Evaluation of global fine-resolution precipitation products and their uncertainty quantification in ensemble discharge simulations

W. Qi[1,2], C. Zhang[1], G. T. Fu[2], C. Sweetapple[2] and H. C. Zhou[1]

[1] School of Hydraulic Engineering, Dalian University of Technology, Dalian 116024, China

[2] Centre for Water Systems, College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Harrison Building, Exeter EX4 4QF, UK

Correspondence to: Chi Zhang (czhang@dlut.edu.cn)

Abstract

The applicability of six fine-resolution precipitation products, including precipitation radar, infrared, microwave and gauge-based products using different precipitation computation recipes, is ~~comprehensively ~~evaluated using statistical and hydrological methods in ~~a usually neglected area (~~northeastern China~~), and~~. In addition, a framework quantifying uncertainty contributions of precipitation products, hydrological models and their interactions to uncertainties in ensemble discharges is proposed. The investigated precipitation products ~~include~~are TRMM3B42, TRMM3B42RT, GLDAS/Noah, APHRODITE, PERSIANN and GSMAP-MVK+. Two hydrological models of different complexities, i.e., a water and energy budget-based distributed hydrological model and a physically-based semi-distributed hydrological model, are employed to investigate the influence of hydrological models on simulated discharges. Results show APHRODITE has high accuracy at a monthly scale compared with other products, and ~~the cloud motion vectors used by ~~GSMAP-MVK+ shows huge advantage and is better than TRMM3B42 in RB, NSE, RMSE, CC, false alarm ratio and

critical success index. These findings could be very useful for validation, refinement and future development of satellite-based products (e.g., NASA Global Precipitation Measurement). Although ~~significant~~large uncertainty exists in heavy precipitation, hydrological models contribute most of the uncertainty in extreme discharges. Interactions between precipitation products and hydrological models contribute ~~significantly~~a lot to uncertainty in discharge simulations and a better precipitation product does not guarantee a better discharge simulation because of interactions. It is also found that a good discharge simulation depends on a good coalition of a hydrological model and a precipitation product, suggesting that, although the satellite-based precipitation products are not as accurate as the gauge-based product, they could have better performance in discharge simulations when appropriately combined with hydrological models. This information is revealed for the first time and very beneficial for precipitation product applications.

1 Introduction

Knowledge of precipitation plays an important role in the understanding of the water cycle, and thus in water resources management (Sellers, 1997;Sorooshian et al., 2005;Wang et al., 2005;Ebert et al., 2007;Buarque et al., 2011;Tapiador et al., 2012;Yong et al., 2012;Gao and Liu, 2013;Peng et al., 2014a;Peng et al., 2014b). However, ~~there are little or no~~ precipitation data are not available in many regions~~ throughout the world~~, particularly ~~in developing countries,~~ mountainous districts and rural areas in developing countries. For example, Northeast China, which plays an important role in food production to support the country's population and is also an industrial region with many heavy industries, frequently suffers from drought, posing a threat to regional sustainable development. In such areas, due to insufficient gauge observations, alternative precipitation data are required for efficient water resources management.

In recent years, implementation of gauge-based and remote satellite-based precipitation products has become popular, particularly for ungauged catchments (Artan et al., 2007;Jiang et al., 2012;Li et al., 2013;Maggioni et al., 2013;Müller and Thompson, 2013;Xue et al., 2013;Kneis et al., 2014;Meng et al., 2014;Ochoa et al., 2014). Numerous precipitation products have been developed to estimate rainfall, for example:

- Tropical Rainfall Measuring Mission (TRMM) products (Huffman et al., 2007)
- Global Land Data Assimilation System (GLDAS) precipitation products (Kato et al., 2007)
- Asian Precipitation - Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) (Xie et al., 2007;Yatagai et al., 2012)
- Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) (Sorooshian et al., 2000;Sorooshian et al., 2002)
- Global Satellite Mapping of Precipitation product (GSMAP) (Kubota et al., 2007;Aonashi et al., 2009)

There are uncertainties in these products. Several studies have been carried out to analyze the uncertainty of TRMM in high latitude regions (Yong et al., 2010;Yong et al., 2012;Chen et al., 2013a;Yong et al., 2014;Zhao and Yatagai, 2014), but studies in northeast China are few. Evaluation of GLDAS data has generally been limited to the United States and other observation-rich regions of the world (Kato et al., 2007); assessments and applications in other regions are rare (Wang et al., 2011;Zhou et al., 2013). The APHRODITE, PERSIANN and GSMAP products are seldom evaluated in northeast China using basin scale gauge data (Zhou et al., 2008). Owing to the high heterogeneity of rainfall across a variety of spatiotemporal scales, the uncertainty characteristics of precipitation products are variable (Asadullah et al., 2008;Dinku et al., 2008;Nikolopoulos et al., 2010;Pan et al., 2010). Thus, in northeast China, it is essential to completely evaluate the applicability of these precipitation products. In addition, it is also worth comparing the performance of different precipitation computation recipes: for example, the artificial neural network function used in PERSIANN, the histogram matching approach used in TRMM3B42, and the cloud motion vectors used in GSMAP-MVK+, because the inter-comparison could reveal the strategies that could be used to obtain more accurate precipitation data.

Many researchers have implemented precipitation products in discharge simulations and reported discharge uncertainties (Hong et al., 2006;Pan et al., 2010;Serpetzoglou et al., 2010). Also, many uncertainty analysis approaches have been introduced to quantify the uncertainty (Beven and Binley, 1992;Freer et al., 1996;Kuczera and Parent, 1998;Beven and Freer, 2001b;Peters et al., 2003;Heidari et al., 2006;Kuczera et al., 2006;Tolson and Shoemaker, 2007;Blasone et al., 2008;Vrugt et al., 2009a;Vrugt et al., 2009b). In these prior approaches, one of the popular methods is the generalized likelihood uncertainty estimation (GLUE) technique, introduced by Beven and Binley (1992). This approach outputs probability distributions of model parameters conditioned on observed data, and the uncertainties in model inputs are represented by uncertain parameters. Similar to GLUE, Hong et al. (2006) proposed a Monte Carlo based method to quantify uncertainty in hydrological simulations using satellite precipitation data, in which flow simulation uncertainty is represented by ensemble simulation results.

In addition to ~~uncertainty resulting~~individual contributions from ~~the~~ hydrological models~~,~~ and precipitation data, the interactions between precipitation products and hydrological models

~~and uncertainty in precipitation data~~ also contribute to uncertainty in simulated discharges. However, to the best of our knowledge, the previous studies have not quantified the respective contributions of precipitation products, hydrological models and their interactions to the total discharge simulation uncertainty.

~~The overall objectives of this paper are: (1) to investigate the applicability of six fine-resolution precipitation products using both statistical and hydrological evaluation methods in the usually-neglected area - northeast China; (2) to propose~~The overall aim of this paper is to develop a framework to quantify the contributions of uncertainties from precipitation products, hydrological models and their interactions to uncertainty in simulated discharges. ~~The~~To achieve the aim, the first step is to understand the performance of the selected precipitation products ~~investigated include~~including TRMM3B42, TRMM3B42RT, GLDAS/Noah (GLDAS_Noah025SUBP_3H), APHRODITE, PERSIANN and GSMAP-MVK+, when applied to the chosen hydrological models. Two hydrological models ~~with~~of different complexities - a water and energy budget-based distributed hydrological model (WEB-DHM) (Wang et al., 2009a;Wang et al., 2009b;Wang et al., 2009c) and a physically-based semi-distributed hydrological model TOPMODEL (Beven and Kirkby, 1979) - were employed to investigate the influence of hydrological models on discharge simulations. ~~A series of 8-year data was employed.~~ Building on the assessment of the precipitation products, the second step is to quantify the respective uncertainties from the precipitation products and hydrological models, and the combined uncertainties from the interactions between products and models. This is achieved using a global sensitivity analysis approach, i.e., the analysis of variance approach (ANOVA). A river basin in northern China with a series of 8-year data is used to demonstrate the methodology.

The paper is organized as follows. Section 2 introduces the study region, precipitation products, hydrological models and the proposed framework. Section 3 presents the statistical evaluation results. Hydrological evaluations and the implementation of the proposed framework are given in section 4. Discussion is given in section 5. Summary and conclusions are presented in section 6.

2 Materials and methodology

2.1 Biliu basin

Biliu basin (2814 km^2), located in the coastal region between the China Bohai Sea and the China Huanghai Sea, covers longitudes 122.29°E to 122.92°E and latitudes 39.54°N to 40.35°N. This basin is characterized by a snow - winter dry - hot summer climate (Koppen climate classification) and the average annual temperature is 10.6°C. Summer (July to September) is the major rainy season. There are 11 rainfall stations and one discharge gauge which have historical data from January 2000 to December 2007. ~~and the average annual temperature is 10.6°C~~ The average elevation is 240 meters.~~–~~ The gauge distribution in Biliu is shown in Fig. 1.

2.2 Precipitation products

The selected precipitation products are shown in Table 1. These data are all freely available. In these selected precipitation products, APHRODITE is fully based on gauge data; TRMM3B42 and GLDAS are remote satellite estimation with gauge data corrections; while others are remote satellite estimation without gauge data corrections. Remote-based precipitation estimation has many weaknesses, e.g., microwave estimation could miss convective rainfall and typhoon rain because of its sparse time interval resolution; infrared estimation has a higher time interval resolution, but it cannot penetrate thick clouds. Ground rain gauge-based interpolation products are limited by interpolation algorithms, gauge density and gauge data quality (Xie et al., 2007). The details of data sources used in each precipitation product can be found in Table 1. The detailed introductions of these products are as follows.

TRMM is a joint mission between NASA and Japan Aerospace Exploration Agency designed to monitor and study tropical rainfall (Kummerow et al., 2000;Huffman et al., 2007). Three instruments - a visible infrared radiometer, a TRMM microwave imager and a precipitation radar - are employed to obtain accurate precipitation estimation. The TRMM precipitation radar is the first space-based precipitation radar and operates between 35°N and 35°S. Outside this band, the microwave imager is used between 40°N and 40°S, and the visible infrared radiometer data are used between 50°N to 50°S. Usually the precipitation radar is considered to give the most accurate estimation from satellite, and data from it are often used for calibration of passive microwave data from other instruments (Ebert et al., 2007). The post-real-time product used in this study is the TRMM3B42, which utilizes three data sources: the TRMM combined instrument estimation using data from both TRMM precipitation radar and the microwave imager; the GPCP monthly rain gauge analysis developed by the Global Precipitation Climatology Center; and the Climate Assessment and Monitoring System monthly rain gauge analysis. TRMM3B42 applies an infrared to rain rate relationship using histogram matching, while TRMM3B42RT merges microwave and infrared precipitation estimation.

PERSIANN is a product that, using an artificial neural network function, estimates precipitation by combining infrared precipitation estimation and the TRMM combined instrument estimation (which assimilates with TRMM precipitation radar and microwave data). GSMAP-MVK+ uses microwave and infrared precipitation data together and combines cloud motion vectors to generate fine-resolution precipitation estimation.

The Global Land Data Assimilation System (GLDAS) project is an extension of the existing and more mature North American Land Data Assimilation System (Rodell et al., 2004). It integrates satellite- and ground-based data sets for parameterizing, forcing and constraining a few offline land surface models for generating optimal fields of land surface states and fluxes. At present, GLDAS drives four Land Surface Models: Mosaic (Koster and Suarez, 1992), Noah (Chen et al., 1996;Betts et al., 1997;Koren et al., 1999;Ek, 2003), the Community Land Model (Dai et al., 2003) and Variable Infiltration Capacity model (Liang et al., 1994). Among them, the GLDAS/Noah Land Surface Model product (GLDAS_NOAH025SUBP_3H) has a 3-h 0.25° × 0.25° resolution, which is desirable for basin scale research. The GLDAS precipitation data combine microwave and infrared, and also assimilate gauge observations.

~~The inter-comparison schemes which mainly include five experiments (Exp 1 – Exp 5) are shown in 2. The five experiments were set up based on the differences among the precipitation products in data types (including data sources and recipes) listed in table 1: thus the differences in precipitation amounts can reflect the differences in data types. The inter-comparison results could be potentially used to improve products. Exp 1 is to compare TRMM3B42 and TRMM3B42RT. Exp 2 is to compare the differences between TRMM3B42 and PERSIANN. Exp 3 is to compare the most popular satellite product TRMM3B42 and the fully gauge-based product APHRODITE. Exp 4 is to compare TRMM3B42 and GSMAP-MVK+. Exp 5 is to compare GSMAP-MVK+ and GLDAS/Noah.~~

2.3 Criteria for accuracy assessment

Uncertainties of precipitation products are evaluated on the basis of basin-averaged rainfall observations. Four evaluation criteria are used in rainfall amount error assessment: correlation coefficient (CC), root mean square error (RMSE), Nash-Sutcliffe coefficient of efficiency (NS~~C~~E) and relative bias (RB). These are calculated as follows:

$$\mathrm{RMSE} = \left(\frac{\sum_{i=1}^{n} \left(X_{pi} - X_{oi} \right)^2}{n} \right)^{\frac{1}{2}} \tag{1}$$

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n} \left(X_{pi} - X_{oi} \right)^2}{\sum_{i=1}^{n} \left(X_{oi} - \overline{X}_o \right)^2} \tag{2}$$

$$\mathrm{RB} = \frac{\sum_{i=1}^{n} X_{pi} - \sum_{i=1}^{n} X_{oi}}{\sum_{i=1}^{n} X_{oi}} \times 100\% \tag{3}$$

where X_{oi} represents observed data; X_{pi} represents estimated data; n is the total number of data points. A perfect fit should have CC and NS~~C~~E values of one. The lower the RMSE and RB, the better the estimation. These comparison criteria have been used by many studies (Ebert et al., 2007;Wang et al., 2011;Yong et al., 2012)~~. In discharge simulation, RMSE and RB~~, so they are used in this study.

Probability distributions by occurrence and volume are also analyzed, which can provide us with the information on the frequency and on the product error dependence on precipitation intensity (Chen et al., 2013a;Chen et al., 2013b). The critical success index (CSI), probability of detection (POD) and false alarm ratio (FAR) are used to quantify the ability of precipitation products to detect observed rainfall events. These are defined as follows:

$$\mathrm{CSI} = \frac{H}{H + M + F} \tag{4}$$

$$\mathrm{POD} = \frac{H}{H + M} \tag{5}$$

$$\mathrm{FAR} = \frac{F}{H + F} \tag{6}$$

where H is the total number of hits; M is the total number of misses; F is the total number of false alarms (Ebert et al., 2007;Su et al., 2008). A perfect detection should have CSI and POD values equal to one and a FAR value of zero.

2.4 Hydrological models and data

2.4.1 WEB-DHM

The distributed biosphere hydrological model, WEB-DHM (Wang et al., 2009a;Wang et al., 2009b;Wang et al., 2009c), was developed by coupling a simple biosphere scheme (Sellers et al., 1986) with a geomorphology-based hydrological model (Yang, 1998) to describe water, energy and CO_2 fluxes at a basin scale. WEB-DHM has been used in several evaluations and applications (Wang et al., 2010a;Wang et al., 2010b;Wang et al., 2012;Shrestha et al., 2013).

WEB-DHM input data include precipitation, temperature, downward solar radiation, long wave radiation, air pressure, wind speed and humidity. With the exception of precipitation, all input data are obtained from automatic weather stations. There are three automatic weather stations near Biliu, and observations from these are obtained from the China Meteorological Data Sharing Service System (downloaded from http://cdc.cma.gov.cn/home.do). Hourly precipitation data are downscaled from daily rain gauge observations using a stochastic method (Wang et al., 2011). Hourly temperatures are calculated from daily maximum and minimum temperatures using the TEMP model (Parton and Logan, 1981). The estimated temperatures are also further evaluated using daily average temperature. Downward solar radiation is estimated from sunshine duration, temperature and humidity using a hybrid model (Yang et al., 2006). Long wave radiation is obtained from the GLDAS/Noah (Rodell et al., 2004). Air pressure is estimated according to altitude (Yang et al., 2006). These meteorological data are then interpolated to 300 m × 300 m model cells through an inverse-distance weighting approach. Because of the elevation differences ~~between~~among model cells and meteorological gauges~~.~~, the interpolated surface air temperatures are further modified with a lapse rate of 6.5K/km. Gauge rainfall data are also interpolated to 300 m × 300 m model cells~~.~~ and basin-averaged gauge rainfall data are calculated on the basis of interpolation results. In addition to the above, the leaf area index and fraction of photosynthetically active radiation data are obtained from level-4 MODIS global products-MOD11A2. Digital Elevation Model (DEM) is from the NASA SRTM (Shuttle Radar Topographic Mission) with a resolution of 30 m × 30 m. We resampled the resolution to 300 m in model calculation to reduce computation cost, while the model processed finer DEM (30 m grid) to generate sub-grid parameters (such as hillslope angle and length). The grid slopes vary from 0 to 38 degrees. Land-use data are obtained from the USGS (http://edc2.usgs.gov/glcc/glcc.php). The land-use types have been reclassified to SiB2 land-use types for this study (Sellers et al., 1996). There are six land-use types, with broadleaf and needle leaf trees and short vegetation being the main types. Soil data are obtained from the Food and Agriculture Organization (FAO) (2003) Global data product, and there are two types of soil in the basin: clay loam-luvisols and loam-phaeozems.

~~WEB-DHM was calibrated against observed discharges of Biliu. Six main parameters were~~

~~selected to calibrate using a trial and error approach given consideration of the model computational burden. What we calibrated were parameter multipliers, similar to the study by Wang et al. (2011). The 'Trial and error' approach has two steps. First, all the multiplier values are set to 1 which represents the default parameter values from Food and Agriculture Association (FAO) (2003) and SiB2 model. Second, varying the multiplier values until acceptable discharge simulation accuracy is obtained. The calibrated parameter values are listed in Table 2. The simulated daily, monthly and inter-annual results are shown in Figs. 3a, 3c and 3e.~~

2.4.2 TOPMODEL

TOPMODEL is a physically-based, variable contributing area model of basin hydrology which attempts to combine the advantages of a simple lumped parameter model with distributed effects (Beven and Kirkby, 1979). Fundamental to TOPMODEL's parameterization are three assumptions: (1) saturated-zone dynamics can be approximated by successive steady-state representations; (2) hydrological gradients of the saturated zone can be approximated by the local topographic surface slope; and (3) the transmissivity profile whose form declines exponentially with increasing vertical depth of the water table or storage is spatially constant. On the basis of the above mentioned assumptions, the index of hydrological similarity is represented as the topographic index, $\ln(a/\tan\beta)$, for which a is the area per unit contour length and ß is the local slope angle. More detailed descriptions of TOPMODEL and its mathematical formulation can be found in Beven et al. (1979). TOPMODEL has been popularly utilized in research across the world (Blazkova and Beven, 1997;Cameron et al., 1999;Hossain and Anagnostou, 2005;Bastola et al., 2008;Gallart et al., 2008;Bouilloud et al., 2010;Qi et al., 2013), because of its relatively simple model structure.

The input data of TOPMODEL mainly includes basin averaged precipitation and topographic data which can be estimated from DEM.

~~Six parameters of TOPMODEL were calibrated using the dynamically dimensioned search algorithm (Tolson and Shoemaker, 2007), and the results are given in Table 3. The ranges were defined based on experience. The simulated daily, monthly and inter-annual results are shown in Figs. 3b, 3d and 3f.~~

2.5 The proposed framework

Fig. ~~4~~2 shows the diagrammatic flowchart of the proposed framework for quantification of uncertainty contributions to ensemble discharges simulated using precipitation products. This framework includes four parts: (a) selection of precipitation products; (b) selection of hydrological models; (c) ensemble discharge simulations using the hydrological models and precipitation products; and (d) quantification of individual and interactive contributions using the analysis of variance (ANOVA) approach including contributions from precipitation products, hydrological models and interactions between models and products. Because the spatial resolution of selected precipitation products does not correspond with WEB-DHM model cells, the following procedures were carried out for basin averaged rainfall calculations: (1) resampling 0.25^{o} or 0.1^{o} precipitation product grids into 300 m × 300 m cells (the grid size used in WEB-DHM simulations); (2) calculating basin-averaged precipitation using 300 m precipitation product grids located within the basin boundary. Diagrammatic descriptions of these procedures are shown in Fig. 1d. Because WEB-DHM needs hourly input data, for the 3-hour resolution precipitation products, we assumed rainfall is uniformly distributed within each 3-hour period. For daily resolution products, we used the same approach as downscaling observed precipitation data. This downscaling approach may affect uncertainty in simulated discharge. However, Wang et al. (2011) have already successfully applied the downscaling approach, and showing that the influence is negligible.

