ISEIS

Journal of Environmental Informatics 38(2) 131-144 (2021)

Journal of Environmental Informatics

www.iseis.org/jei

Parameter Uncertainty and Sensitivity Evaluation of Copula-Based Multivariate Hydroclimatic Risk Assessment

K. Huang[1] and Y. R. Fan[2 *]

[1] *Faculty of Engineering and Applied Science, University of Regina, Regina, SK S4S 0A2, Canada*
[2] *Department of Civil and Environmental Engineering, Brunel University, London, Uxbridge, Middlesex UB8 3PH, United Kingdom*

Received 08 April 2020; revised 20 May 2020; accepted 17 June 2020; published online 03 March 2021

ABSTRACT. Extensive uncertainties exist in hydroclimatic risk analysis. Especially in multivariate hydrologic risk inferences, uncertainties in individual hydroclimatic extremes such as floods and their dependence structure may lead to bias and uncertainty in future hydrologic risk predictions. In this study, a parameter uncertainty and sensitivity evaluation (PUSE) framework is proposed to quantify parameter uncertainties and then reveal their contributions to the multivariate hydroclimatic risk predictions. The predictive risks are finally generated by "integrating" the values over the posterior distributions of the parameters. The proposed approach was applied for bivariate risk analysis of compound floods at the Xiangxi River to characterize the concurrence probabilities of flood peaks and volumes. The results demonstrate that the proposed approach can quantify uncertainties in a copula-based multivariate risk analysis and characterize effects and contributions of parameters in marginal and dependence structures on the multivariate hydroclimatic risk predictions. In terms of the bivariate risk for flood peak and volume at the Xiangxi River, uncertainties in model parameters would lead to noticeable uncertainties even for moderate floods. The performances of the copula model for flood peak-volume at Xiangxi River are mainly affected by the uncertainties in location parameters of the two individual flood variables. Also, parameter uncertainty in the dependence structure (i.e., copula) would also poses explicit impacts on performance of the copula-based risk analyses model. These uncertainties would result into higher bivariate predictive risks than the values obtained by "optimal/deterministic" predictions. This indicates that uncertainties are required to be considered to provide reliable multivariate hydroclimatic risk predictions.

Keywords: Hydroclimatic risk, Copula, MCMC, factorial analysis, global sensitivity analysis

1. Introduction

1.1. Importance and Motivation

Hydroclimatic extremes are expected to increasingly occur especially under climate change. Flood, as one of the most frequently occurred hydroclimatic hazards, has taken a devastating societal and economic toll over the world, leading to a large number of fatalities and property losses. (Fan et al., 2015, 2016a, 2016b). Consequently, design of water infrastructure projects, such as dam spillways, urban drainage systems, needs characterize the occurrence probabilities of floods (Tan et al., 2011; Zhai et al., 2021; Dong et al., 2021). Moreover, flood risk is often not only determined by peak discharge, but is a multidimensional problem (Dung et al., 2015). Multivariate flood risk analysis is desired to provide a full screen for a flood. The applications of multivariate flood risk analyses are growing dramatically since the introduction of copulas in hydrology and geosciences (Serinaldi, 2013). De Michele and Salvador (2003) initially introduced the concept of copulas into hydrological simulation, which described the dependence between storm duration and average rainfall intensity by means of a suitable 2-Copula. After that, a great number of research works have been proposed for multivariate hydrologic simulation through copula functions, such as multivariate flood frequency analysis (e.g., Zhang and Singh, 2006; Li et al., 2008; Sraj et al., 2015; Xu et al., 2017), drought assessments (e.g., Kao and Govindaraju, 2010; Song and Singh, 2010; Sun et al., 2019), storm or rainfall dependence analysis (e.g., Zhang and Singh, 2007; Vandenberghe et al., 2010), climate downscaling (e.g., Zhou et al., 2018; Sun et al., 2021), and so on. Copula functions can estimate the marginal distributions and the joint dependence models in two separate processes, giving additional flexibility in choosing different marginal and joint probability functions (Zhang and Singh, 2006; Genest and Favre, 2007; Karmakar and Simonovic, 2009; Li et al., 2010; Sraj et al., 2015; Huang et al., 2017).

One major issue in hydroclimatic risk analysis is the presence of uncertainties, resulting from model selection and parameter estimation. There are two primary sources of uncertainty: (1) natural uncertainty stemming from variability of the underlying stochastic process, and (2) epistemic uncertainty coming from incomplete knowledge about the system under study (Merz and Thieken, 2005). Dung et al. (2015) proposed a bootstrapping based algorithm to investigate parameter estimation method uncertainty, model selection uncertainty and sampling uncertainty. The results showed that bivariate flood frequency analy-

* Corresponding author. Tel.: +44 (0)1895 265717.
E-mail address: yurui.fan@brunel.ac.uk (Y. R. Fan).

ISSN: 1726-2135 print/1684-8799 online
 doi:10.3808/jei.202100462

sis is expected to carry significant uncertainty and that the quantification and reduction of uncertainty merit greater attention (Dung et al., 2015). However, the research works on uncertainty assessment for multivariate hydroclimatic risk analysis are still limited. Parameter uncertainty can be well quantified through some Monte Carlo based algorithms and uncertainty predictions are then obtained. But two issues may arise: (i) in a multivariate framework, parametric uncertainties in the marginal and dependence structures may lead to great variations in multivariate hydroclimatic risk analysis, but few research is reported to reveal how the interactions of parameter uncertainties affect the performance of the multivariate model; (ii) the aim of a statistical inference is sometimes (even frequently) not parameter estimation, but rather prediction of an unobserved value, but only limited studies are proposed to explore the predictive multivariate hydroclimatic inferences under parameter uncertainties.

Consequently, as an extension of previous research works, this study aims to propose a parameter uncertainties and sensitivities evaluation (PUSE) framework for copula-based multivariate hydroclimatic risk assessment. The PUSE method integrates the capabilities of copula method, Bayesian inference, factorial and sensitivity analyses to quantify parametric uncertainties in multivariate flood risk inferences and further characterize the impacts of parameter uncertainties and generate reliable risk predictions. The proposed PUSE approach will apply for multivariate flood risk assessment in the Xiangxi River located in the Three Gorges Reservoir area in China.

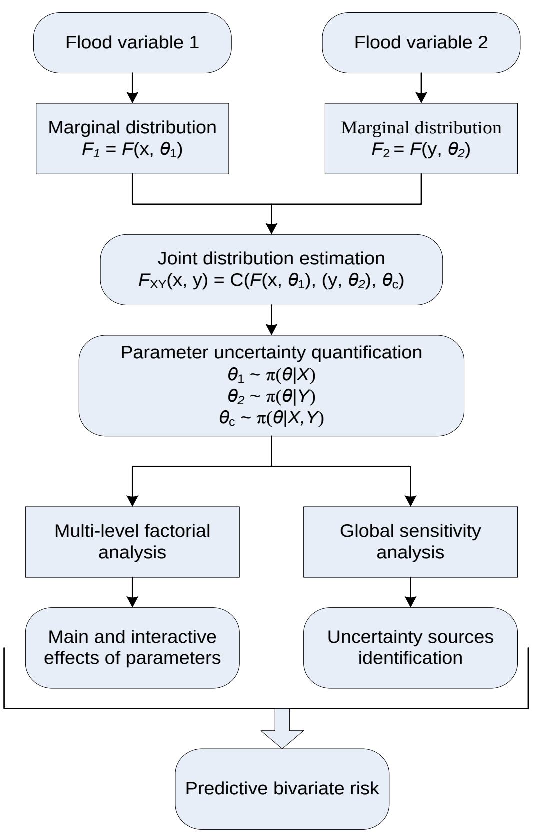

Figure 1. Framework of PUSE.

2. Methodology

Figure 1 presents the detailed procedures for the proposed PUSE framework. In detail, the parameter uncertainties are quantified through a Bayesian inference method. The interactions of parameter uncertainties are explored through a multilevel factorial analysis approach and the contributions of parameter uncertainties are analyzed through Sobol's based sensitivity analysis method. The predictive risks will be finally generated through "integrating out" the parameters of the multivariate risk model.

