Journal of Hydrology 547 (2017) 196–207

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Integrating weather and climate predictions for seamless hydrologic ensemble forecasting: A case study in the Yalong River basin

Aizhong Ye *, Xiaoxue Deng, Feng Ma, Qingyun Duan, Zheng Zhou, Chao Du

Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
State Key Laboratory of Earth Surface and Ecological Resources, Beijing Normal University, Beijing 100875, China

ARTICLE INFO

Article history:
Available online 2 February 2017

Keywords:
Ensemble forecasts
Precipitation
Streamflow
Canonical events
Schaake shuffle

ABSTRACT

Despite the tremendous improvement made in numerical weather and climate models over the recent years, the forecasts generated by those models still cannot be used directly for hydrological forecasting. A post-processor like the Ensemble Pre-Processor (EPP) developed by U.S. National Weather Service must be used to remove various biases and to extract useful predictive information from those forecasts. In this paper, we investigate how different designs of canonical events in the EPP can help post-process precipitation forecasts from the Global Ensemble Forecast System (GEFS) and Climate Forecast System Version 2 (CFSv2). The use of canonical events allow those products to be linked seamlessly and then the post-processed ensemble precipitation forecasts can be generated using the Schaake Shuffle procedure. We used the post-processed ensemble precipitation forecasts to drive a distributed hydrological model to obtain ensemble streamflow forecasts and evaluated those forecasts against the observed streamflow. We found that the careful design of canonical events can help extract more useful information, especially when up-to-date observed precipitation is used to setup the canonical events. We also found that streamflow forecasts using post-processed precipitation forecasts have longer lead times and higher accuracy than streamflow forecasts made by traditional Extend Streamflow Prediction (ESP) and the forecasts based on original GEFS and CFSv2 precipitation forecasts.

1. Introduction

An accurate and reliable hydrological forecast can provide useful information for emergency and water resources managers to deal with hydrologic hazards such as floods and droughts (Pappenberger et al., 2013). Numerical ensemble weather (Gneiting and Raftery, 2005) and climate predictions, which are used as inputs to streamflow prediction models, have great bearings on the accuracy and reliability of a hydrological forecast (Xu et al., 2014).

A number of international initiatives have highlighted the significant improvement in the predictive skills of precipitation forecasts generated by numerical weather prediction (NWP) and climate models in the recent years. For example, a suite of medium-range ensemble meteorological forecast products in the TIGGE database have shown significant skill in forecasting severe weather events such as heavy rains, where TIGGE stands for THORPEX Interactive Grand Global Ensemble, and THORPEX stands for The Observing System Research and Predictability Experiment, which is a World Meteorological Organization sponsored decade-long project started in 2003 to improve medium range forecast of severe weather events (Swinbank et al., 2005). The North American Multi-Model Ensemble (NMME) has assembled nine different climate models to produce global seasonal climate forecasts with a lead time of up to eleven and half months. The preliminary results showed that hydrological forecasts driven by NMME seasonal precipitation forecasts showed meaningful skill (i.e., with the skill better than that of the forecasts by Extended Streamflow Prediction (ESP), which uses climatology as forcing) even six months into the future (Kirtman et al., 2014, Ma et al., 2016). Besides the improved skill in the NWP and climate forecasts, many of those forecasts come with a long retrospective reforecasts (>30 years), which allow forecast users to post-process those forecasts to remove various biases and to downscale them to the application domain of interest.

The Global Ensemble Forecast System (GEFS) and the Climate Forecast System version 2 (CFSv2) are the medium- and long-range meteorological forecasts, respectively, produced by the U.S. National Center for Environmental Prediction (NCEP). Those forecasts are easily accessible online in real-time and both have over 30 years of retrospective reforecasts available. Like other forecast products, the GEFS and CFSv2 forecasts are plagued by various

* Corresponding author at: Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China.

E-mail address: Azye@bnu.edu.cn (A. Ye).

http://dx.doi.org/10.1016/j.jhydrol.2017.01.053
0022-1694/

uncertainties and cannot be applied directly to hydrological forecasting. They need to be post-processed before they can be used to drive a hydrological model (Li et al., 2009a; Liu et al., 2013).

Post-processing serves numerous purposes, including removing systematic and spread biases, downscaling, and generating spatio-temporal time series needed by a hydrological model. There are many post-processing methods that have been used in hydrological forecasting, such as Perfect Prognosis (Klein et al., 1959), Model Conditional Processor (MCP) (Todini, 2008), Model Output Statistics (MOS) (Glahn et al., 2009; Glahn and Lowry, 1972), Multi-model Bayesian methods (Ajami et al., 2007; Duan et al., 2007; Coccia and Todini, 2011), wavelet transformation autocorrelation methods (Bogner and Kalas, 2008) and Ensemble Pre-Processor (EPP) (Schaake et al., 2007). Those post-processing has shown to be very effective in improving the raw meteorological (or hydrological) forecasts from NWP and climate (or hydrological) models (Olsson et al., 2016).

This study uses the EPP method as a statistical post-processor of precipitation forecasts. EPP first establishes joint probability distributions between the forecasted precipitation and their corresponding observation for predetermined canonical events (Schaake et al., 2007). A canonical event refers to a meteorological event for a given location, with a specific lead time and duration. Canonical events can be divided into single-time step canonical events (SCEs, e.g., events with time steps of one day or less like the first 6-h precipitation, the 2nd day precipitation) and composite time step canonical events (CCEs, e.g., multi-day events like the day 6-day 10 average daily precipitation) (Tao et al., 2014). Once the joint probability distributions between forecasts and observations are established for all canonical events, we can obtain the conditional distributions of the observed precipitation given a specific precipitation forecast. Afterwards, the Schaake Shuffle procedure can be applied to create different ensemble forecast members of precipitation time series to be used to drive a hydrological model (Clark et al., 2004).

The choice of canonical events can influence the quality of post-processing by EPP. If canonical events are designed properly, one can extract the maximal amount of predictive information from the raw forecasts. Due to the chaotic nature of the weather and climate systems, long-term precipitation forecasts generally contain less predictive information than shorter lead-time forecasts (Lorenz, 1963). It is nearly impossible to obtain an exact daily precipitation forecast after 7 days into the future. However, it is relatively easy to obtain a satisfactory monthly precipitation forecast for the next month. The reasons are: (1) there is skill in forecasting total amount of precipitation over a long future period, even though it is impossible to avoid the timing error of a precipitation forecast. For this reason, the forecasted precipitation for a future month or a future season can be treated as a canonical event which contains meaningful predictive information. There are many benefits of using canonical events. For example, they can reduce the calculation error when using the average of forecasts to calculate by establishing a statistical relationship between the historical observations and forecasts to extract useful information. Different canonical event schemes, however, will produce different post-processing results (Liu et al., 2013; Tao et al., 2014). The traditional way of generating canonical events in EPP tends to be arbitrary. Since there are some links between the forecasts and observations from the previous few days, we attempted to make use of this link to design canonical events used in EPP to post-process the GEFS and CFS precipitation forecasts.

The organisation of the paper is as follows: Section 2 describes the EPP methodology; Section 3 introduces the data and study domain; Section 4 presents the results and discussion; and Section 5 provides the conclusions.

2. Method

2.1. Ensemble Pre-Processor (EPP)

The Ensemble Pre-Processor (EPP) (Schaake et al., 2007; Wu et al., 2011) is a post-processing method available in the National Weather Service River Forecast System (NWSRFS) that prepares precipitation and temperature forecast ensemble members for hydrological models. This method is aimed to improve the accuracy of forecasts by downscaling the forecast to match the scale of the hydrological models. The methodology transforms the time series of single-value Quantitative Precipitation Forecasts (QPFs) and Quantitative Temperature Forecasts (QTFs) into corresponding ensemble forecasts of precipitation and temperature, which can then be used as input data for the hydrological models.

Fig. 1 shows a flowchart of the main steps of the EPP method. X and Y denote sets of single-valued QPFs and the corresponding observations. Statistical parameters for observations (Y) and the corresponding forecasts (X) then need to be calculated using the EPP. Because precipitation often occurs as a 0 value event for a long period of time, it will affect the stability of the model when the EPP directly calculates single-value precipitation for the statistical parameters. To ensure a sufficient sample size for the calculations, a time window is used to include forecasts before and after the forecast day (e.g., day 5–30 before and after). The window should ensure a sufficient number of non-zero values when the EPP chooses the length of the window. In the statistical analysis, it is necessary to consider the rain and non-rain events. Given the threshold value (e.g., 97%) of zero precipitation, it is assumed that no precipitation occurred when the precipitation is less than the threshold value.

We can determine the marginal distribution of X and Y using the data from years of observations and forecasts. The EPP can fit X and Y to a probability distribution function (PDF), such as a Gamma, Log Normal, Exponential, or Weibull distribution function. A Gamma distribution function is used in the study, because we found the Gamma distribution can get the better precipitation pre-processing results than other distributions in China. It is difficult to determine the joint distribution of X and Y using this PDF. Therefore, we transform X and Y into Gaussian space using the Normal Quantile Transform (NQT) method (Kelly and Krzysztofowicz, 1997, 1997; Montanari and Grossi, 2008), mapping X and Y into standard normal random variables U and V. Thus, we obtain the joint distribution of U and V, $B_{U,V}(u,v)$, which is assumed to be a bivariate standard normal distribution. We can determine $B_{U|V}(u|v)$ from the given conditional PDF, $F_{Y|X}(y|x)$. From $F_{Y|X}(y|x)$, we can obtain the ensemble size, which is the number of years of observed data. We can obtain the corresponding forecasts using the inverse of NQT through remapping random variables U and V from the normal space.

