

Chapter 8

Uncertainty Analysis of Hydrologic Forecasts Based on Copulas



8.1 Introduction

Hydrologic forecasting is a crucial non-structural flood mitigation measure and provides an essential basis for flood warning, flood control and reservoir operation (Guo et al. 2004; Calvo and Savi 2009; Chen et al. 2014a; Zhang et al. 2015; Fan et al. 2016; Liu et al. 2017; Wu et al. 2017). Forecasting models that are widely used at present are typically deterministic, and model outputs are provided to users in the form of deterministic values (Chen and Yu 2007; Coccia and Todini 2011; Ma et al. 2013; Bergstrand et al. 2014; Li et al. 2014). However, a hydrological forecasting model is only a simulation of the real hydrological processes and is therefore imperfect and not precise (Ravines et al. 2008; Wetterhall et al. 2013). These models accept hydrological input, meteorological input, etc., and utilize conceptualized model parameters; and these complex factors cause inevitably uncertainties in the hydrologic forecasts (Freer et al. 1996; Montanari 2007; Montanari and Grossi 2008; Renard et al. 2010; Chen et al. 2014b). The principle of rational decision-making under uncertainty indicates that when a deterministic forecast turns out to be wrong, the consequences will probably be worse than a situation where no forecast is available (Krzysztofowicz 1999; Wetterhall et al. 2013; Ramos et al. 2013). A rational decision maker who wants to make optimal decisions should therefore take forecast uncertainty explicitly into account (Verkade and Werner 2011; Ramos et al. 2013). Therefore, quantitative assessment of inherent uncertainty is a critical issue. Hydrologic forecasting services are trending toward providing users with probabilistic forecasts, in place of traditional deterministic forecasts.

The transition from a deterministic forecast to a probabilistic forecast is based on quantification of the uncertainty inherent in the deterministic forecast. The Bayesian Forecasting System (BFS) proposed by Krzysztofowicz (1999) provides a general framework to produce probabilistic forecasts via any deterministic hydrological model. Various probabilistic forecasting systems suited to different purposes have

been developed within this framework (Reggiani and Weerts 2008; Calvo and Savi 2009; Biondi et al. 2010; Weerts et al. 2011; Sikorska et al. 2012; Pokhrel et al. 2013).

In the BFS, the total uncertainty is decomposed into input uncertainty and hydrological uncertainty. The hydrological uncertainty processor (HUP) is a component of the BFS that quantifies the hydrological uncertainty and produces probabilistic forecast under the hypothesis that there is no input uncertainty (Krzysztofowicz and Kelly 2000). Through Bayes' theorem, the HUP combines a prior distribution, which describes the natural uncertainty about the realization of a hydrologic process, with a likelihood function which quantifies the uncertainty in model forecasts, and outputs a posterior distribution, conditional upon the deterministic forecasts. This posterior distribution provides a complete characterization of uncertainty, including quantiles, prediction intervals and probabilities of exceedance for specified thresholds which are needed by rational decision makers and information providers who want to extract forecast products for their customers.

The HUP can be implemented in many ways, as different mathematical models for prior distribution and likelihood function can be developed. Krzysztofowicz and Kelly (2000) introduced a meta-Gaussian HUP, which was developed by converting both original observations and model forecasts into a Gaussian space by using the Normal Quantile Transform (NQT). This meta-Gaussian HUP has been widely used by many researchers in the fields of hydrology and meteorology (Chen and Yu 2007; Biondi et al. 2010; Biondi and De Luca 2013; Chen et al. 2013a).

The prior density and likelihood function are conditional probability distributions. It is well known that copula function has an outstanding capability to model joint distributions and gives flexibility in choosing an arbitrary marginal distribution (e.g. non-Gaussian form), nonlinear and heteroscedastic dependence structure. The conditional probability distribution can be expressed in the explicit form using copula function (Favre et al. 2004; Nelsen 2006; Zhang and Singh 2006, 2007a, b, c; Genest and Favre 2007; Bárdossy and Li 2008; Chen et al. 2010; Zhang et al. 2011, 2012). These advantageous characteristics of the copula function motivate us to develop the prior distribution and likelihood function models in the original space directly without a data transformation procedure into Gaussian space. Liu et al. (2017) proposed a post-processor based on copula function for deterministic forecast model to produce probabilistic forecasts within the general framework of the HUP.

Despite the tremendous amount of resources invested in developing more hydrologic models, no one can convincingly claim that any particular model in existence today is superior to other models for all type of applications and under all conditions (Wu et al. 2015; Liu et al. 2016; Ba et al. 2017). Different models are capable of capturing different aspects of the hydrologic processes. The uncertainty of each model arises from parameters calibration, the design of the model structure, and input measurements, which partially brings underlying imprecise influence (Göttinger and 2008; Li et al. 2011; Hemri et al. 2015). One of the primary

techniques to reflect different uncertainties in hydrological forecasts is to create an ensemble of forecast trajectories (Seo et al. 2006; Madadgar et al. 2014).

The Bayesian Model Average (BMA) method introduced by Raftery et al. (2005) follows a statistical technique to combine the advantages of different models. Different from other multi-model methods, the BMA method presents a more realistic description of predictive uncertainty, since the BMA predictive variance can be decomposed into two components: between-model variability and within-model variability (Ajami et al. 2007). The BMA method is a statistical procedure that infers consensus predictions by weighing individual predictions based on their probabilistic likelihood measures, with the better performing predictions receiving higher weights than, the worse performing ones. The method has been explored to improve both the accuracy and reliability of streamflow predictions (Vrugt and Robinson 2007; Liang et al. 2011). Duan et al. (2007) concluded that the combination of multi-model ensemble strategies using the BMA framework could quantify statements on prediction uncertainty and significantly improve verification performances. Zhou et al. (2016) compared the mean prediction of BMA with its individual parameter transfer method (physical similarity approach) and demonstrated that the probabilistic predictions of BMA could reduce the uncertainty with a significant degree. Nevertheless, the standard BMA method imposes lots of pseudo variation requirements, and this influences precise understanding of data variations, which gives rise to further development of this theoretical research (Madadgar and Moradkhani 2014).

Klein et al. (2016) used a mixture of marginal density distribution to estimate the predictive uncertainty of hydrologic multi-model ensembles by using pair-copula construction. Similar researches show that copula technique is an effective tool for reflecting the unclear and complex relationships because it can flexibly choose the arbitrary type of the marginal distributions instead of Gaussian distribution (Carreau and Bouvier 2016; Khajehei and Moradkhani 2017). According to the promising results of using copula functions in post-processing of hydrologic forecasts, Madadgar and Moradkhani (2014) firstly integrated copula functions with BMA to estimate the posterior distribution and found that Copula-BMA (CBMA) is an effective post-processor to relax any assumption on the distribution of conditional probability density function (PDF). The CBMA not only displayed better deterministic skill than BMA but also confirmed the impact of posterior distribution in calculating the weights of individual models by EM algorithm. Results indicated that the predictive distributions are more accurate and reliable. It is also shown that the post-processed forecasts have better correlation with observation after CBMA application. The CBMA method in the meteorological application does not need to assume the shape of the posterior distribution and leaves out the data-transformation procedure and demonstrates that predictive distributions are less bias and more confident with small uncertainty (Möller et al. 2013). Inspired by the ideas of Madadgar and Moradkhani (2014), a general framework of the combination of copula Bayesian processor with BMA (CBP-BMA) is proposed by He et al. (2018).

8.2 Hydrologic Uncertainty Processor Based on Copula Function

8.2.1 Hydrologic Uncertainty Processor

Let predict and H be the observed discharge whose realization h is being forecasted. Let estimator S be the output discharge generated by a corresponding deterministic forecast model whose realization s constitutes a point estimate of H . Let random variable H_0 represent the observed discharge at the time $n = 0$ when the forecast is prepared; then H_n ($n = 1, 2, \dots, N$) is the observed discharge at lead time n ; and S_n ($n = 1, 2, \dots, N$) is the corresponding deterministic forecast discharge at lead time n . What the rational decision maker then needs is not a single number s_n , but the distribution function of predictand H_n , conditional on $H_0 = h_0$ and $S_n = s_n$. The purpose of the HUP is to supply such a conditional distribution function through Bayesian revision (Liu et al. 2017).

The Bayesian procedure for information revision of uncertainty involves two steps. First, the expected conditional density function of deterministic forecast discharge S_n given that $H_0 = h_0$ is derived via the total probability law:

$$\kappa_n(s_n|h_0) = \int_{-\infty}^{+\infty} f_n(s_n|h_0, h_n) \cdot g(h_n|h_0) dh_n \quad (8.1)$$

Second, the posterior density function of predictand H_n conditional on a deterministic forecast $S_n = s_n$ and observed discharge at the forecasting time $H_0 = h_0$, is derived via Bayes' theorem (Krzysztofowicz and Kelly 2000):

$$\phi_n(h_n|h_0, s_n) = \frac{f_n(s_n|h_0, h_n) \cdot g_n(h_n|h_0)}{\kappa_n(s_n|h_0)} \quad (8.2)$$

In concept, Bayes' theorem revises the prior density function $g_n(h_n/h_0)$, which characterizes the prior uncertainty about H_n , given observed discharge at the forecasting time $H_0 = h_0$. The extent of the revision is determined by the likelihood function $f_n(s_n/h_0, h_n)$, which characterizes the degree to which $S_n = s_n$ reduces the uncertainty about H_n . The result of this revision is the posterior density function $\Phi_n(h_n/h_0, s_n)$, which quantifies the uncertainty about H_n that remains after the deterministic forecast model generates forecast $S_n = s_n$.

8.2.2 Meta-Gaussian HUP

As we can see from Eq. 8.2, the posterior density depends on the prior density function and likelihood function. The most widely used technique to describe the

prior density and likelihood functions is the meta Gaussian model. In this model, the NQT method (Bogner et al. 2012) is applied to convert both actual flow H_n and predicted flow S_n into the Gaussian space. Then the transformed $\{h_n/h_0\}$ and $\{s_n/h_0, h_n\}$ are assumed to be linear and normally distributed. Subsequently, linear regression method is then employed to determine the posterior density of H_n in the transformed Gaussian space, from which the posterior density function of H_n in the original space can be found. For the sake of the following comparison, the detailed procedure is presented as follows (Krzysztofowicz and Kelly 2000).

8.2.2.1 Normal Quantile Transform

Specifying and determining marginal distributions of the actual flow H_0 , H_n and predicted flow $\{S_n: n = 1, \dots, N\}$ is the first step. The actual flows $\{H_n: n = 0, 1, \dots, N\}$ are considered as random variables. Given only such a record, there is usually no basis for assigning a probability distribution to flow H_n that differs from the distribution assigned to flow H_0 , for any $n = 1, 2, \dots, N$ within a few days. In other words, there is no a statistical difference between these $1 + N$ flow series (Koutsoyiannis and Montanari 2015). Therefore, we hold the opinion that the variables H_n follow the same marginal cumulative distribution functions (CDF) with H_0 , and thus only the CDF of H_0 needed to be fitted. The predicted flows $\{S_n: n = 1, \dots, N\}$ are considered as different random variables and different CDFs needed to be fitted for variable S_n .

Let Γ and $\bar{\Lambda}_n$ be the CDF of H_0 and S_n with corresponding densities γ and $\bar{\lambda}_n$, respectively. The NQT of a variate is defined as a composition of the inverse of the standard normal distribution Q , and the CDF of the variate is assumed to be strictly increasing. The transformed variates are

$$W_n = Q^{-1}[\Gamma(H_n)], \quad n = 0, 1, \dots, N \quad (8.3)$$

$$X_n = Q^{-1}[\bar{\Lambda}_n(S_n)], \quad n = 0, 1, \dots, N \quad (8.4)$$

where W_n and X_n are the normal quantiles of H_n and S_n , respectively. Q^{-1} is the inverse function of Q .

8.2.2.2 Modeling in the Transformed Space

(1) Prior density

The model for the prior density rests on the assumption that the actual river discharge process in the transformed space is governed by the normal-linear equation

$$W_n = cW_{n-1} + \Xi \quad (8.5)$$

where c is a parameter and Ξ is a variate stochastically independent of W_{n-1} and normally distributed with mean zero and variance $1 - c^2$. Consequently, the conditional mean and variance are

$$E(W_n | W_{n-1} = w_{n-1}) = cw_{n-1} \quad (8.6)$$

$$\text{Var}(W_n | W_{n-1} = w_{n-1}) = 1 - c^2 \quad (8.7)$$

The prior density for lead time n takes the form

$$g_{Q_n}(w_n | w_0) = \frac{1}{(1 - c^{2n})^{1/2}} \cdot q \left[\frac{w_n - c^n w_0}{(1 - c^{2n})^{1/2}} \right] \quad (8.8)$$

where q denotes the standard normal density and subscript Q_n denotes a density in the space of transformed variants.

(2) Likelihood function

The model for the likelihood function rests on the assumption that the stochastic dependence between the transformed variate is governed by the normal-linear equation

$$X_n = a_n W_n + d_n W_0 + b_n \Theta_n \quad (8.9)$$

where a_n , b_n and d_n are parameters and Θ_n is a stochastically independent variate of (W_n, W_0) and normally distributed with mean zero and variance δ_n^2 . Consequently, the conditional mean and variance are

$$E(X_n | W_n = w_n, W_0 = w_0) = a_n w_n + d_n w_0 + b_n \quad (8.10)$$

$$\text{Var}(X_n | W_n = w_n, W_0 = w_0) = \delta_n^2 \quad (8.11)$$

The conditional density function is

$$f_{Q_n}(x_n | w_n, w_0) = \frac{1}{\delta_n} \cdot q \left(\frac{x_n - a_n w_n - d_n w_0 - b_n}{\delta_n} \right) \quad (8.12)$$

(3) Posterior density

The posterior density derived from the prior density and likelihood function takes the form as follows

$$\varphi_{Q_n}(w_n|x_n, w_0) = \frac{1}{T_n} q\left(\frac{w_n - A_n x_n - D_n w_0 - B_n}{T_n}\right) \quad (8.13)$$

In which $A_n = \frac{a_n t_n^2}{a_n^2 t_n^2 + \delta_n^2}$, $B_n = \frac{-a_n b_n t_n^2}{a_n^2 t_n^2 + \delta_n^2}$, $D_n = \frac{c_n \delta_n^2 - a_n d_n t_n^2}{a_n^2 t_n^2 + \delta_n^2}$, $T_n^2 = \frac{t_n^2 \delta_n^2}{a_n^2 t_n^2 + \delta_n^2}$, and $t_n^2 = 1 - c^{2n}$.

8.2.2.3 Posterior Density and Distribution in the Original Space

With all densities in the transformed space belonging to the Gaussian family, all densities in the original space belong to the meta-Gaussian family. The meta-Gaussian posterior density of actual river discharge conditional on model output discharge $S_0 = s_0$ and observed river discharge $H_0 = h_0$ takes the form

$$\begin{aligned} \phi_n(h_n|s_n, h_0) &= \frac{\gamma(h_n)}{T_n \cdot q\{Q^{-1}[\Gamma(h_n)]\}} \\ &\cdot q\left\{\frac{Q^{-1}[\Gamma(h_n)] - A_n \cdot Q^{-1}[\bar{\Lambda}_n(s_n)] - D_n \cdot Q^{-1}[\Gamma(h_0)] - B_n}{T_n}\right\} \end{aligned} \quad (8.14)$$

The corresponding meta-Gaussian posterior distribution takes the form

$$\Phi_n(h_n|s_n, h_0) = Q\left\{\frac{Q^{-1}[\Gamma(h_n)] - A_n \cdot Q^{-1}[\bar{\Lambda}_n(s_n)] - D_n \cdot Q^{-1}[\Gamma(h_0)] - B_n}{T_n}\right\} \quad (8.15)$$

8.2.3 Copula-Based HUP

Copula function is an effective tool used to develop prior distribution and likelihood function models, in which the predictand and the deterministic forecasts are allowed to have distribution functions of any form, along with nonlinear and heteroscedastic dependence structure. Therefore, it can be implemented in the original space directly without a data transformation procedure into Gaussian space. The copula function and theory have been introduced in detail in Chap. 2.