2.1. Concept of Copula

A copula is a multivariate distribution function with uniform marginals on the interval [0, 1], which can set up a link between the joint distribution and its marginal distribution functions (Dung et al., 2015). In detail, a multivariate copula function can be expressed as:

$$F(x_1, x_2, \ldots, x_n) = C(F_{X_1}(x_1), F_{X_2}(x_2), \ldots, F_{X_n}(x_n)) \quad (1)$$

where $F_{X_1}(x_1), F_{X_2}(x_2), \ldots, F_{X_n}(x_n)$ are marginal distributions of the random vector ($X_1, X_2, \ldots, X_n$). If these marginal distributions are continuous, then a single copula function C exists, which can be written as (Sraj et al., 2015):

$$C(u_1, u_2, \ldots, u_n) = F(F_{X_1}^{-1}(u_1), F_{X_2}^{-1}(u_2), \ldots, F_{X_n}^{-1}(u_n)) \quad (2)$$

More details on theoretical background and properties of various copula families can be found in Nelsen (2006).

A number of copula functions are widely used in hydroclimatic risk analyses, mainly including the Archimedean, elliptical, extreme value copulas. In this study, the Archimedean copulas will be employed since they can be easily generated and are capable of capturing wide range of dependence structures with several desirable properties, such as, symmetry and associativity (Ganguli and Reddy, 2013). The Clayton, Gumbel and Frank copulas are considered for probabilistic assessment of flood risk, which belong to the class of Archimedean copula. For a bivariate model, the corresponding Archimedean copula can be formulated as (Nelsen, 2006):

$$C_\theta(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2)) \quad (3)$$

where u_1 and u_2 is a specific value of U_1 and U_2, respectively; $U_1 = F_{X_1}(x_1)$ and $U_2 = F_{X_2}(x_2)$; F_{X_1} and F_{X_2} is the cumulative distribution function (CDF) of random variable X_1 and X_2, respectively; ϕ is the copula generator that is a convex decreasing function with $\phi(1) = 0$ and $\phi^{-1}(.) = 0$ when $u_2 \geq \phi(0)$; the subscript θ of copula C is the parameter hidden in the generating function.

Once the copula function is established, the probability density function (pdf) of the copula function can be expressed as (Fan et al., 2018):

$$c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2} \tag{4}$$

The joint pdf of the two random variables can be obtained as:

$$\begin{aligned} f(x_1, x_2) &= \frac{\partial^2 C(u_1, u_2)}{\partial x_1 \partial x_2} = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2}\frac{\partial u_1}{\partial x_1}\frac{\partial u_2}{\partial x_2} \\ &= f_{X_1}(x_1) f_{X_2}(x_2) c(u_1, u_2) \end{aligned} \tag{5}$$

Consequently, the conditional pdf of X_1, given the value of X_2, can be formulated as:

$$f(x_1 | x_2) = \frac{f(x_1, x_2)}{f_{X_2}(x_2)} = f_{X_1}(x_1) c(u_1, u_2) \tag{6}$$

The conditional pdf of X_2, given the value of X_1, can be expressed as:

$$f(x_2 | x_1) = \frac{f(x_1, x_2)}{f_{X_1}(x_1)} = f_{X_2}(x_2) c(u_1, u_2) \tag{7}$$

Various methods have been developed for estimating the unknown parameters in a copula, such as the method of moments with the use of Kendall correlation coefficient, pseudo maximum likelihood (PML) or canonical maximum likelihood (CML) method, inference from margin method (IFM), and exact maximum likelihood (EML) (Dung et al., 2015). For one parameter copula, the unknown parameter (i.e., θ) can be estimated using the method of moments with the use of Kendall correlation coefficient (Nelsen, 2006). For the copulas with two or more unknown parameters, the maximum likelihood method or maximum pseudo-likelihood method can be selected (Sraj et al., 2015). Table 1 presents some basic characteristics of the applied single-parameter bivariate Archimedean copulas.

If appropriate copula functions are specified to reflect the joint probabilistic characteristics among correlated flood variables, the conditional, primary and secondary return periods can be obtained. Specifically, the joint (primary) return periods called OR and AND can be formulated as (Salvadori et al., 2007, 2011; Graler et al., 2013; Sraj et al., 2015; Fan et al., 2018):

$$T_{u_1, u_2}^{OR} = \frac{\mu}{1 - C_{U_1 U_2}(u_1, u_2)} \tag{13}$$

$$T_{u_1, u_2}^{AND} = \frac{\mu}{1 - u_1 - u_2 + C_{U_1 U_2}(u_1, u_2)} \tag{14}$$

where μ is the mean inter arrival time of the two consecutive flooding events.

The secondary return period, called Kendall's return period, is defined as follows (Salvadori et al., 2011; Sraj et al., 2015):

$$\bar{T}_{u_1, u_2} = \frac{\mu}{1 - K_C(t)} \tag{15}$$

where K_C is the Kendall's distribution, associated with theoretical Copula function C_θ. For Archimedean copulas, K_C can be expressed as (Nelsen, 2006):

$$K_C(t) = t - \frac{\phi(t)}{\phi'(t^+)} \tag{16}$$

where $\phi'(t^+)$ is the right derivative of the copula generator function $\phi(t)$.

Furthermore, the hydrologic risk can be characterized through the joint return periods obtained by copulas. In engineering design of hydrologic infrastructures, risk can be explained as the chance of downstream flood attributable to uncontrolled water release from upstream flood facilities (e.g., a reservoir), leading to life and property losses (Gebregiorgis and Hossain, 2012). Yen (1970) proposed a formulation for the risk of failure associated with the return period of a flood event, which can be expressed as:

$$R = 1 - (1 - p)^n = 1 - q^n = 1 - (1 - 1/T)^n \tag{17}$$

where R is the risk of failure; p and q is the exceedance and nonexceedance probability, respectively; T is the return period of a flood event; n is the design life of the hydraulic structure. In multivariate framework, the multivariate hydrologic risk can be similarly defined. In this study, the joint return period in "AND" case is adopted to define the bivariate risk analysis as follows (Fan et al., 2018):

$$R_{u_1, u_2} = 1 - (1 - \frac{1}{T_{u_1, u_2}^{AND}})^n \tag{18}$$

2.2. Uncertainty in the Copula Model

Extensive uncertainties may be involved in the parametric estimation of a copula function due to: (i) the inherent uncertainty in the flooding process, (ii) uncertainty in the selection of appropriate marginal functions and copulas and (iii) statistical uncertainty or parameter uncertainty within the parameter estimation process (e.g., the availability of samples). In this study, we will analyze the inherent uncertainty in the copula model through Bayesian analysis, which provides the posterior probability distributions for the unknown parameters in the copula model. The Bayesian approach has been widely applied for uncertainty quantification since it can incorporate various sources of information into a singly analysis through Bayes' theorem. Given the prior probability density and observations, the posterior distribution can be derived through Bayes' theorem, which is expressed as:

$$\pi(\theta | X) = \frac{L(\theta | X)\pi_0(\theta)}{\int L(\theta | X)\pi_0(\theta) d\theta} \tag{19}$$

where $\pi_0(\theta)$ signifies the prior parameter distribution, and $L(\theta|X)$ denotes the likelihood function. $\int L(\theta|X)\pi_0(\theta)\mathrm{d}\theta$ is the normalization constant. $\pi(\theta|X)$ is the posterior probability density function. $X = (x_1, x_2, \ldots, x_n)$ is the observation vector ($X = (x_1, x_2)$ in this study).

For the bivariate hydrologic risk analysis through copula, it is necessary to estimate the parameter of the copula, as well as the parameters of the two marginal probability distributions. Let θ_c, θ_1, θ_2 denote the parameters of the copula, and the two marginal probability distributions. The posterior distribution can be derived as follows (Fan et al., 2018):

$$\begin{aligned}&\pi(\theta_c,\ \theta_1,\ \theta_2 \mid X,\ Y) \propto L(\theta_c,\ \theta_1,\ \theta_2 \mid X,\ Y)\pi_0(\theta_c,\ \theta_1,\ \theta_2)\\&= L(\theta_c,\ \theta_1,\ \theta_2 \mid X,\ Y)\pi_0(\theta_c)\pi_0(\theta_1)\pi_0(\theta_2)\end{aligned} \tag{20}$$

The term of $L(\theta_c, \theta_1, \theta_2|X, Y)$ is the likelihood function of observations. Based on the dependence structure between a copula and its marginal distributions, as expressed by Equation (5), the likelihood function can be estimated as (Fan et al., 2018):

$$\begin{aligned}&L(\theta_c,\ \theta_1,\ \theta_2 \mid X,\ Y)\\&= c_{U_1U_2}(F_X(x),\ F_Y(y) \mid \theta_c)\, f_X(x \mid \theta_1)\, f_Y(y \mid \theta_2)\end{aligned} \tag{21}$$

where $c_{U_1U_2}$ is the density of the copula function and f_X and f_Y are the two marginal probability density functions, respectively.

