Improved Flood Forecasting in Ungauged Basins: Constrained Runoff Correction Using Multiple Satellite Products

Yanhong Dou[1], Lei Ye[1], Hoshin V Gupta[2], Hairong Zhang[3], and Huicheng Zhou[1]

[1]Dalian University of Technology
[2]The University of Arizona
[3]China Yangtze Power Co.,Ltd

November 22, 2022

Abstract

Satellite-based precipitation products (SPPs) with short latencies provide a new opportunity for flood forecasting in ungauged basins. However, the larger uncertainties associated with such near-real-time SPPs can influence the accuracy of the resulting flood forecast. Here we propose a real-time updating method, referred to as "Constrained Runoff Correction (CRC-M)" that is based on the use of multi-source SPPs. The method is based on the hypothesis that the range over different near-real-time SPPs provides insight regarding the approximate range in which the true rainfall value lies, during the current period. Accordingly, the constrained runoff correction is performed in such a way as to be consistent with this range, and with the observed value of discharge at the basin outlet. Evaluation using real-data indicates that the new method performs well, with Nash–Sutcliffe (NS) values of 0.85 and 0.91 during calibration and evaluation, respectively. The necessity and value of imposing constraints is demonstrated by comparing CRC-M against a control, referred to as "Unconstrained Runoff Correction" (URC-S). Experiments indicate that the key factors resulting in good performance are 1) wider constraint ranges, and 2) relatively reliable SPPs. Further, inclusion of redundant information may only result in slight improvements to forecast performance, and can even cause the performance to deteriorate. Overall, the CRC-M method can result in accurate and stable flood forecasts for ungauged basins, without the need for increased model complexity (i.e., the numbers of model parameters).

Improved Flood Forecasting in Ungauged Basins: Constrained Runoff Correction Using Multiple Satellite Products

Yanhong Dou[1], Lei Ye[1,*], Hoshin V. Gupta[2], Hairong Zhang [3,4], Huicheng Zhou[1]

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

[2]Department of Hydrology & Atmospheric Sciences, The University of Arizona, Tucson, Arizona, 85721, USA

[3] China Yangtze Power Co., Ltd., Yichang 443133, China

[4] Hubei Key Laboratory of Intelligent Yangtze and Hydroelectric Science, Yichang 443133, China

Email address:

yelei@dlut.edu.cn

Corresponding author:

Name: Lei Ye

Complete postal address: School of Hydraulic Engineering, Dalian University of Technology, Dalian 116024, China

Abstract: Satellite-based precipitation products (SPPs) with short latencies provide a new opportunity for flood forecasting in ungauged basins. However, the larger uncertainties associated with such near-real-time SPPs can influence the accuracy of the resulting flood forecast. Here we propose a real-time updating method, referred to as "Constrained Runoff Correction (CRC-M)" that is based on the use of multi-source SPPs. The method is based on the hypothesis that the range over different near-real-time SPPs provides insight regarding the approximate range in which the true rainfall value lies, during the current period. Accordingly, the constrained runoff correction is performed in such a way as to be consistent with this range, and with the observed value of discharge at the basin outlet. Evaluation using real-data indicates that the new method performs well, with Nash–Sutcliffe (NS) values of 0.85 and 0.91 during calibration and evaluation, respectively. The necessity and value of imposing constraints is demonstrated by comparing CRC-M against a control, referred to as "Unconstrained Runoff Correction" (URC-S). Experiments indicate that the key factors resulting in good performance are 1) wider constraint ranges, and 2) relatively reliable SPPs. Further, inclusion of redundant information may only result in slight improvements to forecast performance, and can even cause the performance to deteriorate. Overall, the CRC-M method can result in accurate and stable flood forecasts for ungauged basins, without the need for increased model complexity (i.e., the numbers of model parameters).

Key points:

- Present a method for improving flood forecasts in ungauged basin using multiple near-real-time satellite-based products.

- Give suggestions on how to select satellite-based products to form reasonable constraints.
- The method can be applied in basins with a variety of sizes and climates without any modifications to the hydrological model.

Key words: Multi-source SPPs; Ungauged basins; Flood forecasting; Runoff correction; IMERG; GSMaP

1. Introduction

Real-time forecasting of the rainfall-runoff process is an effective means to reduce flood risk. Accurate rainfall information plays an essential role in this process, since it is the primary driver for a hydrological model during flood events. Traditionally, ground-based rainfall observations have been widely used for hydrological forecasting due to their temporal resolution. However, their spatial coverage is typically not satisfactory, especially in remote regions, which translates into high forecast uncertainty. Recently, quasi-global satellite-based precipitation products (SPPs) have made it possible to develop spatial maps of rainfall with spatial resolutions that are finer than 0.25° and temporal resolutions that are shorter than daily, thereby provide a new opportunity for flood forecasting. Examples include the PERSIANN system (Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks; *Sorooshian et al., 2000*), the CMORPH system (CMORPH developed by the National Oceanic and Atmospheric Administration Climate Prediction Center; *Joyce et al., 2004*), the TMPA system (Tropical Rainfall Measuring Mission Multi-satellite

Precipitation Analysis; *Huffman et al., 2006*), the IMERG system (Integrated Multi-satellite Retrievals for the Global Precipitation Measurement; *Huffman et al., 2018b*), and the GSMaP system (Global Satellite Mapping of Precipitation; *Kubota et al., 2007*).

Numerous studies have examined the accuracy and hydrological utility of the available SPPs (Behrangi et al., 2011; L. Jiang & Bauer-Gottwein, 2019; S. Jiang et al., 2018; Li et al., 2017; Pan et al., 2010; Sirisena et al., 2018; Yuan et al., 2017). These studies highlight the fact that that even though SPPs with short latencies (e.g., “Early” and “Late” of IMERG, hereinafter referred to as IMERG-E and IMERG-L, respectively) perform worse than post-real-time gauge bias-adjusted SPPs (e.g., “Final” of IMERG, hereinafter referred to as IMERG-F), all of them are able to characterize precipitation trends reasonably well (*Yuan et al., 2019*). In other words, while such near-real-time precipitation estimates have larger uncertainties, they do contain useful information about rainfall. In ungauged basins, such products being the only availabile near-real-time rainfall information can be very valuable, particularly because the short latencies are important to flood forecasting. Therefore, in this paper, we examine the possibility of using these near-real-time SPPs for flood forecasting in ungauged basins.

Several previous studies have attempted to improve the accuracy of flood forecasting driven by SPPs. One way is to recalibrate the hydrologic model using such SPPs as input data *(Ciabatta et al., 2016; Yuan et al., 2018*). However, the model parameters obtained by this method tend to deviate from the ranges corresponding to physically plausible values (Bitew & Gebremichael, 2011; S. Jiang et al., 2012). Further, this method seems to only improve flood forecasting accuracy to a limited extent.

Other than this, methods to further improve the accuracy of flood forecasting driven by SPPs are of three types. The first is to bias-correct the SPPs to compensate for the deviation from ground-based observations before they are used to force the hydrological models *(Deng et al., 2019; Harris et al., 2007)*. The second is to integrate over the discharge estimates computed using a variety of SPPs to force the hydrological models *(Sun et al., 2018; Jiang et al., 2014)*. The third is to perform updating of the real-time discharge forecasts using observed discharge or soil moisture measurements *(Wei et al., 2015; Massari et al., 2019; Massari et al., 2018)*. Of these, given that the ground-based discharge can be directly measured, the method of real-time correction using discharge has been reported to give the best results.*(Deng et al., 2019; Sun et al., 2018; Wei Si, Weimin Bao, 2015)* Further, such an approach is suitable even for basins having no precipitation stations at all. As an important intermediate variable in conceptual rainfall-runoff models, runoff forecasts have been corrected using observed real time discharge data to obtain high flood forecasting accuracy (*Liu et al., 2012*). For example, *Weimin et al. (2014)* proposed a runoff correction method based on a dynamic system response curve, which can significantly improve the accuracy of flood forecast especially when the quality of the driving data is poor.

To the best of our knowledge, there have been no investigations into the use of multi-source near-real-time SPPs for real-time correction of runoff to improve the accuracy of flood forecasts. Here, we propose and evaluate a constrained runoff correction method using multi-source SPPs. The physical basis of this method is the flow concentration process of the conceptual rainfall-runoff model. Conceptually, we assume that the near-real-time SPPs retrieved using different algorithms can be considered to approximately span a range that includes the true value of rainfall. Therefore, the range formed by different runoff predictions, computed using different SPPs, will be treated as the likely range in which the true runoff is expected to occur. Constrained runoff correction will be performed using this range to constrain the discharge simulations at the basin outlet.

The rest part of the paper is organized as follows. Section 2 describes the study area and various data sources used. Section 3 describes the details of the Constrained Runoff Correction method based on multi-source SPPs (CRC-M). To demonstrate the importance of using a range to constrain the runoff correction, a control test named Unconstrained Runoff Correction based on a single SPP (URC-S) is also conducted in this section. Section 4 reports the results and illustrates the importance of using reasonable constraint ranges in this method, followed by conclusions in Section 5.

2. Study area and data

2.1 Study area

The Xiaoergou River, a tributary of the Nenjiang River, originates at the southern foot of the Greater Khingan Range, China (Figure 1). The Xiaoergou Basin, with a drainage area of 16761 km^2, has a complex east-west inclined terrain with elevations ranging from 209 m to 1413 m, and extends from longitudes 121.73 °E to 123.88 °E and latitudes 48.83 °N to 50.61 °N in the northeast of Inner Mongolia in northeast China. Benefiting from its location in the middle and high latitudes on the eastern coast of Eurasia, this area has a typical temperate continental climate with long cold winters, dry and windy springs, hot and rainy summers, a short autumn, and very uneven precipitation in time. In the Xiaoergou Basin, the annual average temperature is approximately -1.2 ℃, the highest temperature in the past year is 40.1℃, the lowest temperature is -35.4℃, the frost-free period is 132 days, the main soil types are black soil, coniferous forest soil, dark brown forest soil and humus fen soil. There is no precipitation station in the basin, and only one hydrologic station is located at the basin outlet. It is the largest ungauged sub-basin of the Nenjiang basin with about 400-500 mm of average annual rainfall. Furthermore, due to the lack of a detention reservoir between the Xiaoergou basin and Qiqihar, which is an important city at downstream of the basin, the flood risk is high. The features of the Xiaoergou Basin make it an ideal region for investigating the use of SPP’s for flood forecasting at middle-high latitudes.

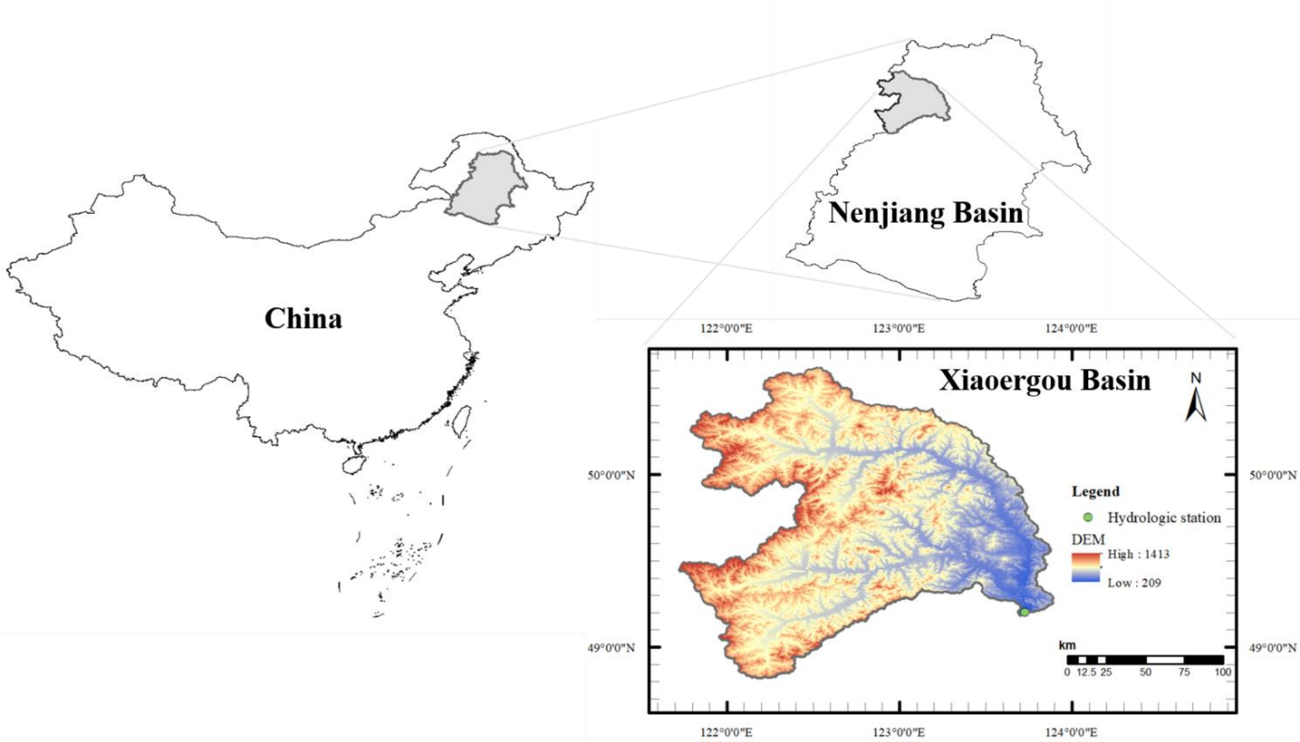

Figure 1. Geographical location, topography and hydrologic station of the Xiaoergou Basin

2.2 Data

2.2.1 Satellite-based products

The satellite-based products used in this study include four near-real-time SPPs, i.e., IMERG-E, IMERG-L, GSMaP-NRT (hereinafter referred to as GSMaP-N), and GSMaP-Gauge-NRT (hereinafter referred to as GSMaP-GN), and two different types of evapotranspiration products, i.e., GLDAS-CLM and GLEAM. The products used in this paper have a daily temporal resolution. Further, the products were spatially averaged over the basin extent to generate the forcing data required by the lumped hydrologic model; this is consistent with previous studies that have shown that aggregating the SPPs over larger spatial regions results in improved characterization of rainfall (Omranian & Sharif, 2018; Tan et al., 2017).

