AMERICAN
METEOROLOGICAL
SOCIETY

Journal of Hydrometeorology

EARLY ONLINE RELEASE

The DOI for this manuscript is doi: 10.1175/JHM-D-14-0105.1

Multi-scale Hydrologic Applications of the Latest Satellite Precipitation Products in the Yangtze River Basin using a Distributed Hydrologic Model

Zhe Li[1], Dawen Yang[1], Bing Gao[2], Yang Jiao[1], Yang Hong[1,3], Tao Xu[4]

[1] State Key Laboratory of Hydroscience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China

[2] School of Water Resources and Environment, China University of Geosciences (Beijing), Beijing 100083, China

[3] Department of Civil Engineering and Environmental Science, University of Oklahoma, Norman, OK 73072, United States

[4] Three Gorges Water Cascade Dispatch and Communication Center, Yichang 443133, China

Revised Manuscript for the ***Journal of Hydrometeorology***

*Corresponding author address:

Dawen Yang, Room 312, New Hydraulic Engineering Building, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China

Tel.: (+86)-10-62796976; Email: yangdw@tsinghua.edu.cn

Abstract:

The present study aims to evaluate three global satellite precipitation products (3B42 V7, 3B42 RT and CMORPH) during 2003-2012 for multi-scale hydrologic applications, including annual water budgeting, monthly and daily streamflow simulation, and extreme floods modeling, via a distributed hydrological model in the Yangtze River basin. The comparison shows the 3B42 V7 data generally has a better performance in annual water budgeting and monthly streamflow simulation, but this superiority is not garneted for daily simulation, especially for floods monitoring. It is also found that, for annual water budgeting, the positive (negative) bias of 3B42 RT (CMORPH) estimate is mainly propagated into the simulated runoff, and simulated evapotranspiration tends to be more sensitive to negative bias. Regarding streamflow simulation, both near-real-time products show region-dependent bias: 3B42 RT tends to overestimate streamflow in the upper Yangtze River; in contrast, CMORPH shows serious underestimation in those downstream sub-basins, while it is able to effectively monitor streamflow into the Three Gorges Reservoir. With 394 selected flood events, the results indicate that 3B42 RT and CMORPH have competitive performances for near real-time floods monitoring in the upper Yangtze, but for those downstream sub-basins, 3B42 RT seems to perform better than CMORPH. Furthermore, the inability of all satellite products to capture some key features of the 2012 July extreme floods reveals the deficiencies associated with them, which will limit their hydrologic utility in local floods monitoring.

Key Words: Satellite Precipitation Products; Hydrologic Applications; Evaluation; Distributed Hydrological Model; the Yangtze River

1. Introduction

The current era of satellite remote sensing has provided unprecedented opportunities for the monitoring and prediction of Earth's terrestrial water conditions (Wood et al., 2011). As a part of the effort to establish Earth Observing System, a growing number of satellite-based quasi-global precipitation datasets have been developed and released routinely during the past decade (Ebert et al., 2007). These multi-sensor blended estimators, including PERSIANN (Sorooshian et al., 2000), CMORPH (Joyce et al., 2004), PERSIANN-CCS (Hong et al., 2004), NRL-Blend (Turk and Miller, 2005), TMPA (Huffman et al., 2007), and GSMap (Kubota et al., 2007), are designed to provide high spatial (≤ 0.25º) and temporal (≤ 3hours) resolution forcing datasets to facilitate regional and global investigations into the weather, climate and hydrology, especially for those natural hazards trigged by precipitation extremes, such as floods, landslides and droughts (Hong et al., 2006; Aragão et al., 2007; Wu et al., 2014).

Despite the continuing great efforts to develop high-resolution satellite precipitation products, the remotely sensed datasets are always subject to significant error sources, arising from indirect measurements, retrieval algorithms and sampling uncertainties (e.g. McCollum et al., 2002; Najssen and Lettemaier, 2004; Villarini et al., 2009). Furthermore, the error characteristics vary in different climatic regions, storm regimes, seasons and surface conditions, suggesting region-specific evaluation of the quality and applicability of these satellite precipitation products is necessary (Sorroshian et al., 2011). In general, the evaluation work can be implemented through two strategies: (1) statistics-based direct quantification, and (2) modeling-based indirect inference. The former one carries out the straightforward comparison of the satellite estimates with the ground rainfall data (e.g. gauge or weather radar) according to a set of meaningful

diagnostic statistics, which has been the subject of numerous validation studies (e.g. Tian et al., 2007; Ebert et al., 2007; Gosset et al., 2013). In contrast, the rationale for the latter approach is application-oriented with a focus on the applicability of satellite products to hydrological modeling. Model, as the surrogate of natural watersheds, can diminish (or amplify) and propagate input errors into other simulated fluxes or states, reflecting useful information to the hydrologists who use them to drive the simulation of terrestrial water cycle (Pan et al., 2010).

With this application-oriented view, many studies have investigated the performance of satellite precipitation products by hydrological modeling in specific watersheds around the world (e.g. Su et al., 2008; Li et al., 2009; Bitew and Gebremichael, 2011; Gebregiorgis et al., 2012; Yong et al., 2012), indicating that there is an increasing potential to use these products in hydrological modeling with continuing upgrades of retrieval algorithms (Su et al., 2008; Yong et al., 2012). In future, it is expected that the recently launched Global Precipitation Measurement (GPM) mission will further improve flood monitoring in medium to large river basins substantially (Hossain and Lettenmair, 2006). However, these products are rarely used in operational hydrologic applications (Bitew and Gebremichael, 2011), perhaps partly due to the suspicion of their accuracy for operational use by the hydrologic prediction community, and partly because of the lack of long-term systematic assessment with well-calibrated modeling system to demonstrate their applicability.

In China, rainstorm is the major cause of riverine floods, which is the most frequent and significant type of natural disaster in summer and autumn, ranging from local to regional scales (Ding and Zhang, 2009). Compared with conventional rain gauges, satellites provide an alternative way to monitor large-scale rainfall dynamics in time.

To our knowledge, many previous studies in China have been conducted to analyze the performance of satellite precipitation data in the statistics-based manner (Zhou et al., 2008; Shen et al., 2010; Gao and Liu, 2013; Hu et al., 2014), while modeling-based investigations are only available in very few medium-sized watersheds in China (Li et al., 2012; Jiang et al., 2012; Yong et al., 2012; Tong et al., 2014). Obviously, without comprehensive hydrological evaluation work emphasizing on the application purpose over diverse regions, especially for the large-scale basins where complicated hydrometeorological regimes and land surface conditions exist, the efforts to enhance disaster predictability using satellite data will not bear results.

The present work focuses on the modeling-based evaluation of satellite precipitation data in the Yangtze River, the largest river of China, to bridge the gap between remote sensing technologies development community and hydrologic prediction community. We extend our previous 4-year statistical analysis (Li et al., 2013) into a decade-long (2003-2012) evaluation via a physically-based distributed modeling framework of the Yangtze, aiming at the following essential goals. First, a physically-based distributed model is established over the whole Yangtze River, to serve as the fundamental tool for quantitative prediction as well as the critical platform for integrating increasing amount of remote sensing data. Second, using popular satellite precipitation products and traditional gauge-based estimates, we compare the simulated hydrological fluxes and states with different forcing data across the Yangtze River basin, and implement a long-term systematic evaluation to discuss the potential and challenges for hydrologic applications of latest satellite precipitation datasets, including annual water balance analysis, streamflow simulation and extreme flood modeling. We believe this decadal investigation will provide valuable information to the end-users who are interested in water resources management, reservoir operation and natural hazards early warnings

(Gosset et al., 2013). Finally, for a better understanding of the error propagation when using satellite product as input, we also discuss this issue via water balance analysis, examining the relationship between the bias of rainfall estimation and the bias of other simulated components of water cycle (e.g. evapotranspiration, runoff), as the water balance analysis is claimed to be another important indicator for testing the validity of precipitation data (Li et al., 2012).

The remaining part of this paper is organized as follows: Section 2 describes the study area and the datasets applied in this study. Section 3 briefly describes the distributed hydrological model, including the setup, calibration and validation of the model, and modeling-based evaluation method. The results are presented and discussed in Section 4. Finally, Section 5 summarizes major conclusions of this study.

2. Study area and data

a. Yangtze River basin

The Yangtze River basin with a drainage area of 1.8 million km^2 is the largest river basin in China. It originates from the Tibetan Plateau, and flows eastward for more than 6300 km before draining into the East China Sea (Fig.1a). The Yangtze River spans near one-fifth of whole China mainland, it experiences diverse landforms and complicated hydroclimatic conditions affected by both East and South Asian monsoon activities (Changjiang Water Resources Commission, 1997). As a consequence, this basin suffers frequent floods during the warm season from April to September.

[Figure 1 is to be inserted here]

In order to evaluate regional performance of satellite precipitation products, we divide the whole basin into eight hydrological sub-regions (Fig.1b), according to its drainage system (Changjiang Water Resources Commission, 1999). From the upstream to the downstream, called Jinsha River, Min and Tuo Rivers, Jialing River, Wu River and Three Gorges Region, Han River, Dongting Lake river system, Poyang Lake river system, and the Middle and Lower Yangtze mainstream, respectively (see Table 1).

b. Ground gauge data

The daily meteorological observations from 1961 to 2012 are obtained from 141 China Meteorological Administration (CMA) gauges (circles in Fig. 1a), including daily precipitation, wind speed, maximum/minimum/mean air temperature, relative humidity, and hours of sunshine.

The daily discharge records (1961-2011) are collected from the Hydrological Year Book published by the Hydrological Bureau of the Ministry of Water Resources of China, and discharge data of 2012 is separately obtained from China Three Gorges Corporation. In this study, 12 streamflow gauges located in the major tributaries and along the Yangtze mainstream are selected (squares indexed with numbers in Fig. 1a; also see Table 1).

[Table 1 is to be inserted here]

c. Satellite precipitation data

According to previous statistical evaluation work in the same basin by Li et al. (2013), we select three sets of satellite products as the forcing data for distributed modeling over the Yangtze River: the latest released version 7 of TMPA data (the post-real-time product and near-real-time product are hereinafter referred to as 3B42 V7 and 3B42 RT, respectively) and CMORPH. All these products are generated by combining

information from both passive microwave (PWM) and infrared (IR) observations with high spatial (0.25^{o}) and temporal (3 hours) resolutions. The three sets of precipitation estimators are all available during the study period of 2003-2012.

d. Other geographical information

The basin-scale geographical information is extracted from a series of global datasets. The digital elevation model (DEM) data is acquired from the USGS HydroSHEDS data (http://hydrosheds.cr.usgs.gov/index.php) at 100-m resolution. For land use/land cover (LULC) data, it is obtained from the Environmental and Ecological Science Data Center of West China (http://westdc.westgis.ac.cn/), consisting of three images from different periods (1980s, 1990s and 2000s) at 100-m resolution. The soil type is obtained from the digital soil map of the world (Food and Agricultural Organization, 2003) at 10-km resolution, while the corresponding soil properties (e.g. the porosity, the saturated hydraulic conductivity) and other soil water parameters can be estimated from IGBP-DIS Global Soil Database (http://www.daac.ornl.gov). The vegetation dynamics are described by monthly Leaf Area Index (LAI), derived from the 8 km Global Inventory Modeling and Mapping Studies (GIMMS) NDVI products, which are also archived from the Environmental and Ecological Science Data Center of West China.

3. Modeling framework

In the present study, a geomorphology-based hydrological model (GBHM; Yang et al., 2003; Cong et al., 2009; Li et al., 2012; Ryo et al., 2014) is applied to the whole Yangtze River basin to simulate the hydrological processes with various precipitation inputs. This distributed modeling framework takes advantages of the geomorphologic

similarities to reduce the spatial-structure complexity within a grid, and characterize the catchment topography by hillslope-stream formulation. In brief, GBHM includes following components: a gridded discretization scheme, a sub-grid parameterization scheme, a hillslope-based hydrological modeling module, and a kinematic wave flow routing module.

a. Model setup

Hillslope is the computational unit for hydrological simulation in GBHM (Fig. 2), and the hillslope-based hydrological modeling includes the following processes: snowmelt, canopy interception, evapotranspiration, infiltration, surface and subsurface flow, and the water exchange between groundwater and the river channel (Yang et al., 2003). A detailed description of GBHM physical representations can be found in several recent papers (Cong et al., 2009; Ryo et al., 2014). All parameters used in the GBHM of the Yangtze River are summarized in Table 2.

[Figure 2 is to be inserted here]

The runoff generated at the hillslope is the lateral inflow into the river channel (Fig. 2) within the flow interval. Flow routing by the kinematic wave approach is then applied to each flow interval sequentially to get the discharge at the outlet of river basin.

[Table 2 is to be inserted here]

Considering the computational capacity and the available geographical datasets, we construct GBHM of the Yangtze River basin using the grid size of 10 km. The whole Yangtze River is then divided into 137 sub-basins organized by the Pfafstetter system (Yang et al., 2003), and all the grids within a sub-basin are classified into a series of flow intervals. Within a flow interval, the grids are grouped into different categories

of hillslopes by the sub-grid parameterization procedure. The hillslopes are identified according to land cover types and topographical characteristics, which are calculated from the geographical information data with finer resolutions (e.g. 100m DEM, 100m LULC).

The 10-km meteorological forcing fields are interpolated from 141 CMA gauges: the daily precipitation, wind speed, relative humidity and sunshine hours are interpolated using an angular distance weighing method (New et al., 2000), whereas the maximum, minimum, and mean temperature are interpolated via an elevation-corrected angular direction weighing method (Yang et al., 2003).