The procedures to derive the posterior distributions are presented as follows:

- Step 1: Set the prior distributions for the unknown parameters. In this study, the unknown parameters in the marginal distributions and the copula function are assumed to be uniformly distributed within intervals.
- Step 2: Use MCMC with Metropolis–Hastings algorithm to derive the posterior probabilities for the parameters in the marginal and joint distributions.
- Step 3: For each MCMC iteration j (if total J iterations are used) after the burn-in period, estimate the quantile set (also, quantile curve QC* in bivariate context) and construct the probability density function of X-Y-pairs in the quantile set.
- Step 4: Sample a large number N of (x, y) pairs from QC* using the individual density functions of coincidence of (x, y).
- Step 5: Generate a random integer number r in $[1, 2, \ldots, J]$ for random selection of the QC* from MCMC results.
- Step 6: Sample a point along the curve QC*.
- Step 7: Repeat steps 8 and 9 N times.

2.3. Interactive and Sensitivity Analysis for Parameters Uncertainties

Due to the uncertainties existing in the unknown parameters for a copula model, the associated risk or the return period for a flooding event may also be uncertain. Few studies are reported to analyze the effect of uncertainties in the copula model on risk evaluation for a flood event. To address the above issue, a multi-level factorial analysis will be employed to assess uncertainties in parameters and their interactive effect on the performance of the proposed copula model. A global sensitivity analysis approach will be further adopted to explore the detailed contributions of uncertainties in the marginal distributions and dependence structure to performances of the copula model.

Factorial designs are the cornerstone of industrial experimentation and used extensively in industrial research and development for process improvement, among which the multi-level factorial design is a powerful statistical technique to study the effects of several independent variables (factors) with multiple levels on a dependent variable (response) (Wang and Huang, 2015; Fan et al., 2020a, 2020b, 2021). A 3^k factorial design is proposed for screening the effects of parameters in a copula model on its performance for evaluating flooding risk. The 3^k factorial design consists of k factors with each factors having three levels. A 3^k factorial design contains 3^k treatment combinations with a degrees of freedom of $3^k - 1$. These treatment combinations can generate sums of squares for k main effects with each having two degrees of freedom, $\binom{k}{2}$ two-factor interactions with each having four degrees of freedom, …, and one k-factor interaction with 2k degrees of freedom (Wang and Huang, 2015). In general, an h-factor has a degrees of freedom of 2h – 1, and can be further partitioned into 2h – 1 orthogonal two-degrees-of-freedom components (Montgomery, 2001). Through the 3k factorial design, the main effect of the model parameters and their interactions can be revealed. In this study, the factors are the unknown parameters of the copula model (both in marginal distributions and the copula function). The three levels for each parameter consist of its 2.5% quantile value, mean and 97.5% quantile value, which are obtained by the sample values from MCMC.

Moreover, the contributions of parameters uncertainties to the uncertainty in model output will be further quantified through the Sobol's global sensitivity analysis (GSA) approach. By using Sobol' GSA, the main effect "first-order sensitivity index" and total effect of the input can be quantified. In Solbol's method, the total variance of outputs attributed to individual model parameters and their interactions can be expressed as (Zhang et al., 2013; Song et al., 2015):

$$V = \sum_{i=1}^{n} V_i + \sum_{i=1}^{n}\sum_{j>i}^{n} V_{ij} + \ldots + V_{1,\,2,\,\ldots,\,n} \tag{22}$$

where V represents the total variance of the model output, V_i denotes the first-order variance due to the ith factor, V_{ij} denotes the interaction between factor i and j. The sensitivity of each factor can be quantified based on their percentage contribution to the total variance:

$$\text{First-order index} \quad S_i = \frac{V_i}{V} \tag{23}$$

$$\text{Total-order index} \quad S_{T_i} = 1 - \frac{V_{\sim i}}{V} \tag{24}$$

Table 1. Basic Properties of Applied Copulas

Copula Name	Function[$C_\theta(u_1, u_2)$]	$\theta \in$	Generating functions [$\phi(t)$]	$\tau = 1 + 4\int_0^1 \frac{\phi(t)}{\phi'(t)} dt$
Ali-Mikhail-Haq	$\frac{u_1 u_2}{[1-\theta(1-u_1)(1-u_2)]}$	[-1, 1)	$\ln(\frac{[1-\theta(1-t)]}{t})$	$\frac{3\theta-2}{\theta} - [\frac{2}{3}(1-\theta^{-1})^2 \ln(1-\theta)]$
Clayton	$[u_1^{-\theta} + u_2^{-\theta} - 1]^{-1/\theta}$	(0, ∞)	$t^{-\theta} - 1$	$\frac{\theta}{\theta+2}$
Gumbel-Hougaard	$\exp\{-[(-\ln u_1)^\theta + (-\ln u_2)^\theta]^{1/\theta}\}$	[1, ∞)	$(-\ln t)^\theta$	$1-\theta^{-1}$
Frank	$-\frac{1}{\theta}\ln\{1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\}$	(-∞, ∞)\{0}	$\ln[\frac{e^{-\theta t}-1}{e^{-\theta}-1}]$	$1-\frac{4}{\theta}[D_1(-\theta)-1]^*$

2.4. Predictive Probability under Uncertainty

The ultimate aim of a statistical inference is sometimes (even frequently) not parameter estimation, but rather prediction of an unobserved value (Renard et al., 2013). In practical hydraulic structure design through the statistical inference model, the operation objective is to identify the acceptable hydrologic event (e.g., flood for dams) rather than only providing the unknown parameter values in the model. For instance, if a dam is designed to prevent a flood with a 100-year return period, the question is what characteristics a 100-year return period would have (e.g., peak flow, volume), especially when the parameters in the statistical model are uncertain? In this study, the predictive distribution proposed by Renard et al. (2013) would be employed to estimate the distribution of an unobserved outcome through "integrating out" the parameters of the statistical model, which can be expressed as:

$$
\begin{aligned}
p(z \mid \boldsymbol{y}) &= \int p(z, \boldsymbol{\theta} \mid \boldsymbol{y}) d\boldsymbol{\theta} = \int p(z \mid \boldsymbol{\theta}, \boldsymbol{y}) p(\boldsymbol{\theta} \mid \boldsymbol{y}) d\boldsymbol{\theta} \\
&= \int p(z \mid \boldsymbol{\theta}) p(\boldsymbol{\theta} \mid \boldsymbol{y}) d\boldsymbol{\theta}
\end{aligned} \tag{25}
$$

where $p(z \mid \boldsymbol{\theta})$ corresponds to the pdf of the assumed distribution, $p(\boldsymbol{\theta} \mid \boldsymbol{y})$ denotes the posterior distribution of model parameters. In a bivariate context, the occurrence probability of a bivariate hydrologic event (i.e., flood peak and volume) can be obtained by:

$$
\begin{aligned}
& p(x^*, y^* \mid X, Y) = \\
& \int c(F_X(x \mid \theta_1), F_Y(y \mid \theta_2) \mid \theta_c) \pi(\theta_c, \theta_1, \theta_2 \mid X, Y) d\theta_1 d\theta_2 d\theta_c
\end{aligned} \tag{26}
$$

where the $c(F_X(x \mid \theta_1), F_Y(y \mid \theta_2) \mid \theta_c)$ is the probability density function for the copula, and $\pi(\theta_c, \theta_1, \theta_2 \mid X, Y)$ is the posterior probability density function for the unknown parameters. Similarly, the non-exceedance probability for a predefine observation pair can be generated by:

$$
\begin{aligned}
& p(X' \le x^*, Y' \le y^* \mid X, Y) = \\
& \int C(F_X(x \mid \theta_1), F_Y(y \mid \theta_2) \mid \theta_c) \pi(\theta_c, \theta_1, \theta_2 \mid X, Y) d\theta_1 d\theta_2 d\theta_c
\end{aligned} \tag{27}
$$

where $C(F_X(x \mid \theta_1), F_Y(y \mid \theta_2) \mid \theta_c)$ is the copula function.