2.2. The "Schaake Shuffle"

Using the method described above, we can obtain the conditional probabilistic forecasts that were calculated by raw forecasts and observed precipitation. For each given canonical event, the corresponding ensemble member is generated to construct ensemble forecasts by using the "Schaake Shuffle" (Clark et al., 2004; Schaake et al., 2007) methodology.

Table 1 shows the "Schaake Shuffle" procedure for single-time step canonical events (SCEs). The "Schaake Shuffle" method needs two sets of data, which are the observations and ensemble members. There are three steps: (1) obtaining the ascending ranks of the observation matrix in each time step; (2) obtaining the sampled ensemble members for SCEs in ascending order;

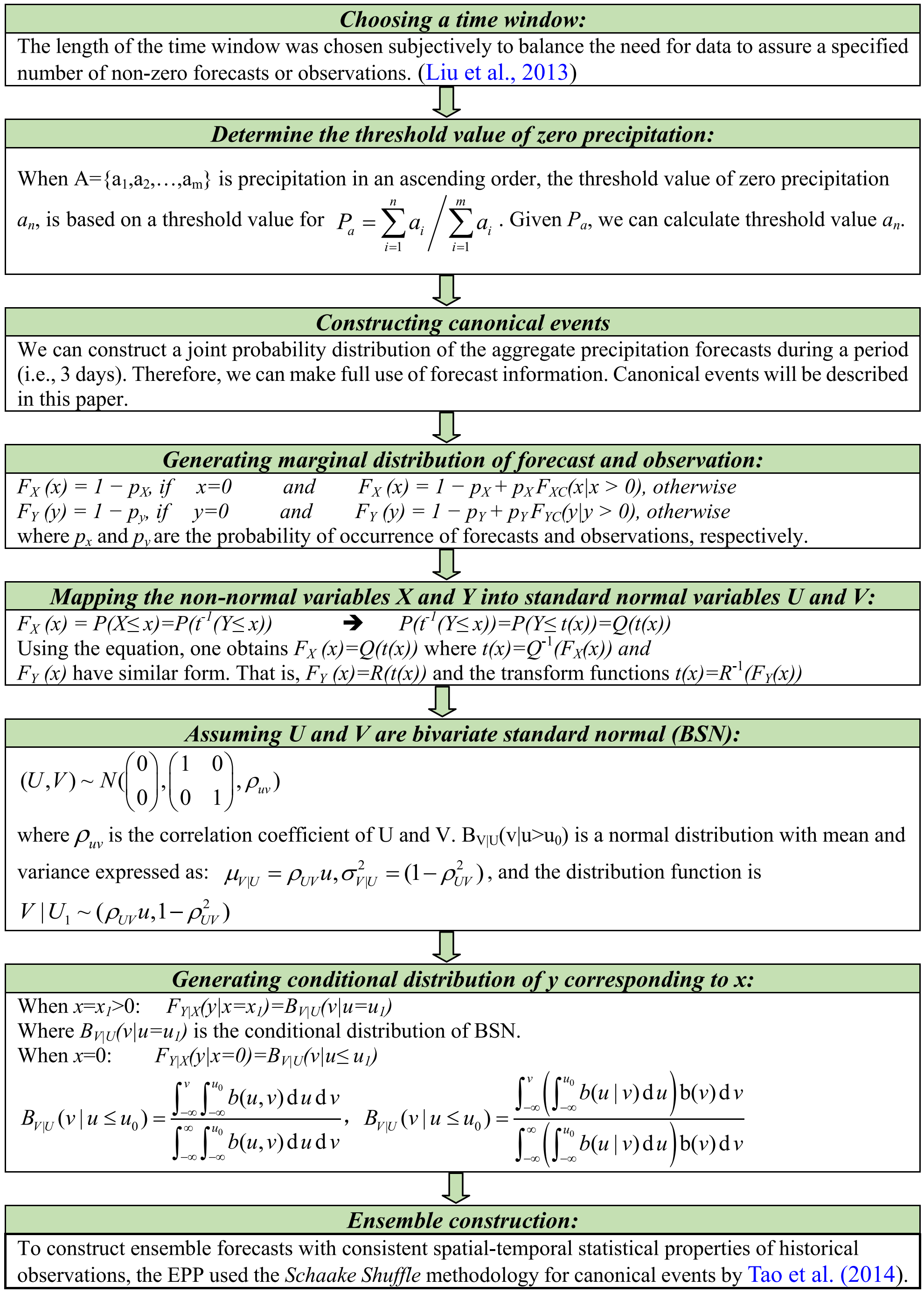

Fig. 1. Flowchart showing the main steps of the EPP method.

and (3) finishing the shuffled ensemble members with the same ranks as the observations. The y matrix is the final output of the EPP.

The composite time step canonical events (CCEs) contain multi-day precipitation (or temperature) information. Expressions for the CCEs and the shuffled ensemble members for use in the CCE formula are as follows:

$$C_{e,j} = \frac{1}{n}\sum\nolimits_{k=1}^{n} a_{e,k}$$

$$y_{i,j} = \frac{x_{e,R_{i,j}} \cdot a_{i,j}}{\frac{1}{n}\sum_{k=1}^{n} a_{e,k}}$$

where i (1–365 days) and k refer to time; j refers to each ensemble member; $R_{i,j}$ refers to the rank of the i,j-th observation; e refers to the canonical event; a refers to the observed precipitation; C refers to the CCE of the observations; x refers to the sampled ensemble members for CCEs in ascending order; and y refers to the shuffled ensemble members for CCEs redistributed into individual time steps according to observed ratios.

Table 2 shows an illustration of the "Schaake Shuffle" procedure for the CCEs. Similar to the SCEs, the CCEs also require observations and ensemble members. There are three steps: (1) generating the CCEs using Eq. (1) and ranking them; (2) re-ordering the sampled ensemble members for the CCEs in an ascending order; and (3) Redistributing the values of shuffled ensemble members for the

Table 1
Illustration of "Schaake Shuffle" for single-time step canonical events (SCEs), where a is the observed precipitation; x is the sampled ensemble members for SCEs in ascending order; and y is the shuffled ensemble members with the same ranks as the observations.

Observation matrix					Sampled ensemble members for SCEs in ascending order			Shuffled ensemble members with the same ranks as the observations		
Year	Time Step 1	Time Step 2	Ranks		Ens. #	Time Step 1	Time Step 2	Ens. #	Time Step 1	Time Step 2
2001	$a_{1,1}$	$a_{2,1}$	5	6	1	$x_{1,1}$	$x_{2,1}$	1	$y_{1,1} = x_{1,5}$	$y_{2,1} = x_{2,6}$
2002	$a_{1,2}$	$a_{2,2}$	4	7	2	$x_{1,2}$	$x_{2,2}$	2	$y_{1,2} = x_{1,4}$	$y_{2,2} = x_{2,7}$
2003	$a_{1,3}$	$a_{2,3}$	1	4	3	$x_{1,3}$	$x_{2,3}$	3	$y_{1,3} = x_{1,1}$	$y_{2,3} = x_{2,7}$
2004	$a_{1,4}$	$a_{2,4}$	6	2	4	$x_{1,4}$	$x_{2,4}$	4	$y_{1,4} = x_{1,6}$	$y_{2,4} = x_{2,2}$
2005	$a_{1,5}$	$a_{2,5}$	7	1	5	$x_{1,5}$	$x_{2,5}$	5	$y_{1,5} = x_{1,7}$	$y_{2,5} = x_{2,1}$
2006	$a_{1,6}$	$a_{2,6}$	3	5	6	$x_{1,6}$	$x_{2,6}$	6	$y_{1,6} = x_{1,3}$	$y_{2,6} = x_{2,5}$
2007	$a_{1,7}$	$a_{2,7}$	2	3	7	$x_{1,7}$	$x_{2,7}$	7	$y_{1,7} = x_{1,2}$	$y_{2,7} = x_{2,3}$
Step1					Step2			Step3		

CCEs into individual time steps according to the observed ratios using Eq. (2). The y matrix is the final output of the EPP.

The order of separating the CCEs into each time step is determined according to the order of the correlation coefficients (observations and forecast), which is from small to large, or the order of the lead time, which is from the future to the present.

2.3. The distributed hydrology model

The distributed time-variant gain model (DTVGM) (Xia et al., 2005; Wang et al., 2009) is used in this study and was established on the basis of the geographic information system (GIS) and Remote sensing (RS) information. Based on the GIS and RS, the model extracts land surface information, such as the slope, flow direction, flow path, river networks, watershed boundaries and land cover (Li et al., 2009b; Ma et al., 2014). The water balance procedure can be expressed by Eq. (3):

$$P_i + W_i = W_{i+1} + g_1\left(\frac{W_{ui}}{WM_u C_j}\right)^{g_2} P_i + W_{ui}\cdot K_r + f_c\cdot\left(\frac{W_{gi}}{WM_g}\right) + Ep_i\cdot\left(\frac{W_{ui}}{WM_u C_j}\right) \tag{3}$$

where W is the soil moisture (mm); W_u is the upper soil moisture at the sub-basin (mm); W_g is the lower soil moisture at the sub-basin (mm); WM_u is the upper saturated soil moisture (mm); u is the 'upper' soil; WM_g is the lower saturated soil moisture (mm); f_c is the soil permeability coefficient (mm/h); g_1 and g_2 are parameters ($0 < g_1 < 1$, $0 < g_2$); g_1 is the runoff coefficient when the soil is saturated; g_2 is the soil moisture parameter; C is the land cover parameter; K_r is the subsurface runoff coefficient; K_g is the groundwater runoff coefficient; K_e is the evaporation coefficient; i is a period of time; and j is the hydrological unit number.

DTVGM evaluations can use three time scales: the monthly scale, daily scale, and hourly scale. An appropriate time scale should be chosen according to need when using the model. A daily scale is used in this study.