(1) Precipitation

IMERG and GSMaP are generally referred to as the new generation products, namely, the products of the Global Precipitation Measurement (GPM) era (*Hou et al., 2014*). This is because IMERG and GSMaP are retrieved from the GPM mission which was officially launched in February 2014 to replace TRMM. IMERG and GSMaP provide more expansive coverage than TRMM, which makes it possible to use SPPs to forecast flood in middle-high latitude basins. IMERG is the level 3 multi-satellite precipitation algorithm of GPM developed by NASA (https://disc.gsfc.nasa.gov), which combines all available constellation observations of the more accurate but infrequent microwave (MW) and more frequent but indirect infrared (IR) (Hou et al., 2014).

The IMERG system is run several times, where the latencies of "Early" (IMERG-E) and "Late" (IMERG-L) are within one day (~4 h and ~12 h after observation time, respectively), so that they can be used to forecast flooding at daily time scales. IMERG-E provides a preliminary estimate using only forward morphing and IMERG-L corrects IMERG-E using both forward and backward morphing as more data arrive. For a more detailed description of IMERG algorithm, readers can refer to (Huffman et al., 2019; Huffman, Bolvin, et al., 2018; Huffman, Gsfc, et al., 2018).

GSMaP is a satellite-based precipitation map algorithm developed by JAXA (http://www.jaxa.jp) for merging the observation of IR sensors and passive MW (PMW) radiometer from GPM Core GMI. The GSMaP products are produced in several steps. Firstly, the instantaneous precipitation rate is retrieved based on the PMW radiometers from different satellite platforms, including GMI, advanced microwave scanning radiometer 2 (AMSR2), TRMM Microwave Imager (TMI), special sensor microwave imager/sounder (SSMIS), advanced microwave sounding unit-A (AMSU-A), and microwave humidity sounder (MHS)(*Zhu et al., 2018*). Then, the gaps between PMW-based estimates are propagated using the cloud motion vectors computed from geo-IR images (*Duan et al., 2016*).

Near-real-time products retrieved by GSMaP include GSMaP-N and GSMaP-GN, where GSMaP-GN is calibrated using gauge-based data, and the latencies of both are 4 hours, so that they can also be used to forecast flooding at daily time scales. Furthermore, since the data sources and retrieval algorithms are different between IMERG and GSMaP, the rainfall values of the four products in the same period are different, which provides a possible range for the rainfall estimate in that period. The distribution of the width of the rainfall range can be seen in Table 3.

Data from four SPPs (IMERG-E, IMERG-L, GSMaP-N and GSMaP-GN), from April 2014 to November 2018, was used in this study. Basic information regarding the four products is provided in Table 1. IMERG-L and GSMaP-GN belong to the "corrected" class, because they use more data than IMERG-E and GSMaP-N, respectively.

Table 1. Overview of the SPPs used in this study

Product		Corrected	Spatial Resolution	Developer	Start Time	Latency
IMERG	Early	No	0.1 °×0.1 °	NASA	June 2000	4 hours
	Late	Yes			June 2000	12 hours
GSMaP	NRT	No	0.1 °×0.1 °	JAXA	March 2000	4 hours
	NRT-gauge	Yes			April 2000	4 hours

(2) Evapotranspiration

To illustrate the importance of various sources of information used in this paper, two kinds of evapotranspiration products with different retrieval algorithms are selected (period April 2014 to November 2018); basic information is provided in Table 2.

Table 2. Overview of the evapotranspiration (ET) products used in this study

Data Sets	**Category**	**Scheme**	**Spatial Resolution**	**Start Time**
GLDAS	LSM	Penman-Monteith	1 °×1 °	January 1979
GLEAM	Diagnostic	Priestley-Taylor	0.25 °× 0.25 °	January 1980

The first product is a simulation generated by the Common Land Model (CLM) V2.0 model as part of the Global Land Data Assimilation System (GLDAS; https://disc.gsfc.nasa.gov). The simulation was forced by a combination of NOAA/GDAS atmospheric analysis fields, spatially and temporally disaggregated NOAA Climate Prediction Center Merged Analysis of Precipitation (CMAP) fields, and observation-based downward shortwave and longwave radiation fields derived using the method of the Air Force Weather Agency's AGRicultural METeorological modeling system (AGRMET). The data is available from January 1979 to present at a 3-hour timestep (*Rodell et al., 2004*). For our purpose, the GLDAS-CLM product was aggregated to daily time-step.

The other product is the version 3.3a dataset of Global Land Evaporation Amsterdam Model (GLEAM; https://www.gleam.eu), which is globally available from 1980 to present at daily temporal resolution and 0.25° spatial resolution. The retrieval algorithm is based on a diagnostic model that takes advantage of the Priestley and Taylor equation. The dataset is based on observations of surface net radiation, near-surface air temperature, precipitation, soil moisture, snow water equivalent and vegetation optical depth *(Miralles et al., 2011; Martens et al., 2017)*. Due to essential difference between GLDAS-CLM and GLEAM, there is a gap between the two products. The distribution of the gap can be seen in the Table 3. Further, the depletion of evapotranspiration from precipitation, as the most common process in conceptual rainfall-runoff models, widens the range of the hydrologic model forcing data , whose distribution can also be seen in the Table 3. Since there are 4 kinds of precipitation estimates and 2 kinds of evapotranspiration estimates, this results in 8 pair-wise combinations of precipitation and evapotranspiration.

Table 3. Distributions of the width of the ranges formed by different satellite-based products for a given period

Variable	**Range [min, max] (mm)**	**Interquartile Range [$\frac{1}{4}$min, $\frac{3}{4}$max] (mm)**	**Median (mm)**
Precipitation (P)	[0, 46.57]	[0.17, 3.54]	1.17
Evapotranspiration (ET)	[0, 5.86]	[0.41, 1.42]	0.97
P-ET	[0.04, 47.89]	[1.26, 5.10]	2.25

2.2.2 Gauged discharge data

Daily discharge data for 2014 to 2018, covering the same period as the GPM SPPs, were obtained from the Xiaoergou hydrologic station. The period includes 14 flood events with various net rainfall and flood peaks (Figure 2), of which the first 9 were used for calibration and the last 5 for verification.

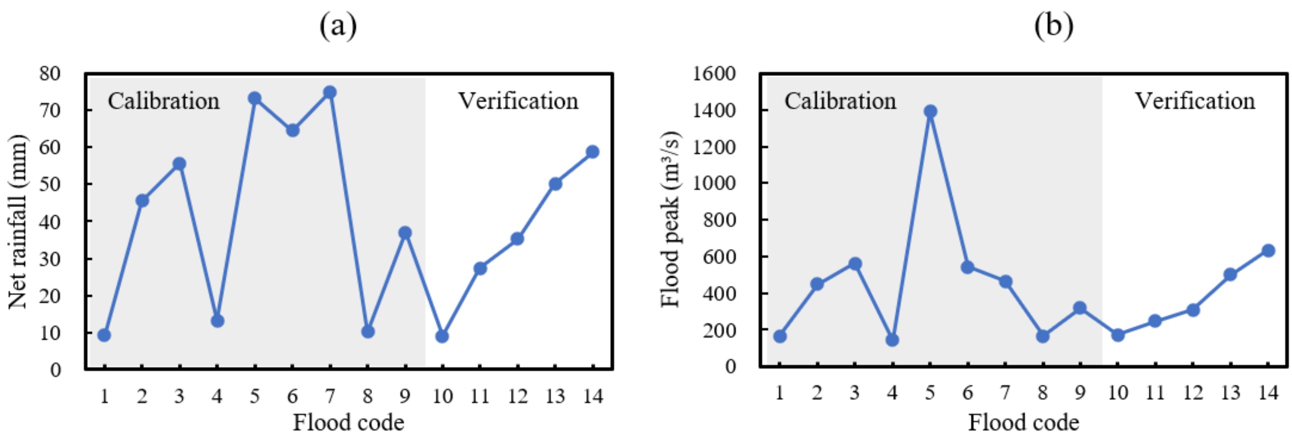

Figure 2. Net rainfall (a) and flood peaks (b) of 14 flood events of the Xiaoergou Basin

3. Methodology

3.1 Conceptual hydrological model

Catchment-scale hydrological models can be conceptualized as shown in Figure 3 and can be expressed mathematically as

$$\boldsymbol{Q} = \boldsymbol{f}[\boldsymbol{P}, \boldsymbol{E}, \boldsymbol{X}, \boldsymbol{\theta}] \tag{1}$$

where $\boldsymbol{Q} = [Q_1, Q_2, Q_3, \cdots, Q_m]^T$ is a time-ordered vector of discharge at the outlet, $\boldsymbol{P} = [P_1, P_2, P_3, \cdots, P_n]^T$ is a corresponding vector of areal mean rainfall from any SPPs, $\boldsymbol{E} = [E_1, E_2, E_3, \cdots, E_n]^T$ is a corresponding vector of areal mean evapotranspiration from any evapotranspiration product, $\boldsymbol{X} = [X_1, X_2, X_3, \cdots, X_n]^T$ is a corresponding vector of state values, and $\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3, \cdots]^T$ is the vector of model parameters.

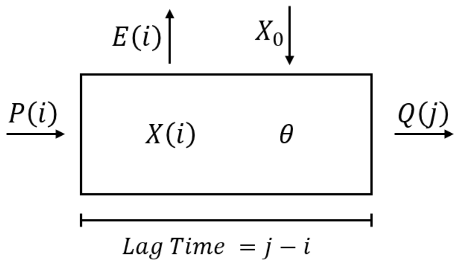

Figure 3. A high-level systems diagram for a hydrological model; *P*(*i*), *E*(*i*) and *X*(*i*) indicate the areal mean rainfall, the areal mean evapotranspiration and the initial state of basin at the *i*th period, and *Q*(*j*) indicates the computed discharge output of the hydrological system at the *j*th period and also the first response to *P*(*i*), *E*(*i*) and *X*(*i*) of the basin outlet. Therefore, the lag time is *j-i*.

Each parameter and time-ordered vector in formula (1) will affect the discharge process at the basin outlet. Once the parameters are determined, they are assumed to remain constant over time, so the above vectors are important factors in the water balance. Runoff is a variable which can synthetically reflect the above vectors, since it is calculated from them. Runoff represents the total amount of water accumulated within a certain period, whereas discharge represents the local value at a particular point and moment in time. To compute runoff, the runoff generation module, a subsystem of the hydrological model, can be expressed as:

$$\boldsymbol{R(i)} = \boldsymbol{f_1}[\boldsymbol{P(i)}, \boldsymbol{E(i)}, \boldsymbol{X(i)}, \boldsymbol{\theta}'] \tag{2}$$

where $R(i)$ is runoff at period i, $X(i) = g_1[\boldsymbol{P}, \boldsymbol{E}, X_0, \boldsymbol{\theta}']$ and $\boldsymbol{\theta}'$ is a subset of $\boldsymbol{\theta}$.

According to formula (2), besides the control exerted by the model parameters, the precision of the runoff is controlled mainly by the precision of rainfall and evapotranspiration, since the state values are calculated mainly from rainfall and evapotranspiration. However, the use of satellite-based products as system drivers results in added uncertainty associated with the runoff generation process. In an attempt to correct this error, this study will focus only on the hydrological model response to $\boldsymbol{R}$. Therefore, formula (1) can be expressed as

$$\boldsymbol{Q} = \boldsymbol{Q}[\boldsymbol{R}, \boldsymbol{\theta}] \tag{3}$$

By inputting rainfall, evapotranspiration and basin state values into the hydrological model at any period, the following can be obtained:

$$\boldsymbol{q} = \boldsymbol{f}[\boldsymbol{P(i)}, \boldsymbol{E(i)}, \boldsymbol{X(i)}, \boldsymbol{\theta}] \tag{4}$$

where $\boldsymbol{q} = [q_1, q_2, q_3, \ldots, q_m]^T$ is the response of hydrological model to $R(i)$, that is, the discharge process of $R(i)$ at the basin outlet. Therefore, formula (4) can also be expressed as

$$\boldsymbol{q} = q[R(i), \boldsymbol{\theta}] \tag{5}$$

$$q(j, i) = q[R(i), \boldsymbol{\theta}, j] \tag{6}$$

where $q[R(i), \boldsymbol{\theta}, j] = 0$ when i $\geqslant$ j.

The accumulation of $\boldsymbol{q}$ in n periods is $\boldsymbol{Q} = [Q(\boldsymbol{R}, \boldsymbol{\theta}, 1), Q(\boldsymbol{R}, \boldsymbol{\theta}, 2), \cdots, Q(\boldsymbol{R}, \boldsymbol{\theta}, m)]^T$, namely the simulated discharge at the basin outlet, as shown in Figure 4, which can be expressed as

$$\begin{cases} Q(\boldsymbol{R},\boldsymbol{\theta},1) = q[R(1),\boldsymbol{\theta},1] + q[R(2),\boldsymbol{\theta},1] + \cdots + q[R(n),\boldsymbol{\theta},1] \\ Q(\boldsymbol{R},\boldsymbol{\theta},2) = q[R(1),\boldsymbol{\theta},2] + q[R(2),\boldsymbol{\theta},2] + \cdots + q[R(n),\boldsymbol{\theta},2] \\ \vdots \\ Q(\boldsymbol{R},\boldsymbol{\theta},m) = q[R(1),\boldsymbol{\theta},m] + q[R(2),\boldsymbol{\theta},m] + \cdots + q[R(n),\boldsymbol{\theta},m] \end{cases} \tag{7}$$

Its matrix form is

$$\boldsymbol{Q} = \boldsymbol{L}\boldsymbol{A} \tag{8}$$

where:

$$\boldsymbol{A} = [1,1,\cdots,1]^T,$$

$$\boldsymbol{L} = \begin{bmatrix} q[R(1),\boldsymbol{\theta},1] & q[R(2),\boldsymbol{\theta},1] & \cdots & q[R(n),\boldsymbol{\theta},1] \\ q[R(1),\boldsymbol{\theta},2] & q[R(2),\boldsymbol{\theta},2] & \cdots & q[R(n),\boldsymbol{\theta},2] \\ \vdots & \vdots & \ddots & \vdots \\ q[R(1),\boldsymbol{\theta},m] & q[R(2),\boldsymbol{\theta},m] & \cdots & q[R(n),\boldsymbol{\theta},m] \end{bmatrix}.$$

The ith column in $\boldsymbol{L}$ is the discharge process of $R(i)$ at the basin outlet. Therefore, to correct the error of $\boldsymbol{R}$, it is necessary to correct $\boldsymbol{L}$ column using real-time updated vector of observed discharge, that is, to retrieve the correction coefficient vector, namely $\boldsymbol{A}' = [\alpha_1,\alpha_2,\cdots,\alpha_{m'}]^T$, which can be used to forecast discharge during the flood.