The total ensemble uncertainty Y is the variance of discharges. To relate Y to the uncertainty sources, the superscripts j and k in $Y^{j,k}$ represent a combination of precipitation product j and hydrological model k

$$Y^{j,k} = P^{j} + M^{k} + PM^{j,k} \tag{7}$$

where P represents the effect of jth precipitation product, M represents the effect of kth hydrological model, and PM represents the interaction effect. In this study, j varies from one to six, and k varies from one to two. Details of the quantification are explained in the follow sections. ~~The chain in which precipitation products and hydrological models are combined is shown in Fig. 5.~~

2.5.1 Subsampling approach

~~It is argued that the~~ ANOVA ~~approach is based on a biased variance estimator that underestimates~~could underestimate variance when the sample size is small (Bosshard et al., 2013). To reduce the effect of the ~~biased estimator on quantification of variance contributions~~sample size, Bosshard et al. (2013) proposed a subsampling method, which was used in this paper. In the subsampling method, the superscript j in Eq. (7) is replaced with $\mathbf{g}(h,i)$. According to Bosshard et al. (2013), in each subsampling iteration i, data from two products should be selected out of all the six products, and thus 15 combinations can be obtained. Therefore, the superscript $\mathbf{g}$ becomes a 2 × 15 matrix:

$$\mathbf{g} = \begin{pmatrix} 1 & 1 & \cdots & 1 & 2 & 2 & \cdots & 4 & 4 & 5 \\ 2 & 3 & \cdots & 6 & 3 & 4 & \cdots & 5 & 6 & 6 \end{pmatrix} \tag{8}$$

2.5.2 Uncertainty contribution decomposition

Based on the ANOVA theory (Bosshard et al., 2013), total error variance (SST) can be divided into sums of squares due to the individual effects as:

$$\mathrm{SST} = \mathrm{SSA} + \mathrm{SSB} + \mathrm{SSI} \tag{9}$$

where SSA is the error contribution of precipitation products, SSB is the error contribution of hydrological models and SSI is the error contribution of their interactions.

The terms can be estimated using the subsampling procedure as follows:

$$\mathrm{SST}_i = \sum_{h=1}^{H}\sum_{k=1}^{K}\left(Y^{\mathbf{g}(h,i),k} - Y^{\mathbf{g}(o,i),o}\right)^2 \tag{10}$$

$$\mathrm{SSA}_i = K \cdot \sum_{h=1}^{H}\left(Y^{\mathbf{g}(h,i),o} - Y^{\mathbf{g}(o,i),o}\right)^2 \tag{11}$$

$$\mathrm{SSB}_i = H \cdot \sum_{k=1}^{K}\left(Y^{\mathbf{g}(o,i),k} - Y^{\mathbf{g}(o,i),o}\right)^2 \tag{12}$$

$$\mathrm{SSI}_i = \sum_{h=1}^{H}\sum_{k=1}^{K}\left(Y^{\mathbf{g}(h,i),k} - Y^{\mathbf{g}(h,i),o} - Y^{\mathbf{g}(o,i),k} + Y^{\mathbf{g}(o,i),o}\right)^2 \tag{13}$$

where symbol ^{o} indicates averaging over a particular index; H is the number of precipitation products (six in this study) and K is the number of hydrological models (two in this study). Then the variation fraction η^2 is calculated as follows:

$$\eta^2_{\mathrm{precipitation}} = \frac{1}{I}\sum_{i=1}^{I}\frac{\mathrm{SSA}_i}{\mathrm{SST}_i} \tag{14}$$

$$\eta^2_{\mathrm{model}} = \frac{1}{I}\sum_{i=1}^{I}\frac{\mathrm{SSB}_i}{\mathrm{SST}_i} \tag{15}$$

$$\eta^2_{\text{interaction}} = \frac{1}{I}\sum_{i=1}^{I}\frac{\text{SSI}_i}{\text{SST}_i} \tag{16}$$

η^2 has a value between 0 and 1, which represent 0% and 100% contributions to the overall uncertainty of simulated discharges respectively. I equals 15 in this study. As shown in Eqs. 14-16, the subsampling approach is necessary because it guarantees that every contributor has the same denominator I. This same denominator makes sure that the inter-comparison among precipitation contribution, model contribution and interaction contribution is free of influence from the sampling number of precipitation products and hydrological models.

3 Statistical evaluations

3.1 Daily and monthly scales

Comparison of precipitation product data and gauge observations at a daily scale is shown in Fig. ~~6~~3. Observations are shown on the *x*-axis and precipitation product data are shown on the *y*-axis. Four criteria, RMSE, CC, NS~~C~~E and RB, are also shown. ~~Most of the selected precipitation products are 0.25° resolutions except GSMAP-MVK+. For 0.25° precipitation products, there are 11 grids locating inside our study basin, and there are 11 in situ rainfall gauges in study basin also. The number is the same, and therefore we used basin average rainfall amount in our evaluations.~~ GSMAP-MVK+ is the best product and PERSIANN ~~is~~has the ~~worst~~poorest performance with respect to RMSE and NS~~C~~E. GSMAP-MVK+ is also the best with respect to CC, while GLDAS ~~is~~has the ~~worst~~poorest performance with a CC value~~s~~ of 0.55. With respect to RB, APHRODITE performs best and GSMAP-MVK+ the second best, while TRMM3B42RT the ~~worst~~least best with an RB value of -38%. None of the products can outperform others in terms of all the statistical criteria~~, which~~. This may be due to the different limitations of satellite sensors and inverse algorithms ~~used to produce the~~of precipitation products. This situation shows that the selection of the best precipitation products is difficult.

TRMM3B42RT and TRMM3B42 underestimate precipitation amounts. This underestimation may be because convective rainfall always happens in summer in northeast China (Shou and Xu, 2007a, b;Yuan et al., 2010), and indicates the limitation of TRMM algorithms in high latitude regions with convective rainfall. This type of rainfall has a large rainfall amount within a short time period and, therefore, cannot be captured by microwave imager. This type of rainfall may also have a thick cloud that is impenetrable by infrared (Ebert et al., 2007). Thus microwave and infrared estimation could underestimate rainfall. Compared with TRMM3B42RT, TRMM3B42 provides an improvement in RB. This improvement may be attributed to the assimilation with ~~precipitation radar,~~ gauge data and histogram matching. Compared with APHRODITE and GSMAP-MVK+, TRMM3B42 has low accuracy as represented by RB. This implies that the retrieval algorithm used by TRMM3B42 still needs to be improved with respect to RB. The reason why APHRODITE outperforms TRMM3B42 is that APHRODITE is a gauge-based product. GSMAP-MVK+ outperforms TRMM3B42 in terms of RB may be due to the cloud motion vectors it uses. ~~This information suggests that TRMM3B42 could be improved with respect to RB by the use of cloud motion vectors or APHRODITE. In addition, the artificial neural network function used by PERSIANN could help to improve RB, because only PERSIANN overestimates rainfall amount among all the products. In terms of RMSE, NSCE and CC, TRMM3B42 may be improved using the cloud motion vectors, since GSMAP-MVK+ provides the lowest RMSE and highest CC and NSCE of all the analyzed products. Compared with GSMAP-MVK+, GLDAS/Noah precipitation shows low accuracy in all the criteria, which indicates the gauge correction method in GLDAS/Noah is not as effective as the cloud motion vectors in GSMAP-MVK+,~~Compared with GSMAP-MVK+, GLDAS/Noah precipitation shows low accuracy in all the criteria even though they use similar data sources: IR and MW.

Comparison of precipitation product data and gauge observations at a monthly scale is shown in Fig. 4. Here, the APHRODITE product (Fig. 4d) performs best based on RMSE, CC, NS~~C~~E and RB. GLDAS/Noah is the poorest in terms of RMSE and NS~~C~~E. With respect to CC, GLDAS and TRMM3B42 are ~~equal worst~~equally poor, with CC values of 0.81. The results also show that PERSIANN overestimates precipitation amount, while Li et al. (2013) found PERSIANN underestimates rainfall in south China. This may be attributed to the different latitudes of the study regions.

Fig. 5 shows time series of average monthly precipitation data against gauge observations during the period 2000-2007. Each curve represents a different precipitation product. GLDAS data (Fig. 5a) seriously underestimate high rainfall. Similarly, TRMM3B42RT underestimates peak precipitation intensity also. Comparatively, APHRODITE, PERSIANN, TRMM3B42 and GSMAP-MVK+ have better performances.

3.2 Inter-annual evaluations

Fig. ~~9~~6 shows the inter-annual average monthly precipitation. Each curve represents a different product data. PERSIANN overestimates in all the 12 months, while others underestimate, especially during the summer. This may result from the artificial neural network function and limitations of infrared and microwave estimation. APHRODITE data are relatively close to observations. Compared with TRMM3B42RT, TRMM3B42 is better, which indicates the ~~precipitation radar,~~ gauge corrections and histogram matching used by

TRMM3B42 impact positively on accuracy. During the summer~~ (Jul and Aug),~~, discrepancies between products become larger~~, implying that uncertainties in rainfall estimates from different products increase.~~. With a decrease of rainfall magnitude, the discrepancies between products reduce~~, indicating that uncertainties in rainfall estimates decrease~~. This information implies that the differences in precipitation estimation algorithms are related to precipitation magnitudes: the larger the rainfall magnitudes, the greater the differences.

3.3 Probability distribution evaluations

Fig. ~~10~~7 shows cumulative probability distribution functions (CDF) by occurrence (CDFc) and by volume (CDFv) for precipitation products. Probabilities are shown on the *y*~~-~~ axis, and the *x*~~-~~ axis shows rainfall intensity with a 1 mm/day interval log space.

PERSIANN is the best by both occurrence and volume. However, for CDFc, TRMM3B42RT is the ~~worst~~least best, and, for CDFv, TRMM3B42RT and GLDAS/Noah are comparable and worse than others. All precipitation products overestimate ~~probabilities by~~ occurrence and ~~by~~ volume probabilities except ~~larger~~ rainfall ~~intensity.~~intensities of larger than 63mm/day and 53mm/day for occurrence and volume probabilities, respectively. This may be because the precipitation products ~~underestimate rain intensity except that PERSIANN overestimates high rain intensity~~overestimate the intensity of some heavy rainfall (recall the results in section 3.1). The results differ from those of Li et al. (2013), in which PERSIANN ~~performs~~has the ~~worst.~~poorest performance. This ~~results~~may result from differences in ~~latitudes~~study region (in the study of Li et al. (2013)~~, south China was studied). This difference confirms the finding in section 3.1 that the artificial neural network function of PERSIANN may be incapable of reducing the geographical dependence of TRMM~~

~~microwave imager error, and the precision of PERSIANN is influenced by geographical locations.~~, south China was studied).

3.4 Contingency statistics

Fig. ~~11~~8 shows the false alarm ratio, probability of detection~~, false alarm ratio~~ and critical success index for each precipitation product.

PERSIANN has the highest false alarm ratio among the products, while TRMM3B42RT has the lowest. The false alarm ratio of TRMM3B42 is larger than TRMM3B42RT, which indicates that the ~~precipitation radar,~~ gauge corrections and histogram matching used by TRMM3B42 do not provide positive effects on false alarm ratio and may give rise to uncertainty in false alarm ratio. ~~To improve the false alarm ratio of TRMM3B42, using the artificial neural network function would be ineffective;~~ GSMAP-MVK+ has a lower false alarm ratio than TRMM3B42~~, indicating that use of cloud motion vectors could potentially improve TRMM3B42 with respect to the false alarm ratio.~~ .

No obvious trends are observed for the false alarm ratio overall (compared with the probability of detection and critical success index), which means the false alarm ratio dependence on rainfall magnitude is weak. However, Chen et al. (2013a) found the false alarm ratios of TRMM3B42 and TRMM3B42RT to increase with an increase in rainfall intensity. The differences are attributed mainly to observed data. In the study of Chen et al. (2013a), national rain gauge data were employed, whereas in this study more detailed basin data are used.

Among all selected products, GLDAS/Noah has the lowest probability of detection and critical success index during periods of high rainfall intensity, while APHRODITE retains a high probability of detection and critical success index. This is because APHRODITE uses gauge observations, and implies that the APHRODITE algorithm is effective. PERSIANN has comparable probability of detection with APHRODITE~~, which indicates that the artificial neural network function used by PERSIANN would be a good tool for probability of detection.~~. The critical success index of GSMAP-MVK+ is also comparable with APHRODITE~~, suggesting that the cloud motion vectors used by GSMAP-MVK+ would be useful for improving the critical success index~~. Compared with TRMM3B42RT, TRMM3B42 has greater probability of detection and comparable critical success index. This information implies that retrieval algorithm of TRMM3B42 provides positive effects on probability of detection, but ~~not significant~~no obvious positive impacts on critical success index.

Decreasing trends are observed for all products in terms of probability of detection and critical success index, matching the results of Chen et al. (2013a) for TRMM3B42 and TRMM3B42RT. This indicates that probability of detection and critical success index have relatively strong dependence on rainfall magnitude, and implies microwave and infrared precipitation estimation may have relatively strong dependence on rainfall magnitude in terms of probability of detection and critical success index.

4 Hydrological evaluations

4.1 Assessment of hydrological models

WEB-DHM was calibrated against observed discharges of Biliu. Six main parameters were selected to calibrate using a trial and error approach due to the model's computational burden. Model parameter multipliers were calibrated, similar to the study by Wang et al. (2011). The

'Trial and error' approach has two steps. First, all the multiplier values are set to 1 which represents the default parameter values from Food and Agriculture Organization (FAO) (2003) and SiB2 model. Second, varying the multiplier values until acceptable discharge simulation accuracy is obtained. The calibrated parameter values are listed in Table 2. The simulated daily, monthly and inter-annual results are shown in Figs. 9a, 9c and 9e.

TOPMODEL uses basin-averaged parameter values, and these parameter values are estimated by experience or observation. However, these methods do not give precise parameter values. Therefore, the parameter values are considered as uncertain and provided with ranges based on experience (Beven and Kirkby, 1979;Beven and Freer, 2001a, b;Peters et al., 2003). Six parameters of TOPMODEL were calibrated using the dynamically dimensioned search algorithm (Tolson and Shoemaker, 2007), and the results are given in Table 3. The simulated daily, monthly and inter-annual results are shown in Figs. 9b, 9d and 9f.

Note that the parameters of TOPMODEL and WEB-DHM were calibrated using observed precipitation data, and the accuracy of simulated discharges has been validated using gauge observations. Comparison with the parameter values reported in previous research shows the parameter values are appropriate (Beven and Freer, 2001a;Peters et al., 2003;Qi et al., 2015).

4.2 Daily scale discharges

Figs. ~~12~~10 and ~~13~~11 display scatterplots of discharges during the period 2000-2007 simulated using WEB-DHM and TOPMODEL against gauge observations at a daily scale. Two criteria, NS~~C~~E and RB, are shown. It should been noted that the start dates are different for precipitation products, and observed data were used when product data are not available: from 1 January 2000 to 29 February 2000 for TRMM3B42RT, GSMAP-MVK+ and

PERSIANN; from 1 January 2000 to 23 February 2000 for GLDAS/Noah. These time periods were not considered for accuracy comparison.

In the case of WEB-DHM simulations, the best NS~~C~~E (0.41) corresponds with APHRODITE (Fig. 1~~2~~0d), while the best value for RB (1%) corresponds with GLDAS/Noah. In the case of TOPMODEL simulations, the best NS~~C~~E (0.41) corresponds with APHRODITE, and the best value for RB (-24%) corresponds with APHRODITE also. Although the best NS~~C~~E is the same for both WEB-DHM and TOPMODEL simulations and corresponding product is also the same, there is a large difference in the best RB values. ~~Big differences exist in large discharges as well, and TOPMODEL underestimates almost all larger discharges except in a simulation using PERSIANN.~~ At the daily scale precipitation amount evaluation, the ~~worst~~least best RB is -38%, corresponding with TRMM3B42RT (Fig. ~~6c~~3c). However, in WEB-DHM discharge simulation, the ~~worst~~least best RB (218%) corresponds with PERSIANN, and, in TOPMODEL simulation, the ~~worst~~least best RB (-62%) corresponds with TRMM3B42RT. These differences stem from differences in hydrological models and interactions between hydrological models and precipitation product data.

All RB criteria at the daily scale precipitation evaluations (recall the results in Fig. ~~6~~3) are ~~nonlinearly~~ amplified by TOPMODEL, while in the case of WBE-DHM, some are amplified and the others are decreased. For example, for GLDAS and PERSIANN, the RB criteria at the daily scale precipitation evaluations are -27% and 28%, but they are -50% and 31% in TOPMODEL simulations; they are 1% and 218% in WEB-DHM simulations. These differences result from the influence of hydrological models and interactions between precipitation products and hydrological models. These results reveal that a hydrological model can amplify uncertainties in input data but also reduce uncertainties, ~~and that the~~

~~propagation of uncertainties from precipitation data to discharge simulations is nonlinear,~~ which may be due to the nonlinear runoff generation process in hydrological models. This finding is consistent with the research by Yong et al. (2010).

4.3 Monthly scale discharges

Figs. ~~14~~12 and ~~15~~13 display scatterplots of discharges during the period 2000-2007 simulated using WEB-DHM and TOPMODEL against gauge observations at a monthly scale. ~~Two criteria, NSCE and RB, are shown.~~

In the case of WEB-DHM, the best NS~~C~~E and RB values are 0.73 and 1%, which corresponding with TRMM3B42 and GLDAS respectively. In the case of TOPMODEL, they are 0.58 and -24%, corresponding with PERSIANN and APHRODITE respectively. ~~In the case of the WEB-DHM simulation using PERSIANN data, almost all peak discharges are overestimated, but in the case of TOPMODEL simulations, many peak discharges are underestimated using PERSIANN data. These differences are mainly because of hydrological model influence and interactive influence between precipitation data and models.~~

The combination of WEB-DHM and TRMM3B42 shows a great performance, with NS~~C~~E and RB values of up to 0.73 and -7%, even though TRMM3B42 is not the best in monthly scale precipitation data evaluation. This reveals the influence of different characterizations of hydrological processes on the selection of precipitation data, implying that accurate discharge simulation does not solely depend on the accuracy of a precipitation product.

At the monthly scale, although APHRODITE is the best precipitation product and

WEB-DHM model has better performance than TOPMODEL in calibration (Figs. ~~3c~~9c and ~~3d~~9d), the combination of APHRODITE and WEB-DHM is not better in the discharge simulation, which can be shown by comparing Fig. ~~14d with Fig. 15d. This is~~12d with Fig. 13d (the RB and NSE of WEB-DHM and APHRODITE combination are -37% and 0.5, but they are -24% and 0.51 for the combination of TOPMODEL and APHRODITE). This could be due to the interactive influence between hydrological models and precipitation products, and implies that the interactions between models and products could be ~~significant~~large and have a big influence on discharge simulations. In addition, comparison of Figs. 1~~4~~2d and 1~~4~~2b shows that discharge simulation of APHRODITE is worse than TRMM3B42, even though APHRODITE is the best precipitation product in terms of all the selected criteria at a monthly scale precipitation amount evaluation. This information shows that better precipitation products do not guarantee better discharge simulations. ~~This is counter intuitive because the influence of interactions between precipitation products and hydrological models is usually ignored: thus our result provides improved understanding about discharge simulations using precipitation products.~~ These results imply that, although the satellite-based precipitation products are not as accurate as gauge-based products in rainfall amount estimation, they could have a better performance in discharge simulations if the combination of precipitation product and hydrological model is good.