DTVGM used the degree–day method to compute the depth of snowmelt (Bormann et al., 2014). The snow melted water was added to precipitation in DTVGM.

$$S = \begin{cases} \alpha\cdot(T_a - T_m) & T_a > T_m \text{ and } \alpha\cdot(T_a - T_m) < H_s \\ H_s & T_a > T_m \text{ and } \alpha\cdot(T_a - T_m) > H \\ 0 & T_a > T_m \end{cases} \tag{4}$$

$$\alpha = 11\cdot\frac{\rho_s}{\rho_w}$$

where S is melted water (mm), α is degree-day factor (DDF) (mm/°Cd), T_a is average air temperature (°C), T_m is the critical

Table 2
Illustration of the "Schaake Shuffle" procedure for composite time step canonical events (CCEs) (Tao et al., 2014).

Observation matrix								Sampled ensemble members for CCEs in ascending order		
Year	Time Step 1	Time Step 2	Time Step 3	CCEs		Ranks		Ens. #	Time Step 1 + 2	Time Step 1 + 2 + 3
2001	$a_{1,1}$	$a_{2,1}$	$a_{3,1}$	$C_{1,1} = (a_{1,1} + a_{2,1})/2$	$C_{2,1} = (a_{1,1} + a_{2,1} + a_{3,1})/3$	5	6	1	$x_{1,1}$	$x_{2,1}$
2002	$a_{1,2}$	$a_{2,2}$	$a_{3,2}$	$C_{1,2} = (a_{1,2} + a_{2,2})/2$	$C_{2,2} = (a_{1,2} + a_{2,2} + a_{3,2})/3$	4	7	2	$x_{1,2}$	$x_{2,2}$
2003	$a_{1,3}$	$a_{2,3}$	$a_{3,3}$	$C_{1,3} = (a_{1,3} + a_{2,3})/2$	$C_{2,3} = (a_{1,3} + a_{2,3} + a_{3,3})/3$	1	4	3	$x_{1,3}$	$x_{2,3}$
2004	$a_{1,4}$	$a_{2,4}$	$a_{3,4}$	$C_{1,4} = (a_{1,4} + a_{2,4})/2$	$C_{2,4} = (a_{1,4} + a_{2,4} + a_{3,4})/3$	6	2	4	$x_{1,4}$	$x_{2,4}$
2005	$a_{1,5}$	$a_{2,5}$	$a_{3,5}$	$C_{1,5} = (a_{1,5} + a_{2,5})/2$	$C_{2,5} = (a_{1,5} + a_{2,5} + a_{3,5})/3$	7	1	5	$x_{1,5}$	$x_{2,5}$
2006	$a_{1,6}$	$a_{2,6}$	$a_{3,6}$	$C_{1,6} = (a_{1,6} + a_{2,6})/2$	$C_{2,6} = (a_{1,6} + a_{2,6} + a_{3,6})/3$	3	5	6	$x_{1,6}$	$x_{2,6}$
2007	$a_{1,7}$	$a_{2,7}$	$a_{3,7}$	$C_{1,7} = (a_{1,7} + a_{2,7})/2$	$C_{2,7} = (a_{1,7} + a_{2,7} + a_{3,7})/3$	2	3	7	$x_{1,7}$	$x_{2,7}$
Step 1								Step 2		

Shuffled ensemble members for CCEs redistributed into individual time steps according to observed ratios			
Ens. #	Time Step 1	Time Step 2	Time Step 3
1	$y_{1,1} = x_{1,5} * a_{1,1}/C_{1,1}$	$y_{2,1} = x_{1,5} * a_{2,1}/c_{1,1}$	$y_{3,1}=x_{2,6} * a_{3,1}/c_{2,1}$
2	$y_{1,2} = x_{1,4} * a_{1,2}/C_{1,2}$	$y_{2,2} = x_{1,4} * a_{2,2}/c_{1,2}$	$y_{3,2} = x_{2,7} * a_{3,2}/c_{2,2}$
3	$y_{1,3} = x_{1,1} * a_{1,3}/C_{1,3}$	$y_{2,3} = x_{1,1} * a_{2,3}/c_{1,3}$	$y_{3,3} = x_{2,4} * a_{3,3}/c_{2,3}$
4	$y_{1,4} = x_{1,6} * a_{1,4}/C_{1,4}$	$y_{2,4} = x_{1,6} * a_{2,4}/c_{1,4}$	$y_{3,4} = x_{2,2} * a_{3,4}/c_{2,4}$
5	$y_{1,5} = x_{1,7} * a_{1,5}/C_{1,5}$	$y_{2,5} = x_{1,7} * a_{2,5}/c_{1,5}$	$y_{3,5} = x_{2,1} * a_{3,5}/c_{2,5}$
6	$y_{1,6} = x_{1,3} * a_{1,6}/C_{1,6}$	$y_{2,6} = x_{1,3} * a_{2,6}/c_{1,6}$	$y_{3,6} = x_{2,5} * a_{3,6}/c_{2,6}$
7	$y_{1,7} = x_{1,2} * a_{1,7}/C_{1,7}$	$y_{2,7} = x_{1,2} * a_{2,7}/c_{1,7}$	$y_{3,7} = x_{2,3} * a_{3,7}/c_{2,7}$
Step 3			

Table 3
Description of the common verification measures used in the study. The x is observed data, y is simulated data, and z is benchmark data.

Verification Measures	Formulas	Descriptions	Perfect/ No skill
Nash-Sutcliffe efficiency value (NSE)	$NSE = 1 - \frac{\sum_{i=1}^{N}(x_i - y_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$	Assessing the predictive power of hydrological models; quantitatively describe the accuracy between forecasts and observations	$1/\leqslant 0$
NSE calculated on inverse transformed flows	$NSE_I = 1 - \frac{\sum_{i=1}^{N}\left(\frac{1}{x_i} - \frac{1}{y_i}\right)^2}{\sum_{i=1}^{N}\left(\frac{1}{y_i} - \overline{\frac{1}{y_i}}\right)^2}$		
NSE calculated on benchmark model	$NSE_B = 1 - \frac{\sum_{i=1}^{N}(x_i - y_i)^2}{\sum_{i=1}^{N}(y_i - z_i)^2}$		
Root Mean Square Error (RMSE)	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2}$	Association of forecasts and observations over a long time period	$0/\infty$
Pearson Correlation Coefficient	$R = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}}$	Linear dependency between forecasts and observations	$1/\leqslant 0$
rBias	$rBias = \left(\sum_{i=1}^{N}x_i / \sum_{i=1}^{N}y_i - 1\right) \cdot 100\%$	Relative difference between forecasts and observations	$0/\infty$
BSS	$BS = \left[flag(y,t) - \frac{1}{n}\sum_{i=1}^{n}flag(x_i,t)\right]^2$ $flag(x,t) = \begin{cases} 1, x \geqslant t \\ 0, x < t \end{cases} BSS = \left(1 - \frac{BS}{BS_{ref}}\right), x_j < y < x_{j+1}$	Brier Skill Score. t is the threshold. Ref. is a reference forecast (e.g., climatology)	1/0
CRPSS	$CRPS = \sum_{i=1}^{j-1}P_i^2 \cdot (x_{i+1} - x_i) + P_j^2 \cdot (y - x_j) + (P_j - 1)^2 \cdot (x_{j+1} - y)$ $+\sum_{i=j+1}^{n}(P_i - 1)^2 \cdot (x_{i+1} - x_i) CRPSS = \left(1 - \frac{CRPS}{CRPS_{ref}}\right), x_j < y < x_{j+1}$	Continuous Rank Probability Skill Score. P is the probability that was forecast.	1/0

temperature of snowmelt, ρ_s is snow density, ρ_w is water density, H_s is snow cover depth.

DTVGM used the kinematic wave model for routing (Ye et al., 2013). The flow direction and serial number of sub-basins were defined by an automatic drainage network extraction method (Ye et al., 2005). The routing is calculated from upstream to downstream on each sub-basin, and to the basin outlet.

2.4. The seamless hydrologic ensemble forecasting

The seamless hydrologic ensemble forecasting is explored through integrating weather forecasts and seasonal climate predictions (Yuan et al., 2014). Weather forecasts are from GEFS and climate predictions are from CFSv2. The integrating method is to use the EPP based on canonical events. There are four steps: (1) convert the raw GEFS and CFSv2 time series into daily time series; (2) combine the raw GEFS (1st –8th day) data and the raw CFSv2 (9th-) data to form a new forecast data (GEFS + CFSv2); (3) Use the GEFS + CFSv2 data to design the canonical events for the EPP; (4) the final ensemble forecasts from the EPP was then used as inputs to the hydrological model.

2.5. Model performance measures

To evaluate the daily scale EPP and DTVGM, we considered the following model performance measures (see Table 3): the Nash-Sutcliffe efficiency value (NSE) (Nash and Sutcliffe, 1970); the NSE value calculated for the inverse transformed flows (NSE_I) (Pushpalatha et al., 2012); the NSE value calculated for the benchmark model (NSE_B) (Schaefli and Gupta, 2007); the correlation coefficient (R); the relative bias (rBias); the Root Mean Square Error (RMSE); the Brier Skill Score (BSS) and the Continuous Rank Probability Skill Score (CRPSS) (Brown et al., 2010). The BSS and CRPSS scores are used to evaluate ensemble forecast performance.

We denote x_i and y_i as the forecast and the corresponding observation, respectively, at time i, while N is the number of pairs of forecast and observations. Similarly, $\bar{x}$ and $\bar{y}$ are denoted as the forecast average and the observation average, respectively.

