for streamflow simulations using gauge-calibrated parameters over the URRB.

Figure 7. Daily discharge flow duration curves (FDCs) for observations, rain gauge simulation, TRMM 3B42 simulation, CMORPH_CRT simulation, and PERSIANN_CDR simulation under Scenario I.

Figure 8. Simulated (red) and observed (black) monthly streamflow under Scenario I from: (a) rain gauge data; (b) TRMM 3B42; (c) CMORPH_CRT; and (d) PERSIANN_CDR. Blue bars show precipitation data.
3.2.2. Daily and Monthly Streamflow Simulations under Scenario II

Table 6 shows the model performance results for the three SPPs when using their own optimal parameters calibrated under Scenario II. The simulation results under Scenario II were clearly improved when compared with those under Scenario I, which is consistent with former studies [5,10,18,78]. For the daily scale (Table 6), the NSE values of streamflow simulated by the TRMM 3B42, CMORPH_CRT, and PERSIANN_CDR products increased to 0.76, 0.77, 0.63, and the CC kept its high values of 0.92, 0.92, and 0.88, respectively. Furthermore, the Bias values were also significantly reduced, except for the PERSIANN_CDR forced simulation. Even so, the underestimation problem for peak flow (e.g., 2001 and 2007) remained after the recalibration (Figure 9). Figure 10 shows the FDCs for the three SPP-driven simulations on a daily scale from 1999 to 2010 under Scenario II. The three individual simulations all agreed well with the FDC produced by observations.

Table 6. Comparison of daily and monthly observed and simulated streamflow under Scenario II.

<table>
<thead>
<tr>
<th>Precipitation Products</th>
<th>Daily</th>
<th>Monthly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NSE</td>
<td>CC</td>
</tr>
<tr>
<td>TRMM 3B42</td>
<td>0.76</td>
<td>0.92</td>
</tr>
<tr>
<td>CMORPH_CRT</td>
<td>0.77</td>
<td>0.92</td>
</tr>
<tr>
<td>PERSIANN_CDR</td>
<td>0.63</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Figure 9. Simulated (red) and observed (black) daily streamflow under Scenario II from: (a) TRMM 3B42; (b) CMORPH_CRT; and (c) PERSIANN_CDR. Blue lines show precipitation data.

For the monthly scale (Table 6 and Figure 11), the NSE values for the three SPP-forced simulations also increased (0.86 for TRMM 3B42, 0.83 for CMORPH_CRT, and 0.79 for PERSIANN_CDR). Furthermore, the high CC values (>0.94) also indicated the good agreements between the three simulations and the observed data. However, only the TRMM 3B42 forced simulation saw a significant reduction in terms of Bias (from 24.2% to 0.8%). Additionally, it was found that the TRMM 3B42 forced simulation had the greatest improvements in NSE and Bias values from Scenario I to Scenario II for both the daily and monthly scales. Overall, the three SPP forced daily and monthly simulations all exhibited good performance, matching well with observations in this study area. The better agreement was achieved using TRMM 3B42 and CMORPH_CRT products, making them more suitable
for performing hydrologic simulations with inadequate surface precipitation observations (e.g., in ungauged catchments).

![Streamflow Graph](image)

**Figure 10.** Daily discharge flow duration curves (FDCs) for observations, TRMM 3B42 simulation, CMORPH_CRT simulation, and PERSIANN_CDR simulation under Scenario II.

![Flow Exceedance Percentile Graph](image)

**Figure 11.** Simulated (red) and observed (black) monthly streamflow under Scenario II from: (a) TRMM 3B42; (b) CMORPH_CRT; and (c) PERSIANN_CDR. Blue bars show precipitation data.