8.2.3.1 Prior Density

The prior CDF of H_n given $H_0 = h_0$ can be expressed as

$$G_n(h_n|h_0) = P(H_n \leq h_n | H_0 = h_0) \quad (8.16)$$

where $G_n(h_n|h_0)$ is the conditional CDF, and P is the non-exceedance probability.

The prior density function $g_n(h_n|h_0)$ is the corresponding probability density function (PDF) of $g_n(h_n|h_0)$ and can be defined as

$$g_n(h_n|h_0) = \frac{dG_n(h_n|h_0)}{dh_n} \quad (8.17)$$

Let H_0 and H_n be random variables with marginal CDFs, $U_1 = F_H(H_0)$ and $U_2 = F_H(H_n)$. Then, U_1 and U_2 are uniformly distributed random variables; and u_1 denotes a specific value of U_1 , and u_2 denotes a specific value of U_2 .

Using the copula function, the joint CDF is expressed by $G_n(h_n, h_0) = C(F_H(h_0), F_H(h_n)) = C(u_1, u_2)$

The conditional CDF $G_n(h_n|h_0)$ and PDF $g_n(h_n|h_0)$ can be rewritten as follows (Zhang and Singh 2006)

$$G_n(h_n|h_0) = P(U_2 \leq u_2 | U_1 = u_1) = \frac{\partial C(u_1, u_2)}{\partial u_1} \quad (8.18)$$

$$g_n(h_n|h_0) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2} \cdot \frac{du_2}{dh_n} = c(u_1, u_2) \cdot f_H(h_n) \quad (8.19)$$

where $c(u_1, u_2)$ is the density function of $C(u_1, u_2)$, and $c(u_1, u_2) = \partial^2 C(u_1, u_2) / \partial u_1 \partial u_2$; $f_H(h_n)$ is the PDF of H_n . Equation 8.19 is the expression of the prior PDF.

8.2.3.2 Likelihood Function

It is considered that S_n is a random variable with marginal CDF $u_3 = F_{S_n}(s_n)$ and PDF $f_{S_n}(s_n)$. The conditional CDF of S_n given $H_0 = h_0$ and $H_n = h_n$ can be expressed as

$$F_n(s_n|h_0, h_n) = P(S_n \leq s_n | H_0 = h_0, H_n = h_n) \quad (8.20)$$

where $F_n(s_n|h_0, h_n)$ is the conditional CDF.

The corresponding PDF of $F_n(s_n|h_0, h_n)$ is defined as

$$f_n(s_n|h_0, h_n) = \frac{dF_n(s_n|h_0, h_n)}{ds_n} \quad (8.21)$$

Using the copula function, the joint CDFs of H_0 , H_n and S_n , denoted as $F_n(h_0, h_n, s_n)$ can be expressed as $F_k(h_0, h_n, s_n) = C(F_{H_0}(h_0), F_{H_n}(h_n), F_{S_n}(s_n)) =$

$C(u_1, u_2, u_3)$. Thus, the conditional CDF $F_n(s_n|h_0, h_n)$ and PDF $f_n(s_n|h_0, h_n)$ are rewritten as follows, (Zhang and Singh 2007c)

$$F_n(s_n|h_0, h_n) = P(U_3 \leq u_3 | U_1 = u_1, U_2 = u_2) = \frac{\partial^2 C(u_1, u_2, u_3) / \partial u_1 \partial u_2}{c(u_1, u_2)} \quad (8.22)$$

$$f_n(s_n|h_0, h_n) = \frac{1}{c(u_1, u_2)} \cdot \frac{\partial^3 C(u_1, u_2, u_3)}{\partial u_1 \partial u_2 \partial u_3} \cdot \frac{du_3}{ds_n} = \frac{c(u_1, u_2, u_3)}{c(u_1, u_2)} \cdot f_{S_n}(s_n) \quad (8.23)$$

where $c(u_1, u_2, u_3) = \partial^3 C(u_1, u_2, u_3) / \partial u_1 \partial u_2 \partial u_3$ is the density function of $C(u_1, u_2, u_3)$. From another point of view, given $H_0 = h_0$ and $S_0 = s_0$, the likelihood function of H_n can be calculated by Eq. 8.23.

8.2.3.3 Posterior Density

Substitute Eqs. 8.19 and 8.23 to Eqs. 8.1 and 8.2, then the posterior density function of H_n can be rewritten as follows

$$\phi_n(h_n|h_0, s_n) = \frac{c(u_1, u_2, u_3)}{\int_0^1 c(u_1, u_2, u_3) du_2} \cdot f_H(h_n) \quad (8.24)$$

For fixed realizations $H_0 = h_0$ and $S_n = s_n$, u_1 and u_3 are constants, while u_2 varies from 0 to 1. Since the denominator $\int_0^1 c(u_1, u_2, u_3) du_2$ cannot be obtained directly by an analytic method, the Monte Carlo sampling technique (Yu et al. 2014; Xiong et al. 2014) is applied by following steps: (1) Generate M random numbers u_2 from uniform distribution $U(0, 1)$; (2) Compute the value of $C(u_1, u_2, u_3)$; (3) The mean value of the M calculated $C(u_1, u_2, u_3)$ equals to the definite integral $\int_0^1 c(u_1, u_2, u_3) du_2$ approximately (Robert and Casella 2013; Kroese et al. 2013). Subsequently, the posterior density function $\phi_n(h_n|h_0, s_n)$ can also be estimated.

8.2.3.4 Candidate Marginal Distributions and Trivariate Copulas

The main purpose of this study aims to extrapolate the extreme events far beyond the observations. The probability distribution of daily flows refers to the flow duration curve, which gives a summary of flow variability at a site and is interpreted as a relationship between any discharge value and the percentage of time that this discharge is equaled or exceeded during a given period (Vogel and Fennessey 1994; Castellarin et al. 2004; Shao et al. 2009). Flow-duration curve has been widely used by engineers and hydrologists around the world in numerous applications, such as

hydropower generation, inflow forecasting, and designing of irrigation systems (Vogel and Fennessey 1995; Yokoo and Sivapalan 2011; Gottschalk et al. 2013).

Even though flow-duration curve can be defined and constructed for different time scales, such as daily, weekly or monthly stream flows, our study will focus on a daily flow-duration curve. If the daily streamflow is assumed to be a random variable, the flow-duration curve may also be viewed as the complement of the cumulative distribution function used in hydrologic frequency analysis when identifying the percentage of time with probability (Castellarin et al. 2004). As a consequence, the flow-duration curve is also a very practical tool used to describe hydrological regimes and represents the relationship between magnitude and frequency of flow (Vogel and Fennessey 1995; Liucci et al. 2014; Xiong et al. 2015).

Six commonly used distributions in hydrology, namely Normal, GMA, Gumbel, P-III, Log-Normal and Log-Weibull, are selected as candidate models for H_0 and S_n ($n = 1, \dots, N$). These univariate probability distributions are summarized in Table 1.1 of Chap. 1. L-moment method is used to estimate the distribution parameters for given data series (Hosking 1990). The Kolmogorov-Smirnov statistic D is used to measure the goodness of fit between the hypothesized distribution and the empirical distribution (Tsai et al. 2001; Arya et al. 2010). In this study, the 95% confidence level is selected to reject or accept a fitted distribution. The probability distribution which provides the minimum D value is chosen as the best fitting distribution.

To estimate of the posterior density functions expressed in Eq. 8.24, three-dimension joint distributions of H_0 , H_n and S_n are needed to be constructed. The symmetric copulas are not considered because the dependence among the three variables pairs (H_0, H_n) , (H_0, S_n) and (H_n, S_n) are not the same, which will be tested against data for the case study. Hence, we use three widely used asymmetric trivariate Archimedean copulas, namely Gumbel-Hougaard, Frank and Clayton as candidates. These three trivariate Archimedean copulas are described in Table 2.2 of Chap. 2. Dependence parameters of the trivariate copula functions are estimated using the maximum pseudo-likelihood method (Zhang and Singh 2007b, c; Chen et al. 2010). The RMSE is used to measure the goodness of fit of the copula distribution (Zhang and Singh 2007a). The copula which has the smallest *RMSE* value is preferred.

8.2.4 Evaluation Criteria

8.2.4.1 Performances of Deterministic Forecasts

Two widely applied criteria, namely Nash-Sutcliffe efficiency (*NSE*) and Relative Error (*RE*) are adopted to evaluate the performance of the deterministic forecast model (Xiong and Guo 1999; Liu et al. 2016).

(1) Nash-Sutcliffe efficiency

The first criterion is the Nash-Sutcliffe efficiency (*NSE*) coefficient (Nash and Sutcliffe 1970) which is defined by

$$NSE = \left[1 - \frac{\sum_{t=1}^T (h_t - s_t)^2}{\sum_{t=1}^T (h_t - \bar{h})^2} \right] \times 100\% \quad (8.25)$$

where t is the time step, T is the total number of time steps; h_t and s_t are the simulated and observed discharges at time t , and \bar{h} is the mean value of the observed discharge. Nash-Sutcliffe efficiency can range from $-\infty$ to 1. An efficiency of 1 ($NSE = 1$) corresponds to a perfect match of simulated discharge to the observed data. An efficiency of 0 ($NSE = 0$) indicates that the model predictions are as accurate as the mean of the observed data, whereas an efficiency less than zero ($NSE < 0$) occurs when the observed mean is a better predictor than the model. Essentially, the closer the model efficiency is to 1, the more accurate the model is.

(2) Relative error

The second criterion used is the relative error (*RE*) of the total runoff amount fit between the observed and simulated discharge series, defined as (Xiong and Guo 1999)

$$RE = \left[1 - \frac{\sum_{t=1}^T (h_t - s_t)}{\sum_{t=1}^T h_t} \right] \times 100\% \quad (8.26)$$

RE represents a systematic error of water balance simulation. A value of *RE* closes to zero indicates a good agreement between observed and simulated runoff volume. In this study, we rank *NSE* as the primary criterion, while *RE* is an auxiliary criterion. Only when simulated discharge series yield the same (higher) *NSE* value, the one with the smaller *RE* value is preferred. Otherwise, the simulation with smaller *RE* value does not reveal any superiority (Liu et al. 2016). For instance, the model with all simulated discharges equal to the mean of observed values can easily provide $RE = 0$. Unfortunately, in this case, the $NSE = 0$, which clearly means an undesired simulation.

8.2.4.2 Performances of Probabilistic Forecasts

The probabilistic forecast technique is expected to provide (a) accurate forecast probabilities, further on named reliability; and (b) narrow forecast intervals, further on the named resolution. Several methods, e.g., predictive quantile-quantile (QQ) plot, α -index and π -index have been proposed in the literatures to evaluate probabilistic forecasts (see e.g. Gneiting et al. 2007; Laio and Tamea 2007; Thyer et al. 2009; Engeland et al. 2010; Renard et al. 2010; Madadgar et al. 2014; Smith et al. 2015) and are used in this study.

(1) Predictive QQ plot

The predictive quantile-quantile (QQ) plot provides an overall assessment of whether the total predictive uncertainty is consistent with the observations. This requires a diagnostic approach that compares a time-varying distribution (the predictive distribution at all times t) to a time series of observations (Thyer et al. 2009; Evin et al. 2014). The predictive QQ plot provides a simple, intuitive and informative summary of the performance of probabilistic prediction frameworks (Gneiting et al. 2007; Laio and Tamea 2007).

The predictive QQ plot is constructed as follows: Let F_t be the CDF of the predictive distribution of runoff at time t , and h_t the corresponding observed runoff. If the hypotheses in the calibration framework are consistent with the data, the observed value h_t should be consistent with the distribution F_t . Hence, under the assumption that the observation h_t is a realization of the predictive distribution, the p -value $F_t(h_t)$ is a realization of a uniform distribution on $[0,1]$. The predictive QQ plot compares the empirical CDF of the sample of p values $F_t(h_t)$ ($t = 1, \dots, T$) with the CDF of a uniform distribution to assess whether the hypotheses are consistent with the observations.

As illustrated in Fig. 8.1, the predictive QQ plot can be interpreted as follows (Thyer et al. 2009): (1) if all points fall on the 1:1 line, the predicted distribution agrees perfectly with the observations; (2) If the observed p values cluster around the mid-range (i.e., a low slope around theoretical quantile 0.4–0.6), the predictive uncertainty is overestimated; (3) If the observed p values cluster around the tails (i.e., a high slope around theoretical quantile 0.4–0.6), the predictive uncertainty is underestimated; (4) If the observed p values at the theoretical median are higher/

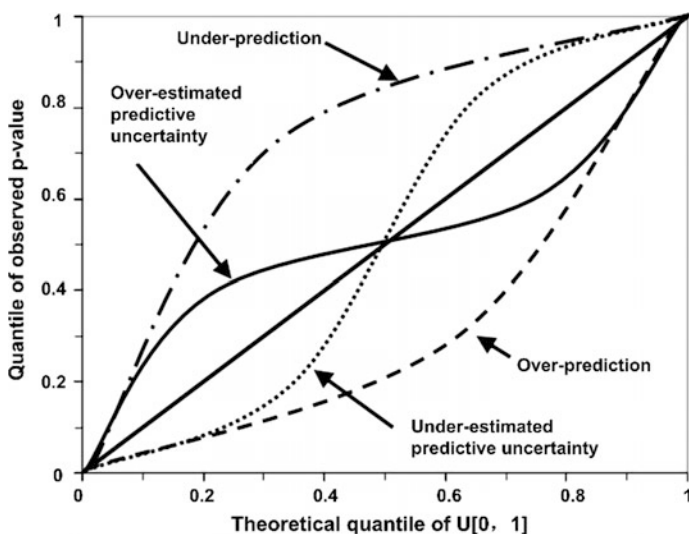


Fig. 8.1 Interpretation of the predictive QQ plot

lower than the theoretical quantiles, the modeled predictions systematically under/over predict the observed data.

Other metrics are the supportive quantitative scores derived from the predictive QQ plot (Laio and Tamea 2007; Thyer et al. 2009; Madadgar et al. 2014). The metrics α -index assesses the reliability of forecasts, and π -index indicates the resolution (precision, sharpness) of the predictive distribution (PD).

(2) Reliability

Reliability means that the forecast should be well calibrated. This can be checked graphically: deviations from the bisector (the 1:1 line) denote interpreted deficiencies (see Fig. 8.1). To simplify the comparison of QQ plots, it is summarized using an index that quantifies the reliability of the PD (Renard et al. 2010; Madadgar et al. 2014):

$$\alpha\text{-index} = 1 - \frac{2}{T} \sum_{t=1}^T [|q^{em}(p_t) - q^{th}(p_t)|] \quad (8.27)$$

where p_t is the observed p -value at time t ; $q^{em}(h_t)$ is the empirical quantile of p_t , $q^{th}(h_t)$ is the theoretical quantile of p_t obtained from the uniform distribution $U[0, 1]$; T is the number of p_t values.

The α -index measures the closeness of quantile plot of the observations to the corresponding uniform quantiles and reflects the overall reliability of the PD. According to Thyer et al. (2009), as the area between the empirical CDF of the observed p -values and the CDF of the uniform distribution in the predictive QQ plot becomes larger, the value of α -index decreases towards zero. It varies between 0 (worst reliability) and 1 (perfect reliability).