To evaluate the reliability of the predictive distributions, the rank histogram (Hamill, 2001; Yuan et al., 2013) is used in this study to diagnose whether the spread of the ensembles is satisfactory for forecasts. A perfect rank histogram would show observations evenly spread across equal probability bins. The perfect rank value is $1/(n+1)$, where n is the number of ensemble members.

3. Data and study domain

3.1. The Yalong River basin

The Yalong River is the largest tributary of the Jinsha River in the southern region of the Tibetan Plateau. The Yalong River (25°12–34°9N, 96°47–102°42E) runs from the northwest to the southeast, with an approximate total basin area of 136,000 km^2 and mainstream length of 1571 km. The maximum altitude is greater than 6000 m, and the minimum altitude is less than 1000 m (Fig. 2). The mean annual temperature is approximately 2.5 °C. The mean annual precipitation is approximately 600–1800 mm, decreasing from the south to the north. The average annual estuarine discharge is 1860 m^3/s. The major vegetation types are forest, shrub and meadow.

3.2. Observed precipitation and discharge data

The observed precipitation (Shen et al., 2010) used in this study is 0.5-degree gridded data between 1957 and 2009 that were generated from the 2416-gauge national network. There are 76 grids in the Yalong River basin (Fig. 2). We collected 32 years (1980–2011) of daily discharge data for the Ganzi station, which is located in the middle reach of the Yalong River basin.

3.3. The NCEP Global Ensemble Forecast System (GEFS) and Climate Forecast System, Version 2 (CFSv2) forecasts

Data from the GEFS and CFSv2 were used as weather and climate forecast data in this study. The GEFS dataset was provided by NOAA's National Centers for Environmental Prediction (NCEP) (Hamill et al., 2013). The real-time forecasts and the reforecasts were both generated using the GEFS model version 9.0.1. For a detailed description, readers are referred to http://www.emc.ncep.noaa.gov/GFS/impl.php. The Reforecast V2 dataset consists of an 11-member ensemble forecasts produced every day,

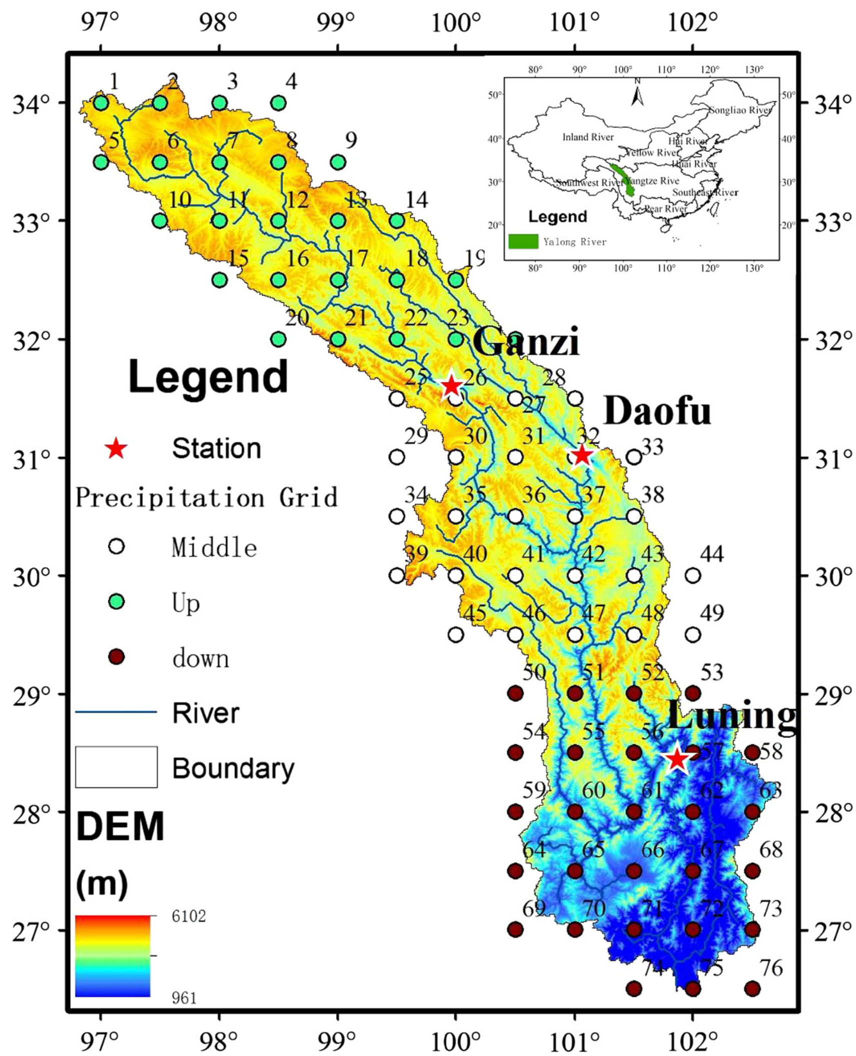

Fig. 2. The illustration of the Yalong River basin.

beginning with 00 UTC initial conditions, from December 1984 to the present. The horizontal resolution of the GEFS forecasts is T254 out to 8 days and T190 from 8 to 16 days. Data are saved at this resolution from day +8 to day +16, which is the end of the GEFS integration period. In this study, the first 8-day precipitation reforecasts, which has a higher spatial resolution, are used. The data format is Grib2 for each day of $1.0^{\circ}\times 1.0^{\circ}$ global coverage (360×181).

The Climate Forecast System, Version 2 (Saha et al., 2014; http://cfs.ncep.noaa.gov) was developed by the Environmental Modeling Center at NCEP. It is a fully coupled model representing the interaction between the Earth's atmosphere, oceans, land and sea ice. This model offers hourly data from around the world for many variables, with a horizontal resolution of 0.5 degrees. CFSv2 uses the latest scientific approaches for collecting or assimilating observations from many data sources: surface observations, upper air balloon observations, aircraft observations, and satellite observations. CFSv2 reforecasts have been extensively used for hydrologic applications (Yuan et al., 2011, 2013). The daily data were used in this study. CFSv2 has updated data on the website from 1985 to the present, with 9-month hindcasts were initiated from every 5th day and run from all 4 cycles of that day, at a spatial resolution of 0.938° by 0.938°.

We used a common period (1985–2009) of GEFS and CFSv2 reforecast datasets for running the hydrological model. Both GEFS and CFSv2 reforecast datasets are bi-linearly interpolated to 0.5 degrees over the Yalong River basin.

4. Results and discussion

4.1. Evaluation of the ensemble means of the GEFS and CFSv2 precipitation forecasts

We evaluated the ensemble means of the GEFS and CFSv2 precipitation forecasts over the 76 grids within the Yalong river basin (Fig. 2) based on different canonical events. The two sets of canonical events for GEFS and CFSv2 are indicated in Table 4, respectively. The GEFS canonical events are daily events (Liu et al., 2013), while the CFSv2 canonical events are multi-day cumulative events starting from the first day.

Figs. 3 and 4 show the correlation coefficients between the observations and the GEFS and CFSv2 forecasts for all canonical events for each day of the year (vertical axis). The horizontal axis denotes different canonical events. Each of subplots within the figures are results for all 76 grids over the Yalong River basin. The orange to red colour in the plots indicates a high correlation, whereas the blue colour indicates a low correlation. Most of correlation coefficients are greater than 0.7, which indicates that the precipitation forecasts have meaningful skill. When comparing

Table 4
Two canonical events to evaluate GEFS and CFSv2.

GEFS: 1986–2009															
Events	1		2		3		4		5		6		7		8
Start	1		2		3		4		5		6		7		8
End	1		2		3		4		5		6		7		8
CFSv2: 1983–2009															
Events	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Start	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
End	1	2	3	4	5	6	7	8	9	10	14	18	22	26	30

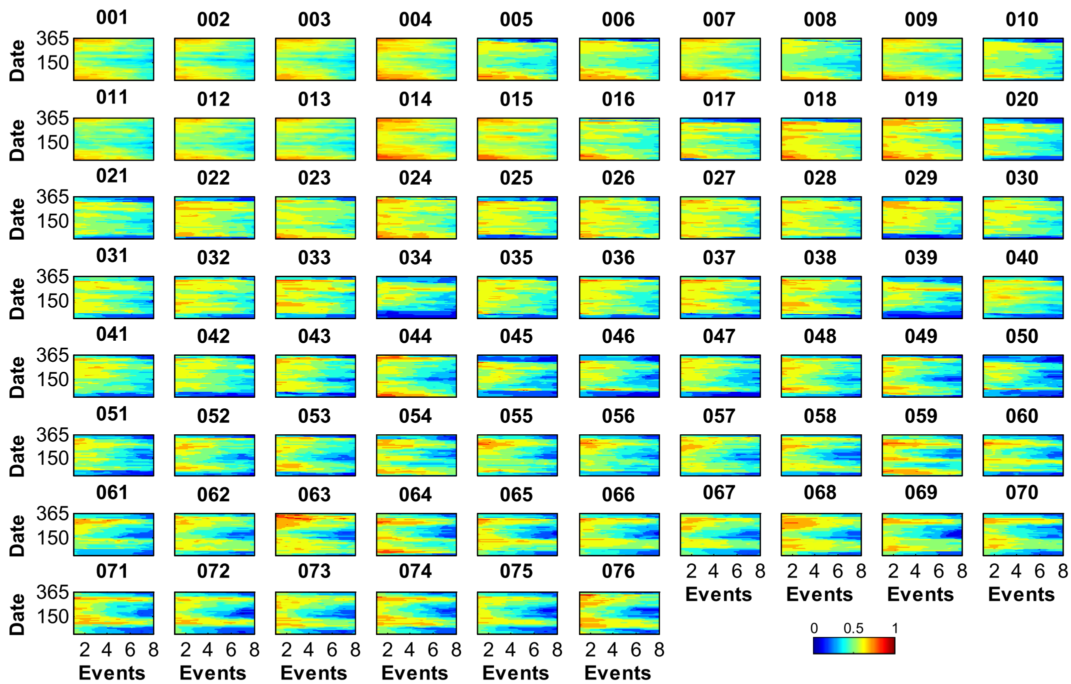

Fig. 3. The correlation coefficients between the raw GEFS forecast and observations in the given canonical events.