It is known that initial condition plays a critical role in hydrologic prediction, and thus we implement a 10-year warm-up simulation to determine it. A predefined initial field is given at the beginning of warm-up period, and then at the end of 10-year warm-up the simulated soil moisture and groundwater table will approximate the real condition, which is recorded to preset the initial condition for future simulations.

b. Model calibration and validation

As shown in Table 2, there are three GBHM parameters in total (snowmelt factor, groundwater hydraulic conductivity, specific yield for unconfined aquifer) that need to be calibrated manually. Considering the available data and recent human activities impacts, we use daily discharge data during the period of 1961-1965 to calibrate the model, and then use data from 1966 to 2002 for a long-term validation. It is noted that simulations during both calibration and validation periods are forced by the CMA gauge interpolated forcing data.

To quantitatively evaluate the model performance, the Nash-Sutcliffe coefficient of efficiency (*NSCE*; Nash and Sutcliffe, 1970) and the relative error (*RE*) are calculated:

$$NSCE = 1 - \frac{\sum (Q_{obs,i} - Q_{sim,i})^2}{\sum (Q_{obs,i} - \bar{Q}_{obs})^2} \quad (1)$$

$$RE = \frac{\sum Q_{sim,i} - \sum Q_{obs,i}}{\sum Q_{obs,i}} \quad (2)$$

where $Q_{obs,i}$ is the observed discharge at the ith day, while $Q_{sim,i}$ is the simulated discharge at the ith day, and $\bar{Q}_{obs}$ is the mean value of the observed discharge series.

Table 3 shows the long-term performance of GBHM model during the period of 1961-2002. It is found that, during calibration and validation periods, *NSCE* value is greater than 0.70 while *RE* is constrained within a reasonable range from -4.0% to 12.0% at most of the streamflow gauges. The result also indicates there is a significant decline of the performance at Huangjiagang during the validation period. This is caused by the impoundment of Danjiangkou Reservoir since 1968 in the upper Han River, and therefore, this gauge is excluded from the following analysis.

[Table 3 is to be inserted here]

c. Modeling-based evaluation

As shown in the results of model calibration and validation, it is demonstrated that the Yangtze River GBHM can provide a reasonable modeling framework to discuss about the pros and cons of satellites precipitation data for prediction applications during the past decade (2003-2012).

The three sets of satellite precipitation products (3B42 V7, 3B42 RT and CMORPH) are integrated from their original 3-hour scale into daily scale first to match the local time, and then daily precipitation fields are projected and resampled into the 10-km modeling grid system.

With exactly the same initial condition (given by the long-term modeling of 1961-2002), GBHM is then driven by the gauge interpolated fields, 3B42 V7, 3B42 RT and CMORPH, respectively. Based on the 10-year inputs and simulated results, including the precipitation, runoff, actual evapotranspiration, and discharge in the river channels, the utilities of the latest satellite precipitation estimators to hydrologic applications are further presented and discussed in the following text.

4. Results and discussions

In this section, for a better understanding of the modeling results with different input datasets, statistical comparison of various precipitation estimates is first revisited; then comprehensive multi-scale evaluation based on hydrological modeling is discussed, focusing on three elements from the perspective of hydrologic applications: (1) water balance analysis for water resources assessment; (2) streamflow simulation for river management; and (3) near real-time flood forecasting for natural disasters warning and mitigation.

a. Comparison of satellite precipitation with the gauge estimates

Figure 3 shows the spatial maps of gauge observed annual precipitation (mm) and mean bias of annual precipitation (mm day^{-1}) between different satellite estimates and gauge estimate during 2003-2012. In general, precipitation increases from the west to the east of the basin, ranging from 348 mm (over the Tibetan Plateau) to 1700 mm (over the Dongting Lake river system and Poyang Lake river system), although there is an obvious enhancement of precipitation over the Sichuan Basin. Looking into the spatial distribution of bias, it is clear that 3B42 V7 shows the closest agreement with the gauge estimate, indicating the critical role of monthly gauge correction algorithm

for bias removal. Comparatively, other two near-real-time estimators without gauge correction (3B42 RT and CMORPH) have evident local bias (Figs. 3c, d).

In the upper Yangtze River, 3B42 RT overestimates (by >1 mm day^{-1}) precipitation while CMORPH presents a mixed pattern with both negative (-0.2 to -1mm day^{-1}, the western part) and positive (0.5 to 2 mm day^{-1}, the eastern part) bias. This bias pattern is consistent with previous studies (Shen et al., 2010; Li et al., 2013), suggesting there are great uncertainties associated with satellite-based precipitation retrievals over the Tibetan Plateau. As 3B42 RT is developed by PMW-calibrated IR technique, perhaps the positive biases can be attributed to the dominant usage of IR images without sufficient calibration by PMW data, considering most PMW overpasses have been screen out due to the snow covers and complex terrains in the Tibetan Plateau (Yong et al., 2014). In contrast, as CMORPH utilizes a different technique which propagates PMW estimates by IR-derived advection vectors, its bias pattern has be contaminated by uncertainties associated with both IR and PMW measurements, and thus it shows more complicated error features over there. When focusing on the middle and lower Yangtze River, CMORPH seems to underestimate precipitation consistently (by -0.2 to -2 mm day^{-1}), but 3B42 RT shows a similar mixed error pattern as 3B42 V7.

[Figure 3 is to be inserted here]

Following the error component decomposition scheme proposed by Tian *et al*. (2009), we also decompose the total bias E into three independent parts (i.e. hit bias H, false precipitation F and missed precipitation $-M$, related by $E=H+F-M$) at seasonal scale to understand the error sources and their evolution. Following our previous study (Li et al., 2013), in the Yangtze River, we define the spring as January–March (JFM), the summer as April–June (AMJ), the autumn as July–September (JAS), and the winter as

October–December (OND). This seasonal division scheme is consistent with the flood season (AMJ and JAS) in the Yangtze and is thus useful for studying the hydrological cycle of this region. Due to the limited space, we only take April-June and October-December to represents the warm season and cold season, respectively (as in the same season, the error pattern tends to be similar during different months, see Li et al., 2013). Figure 4 and Figure 5 present the spatial maps of multi-year (2003-2012) averaged error components for the warm season (AMJ) and the cold season (OND), respectively.

In the warm season (Fig. 4), the TMPA series estimators (3B42 V7 and 3B42 RT) share considerable similarities in their spatial distributions of the total bias and its components over most part of the Yangtze River, except the upper reaches. Over the Jinsha River, 3B42 RT has remarkable overestimation (>100 mm), which are mainly attributed to positive hit bias combined with several false alarms. However, the mixed error pattern in the middle and lower Yangtze is dominated by both positive hit bias (60 to 200 mm) and negative missed precipitation (-60 to -200 mm) for 3B42 RT. Due to the complex precipitation regime itself over the Yangtze River shaped by monsoon climate as well as complex terrains, we suspect the missed events may be caused by the inability to catch warm rainfall or short-lived convective storms, and the positive hit biases perhaps can be related to the overestimation by PMW land algorithm for the convectively active regimes (Tian et al., 2009). In addition, the discrepancies between the hit bias (false precipitation) maps of 3B42 V7 and 3B42 RT demonstrate the effect of gauge-based adjustment, but unfortunately, the correction procedure cannot recover undetected events to further reduce the amounts of missed precipitation. In contrast, CMORPH presents a different error pattern from the TMPA series: its total biases are dominated by underestimation (-30 to -200 mm), with only a part of the upper reaches

along the eastern edge of the Tibetan Plateau covered by overestimations. Considering the complex terrains in that region (Fig. 1a), we speculate this is mostly linked to the anomalously high PMW rainfall contaminated by snow cover in the surrounding high mountains.

For the cold season (Fig. 5), the total bias of all the three products are still dominated by both hit bias and missed precipitation. The most obvious common feature shared by them is that the most serious biases concentrated largely over the lower Yangtze River, as the precipitation bands has "shifted back" to this area due to the movement of the East Asian monsoon during the winter. Over the upper and middle Yangtze, the precipitation amount during this season is very limited, but still, there is a long-lasting overestimation for 3B42 RT and CMORPH, respectively. The former is caused by hit bias, which is similar as the situation in the warm season; while the latter, unlike the warm season, is evidently caused by false alarms. Qualitatively, we also note there are pixels with obviously high false-alarm-rate (the false precipitation amount over these pixels is more than twice the values of their surrounding pixels) over the lower Yangtze for all products (see box 1 and box 2 in Fig. 5; as estimated precipitation is much lower than TMPA series, CMORPH does not show these pixels when its map is scaled into the same color bar). It is found that these pixels correspond to the locations of Poyang Lake and Dongting Lake, and this falsed rainfall-like signal over the inland water bodies is claimed to be caused by the deficiencies of PMW-based retrievals for emissivity characterization (Tian and Peters-Lidard, 2007). Since this effect only exits at a relatively small spatial scale (~1000 km^2), it is believed to have little impacts on basin-scale hydrologic modeling.

[Figure 4 is to be inserted here]

[Figure 5 is to be inserted here]

At daily scale, we further compare basin-scale rainfall interpolated by gauges against that estimated based on satellites in Figure 6. Several statistical metrics are calculated, including relative bias (*RB*), root mean square error (*RMSE*) and Pearson's correlation coefficient (*CC*). Generally, all three satellite estimators show strong correlation with gauges estimation results, interpreted by the high values of *CC* (0.62-0.86). According to the relative bias, 3B42 V7 gives fairly good approximation to gauge estimates with slight overestimation (*RB* lies within the range of 0-10%) over most sub-regions. In summary, estimated biases are smaller in Jinsha River, Min and Tuo River, Wu and Three Gorges and Han River, compared to other sub-regions over the lower Yangtze River. As mixed error pattern exits at these regions (Figs. 3-5), the positive biases and negative ones will cancel each other out and make basin-scale estimates unbiased or less biased.

Without gauge-adjustment, 3B42 RT is found to have overestimated daily basin-scale rainfall over the whole Yangtze River, especially in the upper Yangtze (*RB* is 75.67%, 36.96%, and 18.28% for Jinsha River, Min and Tuo River, and Jialing River, respectively). At the same time, another unadjusted near real-time product, CMORPH, is found to work as the complementary dataset for 3B42 RT, slightly underestimating basin-scale rainfall over the upper Yangtze and seriously underestimating in the lower reaches (*RB* is -31.44%, -38.79%, -35.13% and -31.41% for Wu River and Three Gorges region, Dongting Lake region, Poyang Lake region and the Lower Yangtze, respectively). Combined with the error features characterized at annual and seasonal scale (Figs. 3-5), the daily result also suggests that there is a consistent overestimation (underestimation) across multiple time scales for 3B42 RT (CMORPH).

Additionally, it is demonstrated 3B42 V7 does not always show its superiority over other products at daily scale, in particular CMORPH, in terms of *RMSE* and *CC* (e.g. Figs. 6b-d). This implies, although the monthly satellite-gauge (SG) combination algorithm tends to make 3B42 V7 statistically closer to monthly gauge observation (to minimize the bias), there is no guarantee for the improvement of daily precipitation estimates within a month (Li et al., 2013; Yong et al., 2014).

[Figure 6 is to be inserted here]

b. Analysis of annual water balance simulation

Figure 7 compares averaged simulation results of annual water balance components (basin-scale precipitation, evapotranspiration and runoff) over the 8 hydrological sub-regions in the Yangtze River during the study period (2003-2012).

[Figure 7 is to be inserted here]

First of all, this result clearly provides a general idea on annual water budgeting of the Yangtze River: Precipitation and evapotranspiration increase from the upper reaches to the lower reaches, with an amount of 700-1600 mm and 400-800 mm, respectively. However, runoff, ranging from 300 to 800 mm, shows more complicated pattern as it is controlled by the variations of runoff coefficient. The gauge-based modeling results are also consistent with many previous studies for regional water budget (Gao et al., 2007; Xu et al., 2008).

To quantitatively characterize the error for water balance simulation and understand how the input error propagates into other simulated hydrologic fluxes and states, we calculate the relative bias (%) by dividing the bias of satellite-driven simulation by the gauge-driven modeling result (for each water balance component). As expected from

the comparisons discussed above, 3B42 V7's result shows the closest agreement with gauge-based simulation. Overall, this dataset can provide very reasonable simulated results of evapotranspiration (with relative bias from -1.8% to 4.9%) as well as runoff (with relative bias from -1.9% to 6.8%). It is demonstrated that 3B42 V7 should be the most appropriate dataset for long-term regional water balance studies.

As both 3B42 RT and CMORPH show region-dependent bias for rainfall estimation, and hence the error will be propagated into other simulated variables correspondingly. In summary, 3B42 RT tends to substantially overestimate runoff in those upstream basins (e.g. Jinsha River, Min and Tuo River; Figs. 8a, b) while CMORH seriously underestimates it in the downstream sub-regions (e.g. Wu and Three Gorges Region, Han River, Dongting Lake and Poyang Lake region, Middle and Lower mainstream; Figs. 8d-h). Furthermore, by comparing the relative bias of the estimated precipitation, simulated evapotranspiration and simulated runoff among various products, we can provide implications on errors propagation from the input forcing into other simulated hydrologic fluxes and states.