(3) Resolution

“Resolution” denotes the sharpness (effectively, the “average precision”) of the PD. Note that two inferences can both yield reliable PDs, but with different resolutions. Sharpness refers to the spread of the forecast PDFs and is a property of the predictions only. The more concentrated the forecast PDF, the sharper the forecast, and the sharper the better, subject to calibration (Gneiting et al. 2005). In this paper, the resolution is quantified by π -index defined as the average relative precision of the predictions (Renard et al. 2010; Madadgar et al. 2014):

$$\pi\text{-index} = \frac{1}{T} \sum_{t=1}^T \frac{E[H_t]}{Sdev[H_t]} \quad (8.28)$$

where $E[H_t]$ and $Sdev[H_t]$ are the expected value and standard deviation of H_t obtained from the predictive distribution at time t .

Greater value of π -index indicates greater resolution (lower uncertainty) of forecasts. However, comparison of sharpness may not be a meaningful approach when the employed methods do not primarily perform equally in the α -index

metric. Assuming that precision has lower priority than reliability, given similar forecast reliability, the method with greater resolution (lower uncertainty) is preferred; otherwise, the method with higher resolution does not reveal any superiority. Most of literature rank reliability as the primary criterion, while sharpness is secondary to reliability (Madadgar et al. 2014).

(4) Continuous rank probability score

The goal of probabilistic forecasting is to maximize the sharpness of the forecast PDFs subject to calibration. However, the trade-off between reliability and sharpness have been discussed in previous researches (Xiong et al. 2009; Li et al. 2010a; Kasiviswanathan et al. 2013), which show that these two desirable objectives could not be achieved simultaneously. It is not adequate to judge the performances of probabilistic forecasts only by reliability or sharpness. The continuous rank probability score (*CRPS*) is a standard measure that combines reliability and sharpness (Hersbach 2000; Gneiting et al. 2005) and is used for selecting the preferred model.

The *CRPS* measures the average distance between the predicted and the observed CDFs over the entire period. It is the integral of the Brier scores at all possible threshold values r for the continuous predictand (Hersbach 2000). Specifically, if F is the predictive CDF and h_t is the verifying observation, the *CRPS* is defined as (Hersbach 2000; Gneiting et al. 2007; Pappenberger et al. 2015)

$$CRPS = \frac{1}{T} \sum_{t=1}^T \int_{-\infty}^{+\infty} [F_t(r) - H_s(r - h_t)]^2 dr \quad (8.29)$$

where $H_s(r - h_t)$ denotes the Heaviside step function and takes the value 0 when $r < h_t$ and the value one otherwise.

For a deterministic forecast system, the *CRPS* reduces to the mean absolute error (*MAE*). Thus, the *CRPS* is sometimes interpreted as a generalized version of the *MAE* (Zhao et al. 2015). This is an advantage of *CRPS* and consequently allows the comparison of deterministic and probabilistic forecasts (Gneiting et al. 2007; Pappenberger et al. 2015). The smaller the *CRPS* value is, the better the prediction performance. Its minimal value of zero is only achieved in the case of a perfect deterministic forecast.

8.2.5 Case Studies

8.2.5.1 Study Area and Data

Three Gorges Reservoir (TGR) is a vitally important and back-bone project in the development and harnessing of the Yangtze River in China. The annual average discharge and runoff volume at the dam site are 14,300 m³/s and 4510 × 10⁸ m³, respectively. The total storage capacity of the TGR is 393 × 10⁸ m³, of which

$221.5 \times 10^8 \text{ m}^3$ is flood control storage. The reservoir has a surface area of about 1080 km^2 , an average width of about 1100 m, a mean depth of about 70 m and a maximum depth near the dam of about 170 m. With all the profiles being narrow and deep, the TGR retains the long narrow belt shape of the original river section and is a typical river channel-type reservoir.

As shown in Fig. 8.2, the intervening basin of TGR has a catchment area of $55,907 \text{ km}^2$, about 5.6% of the upstream Yangtze River basin. There are 40 rainfall gauged stations in the intervening basin and two hydrological stations (Cuntan and Wulong), which control the upstream inflow and tributary inflow, respectively. The data set for TGR inflow forecasting includes the daily runoff data of the Cuntan, Wulong and Yichang hydrological stations, arithmetic mean of observed rainfall data in the intervening basin during the flood period (June 1–September 30) from 2003 to 2009. The period 2003–2007 is used for deterministic forecast model calibration and 2008–2009 is used for validation (Li et al. 2010b; Chen et al. 2015).

8.2.5.2 Deterministic Inflow Forecasts of the TGR

The inflow of TGR consists of three components, i.e., the main upstream inflow, the tributary inflow from the Wu River, and the lateral flow from the TGR intervening basin as shown in Fig. 8.2. A multiple-input single-output linear systematic model is chosen for the inflow forecasting of the TGR (Liang et al. 1992). The total inflow to the TGR can be expressed by the following equation

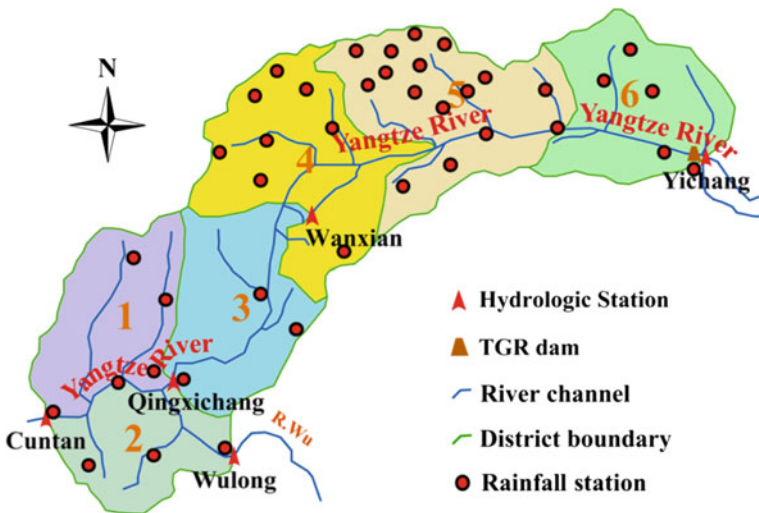


Fig. 8.2 Sketch map of the TGR’s intervening basin

$$\widehat{Q}_t = A \sum_{j=1}^{m_1} R_{t-j+1}^{(1)} h_j^{(1)} + \sum_{j=1}^{m_2} R_{t-j+1}^{(2)} h_j^{(2)} \quad (8.30)$$

where $R_j^{(1)}$ is the lateral flow from the TGR intervening basin which is calculated via the Xinanjiang model (Zhao 1992). $R_j^{(2)}$ is the upstream inflow (inflow at Wulong is added to the inflow at Cuntan). A is the area of the TGR intervening basin, m_1, m_2 are the memory length of the system corresponding to $R_j^{(1)}$ and $R_j^{(2)}$, $h_j^{(1)}$ and $h_j^{(2)}$ are the j th ordinates of the pulse response functions relating inputs $R_j^{(1)}$ and $R_j^{(2)}$, which are calculated by the Nash model as follows

$$h_j^{(1)} = \frac{1}{T} \int_{(j-1)T}^{jT} [S_i(t) - S_i(t-T)]/T dt \quad (i = 1, 2) \quad (8.31)$$

$$S_i(t) = \int_0^t \frac{1}{NK_i \Gamma(NK_i)} e^{-(\tau/NK_i)} (\tau \backslash NK_i)^{NK_i-1} d\tau \quad (i = 1, 2) \quad (8.32)$$

where $S_i(t)$ is the step response function of the i th input, N_i and NK_i are the parameters, and T is the time-step. $\Gamma(\bullet)$ is the gamma function.

The Xinanjiang model was developed in the middle 1970s for forecasting flows in the Xinanjiang reservoir, China. The model has been widely applied for flood forecasting in a large number of basins all over the world, especially in China. Until now, this model is the most popular rainfall-runoff hydrologic model in China for streamflow forecasting in humid and semi-humid areas. Its main feature is the concept of runoff formation on the repletion of storage, which denotes that runoff is not produced until the soil moisture content of the aeration zone reaches field capacity (Zhao 1992; Xu et al. 2013). The Xinanjiang model includes two components, namely, runoff generation and runoff routing. It has 17 parameters that include seven runoff generating component parameters and 10 runoff routing parameters. These parameters are abstract conceptual representations of non-measurable watershed characteristics that have to be calibrated by an optimization method. Figure 8.3 shows the flowchart of the Xinanjiang model for three water sources. All symbols inside the blocks are variables including inputs, outputs, state variables and internal variables while those outside the block are parameters (Zhao 1992; Cheng et al. 2006; Li et al. 2011; Lin et al. 2014; Si et al. 2015).

The deterministic forecast model is calibrated respectively by taking NSE and RE as objective functions via automatic calibration methods with multiple objectives (Madsen 2000). Table 8.1 presents the calibrated parameters obtained for the Xinanjiang model of the TGR intervening basin and multiple-input single-output linear systematic model for TGR. The simulation results for the NSE and RE in the calibration period are 97.72 and -1.04% , respectively. Meanwhile, in the

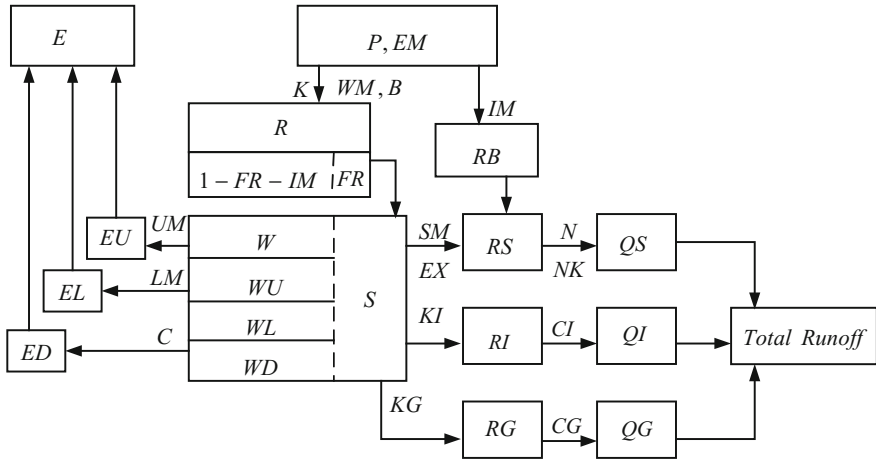


Fig. 8.3 The flow chart of Xinanjiang model for three water sources

verification period, the NSE and RE are 95.84 and -0.21% , respectively. These results show that the deterministic forecast model is proved to be quite efficient in simulating the inflow series for the TGR. The deterministic forecasts obtained from the well-calibrated deterministic model are subsequently applied to produce the probabilistic forecasts through the meta-Gaussian HUP and copula-based HUP.

8.2.5.3 Determination of Marginal Distributions

In this study, future rainfalls are treated as the case of perfect foreknowledge, rather than using the real forecast rainfalls to obtain simulated flows in the future, when the established deterministic forecast model is operated in the real-time forecasting mode. This is only for the illustration purpose if forecast rainfalls are available in reality and these would be used. The forecast lead times are 24 h ($n = 1$), 48 h ($n = 2$), and 72 h ($n = 3$). Especially, for each forecasting time in the record, the recorded rainfall data of 24-, 48- and 72- later followed by this forecasting time are treated as the “deterministic rainfall forecasts” (i.e., assuming perfectly known rainfalls in the future). Then these perfect forecasts of the rainfalls are input to the well-calibrated deterministic forecast model, which in turn produced model inflows (s_1, s_2, s_3). They are attached to actual inflow (h_0, h_1, h_2, h_3) to obtain one joint realization of the model-actual inflow process. The dataset from 2003 to 2009 are used to calibrate and compare the meta-Gaussian HUP and copula-based HUP.

The sample series of H_0 is taken from June 1 to Sept. 27 every year, S_1 from June 2 to Sept. 28, S_2 from June 3 to Sept. 29, and S_3 from June 4 to Sept. 30, thus all these four variables will have a data length of 833. The parameters of six candidate distributions are estimated by the L-moment method, and the K-S tests are used to verify the null hypothesis. The null hypothesis could not be rejected at

Table 8.1 Estimated parameters of Xinanjiang model in the TGR intervene basin

Parameter	Physical meaning	Estimated value
WM	Mean tension water capacity	149.80
UM	Areal mean water capacity of the upper layer	65.23
LM	Areal mean water capacity of the lower layer	38.64
K	Ratio of potential evapotranspiration to pan evaporation	0.433
B	Parameter in the distribution of tension water capacity	1.471
SM	Areal mean free water storage capacity	25.09
EX	Parameter in the distribution of free water storage capacity	0.984
KI	Coefficient relating RI	0.151
KG	Coefficient relating RG	0.12
IM	Impervious area of the basin	0.184
C	Evapotranspiration coefficient from deep layer	0.287
CI	Interflow reservoir constant	0.832
CG	Groundwater reservoir constant	0.904
m_1	Memory length of TGR intervening basin	10
N_1	Number of cascade linear reservoirs for TGR intervening basin	2.967
NK_1	Scale parameter of cascade linear reservoirs for TGR intervening basin	6.991
m_2	Memory length of upstream inflow	14
N_2	Number of cascade linear reservoirs for upstream inflow	1.241
NK_2	Scale parameter of cascade linear reservoirs for upstream inflow	1.911

Note The unit of WM , UM , LM and SM is mm, the rest of parameters are dimensionless

the 95% confidence level (critical value is 0.0471) for all six candidate distributions except Normal distribution. For the four hydrological variables, Gumbel distribution provides the minimum D value and is chosen as the best fitting distribution, respectively. Figure 8.4 shows the empirical CDF values obtained from the Gringorten plotting-position formula (Zhang and Singh 2006) and theoretical CDF values calculated by the Gumbel distributions. It can be seen that the theoretical values fit the empirical values very well. For comparison purposes, the copula-based HUP used the same marginal distributions as the meta-Gaussian HUP.

8.2.5.4 Calibration of Meta-Gaussian HUP

For the given climatic record of actual flows, the joint sample $\{(h_0, h_1)\}$ is formed of realizations on two consecutive days. Each joint realization (h_0, h_1) is processed through the empirical NQT to obtain the transformed joint sample $\{(w_0, w_1)\}$ and this joint sample is used to estimate the Pearson's correlation coefficient c . The advantage of using the empirical distributions in the NQT (instead of the parametric

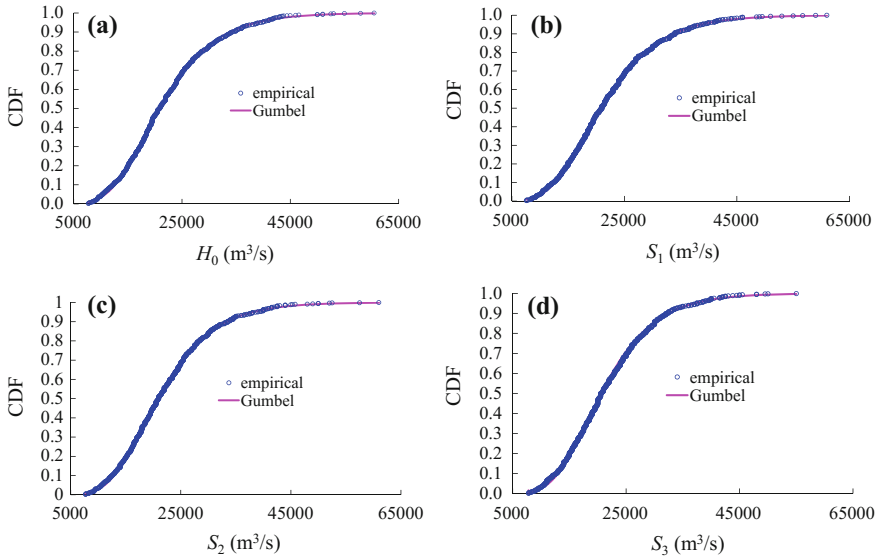


Fig. 8.4 Empirical and theoretical values fitted by Gumbel distributions

distribution) is that the estimate of c remains unaffected by the choice and the goodness of fit of the parametric model. The estimated result of Pearson’s correlation coefficient is 0.951.