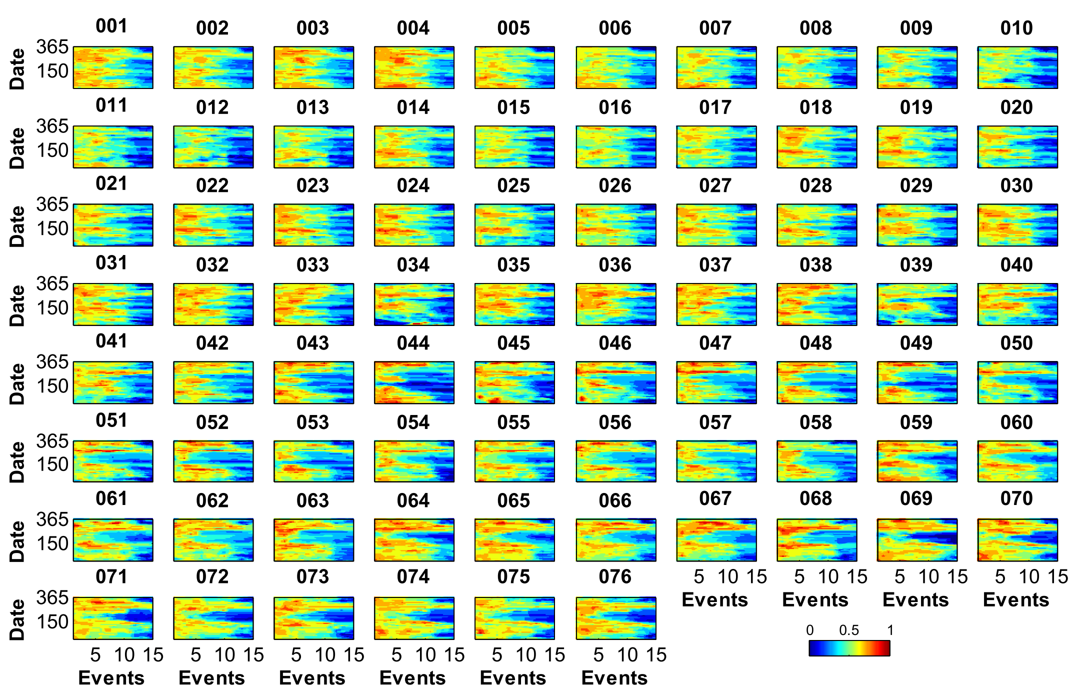

Fig. 4. The correlation coefficients between the raw CFSv2 forecast and observations in the given canonical events.

Table 5
Different scenarios of canonical events in GEFS. '−1' means using 1 day of observed data.

GEFS (1986–2009)	Events	1	2	3	4	5	6	7	8
Scenario 1: daily SCEs	Start	1	2	3	4	5	6	7	8
	End	1	2	3	4	5	6	7	8
Scenario 2: cross CCEs	Start	1	1	1	2	3	4	5	6
	End	1	2	3	4	5	6	7	8
Scenario 3: from 1st CCEs	Start	1	1	1	1	1	1	1	1
	End	1	2	3	4	5	6	7	8
Scenario 4: using 1 day of observed data CCEs	Start	−1	−1	−1	−1	−1	−1	−1	−1
	End	1	2	3	4	5	6	7	8
Scenario 5: using 3 day of observed data CCEs	Start	−3	−3	−3	−3	−3	−3	−3	−3
	End	1	2	3	4	5	6	7	8
Scenario 6: using 5 day of observed data CCEs	Start	−5	−5	−5	−5	−5	−5	−5	−5
	End	1	2	3	4	5	6	7	8

individual grids, we note that the correlation coefficients indicate that the accuracy of the ensemble mean forecasts are decreasing with the lead time. Both GEFS and CFSv2 have higher correlations in cool seasons than in warm seasons for the upstream grids (1–24). For the GEFS daily precipitation forecasts, notable accuracy is indicated mostly for the first few days (Fig. 3). In the winter, the accuracy of daily forecasts can last up to approximately one week. Forecasts of CFSv2 cumulative precipitation have notable accuracy (i.e., the correlation coefficient values of >0.5) for even day 30 for almost all of the grids and for some of the seasons (Fig. 4).

4.2. Schemes for designing canonical events

The EPP calculations were based on canonical events. Canonical events can be designed as different combinations, which may result in different outcomes. We designed six different scenarios for the canonical events for GEFS as shown in Table 5.

In the six scenarios, the length of time for each event was successively extended. Notably, in the fourth scenario, this study included the historical observations of precipitation information in each event to improve the accuracy of the pre-process forecast. Fig. 5 shows the historical periods and future periods used in the improved canonical events in this study. As different canonical events under different scenarios can lead to different forecasts, we chose the most satisfactory results from those scenarios as the input of hydrological model.

In this study, most of the canonical events are CCEs. The CCEs needed to be divided into daily precipitation using the "Schaake Shuffle" method. The order of splitting the CCEs into each daily time step is performed according to the ascending order of the correlation coefficients between observations and forecasts for the

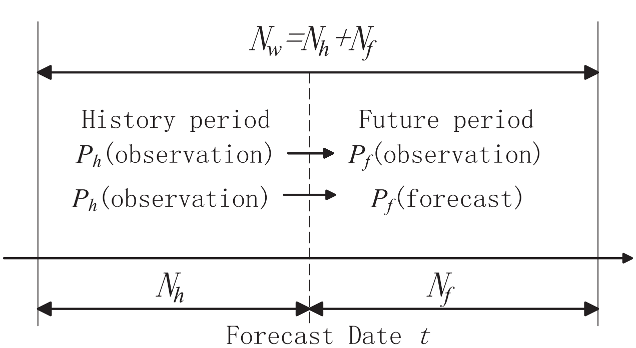

Fig. 5. The data window for a canonical event, where N_h is the number of data points in the history period; N_f is the number of data points in the future period; N_w is the number of data points of a canonical event; P_h is precipitation in the history period; and P_f is precipitation in the future period.

canonical events, which goes from the event with the smallest correlation coefficient to the largest. We provided a comparison of the two schemes below to determine which order of splitting CCEs results in greater improvement for the precipitation forecasts.

Scheme1: The order of splitting CCEs is performed according to the ascending order of the correlation coefficients of the canonical events between observations and the raw forecast. We first split the canonical event with the smallest correlation coefficient into daily precipitation using Eq. (2). Then, we split the event with the next higher correlation coefficient. The following gives an example: Scenario 3 shown in Table 5, assuming the correlation coefficient of the eight events is [0.9, 0.8, 0.6, 0.7, 0.2, 0.4, 0.3, 0.5]. First, we split Event 5, which has the smallest correlation coefficient (0.2), using Eq. (2). Event 5 is the average forecast of day 1 to day 5. Then, we rank the ensemble members of the fifth day forecast ($y_{5,j}$). Upon completion of the initial event, the evaluation of the next event will be successively performed using the same method.

Scheme2: The order of splitting CCEs is performed according to the lead time, which is from the future to the present. We first split the canonical event with the longest lead time into daily precipitation according to Eq. (2). Then, we split the canonical event with the penultimate lead time. The calculation method is similar to Scheme 1.

Fig. 6 displays the RMSE of the rank histograms in six canonical event scenarios for those two schemes. The perfect RMSE is 0. Each figure shows the RMSE of 76 grids (vertical axis) in 8 days (horizontal axis).

When comparing individual schemes, we note that the RMSE of the rank histogram indicates that the reliability of ensemble forecasts is different for the different canonical event scenarios. Fig. 7 shows the RMSE mean of different scenarios. As the canonical events in scenario 1 are SCEs, the RMSE has the highest value, which indicates that the SCEs are the most unreliable canonical events. The CCEs can generate more reliable ensemble forecasts than the SCEs, especially when the historical precipitation observations are included in the event (Scenario 4). However, the RMSEs for Scenarios 5 and 6 are not substantially different than Scenario 4, which indicates that using the first day of historical observed data is sufficient. The main explanation for this is because we would sacrifice useful forecast information if we add too much past observation data.

When comparing the two schemes, the RMSE of Scheme 2 is smaller than Scheme 1 (Fig. 7), which indicates that Scheme 2 is better than Scheme 1 (i.e., the order of splitting CCEs into daily time step is better according to the order of lead times than according to the order of correlation coefficients).