~~Even though WEB-DHM has a better performance than TOPMODEL in calibration (Figs. 3c and 3d), discharge simulation results may not be better than TOPMODEL (for example, the results shown in Fig. 15f are better than those in 14f). This implies that complex and advanced models could lose their advantages if the input data and models were not good coalitions, and models which perform worse in calibration could lead to better discharge simulations due to the interactive influence between models and precipitation products.~~

4.~~3~~4 Inter-annual average monthly discharges

Fig. 14~~16~~ shows inter-annual average monthly discharges of all selected precipitation products during the period 2000-2007. In the case of TOPMODEL, PERSIANN agrees well with gauge observations, and all products underestimate discharges in August. In the case of WEB-DHM, GLDAS data and TRMM3B42 data have a better performance than other data but, with the exception of PERSIANN, all products underestimate peak discharge in August. The simulation results show huge differences even though Figs. ~~3e~~9e and ~~3f~~9f show TOPMODEL and WEB-DHM have almost the same performance using observed data; this is because of the impacts of interactive influence between hydrological models and precipitation products. ~~This shows that similar hydrological model performance and the same precipitation product do not imply similar discharge simulation accuracy.~~

4.~~4~~5 Uncertainty source quantification

All above results suggest that discharge simulations are influenced by precipitation products, hydrological models and interactions between hydrological models and precipitation products. Thus it is essential to quantify the respective influence. Figs. 1~~7~~5a and 1~~7~~5b show contributions of precipitation products, hydrological models and their interactions to uncertainties in monthly average discharges and different flow quantiles respectively. Fig. 1~~7~~5b shows quantiles computed at a daily time step. The contributions of uncertainty sources are represented by stripes.

Fig. 1~~7~~5a shows that precipitation data contribute most of the uncertainty in discharges, and

~~is~~contribute more ~~significant~~ than hydrological models. Interactions between hydrological models and precipitation products have ~~significant~~large contributions, at a similar level to those from hydrological models. In summer (July to September), the contribution of precipitation data is less than most other months except March. However, the uncertainty in precipitation intensity increases in summer (recall the results in section 3.2). In non-summer months except March, the uncertainty contribution from precipitation products is larger than in summer. These differences maybe result from the nonlinear propagation of uncertainty through hydrological models. In March, the contribution of hydrological models is larger than in other months, which may result from the decrease in influences of interactions and precipitation products, and from the nonlinear influence of the hydrological models.

Fig. 1~~7~~5b shows that, for small discharges (smaller than 25% quantile which corresponds to an observed discharge value of 1.79m^3/s) and large discharges~~,~~ (larger than 99% quantile which corresponds to an observed discharge value of 157m^3/s), hydrological models contribute most of the uncertainties. For middle magnitude flows (between 25% and 99% quantiles), precipitation products contribute the majority, and the contribution of interactions is not negligible and of similar magnitude to the contribution from hydrological models. The contribution of interactions is larger for middle magnitude flows than for small and large discharges. ~~This may result from an increase in interaction, implying that interaction effects are influenced by discharge magnitude.~~ The different contributions of interactions for various magnitude flows may be because different magnitude rainfall data could trigger different hydrological processes (Herman et al., 2013). Small discharges mainly come from ~~groundwater flow~~base flows which ~~is~~are relatively stable and do~~es~~ not need much rainfall to be triggered, and large ~~flow is~~discharges are mainly controlled by overland flows when heavy precipitation occurs. Middle magnitude ~~flow, on the other hand, consists~~discharges consist of contributions from ~~groundwater~~base flows, lateral subsurface flows and overland flows. It is more complex and can be triggered by various magnitude rainfalls - thus interactions are more changeable.

Although heavy rainfall data have high uncertainty (recall the results in section 3.1), precipitation products ~~don't~~do not contribute the most uncertainty in large discharges (Fig. 1~~7~~5b). This may be because the nonlinear propagation of uncertainty through hydrological models enlarges the influence of hydrological models, and implies that high uncertainties in extreme rainfall do not mean high uncertainties in extreme discharges. ~~This information suggests that for better implementation of precipitation products, precipitation products may need to provide some information on combinations with hydrological models and on the suitability of the combinations for different magnitude flows.~~

In this study, because hydrological model parameters were calibrated using gauge observations, the hydrological model parameter uncertainty was not considered. Although the uncertainty contribution results in this study may not be transferable to other basins, the proposed framework provides a useful tool for quantifying uncertainty contributions in discharge simulations using precipitation products.

5 Discussion

The spatial distribution of different precipitation data is not considered in this study. The study region is a small river basin, as shown in Fig. 1, there are only 11 grids inside the basin boundary for the precipitation products with a spatial resolution of 0.25 degree. Within a grid of 0.25 degree, there are no differences in precipitation amount between the 300 m × 300 m grids used in hydrological models, and differences exist at the level of 0.25 degree grids only.

Sapriza-Azuri et al. (2015) suggested that the spatial variability of precipitation has little influence on rapidly responding river discharges; this study is the case because the flow transport time from the most upper part of the basin to the downstream discharge gauge is 6 hours, which is shorter than the daily and monthly time steps of discharges investigated. Therefore, the spatial distributions of precipitation products with a resolution of 0.25 degree in the relatively small river basin have little influence on the simulated discharges. However, the assumption of uniform distribution can be regarded as another uncertainty source against spatial variability, and its influence can be assessed using the proposed uncertainty quantification framework. This will allow us to compare the relative contributions of the assumption to those from other sources such as hydrological models, which will be investigated using a much larger river basin in the future work.

In addition to improving the accuracy of precipitation products, a good collation could help to achieve the performance in discharge simulations. Our approach provides a way to assess the different coalitions, i.e., the overall uncertainties in simulated discharges from different combinations of hydrological models and precipitation products. More precipitation products and hydrological models should be included and tested in the future work.

It should be noted that other input data including temperature, downward solar radiation, long wave radiation, air pressure, wind speed and humidity may also have uncertainties. However, Fig. 9 shows that the simulated discharge data are acceptable particularly at monthly and inter-annual scales using these data. Research has shown that the land surface temperatures are highly accurate compared with MODIS satellite land surface temperature observations (Wang et al., 2011;Qi et al., 2015). Thus, the uncertainties from the other inputs are not considered in our case study river basin.

In this study, the parameter values calibrated using gauge observations are not tuned to a specific product. That is, there is little compensation of model parameters for the errors in input precipitation data. The differences in model accuracy mainly results from the different representations of hydrological processes. That is, the errors in precipitation products are primarily compensated by the different representations of model processes.

6 Summary and conclusions

This research ~~comprehensively~~ assesses the applicability of six precipitation products with fine spatial and temporal resolutions at a high latitude region in northeast China~~. This area is usually neglected in precipitation data accuracy assessments. Both~~ using both statistical and hydrological evaluation methods at multi-temporal scales ~~were used in the assessment~~. A framework is proposed to quantify uncertainty contributions of precipitation products, hydrological models and their interactions to simulated discharges. These products ~~include~~are TRMM version 7 products (TRMM3B42 and TRMM3B42RT), GLDAS, APHRODITE, PERSIANN and GSMAP-MVK+. The fully distributed WEB-DHM and semi-distributed TOPMODEL were employed to investigate the influence of hydrological models on simulated discharges. The results show the uncertainty characteristics of the six products, and reveal strategies that could improve precipitation products. This information could be able to provide references for future precipitation product development. The proposed framework can reveal hydrological simulation uncertainties using precipitation products: thus provides useful information on precipitation product applications. The following conclusions are presented on the basis of this study.

First, at daily scale, selecting the best precipitation products is very difficult, while, at a monthly scale, APHRODITE has the best performance in terms of NS~~C~~E, RB, RMSE, and CC, and also retains a high probability of detection and critical success index. This information implies that the APHRODITE algorithm is effective, and APHRODITE could be a very good data set to refine and validate satellite-based precipitation products.

Second, ~~the cloud motion vectors used by~~ GSMAP-MVK+ show huge advantage, and ~~could be used to improve~~is better than TRMM3B42 in RB, NS~~C~~E, RMSE, CC, false alarm ratio and critical success index, while ~~the artificial neural network function used by~~ PERSIANN ~~could be useful for improvements of~~is better than TRMM3B42 in probability of detection~~,~~ and precipitation probability distribution estimation~~ and heavy rainfall estimation~~. At present, the new precipitation estimation mission - NASA Global Precipitation Measurement (GPM) - combines the artificial neural network function of PERSIANN and precipitation radar-matching of TRMM Multi-satellite Precipitation Analysis. However, ~~this~~the above finding implies that incorporating ~~cloud motion vectors~~GSMAP-MVK+ estimation approach into GPM could be ~~better~~useful as well.

Third, it is found that, although high uncertainty exists in heavy rainfall, hydrological models contribute mostly to the uncertainty in extreme discharges. This may result from the nonlinear propagation of uncertainty through hydrological models enlarges the influence of hydrological models, and implies that high uncertainties in extreme rainfall do not mean high uncertainties in extreme discharges.

Fourth, interactions between hydrological models and precipitation products contribute ~~significantly~~a lot to uncertainty in discharge simulations, and interactive impacts are influenced by discharge magnitude. Because of interactive effects, for hydrological models with similar performances in calibration, using the same precipitation products for discharge simulations does not provide a similar level of accuracy in discharge simulations, and in fact ~~significantly~~very different predictions could be obtained. In addition, this finding implies that only considering precipitation products or hydrological model uncertainties could result in overestimation of precipitation product contribution and hydrological model contribution to discharge uncertainty.

Fifth, a good discharge simulation depends on a good coalition of a hydrological model and a precipitation product, and a better precipitation product does not necessarily guarantee a better discharge simulation~~, whereas it is possible that a hydrological model which is worse in calibration may generate a better discharge simulation.~~. This~~ information~~ suggests that, although the satellite-based precipitation products are not as accurate as the gauge-based product, they could have better performance in discharge simulations when appropriately combined with hydrological models. It should be noted that this finding should be further tested with more river basins, in particular large river basins accounting for spatial variability in precipitation products.

In the future, calculating deterministic discharge simulations considering precipitation product uncertainties and hydrological model uncertainties together should be studied because above results show product uncertainties and model uncertainties all are important. In addition, recalibrating hydrological models using precipitation products may reduce the interactive influence between hydrological models and precipitation products on simulated discharges, and this may explain why recalibration can improve discharge simulation accuracy. This should be verified in future work. Further, future ~~work is encouraged to~~

~~develop good collations with hydrological models, as a good collation could help to achieve the same aim as improving the accuracy of precipitation products: precisely simulating discharges. Future~~ research is ~~also~~ encouraged to incorporate ~~cloud motion vectors~~GSMAP-MVK+ estimation approach into GPM because of the good performance of ~~cloud motion vectors.~~GSMAP-MVK+.

Acknowledgements:

This study was supported by the National ~~Basic Research Programme (973 programme)~~Natural Science Foundation of China (Grant No. ~~2013CB036400~~51320105010 and 51279021). The first author gratefully acknowledges the financial support provided by the China Scholarship Council. The APHRODITE data were downloaded from http://www.chikyu.ac.jp/precip/pro-ducts/index.html. The TRMM 3B42 data are downloaded from http://mirador.gsfc.nasa.gov/cgi-bin/mirador/presentNavigation.pl?tree=project&project=TRMM&dataGroup=Gridded. TRMM 3B40RT data are downloaded from ftp://trmmopen.nascom.nasa.gov/pub/merged/combinedMicro/. TRMM 3B41RT data are downloaded from ftp://trmmopen.nascom.nasa.gov/pub/merged/calibratedIR/. TRMM3B42RT data are downloaded from ftp://trmmopen.nascom.nasa.gov/pub/merged/mergeIRMicro/. PERSIANN data are downloaded from http://chrs.web.uci.edu/persiann/data.html. GSMAP_MVK data are downloaded from http://sharaku.eorc.jaxa.jp/GSMaP_crest/. The GLDAS data are downloaded from http://mirador.gsfc.nasa.gov/cgi-bin/mirador/homepageAlt.pl?keyword=GLDAS_NOAH025SUBP_3H. The data of Biliu basin were obtained from the Biliu reservoir administration.

References

Aonashi, K., Awaka, J., Hirose, M., Kozu, T., Kubota, T., Liu, G., Shige, S., Kida, S., Seto, S., Takahashi, N., and Takayabu, Y. N.: GSMaP Passive Microwave Precipitation Retrieval Algorithm: Algorithm Description and Validation, Journal of the Meteorological Society of Japan, 87A, 119-136, 10.2151/jmsj.87A.119, 2009.

Artan, G., Gadain, H., Smith, J. L., Asante, K., Bandaragoda, C. J., and Verdin, J. P.: Adequacy of satellite derived rainfall data for stream flow modeling, Natural Hazards, 43, 167-185, 10.1007/s11069-007-9121-6, 2007.

Asadullah, A., McIntyre, N., and Kigobe, M. A. X.: Evaluation of five satellite products for estimation of rainfall over Uganda, Hydrological Sciences Journal, 53, 1137-1150, 10.1623/hysj.53.6.1137, 2008.

Bastola, S., Ishidaira, H., and Takeuchi, K.: Regionalisation of hydrological model parameters under parameter uncertainty: A case study involving TOPMODEL and basins across the globe, Journal of Hydrology, 357, 188-206, 10.1016/j.jhydrol.2008.05.007, 2008.

Betts, A. K., Chen, F., Mitchell, K. E., and Janjic, Z. I.: Assessment of the land surface and boundary layer models in two operational versions of the NCEP Eta Model using FIFE data, Monthly Weather Review, 125, 2896-2916, 10.1175/1520-0493(1997)125<2896:aotlsa>2.0.co;2, 1997.

Beven, K. J., and Binley, A.: The future of distributed models：model calibration and uncertainty prediction, Hydrological Processes, 6, 279-298, 10.1002/hyp.3360060305,

1992.

Beven, K. J., and Freer, J. E.: A dynamic TOPMODEL, Hydrological Processes, 15, 1993-2011, 10.1002/hyp.252, 2001a.

Beven, K. J., and Freer, J. E.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, Journal of Hydrology, 249, 11–29, 10.1016/S0022-1694(01)00421-8, 2001b.

Beven, K. J., and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology, Hydrological Sciences Bulletin, 24, 43-69, 10.1080/02626667909491834, 1979.

Blasone, R.-S., Vrugt, J. A., Madsen, H., Rosbjerg, D., Robinson, B. A., and Zyvoloski, G. A.: Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling, Advances in Water Resources, 31, 630-648, 10.1016/j.advwatres.2007.12.003, 2008.

Blazkova, S., and Beven, K.: Flood frequency prediction for data limited catchments in the Czech Republic using a stochastic rainfall model and TOPMODEL, Journal of Hydrology, 195, 256-278, 10.1016/S0022-1694(96)03238-6, 1997.

Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M., and Schär, C.: Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections, Water Resources Research, 49, 1523-1536, 10.1029/2011wr011533, 2013.

Bouilloud, L., Chancibault, K., Vincendon, B., Ducrocq, V., Habets, F., Saulnier, G.-M., Anquetin, S., Martin, E., and Noilhan, J.: Coupling the ISBA Land Surface Model and the

TOPMODEL Hydrological Model for Mediterranean Flash-Flood Forecasting: Description, Calibration, and Validation, Journal of Hydrometeorology, 11, 315-333, 10.1175/2009jhm1163.1, 2010.

Buarque, D. C., de Paiva, R. C. D., Clarke, R. T., and Mendes, C. A. B.: A comparison of Amazon rainfall characteristics derived from TRMM, CMORPH and the Brazilian national rain gauge network, Journal of Geophysical Research, 116, 10.1029/2011jd016060, 2011.

Cameron, D. S., Beven, K. J., Tawn, J., Blazkova, S., and Naden, P.: Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty), Journal of Hydrology, 219, 169-187, 10.1016/S0022-1694(99)00057-8, 1999.

Chen, F., Mitchell, K., Schaake, J., Xue, Y., Pan, H.-L., Koren, V., Duan, Q. Y., Ek, M., and Betts, A.: Modeling of land surface evaporation by four schemes and comparison with FIFE observations, Journal of Geophysical Research, 101, 7251, 10.1029/95jd02165, 1996.

Chen, S., Hong, Y., Cao, Q., Gourley, J. J., Kirstetter, P.-E., Yong, B., Tian, Y., Zhang, Z., Shen, Y., Hu, J., and Hardy, J.: Similarity and difference of the two successive V6 and V7 TRMM multisatellite precipitation analysis performance over China, Journal of Geophysical Research: Atmospheres, 118, 13,060-013,074, 10.1002/2013jd019964, 2013a.

Chen, S., Hong, Y., Gourley, J. J., Huffman, G. J., Tian, Y., Cao, Q., Yong, B., Kirstetter, P.-E., Hu, J., Hardy, J., Li, Z., Khan, S. I., and Xue, X.: Evaluation of the successive V6 and V7

TRMM multisatellite precipitation analysis over the Continental United States, Water Resources Research, 49, 8174-8186, 10.1002/2012wr012795, 2013b.

Dai, Y., Zeng, X., Dickinson, R. E., Baker, I., Bonan, G. B., Bosilovich, M. G., Denning, A. S., Dirmeyer, P. A., Houser, P. R., Niu, G., Oleson, K. W., Schlosser, C. A., and Yang, Z.-L.: The Common Land Model, Bulletin of the American Meteorological Society, 84, 1013-1023, 10.1175/bams-84-8-1013, 2003.

Dinku, T., Connor, S. J., Ceccato, P., and Ropelewski, C. F.: Comparison of global gridded precipitation products over a mountainous region of Africa, International Journal of Climatology, 28, 1627-1638, 10.1002/joc.1669, 2008.

Ebert, E. E., Janowiak, J. E., and Kidd, C.: Comparison of Near-Real-Time Precipitation Estimates from Satellite Observations and Numerical Models, Bulletin of the American Meteorological Society, 88, 47-64, 10.1175/bams-88-1-47, 2007.

Ek, M. B.: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, Journal of Geophysical Research, 108, 10.1029/2002jd003296, 2003.

Freer, J. E., Beven, K. J., and Ambroise, B.: Bayesian Estimation of Uncertainty in Runoff Prediction and the Value of Data: An Application of the GLUE Approach, Water Resources Research, 32, 2161-2173, 10.1029/95wr03723, 1996.

Gallart, F., Latron, J., Llorens, P., and Beven, K. J.: Upscaling discrete internal observations for obtaining catchment-averaged TOPMODEL parameters in a small Mediterranean mountain basin, Physics and Chemistry of the Earth, 33, 1090-1094, 10.1016/j.pce.2008.03.003, 2008.

Gao, Y. C., and Liu, M. F.: Evaluation of high-resolution satellite precipitation products using rain gauge observations over the Tibetan Plateau, Hydrology and Earth System Sciences, 17, 837-849, 10.5194/hess-17-837-2013, 2013.

Heidari, A., Saghafian, B., and Maknoon, R.: Assessment of flood forecasting lead time based on generalized likelihood uncertainty estimation approach, Stochastic Environmental Research and Risk Assessment, 20, 363-380, 10.1007/s00477-006-0032-y, 2006.

Herman, J. D., Reed, P. M., and Wagener, T.: Time‐varying sensitivity analysis clarifies the effects of watershed model formulation on model behavior, Water Resources Research, 49, 1400-1414, 10.1002/wrcr.20124, 2013.

Hong, Y., Hsu, K.-l., Moradkhani, H., and Sorooshian, S.: Uncertainty quantification of satellite precipitation estimation and Monte Carlo assessment of the error propagation into hydrologic response, Water Resources Research, 42, 10.1029/2005wr004398, 2006.

Hossain, F., and Anagnostou, E. N.: Assessment of a stochastic interpolation based parameter sampling scheme for efficient uncertainty analyses of hydrologic models, Computers & Geosciences, 31, 497-512, 10.1016/j.cageo.2004.11.001, 2005.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., and Stocker, E. F.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales, Journal of Hydrometeorology, 8, 38-55, 10.1175/jhm560.1, 2007.

Jiang, S., Ren, L., Hong, Y., Yong, B., Yang, X., Yuan, F., and Ma, M.: 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, Journal of Hydrology, 452-453, 213-225, 10.1016/j.jhydrol.2012.05.055, 2012.

Kato, H., Rodell, M., Beyrich, F., Cleugh, H., Gorsel, E. v., Liu, H., and Meyers, T. P.: Sensitivity of Land Surface Simulations to Model Physics, Land Characteristics, and Forcings, at Four CEOP Sites, Journal of the Meteorological Society of Japan, 85A, 187-204, 10.2151/jmsj.85A.187, 2007.

Kneis, D., Chatterjee, C., and Singh, R.: Evaluation of TRMM rainfall estimates over a large Indian river basin (Mahanadi), Hydrology and Earth System Sciences, 18, 2493-2502, 10.5194/hess-18-2493-2014, 2014.

