Evaluating the Applicability of Four Latest Satellite–Gauge Combined Precipitation Estimates for Extreme Precipitation and Streamflow Predictions over the Upper Yellow River Basins in China

Jianbin Su 1, Haishen Lü 1, Jianqun Wang 1, Ali M. Sadeghi 2 and Yonghua Zhu 1

1 State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China; ncuw_hhu@126.com (J.S.); wangjq@hhu.edu.cn (J.W.); zhuyonghua@hhu.edu.cn (Y.Z.)
2 Rm. 104, Bldg. 007, USDA-ARS Hydrology and Remote Sensing Laboratory, Beltsville Agricultural Research Center, Beltsville, MD 20705-2350, USA; Ali.Sadeghi@ars.usda.gov

Received: 7 October 2017; Accepted: 14 November 2017; Published: 21 November 2017

Abstract: This study aimed to statistically and hydrologically assess the performance of the four latest and widely used satellite–gauge combined precipitation estimates (SGPES), namely CRT (CMORPH CRT), BLD (CMORPH BLD), CDR (PERSIANN CDR), 3B42 (TMPA 3B42 version 7) over the upper yellow river basins (UYRB) in China during 2001–2012 time period. The performances of the SGPES were compared with the Chinese Meteorological Administration (CMA) datasets using the hydrologic model called Variable Infiltration Capacity (VIC) which is known as a land surface hydrologic model. Results indicated that irrespective of the slight underestimation in the western mountains and overestimation in the southeast, the four SGPES could generally captured the spatial distribution of precipitation well. Although 3B42 exhibited a better performance in capturing the spatial distribution of daily average precipitation, BLD agreed best with CMA in the time series of watershed average precipitation, which resulted in BLD having a comparable performance to the CMA in the long-term hydrological simulations. Moreover, the potential for disastrous heavy rain mainly occurs in southeastern corner of the basin, and CRT and BLD comparisons showed to be closer to the CMA in the distribution of extreme precipitation events while 3B42 and CDR overestimated the extreme precipitation especially over the southeast of UYRB region. Therefore, CRT and BLD were able to match the high peak discharges very well for the wet seasons, while 3B42 and CDR overrated the high peak discharges. In addition, the four SGPES performed well for the 2005 flood event but exhibited poorly when tested for the 2012 flood event. Results indicate that the application of the four SGPES should be used with caution in simulating massive flood events over UYRB region.

Keywords: precipitation; Hydrometeorology; Variable Infiltration Capacity (VIC) model; Tibetan Plateau; upper yellow river basins

1. Introduction

Floods are among the most frequently occurring and disastrous natural hazards worldwide and have caused tremendous loss of life and property over the past decades [1]. Due to climatic characteristics, conditions of stream banks and channel slope, flood events occur commonly in some parts of the populated zones [2] and have drawn increasing risks with increasing urbanization, growing population, and global warming [3]. Frequent flood events highlight the importance of studies on flood simulation and warning, specially for more vulnerable areas of the world.
Hydrologic models have become important tools for understanding hydrological processes, particularly for forecasting and monitoring flood hydrograph. Among hydrologic models, physically-based distributed hydrologic models characterize complex hydrologic processes in watersheds by using spatialized variables and parameters [4]. In spite of synthetic streamflow generated by distributed hydrologic models being a response to a highly complex and non-linear process, gridded precipitation is a critical input for distributed hydrologic models. The accuracy of precipitation data is essential for reliable hydrologic prediction [5,6]. Conventionally, the gridded precipitation that drives the hydrological model is mainly generated by interpolating ground observations (rain gauge and weather radar networks). However, techniques for making precipitation observations from ground-based measurement networks have limitations in hydrologic modeling because of the large spatial nonuniformity and temporal unavailability in rainfall fields inherently. More severely, there are very few available ground-based observations in inaccessible regions, ungauged basins or heterogeneous terrains [7,8]. With the significant increase in spatial coverage with high spatial and temporal resolutions, satellite-based precipitation estimates (SPEs) are critical and valuable resources in acquiring reliable hydrologic data, particularly for the regions without ground observation networks.

A growing number of quasi-global SPEs with high spatial and temporal resolution have been put into operation and released publicly [9], among which there are many broadly used SPEs such as, the Tropic Rainfall Measurement Mission (TRMM) Multi-Satellite Precipitation Analysis (TMPA) Real-time product (3B42RT) [10], the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC) morphing technique (CMORPH) estimations [11], and Precipitation Estimation from Remotely Sensed Imagery using Artificial Neural Networks (PERSIANN) [12]. More recently, the Integrated Multi-satellite Retrievals for Global Precipitation Measurement mission (IMERG) data products can provide global precipitation estimates at a finer scale of 0.1° x 0.1° spatial resolution and 30 min temporal resolution. Although multi-sensor blended precipitation products have been greatly improved with various sensors flying on a series of satellites and the development of satellite-based precipitation retrieval algorithms [13], the SPEs, as the indirect estimations of precipitation, are also suffering from seasonal and regional systematic biases and random errors [14–16]. The errors of SPEs can be propagated into streamflow predictions through the hydrologic integration processes [17]. In China, numerous validation efforts have been conducted [6,18–21] over Tibet Plateau (TP). Correspondingly, their predictive ability of the streamflow rate has been also validated by the hydrological modeling framework. However, for instance, the authors of [21] verified that 3B42RT and PERSIANN possess limited potential in streamflow simulation over two sub-basins (upper yellow and upper Yangtze river basins) of the TP. Similar conclusions can be found in disparate basins of the TP [22–25].