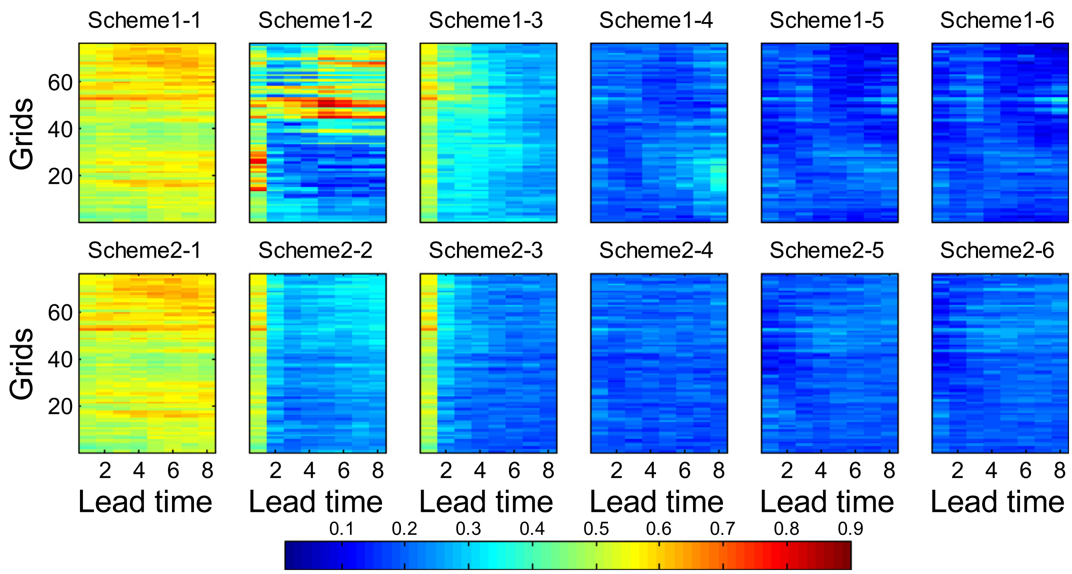

Fig. 6. The RMSE of rank histograms in six canonical event scenarios for two schemes.

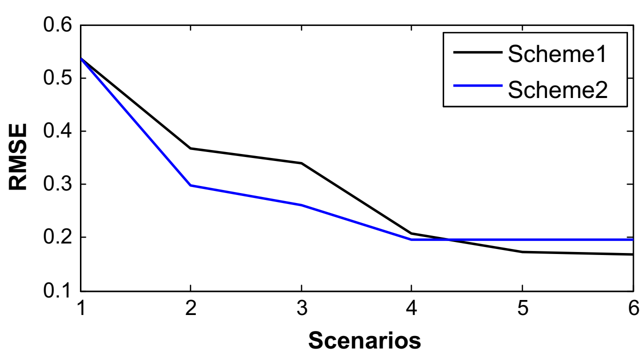

Fig. 7. The RMSE average of the rank histograms in six scenarios for two schemes.

4.3. Verification of the EPP ensemble precipitation forecasts

Through the analysis described above, we found that the best scenarios for the canonical event is Scenario 4 in Scheme 2. Therefore, we selected this scenario to verify the EPP ensemble precipitation forecast. The results indicate that the pre-processing of raw precipitation forecasts has resulted in significant improvement in precipitation forecasts. The Nash-Sutcliffe efficiency value has been substantially improved from 0.1 to 0.4 for all grids. The correlation coefficients have been increased slightly for lead times longer than four days in the southern area of the basin. In addition, the bias indicates a notable decline from 1 to 0.1 or less. Similarly, the RMSE values are also reduced.

Fig. 8 is a rank histogram of grid 001 for Scenario 4 in Scheme 2. It can be used to evaluate the reliability of the EPP ensemble precipitation forecasts. Other grids display similar rank histograms as grid 001 and are not shown here. The rank value in the figure is the frequency of the observed values falling within the rank interval. The perfect rank value is 100/24, as shown by the blue line on Fig. 8. The histogram shows that the rank histogram is relatively flat and that the RMSE value is 0.2687, which indicates that the EPP result is satisfactory and that the ranks are uniform.

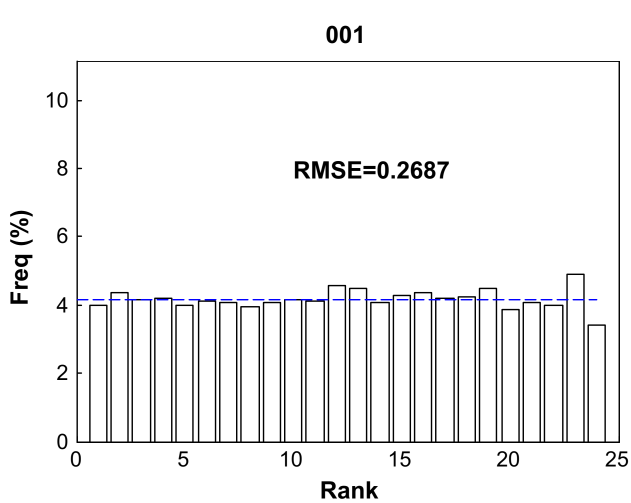

Fig. 8. The rank histogram of the 001 grid for Scenario 4 in Scheme 2.

4.4. Calibration and validation of the DTVGM

To demonstrate the usefulness of the pre-processed GEFS and CFSv2 ensemble precipitation forecasts, we evaluated those forecasts by applying them to a hydrological model, DTVGM. In order to calibrate and validate the DTVGM, we obtained 30 years (1980–2009) of observed precipitation to drive the model. We selected the Ganzi hydro-station (31.6186N, 99.9673E, Fig. 2) to verify the streamflow simulation based on observed precipitation. The total drainage area above the station is 32,575 km^2. The calibration period is 1980–1999, the validation period is 2000–2009.

A manual calibration method was used to calibrate the model parameters, because automatic calibration would be too time consuming as the distributed hydrological model take a long time to run. During the manual calibration, the model was run a few times to ensure that NSE, R and rBias were good.

The R, NSE, NSE_I, NSE_B, rBias and RMSE were shown in Table 6. When computing NSE_B, the daily discharge averaged over the 1980–2009 period was treated as the benchmark daily discharge. All performance indices were reasonably good for a river basin which is humid and exhibits a highly seasonal streamflow pattern (Moriasi et al., 2007). The NSE values of the daily simulated streamflow are close to 0.8 for both the calibration and validation periods, which surpassed the daily streamflow forecast skill standard of 0.6 or above, according to Streamflow Forecasting Manual issued by Chinese Ministry of Water Resources (MWR, 2002). The

Table 6
The performance indices of DTVGM discharge simulation in Ganzi station.

	R	*NSE*	*NSE$_l$*	*NSE$_B$*	*rBias*	*RMSE*
Calibration (1980–1999)	0.899	0.805	0.619	0.455	−4.2%	103.206
Validation (2000–2009)	0.895	0.798	0.739	0.466	−4%	99.831

NSE$_l$ values are in the 0.6–0.7 range, which indicate that low flow simulations are also reasonable for both the calibration and validation periods, even though the NSE$_l$ for measuring the performance of low flows is lower than the NSE value for measuring high flows. The NSE$_B$ values are 0.455 and 0.466 (>0) for the calibration and validation periods, respectively, which suggest that the model simulated daily streamflow are much better than the long-term average daily discharge. The rBias values are −4.2% and −4%, which indicate an under-bias in streamflow simulations. The slight negative rBias value could be due to under-reported precipitation, model structural error or improper model parameters (Zhao et al., 2011), but the under-bias are still within acceptable range of 5% or less (i.e., >−5% and <5%). Figs. 9 and 10 display the daily precipitation-discharge hydrographs in the calibration period and the validation period. We note that the simulated discharge values (i.e., the black dash line) are generally close to the observations (i.e., the solid red line) in all periods. Those results indicate that DTVGM is an acceptable (Moriasi et al., 2007) model for hydrological simulation of this basin.

4.5. Verification of the ensemble means of the streamflow forecasts

We verified the ensemble means of the streamflow forecasts generated by DTVGM, driven by pre-processed GEFS and CFSv2 precipitation forecasts. Five sets of precipitation data are tested as inputs to DTVGM, including the observed precipitation (OBS), precipitation as used in the ESP framework (ESP), the pre-processed GEFS precipitation forecasts (GEFS), the pre-processed CFSv2 precipitation forecasts (SFSv2) and the combined pre-processed precipitation forecasts (GEFS + CFSv2). The GEFS + CFSv2 precipitation data has integrated GEFS for the first 7 days, and CFSv2 from day 8 to one month into the future. In the ESP framework, historical observed precipitation data for forecast window from each year of the observed data period are used as ensemble precipitation members and are used to drive DTVGM.

Fig. 11 shows the daily, weekly and monthly model performance indices of the streamflow simulation. Comparing the daily forecasts, we note that the performance indices indicate that the accuracy of streamflow forecasts decreased with lead time. The

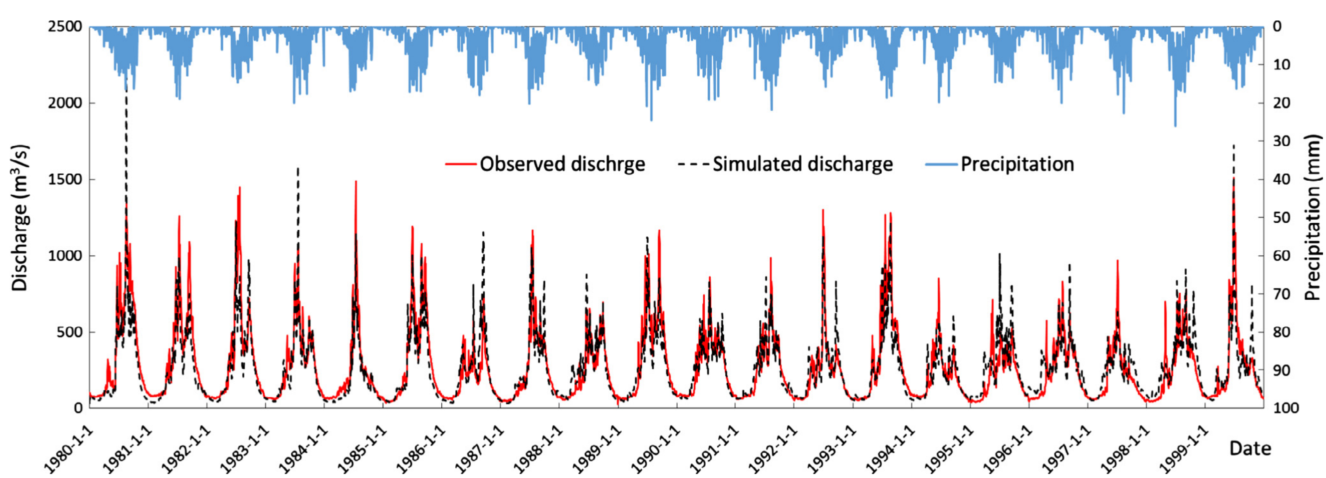

Fig. 9. Daily discharge hydrograph in calibration period (1980–1999).