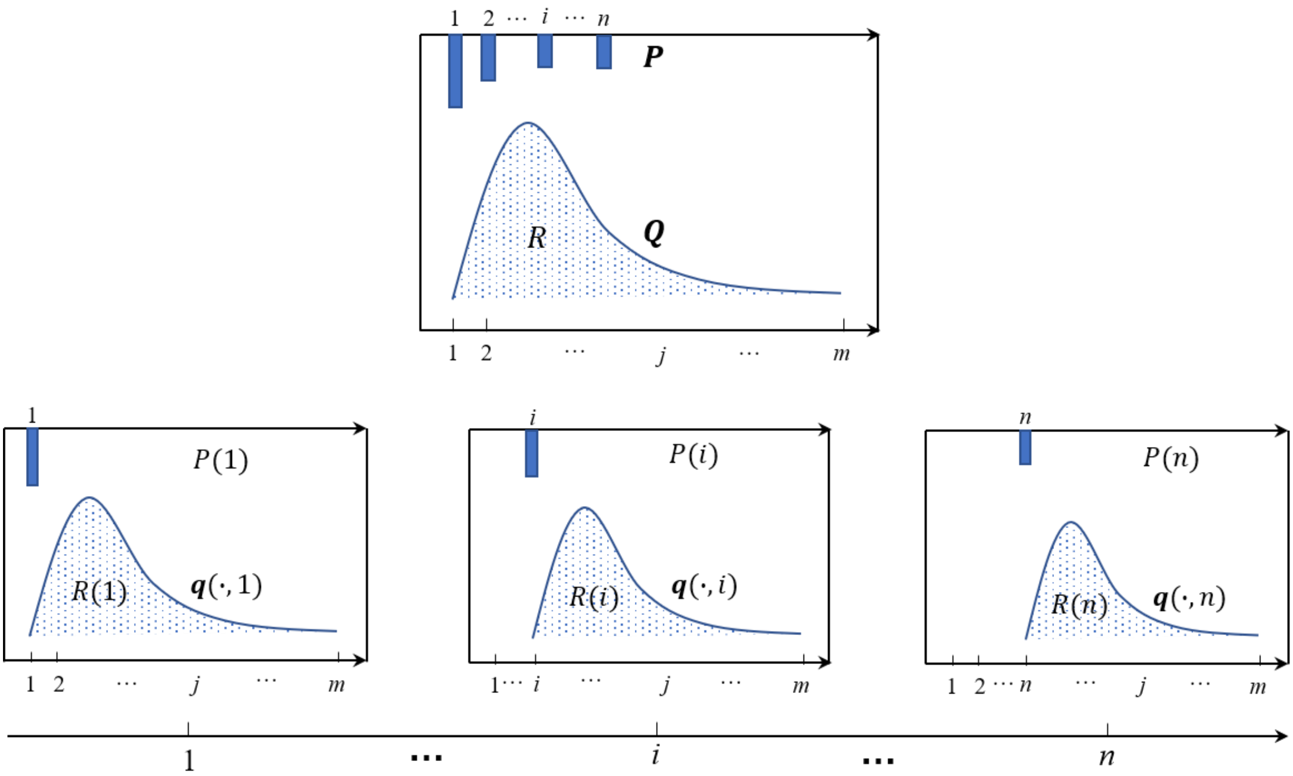

Figure 4. Schematic diagram of formula (2)~(8)

3.2 Constrained runoff correction based on multi-source SPPs

When, for any period, there are h ($h \geq 2,\ h \in \mathrm{Z}$) combinations of rainfall and evapotranspiration products available to force a hydrologic model, h different runoff responses can be calculated via the formula (2), that is, $R_1(i), \cdots, R_h(i)$. This can result in h kinds of responses of hydrological model to $R(i)$, that is, $\boldsymbol{q_1}, \cdots, \boldsymbol{q_h}$, so there are h different $\boldsymbol{L}$s, that is, $\boldsymbol{L_1}, \cdots, \boldsymbol{L_h}$. Accordingly, the upper bound matrix ($\overline{\boldsymbol{L}}$) and lower bound matrix ($\underline{\boldsymbol{L}}$) required for constrained correction can be obtained from formula (9-10). It should be noted that the bigger h is, the wider is the range formed by the $\overline{\boldsymbol{L}}$ and the $\underline{\boldsymbol{L}}$, and vice versa.

$$\bar{\boldsymbol{L}} = \begin{bmatrix} \bar{q}[R(1),\boldsymbol{\theta},1] & \bar{q}[R(2),\boldsymbol{\theta},1] & \cdots & \bar{q}[R(n),\boldsymbol{\theta},1] \\ \bar{q}[R(1),\boldsymbol{\theta},2] & \bar{q}[R(2),\boldsymbol{\theta},2] & \cdots & \bar{q}[R(n),\boldsymbol{\theta},2] \\ \vdots & \vdots & \ddots & \vdots \\ \bar{q}[R(1),\boldsymbol{\theta},m] & \bar{q}[R(2),\boldsymbol{\theta},m] & \cdots & \bar{q}[R(n),\boldsymbol{\theta},m] \end{bmatrix} \tag{9}$$

$$\bar{q}(j,i) = \max\,[q_1(j,i),\cdots,q_h(j,i)]$$

$$\underline{\boldsymbol{L}} = \begin{bmatrix} \underline{q}[R(1),\boldsymbol{\theta},1] & \underline{q}[R(2),\boldsymbol{\theta},1] & \cdots & \underline{q}[R(n),\boldsymbol{\theta},1] \\ \underline{q}[R(1),\boldsymbol{\theta},2] & \underline{q}[R(2),\boldsymbol{\theta},2] & \cdots & \underline{q}[R(n),\boldsymbol{\theta},2] \\ \vdots & \vdots & \ddots & \vdots \\ \underline{q}[R(1),\boldsymbol{\theta},m] & \underline{q}[R(2),\boldsymbol{\theta},m] & \cdots & \underline{q}[R(n),\boldsymbol{\theta},m] \end{bmatrix} \tag{10}$$

$$\underline{q}(j,i) = \min\,[q_1(j,i),\cdots,q_h(j,i)]$$

(1) Correction

As shown in Figure 4, the rainfall event lasts for *n* time periods, and the discharge process generated by this rainfall lasts for *m* time periods. When there are m' $(m' \le n < m)$ values of observed discharge $\boldsymbol{Q_o}$, the $R_1(i),\cdots,R_h(i)$ in m' periods are also known, so that the vector of observed discharge can be implemented through a simple nudging scheme, as in Eq. (11).

$$\boldsymbol{Q_o} = \underline{\boldsymbol{L}}_{(m'\times m')}\boldsymbol{A} + \Delta\boldsymbol{L}_{(m'\times m')}\boldsymbol{A}' + \boldsymbol{\varepsilon} \tag{11}$$

where $\boldsymbol{Q_o} = [Q_o(\boldsymbol{R},\boldsymbol{\theta},1), Q_o(\boldsymbol{R},\boldsymbol{\theta},2),\cdots,Q_o(\boldsymbol{R},\boldsymbol{\theta},m')]^T$; $\Delta\boldsymbol{L}_{(m'\times m')} = \bar{\boldsymbol{L}}_{(m'\times m')} - \underline{\boldsymbol{L}}_{(m'\times m')}$, $\bar{\boldsymbol{L}}_{(m'\times m')}$ and $\underline{\boldsymbol{L}}_{(m'\times m')}$ are the pre-m'-order submatrices of $\bar{\boldsymbol{L}}$ and $\underline{\boldsymbol{L}}$; $\boldsymbol{A}' = [\alpha_1,\alpha_2,\cdots,\alpha_{m'}]^T$ is the constrained correction coefficient vector that has the similar role of the Kalman gain in the classic data assimilation technique, where $0 \le \alpha_i \le 1$, and the greater the value, the higher the weight is given to the upper bound, and lower weight is given to the lower bound. $\alpha_i = 0$ or $\alpha_i = 1$ respectively indicates that the $\boldsymbol{q}$ calculated by $R(i)$ is approximately equal to the lower or upper bound of *h* discharge processes. $\boldsymbol{\varepsilon} = [e_1,e_2,\cdots,e_{m'}]^T$ is the random error vector. According to the least-squares principle, $\boldsymbol{A}'$ satisfies:

$$\min_{\boldsymbol{A}'} \boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} = \min_{\boldsymbol{A}'}\left[\boldsymbol{Q_o} - \underline{\boldsymbol{L}}_{(m'\times m')}\boldsymbol{A} - \Delta\boldsymbol{L}_{(m'\times m')}\boldsymbol{A}'\right]^T\left[\boldsymbol{Q_o} - \underline{\boldsymbol{L}}_{(m'\times m')}\boldsymbol{A} - \Delta\boldsymbol{L}_{(m'\times m')}\boldsymbol{A}'\right] \tag{12}$$

In this way,

$$\boldsymbol{A}' = (\Delta\boldsymbol{L}_{(m'\times m')}{}^T\Delta\boldsymbol{L}_{(m'\times m')})^{-\mathbf{1}}\Delta\boldsymbol{L}_{(m'\times m')}{}^T\left(\boldsymbol{Q_o} - \underline{\boldsymbol{L}}_{(m'\times m')}\boldsymbol{A}\right) \tag{13}$$

In addition, $\alpha_i < 0$ or $\alpha_i > 1$ respectively indicates that the corrected $\boldsymbol{q}$ is below the lower bound or above the upper bound. To constrain correcting $\boldsymbol{R}$, the elements in $\boldsymbol{A}'$ are constrained. So, if $\alpha_i > 1$, then $\alpha_i = 1$; if $\alpha_i < 0$, then $\alpha_i = 0$.

Figure 5 shows the correction and forecasting process of the CRC-M by taking T=1 and m'=3 as an example. Since $\bar{\boldsymbol{L}}_{(m'\times m')}$, $\underline{\boldsymbol{L}}_{(m'\times m')}$ and $\Delta\boldsymbol{L}_{(m'\times m')}$ are strictly lower triangular matrix, as shown in the Figure 5, only the first $(m' - T)$ elements in $\boldsymbol{A}'$ are deterministic solutions, and the rest T elements are not unique, where T is the lag time in Figure 3.

(2) Forecasting

For forecasting, not only the first $(m' - T)$ elements in $\boldsymbol{A}'$ are important but also the last T ones, because the future discharge process is influenced by all the rainfall in the preceding m' periods. Considering the last T elements in $\boldsymbol{A}'$ are the key factors for discharge forecasting, let them be equal to $\alpha_{m'-T}$. Hereby, we can update the time series of forecasting discharge between $(m' + 1)$th and mth period using $\boldsymbol{A}'$ by formula (14).

$$\begin{pmatrix} Q_c(\boldsymbol{R},\boldsymbol{\theta},m'+1) \\ Q_c(\boldsymbol{R},\boldsymbol{\theta},m'+2) \\ \vdots \\ Q_c(\boldsymbol{R},\boldsymbol{\theta},m) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{m'} \underline{q}_{m'+1,i} \\ \sum_{i=1}^{m'} \underline{q}_{m'+2,i} \\ \vdots \\ \sum_{i=1}^{m'} \underline{q}_{m,i} \end{pmatrix} + \begin{bmatrix} \Delta q_{m'+1,1} & \Delta q_{m'+1,2} & \cdots & \Delta q_{m'+1,m'} \\ \Delta q_{m'+2,1} & \Delta q_{m'+2,2} & \cdots & \Delta q_{m'+2,m'} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta q_{m,1} & \Delta q_{m,2} & \cdots & \Delta q_{m,m'} \end{bmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{m'-T} \\ \vdots \\ \alpha_{m'-T} \end{pmatrix} \tag{14}$$

$$\boldsymbol{Q}_{\boldsymbol{c}(m-m')} = \underline{\boldsymbol{L}}_{[(m-m')\times m']}\boldsymbol{A} + \Delta\boldsymbol{L}_{[(m-m')\times m']}\boldsymbol{A}'$$

where $\Delta\boldsymbol{L}_{[(m-m')\times m']} = \bar{\boldsymbol{L}}_{[(m-m')\times m']} - \underline{\boldsymbol{L}}_{[(m-m')\times m']}$, $\bar{\boldsymbol{L}}_{[(m-m')\times m']}$ and $\underline{\boldsymbol{L}}_{[(m-m')\times m']}$ respectively indicate the submatrix of $\bar{\boldsymbol{L}}$ and $\underline{\boldsymbol{L}}$ in rows $m'+1\sim m$, columns $1\sim m'$. $\boldsymbol{Q}_{\boldsymbol{c}(m-m')} = [Q_c(\boldsymbol{R},\boldsymbol{\theta},m'+1), Q_c(\boldsymbol{R},\boldsymbol{\theta},m'+2), \cdots, Q_c(\boldsymbol{R},\boldsymbol{\theta},m)]^T$ is a subset of $\boldsymbol{Q}_{\boldsymbol{c}}$, which is a time-ordered vector of forecasting discharge obtained by this correction process. With the increase of m', the $\boldsymbol{Q}_{\boldsymbol{c}(m-m')}$ keeps updating, leading to the constant update of $\boldsymbol{Q}_{\boldsymbol{c}}$. Once $m' > n$, all the $\boldsymbol{q}$s during n periods are corrected and $\boldsymbol{Q}_{\boldsymbol{c}}$ is no longer updated. Therefore, the final $\boldsymbol{Q}_{\boldsymbol{c}}$ can be obtained after a total of n updates, whose process is shown in Figure 8.