Koren, V., Schaake, J., Mitchell, K., Duan, Q. Y., Chen, F., and Baker, J. M.: A parameterization of snowpack and frozen ground intended for NCEP weather and climate models, Journal of Geophysical Research, 104, 19569, 10.1029/1999jd900232, 1999.

Koster, R. D., and Suarez, M. J.: Modeling the land surface boundary in climate models as a composite of independent vegetation stands, Journal of Geophysical Research-Atmospheres, 97, 2697-2715, 10.1029/91JD01696, 1992.

Kubota, T., Shige, S., Hashizurne, H., Aonashi, K., Takahashi, N., Seto, S., Hirose, M., Takayabu, Y. N., Ushio, T., Nakagawa, K., Wanami, K., Kachi, M., and Okamoto, K.: Global precipitation map using satellite-borne microwave radiometers by the GSMaP project: Production and validation, IEEE Transactions on Geoscience and Remote

Sensing, 45, 2259-2275, 10.1109/tgrs.2007.895337, 2007.

Kuczera, G., Kavetski, D., Franks, S., and Thyer, M.: Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters, Journal of Hydrology, 331, 161-177, 10.1016/j.jhydrol.2006.05.010, 2006.

Kuczera, G., and Parent, E.: Monte Carlo assessment of parameter uncertainty in conceptual catchment models- the Metropolis algorithm, Journal of Hydrology, 211, 69-85, 10.1016/S0022-1694(98)00198-X, 1998.

Kummerow, C., Simpson, J., and Thiele, O.: The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit, Journal of Applied Meteorology and Climatology 39, 10.1175/1520-0450(2001)040<1965:TSOTTR>2.0.CO;2, 2000.

Li, Z., Yang, D., and Hong, Y.: Multi-scale evaluation of high-resolution multi-sensor blended global precipitation products over the Yangtze River, Journal of Hydrology, 500, 157-169, 10.1016/j.jhydrol.2013.07.023, 2013.

Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J.: A simple hydrologically based model of land-surface water and energy fluxes for general-circulation models, Journal of Geophysical Research-Atmospheres, 99, 14415-14428, 10.1029/94jd00483, 1994.

Maggioni, V., Vergara, H. J., Anagnostou, E. N., Gourley, J. J., Hong, Y., and Stampoulis, D.: Investigating the Applicability of Error Correction Ensembles of Satellite Rainfall Products in River Flow Simulations, Journal of Hydrometeorology, 14, 1194-1211, 10.1175/jhm-d-12-074.1, 2013.

Meng, J., Li, L., Hao, Z., Wang, J., and Shao, Q.: Suitability of TRMM satellite rainfall in driving a distributed hydrological model in the source region of Yellow River, Journal of Hydrology, 509, 320-332, 10.1016/j.jhydrol.2013.11.049, 2014.

Müller, M. F., and Thompson, S. E.: Bias adjustment of satellite rainfall data through stochastic modeling: Methods development and application to Nepal, Advances in Water Resources, 60, 121-134, 10.1016/j.advwatres.2013.08.004, 2013.

Nikolopoulos, E. I., Anagnostou, E. N., Hossain, F., Gebremichael, M., and Borga, M.: Understanding the Scale Relationships of Uncertainty Propagation of Satellite Rainfall through a Distributed Hydrologic Model, Journal of Hydrometeorology, 11, 520-532, 10.1175/2009jhm1169.1, 2010.

Ochoa, A., Pineda, L., Crespo, P., and Willems, P.: Evaluation of TRMM 3B42 precipitation estimates and WRF retrospective precipitation simulation over the Pacific–Andean region of Ecuador and Peru, Hydrology and Earth System Sciences, 18, 3179-3193, 10.5194/hess-18-3179-2014, 2014.

Pan, M., Li, H., and Wood, E.: Assessing the skill of satellite-based precipitation estimates in hydrologic applications, Water Resources Research, 46, 10.1029/2009wr008290, 2010.

Parton, W. J., and Logan, J. A.: A model for diurnal variation in soil and air temperature, Agricultural Meteorology, 23, 205-216, 10.1016/0002-1571(81)90105-9, 1981.

Peng, Z., Wang, Q. J., Bennett, J. C., Pokhrel, P., and Wang, Z.: Seasonal precipitation forecasts over China using monthly large-scale oceanic-atmospheric indices, Journal of Hydrology, 519, 792-802, 10.1016/j.jhydrol.2014.08.012, 2014a.

Peng, Z., Wang, Q. J., Bennett, J. C., Schepen, A., Pappenberger, F., Pokhrel, P., and Wang, Z.:

Statistical calibration and bridging of ECMWF System4 outputs for forecasting seasonal precipitation over China, Journal of Geophysical Research: Atmospheres, 119, 7116-7135, 10.1002/2013jd021162, 2014b.

Peters, N. E., Freer, J., and Beven, K.: Modelling hydrologic responses in a small forested catchment (Panola Mountain, Georgia, USA): a comparison of the original and a new dynamic TOPMODEL, Hydrological Processes, 17, 345-362, 10.1002/hyp.1128, 2003.

Qi, W., Zhang, C., Chu, J., and Zhou, H.: Sobol''s sensitivity analysis for TOPMODEL hydrological model: A case study for the Biliu River Basin, China, Journal of Hydrology and Environment Research, 1, 1-10, 2013.

Qi, W., Zhang, C., Fu, G., and Zhou, H.: Global Land Data Assimilation System data assessment using a distributed biosphere hydrological model, Journal of Hydrology, 528, 652-667, 10.1016/j.jhydrol.2015.07.011, 2015.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin*, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, Bulletin of the American Meteorological Society, 85, 381-394, 10.1175/bams-85-3-381, 2004.

Sapriza-Azuri, G., Jódar, J., Navarro, V., Slooten, L. J., Carrera, J., and Gupta, H. V.: Impacts of rainfall spatial variability on hydrogeological response, Water Resources Research, 51, 1300-1314, 10.1002/2014wr016168, 2015.

Sellers, P. J.: Modeling the Exchanges of Energy, Water, and Carbon Between Continents and the Atmosphere, Science, 275, 502-509, 10.1126/science.275.5299.502, 1997.

Sellers, P. J., Mintz, Y., Sud, Y. C., and Dalcher, A.: A Simple Biosphere Model (SIB) for Use within General Circulation Models, Journal of the Atmospheric Sciences, 43, 505-531, 10.1175/1520-0469(1986)043<0505:ASBMFU>2.0.CO;2, 1986.

Sellers, P. J., Randall, D. A., Collatz, G. J., Berry, J. A., Field, C. B., Dazlich, D. A., Zhang, C., Collelo, G. D., and Bounoua, L.: A Revised Land Surface Parameterization (SiB2) for Atmospheric GCMS. Part I: Model Formulation, Journal of Climate, 9, 676-705, 10.1175/1520-0442(1996)009<0676:ARLSPF>2.0.CO;2, 1996.

Serpetzoglou, E., Anagnostou, E. N., Papadopoulos, A., Nikolopoulos, E. I., and Maggioni, V.: Error Propagation of Remote Sensing Rainfall Estimates in Soil Moisture Prediction from a Land Surface Model, Journal of Hydrometeorology, 11, 705-720, 10.1175/2009jhm1166.1, 2010.

Shou, Y., and Xu, J.: The rainstorm and mesoscale convective systems over northeast China in june 2005 I: A synthetic analysis of mcs by conventional observations and satellite data (in Chinese), Acta Meteorologica Sinica, 65, 160-170, 2007a.

Shou, Y., and Xu, J.: The rainstorm and mesoscale convective systems over northeast China in june 2005 II: A synthetic analysis of mcs's dynamical structure by radar and satellite observations (in Chinese), Acta Meteorologica Sinica, 65, 171-182, 2007b.

Shrestha, M., Wang, L., Koike, T., Tsutsui, H., Xue, Y., and Hirabayashi, Y.: Correcting basin-scale snowfall in a mountainous basin using a distributed snowmelt model and remote sensing data, Hydrol. Earth Syst. Sci. Discuss., 10, 11711-11753, 10.5194/hessd-10-11711-2013, 2013.

Sorooshian, S., Gao, X., Hsu, K., Maddox, R. A., Hong, Y., Gupta, H. V., and Imam, B.: Diurnal Variability of Tropical Rainfall Retrieved from Combined GOES and TRMM Satellite Information, Journal of Climate, 15, 983-1001, 10.1175/1520-0442(2002)015<0983:DVOTRR>2.0.CO;2, 2002.

Sorooshian, S., Hsu, K. L., Gao, X., Gupta, H. V., Imam, B., and Braithwaite, D.: Evaluation of PERSIANN system satellite-based estimates of tropical rainfall, Bulletin of the American Meteorological Society, 81, 2035-2046, 10.1175/1520-0477(2000)081<2035:eopsse>2.3.co;2, 2000.

Sorooshian, S., Lawford, R. G., Try, P., Rossow, W., Roads, J., Polcher, J., Sommeria, G., and Schifer, R.: Water and energy cycles: Investigating the links, WMO Bull. , 54, 2005.

Su, F., Hong, Y., and Lettenmaier, D. P.: Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and Its Utility in Hydrologic Prediction in the La Plata Basin, Journal of Hydrometeorology, 9, 622-640, 10.1175/2007jhm944.1, 2008.

Tapiador, F. J., Turk, F. J., Petersen, W., Hou, A. Y., García-Ortega, E., Machado, L. A. T., Angelis, C. F., Salio, P., Kidd, C., Huffman, G. J., and de Castro, M.: Global precipitation measurement: Methods, datasets and applications, Atmospheric Research, 104-105, 70-97, 10.1016/j.atmosres.2011.10.021, 2012.

Tolson, B. A., and Shoemaker, C. A.: Dynamically dimensioned search algorithm for computationally efficient watershed model calibration, Water Resources Research, 43, 10.1029/2005wr004723, 2007.

Vrugt, J. A., ter Braak, C. J. F., Gupta, H. V., and Robinson, B. A.: Equifinality of formal

(DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?, Stochastic Environmental Research and Risk Assessment, 23, 1011-1026, 10.1007/s00477-008-0274-y, 2009b.

Wang, D., Wang, G., and Anagnostou, E. N.: Use of Satellite-Based Precipitation Observation in Improving the Parameterization of Canopy Hydrological Processes in Land Surface Models, Journal of Hydrometeorology, 6, 745-763, 10.1175/JHM438.1, 2005.

Wang, F., Wang, L., Koike, T., Zhou, H., Yang, K., Wang, A., and Li, W.: Evaluation and application of a fine-resolution global data set in a semiarid mesoscale river basin with a distributed biosphere hydrological model, Journal of Geophysical Research, 116, 10.1029/2011jd015990, 2011.

Wang, F., Wang, L., Zhou, H., Saavedra Valeriano, O. C., Koike, T., and Li, W.: Ensemble hydrological prediction-based real-time optimization of a multiobjective reservoir during flood season in a semiarid basin with global numerical weather predictions, Water Resources Research, 48, 10.1029/2011wr011366, 2012.

Wang, L., Koike, T., Yang, D. W., and Yang, K.: Improving the hydrology of the Simple Biosphere Model 2 and its evaluation within the framework of a distributed hydrological model, Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 54, 989-1006, 10.1623/hysj.54.6.989, 2009a.

Wang, L., Koike, T., Yang, K., Jackson, T. J., Bindlish, R., and Yang, D.: Development of a distributed biosphere hydrological model and its evaluation with the Southern Great Plains Experiments (SGP97 and SGP99), Journal of Geophysical Research, 114, 10.1029/2008jd010800, 2009b.

Wang, L., Koike, T., Yang, K., Jin, R., and Li, H.: Frozen soil parameterization in a distributed biosphere hydrological model, Hydrol. Earth Syst. Sci., 14, 557-571, 10.5194/hess-14-557-2010, 2010a.

Wang, L., Koike, T., Yang, K., and Yeh, P. J.-F.: Assessment of a distributed biosphere hydrological model against streamflow and MODIS land surface temperature in the upper Tone River Basin, Journal of Hydrology, 377, 21-34, 10.1016/j.jhydrol.2009.08.005, 2009c.

Wang, L., Wang, Z., Koike, T., Yin, H., Yang, D., and He, S.: The assessment of surface water resources for the semi-arid Yongding River Basin from 1956 to 2000 and the impact of land use change, Hydrological Processes, 24, 1123-1132, 10.1002/hyp.7566, 2010b.

Xie, P., Chen, M., Yang, S., Yatagai, A., Hayasaka, T., Fukushima, Y., and Liu, C.: A Gauge-Based Analysis of Daily Precipitation over East Asia, Journal of Hydrometeorology, 8, 607-626, 10.1175/jhm583.1, 2007.

Xue, X., Hong, Y., Limaye, A. S., Gourley, J. J., Huffman, G. J., Khan, S. I., Dorji, C., and Chen, S.: Statistical and hydrological evaluation of TRMM-based Multi-satellite Precipitation Analysis over the Wangchu Basin of Bhutan: Are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins?, Journal of Hydrology, 499, 91-99, 10.1016/j.jhydrol.2013.06.042, 2013.

Yang, D.: Distributed hydrological model using hillslope discretization based on catchment area function: development and applications, PHD, University of Tokyo, Tokyo, 1998.

Yang, K., Koike, T., and Ye, B.: Improving estimation of hourly, daily, and monthly solar radiation by importing global data sets, Agricultural and Forest Meteorology, 137, 43-55, 10.1016/j.agrformet.2006.02.001, 2006.

Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N., and Kitoh, A.: APHRODITE: Constructing a Long-Term Daily Gridded Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges, Bulletin of the American Meteorological Society, 93, 1401-1415, 10.1175/bams-d-11-00122.1, 2012.

Yong, B., Chen, B., Gourley, J. J., Ren, L., Hong, Y., Chen, X., Wang, W., Chen, S., and Gong, L.: Intercomparison of the Version-6 and Version-7 TMPA precipitation products over high and low latitudes basins with independent gauge networks: Is the newer version better in both real-time and post-real-time analysis for water resources and hydrologic extremes?, Journal of Hydrology, 508, 77-87, 10.1016/j.jhydrol.2013.10.050, 2014.

Yong, B., Hong, Y., Ren, L.-L., Gourley, J. J., Huffman, G. J., Chen, X., Wang, W., and Khan, S. I.: Assessment of evolving TRMM-based multisatellite real-time precipitation estimation methods and their impacts on hydrologic prediction in a high latitude basin, Journal of Geophysical Research, 117, 10.1029/2011jd017069, 2012.

Yong, B., Ren, L.-L., Hong, Y., Wang, J.-H., Gourley, J. J., Jiang, S.-H., Chen, X., and Wang, W.: Hydrologic evaluation of Multisatellite Precipitation Analysis standard precipitation products in basins beyond its inclined latitude band: A case study in Laohahe basin, China, Water Resources Research, 46, 10.1029/2009wr008965, 2010.

Yuan, M., Li, Z., and Zhang, X.: Analysis of a meso scale convective system during a brief torrential rain event in Northeast China (in Chinese), Acta Meteorologica Sinica, 68, 125-136, 10.11676/qxxb2010.013, 2010.

Zhao, T., and Yatagai, A.: Evaluation of TRMM 3B42 product using a new gauge-based analysis of daily precipitation over China, International Journal of Climatology, 34, 2749-2762, 10.1002/joc.3872, 2014.

Zhou, T., Yu, R., Chen, H., Dai, A., and Pan, Y.: Summer Precipitation Frequency, Intensity, and Diurnal Cycle over China: A Comparison of Satellite Data with Rain Gauge Observations, Journal of Climate, 21, 3997-4010, 10.1175/2008jcli2028.1, 2008.

Zhou, X. Y., Zhang, Y. Q., Yang, Y. H., Yang, Y. M., and Han, S. M.: Evaluation of anomalies in GLDAS-1996 dataset, Water Science and Technology, 67, 1718-1727, 10.2166/wst.2013.043, 2013.

Table 1 Precipitation products

Product	Spatial resolution	Temporal resolution	Areal coverage	Start date	Type
TRMM3B42	0.25°	3h	Global 50°N-S	1 Jan 1998	PR+IR+MW+gauge +HM
TRMM3B42RT	0.25°	3h	Global 50°N-S	1 Mar 2000	IR+MW
GLDAS/Noah	0.25°	3h	Global 90°N-60°S	24 Feb 2000	IR+MW+gauge
GSMAP-MVK+	0.1°	1h	Global 60°N-S	1 Mar 2000	IR+MW+CMV
PRRSIANN	0.25°	3h	Global 60°N-S	1 Mar 2000	PR+IR+MW+ANN
APHRODITE	0.25°	1day	60°E-150°E, 15°S-55°N	1 Jan 1961 to 2007	gauge

PR: precipitation radar; IR: infrared estimation; MW: microwave estimation; HM: histogram matching; CMV: cloud motion vectors; ANN: artificial neural network.

Table 2 WEB-DHM parameters

Symbol (units)	Brief description	Basin-averaged value
KS (mm/h)	Saturated hydraulic conductivity for soil surface	26.43
Anik	Hydraulic conductivity anisotropy ratio	11.49
Sstmax (mm)	Maximum surface water storage	42.75
Kg (mm/h)	Hydraulic conductivity for groundwater	0.36
alpha	van Genuchten parameter	0.01
n	van Genuchten parameter	1.88

Table 3 TOPMODEL parameters

Name (units)	Description	Lower bound	Upper bound	Calibration
SZM (m)	form of the exponential decline in conductivity	0.01	0.04	0.019
LNT0 (m^2 h^{-1})	log value of effective lateral saturated transmissivity	-25	1	-11.911
RV (m^2 h^{-1})	hill slope routing velocity	2000	5000	2608.4
SR_{max} (m)	maximum root zone storage	0.001	0.01	0.006
SR_0 (m)	initial root zone deficit	0	0.01	0.005
TD (m h^{-1})	unsaturated zone time delay per unit deficit	2	4	2.885

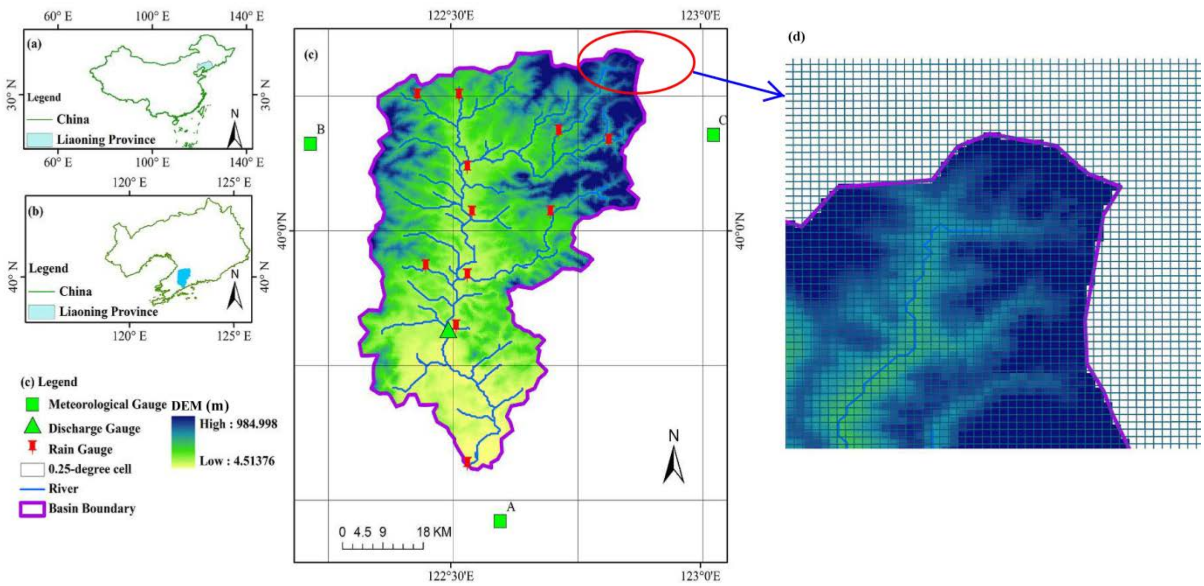


Fig. 1 Biliu basin: (a) the location of Liaoning province within China; (b) the location of Biliu basin within Liaoning province; (c) the distributions of rain gauges, discharge gauge, automatic weather stations, digital elevation model, and diagrammatic 0.25-degree precipitation cells; and (d) diagrammatic description of downscaling the 0.25-degree precipitation cells to 300 m × 300 m cells, and retrieving the 300 m × 300 m cells located within the basin boundary.