Comparing modeling results from 3B42 RT and gauge, we find that basin-scale actual evapotranspiration presents no substantial change (relative bias varies between -1.1% and 4.4%) with increased precipitation over all regions, reflecting the energy-control nature without water stress for evapotranspiration over the whole Yangtze during the warm season. As the total precipitation bias is dominated by errors in summer (Li et al., 2013), thus the positive bias in 3B42 RT estimate has been mainly propagated into simulated runoff bias. Looking into the relative biases of precipitation and runoff, we can also find the relative bias has been enhanced through hydrologic modeling, due to the nonlinear nature of watershed rainfall-runoff processes. Roughly, from rainfall to

runoff, the bias has been multiplied by a factor of 2. At the same time, comparison between the results of gauge and CMORPH indicates that reduced inputs will cause decreases in simulated runoff and in evapotranspiration simultaneously, suggesting evapotranspiration tends to be more sensitive to negative bias in precipitation over the Yangtze. Similarly, the negative relative bias in precipitation has also been enhanced by a factor around 1.5 for most sub-regions.

c. Evaluation of streamflow simulation

Figure 8 presents the simulated monthly streamflow against observations over major tributaries of the Yangtze, and the performance of both monthly and daily simulations is also quantified by *NSCE* and *RE* in Table 4. It is evident that, 3B42 V7, the gauge-adjusted rainfall data, works fairly well to capture the general streamflow dynamics of all tributary basins with a medium size (8-46×10^4 km^2). Comparing with gauge-driven results in Table 4, we can find that there is a substantial decline of model performance when we use 3B42 V7 as the forcing data for daily simulation in the downstream tributaries (e.g. Yuan River, Gan River), though it shows good agreement with gauge-based result at monthly scale. This can be explained by Figure 6, which shows more scattered and less correlated daily rainfall estimates by 3B42 V7 in the downstream sub-basins (Figs. 6f-h), indicating the challenge for satellite data to adequately capture short-lived heavy rainstorms there. As for another two sets of near-real-time product, 3B42 RT and CMORPH, their streamflow simulation results also present remarkable region-dependent errors, which have close correspondence to the local bias contained in precipitation fields (Figs. 3-5). 3B42 RT performs to have seriously overestimated streamflow in the upstream sub-basins (e.g. Jinsha River, Min and Tuo River) during the warm season, which can be traced back to large positive total bias in precipitation (Fig. 4); as a result of the significantly underestimated precipitation, CMORPH cannot

reproduce flood events in the summer for the sub-basins over the middle and lower Yangtze (e.g. Wu and Three Gorges region, Yuan River, Gan River). In other words, this suggests 3B42 RT can get better streamflow modeling results in those midstream and downstream sub-basins while CMORPH should be applied to the upstream sub-basins to get reasonable simulations. However, considering the mixed error pattern in the Yangtze River discussed above (for CMORPH, mainly in Jinsha River and Min and Tuo River; and for 3B42 RT, mainly in the middle and lower Yangtze, see Fig. 3 and Fig. 4), therefore, special caution should be taken when 3B42 RT and CMORPH are applied to streamflow modeling at catchments with smaller scale compared to sub-basins discussed in this study, since the local positive bias and negative bias perhaps cannot cancel each other out. For instance, in the very upstream of the Yangtze River, CMORPH still shows a dominant negative bias (Fig.3d), and as a result, the simulated streamflow is largely underestimated over the source region of the Yangtze (Tong et al., 2014).

[Figure 8 is to be inserted here]

[Table 4 is to be inserted here]

In addition, Figure 9 compares the simulated monthly streamflow with observations along the mainstream. This result, combined with those summary statistics (Table 4), shows again the performance of 3B42 V7 is in exceptionally good agreement with the observation from the upper to the lower Yangtze mainstream. As the drainage area becomes larger for these mainstream stations (87-170×10^4 km^2), the simulation results indicate that modeling performance will be better with larger watershed size. As the general pattern of 3B42 RT and CMORPH estimated rainfall is dominated by positive and negative bias, respectively (Fig. 3), and therefore, 3B42 RT tends to generally

overestimate streamflow (with *RB* from 80.7% to 144.3%) along the main Yangtze while CMORPH underestimates it overall (with *RB* from -23.1% to -45.7%). It is also noted that, at Cuntan station in the upper Yangtze, CMORPH is able to get reasonable results with moderate underestimation for streamflow modeling (*NSCE* value is 0.64 and *RB* is around -20% for daily results). As Cuntan is the inflow gauge to the Three Gorges Reservoir, thus CMORPH offers an alternative way to effectively monitor the streamflow poured into the reservoir in a near real-time manner, which can potentially provide useful information for the regulation of Three Gorges Dam.

[Figure 9 is to be inserted here]

To get a better understanding of the full spectrum of daily streamflow regime, we further compare the simulated streamflow against observation at several gauges in the form of flow duration curves (Fig. 10). Evidently, this figure provides illustration on how the simulated flow regime changes with different river flow levels among various input datasets.

Similarly, the results confirm the region-dependent bias for 3B42 RT and CMORPH again. However, additional information can be found: the discrepancies of CMORPH-driven simulation over the upstream basins (Figs. 10a, b) mainly come from high flow conditions (when the exceedance probability falls below 40%), but it works similarly as gauge for simulation of low flow conditions; in contrast, 3B42RT tends to seriously overestimate it at full range of the streamflow regime in the upper Yangtze River. For the downstream basins, regime reproduced by 3B42 RT becomes very similar to that simulated by both 3B42 V7 and gauge (Figs. 10c, d), which can be inferred from their similar spatial error pattern over this region (Figs. 3-5), but the significant negative biases attached to CMORPH estimations will distort the flow regime seriously at full

range. Moreover, we also noted that, for the stations over Dongting River region and Poyang Lake region (taking WZ as an example here, Fig. 10d), all satellite estimates as well as the gauge estimates, cannot reproduce the observed flow regime accurately for the medium to low flow conditions.

[Figure 10 is to be inserted here]

d. Applicability for near real-time flood forecasting

Finally, we extend our investigations into the applicability of satellite rainfall products to flood forecasting, especially for near real-time monitoring of floods in the Yangtze River. As suggested by previous global evaluation research (Wu et al., 2012), we also apply the percentile-based (95th percentile value of the flow duration curves) method to the observed streamflow data, and eventually select 394 typical flood events in total for this study.

Figure 11 summarizes the diagnostic statistics calculated by comparing the observed flood events and corresponding simulations. Here we flag the flood events as detected only when the absolute relative bias of simulated discharge is no more than 20%, and the ratio of the detected events to total flood events is then defined as the probability of detection. Obviously, 3B42 V7 outperforms other two satellite products for flood forecasting in terms of all three statistics, especially regarding that it has remarkably reduced and stabilized the bias to yield almost exactly the same performance as the gauge (Fig. 11a). In addition, we can find the probability of detection for both gauge and 3B42 V7 will drop from ~0.7 to ~0.3 as the drainage area decreases (Fig. 11a), indicating the flood detection capabilities for both gauge and 3B42 V7 seems scale-dependent, as they have a better performance over large-size watersheds.

Considering the potential to further utilize satellite estimates for near real-time floods monitoring, we also compare the performance of 3B42 RT and CMORPH during the selected flood events. Clearly, we can find distinct features for the two data (Fig. 11a): from the upper to the lower Yangtze mainstream, the negative bias in CMORPH's simulation will accumulate (-40% to -50%) while the positive bias in 3B42 RT's simulation will diminish slightly (70% to 60%); over the tributaries, CMORPH's simulations also present increased biases from the west to the east while the bias of 3B42 RT's simulation drops down and stays around 10%. According to other two statistics, it is also found that the performance of 3B42 RT approximates 3B42 V7's over downstream tributaries. In summary, events-based analysis shows 3B42 RT and CMORPH should have competitive performances for near real-time flood forecasting in the upper Yangtze River, but over the downstream tributaries, 3B42 RT definitely performs better than CMORPH. It is emphasized that 3B42 RT has higher probability of detection for flood events, but it generates more false-alarm events simultaneously due to overestimation. This performance is claimed to be better than CMORPH's for near real-time floods monitoring, since false alarmed events have greater values than "undetected" events in the context of disaster early warnings and risk management.

[Figure 11 is to be inserted here]

In 2012 July, the Three Gorges Dam was hit by a record flood (with the largest inflow of 71,200 $m^3 s^{-1}$), which was reported as the biggest one since the establishment of the dam in 2003, even with a higher local river flows than the devastating Yangtze floods in 1954 and 1998 (Wang and Zhang, 2013). As a case study of extreme floods in the Yangtze River, we analyze the performance of satellite precipitation products when they are applied into the modeling of this extreme event. Figure 12 shows the results

driven by different inputs at GC (Min River) and BB (Jialing River), the two flooding source areas of this event, as well as the result at CT, which is the inflow gauge to the Three Gorges Reservoir.

[Figure 12 is to be inserted here]

In general, all the three sets of remote sensing products can reproduce the key pattern of this extreme flood in two tributaries (GC and BB), however, gauge-adjusted 3B42 V7 obviously no longer shows its superiority over other two near real-time satellite products during this event. This result suggests that monthly gauge-based correction scheme doesn't promise a better result for floods simulation as well, and thus calls for a new strategy to calibrate satellite data in a near real-time manner to improve their performance for flood forecasting. Meanwhile, it is apparent that all satellite estimates have missed the July 15-19 floods in Min River, indicating the inherent deficiencies associated with all satellite data. Our investigation on precipitation map shows, this is caused by a localized rainstorm at July 15-16 which has not been captured by satellite products (figure is not shown here). Therefore, it seems that, due to the indirect way of measurement and the sampling issue, there is still great challenge for application of current satellite-based estimates into local floods monitoring.

e. Discussions on gauge data, model calibration and human activities

In the present study, the gridded "ground truth" of rainfall is obtained by interpolation of gauge data. Therefore, uncertainties may arise due to: (1) inadequate sampling by gauge measurement; and (2) interpolation method applied to infer the spatial fields of rainfall. The inherent scale-mismatching associated with gauge-based validation can make the comparison and evaluation problematic potentially (Gebremichael et al., 2003), however, we believe this problem is unlikely to cause substantial errors in the

investigations discussed above, since we mainly focus on the basin-scale hydrologic simulation. It is also found that large temporal integration scales (e.g. monthly, daily) will depress the spatial sampling error and then get reasonable approximation to areal true rainfall (Villarini et al., 2008). On the other hand, we would like to highlight the rationale of "modeling-based evaluation", which emphasizes on using gauge-driven simulations as the benchmark to investigate the applicability of satellite precipitation products for regional hydrological modeling. In practice, these CMA gauges are the most widely used dataset to force hydrological modeling in China; therefore, in this study, we only provide a "relative judgment" on the skills of newly available multi-sensor merged rainfall estimates rather than state the "absolute advantages" of satellite estimates.

As the basic tool for evaluation, GBHM is calibrated by historical gauge observed rainfall and streamflow data, and then all the model parameters are fixed during both validation period (1966-2002) and evaluation period (2003-2012). However, many previous studies (e.g. Yilmaz et al., 2005; Yong et al., 2012; Xue et al., 2013) also suggested a satellite-based recalibration to improve the performance of hydrological modeling when satellite product is applied as input. Nevertheless, we didn't adopt this strategy due to the following reasons: (1) firstly, this study aims at comparatively evaluating hydrological simulations forced by various rainfall inputs. Hence, the same model with fixed parameters are the prerequisites for a meaningful comparison; (2) Second, long-term gauge data records in history promise a wide range of hydrologic conditions to ensure a model's validity (Su et al., 2008); (3) Finally, it is also reported that the improved performance of streamflow modeling by recalibration will work at the expense of distorting other simulated fluxes, such as evapotranspiration (Bitew and Gebremichael, 2011).

Another critical issue needs to be discussed is the impacts of the reservoir projects in the Yangtze River basin and their implications to our evaluation work. For the major tributaries of the Yangtze River basin, we can only identify obvious alternation of the natural river flow in the Han River due to the impoundment of Danjiangkou Reservoir since 1968, thus this sub-region has been excluded from our analysis of streamflow simulation. For the Yangtze mainstream, with a limited literature (Gao et al., 2012), we knew that only the Three Gorges Dam impacts on the lower mains' flow regime from October through next February. The greatest effect was found at the YC gauge (very close to the Three Gorges Dam), and the effect became less at the HK and DT gauges where more water flows into the mainstream from the middle and downstream tributaries. In summary, the overall impacts of the reservoir projects should be limited, especially over those major tributaries and the upper mainstream. However, as human activities exerted increasing impacts on natural hydrologic processes in the Yangtze River, we would like to leave it as an open question to be addressed quantitatively in future studies.

5. Conclusions

With a decade-long (2003-2012) observation datasets, this thorough evaluation aims at assessing the multi-scale hydrologic utilities for three sets of the most popular high-resolution multi-sensor blended global precipitation products (3B42 V7, 3B42 RT and CMORPH) via a physically-based distributed hydrological modeling framework over the Yangtze River, the largest watershed in China. To accomplish this application-oriented evaluation work, we first establish and validate a 10-km distributed GBHM model, then statistically compare different precipitation estimates at multiple temporal scales, and finally examine their utilities in terms of various hydrologic applications,

including annual water balance simulation, streamflow modeling, and near real-time flood monitoring. The major conclusions of this modeling-based evaluation work can be summarized as follows:

(1) For comparisons of precipitation input datasets: In summary, 3B42 V7 shows the closest agreement with gauge estimates in terms of the bias, and comparatively, other two near-real-time estimators present evident local bias. In the upper Yangtze, 3B42 RT seriously overestimates precipitation while CMORPH has a mixed error pattern, all indicating there are great uncertainties for satellite-based precipitation retrievals over the Tibetan Plateau; for the middle and lower Yangtze, 3B42 RT shows a similar mixed error pattern as 3B42 V7's, but CMORPH tends to underestimate precipitation substantially. By decomposition scheme, it is found that the total bias for all satellite estimators are dominated by hit bias and missed precipitation during both warm and clod seasons. Additionally, daily comparison implies 3B42 V7 does not always show superiority over other products at daily scale, suggesting the monthly SG combination algorithm provides no guarantee for improvement of daily precipitation estimates.