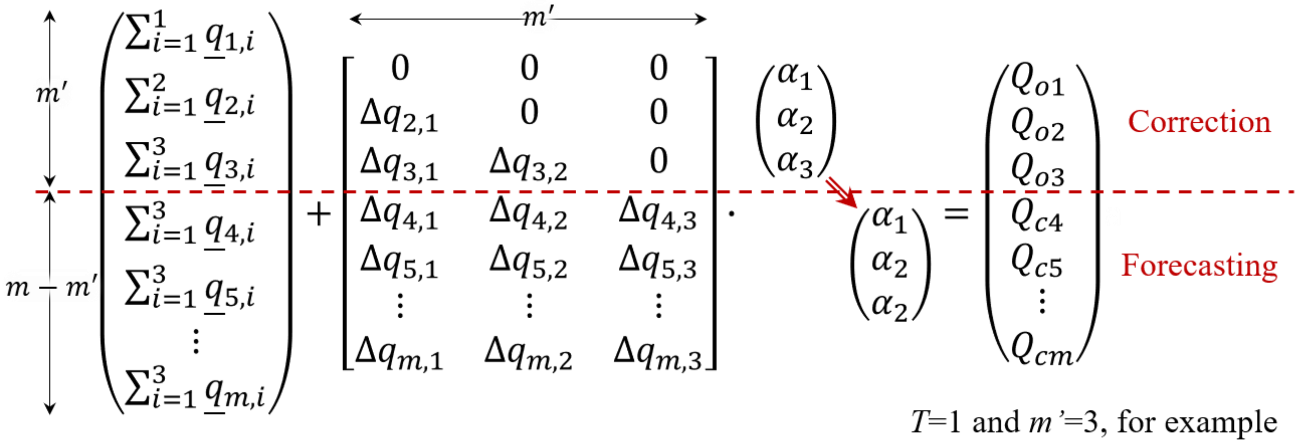

Figure 5. The correction and forecasting process of the CRC-M by taking T=1 and m'=3 as an example, where $\Delta q_{j,i} = \bar{q}_{j,i} - \underline{q}_{j,i}$.

3.3 Unconstrained runoff correction based on a single SPP

Considering the widespread tendency of SPP's to underestimate extreme precipitation *(Deng et al.,* *2019; Zhang et al., 2019; Su et al., 2019)*, by using them to correct runoff with constraints it may be difficult to characterize extreme runoff. In order to demonstrate the necessity of the constraint in the process of runoff correction, a control test is set up, in which unconstrained runoff correction is conducted based on a single satellite-based product (URC-S) that still modifies the $\boldsymbol{L}$ column by column by scaling. Figure 6 shows the correction and forecasting process of the URC-S by taking T=1 and m'=3 as an example.

(1) Correction

When there are m' $(m' \leq n < m)$ values of observed discharge, the $R(i)$ in m' periods are also known, then

$$\begin{cases} Q_o(\boldsymbol{R},\boldsymbol{\theta},1) \approx q[R(1),\boldsymbol{\theta},1]\alpha_1 + q[R(2),\boldsymbol{\theta},1]\alpha_2 + \cdots + q[R(m'),\boldsymbol{\theta},1]\alpha_{m'} \\ Q_o(\boldsymbol{R},\boldsymbol{\theta},2) \approx q[R(1),\boldsymbol{\theta},2]\alpha_1 + q[R(2),\boldsymbol{\theta},2]\alpha_2 + \cdots + q[R(m'),\boldsymbol{\theta},2]\alpha_{m'} \\ \vdots \\ Q_o(\boldsymbol{R},\boldsymbol{\theta},m') \approx q[R(1),\boldsymbol{\theta},m']\alpha_1 + q[R(2),\boldsymbol{\theta},m']\alpha_2 + \cdots + q[R(m'),\boldsymbol{\theta},m']\alpha_{m'} \end{cases} \tag{15}$$

Its matrix form is

$$\boldsymbol{Q_o} = \boldsymbol{L}_{(m' \times m')}\boldsymbol{A}' + \boldsymbol{\varepsilon} \tag{16}$$

where $\boldsymbol{L}_{(m' \times m')}$ is the pre-m'-order submatrix of $\boldsymbol{L}$; $\boldsymbol{A}' = [\alpha_1, \alpha_2, \cdots, \alpha_{m'}]^T$ is the unconstrained correction coefficient vector, where α_i is a member of the real number. According to the least square principle, $\boldsymbol{A}'$ satisfies:

$$\boldsymbol{A}' = (\boldsymbol{L}_{(m' \times m')}^T \boldsymbol{L}_{(m' \times m')})^{-1} \boldsymbol{L}_{(m' \times m')}^T \boldsymbol{Q_o} \tag{17}$$

Since $\boldsymbol{L}_{(m' \times m')}$ is also strictly lower triangular matrix, as shown in the Figure 6, only the first $(m' - T)$ elements in $\boldsymbol{A}'$ are deterministic solutions, where T is the lag time in Figure 3.

(2) Forecasting

By the same logic as the previous method, let the last T elements in $\boldsymbol{A}'$ be equal to $\alpha_{m'-T}$. Hereby, we can update the time series of forecasting discharge between $(m'+1)$th and mth period using $\boldsymbol{A}'$ by formula (18).

$$\begin{cases} Q_c(\boldsymbol{R},\boldsymbol{\theta},m'+1) = q[R(1),\boldsymbol{\theta},m'+1]\alpha_1 + q[R(2),\boldsymbol{\theta},m'+1]\alpha_2 + \cdots + q[R(m'),\boldsymbol{\theta},m'+1]\alpha_{m'} \\ Q_c(\boldsymbol{R},\boldsymbol{\theta},m'+2) = q[R(1),\boldsymbol{\theta},m'+2]\alpha_1 + q[R(2),\boldsymbol{\theta},m'+2]\alpha_2 + \cdots + q[R(m'),\boldsymbol{\theta},m'+2]\alpha_{m'} \\ \vdots \\ Q_c(\boldsymbol{R},\boldsymbol{\theta},m) = q[R(1),\boldsymbol{\theta},m]\alpha_1 + q[R(2),\boldsymbol{\theta},m]\alpha_2 + \cdots + q[R(m'),\boldsymbol{\theta},m]\alpha_{m'} \end{cases}$$

(18)

Its matrix form is

$$\boldsymbol{Q}_{c(m-m')} = \boldsymbol{L}_{[(m-m')\times m']}\boldsymbol{A}' \tag{19}$$

$\boldsymbol{L}_{[(m-m')\times m']}$ indicates the submatrix of $\boldsymbol{L}$ in rows $m'+1 \sim m$, columns $1 \sim m'$. Once $m' > n$, all the $\boldsymbol{q}$s during n periods are corrected and $\boldsymbol{Q_c}$ is no longer updated. Therefore, the final $\boldsymbol{Q_c}$ can be obtained after a total of n updates, whose process is shown in Figure 8.

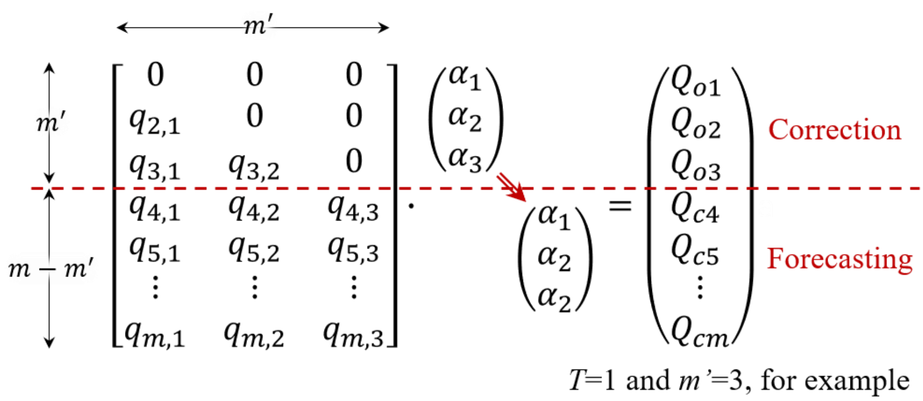

Figure 6. The correction and forecasting process of the URC-S by taking T=1 and m'=3 as an example.

3.4 Parameter calibration and verification

3.4.1 Forecast performance metrics

Several statistical metrics including the Nash–Sutcliffe coefficient (NS), the relative error of peak flow (RPF), the error of peak time (EPT) and the relative error of runoff depth (RRD) are employed to evaluate the forecast performance of the above two correction methods based on $\boldsymbol{Q_c}$ and $\boldsymbol{Q_o}$. The formulas for NS, RPF, EPT and RRD are given by following (*Yapo et al., 1996*).

$$NS = 1 - \sum_{j=1}^{m} [Q_c(j) - Q_o(j)]^2 / \sum_{j=1}^{m} \left[Q_o(j) - \overline{Q_o}\right]^2 \tag{20}$$

$$RPF = \text{mean}\left(\frac{\left|Q_{cf} - Q_{of}\right|}{Q_{of}} \times 100\%\right) \tag{21}$$

$$EPT = \text{mean}\left(\left|T_{cf} - T_{of}\right|\right) \tag{22}$$

$$RRD = \text{mean}\left(\frac{R_{cf} - R_{of}}{R_{of}} \times 100\%\right) \tag{23}$$

where Q_{cf} and Q_{of} are the peaks in the forecast and observed discharge sequences of each flood event, respectively; T_{cf} and T_{of} are the peak time in the forecast and observed discharge sequences of each flood event, respectively; R_{cf} and R_{of} respectively indicates the forecast and observed runoff of each flood event. In view of the fact that $\boldsymbol{\theta}$ is an unknown vector in the calculation of $\boldsymbol{Q_c}$, the above four evaluation indexes are all functions of $\boldsymbol{\theta}$ which is composed of different elements according to the hydrologic model and needs to be calibrated by optimization.

3.4.2 Xin'anjiang model

The Xin'anjiang model (XAJ) is employed as the hydrologic forecasting model in this paper. It is a widely used conceptual hydrological model with excellent performance in flood forecasting (*Li et al., 2013b; Si et al., 2015; Zhao, 1992; Jiang et al., 2014)*. XAJ consists of four modules: evapotranspiration, runoff generation, runoff separation and runoff concentration. The first two modules can be considered as a process for calculating runoff and soil moisture from rainfall and evapotranspiration, namely f_1 and g_1 in formula (2) and also the blue part in Figure 7 (hereinafter referred to as Process 1). The basic principle of this process is: when $P - E < 0$, soil moisture is calculated according to the 3-layer evapotranspiration module； when $P - E > 0$, runoff is calculated according to the theory of conceptual runoff generation under saturated condition, meaning all rainfall is stored in the soil until the soil moisture content reaches field capacity; thereafter, the net rainfall drains out in the form of runoff without further loss.

The main function of the last two modules of XAJ, namely runoff separation and runoff concentration and also the black part in Figure 7 (hereinafter referred to as Process 2), is calculating discharge from runoff. Therefore, the parameters, namely $\boldsymbol{\theta}$ in this paper, to be calibrated are those in the XAJ model. Please refer to the literature (*Zhao, 1995*) for the parameter value range and other detailed information of XAJ model.

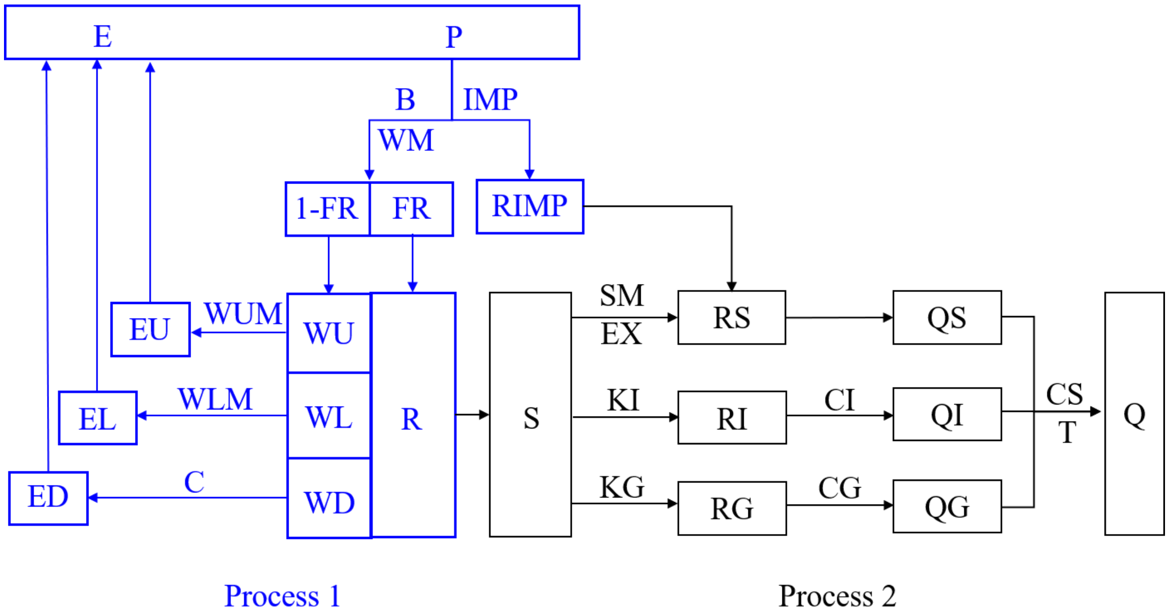

Figure 7. A structural diagram of XAJ

3.4.3 Calibration method

For model calibration, the Particle Swarm Optimization (PSO) method was employed. PSO is an evolutionary computing technique developed by *Kennedy and Eberhart (1995)*, derived from the simulation of a simplified social behavior model. PSO has previously been applied to the parameter calibration of hydrological models, and is considered to be able to efficiently search for the global optimal solution in the search-space. Since PSO is a single-objective optimization method, the information expressed by the NS, RPF, EPT and RRD performance criteria should be converted into a relative membership degree to obtain a single objective value to guide the optimization process. The relative membership degree (Shouyu, 1998, 2005; Shouyu & Yu, 2006) is computed as follows:

$$RMD = \frac{1}{1 + \frac{\sum_{k=1}^{4}[w_k(g_k - r_k)]^2}{\sum_{k=1}^{4}[w_k(r_k - b_k)]^2}} \tag{24}$$

where RMD indicates relative membership degree, such that larger RMD indicates better values for

$\boldsymbol{\theta}$; w_k is the weight of the kth optimization objective, and NS, RPF, EPT and RRD are considered equally important so that w_k=0.25; r_k is the kth optimization objective after normalization; g_k and b_k are the upper and lower bounds of r_k, respectively. Figure 8 briefly illustrates the structure of the parameter calibration process.