### 3.2.3. Capability of Simulating Extreme Events

To further investigate the capability of the three SPPs to simulate extreme events, we defined the observed daily streamflow, exceeding its 90% quantile as high flow, and that less than its 50% quantile as low flow [79]. The observed high-flow and low-flow data were compared with the corresponding simulated flow with various precipitation inputs. In this section, the three simulations under Scenario II, as well as the rain gauge simulation under Scenario I, were evaluated. The model performance results of high-flow and low-flow simulations were presented in Table 7. For the high-flow simulations, the gauged-based simulation generally had the best performance with desirable NSE (0.50), CC (0.67), and Bias (−11.6%) values. TRMM 3B42 and CMORPH_CRT products exhibited comparable performance of similar NSE (0.31 and 0.36), CC (0.65 and 0.68), and Bias values (−19.1% and −18.1%), whereas PERSIANN_CDR performed poorly with the NSE <0 and the largest underestimation (28.0%). For the low-flow simulations, the model performances of the four precipitation inputs were
unacceptable, with the NSE values all below 0, indicating their poor value for low-flow simulation in this region. Additionally, there was an underestimation of the four precipitation inputs for the high-flow simulations, and an overestimation for the low-flow simulations, which is consistent with many other studies [5,10,78]. This is largely due to the underestimation or overestimation of different precipitation products during extreme rainfall events. As discussed, all three SPPs have poor modeling capabilities for extreme hydrologic events over the URRB.

Table 7. Comparison of daily high-flow and low-flow simulations using four precipitation inputs against observations.

<table>
<thead>
<tr>
<th>Precipitation Product</th>
<th>High flow</th>
<th>Low flow</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NSE</td>
<td>CC</td>
</tr>
<tr>
<td>Rain gauge</td>
<td>0.50</td>
<td>0.67</td>
</tr>
<tr>
<td>TRMM 3B42</td>
<td>0.31</td>
<td>0.65</td>
</tr>
<tr>
<td>CMORPH_CRT</td>
<td>0.36</td>
<td>0.68</td>
</tr>
<tr>
<td>PERSIANN_CDR</td>
<td>−0.06</td>
<td>0.52</td>
</tr>
</tbody>
</table>

4. Discussion

SPPs have difficulty representing precipitation in mountainous regions, where precipitation is controlled by the orography and characterized by high spatiotemporal variability [8,80]. PMW data measured from low-orbit satellites and IR data from geostationary satellites are the main data sources for SPPs [17,81]. Specifically, PMW provides direct and accurate precipitation estimates at the cost of coarse temporal resolution, while IR provides useful information mainly about storm clouds based on the low temperatures of the top of these clouds: these data has fine temporal resolutions, but less accuracy [10,60]. In the mountainous regions, IR retrievals generally fail to capture light precipitation events and underestimate orographic rains, whereas PMW retrievals face challenges detecting orographic precipitation, especially in the cold season [82]. Therefore, all SPPs suffer from systematical overestimations or underestimations over the URRB, which is explained by the negative and positive values of Bias (Figure 2, Figure 3, and Table 4). In addition, SPPs perform better in the wet season than in the dry season.

According to our results of accuracy analyses, PERSIANN_CDR performs unsatisfactorily compared with the other two products (Figures 2 and 3), which is probably because PERSIANN_CDR is mainly based on IR data, while TRMM 3B42 and CMORPH_CRT combine PMW and IR data. The main difference between TRMM 3B42 and CMORPH_CRT is the gauge adjustment algorithm that is adopted. To develop CMORPH, PDF matching against the CPC unified daily gauge analysis was used to adjust the biases. Monthly GPCC rain gauge analyses and inverse-error variance weighting were used for the TRMM to adjust the biases [60]. However, the accuracy of TRMM generally outperforms that of CMORPH (Figures 2 and 3), which is probably due to the monthly gauge adjustment algorithm that is used in TRMM being superior to the PDF matching adopted in CMOPRH over this region. A similar finding was obtained in the previous studies over the Huaihe River basin in China [18]. At the basin scale, as expected, the CC and MAE values for SPPs improved, and the relative performances of the SPPs are similar to those at the grid scale (Table 4). Some distinct exceptions are that the TRMM 3B42 showed the largest overestimation in precipitation, while the PERSIANN_CDR showed underestimation in precipitation. This result indicated that the performance of these SPPs are not uniform with an increasing spatial scale, as a consequence of topographical variations [83]. Mei et al. [84] also stated that the size of catchments influences the satellite rainfall errors. Another possible reason is that there may not be enough rain gauges in the area to provide accurate estimates of rainfall for comparisons with satellite estimates at the basin scale.