(2) For annual water balance simulation: As the most appropriate dataset for regional water budgeting study, 3B42 V7 works fairly well to get results comparable to gauge-driven simulations (with relative bias of -1.8% to 4.9% for evapotranspiration, and -1.9% to 6.8% for runoff). Comparing the results of 3B42 RT and CMORPH, it is also found the bias in precipitation estimates has been mainly propagated into simulated runoff, and simulated evapotranspiration tends to be more sensitive to negative bias.

(3) For streamflow modeling: It is found that 3B42 V7-driven simulation shows fair agreement with observations in those upstream sub-basins at both monthly and daily scale, but its performance declines obviously for daily modeling over the downstream

basins, reflecting the challenges for satellite estimators to adequately capture heavy rainstorms over the lower Yangtze. The results also suggest that 3B42 RT tends to get better modeling results in the midstream and downstream sub-basins while CMORPH can be applied to the upstream sub-basins. However, as the mixed error pattern exits, special caution should be taken when we apply 3B42 RT and CMORPH to modeling at catchments with smaller scale compared to sub-basins discussed in this study, since local positive bias and negative bias perhaps cannot cancel each other out.

(4) For near real-time flood monitoring: with selected 394 flood events during the study period, we find 3B42 RT and CMORPH should have competitive performances for near real-time flood monitoring in the upper Yangtze River, but in the downstream tributaries 3B42 RT performs better than CMORPH. During the extreme flood event in 2012 July, the inability of all products to reproduce key features suggests there are inherent deficiencies associated with current satellite rainfall products when they are applied to the monitoring and warning of local floods.

With the four aspects discussed above, we believe the present study will promote better utilization of satellite precipitation products in various hydrologic applications over the Yangtze River. Clearly, satellite precipitation products provide valuable information to regional water resources assessment, river management and natural hazards warning, and this paper also initially illustrates a demo of a physically-based distributed modeling and forecasting framework over the Yangtze River. Future efforts will be made to complement comprehensive evaluation for floods prediction in small- to medium-size basins, as well as to develop multi-scale multi-source merging techniques to effectively combine ground observations and remote sensing estimates. Furthermore, as the Global Precipitation Mission (GPM) Core Observatory was

successfully launched in February 2014, the modeling framework presented here can be readily employed to benchmark the upcoming GPM-era satellite precipitation data into a regional operational hydrological prediction system over the Yangtze River.

Acknowledgments. This research was supported by the National Natural Science Funds for Distinguished Young Scholar (Project No. 51025931) and the National Natural Science Foundation of China (Project No. 51190092). The first author also acknowledges the HyDROS Lab (http://hydro.ou.edu) at the National Weather Center, Norman, OK for their support during his visiting. The authors wish to thank the two reviewers and Dr. Christa D. Peters-Lidard for their insightful comments.

Reference:

Aragão, L. E. O.C., Malhi, Y., Roman-Cuesta,R. M., Saatchi, S., Anderson, L.O., Shimabukuro, Y. E., 2007. Spatial patterns and fire response of recent Amazonian droughts. *Geophys. Res. Lett.*, **34**, L07701, doi:10.1029/2006GL028946.

Bitew, M. M., and M. Gebremichael., 2011. Evaluation of satellite rainfall products through hydrologic simulation in a fully distributed hydrologic model, *Water Resour. Res.*, **47**, W06526, doi:10.1029/2010WR009917.

Changjiang Water Resources Commission., 1997. Hydrology Research for Three Gorges Project (in Chinese). Hubei science and Technology Press, 313pp.

Changjiang Water Resources Commission., 1999. Atlas of the Changjiang River basin (in Chinese). SinoMaps Press, 286pp.

Cong, Z. T., D. W. Yang, B. Gao, H. B. Yang, and H. P. Hu. 2009. Hydrological trend analysis in the Yellow River basin using a distributed hydrological model, *Water Resour. Res.*, **45**, W00A13.

Ding Y.H. and J. Y. Zhang, 2009. Rainstorms and Floods in China (in Chinese). China Meteorological Press, 290pp.

Ebert, E. E., Janowiak, J. E., Kidd, C., 2007. Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Amer. Meteor. Soc.*, **88**, 47–64.

Food and Agricultural Organization of the U. N. 2003. Digital soil map of the world and derived soil properties, Land Water Digital Media Ser. Rev. 1, Rome.

Gao, G., D. Chen, C.-Y. Xu, and E. Simelton, 2007, Trend of estimated actual evapotranspiration over China during1960–2002, *J. Geophys. Res.*, 112, D11120, doi:10.1029/2006JD008010.

Gao, B., D. Yang, T. Zhao, and H. Yang, H, 2012. Changes in the eco-flow metrics of the Upper Yangtze River from 1961 to 2008. *J. of Hydrol.*, 448, 30-38.

Gao, Y., and M., Liu, 2013.Evaluation of high-resolution satellite precipitation products using rain gauge observations over the Tibetan Plateau, *Hydrol. Earth Syst. Sci*., **17**, 837-849.

Gebregiorgis, A. S., Y.Tian, C. D.Peters-Lidard, and F.Hossain., 2012. Tracing hydrologic model simulation error as a function of satellite rainfall estimation bias components and land use and land cover conditions, *Water Resour. Res.*, **48**, W11509, doi:10.1029/2011WR011643.

Gebremichael, M., Krajewski, W. F., Morrissey, M., Langerud, D., Huffman, G. J., and Adler, R., 2003. Error Uncertainty Analysis of GPCP Monthly Rainfall Products: A Data-Based Simulation Study. *J. Appl. Meteor.*, **42(12)**, 1837-1848.

Gosset, M., Viarre, J., Quantin, G. and Alcoba, M., 2013. Evaluation of several rainfall products used for hydrological applications over West Africa using two high-resolution gauge networks. *Q.J.R. Meteorol. Soc.*, **139**, 923–940. doi: 10.1002/qj.2130

Hong, Y., Hsu, K. L., Gao, X., Sorooshian, S., 2004. Precipitation estimation from remotely sensed imagery using artificial neural network-cloud classification system (PERSIANN-CCS). *J. Appl. Meteorol. Climatol.*, **43**, 1834–1853.

Hong, Y., Adler, R.F., Huffman, G.J., 2006. Evaluation of the potential of NASA multi-satellite precipitation analysis in global landslide hazard assessment. *Geophys. Res. Lett*., **33**, L22402, doi:10.1029/2006GL028010.

Hu, Q., Yang, D., Li, Z., Mishra, A., Wang, Y., Yang, H., 2014, Multi-scale evaluation of six high-resolution satellite monthly rainfall estimates over a humid region in China with dense rain gauges, *Int. J. Remote Sens.*, 35, 1272-1294.

Huffman, G. J., Adler, R. F., Bolvin, D.T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., Wolff, D. B., 2007. The TRMM Multi-satellite Precipitation Analysis (TMPA): Quasi-Global, multi-year, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, **8**, 38–55.

Jiang, S., Ren, L.-L., Hong Y., Yong B., Yang X., Yuan F., Ma M., 2012.Comprehensive evaluation of multi-satellite precipitation products with a dense rain gauge network and optimally merging their simulated hydrological flows using the Bayesian model averaging method. *J. Hydrol.*, **452-453**, 213-225.

Joyce, R. J., Janowiak, J. E., Arkin, P. A., Xie, P., 2004. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.*, **5**, 487–503.

Kubota, T., Shige, S., Hashizume, H., Aonashi, K., Takahashi, N., Seto, S., Takayabu, Y.N., Ushio, T., Nakagawa, K., Iwanami, K., Kachi, M., Okamoto, K., 2007. Global precipitation map using satellite-borne microwave radiometers by the GSMaP Project: Production and validation. *IEEE Trans. Geosci. Remote Sens.*, **45**, 2259–2275.

Li L., Y. Hong, J. Wang, R. Adler, F. Policelli, S. Habib, S., D. Irwn, T. Korme, and L. Okello, 2009. Evaluation of the real-time TRMM-based multi-satellite precipitation analysis for an operational Flood Prediction System in Nzoia Basin, Lake Victoria, Africa. *Nat. Hazards.*, **50**, 109-123.

Li, M., D. Yang, J. Chen, and S. S. Hubbard.2012. Calibration of a distributed flood forecasting model with input uncertainty using a Bayesian framework, *Water Resour. Res.*, **48**, W08510.

Li, X. H., Zhang, Q., Xu, C. Y., 2012. Suitability of the TRMM satellite rainfalls in driving a distributed hydrological model for water balance computations in Xinjiang catchment, Poyang lake basin. *J. Hydrol.*, **426**, 28-38.

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

Maidment, D. R., et al. 1993. Handbook of Hydrology, McGraw-Hill, New York.

McCollum, Jeffrey R., Witold F. Krajewski, Ralph R. Ferraro, Mamoudou B. Ba, 2002: Evaluation of Biases of Satellite Rainfall Estimation Algorithms over the Continental United States. *J. Appl. Meteor.*, **41**, 1065–1080.

Nash, J. E., J. V. Sutcliffe., 1970. River flow forecasting through conceptual models part I-A discussion of principles, *J. Hydrol*., **10**, 282- 290.

New, M., M. Hulme, P. Jones., 2000. Representing twentieth-century space-time climate variability. Part II: Development of a 1961 – 96 monthly grids of terrestrial surface climate, *J. Clim.*, **13**, 2217 – 2238.

Nijssen, B., Lettenmaier, D.P., 2004. Effect of precipitation sampling error on simulated hydrological fluxes and states: Anticipating the Global Precipitation Measurement satellites. *J. Geophys. Res.* **109**, D02103, doi:10.1029/2003JD003497.

Pan, M., H. Li, and E. Wood., 2010. Assessing the skill of satellite-based precipitation estimates in hydrologic applications, *Water Resour. Res.*, **46**, W09535, doi:10.1029/2009WR008290.

Ryo, M., O. C. S. Valeriano, S. Kanae, T. D. Ngoc, 2014. Temporal Downscaling of Daily Gauged Precipitation by Application of a Satellite Product for Flood Simulation in a Poorly Gauged Basin and Its Evaluation with Multiple Regression Analysis. *J. Hydrometeor*, **15**, 563–580.

Shen, Y., Xiong, A., Wang, Y., Xie, P., 2010. Performance of high-resolution satellite precipitation products over China. *J. Geophys. Res.*, **115**, D02114.

Sorooshian, S., Hsu, K. L., Gao, X., Gupta, H., Imam, B., Braithwaite, D., 2000. Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Am. Meteorol. Soc.*, **81**, 2035–2046.

Sorooshian, S., and Coauthors, 2011: Advanced Concepts on Remote Sensing of Precipitation at Multiple Scales. *Bull. Amer. Meteor. Soc.*, **92**, 1353–1357.

Su, F., Y. Hong, D.P. Lettenmaier, 2008. Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and Its Utility in Hydrologic Prediction in the La Plata Basin. *J. Hydrometeor*, **9**, 622–640.

Tian, Y., Peters-Lidard, C.D., Choudhury, B. J., Garcia, M., 2007. Multitemporal analysis of TRMM-based satellite precipitation products for land data assimilation applications. *J. Hydrometeorol.* **8**, 1165–1183.

Tian, Y., and C. D. Peters-Lidard, 2007, Systematic anomalies over inland water bodies in satellite-based precipitation estimates, *Geophys. Res. Lett.*, **34**, L14403, doi:10.1029/2007GL030787

Tian, Y., C. D. Peters-Lidard, J. B. Eylander, R. J. Joyce, G. J. Huffman, R. F. Adler, K. Hsu, F. J. Turk, M. Garcia, and J. Zeng, 2009, Component analysis of errors in satellite-based precipitation estimates, *J. Geophys. Res.*, **114**, D24101, doi:10.1029/2009JD011949.

Tong, K., F. Su, D., Yang, and Z., Hao, 2014, Evaluation of satellite precipitation retrievals and their potential utilities in hydrologic modeling over the Tibetan Plateau. *J. Hydrol.*, **519**, 423-437.

Turk, F. J., Miller, S. D., 2005. Toward improving estimates of remotely sensed precipitation with MODIS/AMSR-E blended data techniques. *IEEE Trans. Geosci. Remote Sens*., **43**, 1059–1069.

Villarini, G., P. V. Mandapaka, W. F. Krajewski, and R. J. Moore, 2008. Rainfall and sampling uncertainties: A rain gauge perspective, *J. Geophys. Res.*, **113**, D11102, doi:10.1029/2007JD009214.

Villarini, G., W. F. Krajewski, and J. A. Smith., 2009. New paradigm for statistical validation of satellite precipitation estimates: Application to a large sample of the TMPA 0.25° 3-hourly estimates over Oklahoma, *J. Geophys. Res.*, **114**, D12106, doi:10.1029/2008JD011475.

Wang, X.F., and S.M. Zhang, 2013. Analysis on characteristics and forecasting of “12.7” Flood in upper Yangtze River Basin, Yangtze River., **44(19)**, 1-4 (in Chinese).

Wood, E. F., et al. 2011. Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, *Water Resour. Res.*, **47**, W05301, doi:10.1029/2010WR010090.

Wu, H., Adler, R. F., Tian, Y., Huffman, G. J., Li, H., and Wang J. 2014, Real-time global flood estimation using satellite-based precipitation and a coupled land surface and routing model, *Water Resour. Res.*,**50**, 2693–2717, doi:10.1002/2013WR014710.

Xu, J., D. Yang, Y. Yi, Z. Lei, J. Chen, and W. Yang, 2008, Spatial and temporal variation of runoff in the Yangtze River basin during the past 40 years. *Quatern. Int.*, **186(1)**, 32-42.

Yang, D., and K. Musiake. 2003. A continental scale hydrological model using the distributed approach and its application to Asia, *Hydrol. Processes*, **17**, 2855–2869.