The procedure for validating the meta-Gaussian dependence structure for the likelihood function parallels the procedure described in the prior density section above. The NQT performs adequately, as the empirical structure of dependence between X_n , W_n and W_0 , appears to be linear and homoscedastic. The meta-Gaussian model for the likelihood function captures the nonlinearity and heteroscedasticity of the dependence structure between S_n , H_n and H_0 .

8.2.5.5 Calibration of Copula-Based HUP

The rank-based correlation (Kendall’s coefficient) matrix of variables, H_0 , H_n and S_n are shown in Table 8.2. It is demonstrated that the dependence among the three variables pairs (H_0, H_n) , (H_0, S_n) and (H_n, S_n) are not the same. Furthermore, the highest correlation coefficient is exhibited in the variables pair (H_n, S_n) . This result indicates that rather than symmetric, the asymmetric trivariate copula functions may be more appropriate to be used to three-dimension joint distributions of H_0 , H_n and S_n . When constructing the three-dimension joint distributions using the asymmetric copula functions, the structures $(H_n, S_n)H_0$ are applied. Specifically, copula was firstly built for (H_n, S_n) , and then for H_0 and $C(F_{Hn}(h_n), F_{sn}(s_n))$.

The three-dimension joint distributions of H_0 , H_n and S_n ($n = 1, 2, 3$) are constructed using the three candidate trivariate copula functions. Dependence

Table 8.2 Ranked based correlation matrix of the variables

Lead times (h)	Variables	τ
24	H_0, H_1	0.823
	H_0, S_1	0.824
	H_1, S_1	0.929
48	H_0, H_2	0.694
	H_0, S_2	0.711
	H_2, S_2	0.883
72	H_0, H_3	0.600
	H_0, S_3	0.658
	H_3, S_3	0.828

Table 8.3 Estimated parameters of the three candidate copulas

Variables	Gumbel-Hougaard		Frank		Clayton	
	$[\theta_1, \theta_2]$	RMSE	$[\theta_1, \theta_2]$	RMSE	$[\theta_1, \theta_2]$	RMSE
H_0, H_1, S_1	[9.08, 10.65]	0.0116	[25.44, 35.23]	0.0103	[15.16, 20.78]	0.0150
H_0, H_2, S_2	[4.25, 7.58]	0.0141	[13.45, 20.38]	0.0112	[6.49, 14.33]	0.0186
H_0, H_3, S_3	[2.98, 6.57]	0.0144	[9.58, 16.21]	0.0117	[4.58, 9.34]	0.0188

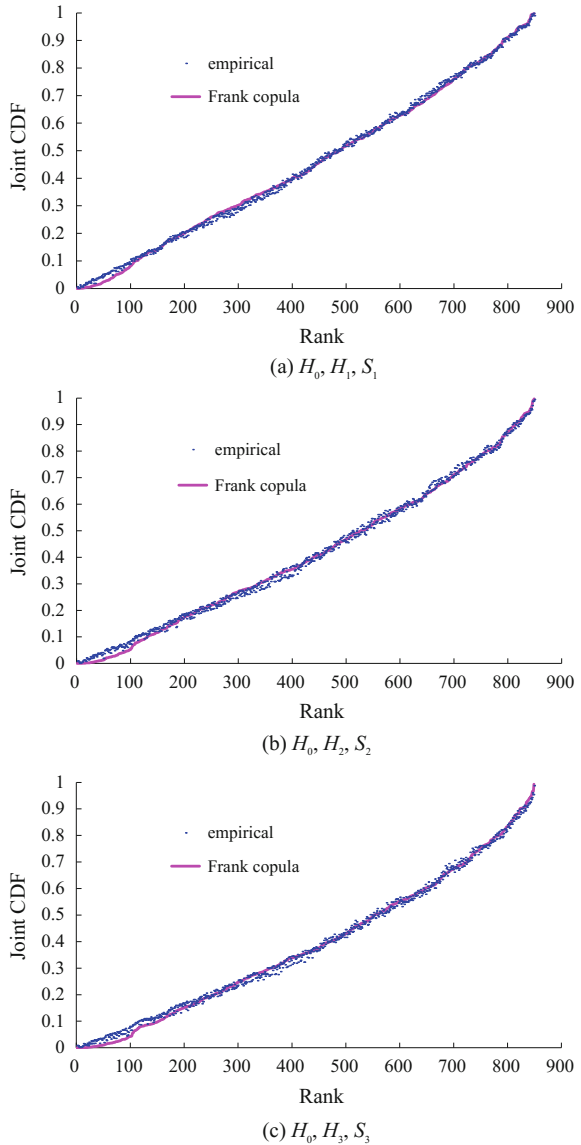
parameters of the trivariate copula functions are estimated using the maximum pseudo-likelihood method, and the results are listed in Table 8.3. It is found that Frank copula performs best with the smallest *RMSE* values for the three joint distributions. Empirical CDFs obtained from the Gringorten plotting-position formula and theoretical CDFs calculated from Frank copula for three joint distributions are plotted in Fig. 8.5. An overall satisfactory agreement between the empirical and theoretical CDF is shown. Hence, the asymmetric trivariate Frank copula functions have good performances in modeling the joint distributions of H_0 , H_n and S_n .

8.2.5.6 Comparison of the Meta-Gaussian HUP and Copula-Based HUP

(1) Posterior median forecasts

For 24-, 48- and 72 h lead times, the model efficiency *NSE* and relative error *RE* calculated by both the deterministic forecast model and posterior median forecasting associated with the meta-Gaussian HUP and copula-based HUP are listed in Table 8.4. It is shown that both the results of the meta-Gaussian HUP and copula-based HUP are slightly better than those of the deterministic forecast model,

Fig. 8.5 Plots of empirical and theoretical values estimated by Frank copulas for three joint CDFs. *Note* Rank represents number of ordered pair, ranked in the ascending order in terms of theoretical joint CDF, respectively



and the copula-based HUP is comparable to the meta-Gaussian HUP. Compared with deterministic forecasts, the NSE and the RE of the copula-based HUP for 24-, 48- and 72 h lead times forecasts are improved by 1.24, 1.26 and 1.26% and reduced by 0.17, 0.57, and 1.72%, respectively. It is also noted that the accuracy of posterior median forecasts of the both HUPs decreases as the lead time increases.

Table 8.4 Comparison of performances evaluation criteria for deterministic forecasts

Lead times (h)	Deterministic model		Meta-Gaussian HUP		Copula-based HUP	
	<i>NSE</i> (%)	<i>RE</i> (%)	<i>NSE</i> (%)	<i>RE</i> (%)	<i>NSE</i> (%)	<i>RE</i> (%)
24	97.55	-0.34	97.65	-0.31	98.79	-0.17
48	94.10	-0.85	94.54	-0.68	95.36	-0.28
72	88.52	-2.51	89.14	-1.14	89.78	-0.79

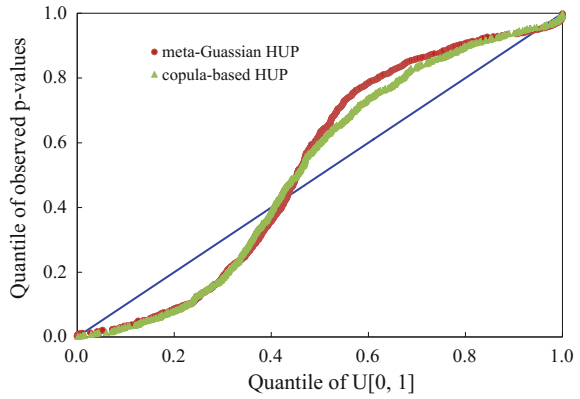
(2) Probabilistic forecasts

The predictive QQ plot, α -index, π -index and *CRPS* are adopted to evaluate the probabilistic forecasts. Figure 8.6 presents the predictive QQ plots regarding the meta-Gaussian HUP and copula-based HUP for 24-, 48- and 72 h lead times. Using Fig. 8.1 as a guide to assess the results, it is clear that the overall performances of all predictive QQ plots are acceptable. Both meta-Gaussian HUP and copula-based HUP systematically under-predict the inflows, since the observed p values at the theoretical median are a bit higher than the theoretical quantiles. In addition, it also shows that the observed p values cluster around the tails (i.e., a high slope around theoretical quantile 0.4–0.6). This finding means that the predictive uncertainty is somewhat underestimated for both HUPs. The overall behaviors of meta-Gaussian HUP and copula-based HUP are found to be similar. The QQ plot for copula-based HUP is slightly closer to the 1:1 line than meta-Gaussian HUP. That is to say, the copula-based HUP performs marginally better regarding reliability. Nonetheless, these underestimations for both meta-Gaussian and copula-based HUPs are in such zones where p values are relatively higher, indicating such differences may not be statistically significant.

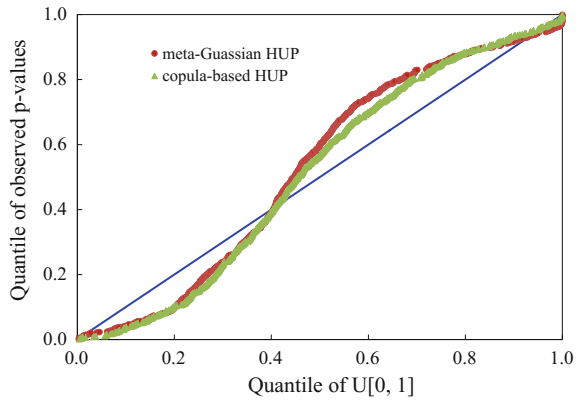
The results of α -index, π -index and *CRPS* are summarized in Table 8.5. For both meta-Gaussian HUP and copula-based HUP, it is clearly shown that the α -index value increases (higher reliability) when the lead time increases. However, it should be noted that this is at the expense of decreasing π -index value (lower resolution). Besides, the copula-based HUP has slightly larger α -index values while smaller π -index values compared with the meta-Gaussian HUP. Regarding *CRPS* value, both HUPs outperform the deterministic forecasts which demonstrate the effectiveness of probabilistic forecasts. Comparison results also indicate that the copula-based HUP is marginally better than the meta-Gaussian HUP. The *CRPS* value of the copula-based HUP for 24-, 48- and 72 h lead times is improved (decreased) by 16.6, 21.2, and 23.3%, respectively.

Although that such marginally better performance does not result for each year, for illustration purposes, the observed and median discharges, and 90% inflow prediction intervals estimated by meta-Gaussian HUP and copula-based HUP in 2004 are presented in Figs. 8.7 and 8.8, respectively. It can be seen that most observed inflows are contained within the 90% prediction intervals. This demonstrates that these 90% prediction intervals can effectively capture the forecast uncertainty and provide more information for decision-making in flood control and reservoir operation. As lead time increases, the 90% prediction intervals become wider (i.e., greater uncertainty).

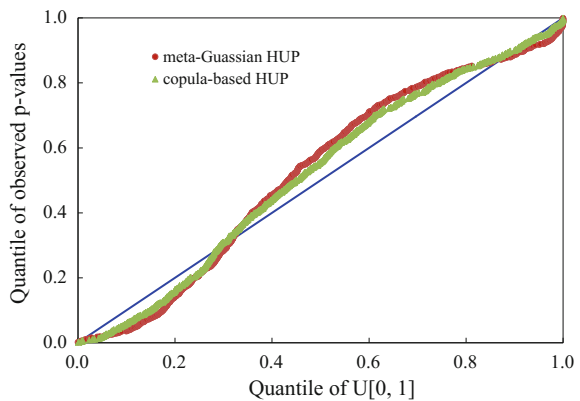
Fig. 8.6 The predictive QQ plots of meta-Gaussian HUP and copula-based HUP



(a) 24h lead time



(b) 48h lead time



(c) 72h lead time

Table 8.5 Comparison of performances evaluation criteria for probabilistic forecasts

Lead times (h)	Deterministic model	Meta-Gaussian HUP			Copula-based HUP		
	<i>CRPS/MAE</i>	α -index	π -index	<i>CRPS</i>	α -index	π -index	<i>CRPS</i>
24	688	0.8028	18.55	608	0.8507	16.39	574
48	1180	0.8555	12.45	975	0.8916	10.42	930
72	1763	0.8879	9.13	1420	0.9184	7.68	1353

Note α -index and π -index are dimensionless; the unit of *CRPS* is m³/s

8.3 Uncertainty Analysis of Hydrological Multi-model Ensembles Based on CBP-BMA Method

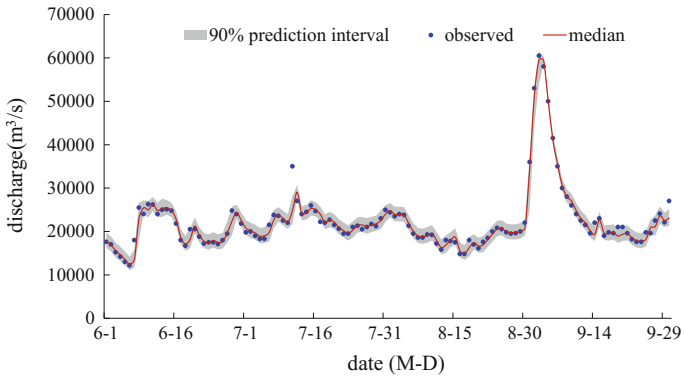
Inspired by the ideas of Madadgar and Moradkhani (2014), a general framework of the combination of copula Bayesian processor with BMA (CBP-BMA) is proposed by He et al. (2018), where the Bayesian theory is applied in the transformation of the posterior distribution. The flowchart of different probability forecast methods based on deterministic models is described in Fig. 8.9.

8.3.1 Description of the Hydrological Models

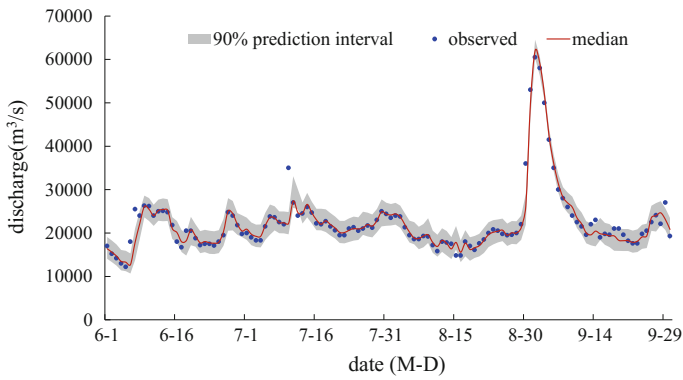
Three world-famous conceptual hydrological models are implemented in the Mumahe catchment, including the Xinanjiang (XAJ), HBV and SIMHYD models. The XAJ model has been used in humid and semi-humid region worldwide (Zhao 1992). It consists of a runoff generation component with seven parameters and a routing component with ten parameters. Those model physical parameters represent the abstract conceptual expression of watershed features. The HBV model is a synthetic flow model with 13 parameters needed to be calibrated. Units of HBV model makes up of the routines for snowmelt accumulation, evapotranspiration and soil routine and response function. The core concept assumes runoff volume changes with soil humidity exponentially (Montero et al. 2016). The SIMHYD model is a lumped conceptual hydrological model which contains seven parameters needed to be calibrated. The model divides runoff into three components: surface flow, interflow and base flow. The surface flow is infiltration excess runoff, inter-flow is estimated as a linear function of the soil wetness, and base flow is simulated as a linear recession from the groundwater store (Chiew et al. 2009; Yu and Zhu 2015). The infiltration rate is a core of the model.