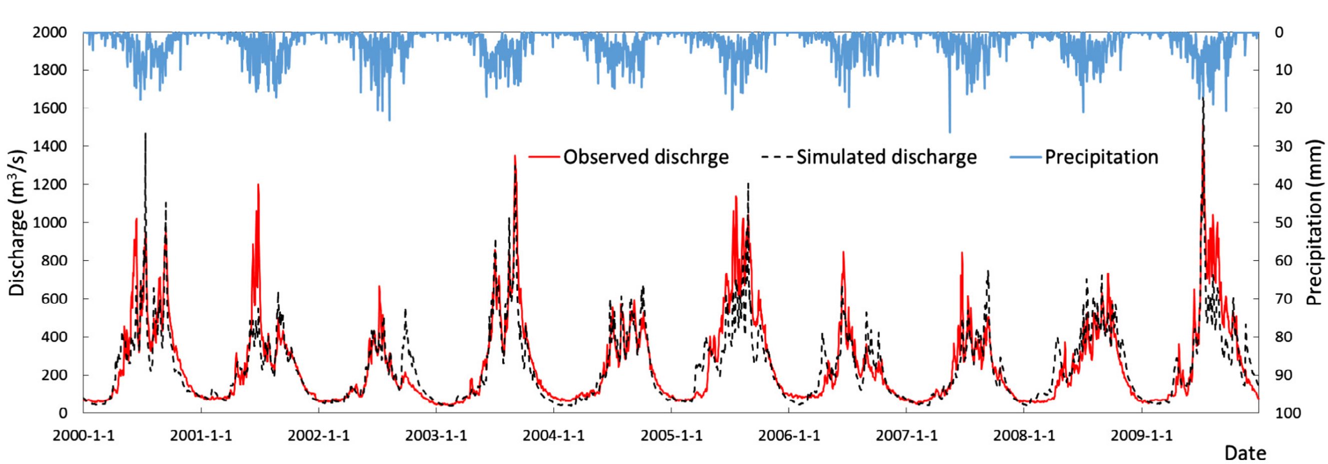

Fig. 10. Daily discharge hydrograph in validation period (2000–2009).

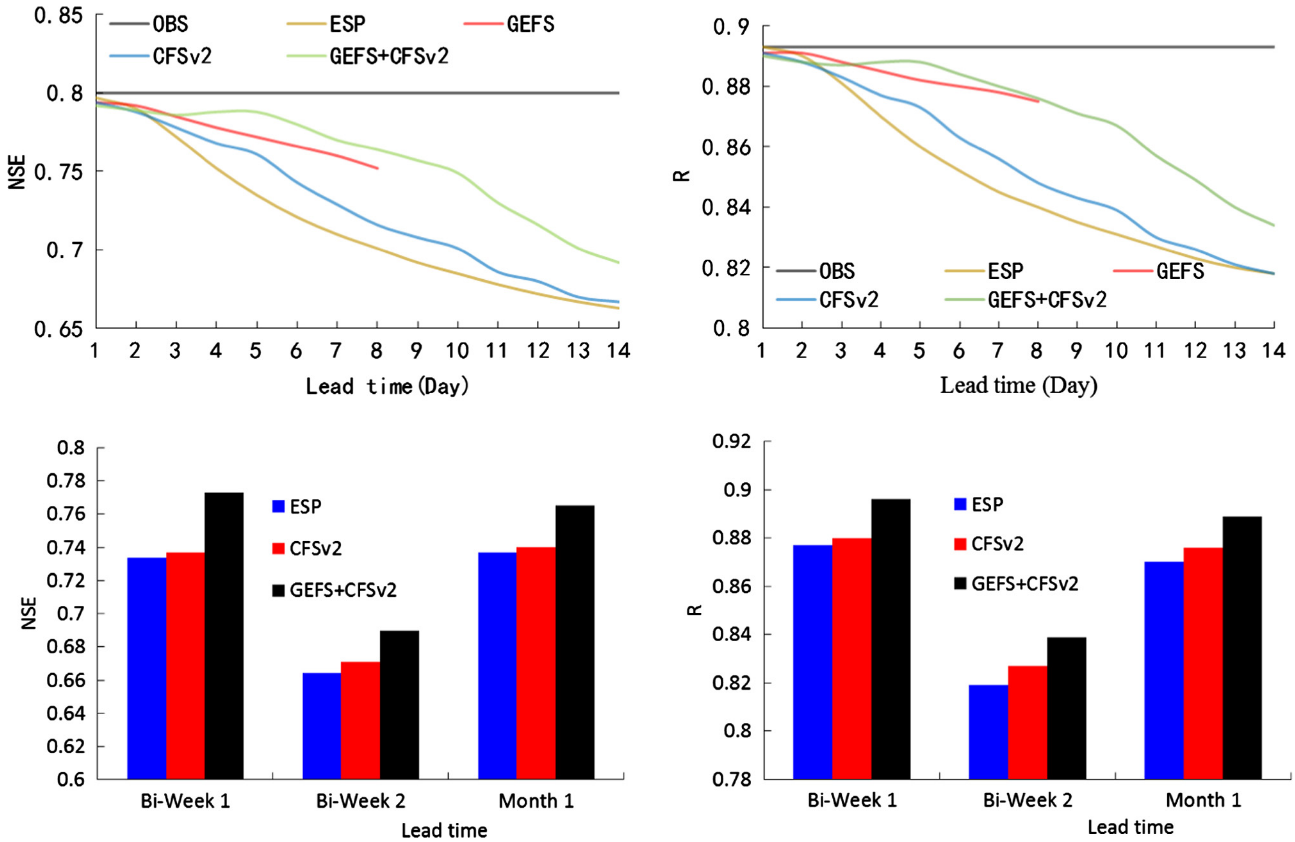

Fig. 11. Daily, weekly and monthly model performance indices of the discharge simulations.

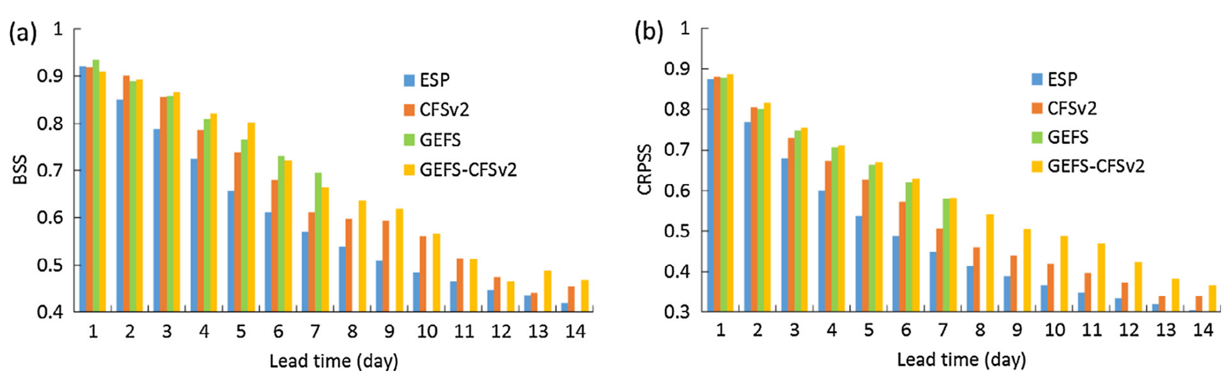

Fig. 12. Brier skill score (BSS) and continuous ranked probability skill score (CRPSS) of the ensemble discharge forecasts using different precipitation forecast (during 1999–2008). The different colors of bars indicate different precipitation forecast.

integrated GEFS and CFSv2 forecast precipitation can achieve the best streamflow forecast. The ESP performance is the weakest as it correspond to the smallest NSE and correlation values for streamflow simulations. The GEFS performance indices are higher than CFSv2 performance indices for the first few days. The GEFS + CFSv2 precipitation forecast also has high performance indices for weekly and monthly streamflow forecasts.

4.6. Ensemble verification of the streamflow forecasts

Fig. 12 displays the Brier skill score (BSS) and the continuous ranked probability skill scores (CRPSS) of the ensemble discharge forecasts relative to climatology using different precipitation forecasts (during 1999–2008). All forecasts exhibit quite significant skills compared to climatology, especially for lead times of day 1 to day 3. Using different precipitation forecasts as inputs, the performance indices of the ensemble streamflow forecasts are similar to the performance indices for the ensemble mean forecast in Section 4.5. The GEFS and GEFS + CFSv2 precipitation forecasts also have high performance indices when the ensemble streamflow forecasts were evaluated over the two-week period. The ESP performance indices are the worst with the smallest NSE values and correlation coefficients.

5. Conclusions

In this study, we presented the EPP and "Schaake Shuffle" methods in the context of canonical events in detail. Six canonical event scenarios and two schemes for designing canonical event were evaluated to determine suitable canonical event setups. We also tested the combined pre-processed GEFS and CFSv2 precipitation forecasts as inputs to the DTVGM model and obtain a seamless hydrological ensemble forecast for the Yalong River basin.

The performance indices indicate that the CCEs that include a day with observed precipitation data prior to day 1 can help extract more useful information. We also found that the optimal order for "Schaake Shuffle" to split CCEs into daily time step is the one based on the order of the lead time from the future to the present.