Ye, X., Q. Zhang, J. Liu, X. Li, and C. Y. Xu, 2013. Distinguishing the relative impacts of climate change and human activities on variation of streamflow in the Poyang Lake catchment, China. *J. of Hydrol.*, **494**, 83-95.

Yong, B., Y. Hong, L.-L. Ren, J. J. Gourley, G. J. Huffman, X. Chen, W. Wang, S. I. Khan, 2012. Assessment of evolving TRMM-based multisatellite real-time precipitation estimation methods and their impacts on hydrologic prediction in a high latitude basin, *J. Geophys. Res*., **117**, D09108.

Yong, B., D. Liu, J. J. Gourley, Y. Tian, G. J. Huffman, L.-L. Ren and Y. Hong, 2014. Global view of real-time TRMM Multi-satellite Precipitation Analysis: implication to its successor Global Precipitation Measurement mission. *Bull. Am. Meteorol. Soc.*, in press

Zhou, T., Yu, R., Chen, H., Dai, A., Pan, Y., 2008. Summer precipitation frequency, intensity, and diurnal cycle over China: A comparison of satellite data with rain gauge observations. *J. Clim*., **21**, 3997 – 4010.

Tables:

Table 1. Streamflow Gauges and Hydrological Regions in the Yangtze River Basin

Location	Gauge Name*[1]	River Name	Drainage Area (10^4 km^2)	Hydrological sub-regions*[2]
Major Tributaries	PS (1)	Jinsha River	45.86	I
	GC (2)	Min River	13.54	II
	BB (3)	Jialing River	15.67	III
	WL (4)	Wu River	8.30	IV
	HJG (5)	Han River	9.52	V
	TY (6)	Yuan River	8.52	VI
	XT (7)	Xiang River	8.16	
	WZ (8)	Gan River	8.09	VII
Mainstream	CT (9)	Yangtze River	86.66	IV
	YC (10)	Yangtze River	100.55	IV
	HK (11)	Yangtze River	148.80	VIII
	DT (12)	Yangtze River	170.54	

***1.** Full names of the gauges: PS: Pingshan; GC: Gaochang; BB: Beibei; WL: Wulong; HJG: Huangjiagang; TY: Taoyuan; XT: Xiangtan; WZ: Waizhou; CT: Cuntan; YC: Yichang; HK: Hankou; DT: Datong.

***2.** Full names of the hydrological sub-regions: I: Jinsha River; II: Min and Tuo River; III: Jialing River; IV: Wu River and Three Gorges Region; V: Han River; VI: Dongting Lake River system; VII: Poyang Lake River system; VIII: Middle and Lower Yangtze Mainstream.

Table 2. Summary of Parameters used in the Yangtze-GBHM

Parameters	Estimation Methods
Vegetation and Land Surface Parameters	
(1) Leaf Area Index (LAI)	Estimated from GIMMS NDVI products
(2) Surface Retention Capacity	Estimated from land use types
(3) Surface Manning Roughness Coefficient	Estimated from land use types
Soil Water Parameters	
(4) Saturated Hydraulic Conductivity	All the parameters are obtained from the IGBP-DIS global soil database
(5) Saturated Volumetric Moisture Content	
(6) Residual Volumetric Moisture Content	
(7) Parameters in van Genuchten's (VG) Equation for soil retention curve	
River Channel Parameters	
(8) Geometry of River Channels	Estimated from the measurement data
(9) River Manning Roughness Coefficient	Refer to *Maidment et al.* (1993)
Other Parameters	
(10) Snowmelt Factor	Calibrated
(11) Groundwater Hydraulic Conductivity	Calibrated
(12) Specific Yield for Unconfined Aquifer	Calibrated

Table 3. The Performance of GBHM during the Calibration and Validation Period

Gauges	Calibration Period (1961-1965)		Validation Period (1966-2000)	
	NSCE	*RB* (%)	*NSCE*	*RB* (%)
PS	0.86	4.8	0.81	12.0
GC	0.89	-3.5	0.86	6.9
BB	0.75	4.6	0.70	9.5
WL	0.77	2.7	0.71	-0.1
HJG	0.72	3.0	-0.29	14.4
TY	0.78	4.5	0.81	-3.3
XT	0.75	4.4	0.75	-2.7
WZ	0.72	4.0	0.73	4.9
CT	0.88	2.7	0.84	11.2
YC	0.82	3.7	0.78	7.8
HK	0.89	4.7	0.88	6.0
DT	0.80	4.0	0.82	5.4

Table 4. Performance of Streamflow Simulation by Different Precipitation Datasets

Gauges	GAUGE		3B42 V7		3B42 RT		CMORPH	
	NSCE (-)	*RB* (%)	*NSCE* (-)	*RB* (%)	*NSCE* (-)	*RB* (%)	*NSCE* (-)	*RB* (%)
				Monthly Simulation				
PS	0.94	8.6	0.95	1.2	-10.71	147.7	0.78	-12.9
GC	0.89	11.6	0.91	2.7	-0.63	67.5	0.75	-5.5
BB	0.92	7.4	0.93	4.1	0.76	25.7	0.57	-39.9
WL	0.61	-1.2	0.67	-11.9	0.62	-2.4	-0.34	-59.8
TY	0.87	-10.0	0.86	-8.2	0.47	6.9	-0.38	-66.2
XT	0.88	-16.3	0.78	-23.3	0.49	-14.0	-0.85	-74.8
WZ	0.91	-2.0	0.90	-0.2	0.36	6.0	-0.34	-65.6
CT	0.96	8.6	0.96	-1.3	-2.10	81.8	0.73	-22.4
YC	0.87	8.8	0.90	-0.5	-1.85	70.6	0.63	-26.9
HK	0.83	2.3	0.84	-4.3	-1.60	43.5	0.00	-42.3
DT	0.80	1.6	0.81	-2.5	-1.61	37.2	-0.33	-45.8
				Daily Simulation				
PS	0.84	8.8	0.84	-0.3	-12.60	144.3	0.65	-15.6
GC	0.80	11.6	0.70	1.8	-1.47	65.9	0.64	-6.4
BB	0.68	7.9	0.67	4.5	0.46	25.0	0.45	-39.3
WL	0.47	-1.7	0.42	-13.4	0.25	-3.5	-0.08	-59.6
TY	0.72	-11.3	0.59	-9.2	-0.09	6.7	0.00	-65.5
XT	0.76	-15.7	0.56	-26.6	0.14	-16.5	-0.37	-75.6
WZ	0.77	-1.8	0.59	-3.8	0.12	3.8	-0.11	-65.6
CT	0.83	8.7	0.85	-2.4	-2.49	80.7	0.64	-23.1
YC	0.68	9.0	0.71	-1.6	-2.37	69.5	0.53	-27.4
HK	0.71	2.8	0.76	-4.7	-1.87	43.4	0.03	-42.2
DT	0.66	2.0	0.66	-3.5	-2.06	37.4	-0.30	-45.7

Figure Captions:

Fig.1. (a) Topography and CMA meteorological gauges; and (b) streamflow gauges with hydrological sub-regions in the Yangtze River basin

Fig.2. Schematic diagram of the geomorphology-based hydrological model (GBHM) over the Yangtze River basin

Fig.3. Spatial maps of (a) annual precipitation (mm) and of bias (mm day^{-1}) for annual mean precipitation between the satellite precipitation products of (b) 3B42 V7 and gauge; (c) 3B42 RT and gauge; (d) CMORPH and gauge estimates during 2003-2012

Fig.4. Spatial maps of multi-year (2003-2012) averaged error components for the warm season (Apr-May-Jun, AMJ)

Fig.5. Same as Figure 4, except for the cold season (Oct-Nov-Dec, OND)

Fig.6 Scatter plots of daily basin-averaged rainfall estimated by satellite versus interpolated rainfall by gauges over 8 hydrological sub-regions in the Yangtze River basin

Fig.7 Comparison of the simulated water balance over 8 sub-regions in the Yangtze River (*RB_P*, *RB_E* and *RB_R* represents relative bias of precipitation, of evapotranspiration, and of runoff, respectively)

Fig.8 Comparison between monthly observed discharge and simulation results forced by different precipitation inputs over major tributaries in the Yangtze River

Fig.9 Comparison between monthly observed discharge and simulation results forced by different precipitation inputs along the mainstream of the Yangtze River

Fig.10 Flow duration curves of the observed and simulated daily discharge at representative gauges in the Yangtze River

Fig. 11 Evaluation statistics for flood events at different streamflow gauges: (a) relative bias; (b) correlation coefficient; and (c) probability of detection

Fig.12 Observed versus simulated daily streamflow during the 2012 July extreme flood event in the upper Yangtze River

Figures:

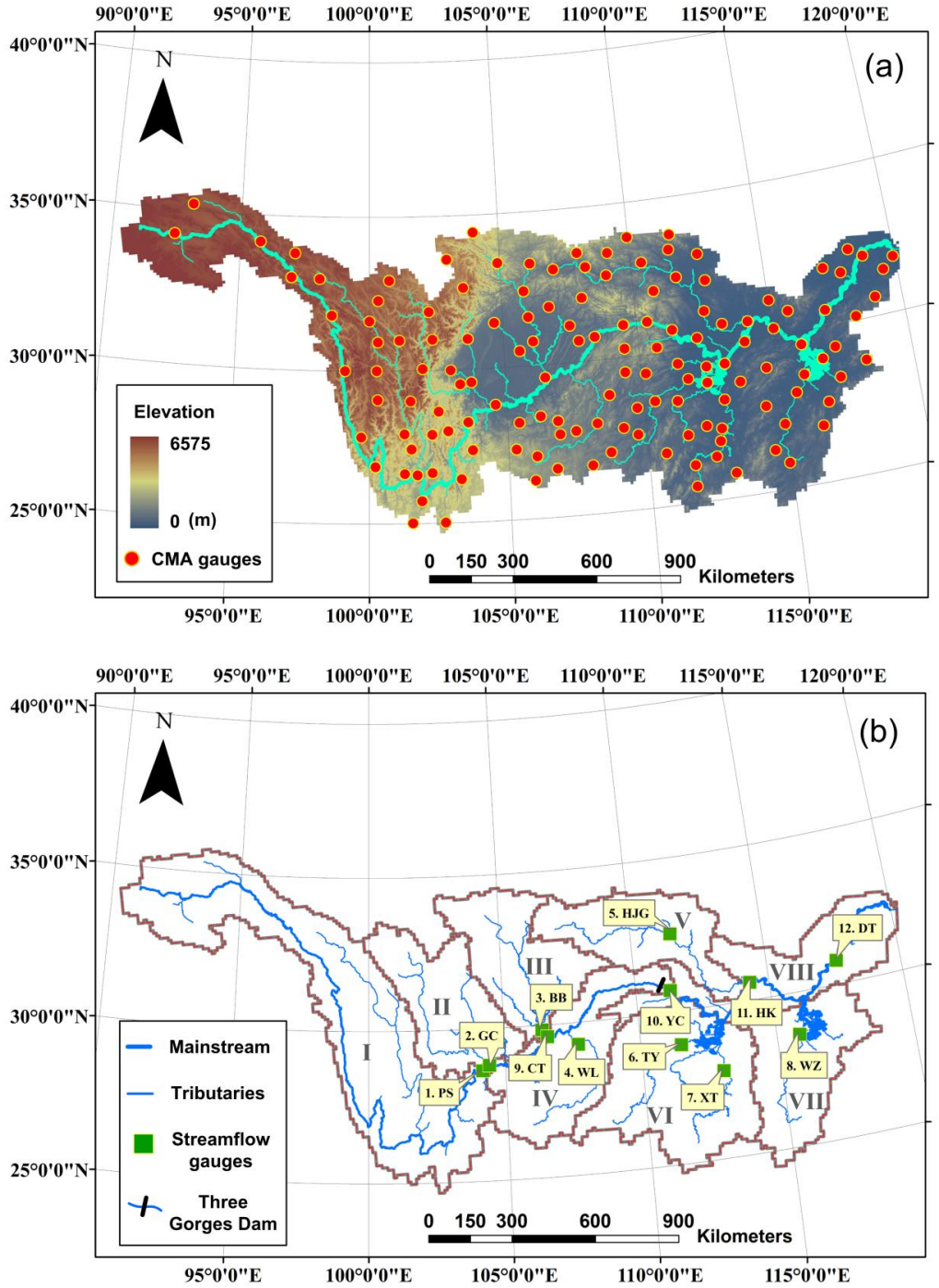

Fig.1 (a) Topography and CMA meteorological gauges; and (b) streamflow gauges with hydrological sub-regions in the Yangtze River basin