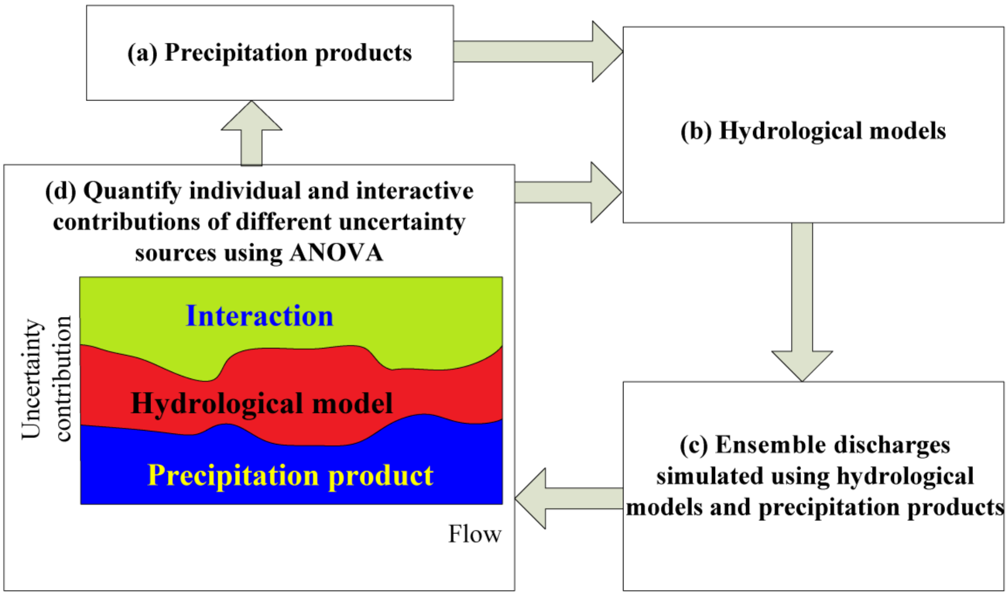


Fig. 2 Diagrammatic flowchart of the proposed framework for quantification of uncertainty contributions to ensemble discharges simulated using precipitation products on the basis of the analysis of variance (ANOVA) approach.

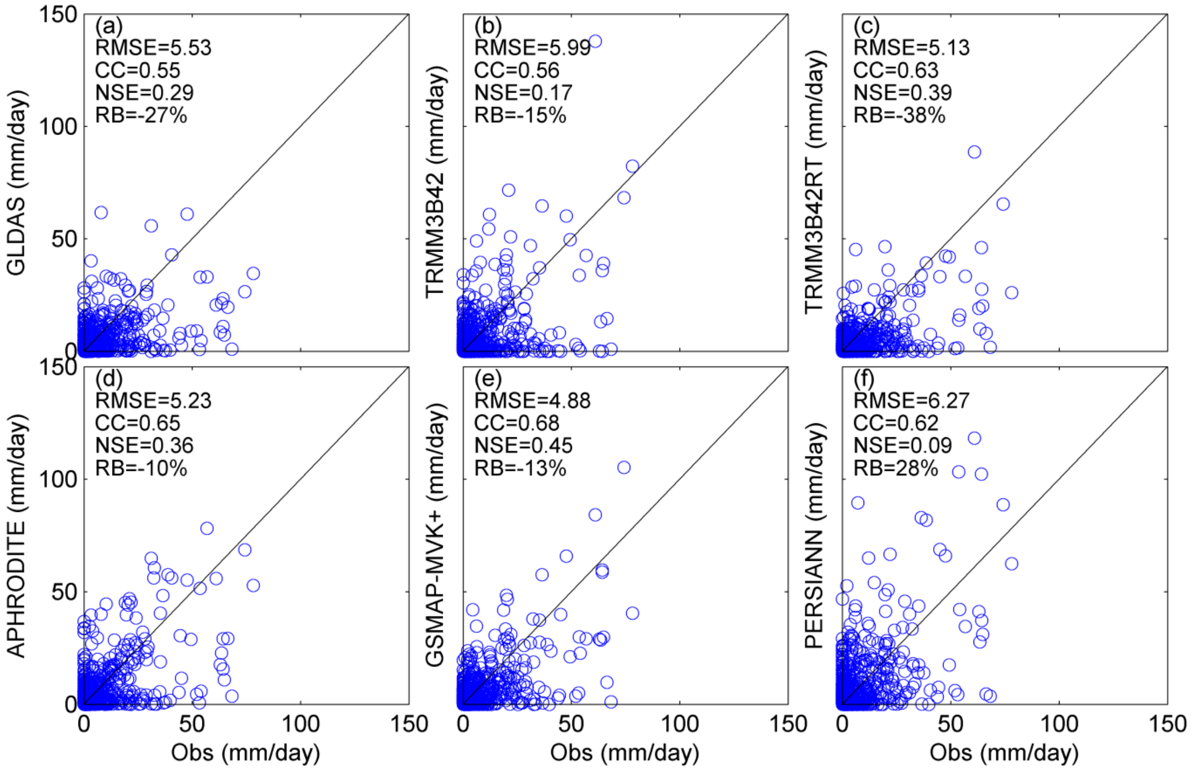


Fig. 3 Scatterplots of basin-averaged precipitation products versus gauge observations at a daily scale.

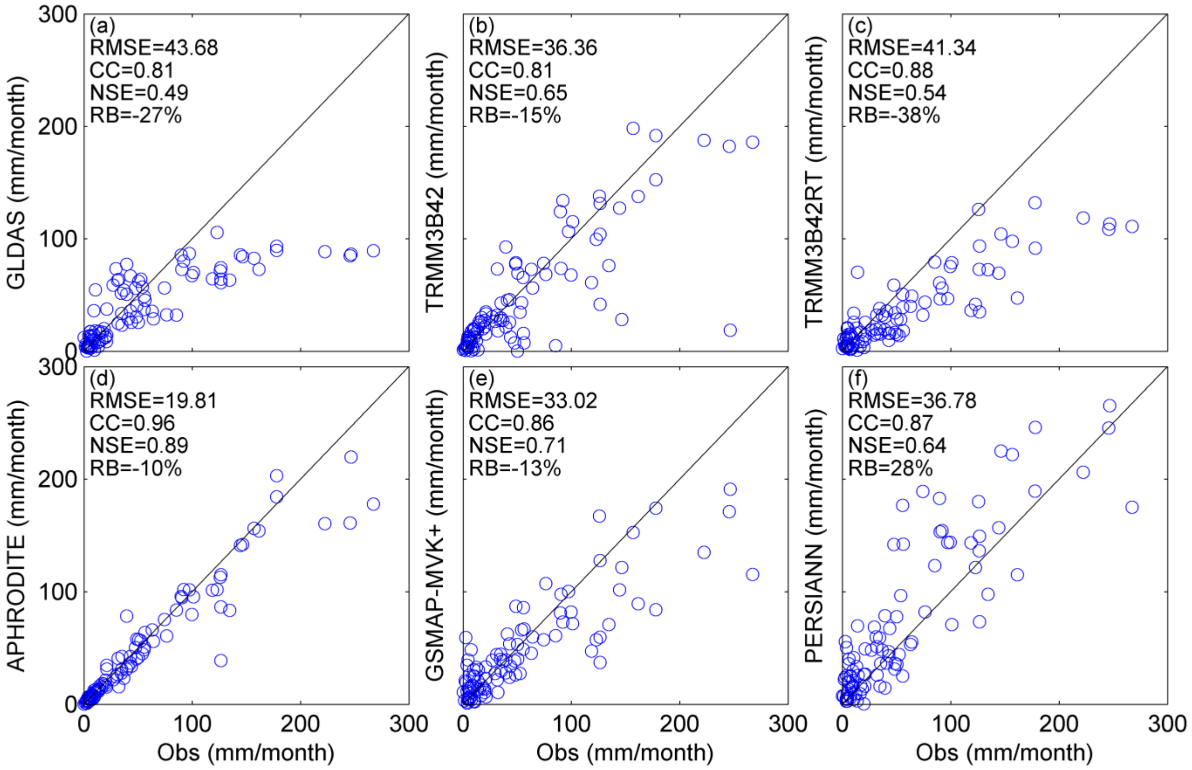


Fig. 4 Scatterplots of basin-averaged precipitation products versus gauge observations at a monthly scale.

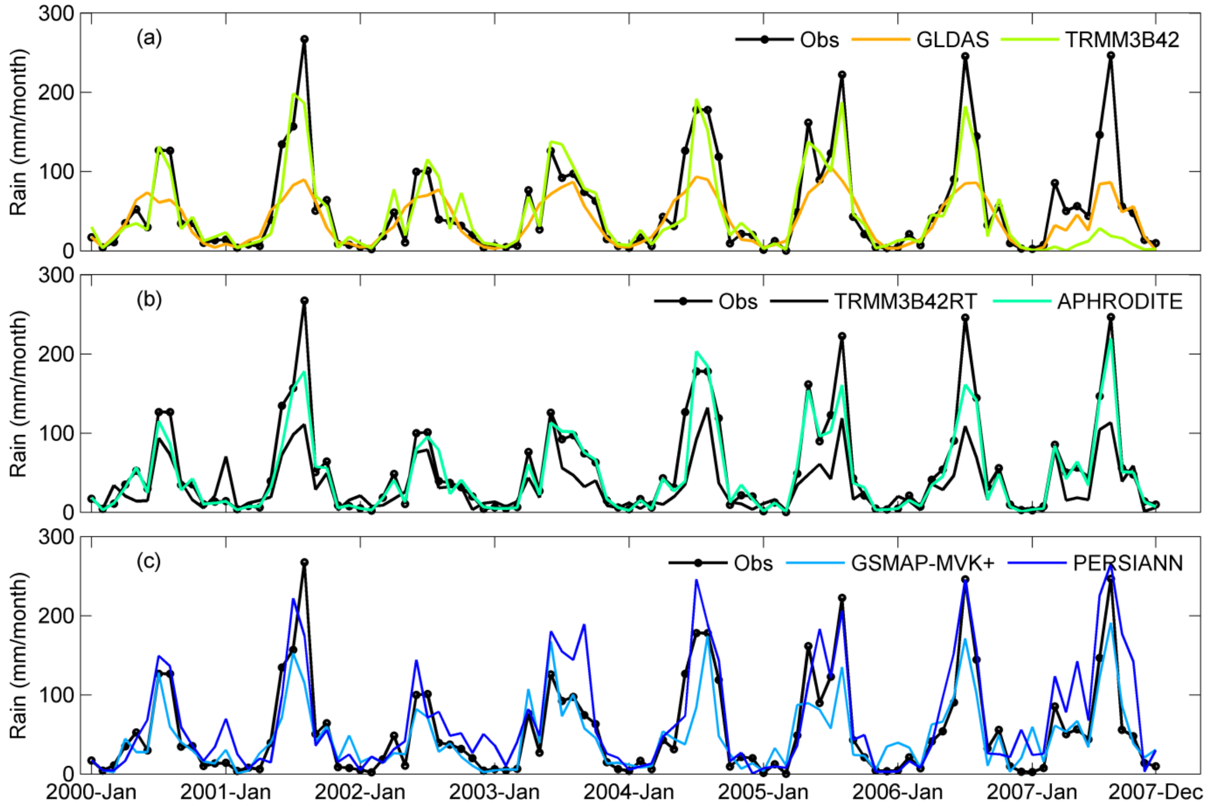


Fig. 5 Time series plots of basin-averaged precipitation product values versus gauge observations at monthly scale.

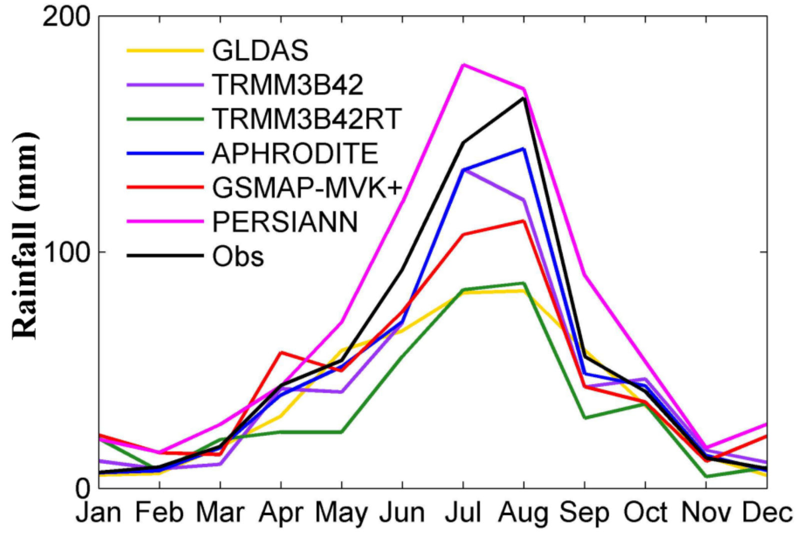


Fig. 6 Inter-annual ~~average~~basin-averaged monthly precipitation.

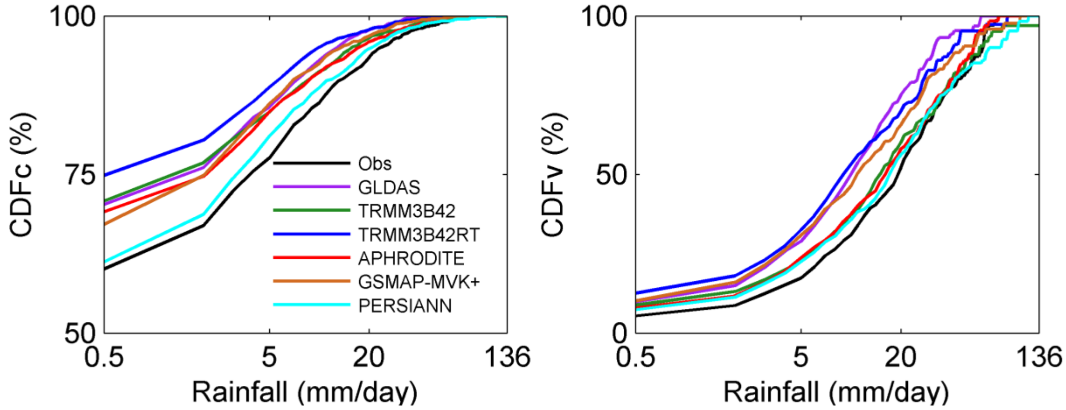


Fig. 7 Probability distributions of the six precipitation products by occurrence (CDFc) and volume (CDFv).

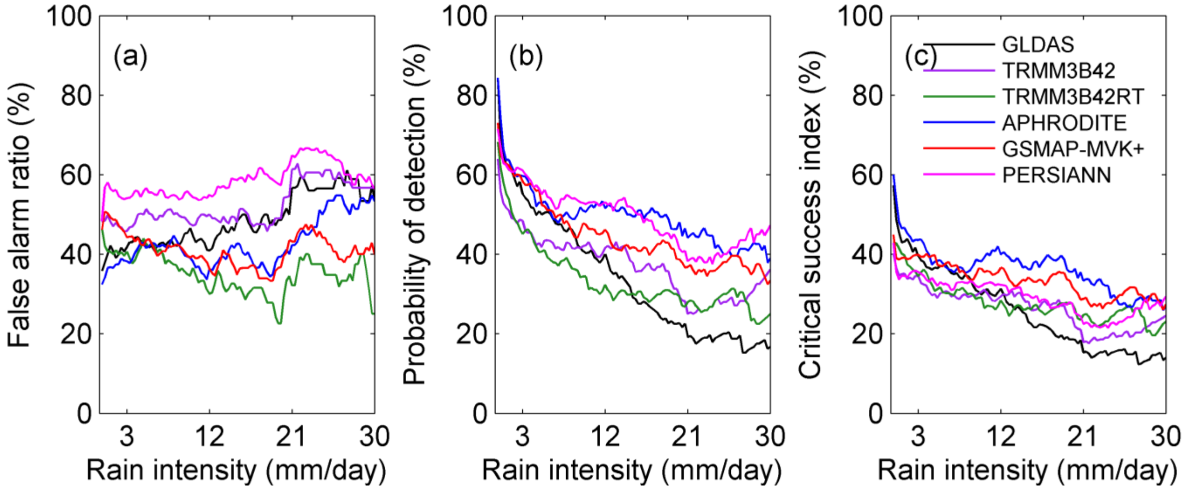


Fig. 8 False alarm ratio, probability of detection~~, false alarm ratio~~ and critical success index for the six precipitation products.

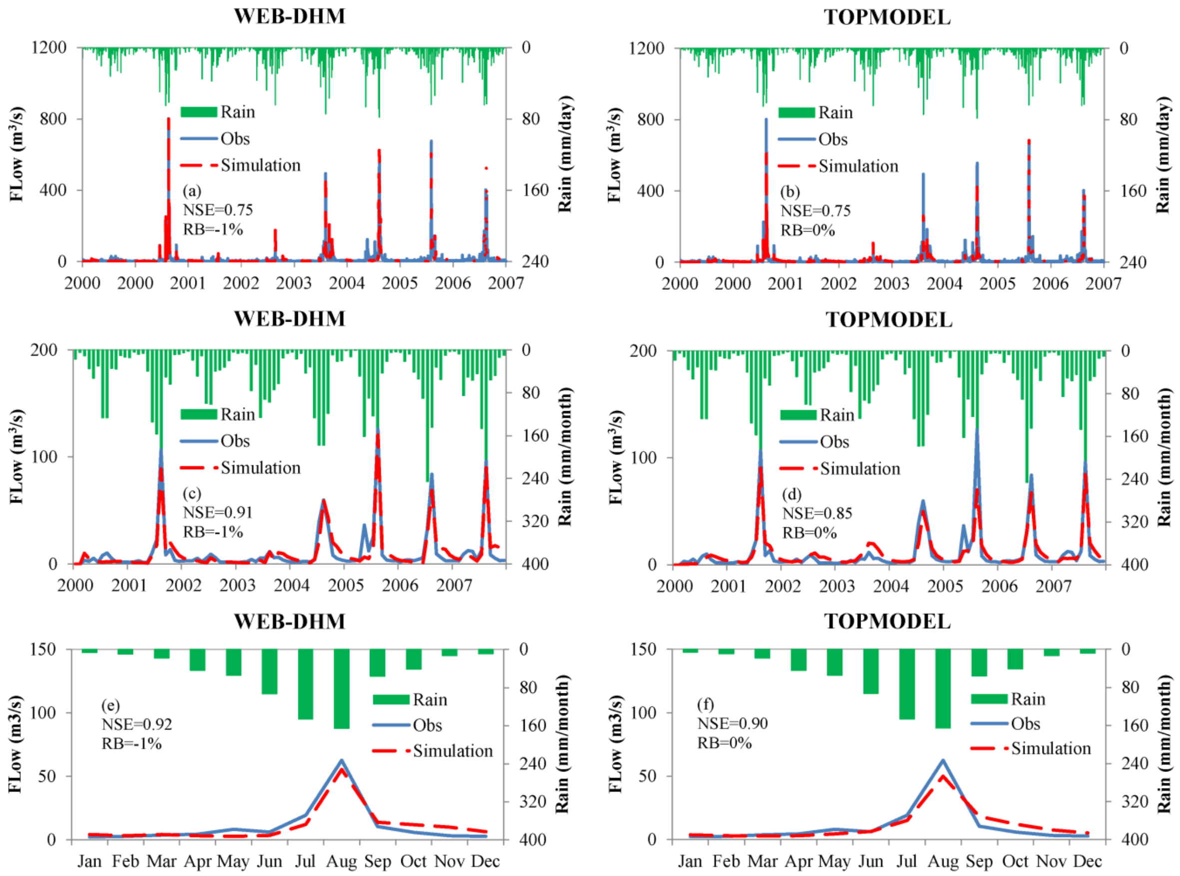


Fig. 9 Observed and simulated flows using WEB-DHM and TOPMODEL from 2000 to 2007: (a), (c) and (e) are daily, monthly and inter-annual simulations using WEB-DHM respectively; (b), (d) and (f) are daily, monthly and inter-annual simulations using TOPMODEL respectively.