Note that ground observation networks are able to provide the most reliable point precipitation observations but lack the spatial representativeness, and SPEs have fine spatial and temporal coverage but suffer from inherent uncertainty. Therefore, the satellite–gauge combined precipitation estimates (SGPEs) taking the pros and cons of SPEs and gauge observations into account have been generated. For example, using gauge analysis from the Global Precipitation Climatology Center (GPCC) to remove the monthly bias of TRMM combined instrument retrievals, the latest post-real-time version (3B42) were released. Based on the CPC unified daily gauge analysis [26], a bias-corrected product (CRT) was generated by NOAA/CPC through adjusting the biases of the CMORPH SPE (CMORPH PAW). Further, another CMORPH’s satellite-gauge blended product (BLD) was generated by combining CMORPH CRT and gauge observations through the Optimal Interpolation (OI) into a unified dataset. Recently, a new satellite-based precipitation product was released by the National Climatic Data Center (NCDC), which was named PERISANN—Climate Data Record (CDR).

Despite the continuing great advances in acquiring accurate precipitation, the errors of SGPEs still exist in the adjusted/merged products because the characteristics of the retrieval errors in different climatic regions, seasons, and surface conditions [27]. The performance of SGPEs is always
changing with the development of retrieval algorithms, data sources and gauge adjustment procedures. The validation efforts can be roughly divided into two categories: the first one focuses on the comparison and evaluation of SGPEs against gauge data and ground-based radar estimates. By this principle, temporal characteristics and spatial distributions of SGPEs are not only investigated but also can be quantitatively analyzed. However, it sustains the scale discrepancy problem while using rain gauge data for validation. The other one is to evaluate SGPEs based on their predictive ability of streamflow rate in a hydrological modeling framework. By this way, the precipitation products are evaluated at the watershed scale with respect to a specific application. Some global and regional validations have been reported for different SGPEs [16,28–30]. For the validation activities for SPEs over the Tibet Plateau (TP)—a region with very complex topography and high elevation, the authors of [21] illustrated that CRT exhibits an encouraging potential for hydrological applications over two sub-basins of the Tibetan Plateau in spite of the general underestimation, while 3B42 shows comparable performance to the China Meteorological Administration data in both monthly and daily streamflow simulations mostly because the monthly gauge adjustment is involved. The authors of [31] insisted that BLD would exhibit higher quality and more stable performance benefited by the probability density function-optimal interpolation (PDF-OI) gauge adjustment procedure. However, BLD features the similar error characteristics of CRT with a positive bias of light precipitation and a negative bias of heavy precipitation owing to the insufficient gauge observations in the merging process. Moreover, BLD was regarded as the product with the best quality of the three CMORPH products [32]. Unfortunately, it has never been assessed over the TP. Besides, the works by [33] showed that the simulated streamflows derived by CDR were closer to observations than those derived by limited gauge-based precipitation interpolation in upper Yangtze River basin of the TP though the Hydroinformatic Modeling System (HIMS) rainfall–runoff model. However, to date, the four SGPEs have not been evaluated thoroughly and systematically over the TP’s basin, especially for their performance on extreme precipitation scenes and hydrologic applicability in capturing the extreme streamflow.

The Yellow River is China’s mother river and is the cradle of Chinese civilization. It originates on the eastern Tibet Plateau flowing across the Loess Plateau and discharges into the Bohai Sea. The upper Yellow River basins (UYRB), above Tangnaihai, is an important component of the National Nature Reserve of Three Rivers Source. Although the UYRB covers a drainage area of approximately 122 \times 10^3 \text{ km}^2, accounting for 16.2\% of the whole Yellow River basin, it is the most important contributing area providing more than 35\% discharge of the Yellow River basin. Since the UYRB is sparsely populated, most of studies over UYRB focus on water resource management and the impact of climate change, there are relatively few studies on flood simulation and prediction. With the development of social economy, flooding prevention cannot be ignored, especially for the 2012 flood event which had brought great pressure to regulation of longyangxia reservoir downstream. Here, the purpose of this paper is to evaluate and compare four SGPEs in daily scale against ground-based precipitation measurements and to assess their hydrologic utility, especially on extreme precipitation and extreme streamflow though the Variable Infiltration Capacity (VIC) model over UYRB, where is poor gauged but has reservoir constructions downstream. The remainder of this paper is organized as follows. Section 2 provides the details of the study area, datasets and methods. Sections 3 and 4 present the main results and the discussion, separately. Finally, a summary of the work is given in the last section.

2. Study Areas, Datasets and Metrics

2.1. Study Area

The representative region selected by this study is located in eastern Tibetan Plateau, China, as shown in Figure 1. The geographical location of UYRB is between 32.17°–36°N, 95.8°–103.5°E. The region under study mostly over 4000 m is featured with meandering drainage river crossing the
towering mountains and the rolling hills. The UYRB can be divided into three sub-catchments by geography, namely, upstream, midstream, and downstream controlled by Jimai, Maqu and Tangnaihai respectively. The upstream is known as the star-studded lakes including the two largest outflows of freshwater lakes namely Eling Lake and Zaling Lake. The midstream is the main watershed of the UYRB mostly because the marsh wetland consumes lots of space of the midstream which can provide replenishment for the trunk stream. The downstream is characterized by steep-sided river valleys and snowmelt runoff. More information of the three sub-catchments can be achieved from Table 1.