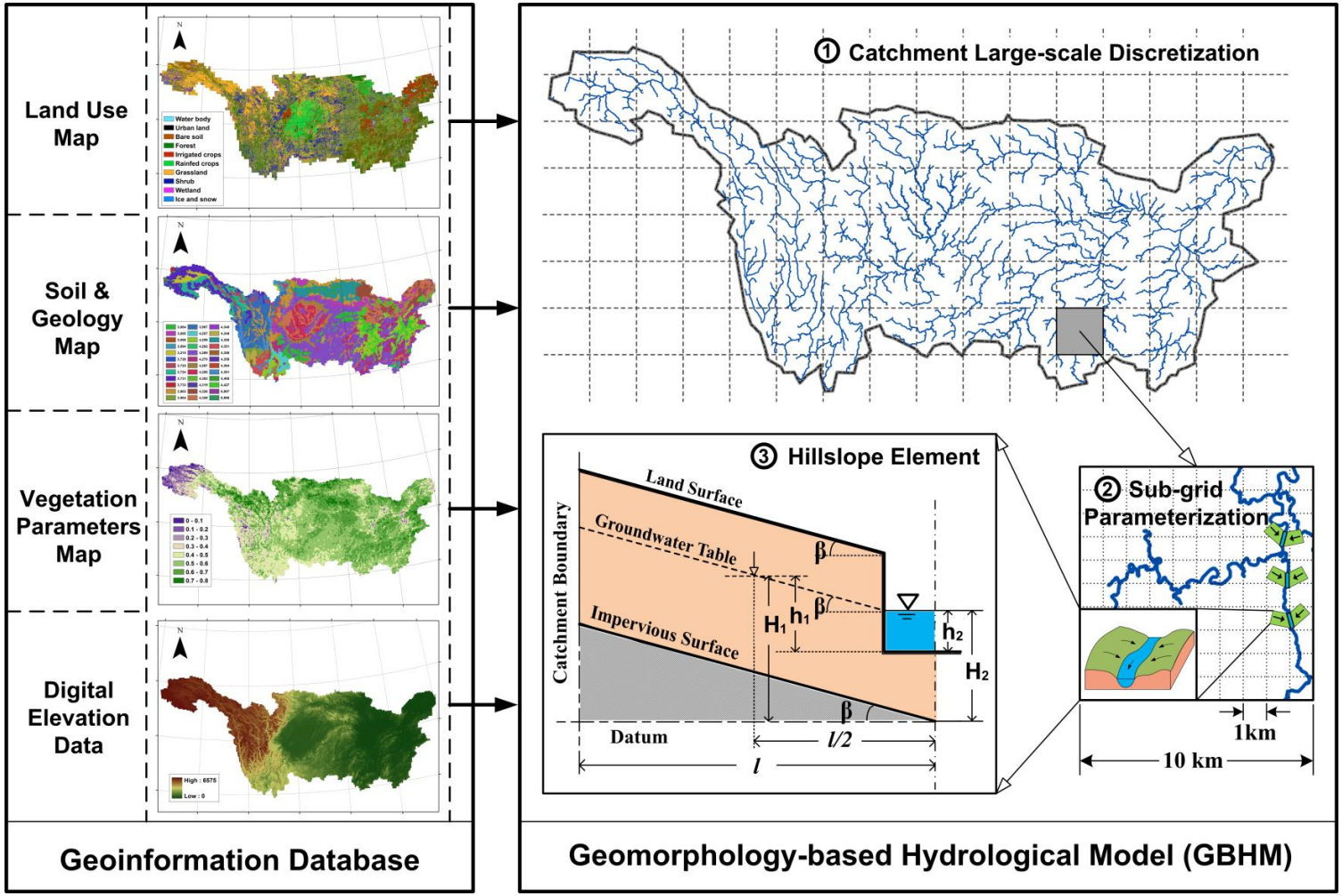

Fig.2 Schematic diagram of the geomorphology-based hydrological model (GBHM) over the Yangtze River basin

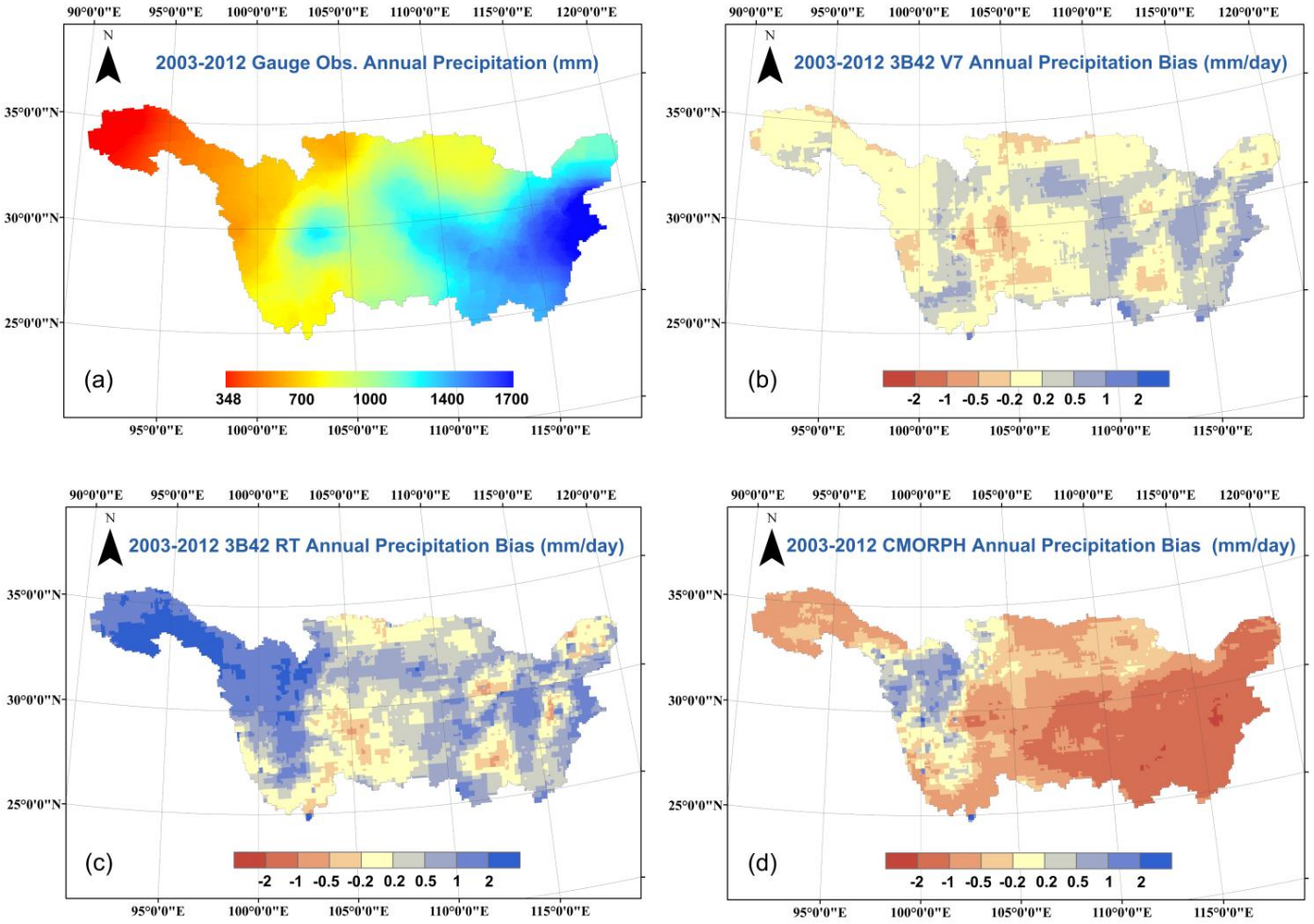

Fig.3 Spatial maps of (a) annual precipitation (mm) and of bias (mm day^{-1}) for annual mean precipitation between the satellite precipitation products of (b) 3B42 V7 and gauge; (c) 3B42 RT and gauge; (d) CMORPH and gauge estimates during 2003-2012

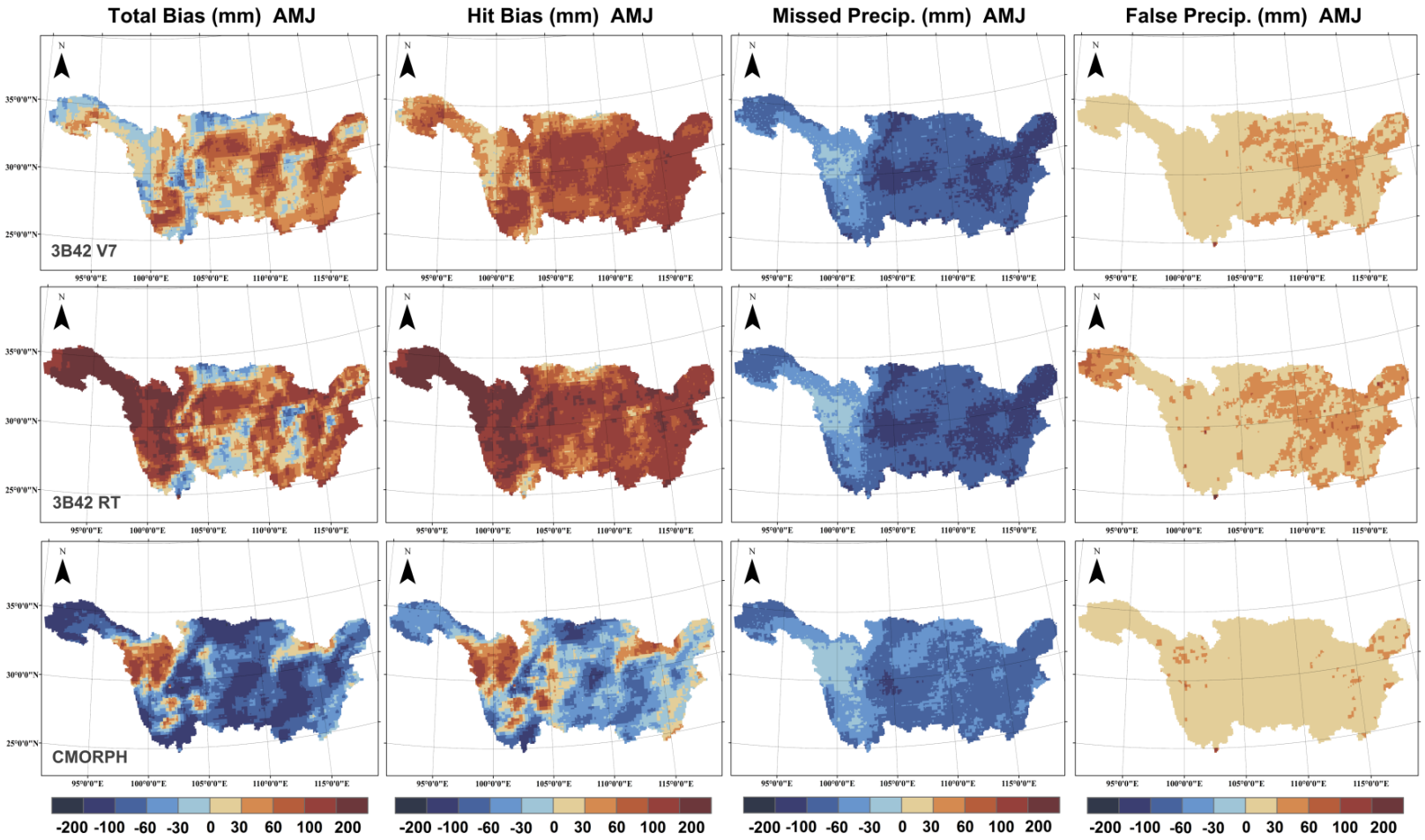

Fig.4 Spatial maps of multi-year (2003-2012) averaged error components for the warm season (Apr-May-Jun, AMJ)

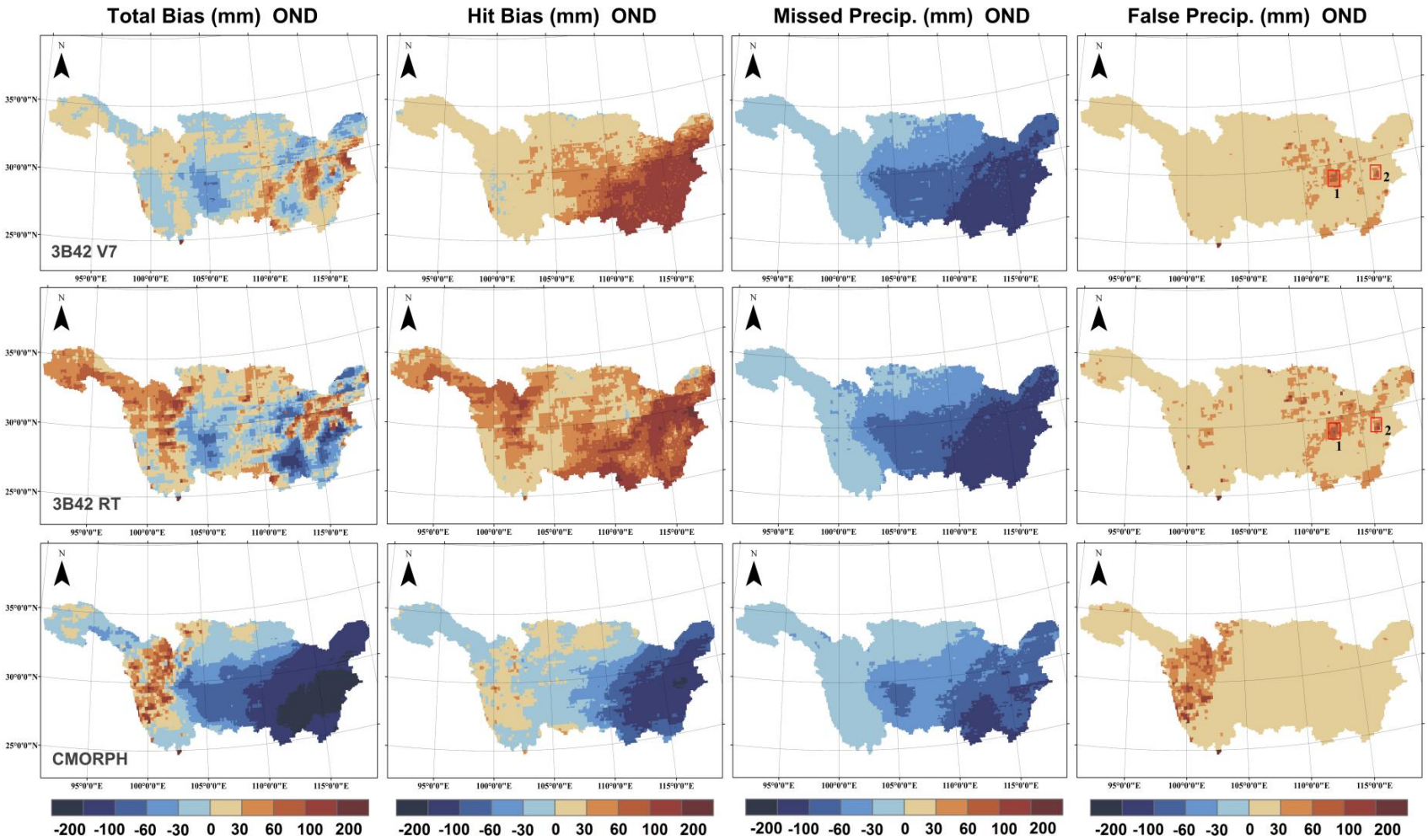

Fig.5. Same as Figure 4, except for the cold season (Oct-Nov-Dec, OND)