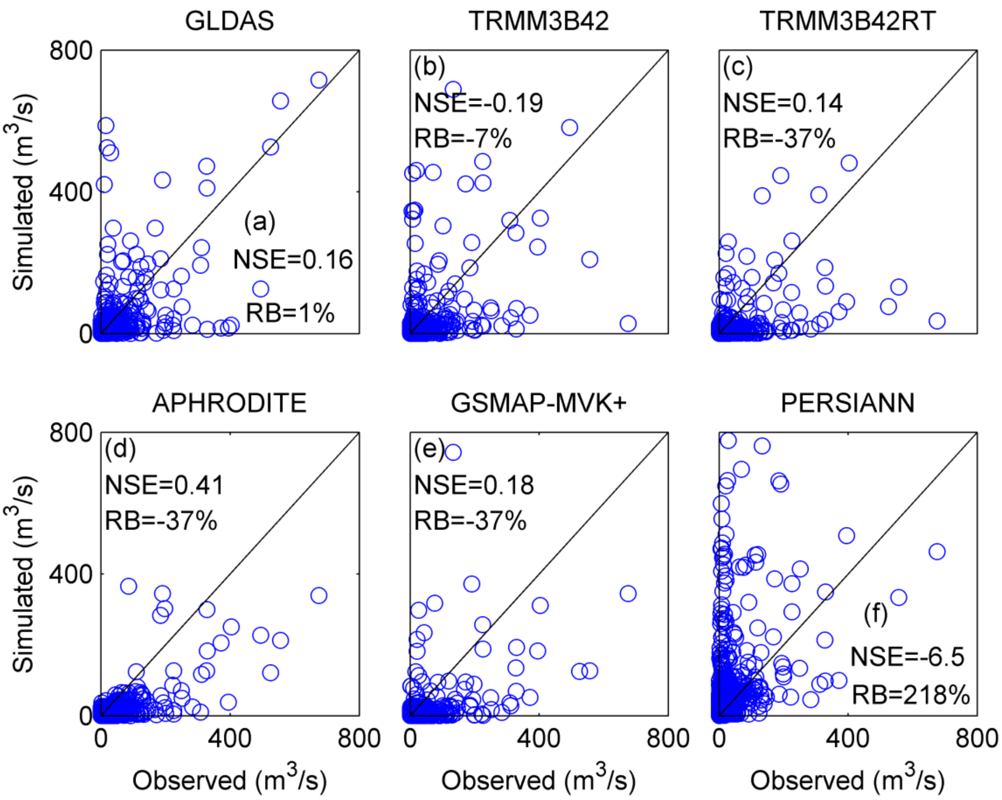


Fig. 10 Scatterplots of simulated discharges with WEB-DHM against gauge observations at a daily scale.

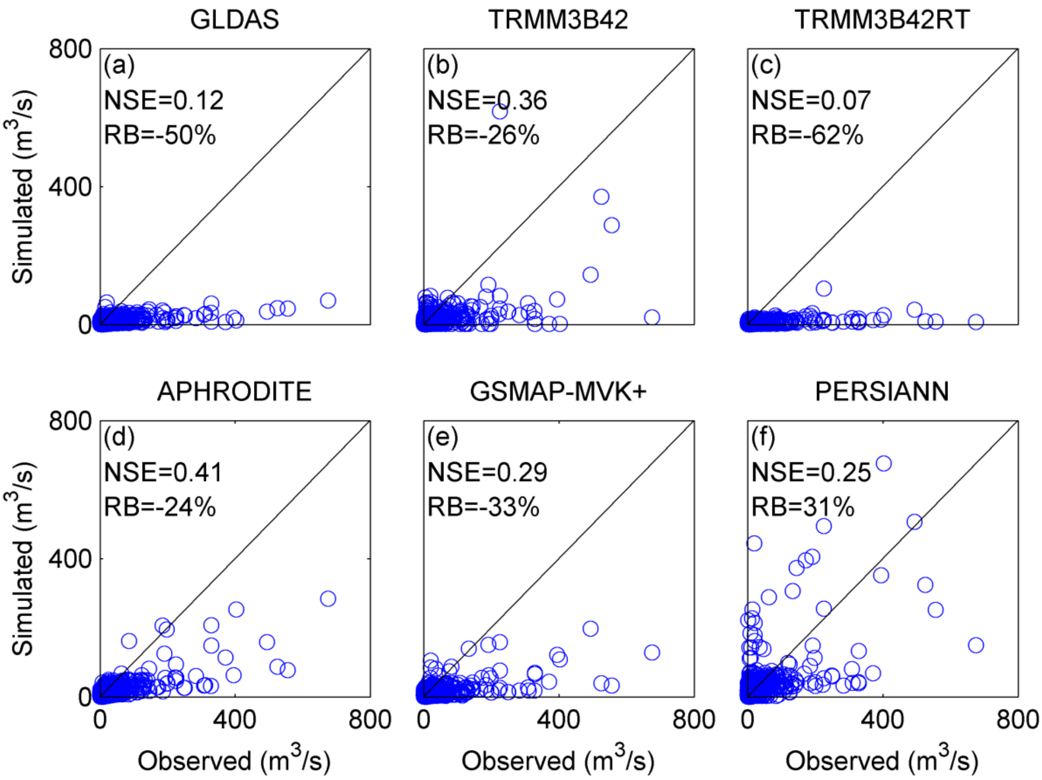


Fig. 11 Scatterplots of simulated discharges with TOPMODEL against gauge observations at a daily scale.

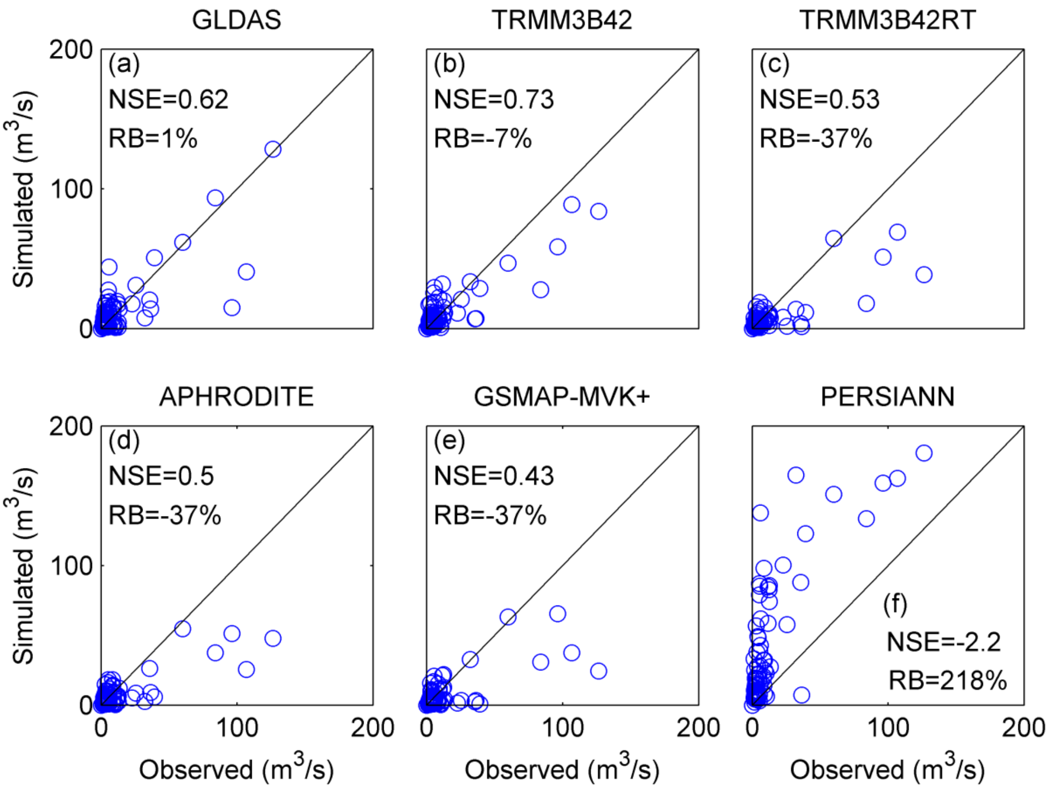


Fig. 12 Scatterplots of simulated flows with WEB-DHM against gauge observations at a monthly scale.

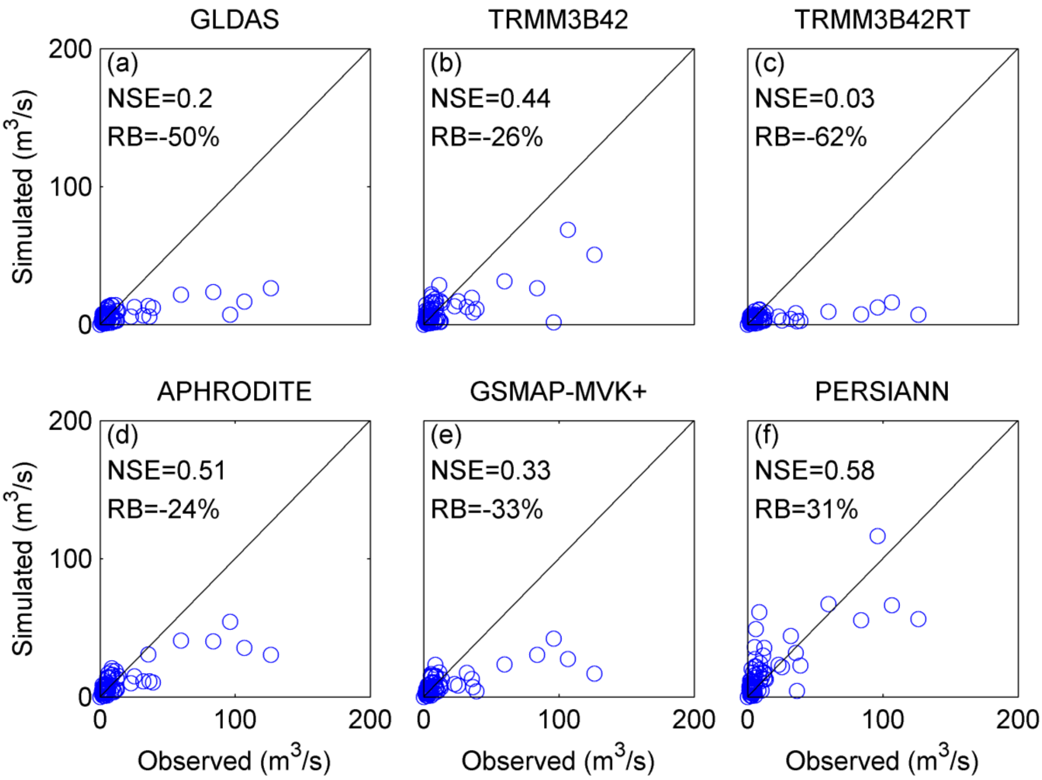


Fig. 13 Scatterplots of simulated discharges with TOPMODEL against gauge observations at a monthly scale.

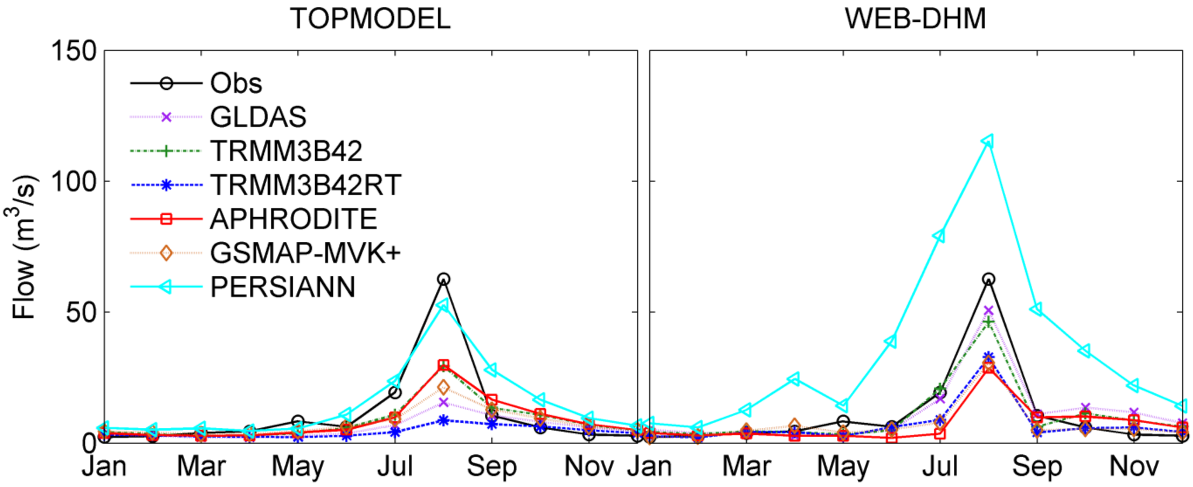


Fig. 14 Inter-annual average monthly discharges.

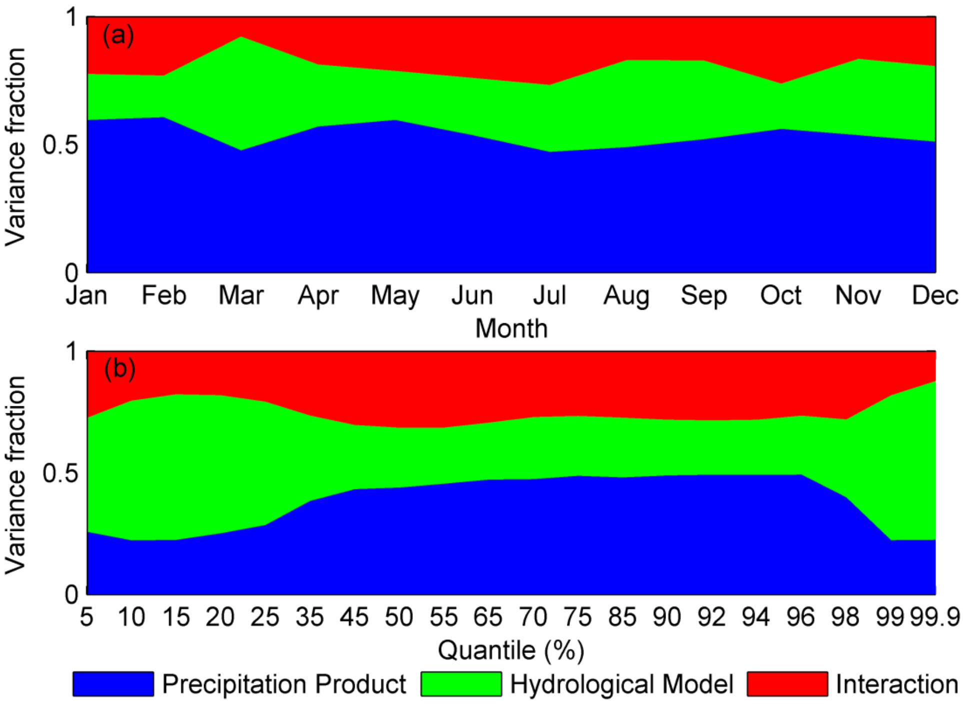


Fig. 15 Contributions of uncertainty sources to (a) average monthly discharges and (b) discharge quantiles based on daily scale simulated results.

We would like to thank the reviewer for the comments and suggestions, which have helped us to improve this paper significantly. Our detailed responses are provided point by point below.

Major comments:

The authors only compared the basin-averaged precipitation, and then the modeled discharges (with TOPMODEL and WEB-DHM) in the basin. Since the spatial patterns of precipitation may also affect the modeled discharges (with semidistributed/distributed hydrological models), I suggest you also investigate how the spatial distributions of different precipitation data (after downscaling to 300 m grids) have affected the modeled discharges.

Response:

Thanks for the suggestion. Discussion has been added in section 5 in the revised paper as below.

‘The spatial distribution of different precipitation data is not considered in this study. The study region is a small river basin, as shown in Fig. 1, there are only 11 grids inside the basin boundary for the precipitation products with a spatial resolution of 0.25 degree. Within a grid of 0.25 degree, there are no differences in precipitation amount between the 300 m × 300 m grids used in hydrological models, and differences exist at the level of 0.25 degree grids only. Sapriza-Azuri et al. (2015) suggested that the spatial variability of precipitation has little influence on rapidly responding river discharges; this study is the case because the flow transport time from the most upper part of the basin to the downstream discharge gauge is 6 hours, which is shorter than the daily and monthly time steps of discharges investigated. Therefore, the spatial distributions of precipitation products with a resolution of 0.25 degree in the relatively small river basin have little influence on the simulated discharges. However, the assumption of uniform distribution can be regarded as another uncertainty source against spatial variability, and its influence can be assessed using the proposed uncertainty quantification framework. This will allow us to compare the relative contributions of the assumption to those from other sources such as hydrological models, which will be investigated using a much larger river basin in the future work.'

In the conclusions section, more explanations have been added:

'It should be noted that this finding should be further tested with more river basins, in particular large river basins accounting for spatial variability in precipitation products.'

Minor comments: Page 9338, line 5: change "usually-neglected area" to small river basin.

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper.

Page 9339, line 21: please confirm that if your reference of APHRODITE data is appropriate. You may add another reference by Dr. Yatagai.

Response:

Thanks for the suggestion. Change has been made accordingly.

Page 9347, line 14: FAO should be "Food and Agriculture Organization".

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper.

Page 9362, lines 9-17: I guess that the spatial distributions of different precipitation products may contribute to the uncertainty in discharge simulations. Therefore, it is better to compare the observed precipitation with each precipitation product in their spatial patterns within the basin.

Response:

Thanks for the suggestion.

As explained above, the spatial distributions of precipitation products with a big resolution in the relatively small case study river basin have little influence on the daily and monthly discharges simulated.

Page 9362, lines 18-24: it is dangerous to draw such a conclusion. You may re-write the conclusion after checking the accuracy of different precipitation products in their spatial distributions, through comparing to gauge observations.

Response:

Thanks for the suggestion.

As explained above, the spatial distributions of precipitation products with very big grids in such a small river basin have little influence on the simulated discharges. Discussion has been added in section 5 in the revised paper, and this conclusion has been re-written as below:

'Fifth, discharge simulation depends on a good coalition of a hydrological model and a precipitation product, and a better precipitation product does not necessarily guarantee a better discharge simulation. This suggests that, although the satellite-based precipitation products are not as accurate as the gauge-based product, they could have better performance in discharge simulations when appropriately combined with hydrological models. It should be noted that this finding should be further tested with more river basins, in particular large river basins accounting for spatial variability in precipitation products.'

Figure 1: the unit of "m" should be given for DEM legend.

Response:

Thanks for the suggestion. Change has been made accordingly.

Figure 4: Please explain ANOVA in the figure caption and also in the body text.

Response:

Thanks for the suggestion.

The caption has been changed to ‘Fig. 2 Diagrammatic flowchart of the proposed framework for quantification of uncertainty contributions to ensemble discharges simulated using precipitation products on the basis of the analysis of variance (ANOVA) approach.’

In the body text, the following sentence has also been added in the first paragraph in section 2.5.

‘(d) quantification of individual and interactive contributions using the analysis of variance (ANOVA) approach including contributions from precipitation products, hydrological models and interactions between models and products.’

Figures 6-9: it is better to mention that the values in comparisons are basin-averaged ones.

Response:

Thanks for the suggestion. Changes have been made accordingly.

Figure 11: The caption should be "False alarm ratio, probability of detection, and critical success index for the six precipitation products."

Response:

Thanks for the suggestion. Changes have been made accordingly.

Figure 17: please indicate (a) and (b) in your figure caption.

Response:

Thanks for the suggestion. Changes have been made accordingly.

Reference:

Sapriza-Azuri, G., Jódar, J., Navarro, V., Slooten, L. J., Carrera, J., and Gupta, H. V.: Impacts of rainfall spatial variability on hydrogeological response, Water Resources Research, 51, 1300-1314, 10.1002/2014wr016168, 2015.

We would like to thank the reviewer for the comments and suggestions, which have helped us to improve this paper significantly. Our detailed responses are provided point by point below.

Specific comments:

1) Page 3, Line 11: the full name of APHRODITE is “Asian PrecipitationâĂŤHighly ̆ Resolved Observational Data Integration Towards Evaluation of Water Resources” instead of “Ground rain gauge-based interpolation products”. It is more appropriate to cite Yatagai et al. 2012 on APHRODITE.