The basin is an alpine climate zone characterized with a dry winter and a rainy summer. In winter, the UYRB is controlled by the Tibetan High which cause a seven-month cold period and most places have no absolute frost-free period. In summer, the UYRB is influenced by the southwest monsoon with thermal low pressure, abundant water vapor and high precipitation accounting for 75–90% of the annual total. In addition, temperature variation of the annual is very small but the temperature variation between day and night is so big. The land use and land cover classes is consisted by the typical Tibet Plateau vegetation belt which is a young derivatives department characterized by high mountain specialization and cold and drought specialization. It consists of evergreen forest, open scrublands, grasslands, and bare ground being the main types. Soil types of the UYRB are dominated by sandy loam, sandy clay and loamy sand being the main types.

![Location of the upper Yellow River basin (UYRB). The hydrological station and the meteorological station used in paper are also shown in the figure.](image)

**Figure 1.** Location of the upper Yellow River basin (UYRB). The hydrological station and the meteorological station used in paper are also shown in the figure.

<table>
<thead>
<tr>
<th>Category</th>
<th>Hydrometric Station</th>
<th>Area (km²)</th>
<th>Average Elevation (m)</th>
<th>Average Runoff (m³/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>upstream</td>
<td>Jimai</td>
<td>45,019</td>
<td>4464</td>
<td>173.75</td>
</tr>
<tr>
<td>midstream</td>
<td>Maqu</td>
<td>41,029</td>
<td>3894</td>
<td>470.92</td>
</tr>
<tr>
<td>downstream</td>
<td>Tangnaihai</td>
<td>35,924</td>
<td>3947</td>
<td>696.70</td>
</tr>
</tbody>
</table>

**2.2. Datasets Description**

In the current study, due to restrictions on the availability of the hydrometric stations observations over the UYRB, a twelve-year period from 1 January 2001 to 31 December 2012 has been selected, of which the period from 2001 to 2005 is treated as the calibration period while 2006–2012 is treated as the validation period. Both satellite and gauged precipitation datasets are available in these periods. The historical records of daily datasets are collected from the Chinese Meteorological Administration.

**Table 1.** The major information about the three hydrometric stations.
(CMA) meteorological stations. Moreover, the ground observations including precipitation, minimum temperature, maximum temperature and wind speeds are interpolated into 0.25° × 0.25° grids to match the spatial distribution of satellite precipitation. Besides, daily observed streamflow records of the UYRB are collected from TangNaiHai hydrological station.

This study aimed to assess four SGPEs, namely CRT, BLD, CDR and 3B42 and investigated their suitability in hydrological application over UYRB region. A general introduction description of these products is summarized as follows. The TMA product’s latest version 7 was released in May 2012 by the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC), in which a near real-time version (3B42RT) and a post-real-time version (3B42) are included. These estimates are produced by merging three types of observations such as PMW, IR, and PR from multiple LEO and Geo satellites and ground observations at a spatial resolution of 0.25° × 0.25°, with a temporal resolution of 3-hourly. The main difference is that the TRMM Combined Instrument (TCI) dataset has been used in 3B42 for calibration; however, it was replaced by the TRMM Microwave Imager (TMI) dataset in 3B42RT. Besides, the 3B42 provides the bias-corrected precipitation estimates by inclusion ground gauge observations from the GPCC and the Climate Assessment and Monitoring System (CAMS) at the monthly scale while 3B42RT is a satellite only precipitation estimates. Therefore, the 3B42RT is available approximately nine hours after realtime observations, while the 3B42 is available about 10–15 days after the end of each month. Here, the 3B42 is employed in daily scale from 2001 to 2012.

The CMORPH product stands for the precipitation product estimated by the NOAA Climate Prediction Center MORPHing technique. The latest CMORPH V1.0 products include RAW, CRT and BLD. The CMORPH RAW is a satellite-only precipitation product by combining existing passive microwave-based precipitation estimates from multiple low orbit satellites and the infrared data from multiple geostationary satellites. Based on the RAW product, the CRT product is generated by adopting the probability density function (PDF) matching to conduct bias reduction over land. The CRT product is further combined with the gauge analysis through an optimal interpolation (OI) technique to generate the BLD product. All three products are available at PERSIANN websites [34]. Here, just the two SGPEs versions are applied in daily scale from 2001 to 2012.

The original PERSIANN is one of the popular global precipitation estimation for estimating historical precipitation from March 2000 to present developed by combining IR and PMW observations from GEO and LEO satellite imagery, respectively. The latest CDR uses the archive of the GridSat-B1 IR data as the input to train PERSIANN model; then the biases in the PERSIANN estimated precipitation is adjusted by the Global Precipitation Climatology Project (GPCP) monthly 2.5° product version 2.2. Unlike the other precipitation estimations, the PERSIANN model parameters were pretrained by using the National Centers for Environmental Prediction (NCEP) stage IV hourly precipitation data instead of PMW data. Currently, this version of PERSIANN CDR is only available with high spatial resolution of 0.25° × 0.25° and daily temporal resolution. It can be acquired from 1 January 1983 to present form CMORPH websites [35]. However, in this paper, subset data from 2001 to 2012 are adopted in daily scale.