The streamflow forecast results are the best when the combined GEFS + CFSv2 precipitation forecasts are used. The resulting seamless streamflow forecasts have longer lead times and higher accuracy than the forecasts based on the original CFSv2 precipitation forecasts and the ESP streamflow forecasts.

The EPP method used in this study was based on a parametric approach. We used the Gamma distribution to describe the daily precipitation amounts. Although this distribution was tested, the choice and calibration of this distribution may introduce some uncertainty in the results. Alternative approaches are the non-parametric ones (Van Steenbergen et al., 2012) and further studies in the future should consider those approaches. Another potential issue with the setup of the canonical events in this study is that the data we used contained only six possible extreme events. The optimal events we found may be the local optimum instead of the global optimum. A more comprehensive evaluation that include other strategies for choosing the canonical events (for example, choosing canonical events based information theory) should also be considered in future studies.

Acknowledgments

This study was supported by the Natural Science Foundation of China (Nos. 41475093, 51279210), the State Key Laboratory of Severe Weather Open Research Program (NO. 2015LASW-A05) and the Beijing Science and Technology Plan Project (No. Z141100003614052).

References

Ajami, N.K., Duan, Q.Y., Sorooshian, S., 2007. An integrated hydrologic Bayesian multimodel combination framework: confronting input, parameter, and model structural uncertainty in hydrologic prediction. Water Resour. Res. 43, 1–19.

Bogner, K., Kalas, M., 2008. Error-correction methods and evaluation of an ensemble based hydrological forecasting system for the Upper Danube catchment. Atmos. Sci. Lett. 9 (2), 95–102. http://dx.doi.org/10.1002/asl.180.

Bormann, K.J., Evans, J.P., McCabe, M.F., 2014. Constraining snowmelt in a temperature-index model using simulated snow densities. J. Hydrol. 517, 652–667.

Brown, J.D., Demargne, J., Seo, D.-J., Liu, Y., 2010. The Ensemble Verification System (EVS): a software tool for verifying ensemble forecasts of hydrometeorological and hydrologic variables at discrete locations. Environ. Model. Softw. 25, 854–872.

Clark, M., Gangopadhyay, S., Hay, L., Rajagopalan, B., Wilby, R., 2004. The Schaake Shuffle: a method for reconstructing space-time variability in forecasted precipitation and temperature fields. J. Hydrometeorol. 5, 243–262.

Coccia, G., Todini, E., 2011. Recent developments in predictive uncertainty assessment based on the model conditional processor approach. Hydrol. Earth Syst. Sci. 15 (10), 3253–3274.

Duan, Q. et al., 2007. Multi-model ensemble hydrologic prediction using Bayesian model averaging. Adv. Water Resour. 30 (5), 1371–1386. http://dx.doi.org/10.1016/j.advwatres.2006.11.014.

Gneiting, T., Raftery, A.E., 2005. Weather forecasting with ensemble methods. Science 310 (5746), 248–249.

Glahn, B., Peroutka, M., Wiedenfeld, J., 2009. MOS uncertainty estimates in an ensemble framework. Mon. Weather Rev. 137, 246–268.

Glahn, H.R., Lowry, D.A., 1972. The use of model output statistics (MOS) in objective weather forecasting. J. Appl. Meteorol. 11, 1203–1211.

Hamill, T.M., 2001. Interpretation of rank histograms for verifying ensemble forecasts. Mon. Weather Rev. 129, 550–560.

Hamill, T.M. et al., 2013. NOAA's second-generation global medium-range ensemble reforecast dataset. Bull. Am. Meteorol. Soc. 94, 1553–1565.

Kelly, K.S., Krzysztofowicz, R., 1997. A bivariate meta-Gaussian density for use in hydrology. Stochastic Hydrol. Hydraul. 11, 17–31.

Kirtman, B. et al., 2014. The North American Multi-Model Ensemble: Phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. Bull. Am. Meteorol. Soc. 95 (4), 585–601. http://dx.doi.org/10.1175/BAMS-D-12-00050.1.

Klein, W.H., Lewis, B.M., Enger, I., 1959. Objective prediction of five-day mean temperatures during winter. J. Meteorol. 16, 672–682.

Krzysztofowicz, R., 1997. Transformation and normalization of variates with specified distributions. J. Hydrol. 197, 286–292.

Li, H., Luo, L., Wood, E.F., Schaake, J., 2009a. The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting. J. Geophys. Res. Atmosp. 114, 1–10 (1984–2012).

Li, L., Xia, J., Xu, C.Y., Chu, J., Wang, R., 2009b. Analyse the sources of equifinality in hydrological model using GLUE methodology. Hydroinform. Hydrol. Hydrogeol. Water Resour. 4, 130–138.

Liu, Y. et al., 2013. Evaluating the predictive skill of post-processed NCEP GFS ensemble precipitation forecasts in China's Huai river basin. Hydrol. Process. 27, 57–74.

Lorenz, E.N., 1963. Deterministic nonperiodic flow. J. Atmos. Sci. 20, 130–141.

Ma, F. et al., 2014. An estimate of human and natural contributions to flood changes of the Huai River. Global Planet. Change 119, 39–50.

Ma, F. et al., 2016. Evaluating the skill of NMME seasonal precipitation ensemble predictions for 17 hydroclimatic regions in continental China. Int. J. Climatol. 36, 132–144. http://dx.doi.org/10.1002/joc.4333.

Montanari, A., Grossi, G., 2008. Estimating the uncertainty of hydrological forecasts: a statistical approach. Water Resour. Res. 44, W00B08.

Moriasi, D.N. et al., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models: Part I – A discussion of principles. J. Hydrol. 10, 282–290.

Olsson, J., Uvo, C.B., Foster, K., Yang, W., 2016. Technical Note: initial assessment of a multi-method approach to spring-flood forecasting in Sweden. Hydrol. Earth Syst. Sci. 20, 659–667.

Pappenberger, F. et al., 2013. Seamless forecasting of extreme events on a global scale. Clim. Land Surf. Changes Hydrol. 359, 3–10.

Pushpalatha, R., Perrin, C., Le Moine, N., Andreassian, V., 2012. A review of efficiency criteria suitable for evaluating low-flow simulations. J. Hydrol. 420, 171–182.

Saha, S. et al., 2014. The NCEP climate forecast system version 2. J. Clim. 27, 2185–2208.

Schaake, J. et al., 2007. Precipitation and temperature ensemble forecasts from single-value forecasts. Hydrol. Earth Syst. Sci. Discuss. 4, 655–717.

Schaefli, B., Gupta, H.V., 2007. Do Nash values have value? Hydrol. Process. 21, 2075–2080.

Shen, Y., Feng, M., Zhang, H., Gao, F., 2010. Interpolation methods of China daily precipitation data. J. Appl. Meteorol. Sci. 21, 279–286.

Swinbank, R., Kyouda, M., Buchanan, P., Froude, L., Hamill, T., Hewson, T., Keller, J., Matsueda, M., Methven, J., Pappenberger, F., Scheuerer, M., Titley, H., Wilson, L., Yamaguchi, M., 2005. The TIGGE project and its achievements. Bull. Amer. Meteor. Soc. http://dx.doi.org/10.1175/BAMS-D-13-00191.1.

Tao, Y. et al., 2014. An evaluation of post-processed TIGGE multimodel ensemble precipitation forecast in the Huai river basin. J. Hydrol. 519, 2890–2905.

The Ministry of Water Resources of People's Republic of China (MWR), 2002. Standard for Hydrological Information and Hydrological Forecasting (SL278—2002). Hydropower and Electrical Press, Beijing (In Chinese).

Todini, E., 2008. A model conditional processor to assess predictive uncertainty in flood forecasting. Int. J. River Basin Manage. 6, 123–137.

Van Steenbergen, N., Ronsyn, J., Willems, P., 2012. A non-parametric data-based approach for probabilistic flood forecasting in support of uncertainty communication. Environ. Model. Softw. 33, 92–105.

Wang, G., Xia, J., Chen, J., 2009. Quantification of effects of climate variations and human activities on runoff by a monthly water balance model: a case study of the Chaobai River basin in northern China. Water Resour. Res. 45, W00A11.

Wu, L.M. et al., 2011. Generation of ensemble precipitation forecast from single-valued quantitative precipitation forecast for hydrologic ensemble prediction. J. Hydrol. 399, 281–298.

Xia, J., Wang, G.S., Ye, A.Z., Niu, C.W., 2005. A distributed monthly water balance model for analyzing impacts of land cover change on flow regimes. Pedosphere 15, 761–767.

Xu, J., Ye, A., Mao, Y., Deng, X.X., 2014. Review of research and application of hydrologic ensemble forecast. South-to-North Water Transfers Water Sci. Technol. 1, 021.

Ye, A., Duan, Q., Zhan, C., Liu, Z., Mao, Y., 2013. Improving kinematic wave routing Land Model scheme in Community Land Model. Hydrol. Res. 44, 886–903.

Ye, A., Xia, J., Wang, G., 2005. Drainage network extraction and subcatchment delineation based on digital elevation model. J. Hydraul. Eng. 36 (5), 531–537.

Yuan, X., Wood, E.F., Liang, M., 2014. Integrating weather and climate prediction: toward seamless hydrologic forecasting. Geophys. Res. Lett. 41, 5891–5896.

Yuan, X., Wood, E.F., Luo, L.F., Pan, M., 2011. A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction. Geophys. Res. Lett. 38.

Yuan, X., Wood, E.F., Roundy, J.K., Pan, M., 2013. CFSv2-based seasonal hydroclimatic forecasts over the conterminous United States. J. Clim. 26, 4828–4847.

Zhao, L., Duan, Q., Schaake, J., Ye, A., Xia, J., 2011. A hydrologic post-processor for ensemble streamflow predictions. Adv. Geosci. 29, 51–59.