Yatagai, A., K. Kamiguchi, O. Arakawa, A. Hamada, N. Yasutomi and A. Kitoh (2012): APHRODITE: Constructing a Long-term Daily Gridded Precipitation Dataset for Asia based on a Dense Network of Rain Gauges, Bulletin of American Meteorological Society , doi:10.1175/BAMS-D-11-00122.1

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper.

2) Page 6, Line 8: It is better to put “the average annual temperature is 10.6 C” right after “This basin is characterized by a temperate monsoon marine climate”

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper.

3) Page 21, Line 6: Simulation is from 2000-2007, start date should be specified, since the first few month of 2000 is not available for some of precipitation products (for instance, TRMM3B42RT is available after 1 Mar 2000).

Response:

Thanks for the suggestion. Changes have been made in the first paragraph in section 4.2 in the revised paper as below.

‘It should been noted that the start dates are different for precipitation products, and observed data were used when product data are not available: from 1 January 2000 to 29 February 2000 for TRMM3B42RT, GSMAP-MVK+ and PERSIANN; from 1 January 2000 to 23 February 2000 for GLDAS/Noah. These time periods were not considered for accuracy comparison.’

4) Page 15, Line 11: observations should be “gauge observations”.

Response:

Thanks for the suggestion. Change has been made accordingly in the revised paper.

5) Page 15, Line 16: “The number is the same, and therefore we used basin average rainfall amount in our evaluations.” There is no causal relationship between “The number is the same” and “we used basin average rainfall amount in our evaluations”. In addition, authors should describe how to get the basin-averaged rainfall for gauge observations and the gridded products.

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper.

The sentence has been deleted.

The basin-averaged rainfall calculation for precipitation product is introduced in the first paragraph in section 2.5 as below.

‘the following procedures were carried out for basin averaged rainfall calculations: (1) resampling 0.25o or 0.1o precipitation product grids into 300 m × 300 m cells (the grid size used in WEB-DHM simulations); (2) calculating basin-averaged precipitation using 300 m precipitation product grids located within the basin boundary.’

The basin-averaged rainfall calculation for gauge observations is introduced in the second paragraph in section 2.4.1 as below.

'Gauge rainfall data are also interpolated to 300 m × 300 m model cells and basin-averaged gauge rainfall data are calculated on the basis of interpolation results.'

6) Page 16, Line 9: "This improvement may be attributed to the assimilation with precipitation radar, gauge data and histogram matching." The Buliu basin (39.54N-40.35N) is beyond the of TRMM TMI/PR coverage (38S-38N). The authors should investigate other causes for the improvement.

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper as below.

'This improvement may be attributed to the assimilation with gauge data and histogram matching.'

7) Page 35, Table 1, start data for APHRODITE should be 1 Jan 1961.

Response:

Thanks for the suggestion. Change has been made accordingly in the revised paper.

8) Page 41, Figure 3: the caption is too brief. It would be better to name each panel with a) – f), and describe what each panel represent.

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper as below.

'Fig. 9 Observed and simulated flows using WEB-DHM and TOPMODEL from 2000 to

2007: (a), (c) and (e) are daily, monthly and inter-annual simulations using WEB-DHM respectively; (b), (d) and (f) are daily, monthly and inter-annual simulations using TOPMODEL respectively.'

9) Page 46, Figure 8: "Same as Figure 7" in caption is not proper words, since Figure 8 is very different from Figure 7.

Response:

Thanks for the suggestion. Change has been made accordingly in the revised paper as below.

'Fig. 5 Time series plots of basin-averaged precipitation product values versus gauge observations at monthly scale.'

10) Page 47, Figure 9: y axis should be rain (or rainfall) in mm.

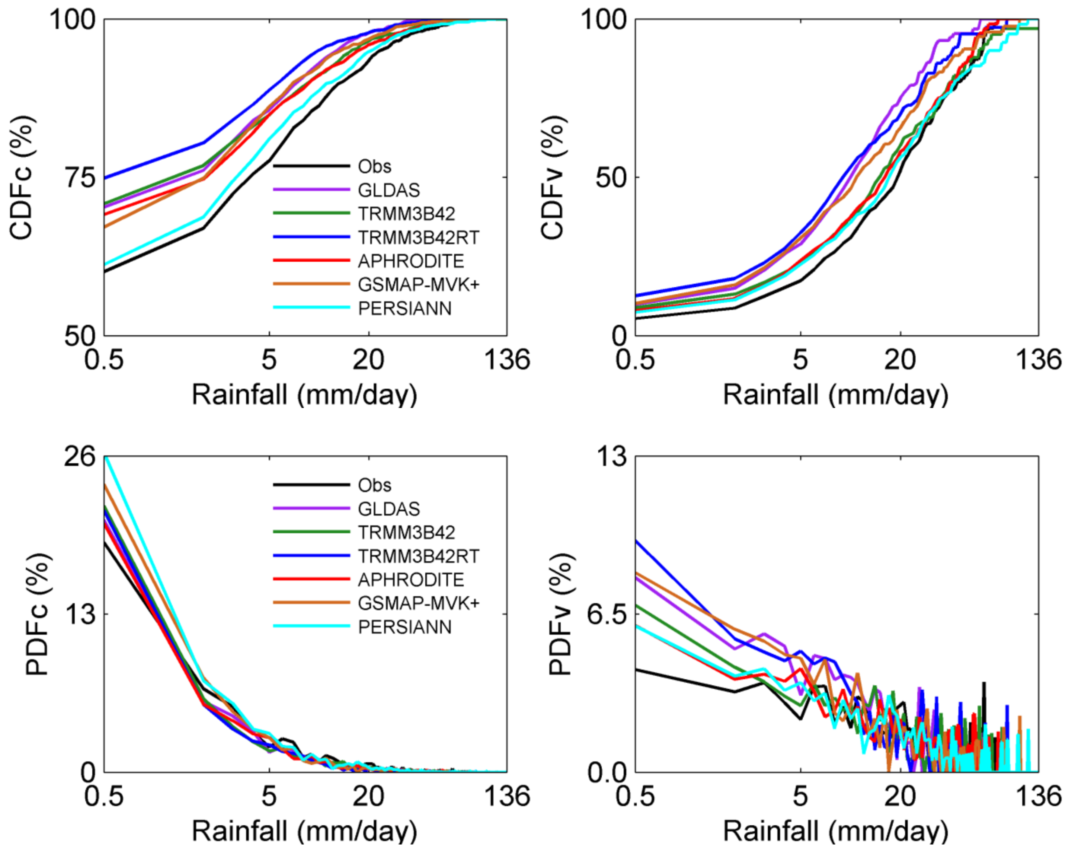
Response:

Thanks for the suggestion. Change has been made accordingly in the revised paper.

11) Page 48, Figure 10: it could be better to present PDF instead of CDF, and to include rainfall intensity less than 1mm/day.

Response:

Thanks for the suggestion. Changes have been made accordingly to include rainfall intensity less than 1mm/day. However, as shown below, the CDF figures are more clear than PDF figures. Thus, we still use the CDF figures in the revised paper.



12) Page 49, Figure 11: it could be more readable if the values on x axis are integers.

Response:

Thanks for the suggestion. Changes have been made accordingly.

13) Page 50, Figure 12: the texts overlap with circles.

Response:

Thanks for the suggestion. Changes have been made accordingly.

We would like to thank the reviewer for the comments and suggestions, which have helped us to improve this paper significantly. Our detailed responses are provided point by point below.

Major issues

1. The authors set up five experiments for the inter-comparison of the precipitation products, each experiment compared two products. However, in anywhere of the manuscript, the accuracy scores and probability distribution of all the six products were calculated and showed together, then, what's your purpose for setting up these independent experiments? I appreciate the amount of work from the authors on data processing analysis and efforts being made for an English journal paper, but the clarity of the manuscript should be further improved.

Response:

Thanks for the suggestion. Because Fig. 2 and related experiments set up are not relevant to this study, Fig. 2 and related experiments has been removed in the revised paper as the request from the reviewer 4.

The overall aim of this paper is to develop a framework to quantify the contributions of uncertainties from precipitation products, hydrological models and their interactions to uncertainty in simulated discharges. This is to answer one question asked by modelers: which product and model I should choose and which is the main uncertainty source as there are so many products and models? To achieve the aim, the first step is to understand the performance of the 6 products selected when applied to the two hydrological models. That is why we assessed the 6 products in terms of accuracy scores and probability distributions. Building on this, the second step is to quantify the respective uncertainties from products and models, and the combined uncertainties from the interactions between products and models.

As suggested by the reviewer, we have revised the second paragraph from the back in the introduction section and added more explanations to clarify the purpose of the experiments.

‘The overall aim of this paper is to develop a framework to quantify the contributions of uncertainties from precipitation products, hydrological models and their interactions to uncertainty in simulated discharges. To achieve the aim, the first step is to understand the performance of the selected precipitation products including TRMM3B42, TRMM3B42RT, GLDAS/Noah (GLDAS_Noah025SUBP_3H), APHRODITE, PERSIANN and GSMAP-MVK+, when applied to the chosen hydrological models. Two hydrological models of different complexities - a water and energy budget-based distributed hydrological model (WEB-DHM) (Wang et al., 2009a;Wang et al., 2009b;Wang et al., 2009c) and a physically-based semi-distributed hydrological model TOPMODEL (Beven and Kirkby, 1979) - were employed to investigate the influence of hydrological models on discharge simulations. Building on the assessment of precipitation products, the second step is to quantify the respective uncertainties from precipitation products and hydrological models, and the combined uncertainties from the interactions between products and models. This is achieved using a global sensitivity analysis approach, i.e., the analysis of variance approach (ANOVA). A river basin in northern China with a series of 8-year data is used to demonstrate the methodology.’

2. For the hydrology models WEB-DHM and TOPMODEL, precipitation data is not the only input. Other input data include temperature, downward solar radiation, long wave radiation, air pressure, wind speed and humidity. These input data can also bring nonnegligible uncertainty to the discharge simulating, especially after heavily processing to meet the model

demand. It’s unreasonable for the authors didn’t consider the uncertainty from other input in uncertainty quantification.

Response:

Thanks for the suggestion. Fig. 9 in the revised paper shows that the simulated discharge data are acceptable particularly at monthly and inter-annual scales using the temperature, downward solar radiation, long wave radiation, air pressure, wind speed and humidity. Research has shown that the land surface temperatures are highly accurate compared with MODIS satellite land surface temperature observations (Wang et al., 2011;Qi et al., 2015). Thus, the uncertainties from other inputs are not considered in our case study river basin.

Having said this, a paragraph in section 5 in the revised paper has been added as below.

‘It should be noted that other input data including temperature, downward solar radiation, long wave radiation, air pressure, wind speed and humidity may also have uncertainties. However, Fig. 9 shows that the simulated discharge data are acceptable particularly at monthly and inter-annual scales using these data. Research has shown that the land surface temperatures are highly accurate compared with MODIS satellite land surface temperature observations (Wang et al., 2011;Qi et al., 2015). Thus, the uncertainties from the other inputs are not considered in our case study river basin.’

Minor issues

1. There are four precipitation products whose start dates were late than 1 Jan 2000, what’s exact date the simulation and evaluation start from?

Response:

Thanks for the suggestion. Changes have been made in the first paragraph in section 4.2 in the revised paper as below.

‘It should been noted that the start dates are different for precipitation products, and observed data were used when product data are not available: from 1 January 2000 to 29 February 2000 for TRMM3B42RT, GSMAP-MVK+ and PERSIANN; from 1 January 2000 to 23 February 2000 for GLDAS/Noah. These time periods were not considered for accuracy comparison.’

2. Figure 5 is really not that informative, and this figure should probably be removed from the manuscript.

Response:

Thanks for the suggestion. The figure has been removed in the revised paper.

3. Sections 3 and 4 has repetitions. Some of it can be edited out for brevity, i:e: it’s useless to say ”Observations are shown on the x axis and precipitation product data are shown on the y axis”, but you mentioned it twice in section 3.1.

Response:

Thanks for the suggestion. Changes have been made accordingly.

4. Page 9359, lines 19-20: Don’t use “significant” if you didn’t do the significance test.

Response:

Thanks for the suggestion. The words ‘significant’ and ‘significantly’ have been replaced in the entire paper where possible.

References:

Beven, K. J., and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology, Hydrological Sciences Bulletin, 24, 43-69, 10.1080/02626667909491834,

1979.

Qi, W., Zhang, C., Fu, G., and Zhou, H.: Global Land Data Assimilation System data assessment using a distributed biosphere hydrological model, Journal of Hydrology, 528, 652-667, 10.1016/j.jhydrol.2015.07.011, 2015.

Wang, F., Wang, L., Koike, T., Zhou, H., Yang, K., Wang, A., and Li, W.: Evaluation and application of a fine-resolution global data set in a semiarid mesoscale river basin with a distributed biosphere hydrological model, Journal of Geophysical Research, 116, 10.1029/2011jd015990, 2011.

Wang, L., Koike, T., Yang, D. W., and Yang, K.: Improving the hydrology of the Simple Biosphere Model 2 and its evaluation within the framework of a distributed hydrological model, Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 54, 989-1006, 10.1623/hysj.54.6.989, 2009a.

Wang, L., Koike, T., Yang, K., Jackson, T. J., Bindlish, R., and Yang, D.: Development of a distributed biosphere hydrological model and its evaluation with the Southern Great Plains Experiments (SGP97 and SGP99), Journal of Geophysical Research, 114, 10.1029/2008jd010800, 2009b.

Wang, L., Koike, T., Yang, K., and Yeh, P. J.-F.: Assessment of a distributed biosphere hydrological model against streamflow and MODIS land surface temperature in the upper Tone River Basin, Journal of Hydrology, 377, 21-34, 10.1016/j.jhydrol.2009.08.005, 2009c.

We would like to thank the reviewer for the comments and suggestions, which have helped us to improve this paper significantly. Our detailed responses are provided point by point below.

Major Comments

1 Compensation of errors

Consistently in the manuscript the authors make an argument summarized by this statement in the abstract: “It is also found that a good discharge simulation depends on a good coalition of a hydrological model and a precipitation product, suggesting that, although the satellite-based precipitation products are not as accurate as the gauge based product, they could have better performance in discharge simulations when appropriately combined with hydrological models”. The authors call this a “comprehensive result” on P9357, L26-28. See also: the 5th point of the Conclusion (P9362).

I have challenges with this point, which is presented as a buzz point of the manuscript. It would appear to me that the models could be compensating for the errors/uncertainty in the input precipitation products. Generally, this compensation manifests in some instances in the form of questionable/unrealistic model parameter values (see Nikolopoulos et al.,(2013)). Looking at Table 1 and 2 one cannot help but notice some potentially odd values that might not compare with general range of values found in most publications. Whilst the authors do not explain the geology and soil type of the study region, except that the study area is a mountainous region, parameter values like: lnTo(why is negative?), RV, saturated hydraulic conductivity and the hydraulic conductivity anisotropy ratio need to be evaluated and justified.

A discussion on where or how the precipitation error/uncertainty was compensated should have been presented!

Unfortunately, this makes the key conclusion or “comprehensive result” of this work stand on shaky ground, if the input error is being compensated in parameter value estimation.

Response:

We agree with the reviewer that the compensation should be better explained in the text. First we would like to explain the ranges of the model parameters. LnTo represents the log values of saturated hydraulic conductivity: thus it can be negative. TOPMODEL is a semi-distributed hydrological model. It uses basin-averaged parameter values, and these parameter values are estimated by experience or observation. Therefore, the parameter values are generally considered as uncertain and different ranges were provided in the previous studies (Beven and Kirkby, 1979;Beven and Freer, 2001a, b;Peters et al., 2003). For example, the parameter values used by Beven and Freer (2001a) are as below.

Name (units)	Lower bound	Upper bound
SZM (m)	0.005	0.06
LNT0 (m^2 h^{-1})	0.1	8
RV (m^2 h^{-1})	1000	5000
SR_{max} (m)	0.005	0.3
SR_0 (m)	0	0.3
TD (m h^{-1})	0.1	120

The parameter values used by Peters et al. (2003) are as below.

Name (units)	Lower bound	Upper bound
SZM (m)	0.01	0.08
LNT0 (m^2 h^{-1})	-7	1
RV (m^2 h^{-1})	1000	5000
SR_{max} (m)	0.015	0.1
SR_0 (m)	0	0.05
TD (m h^{-1})	0.1	120

It can be seen that the parameter values of TOPMODEL in our paper are comparable with the values listed in the above tables. Therefore, we believe the parameter values used in our study are acceptable.

The parameter values of WEB-DHM model have been evaluated on the basis of basin scale water and energy cycles, for example, the accuracy of simulated based-averaged land surface temperature, discharge data and the spatial distributions of land surface temperature (Qi et al., 2015). This validation can be seen in the research by Qi et al. (2015). In addition, in our research, we showed the validation results using discharge data, and the discharge evaluation results are acceptable. Thus, the parameter values of WEB-DHM are acceptable.

In our study, we calibrated the hydrological model parameters using gauge observed precipitation data, and the same set of parameter values are applied to different products. Thus, the parameter values are not tuned to a specific product. That is, there is little compensation of model parameters for the errors in input precipitation data. The 5th point of the conclusions was drawn by comparing the simulation accuracies of WEB-DHM and TOPMDOEL, and the differences in model accuracy mainly results from the different representations of hydrological processes. That is, the errors in precipitation products are primarily compensated by the different representations of model processes.

As explained above, the paragraph in lines 26-28 on P9357 (it is the third paragraph in section 4.3 in the revised paper) has been rephrased as below.

‘The combination of WEB-DHM and TRMM3B42 shows a great performance, with NSE and RB values of up to 0.73 and -7%, even though TRMM3B42 is not the best in monthly scale precipitation data evaluation. This reveals the influence of different characterizations of hydrological processes on the selection of precipitation data, implying that accurate discharge simulation does not solely depend on the accuracy of a precipitation product.’

A paragraph has been added in the third paragraph in section 4.1 in the revised paper as below.

‘Note that the parameters of TOPMODEL and WEB-DHM were calibrated using observed precipitation data, and the accuracy of simulated discharges has been validated using gauge observations. Comparison with the parameter values reported in previous research shows the parameter values are appropriate (Beven and Freer, 2001a;Peters et al., 2003;Qi et al., 2015).’

The fourth paragraph has been added in the section 5 in the revised paper as below.

‘In this study, the parameter values calibrated using gauge observations are not tuned to a specific product. That is, there is little compensation of model parameters for the errors in input precipitation data. The differences in model accuracy mainly results from the different representations of hydrological processes. That is, the errors in precipitation products are primarily compensated by the different representations of model processes.’

The grid slopes, land-use types and soil data have been introduced at the end of the second paragraph in the section 2.4.1 as below.

‘The grid slopes vary from 0 to 38 degrees. Land-use data are obtained from the USGS (http://edc2.usgs.gov/glcc/glcc.php). The land-use types have been reclassified to SiB2 land-use types for this study (Sellers et al., 1996). There are six land-use types, with broadleaf and needle leaf trees and short vegetation being the main types. Soil data are obtained from the Food and Agriculture Organization (FAO) (2003) Global data product, and there are two types of soil in the basin: clay loam-luvisols and loam-phaeozems.’

2 Precipitation performance argumentation

The authors in presenting and discussing their evaluation of the performance of precipitation product take an algorithmic inclined line. Basically, the authors were arguing that, say, if product X outperforms Y, the authors then say to improve Y the algorithm behind X should be taken-up/considered. This line of discussion or argumentation runs through their whole presentation in Section 3. As example on Page 9353, L20, the authors say "Thus, if TRMM3B42 wants to improve heavy precipitation estimation, the artificial neural network function [of PERSIANN] and APHRODITE products could be helpful".

Although in the particular reference cited above, the ANN is recommended when it had overestimated heavy precipitation (why?), my challenge is with overall approach of argumentation. It is not only the main statistical or mathematical equation behind a product that determines its success. It could be influenced by other factors such as the inputs used in the merged products or any internal calibration or merging procedures etc. Without addressing such confounding factors how do the authors draw up their discussions and conclusions along these lines? I did not see sufficient evidence in the manuscript to warrant such a discussion and conclusion to the extent of singling out the statistical or mathematical equation. The cause-effect framework is too simplistic and limited!

I feel the authors could have gotten more purchase by discussing their results with the perspective of use/application of the products! Section 3 was generally not well presented!