2.3. Hydrological Model

The hydrologic model used in this study is the Variable Infiltration Capacity (VIC) model which is known as a macroscale and semidistributed land surface hydrologic model. It was first developed by Liang et al. [36,37] for solving both surface water and energy balances over a grid mesh. Then, various updates of the model have been further described, such as cold land process updates [38], snow model updates [39], lakes and wetlands [40]. The key characteristics of the grid-based VIC are the representation of vegetation heterogeneity, multiple soil layers with variable infiltration, and non-linear base flow. The three-layer VIC model (VIC-3L) framework includes a top thin soil layer to represent quick bare soil evaporation following small rainfall events, a middle soil layer to represent the dynamic response of the soil to rainfall events based on the infiltration capacity curve as described by the Xinanjiang hydrological model [41], and a bottom soil layer (third
soil layer, maximum depth around 1.5 m) to characterize the seasonal soil moisture behavior based on the drainage in the Arno model [42]. In addition, the simulated streamflow at the basin outlet was post-processed by a separate routing model based on the linearized Saint-Venant equation using the VIC grid results as inputs [43]. Besides, the VIC model assigns many of its parameters based on the vegetation type and soil texture. Although most parameters can be directly estimated from the land surface database, several important parameters including the exponent of variable infiltration capacity curve (B), the maximum velocity of base flow ($D_{\text{smmax}}$), the fraction of maximum base flow (D$_{s_{\text{max}}}$), the fraction of maximum soil moisture content of the third layer (W$_{s_{3}}$), the second and the third soil-layer thicknesses ($d_2$ and $d_3$) must be optimized through the model calibration process. For further details, readers can refer to [44]. The VIC model has been widely applied in seasonal hydrological forecasting, climate change impacts studies and water and energy budget studies. Recently, there is a new trend that the VIC model has been popularly used for evaluating the capability of satellite precipitation retrievals in streamflow simulation [21,23,45]. Here, the VIC model will be employed to evaluate the capability of SGPEs in streamflow simulation over UYRB of the TP. To match the resolution of precipitation products from the satellite datasets, the VIC model was set at 0.25° spatial resolution.

### 2.4. Statistical Evaluation Metrics

To quantitatively analyze the overall performance of SGPEs in comparison with gauged precipitation observations, a set of continuous verification metrics are considered, including correlation coefficient (CC), relative bias (RB), and root mean-square error (RMSE). Also, the CC and RB are operated to assess the performance of simulated streamflow. Besides, the Nash-Sutcliff (NS) index is carried out to quantify the performance of the hydrological model. The formulas are given as:

\[
CC = \frac{\sum_{i=1}^{n}(\text{Obs}_i - \overline{\text{Obs}})(\text{Sim}_i - \overline{\text{Sim}})}{\sqrt{\sum_{i=1}^{n}(\text{Obs}_i - \overline{\text{Obs}})^2} \sqrt{\sum_{i=1}^{n}(\text{Sim}_i - \overline{\text{Sim}})^2}}
\]

\[
RB = \frac{\sum_{i=1}^{n}(\text{Obs}_i - \text{Sim}_i)}{\sum_{i=1}^{n}\text{Sim}_i} \times 100\%,
\]

\[
RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\text{Obs}_i - \text{Sim}_i)^2},
\]

\[
NS = 1 - \frac{\sum_{i=1}^{n}(\text{Obs}_i - \overline{\text{Obs}})^2}{\sum_{i=1}^{n}(\text{Sim}_i - \overline{\text{Sim}})^2},
\]

where $\text{Sim}_i$ and $\text{Obs}_i$ represent SGPE (or simulated streamflow) and ground observed precipitation (or observed streamflow), respectively. $\overline{\text{Sim}}$ and $\overline{\text{Obs}}$ are mean values of the corresponding elements.

### 3. Results and Analysis

#### 3.1. Evaluation and Comparison of Satellite-Gauged Precipitation Estimates

We evaluated the applicability of the four SGPEs based on CMA over the study region. Because of the negative effects of their associated uncertainties on the hydrological modeling process, the comparative analysis was divided into two periods, which were the calibration (2001–2005) and the validation (2006–2012) periods. This was to investigate and characterize precipitation patterns and error quantification of the four satellite products over the UYRB region. Data in Figure 2 show the spatial distribution of daily average precipitation obtained from CMA and the four SGPEs (CRT, BLD, CDR, and 3B42) tested, respectively for calibration (first column) and validation (second column) periods. Furthermore, the mean values (MEs) of daily average precipitation over the UYRB region had been depicted in Figure 2. Intuitively, Figure 2 shows that the precipitation patterns derived by CMA and the four SGPEs are visually compatible in the two periods with the precipitation intensities gradually decreasing from the eastern part of the basin to the western part. Meanwhile, a large
amount of precipitation concentrated upon the southeastern region of the UYRB and the spatial variability analysis revealed that the low-altitude regions of the basin are characterized by higher spatial variability of precipitation in comparison with the high mountainous regions partly due to the abundant moisture supply by the Indian Summer Monsoon from the Bay of Bengal and the orographic enhancement effect in the western part.

It should be noted that the precipitation amount in validation periods was higher than that in the calibration periods. Additionally, both CRT and BLD resembled extremely well with the CMA in precipitation amount. However, some precipitation patches were found in the CRT and BLD estimates in the source region of the UYRB that are not observed in CMA datasets. The primary reason is that CMORPH possesses inherently systematic anomalies over inland water bodies [46]. However, these systematic anomalies in validation period were slighter than those in calibration period. Furthermore, compared to CMA datasets, CDR had the highest precipitation estimates in both periods while 3B42 reaches the lowest value.

Figure 2. Spatial distribution of daily average precipitation (mm/day) derived from CMA and SGPEs during calibration (2001–2005) and validation (2006–2012) periods.

The RB of the SGPEs against the CMA data in both calibration and validation periods are shown in Figure 3. In general, the four SGPEs showed an underestimation of the precipitation in the upstream regions dominated by mountains and an overestimation in the midstream and downstream regions with low-altitude during the two periods. Regardless of the overestimation over the area of Zhaling Lake and Eling Lake, CRT exhibited an obvious pattern that the negative biases primarily occurred in the western region of east longitude 99° while the positive biases are located in the eastern region.
during the calibration period and these underestimation and overestimation become more visible in validation period. Compared with the corresponding datasets, CRT, BLD exhibited a similar spatial pattern of RB with more petty negative and positive biases, respectively in the west and the east. This means that BLD is more close to CMA than CRT in precipitation amount. Concurrently, CDR showed a large positive bias in the southeast UYRB and the regions close to Jimai hydrological station while slight negative bias just occurred in source region and outlet region of UYRB region. Correspondingly, 3B42 underestimated the precipitation in source region and outlet region of UYRB but there was no prominent overestimation in the southeast UYRB and the regions close to Jimai hydrological station during the two periods.