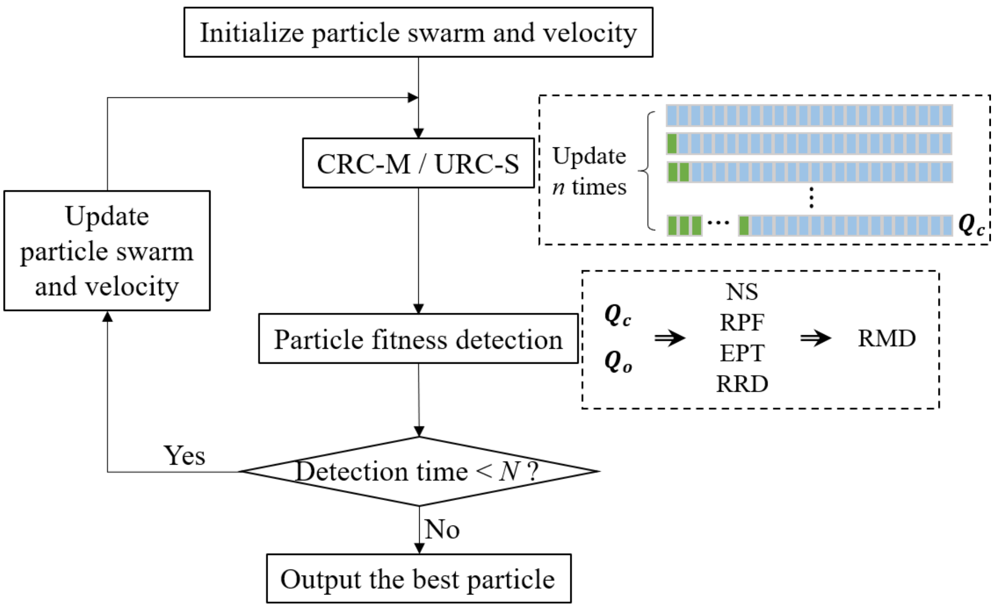

Figure 8. A structural diagram of parameter calibration

4. Results and discussion

4.1 Comparison between results of constrained and unconstrained runoff correction

To demonstrate the necessity of constraints when performing runoff correction based on SPPs, we compare the results of CRC-M and URC-S during calibration and evaluation stages. A combination of precipitation and evapotranspiration products forms the basic input to the conceptual hydrological model. For CRC-M, multiple (at least two) combinations should be input into the model, so as to not only force the hydrological model, but more importantly, form a runoff range to constrain the runoff correction. Since there are 4 kinds of precipitation estimates and 2 kinds of evapotranspiration estimates, we can input up to 8 combinations into the model. For URC-S, only a single combination of precipitation and evapotranspiration is needed as input to the model. Here, we decided to use 8 combinations to force CRC-M (namely h=8) and use combination the single combination of GSMaP-GN and GLDAS-CLM to force URC-S.

The performance metric values, including NS, RPF, EPT, RRD and RMD, obtained using CRC-M and URC-S for the calibration and evaluation stages are listed in Table 4. Of these five metrics reported, RMD provides a summary indicator. As indicated by RMD, whether in the calibration or the evaluation stage, the accuracy of flood forecasting based on CRC-M is higher than it based on URC-S. During calibration it is equal to 0.890 based on CRC-M and 0.764 based on URC-S, and during evaluation it is equal to 0.905 and 0.684 respectively based on CRC-M and URC-S. The consistency of accuracy in the calibration and evaluation periods indicates that the parameters determined by the CRC-M are representative, rather than results obtained by overfitting. As shown in Table 4, NS and RPF of URC-S (0.84 and 22.02%, respectively) are only a little bit worse than those of CRC-M (0.85 and 17.35%, respectively) in the calibration stage, while they are much worse (0.64 and 34.92%, respectively) than those of CRC-M (0.91 and 9.6%, respectively) in the evaluation stage. This implies that the performance of CRC-M is more reliable and stable than that of URC-S. In general, Table 4 shows the superiority of CRC-M; in other words, it supports the importance of constraints when performing runoff correction based on SPPs.

Table 4. Evaluation metrics of CRC-M and URC-S during calibration and validation stages.

Evaluation metrics	Calibration		Validation	
	CRC-M	URC-S	CRC-M	URC-S
NS	0.85	0.84	0.91	0.64
RPF (%)	17.35	22.02	9.6	34.92
EPT (d)	1.22	2.11	1.60	1.40
RRD (%)	13.67	18.74	16.25	11.39
RMD	0.890	0.764	0.905	0.684

Performance metrics comparing CRC-M and URC-S for 14 different flood events are listed in Table 5. The data in red are the RPFs when the flood peak is underestimated. Clearly, there are more red data in the list of CRC-M than in the list of URC-S (10 floods for CRC-M while 2 floods for URC-S). This implies that CRC-M tends to underestimates peak flows while URC-S tends to overestimate them. The possible reason is that since it is common for SPPs to underestimate extreme rainfall, using them to establish the runoff correction range will make it difficult to characterize extreme runoff. On the other hand, for URC-S, due to the lack of constraints in the process of correction, runoff can be infinitely magnified. Therefore, the peak flow estimates based on URC-S tends to be much larger than observation with RPF greater than 50% in 3 of the 14 flood events. Meanwhile, the grey-filled rows in Table 5 indicate the flood events for which all four of the evaluation metrics are inferior to the other method. We note that, based on this perspective, only two flood events based on CRC-M perform worse than those based on URC-S, while 6 flood events based on URC-S perform worse than those based on CRC-M.

Table 5. Evaluation metrics of CRC-M and URC-S during different flood events, where the data in red are the RPFs when the flood peak is underestimated and the grey-filled data are the flood events

whose all four evaluation metrics are inferior to the other method.

		CRC-M				URC-S			
	event	NS	RPF (%)	EPT (d)	RRD (%)	NS	RPF (%)	EPT (d)	RRD (%)
Calibration	1	0.79	-22.93	0	13.93	0.76	-28.36	-3	-25.04
	2	0.91	-25.65	-1	-25.70	0.94	9.18	-1	-12.49
	3	0.95	8.03	0	-9.19	0.94	16.32	-1	-1.32
	4	0.84	21.98	0	3.58	0.92	8.46	0	-19.94
	5	0.76	-22.11	2	4.46	0.92	1.87	-1	-7.93
	6	0.96	-0.09	-2	-10.30	0.52	27.09	4	26.97
	7	0.82	-15.30	-4	-28.02	0.91	9.66	0	-1.43
	8	0.76	-39.71	1	-25.14	0.01	75.69	1	63.74
	9	0.95	-0.34	-1	2.72	0.54	21.51	8	9.84
Validation	10	0.69	1.29	1	36.44	0.97	-10.24	0	-1.95
	11	0.93	-4.63	-1	0.45	0.91	28.63	-1	5.47
	12	0.73	19.06	-3	26.62	0.47	60.78	3	13.30
	13	0.92	-21.87	-1	-6.70	0.93	8.86	2	-1.23
	14	0.97	-1.14	-2	-11.04	0.43	66.12	1	34.99

Values for NS, absolute RPF, absolute ERT and absolute RRD values of CRC-M and URC-S for 14 flood events are shown in Figure 9. Based on NS and absolute RPF, CRC-M generally performs better than URC-S, especially in the evaluation period. Based on values of absolute ERT and absolute RRD, both CRC-M and URC-S show similar performance, except that URC-S does not perform well in some flood events (No.9 for absolute ERT and No.8 for absolute RRD). Based on this assessment, CRC-M tends to give better and more stable performance than URC-S, although with a tendency to underestimates peaks.

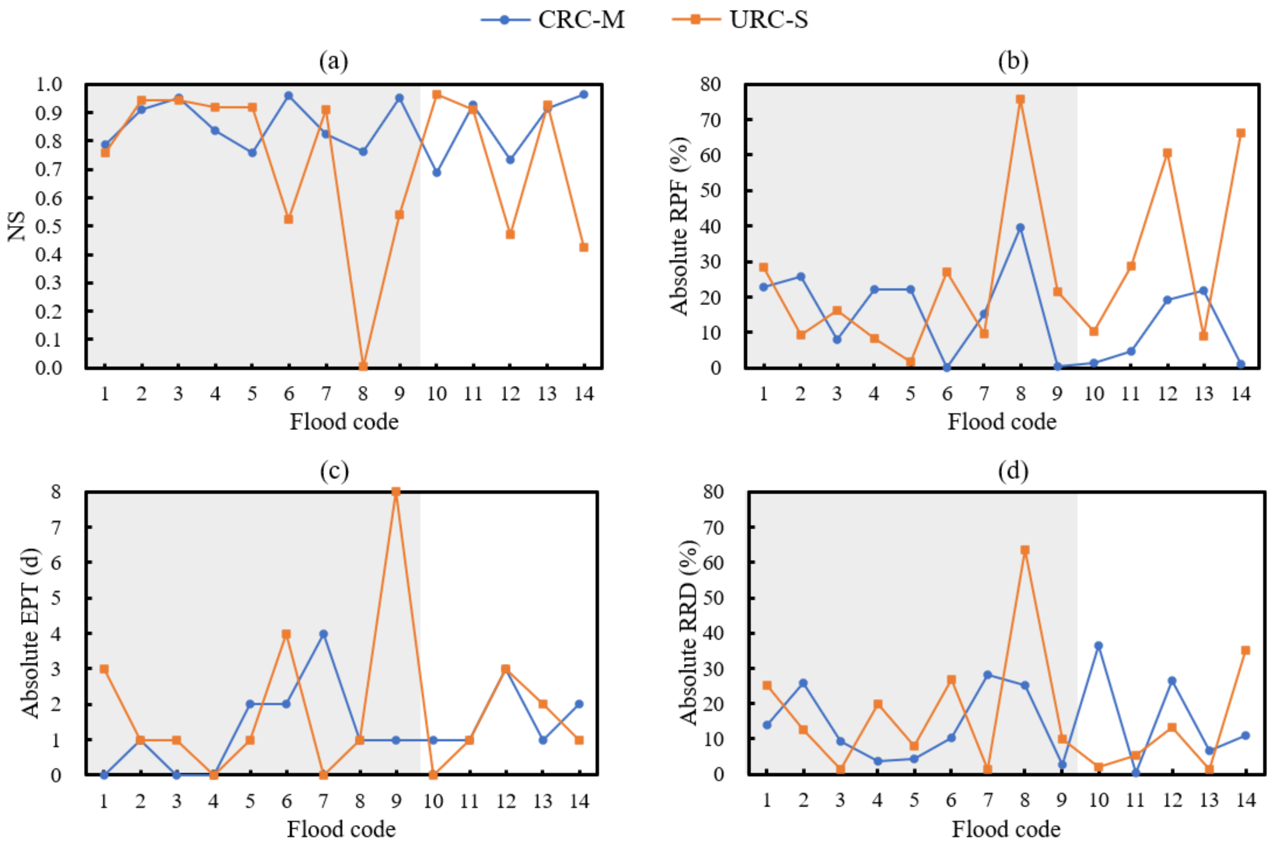

Figure 9. NS, absolute RPF, absolute ERT and absolute RRD values of CRC-M (in blue) and URC-S (in orange) for 14 flood events. The gray section indicates the calibration period and the white section indicates the evaluation period.

4.2 How to form reasonable constraints in CRC-M?

To figure out how to determine reasonable constraints for use in CRC-M, this section compares the flood forecast performance of different input schemes, in other words, of different constraints, based on up to four SPPs. The performance is evaluated using RMD. In addition, the calculation in this section adopts the parameters calibrated by CRC-M in section 4.1, due to their stable performance.

Since the constraints in CRC-M require at least two combinations to be input into the model, there are 11 input schemes to force CRC-M. ($\sum_{h=2}^{4} C_4^h = 11$, using up to 4 SPPs). As shown in Figure 10a, there are 11 input schemes with each evapotranspiration product. For any kind of evapotranspiration product, using all 4 SPPs to drive CRC-M, namely h=4, is the best performing scheme. With the increase of h, RMD of optimal schemes for each h increases approximately linearly. It can be inferred from formulas (9) and (10) that the larger h is, the wider the constraint range is. It might be further inferred that the wider the constraint range, the better the forecast performance. As it tends to the limit, the widest range would be the absence of constraints. However, this inference seems to contradict the conclusion in section 4.1 that CRC-M performs better URC-S which can be viewed as having no constraint range. This is because the excellent performance of CRC-M requires not only a wide range of constraints, but also SPPs with relatively high accuracy which can be seen from the schemes whose h are equal to 3 and 2.

When h=3, the optimal scheme is using IMERG-E, IMERG-L and GSMaP-GN. As can be seen in the Figure 10b, the difference between the unselected GSMaP-N and gauge observation is the largest, which means that use of GSMaP-N results in the worst performance. When h=2, the optimal scheme is using IMERG-L and GSMaP-GN. As can be seen in the Figure 10b, IMERG-E, IMERG-L and GSMaP-GN have similar differences with gauge. The largest pairwise difference of these three SPPs is the difference between IMERG-L and GSMaP-GN, since IMERG and GSMaP are two parallel retrieval algorithms for the GPM mission. As a result, they can form a relatively wider range. Therefore, relatively reliable SPPs and a wide constraint range are both very important for the performance of CRC-M.