Two parameters scenarios were designed to evaluate the effect of precipitation uncertainty on streamflow simulations. Under Scenario I (Table 5), CMORPH_CRT showed the best performance among the SPPs, which is mostly due to the relatively low Bias in precipitation input. In contrast,
the TRMM 3B42-driven simulations overestimate the streamflow by 24.2% (Table 5), which is attributed to the large overestimation in the TRMM 3B42 precipitation input. Another possible explanation is that CMORPH_CRT used daily rain gauge analysis, and all of the bias correction algorithms were conducted at the daily scale directly, while the other two products conducted bias correction with rain gauge analysis at the monthly scale. Therefore, CMORPH_CRT performs better in streamflow simulation using the gauge-calibrated parameters. Previous studies have reported that the errors of SPPs are propagated into hydrological simulations [10,17,79]. An overestimation/underestimation of precipitation estimates can be transformed into a larger overestimation/underestimation in the simulated streamflow. In this study, the basin-scale Bias of TRMM 3B42 (9.6%), CMORPH_CRT (−0.9%), and PERSIANN_CDR (2.2%) resulted in corresponding Bias values for streamflow simulations of 24.2%, −7.5%, and −2.9% at the daily scale, and 24.2%, −0.9%, and 5.5% at the monthly scale, respectively. This comparison between the input and output Bias values in GR models indicates that there is a non-linear error propagation pattern [16].

Under Scenario II, a recalibrated hydrological model using the SPPs can greatly improve streamflow simulation performance, as the different parameter settings can compensate for errors in the satellite rainfall data (Table 6). However, a few studies have indicated that recalibration can sometimes cause parameter values to exceed their reasonable ranges [16,79]. The calibrated model parameters for rain gauge are all within the 80% confidence interval (Tables 2 and 3). However, some model parameters (i.e., x1 and x4) for SPP-forced calibration greatly exceed the 80% confidence interval, which may be attributed to the GR models being sensitive to the precipitation data and the size of catchments. However, for a specific catchment, since the underlying surface condition remains unchanged, the hydrological model parameters largely depend on the input data. If the forcing data change, the sensitive parameters will change accordingly in order to match the streamflow (Table 3). In spite of the influence of cancellation between parameter differences and precipitation bias on streamflow simulation, the three SPPs are able to produce a reasonably good streamflow under scenario II. For instance, TRMM 3B42-driven results presented a satisfactory model efficiency (NSE = 0.86) and smaller Bias (0.8% relative to the observations) than that of the rain gauge data at the monthly scale. However, the uncertainty in satellite-based precipitation data, together with parameter uncertainty and the structural uncertainty of hydrologic models, will result in uncertainty in streamflow predictions [85]. Therefore, a better understanding of parameter uncertainties and a comparison of different hydrological models will be the focus of future research.

Many studies have explored the applicability of different SPPs using hydrological models over basins of different scales. When the studied SPPs included TRMM 3B42, CMORPH_CRT, and PERSIANN_CDR, similar results were obtained as this study: the streamflow simulations of TRMM 3B42 and CMORPH_CRT performed better than that of PERSIANN_CDR. For example, Su et al. [27] found that TRMM 3B42 and CMORPH_CRT products showed acceptable performance in four SPPs, while PERSIANN_CDR showed little potential for streamflow simulations over the upper Yellow River Basin in China. For the Xixian Basin (upstream of the Huai River Basin), streamflow simulations using the Xinanjiang model found that TRMM 3B42 forced simulation fitted best with the observed streamflow series among three post-real-time research products; this was followed by the CMORPH_CRT-based simulation, and then the PERSIANN_CDR-based simulation [72]. Alazzy et al. [46] also drew similar conclusions according to the results of hydrologic simulation, testing four SPPs, including the three used in our study. According to Moriasi et al. [86], the models can be considered satisfactory if the NSE >0.5 and the absolute Bias <25%. In this study, the daily NSE and absolute Bias of PERSIANN_CDR were 0.53 and 2.9% under Scenario I, and 0.63 and 3.4% under Scenario II, indicating that the PERSIANN_CDR product also has the capability to produce acceptable streamflow simulation results by using the GR hydrological model in the URRB.

Inevitably, a few limitations were present in this study. First, compared to the dense rain gauge networks of previous studies [10,18], the rain gauge stations in our study region are relatively sparse and unevenly distributed, which may cause uncertainties in the rain gauge comparison and streamflow
Remote Sens. 2018, 10, 1881

5. Conclusions

This study provides a comprehensive assessment of the three latest SPPs (TRMM 3B42 V7, CMORPH_CRT, and PERSIANN_CDR) based on rain gauge observations over the URRB for the period 1998–2010. The primary conclusions can be summarized as follows.