To check the consistency of precipitation in time series, the basin averaged precipitation from SGPEs versus CMA datasets were scattered in Figure 4 respectively. Meanwhile, the Quantile-Quantile (Q-Q) plot technique was adopted to illustrate more insight into the nature of the differences between the four satellites and the CMA data in both calibration (2001–2005) and validation (2006–2012) periods over the UYRB region. It is well known that if the SGPEs are close to the observed ones, the points in the Q-Q plots should fall close to the 45° reference line. The greater the departure from the reference line or the nonlinearity of the resulting graph is, the greater the evidence of heterogeneity is [47]. It was evident that BLD has the best consistency with CMA in both calibration and validation periods. Simultaneously, the differences in daily average precipitation estimates between SGPEs and CMA became more distinct as the precipitation amount increased. Furthermore, the performance exhibited by CRT was similar to 3B42 in the two periods while CDR tends to overestimate precipitation in both calibration and validation periods.

In addition, the three statistical indexes including CC, RB and RMSE were calculated based on CMA data in two periods as shown in Figure 4. Generally, the best values of CC = 0.66 and
CC = 0.67 from BLD were found in calibration period and validation period, respectively. Also, BLD had the best RB and RMSE displayed as 0.16\%, −4.25\% and 1.86\%, 1.94\% during the two periods, respectively. In contrast, CDR showed the poorest values with CC = 0.45, RB = 12.22\% and RMSE = 2.92 in calibration period and CC = 0.45, RB = 9.39\% and RMSE = 3.27 in validation period. Moreover, the three statistical standards of CRT were very close to that of 3B42 in validation period. The values of CC and RMSE from CRT were approximately equal to those from 3B42 in validation period, but 3B42 tended to underestimate the precipitation amount in calibration period and there is no evidence to support this phenomenon in CRT with RB = 0.2\%.

3.2. Extreme Precipitation Analysis

To reveal the changing features of extreme precipitation in the UYRB region, four extreme precipitation scenarios were chosen according to Jiang et al. [9], which are the 95th percentile of daily precipitation (95\% pr), the maximum 1-day precipitation (H1D pr), the maximum 3-day precipitation total precipitation (H3D pr) and the maximum 7-day total precipitation (H7D pr) for each pixel during 2001–2012. Difference of the SGPEs from the four extreme precipitation scenarios were compared in Figure 5 respectively. The spatial pattern of the 95\% pr from CMA and SGPEs were similar to that of the corresponding daily average precipitation illustrated in Figure 2. It can be observed that the value of 95\% pr is large higher than that of the daily average precipitation. From the 95\% pr analysis, it was found that the spatial heterogeneity of CMA was slightest probably because of the influence of interpolation. In contrast, CDR had the strongest spatial heterogeneity over 15 mm/day in southeast corner and under 3 mm/day in northwest corner. The difference in spatial heterogeneity between CDR and CMA was gradually growing in the scenario of H1D, H3D and H7D. For the scenario of H7D, nearly half of the UYRB had more than 200 mm precipitation in seven days but most regions in CMA was lower than 100 mm in seven days. Besides, 3B42 had much higher value than CMA in the western
UYRB in terms of the four extreme cases of 95%, H1D, H3D and H7D. This differed from the daily average case when 3B42 underestimated the precipitation amount. Without the consideration of the overestimation over inland water bodies, both CRT and BLD enjoyed the similar spatial distribution and precipitation amount estimation to CMA for the four extreme cases, although they overestimate the precipitation amount in the western basin. In general, the extreme precipitation variability varies greatly over the UYRB. The high values were mainly concentrated in the middle basin while less extreme higher precipitation occurred in downstream reaches and the least extreme higher precipitation was found in the source region. Although the four satellite estimates were all overestimated the extreme precipitation amount over western region, compared with CDR and 3B42, both the spatial distribution and the precipitation amount of CRT and BLD were better agreement with the real extreme precipitation in CMA.

3.3. Streamflow Simulations Analysis

Here, the VIC model was firstly calibrated by comparing the observed and simulated streamflow forced by the CMA precipitation data in the calibration period over the UYRB. The parameters subjected in calibration and the sensitivity of those parameters were discussed in detail in [48]. Subsequently, the period 2006–2012 was used for model validation. Lastly, the VIC model was forced by CRT, BLD, CDR and 3B42 as inputs for twelve years (2001–2012) with the model parameters calibrated using CMA data in the calibration period. The comparison of simulated streamflow and observed streamflow in both calibration and validation period are illustrated in Figure 6. It can be seen that the simulated streamflow hydrograph can reasonably well match the observed streamflow, especially for the high peak discharges in wet seasons. In spite of the overestimation in the dry seasons, the performance of the VIC model did not affect subsequent analysis associated with the relatively low discharges in the dry seasons. Besides, the three statistical indicators for calibration, verification and
the whole periods are listed in Table 2, respectively. Overall, the VIC model was capable of capturing the timing and magnitude of the daily observed streamflow.

![Graph showing streamflow comparison](image_url)

**Figure 6.** Observed and simulated streamflow forced by CMA data for the UYRB for the calibration period (2001–2005) and verification period (2006–2012).