8.3.2 Bayesian Model Averaging (BMA)

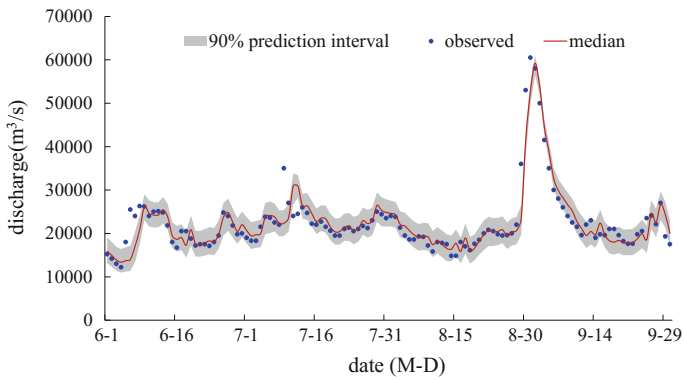
Raftery et al. (2005) successfully extended BMA to statistical post-processing for forecast ensembles. The BMA method addresses total model uncertainty by



(a) 24h lead time

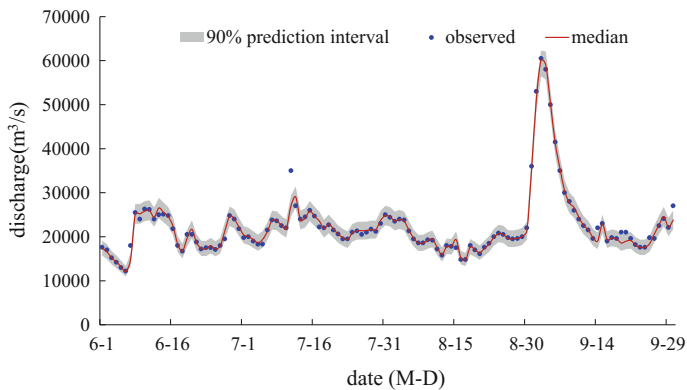


(b) 48h lead time

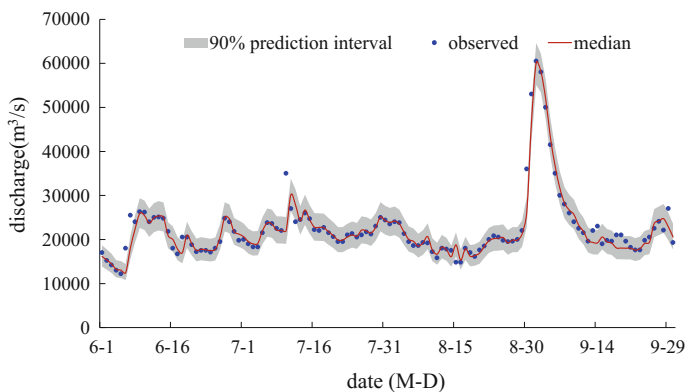


(c) 72h lead time

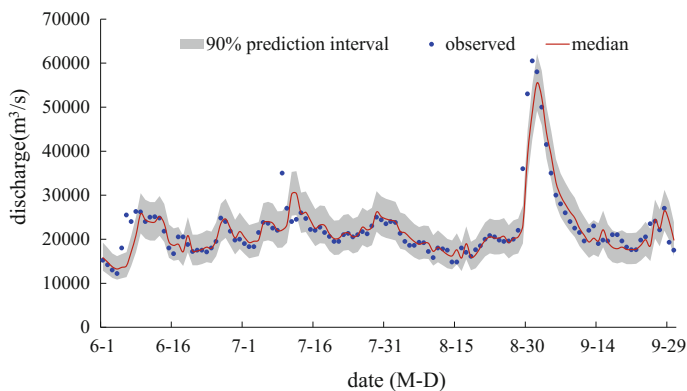
Fig. 8.7 The 90% prediction intervals, median and observed discharges in 2004 (meta-Gaussian HUP)



(a) 24h lead time



(b) 48h lead time



(c) 72h lead times

Fig. 8.8 The 90% prediction intervals, median and observed discharges in 2004 (copula-based HUP)

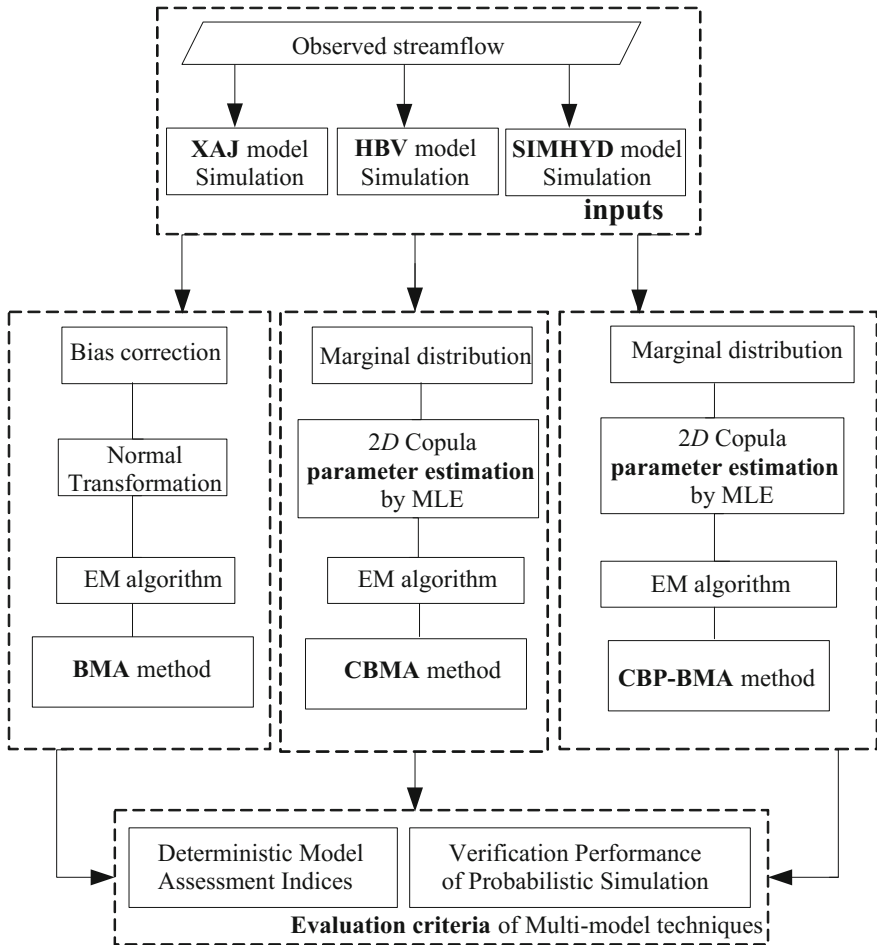


Fig. 8.9 Flowchart of hydrologic multi-model ensembles for uncertainty analysis

conditioning not only on a single outstanding model but on the entire ensemble models. The method was originally proposed as a pathway for method combination of several competing models (Duan et al. 2007; Liang et al. 2011).

According to BMA (Duan et al. 2007), the ensemble predictive density of the actual flow variable q , given the different hydrologic model simulations of K models $[S_1, S_2, \dots, S_K]$ and the observations during the training period, Q , can be expressed in terms of the law of total probability:

$$p(q|S_1, S_2, \dots, S_K, Q) = \sum_{i=1}^K p(S_i|Q) \cdot p_i(q|S_i, Q) \tag{8.33}$$

where $p(S_i|Q)$ is the posterior probability of i th model prediction. This static term can also be expressed as w_i , reflecting how well the ensemble term fits the observation dataset. It ranges from 0 to 1 since the posterior model probabilities add up to one. Before the implantation of BMA algorithm, the expected value of observation and forecast for each model should be equal zero ($E[q - S_i] = 0$). Any bias-correction method, such as linear regression, should be applied to substitute the bias-corrected forecast (f_i) for the original deterministic forecast:

$$f_i = a_i + b_i \cdot S_i \quad (8.34)$$

where $\{a_i, b_i\}$ are the coefficients of the linear regression model.

The term $p_i(q|f_i, Q)$ is the conditional pdf of h based on the bias-corrected simulation f_i and the observation dataset. Moreover, the power Box-cox transformation is taken for the computational convenience of using a Gaussian distribution. The posterior distribution $p_i(q|f_i, Q)$ is mapped to a Gaussian space with mean f_i and variance s_i^2 ; i.e., $p_i(q|f_i, Q) \sim g(q|f_i, \sigma_i^2)$. The BMA predictive mean and variance of q are defined as follows (Raftery et al. 2005):

$$E(q|Q) = \sum_{i=1}^K p(f_i|Q) \cdot E[p_i(q|f_i, Q)] = \sum_{i=1}^K \omega_i f_i \quad (8.35)$$

$$\text{Var}(q|Q) = \sum_{i=1}^K \omega_i \left(f_i - \sum_{i=1}^K \omega_i f_i \right)^2 + \sum_{i=1}^K \omega_i \sigma_i^2 \quad (8.36)$$

Successful application of the BMA method requires estimations of the weight ω_i and variance σ_i^2 of the individual pdf. The log maximum likelihood function rather than the likelihood function is optimized for reasons of both numerical stability and algebraic simplicity. If the BMA parameters are estimated by $\theta = \{\omega_i, \sigma_i, i = 1, 2, \dots, K\}$, the log likelihood function of θ is mathematically denoted as:

$$l(\theta) = \log \left(\sum_{i=1}^K \omega_i \cdot p_i(q|f_i, Q) \right) \quad (8.37)$$

After the completion of BMA parameter estimation by the EM algorithm (Duan et al. 2007), another feature of the BMA method is to make use of Monte Carlo method to derive BMA probabilistic ensemble prediction for any time t (Kuczera and Parent 1998). The procedures are described as follows (Zhou et al. 2016).

- (1) Select the probabilistic ensemble size, M ($M = 100$).
- (2) Randomly generate a value of k from the numbers $[1, 2, \dots, K]$ with probabilities $[\omega_1, \omega_2, \dots, \omega_i]$. The detail processes are shown as follows: (a) Initial the cumulative weight $\omega'_0 = 0$ and compute $\omega'_i = \omega'_{i-1} + \omega_i$ for $i = 1, 2, \dots, K$;

- (b) Generate a random number u between 0 and 1; and (c) If $\omega'_{i-1} \leq u \leq \omega'_i$, then the i th member of the ensemble predictions are chosen.
- (3) Generate a value of q from the pdf of $p_i(q|f_i, \sigma_i^2)$.
- (4) Repeat steps (2) and (3) for M times.

The results are sorted in ascending order, and the 90% confidence interval can be derived within the range of the 5 and 95% quantiles.

8.3.3 The Hybrid Copula-BMA (CBMA)

As illustrated before, the BMA predictive distribution provides a weighted average of simulation pdf which generally complies with a parametric distribution, e.g., Gaussian distribution after the box-cox transformation. Madadgar and Moradkhani (2014) employed copula to estimate the posterior distribution of forecast variables for each model, i.e., $p_i(q|f_i, Q)$, and found that the hydrological forecasts are improved after the integration of copulas and BMA (CBMA). A series of research demonstrates that the procedures of CBMA not only eliminate the prophase bias-correction and the external calculation of variance but also simplify the calculation of the weighted average and the probability model structure by copula (Möller et al. 2013).

Alternatively, in statistical applications, the conditional probability distribution of h given s_i ($i = 1, 2, 3$) is expressed as (Madadgar and Moradkhani 2014):

$$f(q|s_i) = \frac{f(q, s_i)}{f(s_i)} = \frac{c(u, v_i) \cdot f(q) \cdot f(s_i)}{f(s_i)} = c(u, v_i) \cdot f(q) \quad (8.38)$$

where $c(u, v_i)$ is computed for each pair of (u, v_i) , $f(q)$ represents the marginal distribution of actual flow. Although different copula families have been proposed and described in current studies (Chebana and Ouarda 2007), several families of Archimedean copulas, including Frank, Gumbel, and Clayton, have been popular choices for dependence models in hydrologic analyses due to their simplicity and generation properties.

The predictive distribution of CBMA is modified as follows (Madadgar and Moradkhani 2014):

$$f(q|s_1, s_2, \dots, s_K) = \sum_{i=1}^K \omega_i f(q|s_i) = \sum_{i=1}^K \omega_i \cdot c(u, v_i) \cdot f(q) \quad (8.39)$$

It can be seen from Eq. 8.39 that it relaxes any assumption on the type of posterior distribution $f(q|s_i)$, whose term can be directly inferred with the help of copula functions. Once the term $f(q|s_i)$ is defined, their weights are estimated by the EM algorithm with a few adjustments, which can refer to Madadgar and Moradkhani (2014) for details.

The hybrid CBMA model applies the idea of “pair and ensemble”. The pair of observation q and the i th model simulation is established to get the probability task by the well-developed copula theory, while the ensemble is to formulate a consensus probability interval.

8.3.4 Copula Bayesian Processor Associated with BMA (CBP-BMA) Method

8.3.4.1 Copula Bayesian Processor (CBP)

Copula Bayesian processor (CBP) is developed as another component of the probabilistic forecasting system in virtue of the integration of Bayesian theory and copula functions. The CBP procedure generates a probabilistic result and quantifies the hydrologic uncertainty under the assumption that input uncertainty is ignored, which refers to hydrologic uncertainty processor (Krzysztofowicz and Kelly 2000). This method also has the advantage of leaving out a data transformation procedure into Gaussian space. The Bayesian procedure based on the law of total probability involves two parts for information revision of uncertainty (Zhang and Singh 2007a, b, c):

- (1) The expected conditional density function of deterministic simulation, S_i given $Q = q$ is expressed as:

$$\kappa(s_i|q) = \int f(s_i|q) \cdot g(q) dq \quad (8.40)$$

where $f(q|s_i)$ has the same conception as before, $g(q)$ represents the prior density function.

- (2) The posterior density function conditional on a deterministic result $S_i = s_i$ is derived via Bayes' theorem:

$$\phi(q|s_i) = \frac{f(s_i|q) \cdot g(q)}{\kappa(s_i|q)} \quad (8.41)$$

Equations 8.40 and 8.41 could be rewritten by using copula functions, i.e., the CBP form of the right term is mathematically expressed by:

$$\phi(q|s_i) = \frac{f(s_i|q) \cdot g(q)}{\int f(s_i|q) \cdot g(q) dq} = \frac{c(u, v_i)}{\int_0^1 c(u, v_i) du} \cdot g(q) \quad (8.42)$$

The final CBP outputs a posterior distribution of the process, conditional upon the deterministic simulation. Since the analytical solution to the integral term $\int_0^1 c(u, v_i) du$ is very complex, the Monte Carlo technique is used to estimate the posterior density function $\phi(q|s_i)$ (Robert and Casella 2011; Kroese et al. 2013).