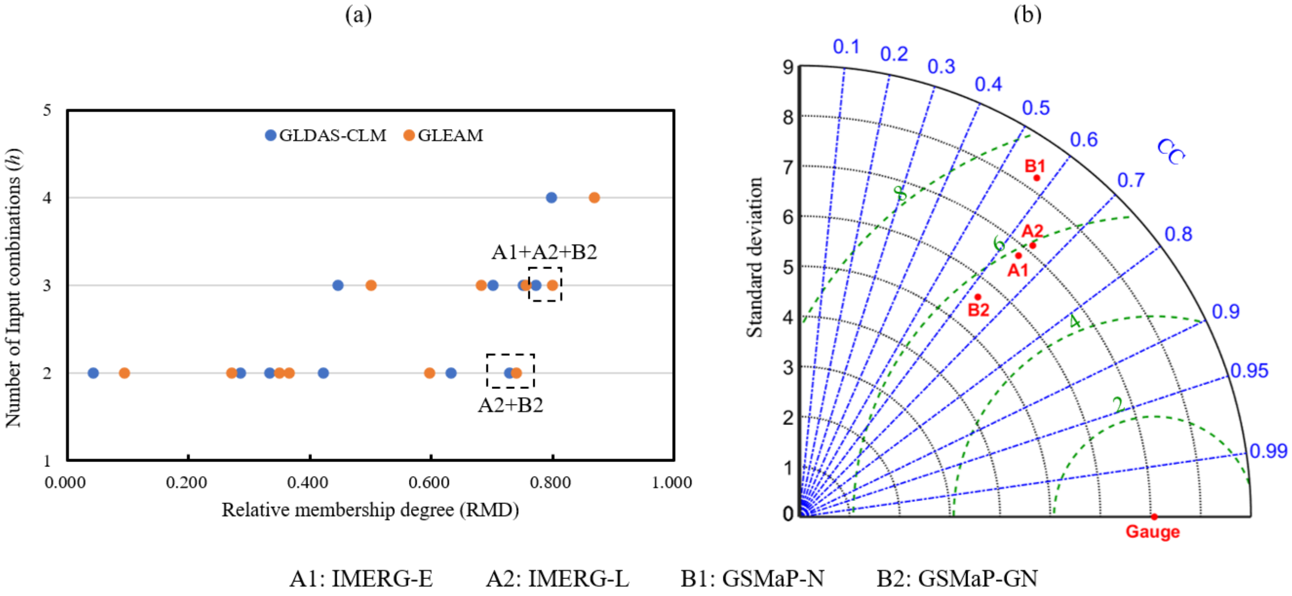

Figure 10. (a) The relationship between number of input combinations (*h*) and relative membership degree (RMD). The blue and orange points indicate input schemes using GLDAS-CLM and GLEAM, respectively. The point in the dotted rectangle is the optimal schemes for the same *h*. (b) Taylor diagram of the four SPPs, where the standard deviation, correlation coefficient (CC) and root mean squared error (RMSE) evaluate the accuracy of the product data at precipitation gauges in the Nenjiang Basin at the similar latitude as Xiaoergou Basin.

4.3 Does more information result in better performance?

To see if providing additional information necessarily results in improved forecasting performance, evapotranspiration products were added to generate different constraint schemes; overall the performance of 8 combinations is compared. Consistent with section 4.2, we also use RMD to evaluate the performance and use the parameters calibrated by CRC-M in section 4.1.

Since there are 4 kinds of precipitation estimates and 2 kinds of evapotranspiration estimates, for an input scheme, we can input 2 ~ 8 combinations into the model, and each combination is displayed in Table 6. Therefore, there are 247 input schemes to force CRC-M ($\sum_{h=2}^{8} C_8^h = 247$). From the analysis in Section 4.2, we have determined that the width of the constraint range is one of the important factors affecting the forecast performance. To quantify the width of constraint range in an input scheme, we used the mean absolute difference (MAD) calculated by $MAD = \frac{1}{m}\sum_{i=1}^{m}\left|\left(\overline{PE_i} - \underline{PE_i}\right)\right|$. PE_i indicates subtraction between precipitation and evapotranspiration at period i in an input scheme, and there are 2 ~ 8 combinations of precipitation and evapotranspiration. $\overline{PE_i}$ and $\underline{PE_i}$ are the largest and smallest values among PE_is of the combinations in an input scheme, respectively.

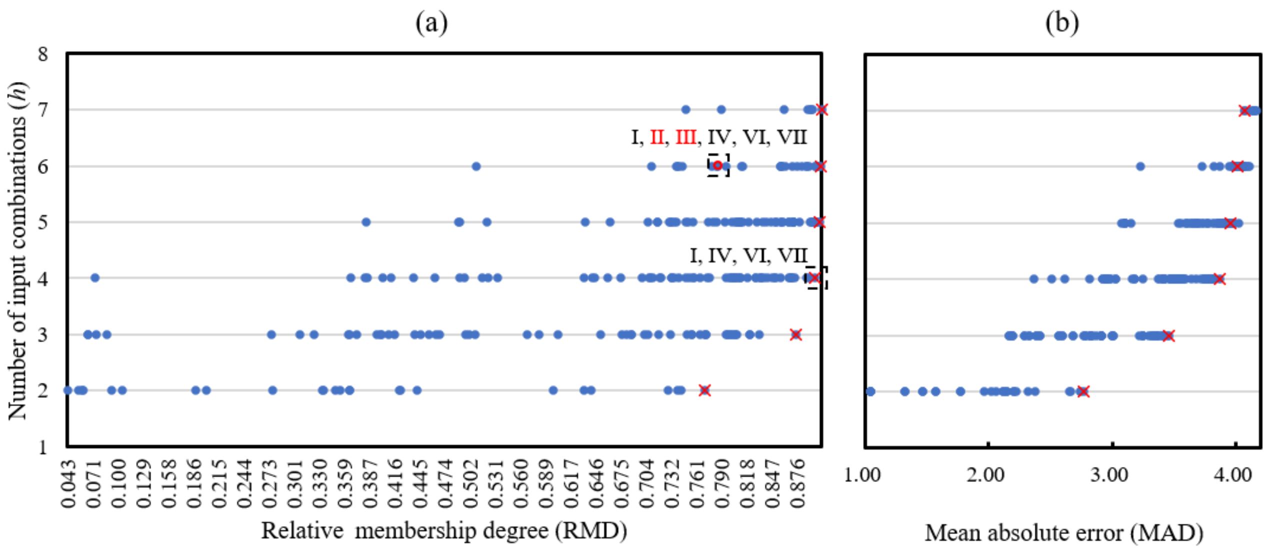

Figure 11. (a) Relative membership degree (RMD) for flood forecasting of different input schemes with various h, where h indicates the number of input combinations. The red crosses indicate the largest RMD scheme for each h. (b) Mean absolute difference (MAD) of each input scheme. The red crosses indicate the MAD of the optimal scheme for each h.

As shown in Figure 11a, in general there is a tendency that more input combinations result in larger RMD which means better performance of flood forecasting, especially for the optimal scheme for each h. But unlike Figure 10(a), the tendency is not linear. With the increase of h, RMD of optimal scheme for each h increases more and more slowly. When h is equal to 4, 5, 6 and 7, RMD of optimal scheme is 0.897, 0.902, 0.904 and 0.904, as shown in Table 7. For each additional combination, RMD of optimal scheme increases by a maximum of 0.005. However, when *h* is equal to 2, 3 and 4, for each additional combination RMD of optimal scheme increases 0.105 and 0.022, respectively. The optimal scheme for *h*=4, which can be approximated as an inflection point, is formed by I, IV, VI and VII. In more detail, as listed in table 6, it is formed by IMERG-E+GLDAS-CLM, GSMaP-GN+GLDAS-CLM, IMERG-L+GLEAM and GSMaP-N+GLEAM, which means that the scheme has already chosen all 4 SPPs and their own relatively appropriate evapotranspiration products. The combinations not included in the scheme are II, III, V and VIII, which have similar information to VI, VII, I and IV respectively, so they constitute redundant information.

Meanwhile, as shown in the Figure 11(b), with the increase of *h*, the tendency of MAD is similar to that of RMD of each scheme, especially for the optimal scheme for each *h*, when *h* is larger than 4, its MAD grows slowly, and vice versa. To some extent, this indicates that the forecasting performance is controlled by the width of constrained range, which is consistent with the conclusion in 4.2. As shown in the Figure 11(b), when $h \leqslant 4$, MAD of each optimal scheme is the largest among the scheme with the same *h*. While *h*>4, it is no longer the maximum within the group with the same *h*. Even when *h* is equal to 7, it is the smallest in the group. So, this suggests that is it no longer possible to perform better as the width of constrained range increases. A wider width would interfere with the forecast.

An example shown in Figure 11(a) can illustrate this further, in which two points circled with two dotted rectangles respectively represent the optimal scheme for *h*=4 and a common scheme for *h*=6. Input combinations of the latter scheme (I, II, III, IV, VI, VII) include all the input combinations of the former scheme (I, IV, VI, VII). However, the RMD of the latter scheme (0.785) is much lower than that of the former (0.897), which indicates that the extra II and III have interfered with the original performance and reduced the forecasting accuracy.

Table 6. Each combination of precipitation and evapotranspiration and its tag

Tag	Precipitation		Evapotranspiration
I	IMERG	Early	GLDAS_CLM
II		Late	
III	GSMaP	NRT	
IV		Gauge-NRT	
V	IMERG	Early	GLEAM
VI		Late	
VII	GsMaP	NRT	
VIII		Gauge -NRT	

Table 7. The composition and RMD of the optimal schemes for the same *h*.

h	The largest RMD	Scheme
2	0.770	（IV, VI）
3	0.875	（IV, V, VI）
4	0.897	（I, IV, VI, VII）
5	0.902	（I, IV, VI, VII, VIII）
6	0.904	（I, III, IV, VI, VII, VIII）
7	0.904	（I, III, IV, V, VI, VII, VIII）

4.4 Test using multiple basins of different sizes and climates

To further test stability and reliability of the CRC-M, we tested the method on 7 basins with a variety of sizes and climates (including the aforementioned Xiaoergou Basin). The area of these basins (Figure

12) is approximately evenly distributed between 2000 and 32000 km^2, of which the smallest is the Biliuhe basin (2061 km^2) and the largest is the Kumotun basin (32087 km^2). The shapes and locations of these basins are shown in Figure 13. The 5 basins at the highest latitudes (②Geni, ③Kehou, ④ Jiagedaqi, ⑤Xiaoergou and ⑦Kumotun) are located in the upper reaches of the Nenjiang River, and have similar climatic conditions to the Xiaoergou Basin (see Section 2.1), with average annual rainfall of about 400-500 mm and average annual temperature of approximately -1.2℃. In addition, the Biliuhe basin (①), which extends from latitude 39.55 °N to 40.35 °N and longitude 122.31 °E to 122.89 °E, experiences an average annual rainfall about 700-800mm and average annual temperature of approximately 10.6℃. The lowest latitude Chishuihe basin (⑥) extends from latitude 27.20 °N to 28.83 °N and longitude 104.72 °E to 106.99 °E，and experiences an average annual rainfall of about 1000-1100mm and average annual temperature of approximately 17.6℃of the. The basins experience different frequencies of flooding, net rainfall amounts, and flood peak magnitudes. All of the flood events data used in this study are at the daily time step.

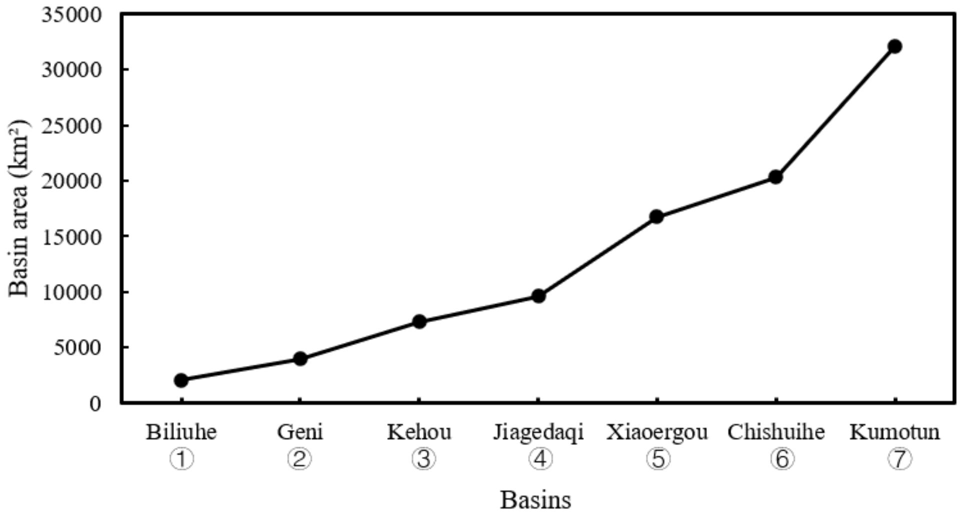

Figure 12. Areas of test basins.

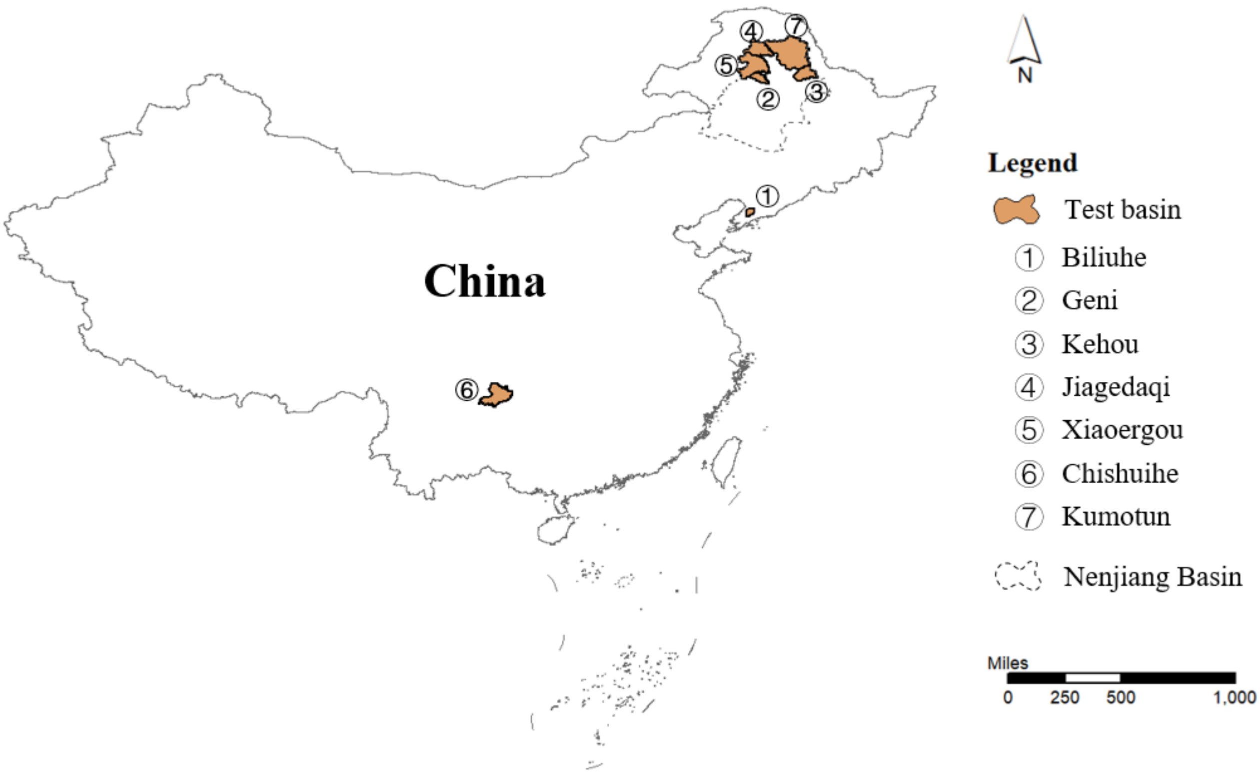

Figure 13. Shapes and locations of test basins, where the number is sorted by basin area.