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Calibration</th>
<th>Verification</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NS</td>
<td>CC</td>
<td>RB (%)</td>
</tr>
<tr>
<td>CMA</td>
<td>0.80</td>
<td>0.90</td>
<td>4.84</td>
</tr>
<tr>
<td>CRT</td>
<td>0.47</td>
<td>0.8</td>
<td>24.76</td>
</tr>
<tr>
<td>BLD</td>
<td>0.72</td>
<td>0.87</td>
<td>13.76</td>
</tr>
<tr>
<td>CDR</td>
<td>&lt;0</td>
<td>0.74</td>
<td>32.01</td>
</tr>
<tr>
<td>3B42</td>
<td>0.62</td>
<td>0.83</td>
<td>3.84</td>
</tr>
</tbody>
</table>

**Table 2.** The statistical indicators for simulated streamflow forced by CMA data and four SGPEs versus the observed streamflow during calibration (2001–2005), verification (2006–2012) and the total (2001–2012) periods.

The calibrated VIC model was then used to evaluate the SGPEs in streamflow simulation over the UYRB, without any further adjustment of parameters. Keeping the same parameters allows us to investigate how differences of SGPEs affect the accuracy or quality of VIC streamflow simulation. Figure 7 shows the daily simulated streamflow with the precipitation forcing from the four SGPEs over the UYRB. Intuitively, the SGPEs driven simulations catch the fluctuation of daily hydrograph well, although underestimation/overestimation of peak floods existed in some cases. However, compared to Figure 6, none of the four SGPEs driven simulations made significant contributions to improve the performance of streamflow simulation in the timing and magnitude of the observed streamflow. In fact, given that the model was calibrated with CMA, that is not surprising. Besides, the streamflow simulation forced by CDR overestimated the flood peak in both the calibration and verification periods, which was probably because CDR overestimated the amount of precipitation in midstream. The BLD simulation showed the best consistent with the observed daily hydrograph, but it tends to overestimated the peak in calibration period (except for 2005) and underestimated the peak in validation period (except for 2006). The overvalued or undervalued scenarios were more noticeable in CRT simulation results. Moreover, excepting 2001 and 2010, the 3B42 simulations could capture the peak very well but there was a large fluctuation in the 3B42 simulations with the snipy peak. In addition, the statistic summary of the comparison between the observed and the SGPEs forced scenarios also be shown in Table 2. The simulation results of hydrological model were directly depend on precipitation input, the more accurate the precipitation input, the better the simulation results. Overall, from Table 2, the CMA simulations have showed the best performance of all, followed by BLD. Due to the overestimation in heavy rain, CDR simulation scenario overrates the flood peaks in in wet seasons and showed poor NS indicators. As for 3B42 simulation scenario, it had the best values of
RB (3.84%, −6.64% and −2.75% respectively) although the NS of 3B42 (0.62, 0.65 and 0.64 respectively) were lower than that of CMA (0.80, 0.80 and 0.81 respectively). CRT simulation scenario exhibited a large overestimation in hydrograph over calibration period, but it showed a similar performance of 3B42 simulations in verification period.

3.4. Extreme Streamflow Analysis

Figure 8 shows the extreme values of observed streamflow and simulated streamflow driven by CMA and SGPEs in the UYRB from 2001 to 2012. The four extreme streamflow scenarios illustrated in Figure 8 were the 95th percentile of daily streamflow, the maximum 1-day streamflow (H1D St), the maximum 3-day streamflow (H3D St) and the maximum 7-day streamflow (H7D St) for each year of study period respectively. The results showed that the CMA could capture the main patterns of the extreme streamflow in the UYRB by using VIC hydrological model with slight underestimation in flood years and hidden overestimation in dry years. CRT and BLD exhibited comparable hydrologic performance to CMA, especially for BLD in H1D St, H3D St and H7D St scenarios. 3B42 also captured the main patterns of the extreme streamflow. However, excepting 2007 and 2012, it had an explicit overestimation in the four scenarios. Similar case occurred in CDR forced scenario, it captured the main patterns of the extreme streamflow but enjoyed the largest overestimation. Besides, the statistic summary of the comparison between the observed and simulated extreme streamflow is shown in Table 3.

**Table 3.** The statistical indicators for simulated extreme streamflow forced by CMA data and four SGPEs versus the observed streamflow during 2001–2012.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>95% St</th>
<th>H1D St</th>
<th>H3D St</th>
<th>H7D St</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NS</td>
<td>CC</td>
<td>RB (%)</td>
<td>NS</td>
</tr>
<tr>
<td>CMA</td>
<td>0.56</td>
<td>0.84</td>
<td>−4.33</td>
<td>0.59</td>
</tr>
<tr>
<td>CRT</td>
<td>0.37</td>
<td>0.69</td>
<td>5.38</td>
<td>0.45</td>
</tr>
<tr>
<td>BLD</td>
<td>0.59</td>
<td>0.83</td>
<td>−0.88</td>
<td>0.63</td>
</tr>
<tr>
<td>CDR</td>
<td>&lt;0</td>
<td>0.70</td>
<td>49.89</td>
<td>&lt;0</td>
</tr>
<tr>
<td>3B42</td>
<td>0.26</td>
<td>0.71</td>
<td>16.89</td>
<td>&lt;0</td>
</tr>
</tbody>
</table>
These two floods have brought great flood control pressure to the deployment of the Longyangxia reservoir downstream. The statistical characteristics of the two floods have been shown in Table 4. Based on the works reported in [49], it was found that floods in UYRB are characterized by rise and fall flat, long duration, low peak and huge flood volume. The capability of satellite-based products in predicting severe flood events was evaluated. In July and August 2012, there was a long and continuous duration heavy rainfall occurred over UYRB. The everlasting torrential rain had led to the most serious flood in UYRB since 1989 [49]. The largest discharge of the 2012 flood event was 3350 m$^3$/s occurring on 24 July 2012. Besides, a high discharge of 2720 m$^3$/s is found on 6 October 2005. These two floods have brought great flood control pressure to the deployment of the Longyangxia reservoir downstream. The statistical characteristics of the two floods have been shown in Table 4.