8.3.4.2 The CBP-BMA Method

The difference between the CBP-BMA and CBMA methods is the estimation procedure of the posterior density function.

$$\phi(q|s_1, s_2, \dots, s_K) = \sum_{i=1}^K \omega_i \phi(q|s_i) = \sum_{i=1}^K \omega_i \frac{c(u, v_i)}{\int_0^1 c(u, v_i) du} g(q) \quad (8.43)$$

It should be rational to assign weights on account of multiple deterministic results. The calculation process of weights is conducted by the EM algorithm (Montanai and Grossi 2008). The three main steps of the presented weights calculating paradigm can be summarized as follows:

$$\begin{aligned} w_i^{Iter} &= \frac{1}{T} \sum_{t=1}^T z_{i,t}^{Iter} \\ z_{i,t}^{Iter} &= \frac{w_i^{Iter-1} \cdot \phi(q_t|s_{i,t})}{\sum_{i=1}^K w_i^{Iter-1} \cdot \phi(q_t|s_{i,t})} = \frac{w_i^{Iter-1} \cdot c(u_t, v_{i,t})g(q_t) / \int_0^1 c(u_i, v_{i,t}) du_t}{\sum_{i=1}^K w_i^{Iter-1} \cdot c(u_t, v_{i,t})g(q_t) / \int_0^1 c(u_i, v_{i,t}) du_t} \\ l(\theta_{Iter}) &= \log \left(\sum_{i=1}^K w_i^{Iter-1} \cdot \sum_{i=1}^K c(u_i, v_{i,t})g(q_t) / \int_0^1 c(u, v_{i,t}) du_t \right) \end{aligned} \quad (8.44)$$

where T is the length of the training period; and z is a latent variable. Compared with the standard BMA method, the calculation of variance and data transformations are eliminated in Eq. 8.44. The posterior probability of q_t is calculated only once while it need be re-calculated every time in the standard BMA method.

8.3.5 Evaluation Criteria for Multi-model Techniques

8.3.5.1 Deterministic Model Assessment Indices

To evaluate the quality of the deterministic model, three metrics are used.

- (1) Nash-Sutcliffe efficiency coefficient (*NSE*), see Eq. 8.25
- (2) Daily root mean square error (*DRMS*)

$$DRMS = \sqrt{\frac{\sum_{i=1}^N (q_o^i - q_m^i)^2}{T}} \quad (8.45)$$

As the second tool employed is sensitive to the differences between observations and simulations, the values of *DRMS* approaching to stand for better performance.

(3) Kling-Gupta efficiency (*KGE*)

$$\begin{aligned}
 KGE &= 1 - \sqrt{(r - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2} \\
 \beta &= \bar{q}_m / \bar{q}_o \\
 \gamma &= CV_m / CV_o = \frac{\sigma_m / \bar{q}_m}{\sigma_o / \bar{q}_o}
 \end{aligned}
 \tag{8.46}$$

where r is the Pearson correlation between the observation and simulation, b is the bias ratio indicator; g is the variability ratio (Kling et al. 2012). All calculative variables are replaced by the expected values of the estimate predictive distributions.

8.3.5.2 Verification of Probabilistic Simulations

With regard to assessment of assessing the uncertainty analysis of simulation interval, Xiong et al. (2009) and Dong et al. (2013) presented multiple verification indices and applied in hydrologic practice. Three main metrics are selected to evaluate the simulation uncertainty intervals generated by the BMA, CBMA and CBP-BMA methods.

(1) Containing ratio (*CR*)

The containing ratio is utilized as a significant index for assessing the goodness of the uncertainty interval. It is defined as the percentage of observed data points that fall between the prediction bounds, directly reflecting the interval performance.

$$CR = \frac{C_{i=1}^N (q_l^i \leq q_o^i \leq q_u^i)}{N} \times 100\%
 \tag{8.47}$$

where q_l^i is denoted as the lower bound corresponding to 5% quantile at time t , q_u^i is denoted as the upper bound corresponding to 95% of the quantile. $C_{i=1}^N$ is the number of the observed data points q_o^i that satisfy the inequality conditions.

(2) Average bandwidth (*BW*)

$$B = \frac{1}{N} \sum_{i=1}^N (q_u^i - q_l^i)
 \tag{8.48}$$

where BW is also an index measuring the average width of estimated uncertainty interval just as the definition name indicates. Smaller values of BW show a greater precision. Consider two forecasts with the same containing ratio; the situation with smaller BW is preferred because it has less uncertainty or greater precision.

(3) Average deviation amplitude (DA)

The average deviation amplitude DA is an index to quantify the average deflection of the curve of the middle points of the prediction bounds from the observed streamflow hydrograph. It is defined as

$$D = \frac{1}{N} \sum_{i=1}^N \left| \frac{1}{2}(q_u^i - q_l^i) - q_o^i \right| \quad (8.49)$$

where the notations are defined previously.

8.3.6 Case Study

The Mumahe catchment (Fig. 8.10), a sub-basin of Hanjiang River basin in China is selected as a case study. The catchment lies in Shanxi Province with an area of 1224 km² and locates in the subtropical monsoon region with a humid climate and fairly plenty of precipitation. The annual mean precipitation and runoff is 1070 and 687 mm, respectively. The available dataset contains daily precipitation, runoff, and evaporation with a length of 11 years (1980–1990). The first year (1980) is used as the spin-up period for each hydrologic model to achieve the best effective model formulation. The remaining years (1981–1990) are divided into two sub-periods, with seven years (1981–1987) for calibration and three years (1988–1990) for validation.

Different multi-model techniques, i.e., BMA, CBMA, and CBP-BMA, are applied to combine the ensemble flow simulation. The structures of three hydrologic models ought to be determined as the deterministic results are crucial to final uncertainty analysis. As mentioned above, the calibration parameters of the first BMA method are ω_k and σ_k^2 ; In the CBMA method, they are the parameters of marginal distributions, weights ω_k and the parameters of the PDF of the copula. In the CBP-BMA method, the Monte Carlo sampling technique is also used to obtain the integral item.

8.3.6.1 Deterministic Hydrologic Model Simulations

The genetic and simplex algorithms are used for model calibration on account of their flexibility and good convergence. The genetic algorithm can acquire the global optimal value with independent of initial parameter values. The simplex algorithm

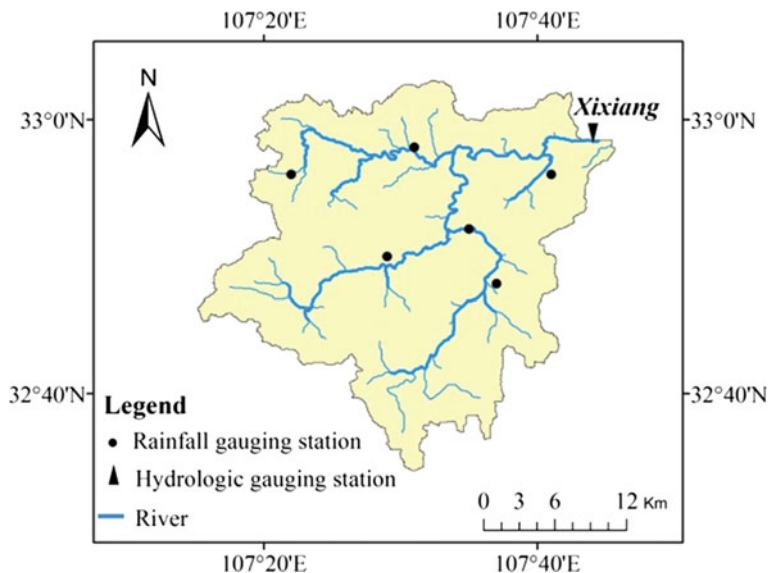


Fig. 8.10 Sketch map of the Mumuhe catchment

is of high accuracy with low convergence rate. With the merit of two methods integrated, the approximately optimal values of model parameters are obtained. Three deterministic assessment indices: *NSE*, *DRMS* and *KGE* scores over the calibration period (1981–1987) and the validation period (1988–1990) are calculated for XAJ, HBV, and SIMHYD models. Table 8.6 indicates that the XAJ model has the best results, the HBV model takes the second place, and the SIMHYD model behaves worst among the three. The reason can be attributed to the dissimilar process for the calibration of each model (Nasonova et al. 2009). In practice, it might partially refer to inaccurate estimation of model parameters as one of the error sources of the model structure, abstract formulation of physical processes, and different sources of forcing data set for each model. In general, these simulation results can be used as the input data of multi-model ensemble in terms of the *NSE* and *KGE* values, which are 85% and higher than 0.82 respectively beside the ill value of *KGE* of HBV.

8.3.6.2 Determination of the Marginal Distributions

The marginal distributions of the random variables of H and S_i ($i = 1, 2, 3$) need to be determined. Five common candidate distributions, namely Normal, Gamma, Gumbel, P-III and Log-Normal, have been fitted to the daily mean streamflow values as well as to the XAJ, HBV and SIMHYD model simulations.

Table 8.6 Deterministic accuracy assessment of different hydrological models

Model	Calibration			Validation		
	<i>NSE</i> (%)	<i>DRME</i>	<i>KGE</i> (%)	<i>NSE</i> (%)	<i>DRME</i>	<i>KGE</i> (%)
XAJ	88.25	30.06	90.59	84.85	24.06	87.08
HBV	84.81	34.16	52.99	82.24	26.06	42.89
SIMHYD	86.25	32.50	82.51	84.98	23.97	85.01
BMA	88.87	33.79	89.14	85.72	24.81	90.62
CBMA	88.93	26.54	90.06	86.07	21.25	88.45
CBP-BMA	89.76	27.63	90.96	86.69	23.39	89.23

Notes Values in bold represent the optimal result

Regarding random variable H , the parameters of five candidate distributions are estimated by the method of L -moment (Hosking 1990), and the parameter values are listed in Table 8.7. The K-S tests are used to verify the null hypothesis, and the corresponding statistic D_{K-S} values are also listed in Table 8.7. It is shown that the null hypothesis could not be rejected at the 95% confidence level (threshold value $D_n = 1.36/\sqrt{N}$, N is the number of sampling points) for Log-Normal distribution with providing the minimum D_{K-S} value. Meanwhile, Fig. 8.11 indicate the Log-Normal is satisfactory on visual inspection that the cumulative distribution function (CDF) plots of the theoretical Log-Normal distributions fitted the empirical CDF values obtained from the Gringorten plotting-position formula (Zhang and Singh 2006) relatively well. The estimation of marginal distributions for S_i had the similar procedures. The Kolmogorov-Smirnov statistics D_{K-S} indicate that the Log-Normal distribution also gives the best fit in this study.

8.3.6.3 Archimedes Copula Selection and Estimation

In the application of the CBMA and CBP-BMA methods, a copula function to link the CDF of observation and model simulations needs to be defined. The Gumbel, Clayton and Frank copula belonging to Archimedes family are chosen to test for flexibility and universality (Madadgar and Moradkhani 2014; Chen et al. 2015).

For Archimedes copula, the Kendall correlation coefficient τ_i ($i = 1, 2, 3$) between observed and different simulated flows is firstly derived. The higher τ_i indicator reflects the stronger correlation between observation and model simulation. The corresponding copula parameter θ_i is calculated by the method based on the inversion of τ_i in Table 2.3 of Chap. 2. The parameter estimators and goodness-of-fit test ($RMSE$ and AIC) are used to determine the best fit copula for integrating the streamflow properties. The results illustrate that copulas have the good performance in exploring the associations of observed and simulated flows. All variables passed the null hypothesis for Gumbel and Frank copulas. Gumbel copula performs with the lowest $RMSE$ and AIC values.

Table 8.7 Estimated parameters and statistic test D_{k-s} of five candidate marginal distributions

Marginal parameter distribution and $K-S$ test		H	S_1	S_2	S_3
Gumbel	σ	40.3	40.6	39.5	37.2
	μ	17.8	14.9	15.1	14.5
	D_{k-s}	0.025	0.024	0.026	0.026
Gamma	α	0.4	0.3	0.3	0.3
	β	103.4	129.7	153.2	106.5
	D_{k-s}	0.148	0.122	0.141	0.153
P-III	α	0.20	0.24	0.20	0.19
	β	0.0056	0.0067	0.0050	0.0063
	γ	6.02	4.43	5.27	4.59
	D_{k-s}	0.034	0.028	0.029	0.031
Log-normal	α	2.83	2.11	2.43	2.57
	γ	1.17	1.46	1.28	1.26
	D_{k-s}	0.016	0.013	0.017	0.015
Normal	α	41.08	38.27	38.43	34.69
	γ	49.57	49.85	41.19	43.03
	D_{k-s}	0.074	0.062	0.075	0.064

Notes Values in bold denote that the distribution model passes the goodness-of-fit test at 0.05 significance levels

8.3.6.4 Deterministic Assessment of Three Ensemble Methods

We check the mean simulation of hydrologic multi-model ensembles using three criteria illustrated in Sect. 8.3.5.1. The effectiveness results of BMA, CBMA and CBP-BMA methods are listed in Table 8.6. The performances of different multi-models are better than that of the individual XAJ model regarding NSE . The BMA method outperforms the reference model at the cast of $DRMS$ and KGE indicators, The CBMA and CBP-BMA methods slightly improve in all aspects during the calibration period, which have excellent properties in the validation period. The reason of the CBMA and CBP-BMA methods enhancing the performance can be attributed to that copula functions are efficient tools to remove bias instead of a simple bias correction such as linear regression in the BMA method (Madadgar and Moradkhani 2014). Especially, copula has reliable parameter estimation prior model average procedure. Another reason might be owed to the weight of each individual model, which is directly influenced by the estimation of posterior distributions.

Figure 8.12 illustrates the bar plots of KGE score and its components. The KGE score might be a little descending through BMA or CBMA application, a little incremental through CBP-BMA application in comparison with the best XAJ model. The correlation coefficients between observation and simulation of individual models are up to 0.93 in the calibration period and 0.92 in the verification

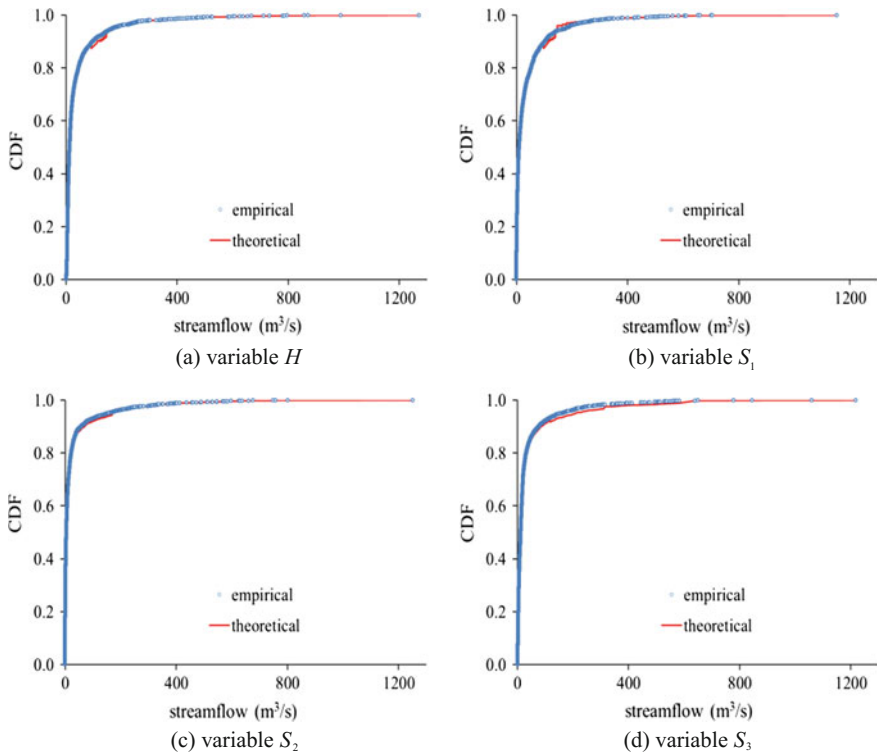


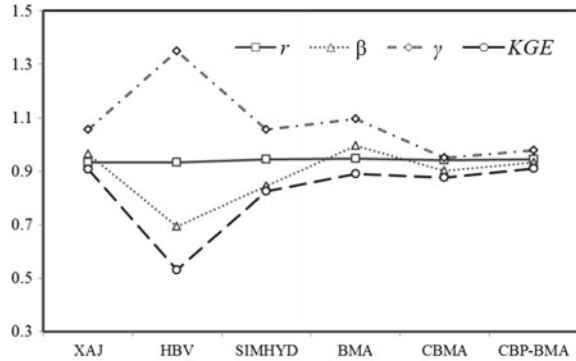
Fig. 8.11 Comparison of the empirical and theoretical cumulative distribution functions

period, which represent stronger correlation for the values are more than 0.9. However, β indicator of deterministic models varies from 0.64 for HBV model to 0.97 for XAJ model. The value less than 1 indicates the total amount of streamflow simulation in any individual model is less than that of observation. It might cause the general underestimation of the mean streamflow (negative bias) in hydrological multi-model ensemble applications. The BMA method is such a promising method for locating simulation to observation for its term β closer to 1. Regarding the variability ratio, all methods except for HBV could perfectly perform, but no particular method is superior to others with all $\gamma \approx 1$.