Compared with the standard approach for prediction through maximizing the posterior pdf, the predictive distribution uses the assumed distribution $p(z \mid \boldsymbol{\theta})$ integrated over possible posterior realizations of $\boldsymbol{\theta}$, which can accounts for uncertainties in prediction model (Renard et al., 2013). In practice, the integrations in Equations (12) and (13) can hardly be obtained analytically. Consequently, they will be approximated using the MCMC samples. Based on the predictive pdf, the associated predictive risk for a predefined hydrological event can be evaluated.

3. Overview of the Studied Watershed

The Xiangxi River is located between 30.96 ~ 31.67 °N and 110.47 ~ 111.13 °E in Hubei part of China Three Gorges Reservoir (TGR) region, draining an area of about 3200 km^2, as shown in Figure 2. The Xiangxi River originates in the Shennongjia Nature Reserve with a main stream length of 94 km, which is one of the main tributaries of the Yangtze River (Han et al., 2014). The Xiangxi River watershed experiences a northern subtropics climate. Annual precipitation is 1100 mm and ranges from 670 to 1700 mm with considerable spatial and temporal variability (Xu et al., 2010). The main rainfall season is May to September, with a flooding season from July to August. Approximately 70% of the precipitation received between May and September, is rainfall and over 80% of the area is mountainous, and the land cover is dominated by mixed needle-leaf and broad-leaf forests (Li et al., 2015). The annual average temperature in this region is 15.6 ℃ and ranges from 12 to 20 ℃.

We investigate the bivariate hydrologic risk for flood peak (Q) and volume (V) at the Xiangxi River, based on the measurements at Xingshan Hydrologic Station. The Xingshan Hydrologic Station (110°45'0''E, 31°13'0''N) is located on the main stem of Xiangxi River, with a drainage area of 1,900 km^2. It is the main control gauge station with an average flow volume of 1.27×10^9 m^3. The measured minimum and maximum flow volume are 5.7×10^8 (in 1966) and 2.15×10^9 m^3 (in 1988), respecttively. The biggest flood occurred in 1935 with an estimated flood volume of 2770 m^3/s. In this study, total fifty years' daily discharge data (1961 ~ 2010) from Xingshan Hydrologic Station would be used for probabilistic assessment of flood risks in Xiangxi River. Based on the daily stream flow data, the flood peak applied in this study is defined as the maximum daily flow during the flood event. The flood volume being considered as the cumulative flow volume during the flood period. Such flood characteristics are obtained based on the annual scale, meaning in each year one flood would occur. The detailed method to identify the flood peak and the associated flood volume can be

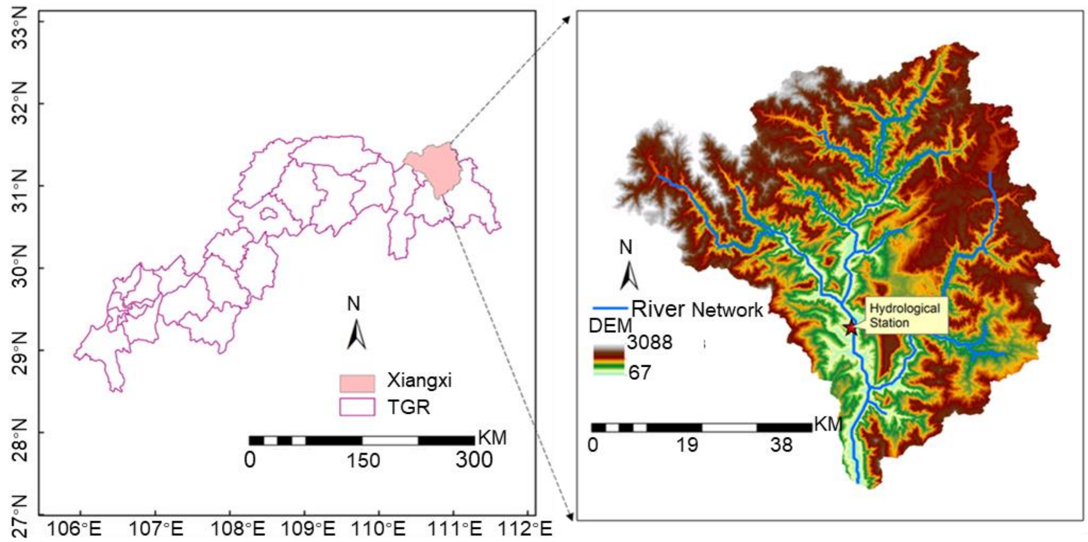

Figure 2. The location of the Xiangxi River basin.

found in Yue (2000, 2001). Table 2 shows some descriptive statistics values of the considered variables (peak discharge, *Q*; hydrograph volume, *V*). The positive values of kurtosis and skewness suggest that the flood variables can be modeled by sharp and right tailed distributions.

Table 2. Statistical Characteristics of Flood Variables

No.			Flood characteristics	
			Peak	Volume
1	Percentile	Minimum	91	72
		25%	324	530.8
		50%	451.5	713.3
		75%	684	1189.5
		Maximum	1050	2430.8
2	Range		959	2358.8
3	Mean		510	920.5
4	Std		243.8	531.9
5	Skewness		0.74	0.959
6	Kurtosis		2.61	3.20

4. Results Analysis

4.1. Marginal Distribution Analysis

The univariate flood frequency analyses would be performed based on the historical flooding records. A number of parametric or non-parametric approaches have been proposed to estimate the distributions of flood variables (Karmakar and Simonovic, 2009; Sun et al., 2019). For the distribution of flood peak, different distributions have been recommended in different regions, such as the general extreme value distribution in the United Kingdom, Log-Pearson Type-III in the U.S. and Pearson III in China (Adamowski, 1989, Kidson and Richards, 2005, Wu et al., 2013). In this study, several parametric distribution functions will be employed to quantify the distributions of the flood peak and volume, including Gamma, generalized extreme value (GEV), Lognormal, Pearson Type III, and Log-Pearson Type III distributions. The expressions of these distribution functions are presented in Table 3. The parameters in these distributions are estimated through maximum likelihood estimation (MLE) method.

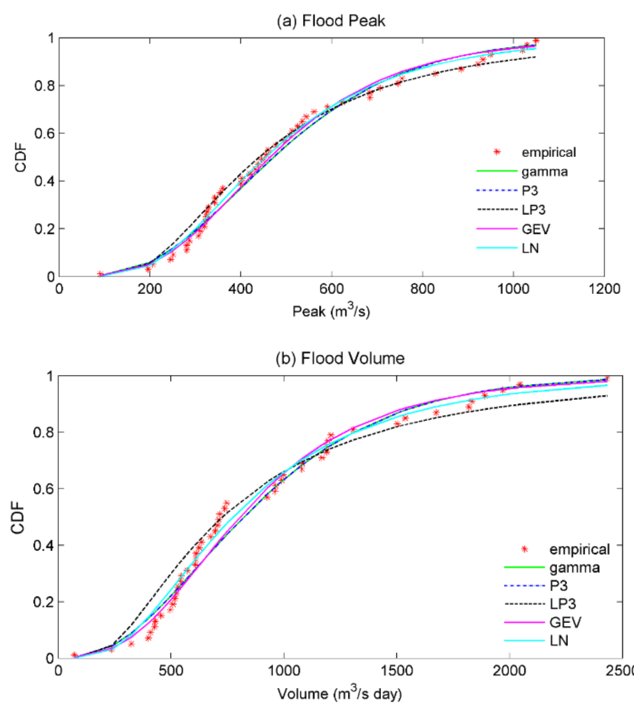

Figure 3. Comparison of different distributions in quantifying the flood variables.