Response:

We agree with the reviewer and all relevant statements have been revised or deleted accordingly in the revised paper. The abstract and the conclusions section have been rephrased. The second point of the conclusions has been revised as below.

'GSMAP-MVK+ show huge advantage, and is better than TRMM3B42 in RB, NSE, RMSE, CC, false alarm ratio and critical success index, while PERSIANN is better than TRMM3B42 in probability of detection and precipitation probability distribution estimation. At present, the new precipitation estimation mission - NASA Global Precipitation Measurement (GPM) - combines the artificial neural network function of PERSIANN and precipitation radar-matching of TRMM Multi-satellite Precipitation Analysis. However, the above finding implies that incorporating GSMAP-MVK+ estimation approach into GPM could be useful as well.'

3 Other argumentation issues:

> Arguments without clear process-level backing: There are instances in the manuscript that the authors make arguments without adequate process-level backing. On P9353, L1-5, that the authors argue that the performance/accuracy of PERSIANN is related to latitude! From a process-level perspective this is difficult to comprehend. Is this process related or it's related to the sensors used? The authors should ensure that there is process-level sufficiency in their statements in the manuscript.

Response:

Thanks for the suggestion. The statements have been deleted accordingly in the revised paper.

> Arguments without backing: On P9352, L1-5: Without seasonal analysis, how do the authors pinpoint summer convective rainfall with the authority that they do in the text?

Response:

Thanks for the suggestion. Three references that study convective precipitation in northeast China have been added to back the statement. The three references are as below.

(Shou and Xu, 2007a, b;Yuan et al., 2010)

P9357, L3-5: How is the non-linear amplification derived and supported?

Response:

The ‘non-linear’ has been deleted and an example has been given after the statements in the revised paper as below.

‘For example, for GLDAS and PERSIANN, the RB criteria at the daily scale precipitation evaluations are -27% and 28%, but they are -50% and 31% in TOPMDOEL simulations; they are 1% and 218% in WEB-DHM simulations.’

> P9356, L20-23 “Big differences. . .using PERSIANN”, where can the reader see this?

Response:

Thanks for the suggestion. The statement has been deleted in the revised paper.

4 Flows – P9360

> L3-5: You should highlight to the reader what “small”, “middle” and “large” discharge means using a value or a quantile and its corresponding discharge value?

Response:

Thanks for the suggestion. The statements have been changed accordingly in the revised paper as below.

‘Fig. 15b shows that, for small discharges (smaller than 25% quantile which corresponds to an observed discharge value of 1.79m^3/s) and large discharges (larger than 99% quantile which corresponds to an observed discharge value of 157m^3/s), hydrological models contribute most of the uncertainties. For middle magnitude flows (between 25% and 99% quantiles), precipitation products contribute the majority, and the contribution of interactions is not negligible and of similar magnitude to the contribution from hydrological models.'

> L7-9: "This may. . .discharge magnitude". What is the meaning of this statement? Explain to the reader what "interaction effects" are – and maybe give an example?

Response:

Thanks for the suggestion. The statement has been deleted in the revised paper. The explanation has been given at the end of the same paragraph as below.

'The different contributions of interactions for various magnitude flows may be because different magnitude rainfall data could trigger different hydrological processes (Herman et al., 2013). Small discharges mainly come from base flows which are relatively stable and do not need much rainfall to be triggered, and large discharges are mainly controlled by overland flows when heavy precipitation occurs. Middle magnitude discharges consist of contributions from base flows, lateral subsurface flows and overland flows. It is more complex and can be triggered by various magnitude rainfalls - thus interactions are more changeable.'

> L 13-15: Whilst it is vague what the authors mean by middle flows numerically, I am not sure that the statement holds physically. I would think contributions from groundwater flows in such a simulation would be baseflows and these would be low flows – in your terminology, what you call "small" discharges?

Response:

Thanks for the suggestion.

Small magnitude flow represents the data that are smaller than 25% quantile which corresponds to an observed discharge value of 1.79m^3/s.

Large magnitude flow represents the data that are larger than 99% quantile which corresponds to an observed discharge value of 157m^3/s.

Middle magnitude flow represents the data that are between 25% and 99% quantiles.

The statements have been revised according as below.

‘Small discharges mainly come from base flows which are relatively stable and do not need much rainfall to be triggered, and large discharges are mainly controlled by overland flows when heavy precipitation occurs. Middle magnitude discharges consist of contributions from base flows, lateral subsurface flows and overland flows. It is more complex and can be triggered by various magnitude rainfalls - thus interactions are more changeable.’

> L22-24: How is this practically feasible that precipitation products provide such information?

Response:

Thanks for the suggestion. The statement has been deleted in the revised paper, and the second paragraph has been added in section 5 in the revised paper as below to explain how to get the combination.

‘In addition to improving the accuracy of precipitation products, a good collation could help to achieve the performance in discharge simulations. Our approach provides a way to assess the different coalitions, i.e., the overall uncertainties in simulated discharges from different combinations of hydrological models and precipitation products. More precipitation products and hydrological models should be included and tested in the future work.’

5 Use appropriate terminology and scientific English. For example the issue of “small”, “medium” and “large” discharges highlighted above. In section 3, avoid using the word “worst” in your comparisons.

Response:

Thanks for the suggestion. These issues have been revised accordingly.

6 The scheme in Figure 2 – Did you use this scheme eventually in the paper? Or it's something you forgot to edit out? In case you used it, on P9344, you will need to justify the experimental set ups you came up with e.g. why not GLDAS and APHRODITE?

Response:

Thanks for the suggestion. The figure has been deleted and corresponding statements have been deleted.

7 Why didn't you apply the same calibration method for both models?

Response:

Thanks for the suggestion. Because TOPMODEL is computationally efficient, but WEB-DHM is computationally expensive. In one single run, WEB-DHM needs 4 hours to simulate one year discharges but TOPMODEL needs 2 seconds only using the same computer.

8 CDFs - (P9354) L14-25: How are you defining "larger rainfall intensity"? Could it be due that there is usually limited "larger intensity" events/observations in those regions of any distribution such that it appears as good performance, but in actual fact it's just spurious good performance?

Response:

Thanks for the suggestion. The statement has been changed in the revised paper as below.

'All precipitation products overestimate occurrence and volume probabilities except rainfall intensities of larger than 63mm/day and 53mm/day for occurrence and volume probabilities, respectively. This may be because the precipitation products overestimate the intensity of some heavy rainfall (recall the results in section 3.1).'

9 P9358, L11-14: Is this really counter-intuitive? It is accepted that good modelling work depends on having good data and a good or appropriate model structure

Response:

Thanks for the suggestion. The statement has been deleted in the revised paper.

10 Without advocating for long and winding papers, generally it appears that the writing leaves a lot of the issues hanging or incompletely presented. As an example in section 1.0, as a reader one does not get the continuity of the text. Please revise the manuscript.

Response:

Thanks for the suggestion. The paper has been revised accordingly. We have revised the second paragraph from the back in the introduction section as below.

'The overall aim of this paper is to develop a framework to quantify the contributions of uncertainties from precipitation products, hydrological models and their interactions to uncertainty in simulated discharges. To achieve the aim, the first step is to understand the performance of the selected precipitation products including TRMM3B42, TRMM3B42RT, GLDAS/Noah (GLDAS_Noah025SUBP_3H), APHRODITE, PERSIANN and GSMAP-MVK+, when applied to the chosen hydrological models. Two hydrological models of different complexities - a water and energy budget-based distributed hydrological model (WEB-DHM) (Wang et al., 2009a;Wang et al., 2009b;Wang et al., 2009c) and a physically-based semi-distributed hydrological model TOPMODEL (Beven and Kirkby, 1979) - were employed to investigate the influence of hydrological models on discharge simulations. Building on the assessment of precipitation products, the second step is to quantify the respective uncertainties from precipitation products and hydrological models, and the combined uncertainties from the interactions between products and models. This is achieved using a global sensitivity analysis approach, i.e., the analysis of variance approach (ANOVA). A river basin in northern China with a series of 8-year data is used to demonstrate the methodology.'

Technical Corrections

P9338, L1-7: Sentence too long. Please revise!

Response:

Thanks for the suggestion. The statement has been changed in the revised paper as below.

'The applicability of six fine-resolution precipitation products, including precipitation radar, infrared, microwave and gauge-based products using different precipitation computation recipes, is evaluated using statistical and hydrological methods in northeastern China. In addition, a framework quantifying uncertainty contributions of precipitation products, hydrological models and their interactions to uncertainties in ensemble discharges is proposed.'

P9338, L5: "usually-neglected area" has no hydrologic value. Remove!

Response:

Thanks for the suggestion. The statement has been removed.

P9339, L5-7: "However. . .rural areas". Revise!

Response:

Thanks for the suggestion. The statement has been changed in the revised paper as below.

'However, precipitation data are not available in many regions, particularly mountainous districts and rural areas in developing countries.'

P9340, L21-23: Add reviews of uncertainty quantification/analysis work by Kuczera et al., (2006), Vrugt et al (2009b), Vrugt et al (2009a), Tolson and Shoemaker (2007).

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper.

P9341, L3-6: Revise! The repeated use of "and" makes the sentence difficult to read.

Response:

Thanks for the suggestion. Changes have been accordingly in the revised paper as below.

'In addition to individual contributions from hydrological models and precipitation data, the interactions between precipitation products and hydrological models also contribute to uncertainty in simulated discharges.'

P9341, L11: Remove "usually-neglected area" P9341, L14: Replace "include" with "are"

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper.

P9344, L17: Use the common-abbreviation NSE instead of NSCE throughout the manuscript.

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper.

P9345, L4-6: So?

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper.

P9346: If you don't have hourly data, why force it?

Response:

Thanks for the suggestion. Because in this study we want to investigate the influences of hydrological models on simulated discharges, and WEB-DHM has to use hourly data. Many studies have shown that using the downscaled hourly data is acceptable in discharge simulations (Wang et al., 2009c;Wang et al., 2011;Wang et al., 2012;Zhou et al., 2015). In addition, in this study we also showed the accuracy of simulated discharges is acceptable particularly at monthly and inter-annual scales.

P9346, L22-24, “Surface air. . . gauges” Not clear. Explain!

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper as below.

‘Because of the elevation differences among model cells and meteorological gauges, the interpolated surface air temperatures are further modified with a lapse rate of 6.5K/km.’

P9347, L14: Secondly instead of second.

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper.

P9347, L26: add ‘the’- “. . . of the above mentioned. . .”

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper.

P9347 – Fig 3 contains results and these should be presented and discussed in the results section and not here! Discuss your parameter values also.

Response:

Thanks for the suggestion. Section 4.1 has been added to present and discuss the hydrological model assessment results and hydrological model parameters. Discussion about hydrological model parameters in section 2.4 has been moved to section 4.1. The contents in section 4.1 are as below.

'WEB-DHM was calibrated against observed discharges of Biliu. Six main parameters were selected to calibrate using a trial and error approach due to the model's computational burden. Model parameter multipliers were calibrated, similar to the study by Wang et al. (2011). The 'Trial and error' approach has two steps. First, all the multiplier values are set to 1 which represents the default parameter values from Food and Agriculture Organization (FAO) (2003) and SiB2 model. Second, varying the multiplier values until acceptable discharge simulation accuracy is obtained. The calibrated parameter values are listed in Table 2. The simulated daily, monthly and inter-annual results are shown in Figs. 9a, 9c and 9e.

TOPMODEL uses basin-averaged parameter values, and these parameter values are estimated by experience or observation. However, these methods do not give precise parameter values. Therefore, the parameter values are considered as uncertain and provided with ranges based on experience (Beven and Kirkby, 1979;Beven and Freer, 2001a, b;Peters et al., 2003). Six parameters of TOPMODEL were calibrated using the dynamically dimensioned search algorithm (Tolson and Shoemaker, 2007), and the results are given in Table 3. The simulated daily, monthly and inter-annual results are shown in Figs. 9b, 9d and 9f.

Note that the parameters of TOPMODEL and WEB-DHM were calibrated using observed precipitation data, and the accuracy of simulated discharges has been validated using gauge observations. Comparison with the parameter values reported in previous research shows the parameter values are appropriate (Beven and Freer, 2001a;Peters et al., 2003;Qi et al., 2015).'

P9349, L14-15: this line is too similar in phrasing to what is in the paper by Bosshard et al.(2013). Revise!

Response:

Thanks for the suggestion. Changes have been accordingly in the revised paper as below.

'ANOVA could underestimate variance when the sample size is small (Bosshard et al., 2013).'

P9351, L21-23: Revise!

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper as below.

'None of the products can outperform others in terms of all the statistical criteria. This may be due to the different limitations of satellite sensors and inverse algorithms of precipitation products.'

P9353, L3: This may be attributed to the different. . .

Response:

Thanks for the suggestion. Changes have been made accordingly in the revised paper.

P9353, L10: Is the "trend" visually assessed?

Response:

We guess this question is about P93555 L10.

The trend is observed when compared with the probability of detection and critical success index. In Fig. 8 in the revised paper, comparatively the false alarm ratio has no obvious trend.

The first sentence in the third paragraph in section 3.4 has been rephrased as below in the revised paper.

'No obvious trends are observed for the false alarm ratio overall (compared with the probability of detection and critical success index), which means the false alarm ratio dependence on rainfall magnitude is weak.'

P9357, L14-23, "In the case. . . and models": Highlight where the peak discharge over/under estimation assessment is coming from? Clarify the reason in simple terms for the reader – what is the hydrological model influence and interactive influence?

Response:

Thanks for the suggestion. This sentence has been deleted.

P9358, L3: Is no better than?

Response:

Thanks for the suggestion. Yes. The RB and NSE of WEB-DHM and APHRODITE combination are -37% and 0.5, but they are -24% and 0.51 for the combination of TOPMDOEL and APHRODITE. This sentence has been added after the statement.

P9358, L4: This could be due to. . .

Response:

Thanks for the suggestion. Change has been made accordingly.

P9358, L19-25: Review and revise! The conclusion on TOPMODEL vs. WEB-DHM appears rushed. It's obvious that using good data and an inappropriate model structure results in poor performance, if parameter physical implications are kept in check! How do we achieve the coalition?

Response:

Thanks for the suggestion. As shown in the 'Major Comments 1', the parameters of TOPMODEL and WEB-DHM were calibrated using observed data, and the accuracy of simulated discharges have been validated using gauge observations. In addition, comparison with the parameter values reported in previous research (Beven and Freer, 2001a;Peters et al., 2003;Qi et al., 2015) shows the parameter values are appropriate.

The coalition can be achieved using a trial and error approach: for example, combining different hydrological models with a used precipitation product and selecting the best combination; combining different precipitation products with a used hydrological model and selecting the combination with best discharge simulation accuracy.

Having said this, the statements have been deleted and two paragraphs have been added in the revised paper as below.

The third paragraph in section 4.1:

'Note that the parameters of TOPMODEL and WEB-DHM were calibrated using observed precipitation data, and the accuracy of simulated discharges has been validated using gauge observations. Comparison with the parameter values reported in previous research shows the parameter values are appropriate (Beven and Freer, 2001a; Peters et al., 2003; Qi et al., 2015).'

The second paragraph in section 5:

'In addition to improving the accuracy of precipitation products, a good collation could help to achieve the performance in discharge simulations. Our approach provides a way to assess the different coalitions, i.e., the overall uncertainties in simulated discharges from different combinations of hydrological models and precipitation products. More precipitation products and hydrological models should be included and tested in the future work.'

P9359, L7-9: "This shows. . . accuracy" - What do you mean?

Response:

Thanks for the suggestion. This sentence has been deleted.

P9361, L1-2: Sentence is incomplete!

Response:

Thanks for the suggestion. The sentence has been rephrased as below.

'This research assesses the applicability of six precipitation products with fine spatial and temporal resolutions at a high latitude region in northeast China using both statistical and hydrological evaluation methods at multi-temporal scales.'

References:

Beven, K. J., and Freer, J. E.: A dynamic TOPMODEL, Hydrological Processes, 15, 1993-2011, 10.1002/hyp.252, 2001a.

Beven, K. J., and Freer, J. E.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, Journal of Hydrology, 249, 11–29, 10.1016/S0022-1694(01)00421-8, 2001b.

Beven, K. J., and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology, Hydrological Sciences Bulletin, 24, 43-69, 10.1080/02626667909491834,

1979.

Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M., and Schär, C.: Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections, Water Resources Research, 49, 1523-1536, 10.1029/2011wr011533, 2013.

Herman, J. D., Reed, P. M., and Wagener, T.: Time‐varying sensitivity analysis clarifies the effects of watershed model formulation on model behavior, Water Resources Research, 49, 1400-1414, 10.1002/wrcr.20124, 2013.

Peters, N. E., Freer, J., and Beven, K.: Modelling hydrologic responses in a small forested catchment (Panola Mountain, Georgia, USA): a comparison of the original and a new dynamic TOPMODEL, Hydrological Processes, 17, 345-362, 10.1002/hyp.1128, 2003.

Qi, W., Zhang, C., Fu, G., and Zhou, H.: Global Land Data Assimilation System data assessment using a distributed biosphere hydrological model, Journal of Hydrology, 528, 652-667, 10.1016/j.jhydrol.2015.07.011, 2015.

Sellers, P. J., Randall, D. A., Collatz, G. J., Berry, J. A., Field, C. B., Dazlich, D. A., Zhang, C., Collelo, G. D., and Bounoua, L.: A Revised Land Surface Parameterization (SiB2) for Atmospheric GCMS. Part I: Model Formulation, Journal of Climate, 9, 676-705, 10.1175/1520-0442(1996)009<0676:ARLSPF>2.0.CO;2, 1996.

Shou, Y., and Xu, J.: The rainstorm and mesoscale convective systems over northeast China in june 2005 I: A synthetic analysis of mcs by conventional observations and satellite data (in Chinese), Acta Meteorologica Sinica, 65, 160-170, 2007a.

Shou, Y., and Xu, J.: The rainstorm and mesoscale convective systems over northeast China in june 2005 II: A synthetic analysis of mcs's dynamical structure by radar and satellite observations (in Chinese), Acta Meteorologica Sinica, 65, 171-182, 2007b.

Tolson, B. A., and Shoemaker, C. A.: Dynamically dimensioned search algorithm for computationally efficient watershed model calibration, Water Resources Research, 43, 10.1029/2005wr004723, 2007.

Wang, F., Wang, L., Koike, T., Zhou, H., Yang, K., Wang, A., and Li, W.: Evaluation and application of a fine-resolution global data set in a semiarid mesoscale river basin with a distributed biosphere hydrological model, Journal of Geophysical Research, 116, 10.1029/2011jd015990, 2011.

Wang, F., Wang, L., Zhou, H., Saavedra Valeriano, O. C., Koike, T., and Li, W.: Ensemble hydrological prediction-based real-time optimization of a multiobjective reservoir during flood season in a semiarid basin with global numerical weather predictions, Water Resources Research, 48, W07520, 10.1029/2011wr011366, 2012.

Wang, L., Koike, T., Yang, D. W., and Yang, K.: Improving the hydrology of the Simple Biosphere Model 2 and its evaluation within the framework of a distributed hydrological model, Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 54, 989-1006, 10.1623/hysj.54.6.989, 2009a.

Wang, L., Koike, T., Yang, K., Jackson, T. J., Bindlish, R., and Yang, D.: Development of a distributed biosphere hydrological model and its evaluation with the Southern Great Plains Experiments (SGP97 and SGP99), Journal of Geophysical Research, 114, 10.1029/2008jd010800, 2009b.

Wang, L., Koike, T., Yang, K., and Yeh, P. J.-F.: Assessment of a distributed biosphere hydrological model against streamflow and MODIS land surface temperature in the upper Tone River Basin, Journal of Hydrology, 377, 21-34, 10.1016/j.jhydrol.2009.08.005, 2009c.

Yuan, M., Li, Z., and Zhang, X.: Analysis of a meso scale convective system during a brief torrential rain event in Northeast China (in Chinese), Acta Meteorologica Sinica, 68, 125-136, 10.11676/qxxb2010.013, 2010.

Zhou, J., Wang, L., Zhang, Y., Guo, Y., Li, X., and Liu, W.: Exploring the water storage changes in the largest lake (Selin Co) over the Tibetan Plateau during 2003-2012 from a basin-wide hydrological modeling, Water Resources Research, 10.1002/2014wr015846, 2015.