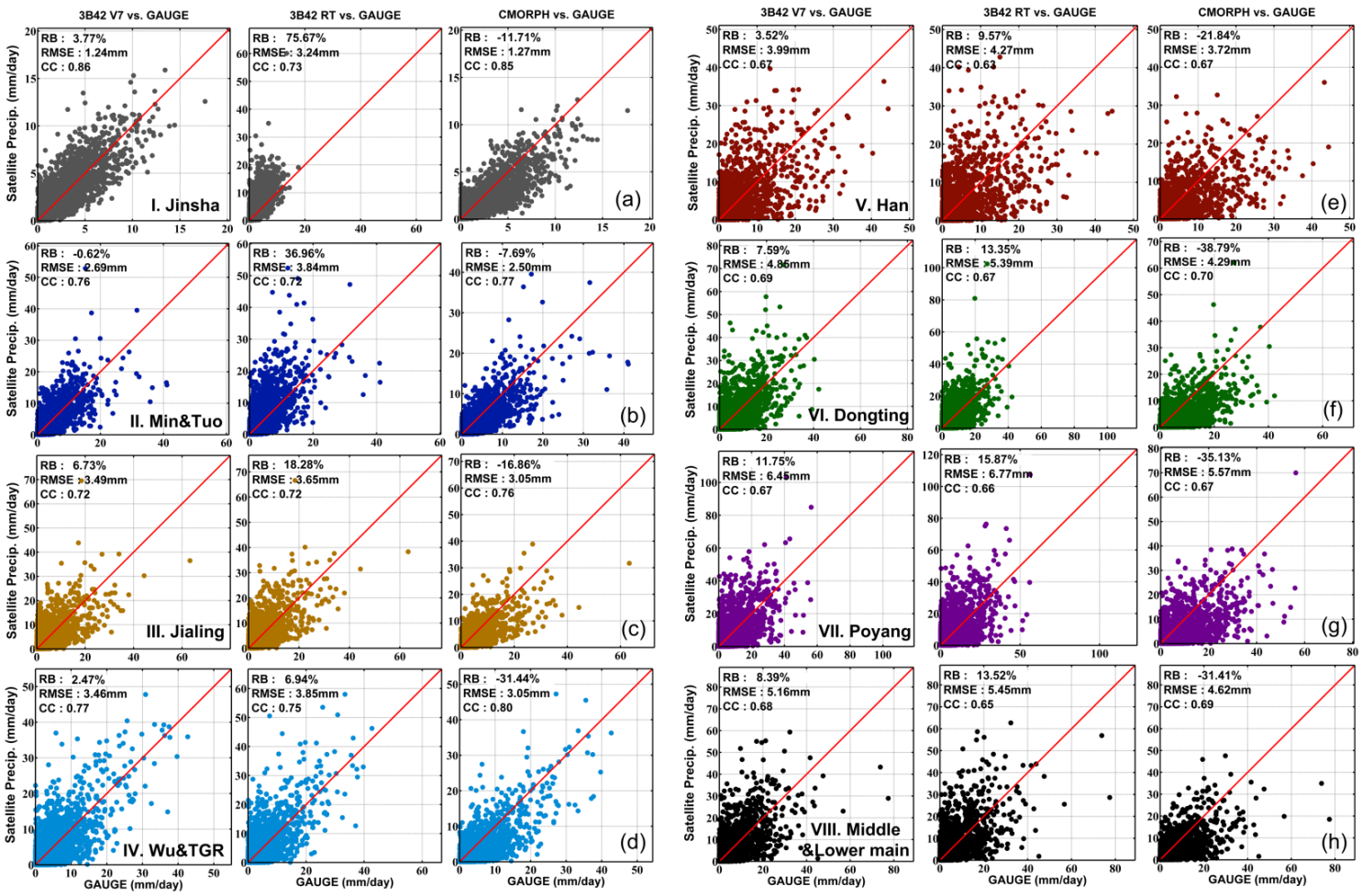

Fig.6 Scatter plots of daily basin-averaged rainfall estimated by satellite versus interpolated rainfall by gauges over 8 hydrological sub-regions in the Yangtze River basin

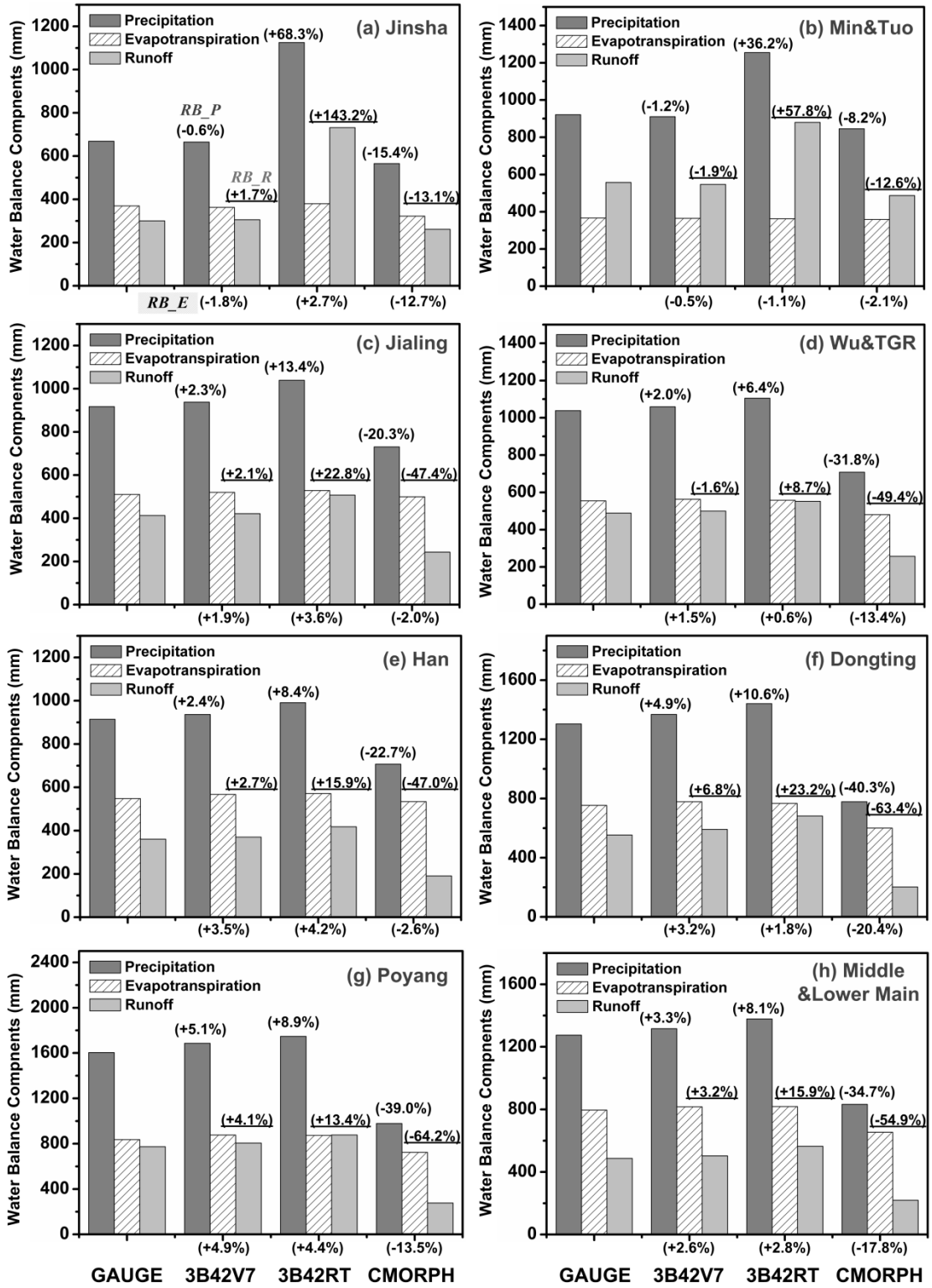

Fig.7 Comparison of the simulated water balance over 8 sub-regions in the Yangtze River (*RB_P*, *RB_E* and *RB_R* represents relative bias of precipitation, of evapotranspiration, and of runoff, respectively)

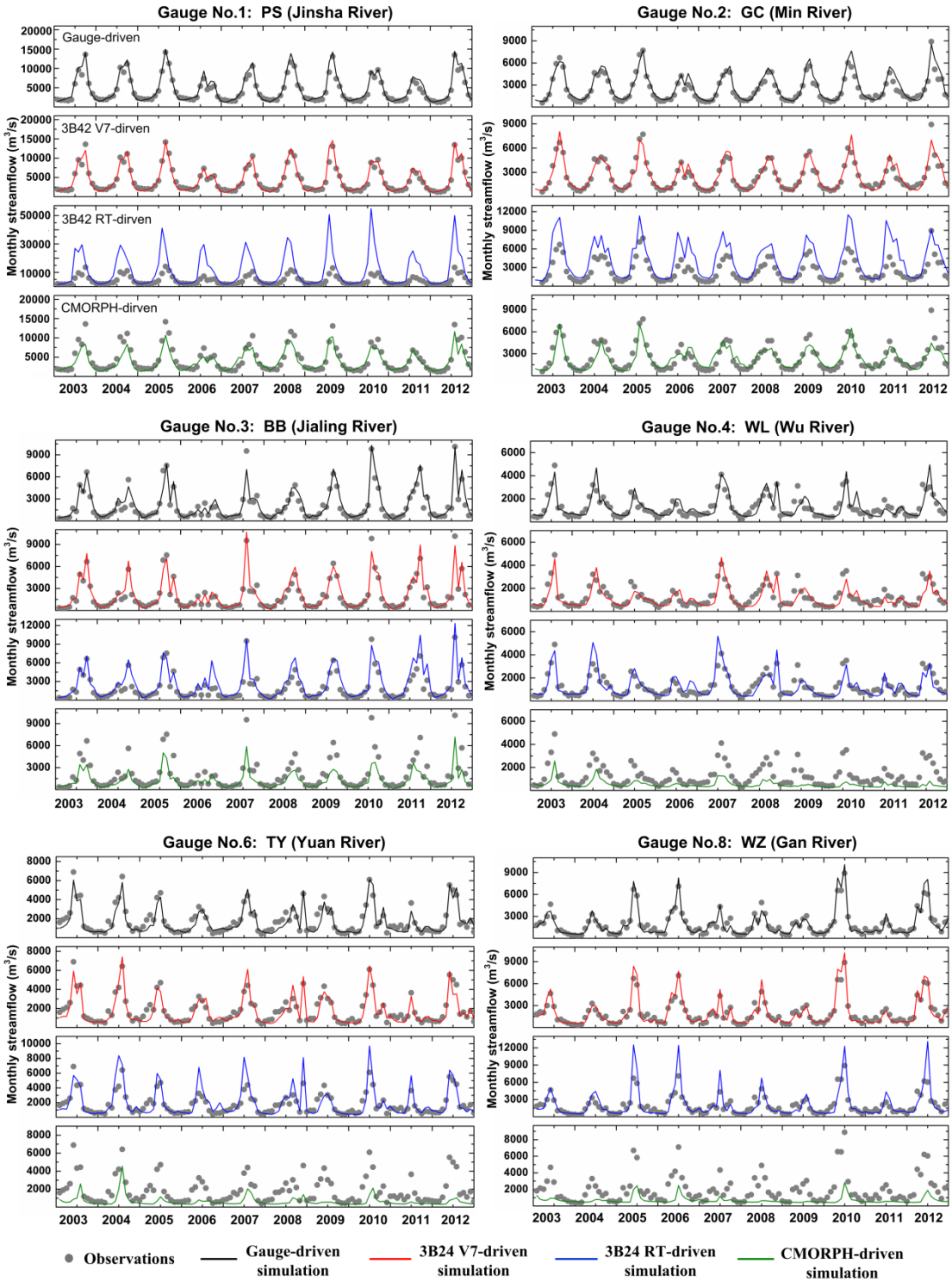

Fig.8 Comparison between monthly observed discharge and simulation results forced by different precipitation inputs over major tributaries in the Yangtze River