(1) On the grid scale, TRMM 3B42 performs the best, while PERSIANN_CDR performs the worst. Moreover, monthly SPP data have a much better correlation with gauge rainfall data than daily SPP data. Similar results are obtained at the basin scale, but with a high Bias for TRMM 3B42 (9.6%) and a much-improved Bias for CMORPH_CRT (−0.9%) and PERSIANN_CDR (2.2%).

(2) For the detection capability of precipitation events, TRMM 3B42 performs the best, while PERSIANN_CDR exhibits the worst performance. By comparison, CMORPH_CRT shows relatively better capability, but with larger fluctuation among different rain gauge stations.

(3) To different degrees, all three SPPs overestimate or underestimate no-rain (0 mm) and light rainfall (0–1 mm) events. Additionally, there is an overestimation of moderate rainfall events (1–25 mm) and an underestimation of heavy and hard rainfall events (>25 mm), indicating their poor ability to reflect extreme precipitation. For the distribution of relative contribution, the PERSIANN_CDR product deviates the most from gauge data.

(4) During hydrologic validation under Scenario I (gauge-calibrated parameters), the CMORPH_CRT product had the best consistency with observed streamflow series at both daily (NSE = 0.73) and monthly scales (NSE = 0.82), while TRMM 3B42 showed obvious overestimation (24.2%) for both daily and monthly streamflow simulations. Under Scenario II (individual-calibrated parameters), the performance of the recalibrated models significantly improved (NSE > 0.63 for daily, NSE > 0.79 for monthly); TRMM 3B42 and CMORPH_CRT performed better than PERSIANN_CDR. All three SPP-forced simulations showed underestimation for high-flow (18.1%–28%) and overestimation for low-flow (18.9–49.4%).

These findings clearly show the great potential for TRMM 3B42 and CMORPH_CRT products in hydrological applications over poorly gauged and inaccessible transboundary river basins in Southwest China, particularly for monthly time intervals, which are suitable for water resource management. However, all three SPPs underestimate and overestimate the occurrence frequency of daily precipitation for some rain intensity classes. Therefore, the local calibration of satellite-derived rainfall estimates and the merging of satellite estimates with rain gauge observations could be employed to alleviate these problems [87,88]. Future work will focus on the validation of higher-resolution SPPs (i.e., GPM), error corrections, spatial downscaling techniques, and their application in distributed hydrological modeling [9,89].


Author Contributions: Y.L. conceived and designed the study. Y.Z. performed the calculations and assessments. X.J., X.L., and X.L. participated in data processing. Y.Z. wrote the manuscript. All authors contributed to editing and reviewing the manuscript.

Funding: This work was funded by the Applied Basic Research Programs of Yunnan Province (Grant Number 2017FB071) and the National Key Research and Development Program of China (Grant Number 2016YFA0601601).

Acknowledgments: The authors would like to acknowledge the TRMM, CMORPH, and PERSIANN research communities for making the data available to international users. Also, the authors acknowledge the GR model developers for making the suite of lumped GR hydrological models available in an R package. The authors would...
also like to extend their acknowledgements to the editors and anonymous reviewers for their valuable comments and advice.

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

References


48. Li, D.; Christakos, G.; Ding, X.; Wu, J. Adequacy of TRMM satellite rainfall data in driving the SWAT modeling of the Tiaoxi catchment (Taidu lake basin, China). *J. Hydrol.* 2018, 556, 1139–1152. [CrossRef]


60. Duan, Z.; Liu, J.; Tuo, Y.; Chiogna, G.; Disse, M. Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales. *Sci. Total Environ.* 2016, 573, 1536–1553. [CrossRef] [PubMed]


68. Perrin, C.; Michel, C.; Andréassian, V. Improvement of a parsimonious model for streamflow simulation. J. Hydrol. 2003, 279, 275–289. [CrossRef]


77. 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]


82. Derin, Y.; Yilmaz, K.K. Evaluation of Multiple Satellite-Based Precipitation Products over Complex Topography. J. Hydrometeorol. 2014, 15, 1498–1516. [CrossRef]


88. Hashemi, H.; Nordin, M.; Lakshmi, V.; Huffman, G.J.; Knight, R. Bias correction of long-term satellite monthly precipitation product (TRMM 3B43) over the conterminous United States. *J. Hydrometeorol.* 2017, 18. [CrossRef]


© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).