Figure 3 illustrates the fitted marginal distributions for the flood peak and volume through Gamma, GEV, Lognormal, Pearson Type III and Log-Pearson Type IIT distribution func-

Table 3. Parameters of Marginal Distribution Functions of Flood Variables

Name	Probability density function		Parameters		
			Peak	Volume	Duration
Gamma	$\frac{1}{b^a\Gamma(a)}x^{a-1}e^{-\frac{x}{b}}$, $\Gamma(a)=\int_0^\infty u^{a-1}e^{-u}du$	a	4.50	3.06	8.72
		b	113.26	301.24	0.75
GEV	$(\frac{1}{\sigma})\exp(-(1+k\frac{(x-\mu)}{\sigma})^{-\frac{1}{k}})(1+k\frac{(x-\mu)}{\sigma})^{-1-\frac{1}{k}}$	k	0.032	0.099	0.054
		μ	185.0	373.16	1.735
		σ	396.15	664.96	5.427
Lognormal	$\frac{1}{x\sqrt{2\pi\sigma_y}}\exp(-\frac{(y-\mu_y)^2}{2\sigma_y^2})$ $y=\log(x)$, $x>0$, $-\infty<\mu_y<\infty$, $\sigma_y>0$	μ_y	6.12	6.65	1.82
		σ_y	0.50	0.63	0.34
Pearson Type III	$\frac{1}{b^a\Gamma(a)}(x-\alpha)^{a-1}e^{-\frac{(x-\alpha)}{b}}$, $\Gamma(a)=\int_0^\infty u^{a-1}e^{-u}du$	a	4.28	3.14	2.42
		b	116.73	301.28	1.54
		α	10.91	-20.24	2.80
Log-Pearson Type III	$\frac{1}{xb^a\Gamma(a)}(\log(x)-\alpha)^{a-1}e^{-\frac{(\log(x)-\alpha)}{b}}$	a	54.66	75.31	17.61
		b	0.071	0.079	0.083
		α	2.25	0.676	0.355

Table 4. Statistical Test Results for Marginal Distribution Estimation

Flooding variables	Marginal distribution	A-D test		RMSE	AIC
		T	P-value		
Peak	Gamma	0.4888	0.7574	0.0370	-325.5954
	Pearson Type III	0.4731	0.7735	0.0361	-326.0715
	Log Pearson Type III	0.4133	0.8348	0.0273	-354.0118
	GEV	0.4142	0.8339	0.0329	-335.3050
	Lognormal	0.3356	0.9088	0.0260	-358.9344
Volume	Gamma	0.5863	0.6602	0.0436	-309.2290
	Pearson Type III	0.5996	0.6475	0.0448	-304.5710
	Log Pearson Type III	0.8073	0.4752	0.0428	-309.0321
	GEV	0.5084	0.7374	0.0415	-312.2103
	Lognormal	0.5249	0.7207	0.0353	-328.2738

tions. The results indicate that the five distribution functions can fit the empirical probabilities well. Moreover, the goodness-of-fit test for the used distributions are tested through the Kolmogorov-Smirnov test. The results presented in Table 4 indicate that all the five parametric distributions can produce satisfactory results, with all the *p*-values larger than 0.05. Furthermore, the performance of each marginal distribution is evaluated against the empirical non-exceedance probability, using root mean square error (RMSE) and Akaike Information Criterion (AIC) criteria, which aims to identify the most appropriate distribution function for flood peak and volume. The results are presented in Table 4. It can be concluded that the Lognormal distributions would be most appropriate in modeling the distributions of flood peak and volume, which leads to lowest RMSE and AIC values.

4.2. Joint Probability Distribution

Three Archimedean copulas, including Cook-Johnson (Clayton), Gumbel-Hougaard and Frank copulas are employed to model the dependence among flood peak-volume. The Ali-Mikhail-Haq copula is excluded in this study since it can only be applicable when the Kendall's tau coefficient between the two random variables is located within [-0.18, 0.33]. However, the Kendall's tau coefficient between flood peak and volume is 0.75. The unknown parameter in these copulas can be estimated by method-of-moments-like (MOM) estimator based on inversion of Kendall's tau. The goodness-of-fit statistics is performed based on the Cramér von Mises statistic proposed by Genest et al. (2009). The root mean square error (RMSE) and Akaike information criterion (AIC) are further used to evaluate the performance of the obtained copulas and identify the most appropriate one.

Table 5 shows statistical test results for the three copulas based on the method proposed by Genest et al. (2009). The results show that the Frank and Gumbel copulas can generated satisfactory results, with the *p*-values being 0.42 and 0.45, respectively. In comparison, the Clayton copula produces unsatisfactory results in modelling the dependence between flood peak and volume, with the *p*-value of goodness-of-fit test less than 0.05. Moreover, among the Frank and Gumbel copulas, the Frank copula performs better than the Gumbel copula in modeling the dependence of flood peak and volume, leading to lower RMSE and AIC values. Consequently, the Frank copula will be chosen in this study to further characterize the uncertainty in model parameters and the resulting risks.

Table 5. Comparison of RMSE and AIC Values for Joint Distributions through Different Copulas

	RMSE	AIC	Sn	*p*-value
Frank	0.0293	-350.8780	0.0245	0.4211
Clayton	0.0393	-321.6217	0.0691	0.01149
Gumbel	0.0308	-346.0389	0.0223	0.454

4.3. Uncertainty Analysis

In this study, the Bayesian-based MCMC approach is employed to explore the inherent uncertainty both the marginal distributions for flood peak and volume and their dependence structure. In detail, the parallelizing MCMC algorithm is applied to quantify the posterior distributions for the parameters in the marginal and joint distributions, in which five Markov chains are generated by the Metropolis–Hastings algorithm. Total 10,000 iterations are implemented with the last 50% being considered as the posterior samples.

4.3.1. Uncertainty in marginal distributions

Figure 4 compares the distributions of parameters in the marginal distributions for flood peak and volume. In this study, the lognormal distribution would be employed to quantify the probabilistic characteristics in flood peak and volume. Consequently, for each flood variable, two pdfs would be obtained to indicate the location and shape parameter, respectively. Based on the Bayesian method, the sample values for each parameter are generated through their posterior distributions, and the associated pdf is then obtained through kernel smoothing method. In detail, for the location parameters of flood peak and volume, their distributions show similar features, which approximately obey normal distributions, as presented in Figures (4a) and (4c). However, the MCMC approach produces a slightly bimodal distribution for the location parameter of the flood volume. For shape parameters in the distributions of flood peak and volume (i.e., Figures (4b) and (4d)), the posterior distributions are approximately symmetric, following normal-like distributions.

Table 6 compares the optimal parameter values of the marginal distribution obtained by the maximum likelihood estimation (MLE) and Bayesian methods. Based on the sample values from the MCMC algorithm, the mean values and the associated predictive intervals (PIs) can be obtained. In this study, these mean values and PIs are provided based on the 5,000 posterior samples obtained by MCMC. The mean values can serve as the deterministic estimation and the PIs characterizes the inherent uncertainty in model estimation process. As presented in Table 6, the mean values from MCMC shows good agreement with the estimated values from MLE, suggesting the well performance of the proposed method in identifying the deterministic parameter estimations. Moreover, the 95% PIs can be obtained through the MCMC method, indicating the inherent uncertainty in the pdfs of flood peak and volume. For instance, for the parameter σ in the lognormal distribution of flood volume, the Bayesian approach produces a 95% PI of [0.4130, 0.6261]. Figure 5 shows the comparison between the observations and quantile estimations from MCMC. This figure indicates that the obtained confidence intervals can well bracket the observed probabilities.

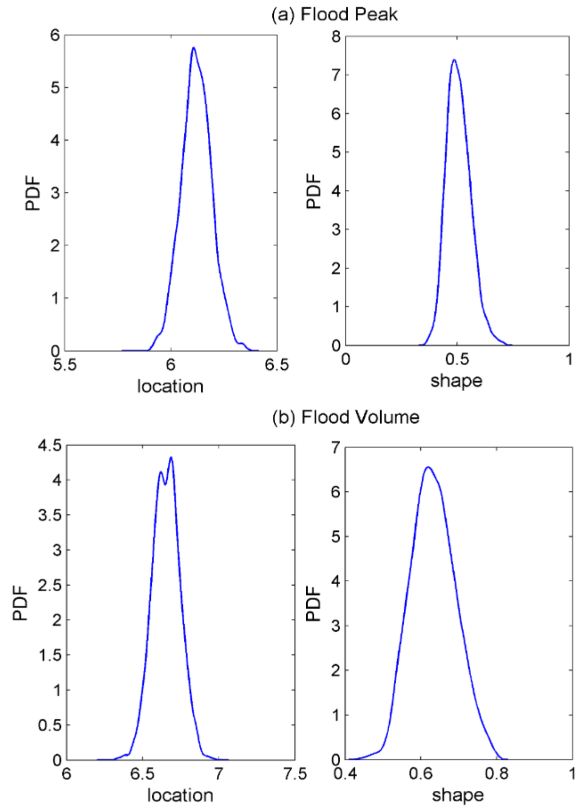

Figure 4. Parameter uncertainty for marginal distributions.