The results obtained by application of the CRC-M to these 7 basins are shown in Tables 8 and 9, and in Figure 14. The accuracy of flood forecasts based on the CRC-M method in these 7 basins is high. The RMD is above 0.810 (with a maximum of 0.988) and the NS is above 0.73 (with a maximum of 0.92), indicating that the method provides reliable results across a wide variety of basin conditions. Further, the similarity of the calibration and evaluation period metrics across basins indicates that the CRC-M method provides stable performance, and should therefore be applicable for flood forecasting in other ungauged basins.

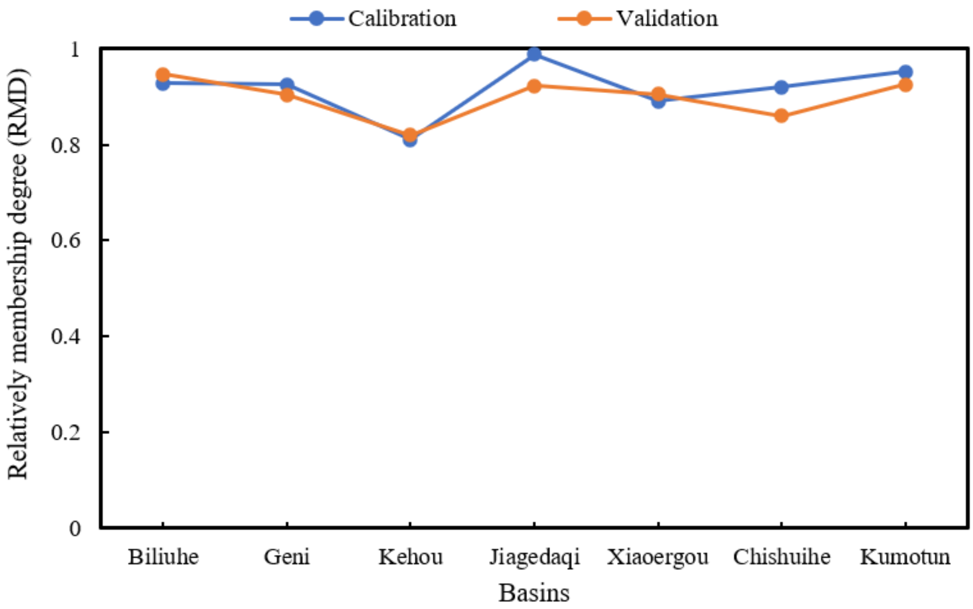

Figure 14. Relatively membership degree of 7 basins during calibration (in bule) and validation (in orange) based on CRC-M

Table 8. Performance metric values obtained for the 7 basins during the calibration period.

Basins	NS	RPF (%)	EPT (d)	RRD (%)	RMD
Biliuhe	0.80	16.66	0.60	5.92	0.928
Geni	0.84	16.74	1.25	8.16	0.925
Kehou	0.73	13.48	1.88	14.34	0.810
Jiagedaqi	0.83	3.54	0.40	2.90	0.988
Xiaoergou	0.85	17.35	1.22	13.67	0.890
Chishuihe	0.79	13.96	0.25	16.43	0.919
Kumotun	0.92	14.55	1.00	5.52	0.952

Table 9. Performance metric values obtained for the 7 basins during the evaluation period.

Basins	NS	RPF (%)	EPT (d)	RRD (%)	RMD
① Biliuhe	0.92	8.65	0.40	6.97	0.947
② Geni	0.85	17.02	1.50	10.95	0.904

③ Kehou	0.82	10.91	1.75	23.17	0.820
④ Jiagedaqi	0.84	14.51	0.60	16.92	0.923
⑤ Xiaoergou	0.91	9.60	1.60	16.25	0.905
⑥ Chishuihe	0.81	20.22	0.50	20.78	0.859
⑦ Kumotun	0.70	8.32	1.40	4.46	0.925

5. Conclusions

This paper proposes and evaluates a constrained runoff correction method (CRC-M) based on the use of multi-source satellite-based products to facilitate flood forecasting. The basic principle is to correct runoff using the real-time updated outlet discharge with constraints calculated by SPPs. To illustrate this, four near-real-time SPPs (IMERG-E, IMERG-L, GSMaP-N and GSMaP-GN) were employed to calculate the constraint range to be applied to runoff correction. The method was first tested in an ungauged basin and its performance discussed in terms of the importance of constraint range and characteristics of good constraints. Then, the method was applied to 7 basins with different sizes and hydroclimatic conditions. The primary conclusions can be summarized as follows:

1. Application of CRC-M to the study site resulted in NS forecasting performance of 0.85 and 0.91 during calibration and evaluation respectively. CRC-M resulted in better and more stable performance compared to a control test (URC-S), thereby demonstrating the importance of using a constraint range. While CRC-M underestimates the peak flow slightly, the unconstrained URC-S approach tends to overestimate it severely. The results indicate that use of a runoff constraint range determined using different SPPs results in more stable and accurate runoff correction.

2. We find that two characteristics result in better forecasting performance; wider constraint ranges and relatively reliable SPPs. For any kind of evapotranspiration product, using all 4 SPPs to drive CRC-M (meaning the widest constraint range) is the best performing scheme. Meanwhile, when *h* is limited, SPPs with higher accuracy are preferentially selected.

3. Adding more information, in the form of evapotranspiration products, allows us to generate additional schemes. With increasing of *h*, we see progressively marginal improvements in RMD and MAD of the optimal scheme, especially when *h* is larger than 4. The optimal scheme for *h*=4, which includes all of the satellite products mentioned in this paper, can be approximated as an inflection point. After the inflection point, the increased information tends to be redundant (or even detrimental), and only improves the forecast performance slightly (or can even cause the performance to deteriorate).

4. Application of CRC-M to a variety of basins having different sizes and hydroclimatic conditions resulted in RMD values above 0.8 and consistently similar calibration and evaluation period performance, suggesting that the CRC-M method can reliably be used for flood forecasting in a variety of other ungauged basins.

Overall, the CRC-M method performs well in ungauged basins even though the accuracy of the near-real-time SPPs used is not very high. Its performance is improved under wider constraint conditions and with more accurate SPPs, as long as redundancy and interference are minimal. The data and computer codes used in this study can be obtained from the corresponding author upon request. As always, we invite discussion and collaboration on this and related aspects of modeling and hydrological forecasting.

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China (nos. 51709033, 51909010) and the Fundamental Research Funds for the Central Universities (no. DUT20RC(3)019). The precipitation data can be obtained from https://disc.gsfc.nasa.gov and http://www.jaxa.jp. The evapotranspiration data can be obtained from https://www.gleam.eu and https://disc.gsfc.nasa.gov. Discharge data and model program used in this study are available upon request from the corresponding author.

References

Behrangi, A., Khakbaz, B., Jaw, T. C., AghaKouchak, A., Hsu, K., & Sorooshian, S. (2011). Hydrologic evaluation of satellite precipitation products over a mid-size basin. *Journal of Hydrology*, *397*(3–4), 225–237. https://doi.org/10.1016/j.jhydrol.2010.11.043

Bitew, M. M., & Gebremichael, M. (2011). Evaluation of satellite rainfall products through hydrologic simulation in a fully distributed hydrologic model. *Water Resources Research*, *47*(6), 1–11. https://doi.org/10.1029/2010WR009917

Ciabatta, L., Brocca, L., Massari, C., Moramarco, T., Gabellani, S., Puca, S., & Wagner, W. (2016). Rainfall-runoff modelling by using SM2RAIN-derived and state-of-the-art satellite rainfall products over Italy. *International Journal of Applied Earth Observation and Geoinformation*, *48*, 163–173. https://doi.org/10.1016/j.jag.2015.10.004

Deng, P., Zhang, M., Bing, J., Jia, J., & Zhang, D. (2019). Evaluation of the GSMaP_Gauge products using rain gauge observations and SWAT model in the Upper Hanjiang River Basin. *Atmospheric Research*, *219*(December 2018), 153–165. https://doi.org/10.1016/j.atmosres.2018.12.032

Duan, Z., Liu, J., Tuo, Y., Chiogna, G., & Disse, M. (2016). Evaluation of eight high spatial resolution gridded precipitation products in Adige Basin (Italy) at multiple temporal and spatial scales. *Science of The Total Environment*, *573*, 1536–1553. https://doi.org/10.1016/j.scitotenv.2016.08.213

Harris, A., Rahman, S., Hossain, F., Yarborough, L., Bagtzoglou, A., & Easson, G. (2007). Satellite-based Flood Modeling Using TRMM-based Rainfall Products. *Sensors*, *7*(12), 3416–3427. https://doi.org/10.3390/s7123416

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. d., Kojima, M., et al. (2014). The Global Precipitation Measurement Mission. *Bulletin of the American Meteorological Society*, *95*(5), 701–722. https://doi.org/10.1175/BAMS-D-13-00164.1

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., et al. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *Journal of Hydrometeorology*, *8*(1), 38–55. https://doi.org/10.1175/JHM560.1

Huffman, G. J., Gsfc, N., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., et al. (2018). Algorithm Theoretical Basis Document (ATBD) NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG). *Algorithm Theoretical Basis Doc., Version 5.2*, (February). Retrieved from https://pmm.nasa.gov/sites/default/files/ document_files/IMERG_ATBD_V5.2_0.pd

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Stocker, E. F., & Tan, J. (2018). V05 IMERG Final Run Release Notes. *NASA Goddard Earth Sciences Data and Information Services Center: Greenbelt, MD, USA*, (February), 1–8. Retrieved from https://pmm.nasa.gov/sites/default/files/document_files/IMERG_FinalRun_V05_ release_notes-rev3.pdf

Huffman, G. J., Bolvin, D. T., & Nelkin, E. J. (2019). Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation. *NASA/GSFC Code*. Retrieved from https://docserver.gesdisc.eosdis.nasa.gov/public/project/GPM/IMERG_doc.06.pdf

Jiang, L., & Bauer-Gottwein, P. (2019). How do GPM IMERG precipitation estimates perform as hydrological model forcing? Evaluation for 300 catchments across Mainland China. *Journal of Hydrology*, *572*(March), 486–500. https://doi.org/10.1016/j.jhydrol.2019.03.042

Jiang, S., Ren, L., Hong, Y., Yong, B., Yang, X., Yuan, F., & Ma, M. (2012). Comprehensive evaluation of multi-satellite precipitation products with a dense rain gauge network and optimally merging their simulated hydrological flows using the Bayesian model averaging method. *Journal of Hydrology*, *452–453*, 213–225. https://doi.org/10.1016/j.jhydrol.2012.05.055

Jiang, S., Ren, L., Hong, Y., Yang, X., Ma, M., Zhang, Y., & Yuan, F. (2014). Improvement of Multi-Satellite Real-Time Precipitation Products for Ensemble Streamflow Simulation in a Middle Latitude Basin in South China. *Water Resources Management*, *28*(8), 2259–2278. https://doi.org/10.1007/s11269-014-0612-4

Jiang, S., Ren, L., Xu, C.-Y., Yong, B., Yuan, F., Liu, Y., et al. (2018). Statistical and hydrological evaluation of the latest Integrated Multi-satellitE Retrievals for GPM (IMERG) over a midlatitude humid basin in South China.

Atmospheric Research, *214*(June), 418–429. https://doi.org/10.1016/j.atmosres.2018.08.021

Joyce, R. J., Janowiak, J. E., Arkin, P. A., & Xie, P. (2004). CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, *5*(3), 487–503. https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2

Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks* (Vol. 4, pp. 1942–1948). IEEE. https://doi.org/10.1109/ICNN.1995.488968

Kubota, T., Shige, S., Hashizume, H., Aonashi, K., Takahashi, N., Seto, S., et al. (2007). Global Precipitation Map Using Satellite-Borne Microwave Radiometers by the GSMaP Project: Production and Validation. *IEEE Transactions on Geoscience and Remote Sensing*, *45*(7), 2259–2275. https://doi.org/10.1109/TGRS.2007.895337

Li, N., Tang, G., Zhao, P., Hong, Y., Gou, Y., & Yang, K. (2017). Statistical assessment and hydrological utility of the latest multi-satellite precipitation analysis IMERG in Ganjiang River basin. *Atmospheric Research*, *183*, 212–223. https://doi.org/10.1016/j.atmosres.2016.07.020

Liu, Y., Weerts, A. H., Clark, M., Hendricks Franssen, H.-J., Kumar, S., Moradkhani, H., et al. (2012). Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities. *Hydrology and Earth System Sciences*, *16*(10), 3863–3887. https://doi.org/10.5194/hess-16-3863-2012

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., et al. (2017). GLEAM v3: satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, *10*(5), 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017

Massari, C., Camici, S., Ciabatta, L., & Brocca, L. (2018). Exploiting satellite-based surface soil moisture for flood forecasting in the Mediterranean area: State update versus rainfall correction. *Remote Sensing*, *10*(2), 292. https://doi.org/10.3390/rs10020292

Massari, C., Maggioni, V., Barbetta, S., Brocca, L., Ciabatta, L., Camici, S., et al. (2019). Complementing near-real time satellite rainfall products with satellite soil moisture-derived rainfall through a Bayesian Inversion approach. *Journal of Hydrology*, *573*(March), 341–351. https://doi.org/10.1016/j.jhydrol.2019.03.038

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., & Dolman, A. J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences*, *15*(2), 453–469. https://doi.org/10.5194/hess-15-453-2011

Omranian, E., & Sharif, H. O. (2018). Evaluation of the Global Precipitation Measurement (GPM) Satellite Rainfall Products over the Lower Colorado River Basin, Texas. *JAWRA Journal of the American Water Resources Association*, *54*(4), 882–898. https://doi.org/10.1111/1752-1688.12610

Pan, M., Li, H., & Wood, E. (2010). Assessing the skill of satellite-based precipitation estimates in hydrologic applications. *Water Resources Research*, *46*(9), 1–10. https://doi.org/10.1029/2009WR008290

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., et al. (2004). The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, *85*(3), 381–394. https://doi.org/10.1175/BAMS-85-3-381

Shouyu, C. (1998). Theory and application engineering fuzzy sets. *National Defense Industry Press, Beijing (in Chinese)*.