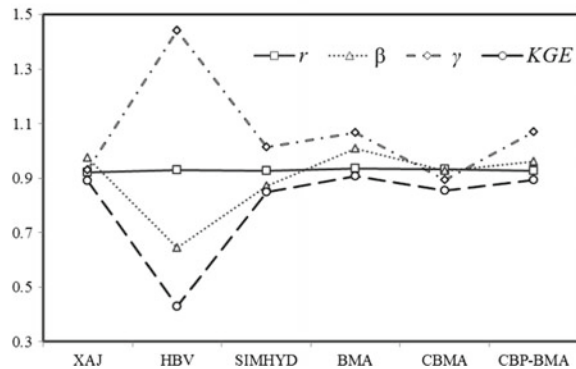
8.3.6.5 Probabilistic Verification of Three Ensemble Methods

For probabilistic verification of simulation, Figs. 8.13 and 8.14 describe the uncertainty bands of different methods for the representative year during calibration and verification periods with a visual inspection. These two plots indicate that the observed values approximately fall within the 5–95% uncertainty range and fit the

Fig. 8.12 The simulation results of *KGE* score and its components



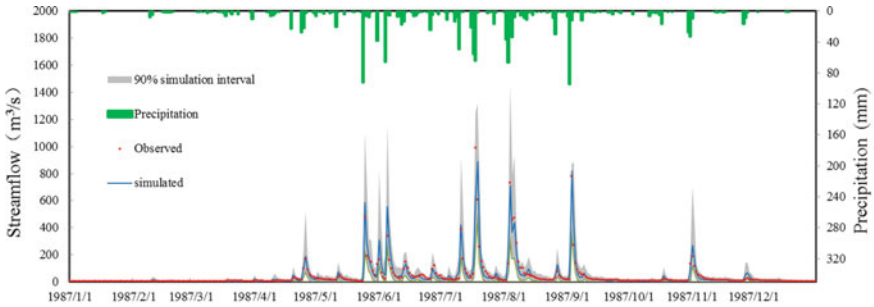
(a) Calibration period



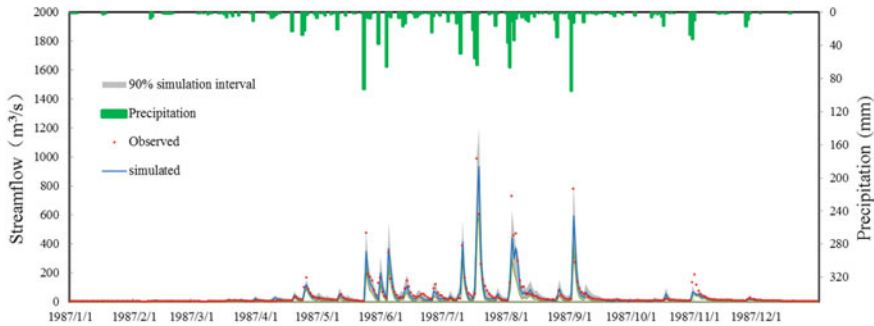
(b) Validation period

mean flow hydrograph for all multi-model ensembles. In this case, the 90% confidence interval could capture the flood peaks but miss more low flow values.

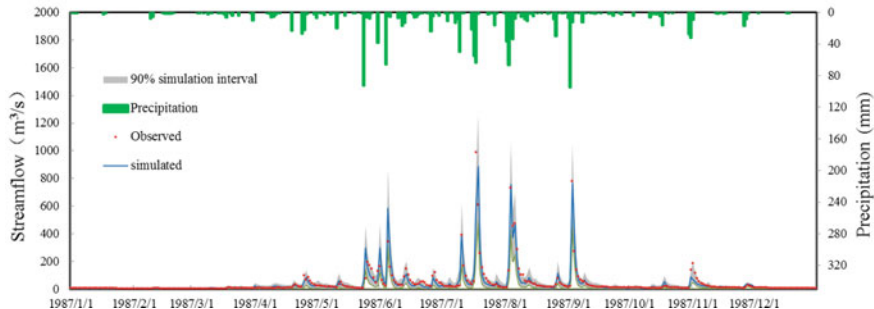
Three probabilistic verification measurements (*CR*, *BW*, *DA*) are presented in Table 8.8. It can be seen from these quantitative indices that they have a good performance regarding containing ratio, which is corresponding to the confidence interval. The probability of observed value falling in the range should be in accord with the percentage of confidence interval containing points through many independent statistical experiments. The CBP-BMA method performs better than CBMA method regarding *CR* index because it roughly covers 91% of the sample points, which is more than CBMA does. A combination of *CR* and *BW* possess the power to make a decision on model probabilistic performance. The comparison between the CBMA (and CBP-BMA) and BMA methods exactly illustrates that the CBMA method outperforms the BMA method, either *CR*, *BW* or *DA*, especially, the containing ratios of CBP-BMA method in different periods are up to 91.17 and 91.33%, respectively. Referring to the smaller *BW* result in the CBMA and CBP-BMA methods, the total predictive variance is reduced by relaxing the PDF generated by copula functions rather than the Gaussian posterior distribution via



(a) BMA



(b) CBMA

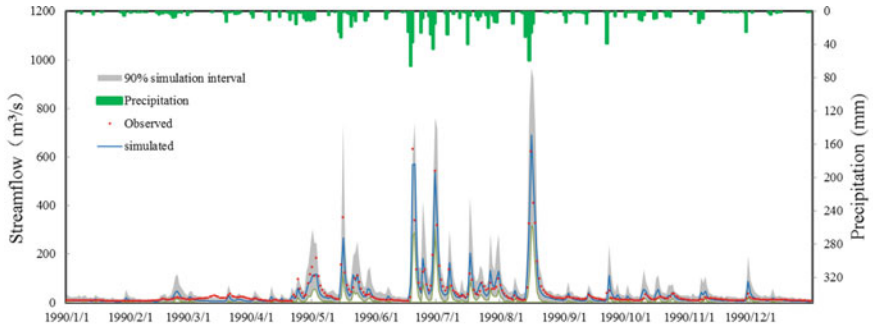


(c) CBP-BMA

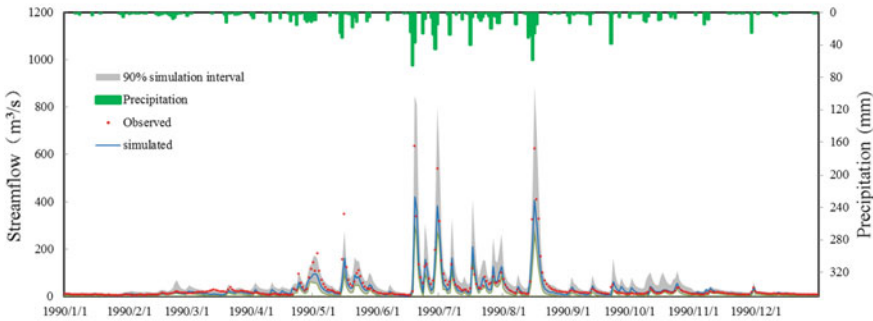
Fig. 8.13 The 90% uncertainty interval, observed, mean simulation for the Mumahe catchment in 1987 during the calibration period

box-cox transformation. Since the between-model variance keeps identical after using the same EM algorithm in all three methods, it is inferred that the reduction of within-model variance works.

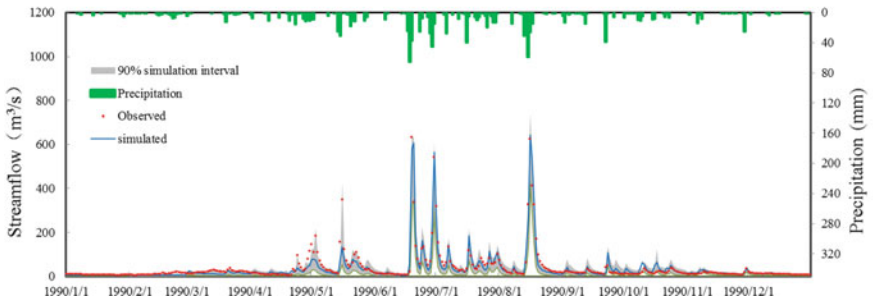
The CBMA and CBP-BMA methods are two flexible and robust approaches to estimate uncertainty regarding the optimal bandwidth and average deviation amplitude. They have an intuitive and simple structure conditional on several model



(a) BMA



(b) CBMA



(c) CBP-BMA

Fig. 8.14 The 90% uncertainty interval, observed, mean simulation for the Mumahé catchment in 1990 during the validation period

simulations by the integration of BMA and copula tools, which makes this method promising to derive uncertainty. The difference between them reflected in the procedure of processing posterior distribution. Further improvement might be realized through the weight allocation for each model or the nonparametric posterior distribution.

Table 8.8 Uncertainty assessment of different hydrological multi-model ensembles

Model	Calibration			Validation		
	<i>CR (%)</i>	<i>BW</i>	<i>DA</i>	<i>CR (%)</i>	<i>BW</i>	<i>DA</i>
BMA	87.34	56.79	21.26	88.53	58.14	18.79
CBMA	89.23	38.75	16.52	89.76	40.28	12.26
CBP-MA	91.17	45.28	12.35	91.33	42.35	10.99

Note Values in bold represent the optimal result

8.4 Conclusion

Hydrological forecasting services are trending toward providing users with probabilistic forecasting, and adequate assessment of uncertainty forecasts is an important issue and task. A copula-based HUP for probabilistic forecasting and CBP-BMA method for evaluating uncertainties of hydrologic multi-model ensembles are proposed. Three Gorges Reservoir (TGR) and Mumahe basins are selected as case studies. The main conclusions are summarized as follows:

- (1) The output of the HUP is a posterior distribution of the process, conditional upon the deterministic forecast. This posterior distribution provides the complete and well-calibrated characterization of uncertainty needed by rational decision makers who use formal decision models and by information providers who want to extract various forecast products for their customers (e.g., quantiles with specified exceedance probabilities, prediction intervals with specified inclusion probabilities, probabilities of exceedance for specified thresholds).
- (2) Based on copula function, the prior density and likelihood function of the HUP are explicitly expressed, and the corresponding posterior density and distribution can be obtained using the Monte Carlo sampling technique. This copula-based HUP can be implemented in the original space directly without a data transformation procedure into Gaussian space and allows for any form of marginal distribution of predictand and the deterministic forecast variable, and a nonlinear and heteroscedastic dependence structure.
- (3) The proposed copula-based HUP is comparable to the meta-Gaussian HUP regarding the posterior median forecasts. It is also shown that probabilistic forecasts produced by the copula-based HUP have slightly higher reliability and lower resolution compared with the meta-Gaussian HUP. According to the CRPS value, it is found that both HUPs are superior to deterministic forecasts which highlight the effectiveness of probabilistic forecasts, and the copula-based HUP is marginally better than the meta-Gaussian HUP.
- (4) Deterministic results of different multi-model ensembles outperform those of the individual model. The CBMA and CBP-BMA methods slightly outperform BMA method regarding *NSE*, *DRMS*, and *KGE*. When the CBMA method is used as a reference, the CBP-BMA method can improve the *NSE* and *KGE* and enlarge *DRMS* values. Underestimation of all individual models may cause negative bias of ensemble multi-model.

- (5) The combination of containing ratio and bandwidth index demonstrates the probabilistic model performance with the auxiliary index-average deviation amplitude. It is found that containing ratio is approximately equal to the percentage of the confidence interval. The CBMA or CBP-BMA methods outperform BMA method regarding evaluation criteria with a high containing ratio, small uncertainty, and average deviation amplitude.

References

- Ajami NK, Duan QY, 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):W01403
- Arya DS, Goel NK, Dharmy AP (2010) Design flow and stage computations in the Teesta River, Bangladesh, using frequency analysis and MIKE 11 modeling. *J Hydrol Eng* 16(2):176–186
- Ba HH, Guo SL, Wang Y, Hong XJ, Zhong YX (2017) Improving ANN model performance in runoff forecasting by adding soil moisture input and using data preprocessing techniques. *Hydrol Res*
- Bárdossy Göttinger J (2008) A Generic error model for calibration and uncertainty estimation of hydrological models. *Water Resour Res* 44(12):1393–1442
- Bárdossy A, Li J (2008) Geostatistical interpolation using copulas. *Water Resour Res* 44:W07412
- Bergstrand M, As SS, Lindström G (2014) Nationwide hydrological statistics for Sweden with high resolution using the hydrological model S-HYPE. *Hydrol Res* 45(3):349–356
- Biondi D, De Luca DL (2013) Performance assessment of a Bayesian forecasting system (BFS) for real-time flood forecasting. *J Hydrol* 479(1):51–63
- Biondi D, Versace P, Sirangelo B (2010) Uncertainty assessment through a precipitation dependent hydrologic uncertainty processor: an application to a small catchment in southern Italy. *J Hydrol* 386(1):38–54
- Bogner K, Pappenberger F, Cloke HL (2012) Technical note: the normal quantile transformation and its application in a flood forecasting system. *Hydrol Earth Syst Sci* 16(4):1085–1094
- Calvo B, Savi F (2009) Real-time flood forecasting of the Tiber River in Rome. *Nat Hazards* 50(3):461–477
- Carreau J, Bouvier C (2016) Multivariate density model comparison for multi-site flood-risk rainfall in the French Mediterranean area. *Stoch Env Res Risk Assess* 30(6):1591–1612
- Castellarin A, Vogel RM, Brath A (2004) A stochastic index flow model of flow duration curves. *Water Resour Res* 40(3). <https://doi.org/10.1029/2003wr002524>
- Chebana F, Ouarda TBMJ (2007) Multivariate L-moment homogeneity test. *Water Resour Res* 43(8):199–212
- Chen ST, Yu PS (2007) Real-time probabilistic forecasting of flood stages. *J Hydrol* 340:63–77
- Chen L, Guo SL, Yan B, Pan L, Fang B (2010) A new seasonal design flood method based on bivariate joint distribution of flood magnitude and date of occurrence. *Hydrol Sci J* 55(8):1264–1280
- Chen FJ, Jiao MY, Chen J (2013) The meta-Gaussian Bayesian processor of forecasts and associated preliminary experiments. *Acta Meteorologica Sinica* 27:199–210
- Chen L, Singh VP, Guo SL, Zhou J, Ye L (2014a) Copula entropy coupled with artificial neural network for rainfall-runoff simulation. *Stoch Env Res Risk Assess* 28(7):1755–1767
- Chen L, Ye L, Singh VP, Asce F, Zhou J, Guo SL (2014b) Determination of input for artificial neural networks for flood forecasting using the copula entropy method. *J Hydrol Eng* 19(11):04014021