4.3.2. Uncertainty in joint distribution

The Frank copula is employed to model the dependence between flood peak and volume, due to its lowest RMSE and AIC values, as presented in Table 5. The pdf of its parameter (i.e., θ) is also approximated through the kernel method based on the samples from MCMC method. Figure 6 shows the obtained posterior pdf of the dependence parameter between flood peak and volume. It indicates that the posterior pdf from MCMC approximately follows a norm distribution. Table 7 shows the comparison of the estimated parameter values through different method. The predictive mean provided by MCMC is different with the optimal value from MLE. This means that the uncertainty in hydrologic series pose significant impact in the estimation of the dependence parameter. Moreover, the 95% PI for the dependence parameter is much wider when compared with the 95% PIs for the parameters in the marginal distributions. As presented in Table 7, the 95% PI from MCMC ranges within [3.90, 8.72], while the 95% PI for the location parameter of flood peak is [5.98, 6.26].

Table 6. Comparison of Parameter Estimation for Marginal Distributions through Different Methods.

Flood Variable	Distribution Family	Estimation Method	Estimation Type	Estimated Parameters	
				μ (mean)	σ (Standard Deviation)
Q	LN	MLE	Parameter estimation	6.1192	0.4952
		Bayesian	Mean	6.1209	0.5063
			95% PI	[5.9823, 6.2644]	[0.4130, 0.6261]
V	LN	MLE	Parameter estimation	6.6525	0.6272
		Bayesian	Mean	6.6540	0.6332
			95% PI	[6.4806, 6.8339]	[0.5280, 0.7544]

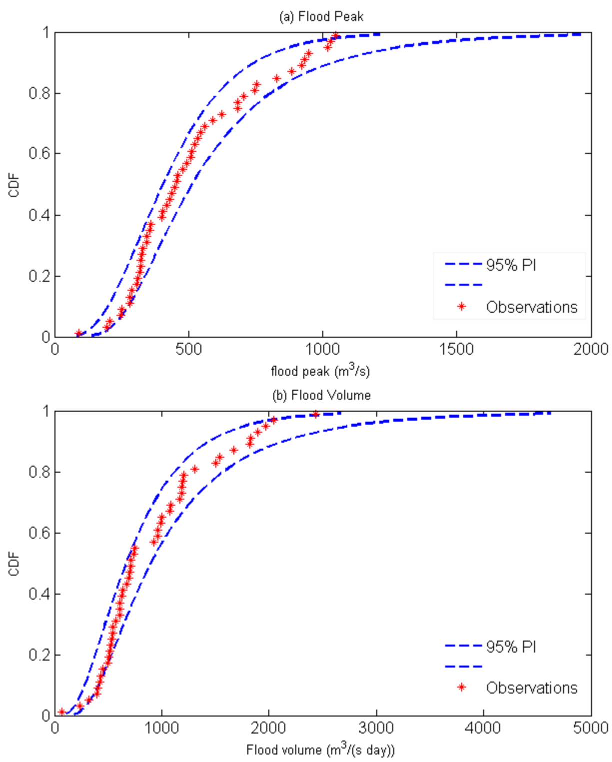

Figure 5. Quantile estimation of peak flow and volume

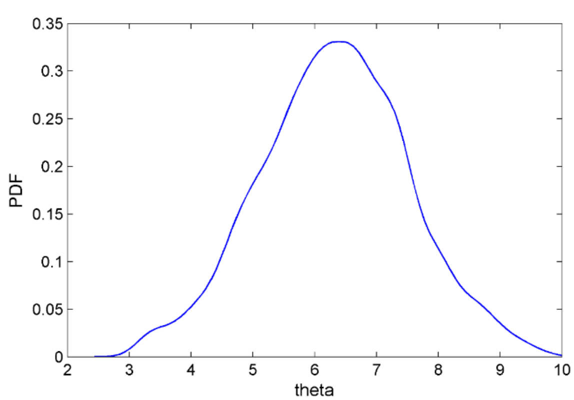

Figure 6. Parameter uncertainty of the dependence parameter.

To illustrate the impacts of uncertainty in model parameters, the bivariate quantile would be obtained compared. The method proposed by Zhang et al. (2015) would be employed to characterize the confidence regions (CRs) for several hazard levels. In this study, the CRs for the joint return period in "AND" is obtained under the return periods of 10, 50 and 100 years. Figure 7 shows confidence regions (CIs or CRs) obtained from MCMC. It shows quite large uncertainties even for the moderate flood with a 50-year return period, ranging between the value of the flood event with a return period of 10 years and that with a return period more than 100 years. Moreover, as presented in Figure 7, the increase in the return period for a flood event would lead to larger uncertainty. For instance, a flood event with a joint return period of 50 years would show a return period range of [21.5, 181.4] under parameter uncertainties, while a flood event with a 100-year joint return period presents a return period range of [36.3, 473.3] due to the parameter uncertainties.

4.4. Evaluation of parameter sensitivities

Due to the uncertainties in the marginal and joint distributions, the return period for one historical flood event would also be uncertain. However, how do the uncertainties in marginal distributions and the dependence structure contribute to the uncertainty of model output? To address the above issue, parameter sensitivities will be conducted. In detail, a multilevel factorial analysis is proposed to characterize the effect of parameter uncertainties and their interactions; a variance-based global sensitivity analysis is further employed to reveal the contributions of the uncertain parameters to the model outputs.

4.4.1. Factorial analysis

Figure 8 presents the main effects plot for four parameters in marginal distributions and one parameter in the dependence structure at three levels. The three levels are generated from the parameter distributions obtained by MCMC, in which the low, medium, and high levels are the 2.5% quantile value, mean and 97.5% quantile values. In the main effects plot, the evaluation index is the RMSE value at the various levels of each factor, with a reference line drawn at the grand mean of RMSE values. This plot reveals that the location parameters (i.e. P-mu and V-mu) for flood peak and volume has the greatest magnitude of the main effect upon RMSE value. In particular, the RMSE value would reach a lowest value of about 0.006 when the two location parameters get their medium values. This is because the medium values of the two location parameters are close to

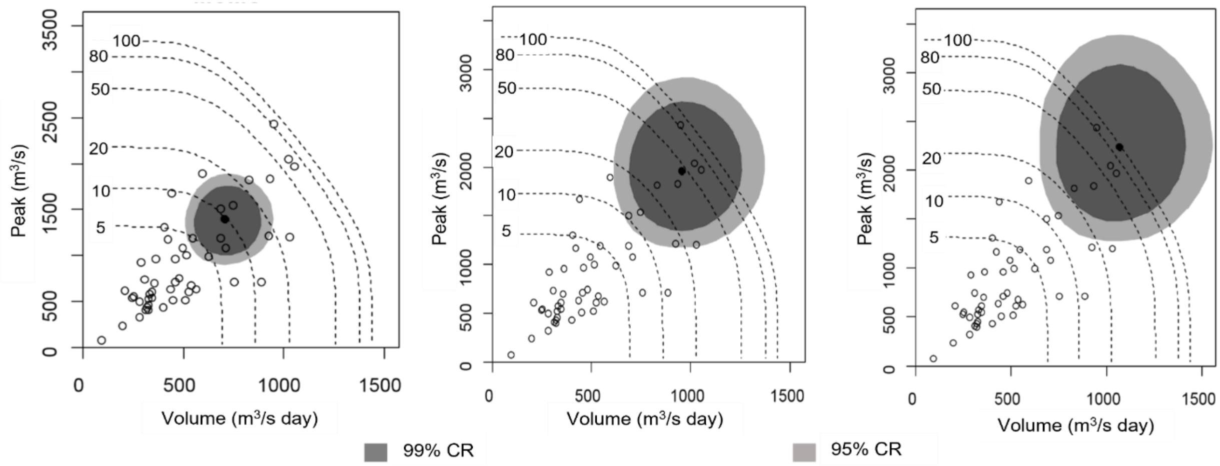

Figure 7. Comparison of bivariate quantile estimation.

their optimal estimation, leading to good performance of the proposed model. Contrarily, the two shape parameters (i.e., P-shape and V-shape) show less contributions to the variability of the RMSE value, and the parameter of the dependence structure shows lowest impact.