Table 4. Basic data of the two selected flood events.

<table>
<thead>
<tr>
<th>Event</th>
<th>Start</th>
<th>End</th>
<th>Peak Discharge (m$^3$/s)</th>
<th>Flood Volume (10$^9$ m$^3$)</th>
<th>Last (Day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>18 September</td>
<td>31 October</td>
<td>2720</td>
<td>59.7</td>
<td>44</td>
</tr>
<tr>
<td>2012</td>
<td>26 June</td>
<td>13 September</td>
<td>3350</td>
<td>149.3</td>
<td>80</td>
</tr>
</tbody>
</table>

The observed streamflows and simulated streamflows forced by CMA and four SGPEs are illustrated in Figure 9. As the simulation hydrographs are cohesive sets of the corresponding precipitation input, the time series of precipitation during the flood events are also plotted in Figure 9. In addition, the three statistical indicators (RB, CC and RMSE) for SGPEs against CMA data during the flood period over UYRB have been analyzed in Figure 10. Owing to the perfect match in magnitude and timing of the event 2005 flood, CMA is further believed to reflect the spatial and temporal distribution of precipitation during the 2005 flood. Compared to CMA, CRT and BLD both enjoy a negative RB during the event 2005 flood, but BLD had a higher CC and a lower RMSE than CRT. Hence, BLD forced simulation scenario showed a better performance than that of CRT in 2005 flood period in spite of the underestimation in flood peak. Combinative comparison of Figures 9 and 10 showed that CDR possessed a positive RB but the CDR forced simulation scenario matches the flood peak best in the 2005 flood period. This illustrates that CDR had great uncertainty in space-time distribution and the median and the mean values should not be used exclusively as measures of location when there are evident outliers. Moreover, the lower flood peaks were found in SGPEs based simulation results but not found in observed streamflow and CMA forced scenario, which is mainly because of the unreliable antecedent precipitation in SGPEs.
Figure 9. The observed and simulated streamflows for both the event 2005 and event 2012 flood scenarios.

Figure 10. The boxplot of the statistical indicators of SGPEs against CMA data over the UYRB. Five lines from bottom to top for one box represent minimum value, 25th percentile, 50th percentile, 75th percentile, and maximum value, respectively.

Unfortunately, the CMA based simulation results significantly underestimated the flood events of 2012, and the CMA based flood peak lags behind the observed values. This indicates that the veracity of the CMA based precipitation is questionable in amount and occurrence time in the 2012 flood period. So the quantitative assessment using CMA as the reference is unreliable over the 2012
Unfortunately, the CMA based simulation results significantly underestimated the flood events of 2012, and the CMA based flood peak lags behind the observed values. This indicates that the veracity of the CMA based precipitation is questionable in amount and occurrence time in the 2012 flood period. So the quantitative assessment using CMA as the reference is unreliable over the 2012 flood period. However, the accurate spatial and temporal precipitation distribution plays an important role in generating realistic streamflow through distributed hydrologic models. The performance of precipitation can be indirectly analyzed from the simulation results. Three major flood peaks continuously happened on 12 July, 24 July and 24 August, and the middle flood event was the heaviest. Although CRT accurately captured the magnitude and time of the third flood peak, but CRT and BLD forced simulations greatly underestimate the previous two flood peaks. Low quality in reproducing 2012 flood event reflects once again that both CRT and BLD is incompetent in detecting extreme precipitation during the flood 2012 event. For 3B42, it was observed that there exist remarkable magnitude differences with underrating the first two peaks and overrating the third one, and there was a time delay for middle flood. These observation show that 3B42 underestimated the inchoate precipitation but overestimated the later precipitation during the 2012 flood period. CDR forced hydrograph missed the flood peak and had large jumps which is mainly because of that CDR better depicted the total areal rainfall but misrepresented the heavy precipitation’s pattern and occurred time.

4. Discussion

By merging ground-based observations, SGPEs could reduce the errors of SPEs and showed a new opportunity for basin-scale hydrological applications. However, the adjusted/merged procedures in the production of SGPEs greatly affected their quality of precipitation estimation. Despite the adjustment by GPCP monthly precipitation, the CDR product with the IR data as the input to PERSIANN model showed an intrinsic deficiency over UYRB. Given the longer time series (1983 to present), the CDR product is still extremely valuable in investigating climate change with high resolution. As for the two CMORPH-based SGPEs, benefiting from the two step bias correction method (PDF-OI merging algorithm), BLD showed more stable performances than CRT whose overall quality was comparable with 3B42. Since CRT use PDF matching to adjust daily bias of CMORPH RAW data while 3B42 only handles the monthly biases, on the other hand, the OI procedure used in BLD is considered to be a more effective method in improving precipitation estimates. However, further evaluations in extreme precipitation analysis indicated that the CRT showed great advantage over 3B42, which showed the fine time-scale modifications can better capture wickedly heavy precipitation.