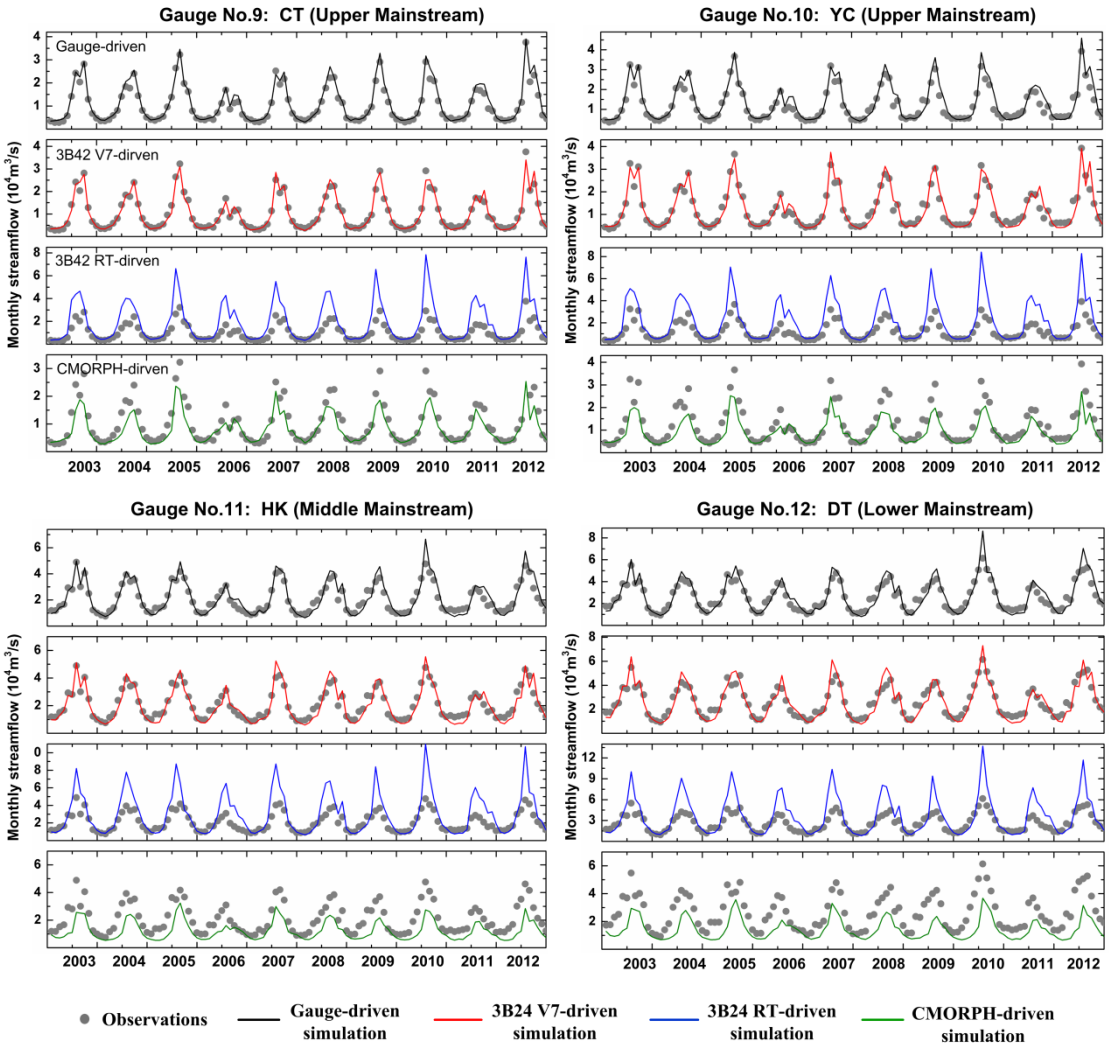

Fig.9 Comparison between monthly observed discharge and simulation results forced by different precipitation inputs along the mainstream of the Yangtze River

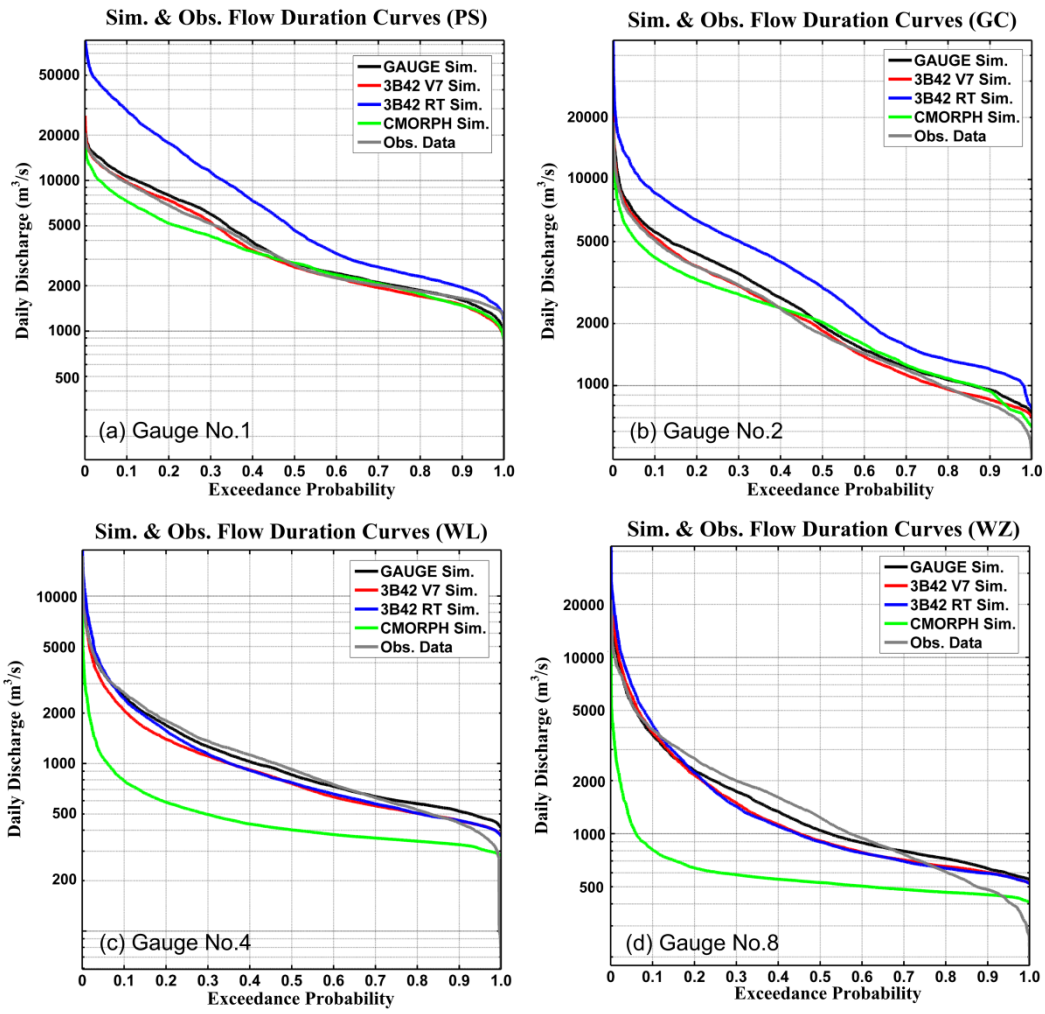

Fig.10 Flow duration curves of the observed and simulated daily discharge at representative gauges in the Yangtze River

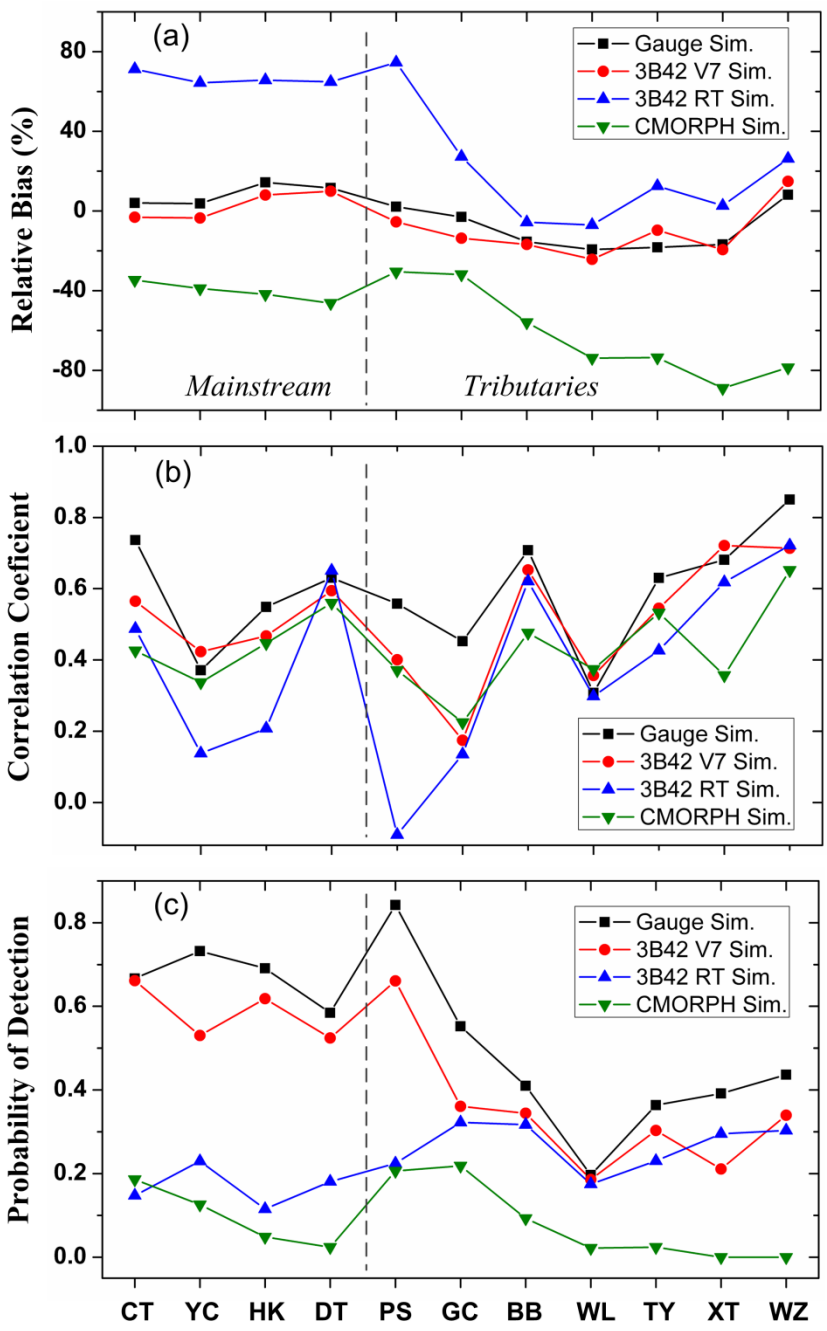

Fig.11 Evaluation statistics for flood events at different streamflow gauges: (a) relative bias; (b) correlation coefficient; and (c) probability of detection.

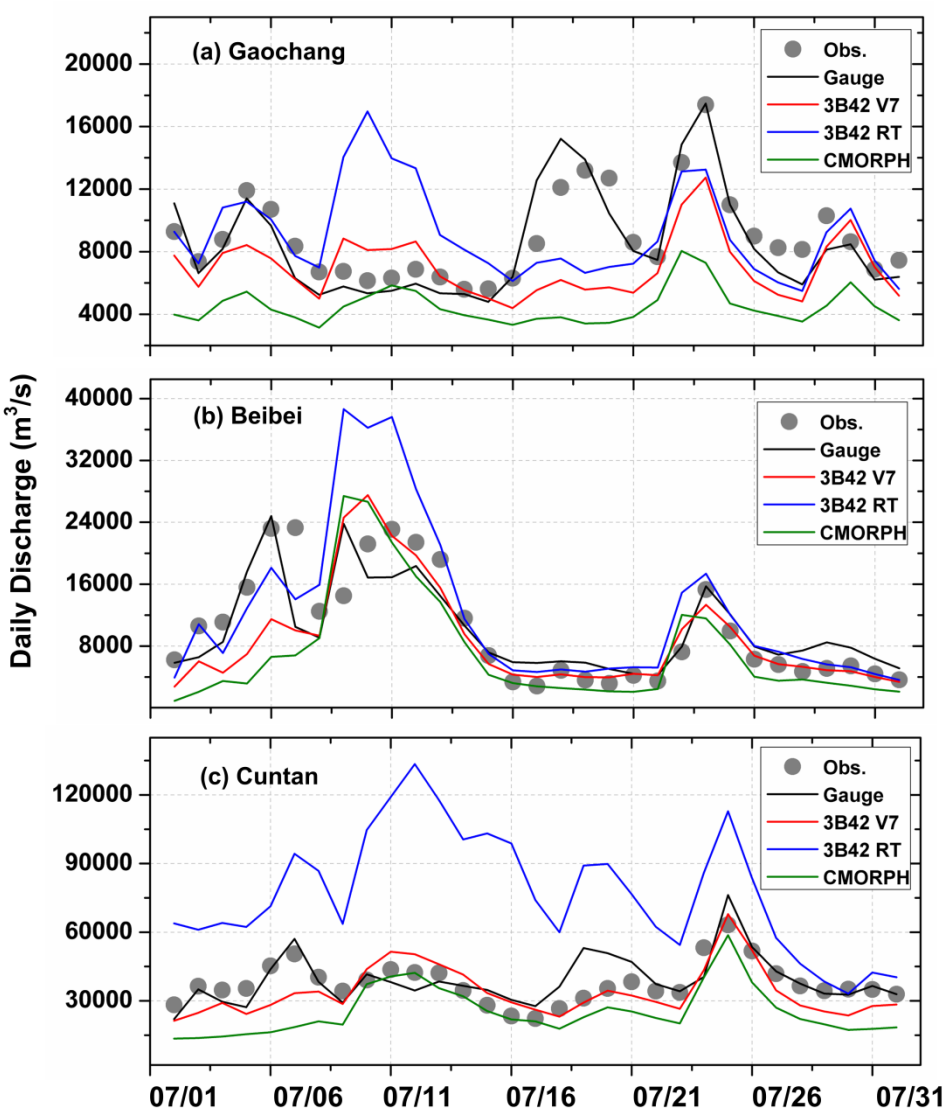

Fig.12 Observed versus simulated daily streamflow during the 2012 July extreme flood event in the upper Yangtze River