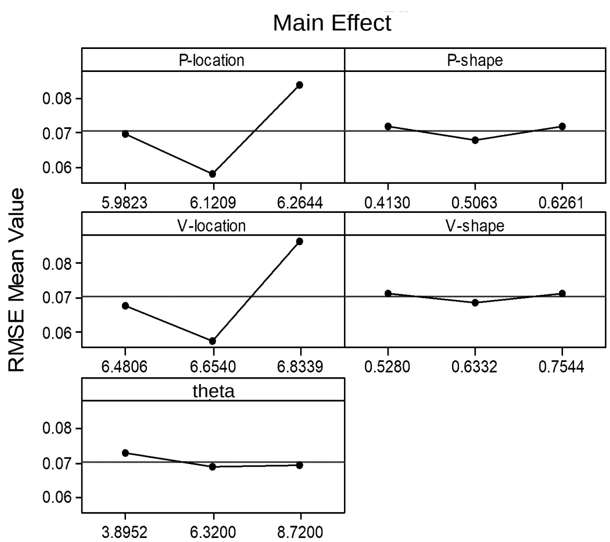

Figure 8. Main effects for the uncertainty parameters.

Table 7. Comparison of Parameter Estimation for the Copula through Different Methods.

Flood Variable	Distribution Family	Estimation Method	Estimation Type	θ
Q-V	Frank	MLE	Parameter estimation	6.9595
		Bayesian	Mean	6.32
			95% PI	[3.8952, 8.7200]

The interaction plot for the five factors at three levels is presented in Figure 9. It reveals that the change in the RMSE value differs across the three levels of one factor depending on the level of the other factors, implying that interactions between these factors occurs that their effects are dependent upon each other. Generally, the copula-based model get lowest RMSE values at the medium values of the five parameters. Specifically, the interactions between the shape parameters (i.e., P-shape and V-shape) and the dependence parameter (theta) seem to be insignificant since the three lines are nearly parallel.

4.4.2. Global Sensitivity Analysis

Figure 10 shows the first-order and total-order sensitivities for the five parameters obtained by the Sobol's sensitivity indices. In this study, the distributions of the five parameters are smoothed through the kernel distribution estimation method. Total 100,000 samples are drawn from the distributions of the parameters. The first-order and the total-order sensitivities are generated based on the selected samples through Equations (23) and (24). Figure (10) presents the contributions of the parameter uncertainties quantified by MCMC to performances of the copula model. It shows that the location parameter for flood volume contribute most to the performance variability of copula-based risk analysis models, followed by the location parameter of flood peak. Moreover, for the parameter of the dependence structure, the first-order sensitivity is not significantly, with a sensitivity of 2.29%. But its total-order sensitivity would reach 7.56%, pose the third contribution to the performance variability of copula models.

As presented in Table 6, the predictive intervals (PIs) for the location parameters of the marginal distributions are quite narrow. The 95% PIs of the location parameters are [5.98, 6.26] and [6.48, 6.83], respectively, showing uncertainty degrees of 4% and 5%. But the uncertainty in the shape parameters is more significant, with uncertainty degrees of 41% for flood peak and 35% for flood volume. Moreover, the dependence structure shows most significant uncertainty. The 95% PI for the dependence parameter ranges from 3.90 to 8.72, with an uncertainty degree of 76, as presented in Table 7. However, in terms of parameter sensitivities, slight variations in the location parameters will

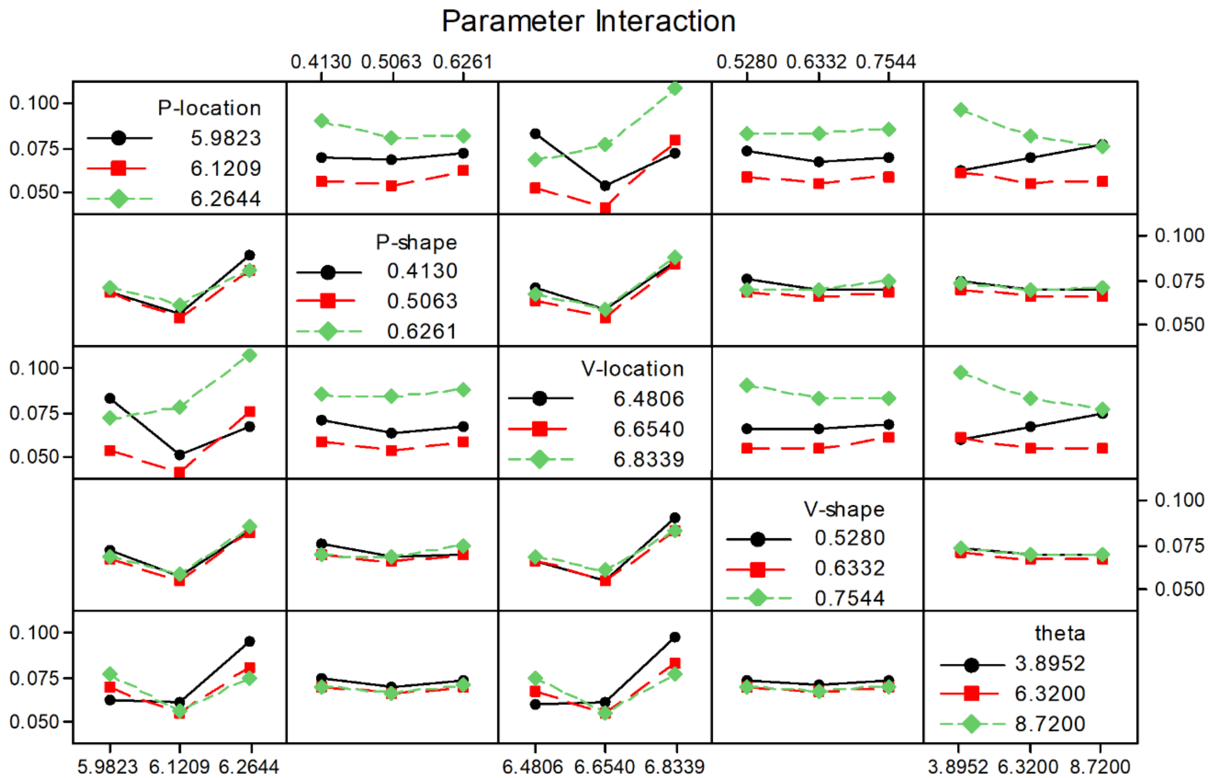

Figure 9. Interaction effects for the uncertainty parameters.

Table 8. Comparison for the Univariate and Joint Return Periods for Flood Characteristics.

	T	5	10	20	50	100
	MLE*	689.5	857.4	1026.4	1256.7	1438.4
P	PM**	699.8	875.6	1054.2	1299.7	1494.8
	95% PI***	[595.9, 840.6]	[729.6, 1087.7]	[857.6, 1348.7]	[1020.3, 1725.1]	[1144.6, 2040.1]
	MLE	1313.5	1730.9	2173.9	2809.5	3333.4
V	PM	1329.2	1759.2	2218.5	2881.8	3432.0
	95% PI	[1088.7, 1632.8]	[1403.8, 2223.2]	[1719.4, 2889.5]	[2161.5, 3898.7]	[2511.7, 4751.2]
	MLE	3.4	6.1	11.3	26.4	51.5
Tor	PM	3.5	6.2	11.4	26.5	51.5
	95% PI	[3.2, 3.7]	[5.8, 6.6]	[10.9 11.8]	[26.0, 27.0]	[51.0, 52.1]
	MLE	8.4	24.0	77.0	408.3	1534.9
Tand	PM	8.9	26.0	85.1	459.0	1737.8
	95% PI	[7.6, 11.0]	[21.1, 34.7]	[65.4, 119.9]	[336.0, 677.4]	[1245.9, 2612.7]
	MLE	6.4	16.0	45.9	221.7	801.7
$\overline{T}$	PM	6.6	17.0	49.9	246.9	903.0
	95% PI	[6.0, 7.5]	[14.7, 21.1]	[40.3, 67.0]	[185.8, 355.7]	[657.4, 1339.9]

Note: *: Results obtained through maximum likelihood estimation; **: the predictive mean; ***: 95% predictive interval

pose significant effects on the performance of the obtained copula, but the extensive uncertainty in the shape parameters of the marginal distribution and the dependence parameter will not show explicit contributions to performance variations from the copula. Therefore, in practical multivariate hydrologic risk analysis, the location parameters in the flood variables play the most vital role, followed by the dependence parameter and the shape parameters in order to improve the performance on modelling the dependence between flood peaks and volumes.