For streamflow simulating, the hydrological utility of SGPEs was associated with large uncertainties form parameters and model structures. Theoretically, the parameters of hydrological model should be calibrated using the “perfect” observed precipitation data in a dense gauge network, so that they can be the best possible approximation of watershed hydrological features. Nevertheless, restricted to the high altitude, the gauge network’s distribution over UYRB is sparse. A few studies [33] suggested that recalibrated the hydrological model using the SGPEs can greatly improve the streamflow simulation performance. The uncertainty of parameters and the comparison of different models are being carried out in the following research. Moreover, since hydrological processes in distributed hydrologic models were sensitive to the total precipitation amount as well as rainfall intensity distribution, the errors of SGPEs were also propagated into hydrologic simulations. Using the parameters calibrated with CMA, the general streamflow pattern forced by the four SGPEs exhibits fair agreement with observation. Note that CRT and BLD were remarkably superior to 3B42 and CDR in flood peak simulation, this indicates again the superiority of fine time-scale correction. However, in the streamflow simulations analysis especially for the two flood events, the simulation results forced by the four SGPEs are not satisfactory. This is mostly because the rain gauge observations used in SGPEs are limited to GTS reports and it is hard to capture heavy rain center effectively through sparse gauge network. Despite of the same PDF-OI merging algorithm, the work in [31] indicated
that the newly developed SGPE which combines the CMORPH RAW data with daily precipitation observations from about 2400 gauges in China had a better performance than BLD over Huaihe river basin. With more available rain gauge observations, there are great advantages of regional or national scale modifications in improving the quality of precipitation data. Besides, the meteorological stations over UYRB region are mainly located in the midstream and downstream regions, more efforts are therefore urgently needed to build a denser rain gauge observation network in the upstream regions where are characterized by complex terrain and a rigid climate.

5. Summary and Conclusions

In this paper, we have made a comprehensive analysis of the applicability and reliability of the latest SGPEs (CRT, BLD, CDR, and 3B42) against gauge-based datasets (CMA) in UYRB where it is a most important watershed of the TP during January 2001–December 2012. Subsequently, their utility in streamflow prediction was compared by driving VIC model. Last, the changes of extreme precipitation and extreme streamflow were analyzed.

The main results are summarized as follows:

1. Compared to the CMA precipitation, the four SGPEs could generally captured the spatial distribution of precipitation well in spite of the underestimation in the western mountains and overestimation in the southeast which is located in a lower elevation. Overall, CDR overrated the precipitation in basin scale while 3B42 performs best. However, the two CMORPH (CRT and BLD) agreed well with CMA in time series of watershed average precipitation in both the calibration and verification periods.

2. The spatial pattern of the extreme precipitation was similar to that of daily average precipitation with the precipitation amount increasing from the northwest to the southeast. The disastrous heavy rain mainly occurred in the southeast corner of the basin. Also, 3B42 and CDR overestimated the extreme precipitation, especially in the southeast, while CRT and BLD were closer to CMA in the distribution of extreme precipitation.

3. Notice that none of the four SGPEs performed better than CMA in hydrologic utility. However, BLD performed fairly well and showed comparable hydrologic utility with CMA over UYRB, and CRT and 3B42 showed an acceptable performance. In contrast, CDR is equipped with little potential for the streamflow simulation with wildly overrating the discharge in flood season. This is closely related to the overreaction of the extreme precipitation over the southeastern part of UYRB.

4. It can be seen that the four SGPEs showed well performance in the 2005 floods event, while they all exhibited poor performance in matching the hydrograph of the 2012 flood event which was a disastrous flood for the longyangxia reservoir. The simulation results of 2012 flood event indicated that maybe there exist large errors in SGPEs for a few rather large torrential rain events which could generate errors in estimating flood peaks, peak times and flood volume. Hence, it should be used with caution for the SGPEs in simulating massive flood events over UYRB.

Acknowledgments: The authors wish to acknowledge the reviewers of the manuscript for their constructive comments and helpful suggestions. This research is supported by National Key Research and Development Program (grant number: 2016YFC0400909, 2016YFA0601504); NNSF (grant numbers: 41371049, 41571015, 41323001, and 51539003); and the Project of the State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, China (grant number: 20165042612).

Author Contributions: All authors contributed extensively to the work presented in this paper. Haishen Lu and Yonghua Zhu designed the framework of this study. Jianbin Su analyzed the data and wrote the paper. Haishen Lu and Yonghua Zhu revised the paper. Jianqun Wang and Ali M. Sadeghi helped check the English language and style.

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


3. Rozalis, S.; Morin, E.; Yair, Y.; Price, C. Flash flood prediction using an uncalibrated hydrological model and radar rainfall data in a Mediterranean watershed under changing hydrological conditions. J. Hydrol. 2010, 394, 245–255. [CrossRef]


17. Mei, Y.; Nikolopoulos, E.I.; Anagnostou, E.N.; Borga, M. Evaluating satellite precipitation error propagation in runoff simulations of mountainous basins. J. Hydrometeorol. 2015, 17. [CrossRef]

18. Gao, Y.; Liu, M. Evaluation of high-resolution satellite precipitation products using rain gauge observations over the Tibetan plateau. Hydrol. Earth Syst. Sci. 2013, 17, 837–849. [CrossRef]

32. Duan, Z.; Liu, J.; Tuo, Y.; Chioagn, 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]
44. Xie, Z.; Yuan, F.; Duan, Q.; Zheng, J.; Liang, M.; Chen, F. Regional parameter estimation of the vic land surface model: Methodology and application to river basins in China. *J. Hydrometeorol.* 2007, 8, 447. [CrossRef]
47. Wang, W.; Chen, X.; Shi, P.; Phajmvan, G. Detecting changes in extreme precipitation and extreme streamflow in the dongjiang river basin in southern China. *Hydrol. Earth Syst. Sci.* 2008, 12, 207–221. [CrossRef]

© 2017 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/).