- Chen L, Zhang Y, Zhou J, Singh VP, Guo SL, Zhang J (2015) Real-time error correction method combined with combination flood forecasting technique for improving the accuracy of flood forecasting. *J Hydrol* 521:157–169
- Cheng C, Zhao M, Chau K, Wu X (2006) Using genetic algorithm and TOPSIS for Xinanjiang model calibration with a single procedure. *J Hydrol* 316(1):129–140
- Chiew FHS, Teng J, Vaze J, Post DA, Perraud JM, Kirono DGC, Viney NR (2009) Estimating climate change impact on runoff across southeast Australia: method, results, and implications of the modeling method. *Water Resour Res* 45(10):82–90
- 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
- Dong LH, Xiong LH, Yu KX (2013) Uncertainty analysis of multiple hydrologic models using the bayesian model averaging method. *J Appl Math* 2013:1–11
- Duan QY, Ajami NK, Gao X, Sorooshian S (2007) Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Adv Water Resour* 30(5):1371–1386
- Engeland K, Renard B, Steinsland I, Kolberg S (2010) Evaluation of statistical models for forecast errors from the HBV model. *J Hydrol* 384(1):142–155
- Evin G, Thyer M, Kavetski D, McInerney D, Kuczera G (2014) Comparison of joint versus postprocessor approaches for hydrological uncertainty estimation accounting for error autocorrelation and heteroscedasticity. *Water Resour Res* 50(3):2350–2375
- Fan YR, Huang GH, Li YP, Wang XQ, Li Z (2016) Probabilistic prediction for monthly streamflow through coupling stepwise cluster analysis and quantile regression methods. *Water Resour Manag* 30(14):5313–5331
- Favre AC, Adlouni SE, Perreault L, Thiémondge N, Bobée B (2004) Multivariate hydrological frequency analysis using copulas. *Water Resour Res* 40W01101. <https://doi.org/10.1029/2003wr002456>
- Freer J, Beven K, Ambrose B (1996) Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach. *Water Resour Res* 32(7):2161–2173
- Genest C, Favre A (2007) Everything you always wanted to know about copula modeling but were afraid to ask. *J Hydrol Eng* 12(4):347–368
- Gneiting T, Raftery AE, Westveld AH, Goldman T (2005) Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon Weather Rev* 133(5):1098–1118
- Gneiting T, Balabdaoui F, Raftery AE (2007) Probabilistic forecasts, calibration and sharpness. *J Roy Stat Soc: Ser B (Stat Methodol)* 69(2):243–268
- Gottschalk L, Yu K, Leblais E, Xiong L (2013) Statistics of low flow: theoretical derivation of the distribution of minimum streamflow series. *J Hydrol* 481:204–219
- Guo SL, Zhang H, Chen H, Peng D, Liu P, Pang B (2004) A reservoir flood forecasting and control system for China. *Hydrol Sci J* 49(6):959–972
- He SK, Guo SL, Liu ZJ, Yin JB, Chen KB, Wu XS. (2018) Uncertainty analysis of hydrological multi-model ensembles based on CBP-BMA method. *Hydrol Res* (in press)
- Hemri S, Lisniak D, Klein B (2015) Multivariate post-processing techniques for probabilistic hydrological forecasting. *Water Resour Res* 51(9):7436–7451
- Hersbach H (2000) Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather Forecast* 15(5):559–570
- Hosking JRM (1990) L-moments: analysis and estimation of distributions using linear combinations of order statistics. *J Roy Stat Soc* 52(1):105–124
- Kasiviswanathan KS, Cibin R, Sudheer KP, Chaubey I (2013) Constructing prediction interval for artificial neural network rainfall runoff models based on ensemble simulations. *J Hydrol* 499:275–288
- Khajehei S, Moradkhani H (2017) Towards an improved ensemble precipitation forecast: a probabilistic post-processing approach. *J Hydrol* 546:476–489
- Klein B, Meissner D, Kobialka HU, Reggiani P (2016) Predictive uncertainty estimation of hydrological multi-model ensembles using pair-copula construction. *Water* 8(4):1–22

- Kling H, Fuchs M, Paulin M (2012) Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J Hydrol* 424–425:264–277
- Koutsoyiannis D, Montanari A (2015) Negligent killing of scientific concepts: the stationarity case. *Hydrol Sci J* 60(7–8):1174–1183
- Kroese DP, Taimre T, Botev ZI (2013) *Handbook of Monte Carlo methods*. Wiley, New York
- Krzysztofowicz R (1999) Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resour Res* 35(9):2739–2750
- Krzysztofowicz R, Kelly KS (2000) Hydrologic uncertainty processor for probabilistic river stage forecasting. *Water Resour Res* 36(11):3265–3277
- Kuczera G, Parent E (1998) Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the metropolis algorithm. *J Hydrol* 211(1–4):69–85
- Lai F, Tamea S (2007) Verification tools for probabilistic forecasts of continuous hydrological variables. *Hydrol Earth Syst Sci* 11(4):1267–1277
- Li L, Xia J, Xu CY, Singh VP (2010a) Evaluation of the subjective factors of the GLUE method and comparison with the formal Bayesian method in uncertainty assessment of hydrological models. *J Hydrol* 390(3):210–221
- Li X, Guo SL, Liu P, Chen G (2010b) Dynamic control of flood limited water level for reservoir operation by considering inflow uncertainty. *J Hydrol* 391:124–132
- Li Z, Xin P, Tang J (2011) Study of the Xinanjiang model parameter calibration. *J Hydrol Eng* 18(11):1513–1521
- Li H, Beldring S, Xu CY (2014) Implementation and testing of routing algorithms in the distributed HBV model for mountainous catchments. *Hydrol Res* 45(3):322–333
- Liang G, Kachroo RK, Kang W, Yu X (1992) River flow forecasting. part 4. applications of linear modelling techniques for flow routing on large catchments. *J Hydrol* 133(1):99–140
- Liang Z, Wang D, Guo Y, Zhang Y, Dai R (2011) Application of Bayesian model averaging approach to multimodel ensemble hydrologic forecasting. *J Hydrol Eng* 18(11):1426–1436
- Lin K, Lv F, Chen L, Singh VP, Zhang Q, Chen X (2014) Xinanjiang model combined with Curve Number to simulate the effect of land use change on environmental flow. *J Hydrol* 519:3142–3152
- Liu ZJ, Guo SL, Zhang HG, Liu DD, Yang G (2016) Comparative study of three updating procedures for real-time flood forecasting. *Water Resour Manag* 30(7):2111–2126
- Liu ZJ, Guo SL, Xiong LH, Xu CY (2017) Hydrological uncertainty processor based on a copula function. *Hydrol Sci. J.* <https://doi.org/10.1080/02626667.2017.1410278>
- Liucci L, Valigi D, Casadei S (2014) A new application of flow duration curve (FDC) in designing run-of-river power plants. *Water Resour Manag* 28(3):881–895
- Ma Z, Li Z, Zhang M, Fan Z (2013) Bayesian statistic forecasting model for middle-term and long-term runoff of a hydropower station. *J Hydrol Eng* 18(11):1458–1463
- Madadgar S, Moradkhani H, Garen D (2014) Towards improved post-processing of hydrologic forecast ensembles. *Hydrol Process* 28(1):104–122
- Madsen H (2000) Automatic calibration of a conceptual rain-fall model using multiple objectives. *J Hydrol* 235:276–288
- Möller A, Lenkoski A, Thorarinsdottir TL (2013) Multivariate probabilistic forecasting using ensemble Bayesian model averaging and copulas. *J Roy Meteorol Soc* 139(673):982–991
- Montanari A (2007) What do we mean by ‘uncertainty’? The need for a consistent wording about uncertainty assessment in hydrology. *Hydrol Process* 21(6):841–845
- Montanari A, Grossi G (2008) Estimating the uncertainty of hydrological forecasts: a statistical approach. *Water Resour Res* 44:W00B08. <https://doi.org/10.1029/2008-wr006897>
- Montero RA, Schwanenberg D, Krahe P, Lisniak D, Sensoy A, Sorman AA, Akkol B (2016) Moving horizon estimation for assimilating H-SAF remote sensing data into the HBV hydrological model. *Adv Water Resour* 92:248–257
- Nash J, Sutcliffe JV (1970) River flow forecasting through conceptual models part I-A discussion of principles. *J Hydrol* 10(3):282–290

- Nasonova ON, Gusev YM, Kovalev YE (2009) Investigating the ability of a land surface model to simulate streamflow with the accuracy of hydrological models: a case study using MOPEX materials. *J Hydrometeorol* 10(5):1128–1150
- Nelsen RB (2006) An introduction to copulas, 2nd edn. Springer, New York
- Pappenberger F, Ramos MH, Cloke HL, Wetterhall F, Alfieri L, Bogner K (2015) How do I know if my forecasts are better? Using benchmarks in hydrological ensemble prediction. *J Hydrol* 522:697–713
- Pokhrel P, Robertson D, Wang QJ (2013) A Bayesian joint probability post-processor for reducing errors and quantifying uncertainty in monthly streamflow predictions. *Hydrol Earth Syst Sci* 17(2):795–804
- Rafferty AE, Gneiting T, Balabdaoui F, Polakowski M (2005) Using Bayesian model averaging to calibrate forecast ensembles. *Mon Weather Rev* 133(5):1155–1174
- Ramos MH, van Andel SJ, Pappenberger F (2013) Do probabilistic forecasts lead to better decisions? *Hydrol Earth Syst Sci* 17(6):2219–2232
- Ravines RR, Schmidt AM, Migon HS, Rennó CD (2008) A joint model for rainfall-runoff: the case of Rio Grande Basin. *J Hydrol* 353(1):189–200
- Reggiani P, Weerts AH (2008) A Bayesian approach to decision-making under uncertainty: an application to real-time forecasting in the river Rhine. *J Hydrol* 356(1):56–69
- Renard B, Kavetski D, Kuczera G, Thyer M, Franks SW (2010) Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resour Res* 46(5) <https://doi.org/10.1029/2009wr008328>
- Robert C, Casella G (2013) Monte Carlo statistical methods. Springer, New York
- Seo DJ, Herr HD, Schaake JC (2006) A statistical post-processor for accounting of hydrologic uncertainty in short-range ensemble streamflow prediction. *Hydrol Earth Syst Sci* 3(4):1987–2035
- Shao QX, Zhang L, Chen YD, Singh VP (2009) A new method for modelling flow duration curves and predicting streamflow regimes under altered land-use conditions. *Hydrol Sci J* 54(3):606–622
- Si W, Bao W, Gupta HV (2015) Updating real-time flood forecasts via the dynamic system response curve method. *Water Resour Res* 51(7):5128–5144
- Sikorska AE, Scheidegger A, Banasik K, Rieckermann J (2012) Bayesian uncertainty assessment of flood predictions in ungauged urban basins for conceptual rainfall-runoff models. *Hydrol Earth Syst Sci* 16(4):1221–1236
- Smith LA, Suckling EB, Thompson EL, Maynard T, Du H (2015) Towards improving the framework for probabilistic forecast evaluation. *Clim Change* 132(1):31–45
- Thyer M, Renard B, Kavetski D, Kuczera G, Franks SW, Srikanthan S (2009) Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: a case study using Bayesian total error analysis. *Water Resour Res* 45(12). <https://doi.org/10.1029/2008wr006825>
- Tsai CN, Adrian DD, Singh VP (2001) Finite Fourier probability distribution and applications. *J Hydrol Eng* 6(6):460–471
- Verkade JS, Werner MGF (2011) Estimating the benefits of single value and probability forecasting for flood warning. *Hydrol Earth Syst Sci* 15(12):3751–3765
- Vogel RM, Fennessey NM (1994) Flow-duration curves. I: new interpretation and confidence intervals. *J Water Resour Plann Manag* 120(4):485–504
- Vogel RM, Fennessey NM (1995) Flow-duration curves. II: a review of applications in water resources planning. *J Am Water Resour Assoc* 31(6):1029–1039
- Vrugt JA, Robinson BA (2007) Treatment of uncertainty using ensemble methods: comparison of sequential data assimilation and Bayesian model averaging. *Water Resour Res* 43(1):223–228
- Weerts AH, Winsemius HC, Verkade JS (2011) Estimation of predictive hydrological uncertainty using quantile regression: examples from the National Flood Forecasting System (England and Wales). *Hydrol Earth Syst Sci* 15(1):255–265
- Wetterhall F, Pappenberger F, Cloke HL, Pozo JT (2013) Forecasters priorities for improving probabilistic flood forecasts. *Hydrol Earth Syst Sci* 17(11):4389–4399

- Wu J, Zhou JZ, Chen L, Ye L (2015) Coupling forecast methods of multiple rainfall–runoff models for improving the precision of hydrological forecasting. *Water Resour Manage* 29 (14):5091–5108
- Wu XS, Wang ZL, Guo SL, Liao W, Zeng Z, Chen X (2017) Scenario-based projections of future urban inundation within a coupled hydrodynamic model framework: a case study in Dongguan City, China. *J Hydrol* 547:428–442
- Xiong LH, Guo SL (1999) A two-parameter monthly water balance model and its application. *J Hydrol* 216(1):111–123
- Xiong LH, Min W, Wei XJ, O'Connor KM (2009) Indices for assessing the prediction bounds of hydrological models and application by generalised likelihood uncertainty estimation. *Hydrol Sci J* 54(5):852–871
- Xiong LH, Yu KX, Gottschalk L (2014) Estimation of the distribution of annual runoff from climatic variables using copulas. *Water Resour Res.* <https://doi.org/10.1029/2008WR006897>
- Xiong L, Du T, Xu CY, Guo SL, Jiang C, Gippel CJ (2015) Non-Stationary annual maximum flood frequency analysis using the norming constants method to consider non-stationarity in the annual daily flow series. *Water Resour Manag* 29(10):3615–3633
- Xu H, Xu CY, Chen H, Zhang Z, Li L (2013) Assessing the influence of rain gauge density and distribution on hydrological model performance in a humid region of China. *J Hydrol* 505:1–12
- Yokoo Y, Sivapalan M (2011) Towards reconstruction of the flow duration curve: development of a conceptual framework with a physical basis. *Hydrol Earth Syst Sci* 15(9):2805–2819
- Yu B, Zhu Z (2015) A comparative assessment of AWBM and SimHyd for forested watersheds. *Hydrol Sci J* 60(7):1–13
- Yu KX, Xiong LH, Gottschalk L (2014) Derivation of low flow distribution functions using copulas. *J Hydrol* 508:273–288
- Zhang L, Singh VP (2006) Bivariate flood frequency analysis using the copula method. *J Hydrol Eng* 11(2):150–164
- Zhang L, Singh VP (2007a) Bivariate rainfall frequency distributions using Archimedean copulas. *J Hydrol* 332(1):93–109
- Zhang L, Singh VP (2007b) Gumbel-Hougaard copula for trivariate rainfall frequency analysis. *J Hydrol Eng* 12(4):409–419
- Zhang L, Singh VP (2007c) Trivariate flood frequency analysis using the Gumbel-Hougaard copula. *J Hydrol Eng* 12(4):431–439
- Zhang Q, Chen YD, Chen X, Li J (2011) Copula-based analysis of hydrological extremes and implications of hydrological behaviors in the Pearl River basin, China. *J Hydrol Eng* 16 (7):598–607
- Zhang Q, Li J, Singh VP (2012) Application of Archimedean copulas in the analysis of the precipitation extremes: effects of precipitation changes. *Theor Appl Climatol* 107(1–2):255–264
- Zhang J, Chen L, Singh VP, Cao W, Wang D (2015) Determination of the distribution of flood forecasting error. *Nat Hazards* 75(2):1389–1402
- Zhao RJ (1992) The Xinanjiang model applied in China. *J Hydrol* 135(1–4):371–381
- Zhao T, Wang QJ, Bennett JC, Robertson DE, Shao Q, Zhao J (2015) Quantifying predictive uncertainty of streamflow forecasts based on a Bayesian joint probability model. *J Hydrol* 528:329–340
- Zhou YL, Guo SL, Xu CY, Chen H, Guo J, Lin K (2016) Probabilistic prediction in ungauged basins (PUB) based on regional parameter estimation and Bayesian model averaging. *Hydrol Res* 47(6):1087–1103