Shouyu, C. (2005). Theories and methods of variable fuzzy sets in water resources and flood control system. *Dalian University of Technology Press, Dalian (in Chinese)*.

Shouyu, C., & Yu, G. (2006). Variable fuzzy sets and its application in comprehensive risk evaluation for flood-control engineering system. In *Fuzzy Optimization and Decision Making* (Vol. 5, pp. 153–162). https://doi.org/10.1007/s10700-006-7333-y

Sirisena, T. A. J. G., Maskey, S., Ranasinghe, R., & Babel, M. S. (2018). Effects of different precipitation inputs on streamflow simulation in the Irrawaddy River Basin, Myanmar. *Journal of Hydrology: Regional Studies*, *19*(April), 265–278. https://doi.org/10.1016/j.ejrh.2018.10.005

Sorooshian, S., Hsu, K. L., Gao, X., Gupta, H. V, Imam, B., & Braithwaite, D. (2000). Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bulletin of the American Meteorological Society*, *81*(9), 2035–2046. https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2

Su, J., Lü, H., Zhu, Y., Cui, Y., & Wang, X. (2019). Evaluating the hydrological utility of latest IMERG products over the Upper Huaihe River Basin, China. *Atmospheric Research*, *225*(January), 17–29. https://doi.org/10.1016/j.atmosres.2019.03.025

Sun, R., Yuan, H., & Yang, Y. (2018). Using multiple satellite-gauge merged precipitation products ensemble for hydrologic uncertainty analysis over the Huaihe River basin. *Journal of Hydrology*, *566*(May), 406–420. https://doi.org/10.1016/j.jhydrol.2018.09.024

Tan, J., Petersen, W. A., Kirstetter, P.-E., & Tian, Y. (2017). Performance of IMERG as a Function of Spatiotemporal

Scale. *Journal of Hydrometeorology*, *18*(2), 307–319. https://doi.org/10.1175/JHM-D-16-0174.1

Wei Si, Weimin Bao, and H. V. G. (2015). Updating real-time flood forecasts via the dynamic system response curve method. *Water Resources Research*, 2498–2514. https://doi.org/10.1002/2015WR017200.A

Weimin, B., Wei, S., & Simin, Q. (2014). Flow Updating in Real-Time Flood Forecasting Based on Runoff Correction by a Dynamic System Response Curve. *Journal of Hydrologic Engineering*, *19*(4), 747–756. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000848

Yapo, P. O., Gupta, H. V., & Sorooshian, S. (1996). Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data. *Journal of Hydrology*, *181*(1–4), 23–48. https://doi.org/10.1016/0022-1694(95)02918-4

Yuan, F., Zhang, L., Win, K., Ren, L., Zhao, C., Zhu, Y., et al. (2017). Assessment of GPM and TRMM Multi-Satellite Precipitation Products in Streamflow Simulations in a Data-Sparse Mountainous Watershed in Myanmar. *Remote Sensing*, *9*(3), 302. https://doi.org/10.3390/rs9030302

Yuan, F., Wang, B., Shi, C., Cui, W., Zhao, C., Liu, Y., et al. (2018). Evaluation of hydrological utility of IMERG Final run V05 and TMPA 3B42V7 satellite precipitation products in the Yellow River source region, China. *Journal of Hydrology*, *567*(May), 696–711. https://doi.org/10.1016/j.jhydrol.2018.06.045

Yuan, F., Zhang, L., Soe, K., Ren, L., Zhao, C., Zhu, Y., et al. (2019). Applications of TRMM- and GPM-Era Multiple-Satellite Precipitation Products for Flood Simulations at Sub-Daily Scales in a Sparsely Gauged Watershed in Myanmar. *Remote Sensing*, *11*(2), 140. https://doi.org/10.3390/rs11020140

Zhang, A., Xiao, L., Min, C., Chen, S., Kulie, M., Huang, C., & Liang, Z. (2019). Evaluation of latest GPM-Era high-resolution satellite precipitation products during the May 2017 Guangdong extreme rainfall event. *Atmospheric Research*, *216*(May 2018), 76–85. https://doi.org/10.1016/j.atmosres.2018.09.018

Zhu, Z., Yong, B., Ke, L., Wang, G., Ren, L., & Chen, X. (2018). Tracing the Error Sources of Global Satellite Mapping of Precipitation for GPM (GPM-GSMaP) Over the Tibetan Plateau, China. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(7), 2181–2191. https://doi.org/10.1109/JSTARS.2018.2825336

Table captions:

Table 1. Overview of the SPPs used in this study

Table 2. Overview of the evapotranspiration (ET) products used in this study

Table 3. Distributions of the width of the ranges formed by different satellite-based products for a given period

Table 4. Evaluation metrics of CRC-M and URC-S during calibration and validation stages.

Table 5. Evaluation metrics of CRC-M and URC-S during different flood events, where the data in red are the RPFs when the flood peak is underestimated and the grey-filled data are the flood events whose all four evaluation metrics are inferior to the other method.

Table 6. Each combination of precipitation and evapotranspiration and its tag

Table 7. The composition and RMD of the optimal schemes for the same *h*.

Table 8. Evaluation metrics obtained for 7 basins during calibration period.

Table 9. Evaluation metrics obtained for 7 basins during validation period.

Figure captions:

Figure 1. Geographical location, topography and hydrologic station of the Xiaoergou Basin

Figure 2. Net rainfall (a) and flood peaks (b) of 14 flood events of the Xiaoergou Basin

Figure 3. A high-level systems diagram for a hydrological model; *P*(*i*), *E*(*i*) and *X*(*i*) indicate the areal mean rainfall, the areal mean evapotranspiration and the initial state of basin at the *i*th period, and *Q*(*j*) indicates the computed discharge output of the hydrological system at the *j*th period and also the first response to *P*(*i*), *E*(*i*) and *X*(*i*) of the basin outlet. Therefore, the lag time is *j*-*i*.

Figure 4. Schematic diagram of formula (2)~(8)

Figure 5. The correction and forecasting process of the CRC-M by taking *T*=1 and *m'*=3 as an example, where $\Delta q_{j,i} = \overline{q}_{j,i} - \underline{q}_{j,i}$.

Figure 6. The correction and forecasting process of the URC-S by taking *T*=1 and *m'*=3 as an

example.

Figure 7. A structural diagram of XAJ

Figure 8. A structural diagram of parameter calibration

Figure 9. NS, absolute RPF, absolute ERT and absolute RRD values of CRC-M (in blue) and URC-S (in orange) for 14 flood events. The gray section indicates the calibration period and the white section indicates the evaluation period.

Figure 10. (a) The relationship between number of input combinations (*h*) and relative membership degree (RMD). The blue and orange points indicate input schemes using GLDAS-CLM and GLEAM, respectively. The point in the dotted rectangle is the optimal schemes for the same *h*. (b) Taylor diagram of the four SPPs, where the standard deviation, correlation coefficient (CC) and root mean squared error (RMSE) evaluate the accuracy of the product data at precipitation gauges in the Nenjiang Basin at the similar latitude as Xiaoergou Basin.

Figure 11. (a) Relative membership degree (RMD) for flood forecasting of different input schemes with various *h*, where *h* indicates the number of input combinations. The red crosses indicate the largest RMD scheme for each *h*. (b) Mean absolute difference (MAD) of each input scheme. The red crosses indicate the MAD of the optimal scheme for each *h*.

Figure 12. Areas of test basins.

Figure 13. Shapes and locations of test basins, where the number is sorted by basin area.

Figure 14. Relatively membership degree of 7 basins during calibration (in bule) and validation (in orange) based on CRC-M

Table 1. Overview of the SPPs used in this study

Product		Corrected	Spatial Resolution	Developer	Start Time	Latency
IMERG	Early	No	0.1°×0.1°	NASA	June 2000	4 hours
IMERG	Late	Yes	0.1°×0.1°	NASA	June 2000	12 hours
GSMaP	NRT	No	0.1°×0.1°	JAXA	March 2000	4 hours
GSMaP	NRT-gauge	Yes	0.1°×0.1°	JAXA	April 2000	4 hours

Table 2. Overview of the evapotranspiration (ET) products used in this study

Data Sets	Category	Scheme	Spatial Resolution	Start Time
GLDAS	LSM	Penman-Monteith	1°×1°	January 1979
GLEAM	Diagnostic	Priestley-Taylor	0.25°× 0.25°	January 1980

Table 3. Distributions of the width of the ranges formed by different satellite-based products for a given period

Variable	Range [min, max] (mm)	Interquartile Range $[\frac{1}{4}\text{min}, \frac{3}{4}\text{max}]$ (mm)	Median (mm)
Precipitation (P)	[0, 46.57]	[0.17, 3.54]	1.17
Evapotranspiration (ET)	[0, 5.86]	[0.41, 1.42]	0.97
P-ET	[0.04, 47.89]	[1.26, 5.10]	2.25

Table 4. Evaluation metrics of CRC-M and URC-S during calibration and validation stages.

Evaluation metrics	Calibration		Validation	
	CRC-M	URC-S	CRC-M	URC-S
NS	0.85	0.84	0.91	0.64
RPF (%)	17.35	22.02	9.6	34.92
EPT (d)	1.22	2.11	1.60	1.40
RRD (%)	13.67	18.74	16.25	11.39
RMD	0.890	0.764	0.905	0.684

Table 5. Evaluation metrics of CRC-M and URC-S during different flood events, where the data in red are the RPFs when the flood peak is underestimated and the grey-filled data are the flood events whose all four evaluation metrics are inferior to the other method.

		CRC-M				URC-S			
	event	NS	RPF (%)	EPT (d)	RRD (%)	NS	RPF (%)	EPT (d)	RRD (%)
Calibration	1	0.79	-22.93	0	13.93	0.76	-28.36	-3	-25.04
	2	0.91	-25.65	-1	-25.70	0.94	9.18	-1	-12.49
	3	0.95	8.03	0	-9.19	0.94	16.32	-1	-1.32
	4	0.84	21.98	0	3.58	0.92	8.46	0	-19.94
	5	0.76	-22.11	2	4.46	0.92	1.87	-1	-7.93
	6	0.96	-0.09	-2	-10.30	0.52	27.09	4	26.97
	7	0.82	-15.30	-4	-28.02	0.91	9.66	0	-1.43
	8	0.76	-39.71	1	-25.14	0.01	75.69	1	63.74
	9	0.95	-0.34	-1	2.72	0.54	21.51	8	9.84
Validation	10	0.69	1.29	1	36.44	0.97	-10.24	0	-1.95
	11	0.93	-4.63	-1	0.45	0.91	28.63	-1	5.47
	12	0.73	19.06	-3	26.62	0.47	60.78	3	13.30
	13	0.92	-21.87	-1	-6.70	0.93	8.86	2	-1.23
	14	0.97	-1.14	-2	-11.04	0.43	66.12	1	34.99

Table 6. Each combination of precipitation and evapotranspiration and its tag

Tag	Precipitation		Evapotranspiration
I	IMERG	Early	GLDAS_CLM
II		Late	
III	GSMaP	NRT	
IV		Gauge-NRT	
V	IMERG	Early	GLEAM
VI		Late	
VII	GsMaP	NRT	
VIII		Gauge -NRT	

Table 7. The composition and RMD of the optimal schemes for the same *h*.

h	The largest RMD	Scheme
2	0.770	（IV, VI）
3	0.875	（IV, V, VI）
4	0.897	（I, IV, VI, VII）
5	0.902	（I, IV, VI, VII, VIII）
6	0.904	（I, III, IV, VI, VII, VIII）
7	0.904	（I, III, IV, V, VI, VII, VIII）

Table 8. Evaluation metrics obtained for 7 basins during calibration period.

Basins	NS	RPF (%)	EPT (d)	RRD (%)	RMD
Biliuhe	0.80	16.66	0.60	5.92	0.928
Geni	0.84	16.74	1.25	8.16	0.925
Kehou	0.73	13.48	1.88	14.34	0.810
Jiagedaqi	0.83	3.54	0.40	2.90	0.988
Xiaoergou	0.85	17.35	1.22	13.67	0.890
Chishuihe	0.79	13.96	0.25	16.43	0.919

Kumotun	0.92	14.55	1.00	5.52	0.952

Table 9. Evaluation metrics obtained for 7 basins during validation period.

Basins	NS	RPF (%)	EPT (d)	RRD (%)	RMD
① Biliuhe	0.92	8.65	0.40	6.97	0.947
② Geni	0.85	17.02	1.50	10.95	0.904
③ Kehou	0.82	10.91	1.75	23.17	0.820
④ Jiagedaqi	0.84	14.51	0.60	16.92	0.923
⑤ Xiaoergou	0.91	9.60	1.60	16.25	0.905
⑥ Chishuihe	0.81	20.22	0.50	20.78	0.859
⑦ Kumotun	0.70	8.32	1.40	4.46	0.925