4.5. Predictive risk

4.5.1. Return Period Characterization

The concurrence probabilities of various combinations of flood variables can be revealed through the obtained copulas. As presented in Equations (13) ~ (16), the joint return period and second return period can be derived based on the selected copula functions. However, due to the uncertainties in the location and shape parameters of the univariate flood variable distributions (e.g., flood peak and volume) and the dependence structure, the primary and secondary joint return periods will also exhibit uncertainty. Table 8 shows the deterministic and predictive values for the univariate and joint return periods for flood characteristics. For the deterministic return periods, they are generated through the copulas with parameters estimated by MLE and the predictive mean based on Equation (27). The results suggest that for the univariate return periods, the MLE approach provides underestimations when compared with the predictive means. For instance, the deterministic approach quantifies a peak value of 1438.5 m^3/s for a flood with the 100-year return period, while under consideration of uncertainty, the predictive mean for such a flood peak would be 1494.8 m^3/s. Similarly, the primary and secondary return periods obtained from MLE are also underestimated when compared with their predictive means. Moreover, the 95% predictive intervals can bracket deterministic predictions from MLE, indicating the effectiveness of the MCMC approach in quantifying the uncertainty in the copula model. Specifically, as presented in Table 8, the width of the 95% PIs for the joint return period in "OR" case do not change significantly. On the contrary, the 95% PIs for the return period in "AND" case show more uncertainty as the increase of the return period, which is consistent with the results in Figure 7. Similar features can also be characterized for the secondary return period.

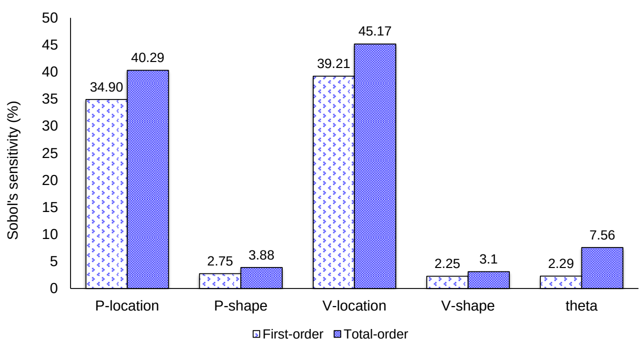

Figure 10. First-order, total-order sensitivities of the five parameters.

4.5.2. Bivariate risk analysis

The damages caused by a flood, such as the failure of hydraulic structures, mainly due to the high peak flow of the flood. The annual maximum peak discharge is the central issue to be considered for hydrologic risk analysis. Moreover, the flood discharge volume may also be under consideration in practical flood control and mitigation. The flood volume is related to flood diversion practices. Consequently, multivariate flood risk analysis, which involves more flood variables than just considering flood peak, is more helpful for actual flood control. The bivariate hydrologic risk defined by Equation (18) is able to reflect the interactive effects of flood variables on the occurrence probability of a flood. In this study, the designated peak flow is 1500 m^3/s and two service time scenarios, i.e., 30 and 50 years are under consideration.

Figure 11 provides an illustration of the predictive risk and compares it to more standard predictions using an "optimal" predictor through MLE. The results indicate that both estimates are similar since the uncertainties in the posterior probabilities of the model parameters are relatively small. Integrating them do not dramatically results into remarkable deviation to the deterministic predictions. However, as shown in Figure 11, the "optimal" predictions through MLE provides underestimation than the predictive risk under consideration of parameter uncertainties. This may result from the underestimation for the univariate return period from the "optimal" parameter values. In this study, a flood peak value of 1500 m^3/s is considered as the design standard for the river levee around the Xingshan station. Such a flood peak will have a return period more than 125 years under "optimal" prediction while the predictive mean return period is only about 90 years from the Bayesian inference. Such an overestimation for the return period of the designated flood peak leads to an underestimation of the bivariate hydrologic risk.

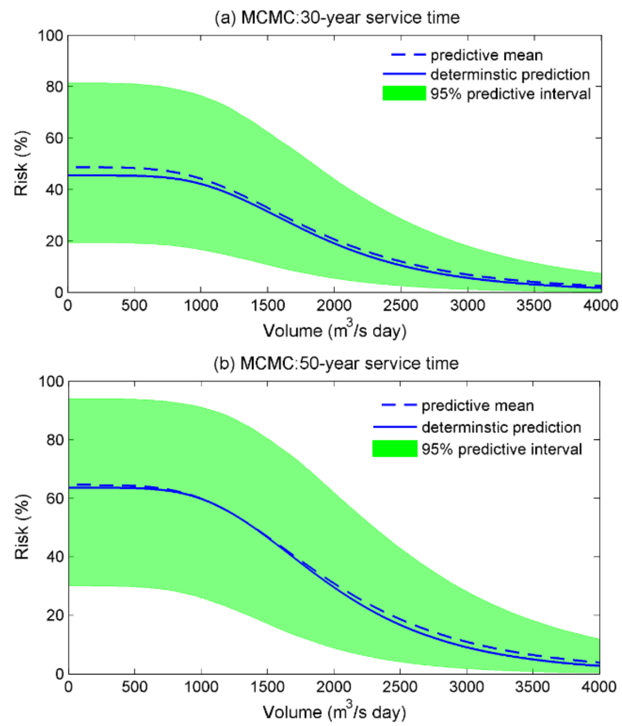

Figure 11. The predictive bivariate risk under uncertainty.

The quantile curve is a standard representation used in extreme values analyses (Renard et al., 2013). Figure 11 also shows 95% predictive interval for the bivariate hydrologic risk bracketed by the 2.5% and 97.5% quantile estimations resulting from Bayesian inference. In this figure, the quantile curves are generated based on the posterior distributions of the model parameters quantified by MCMC. The results show that the bivariate hydrologic risk will decrease as the increase of the flood volume, indicating the low concurrence probabilities of high peak flow and high flood volume. Moreover, the width of the predictive intervals will become narrow as the decrease of the bivariate risk values. However, uncertainties exist in the predictive bivariate hydrologic risk resulting from parameter uncertainties. Particularly, the bivariate flood risk increases as the increase of the service time of the hydraulic infrastructures, and the uncertainties in predictive risk also increase at the same time.

5. Conclusions

Extensive uncertainties exist in multivariate hydroclimatic risk analysis resulting from a variety of sources such as limited sample sizes, different model selection and parameter estimation techniques. In this study, parameter uncertainty and sensitivity evaluation (PUSE) framework was proposed to quantify uncertainty in multivariate hydroclimatic risk analysis, characterize impacts of parameter uncertainties on model output variations and finally provide predictive risk predictions. Uncertainties in multivariate hydroclimatic risk analysis framework was quantified through coupling Bayesian inference into the copula framework. The posterior distributions for the parameters in marginal distributions and dependence structure were quantified through a Markov chain Monte Carlo (MCMC) technique. The main effects of model parameters and their interactions on model performance were revealed through a multi-level factorial analysis approach. The contributions of parameter uncertainties on model output variations were characterized through Sobol's based global sensitivity analysis method. The predictive risks and the associated quantile estimations were generated based on the posterior distributions of the parameters.

The proposed PUSE approach are applied to the bivariate flood risk analysis in the Xiangxi River, in which the flood peak and volume are under consideration. The results show that uncertainties in multivariate hydroclimatic risk analysis can be well quantified through Bayesian-based Markov Chain Monte Carlo approach. But parameter uncertainties lead to quite large uncertainties even for the moderate flood with a 50-year return period. Moreover, the multi-level factorial analysis results show that the location parameters of the flood variables have the greatest magnitude of the main effect, but the interactions between the location parameters and the dependence parameter are insignificant. Particularly, the two location parameters contribute most to the modelling performances, followed by the dependence and shape parameters, as revealed by the global sensitivity analysis. Under uncertainty, the predictive bivariate risks show similar features with the "optimal" predictions. But the "optimal" predictions present underestimations when compared with the predictive risks through Bayesian inference. Therefore, in practical hydroclimatic risk analysis, the inherent uncertainty needs to be well quantified especially for the uncertainty in the location parameters of the flood variables. Well uncertainty quantification may provide more reliable hydroclimatic risk predictions.

Acknowledgement. This research was supported by the National Key Research and Development Plan (2016YFA0601502), and the Royal Society International Exchanges Program (No. IES\R2\202075).

References

Adamowski, K. (1989). A Monte Carlo comparison of parametric and nonparametric estimation of flood frequencies. *J. Hydrol*, 108, 295-308. https://doi.org/10.1016/0022-1694(89)90290-4

Dong, C., Huang, G.H, and Cheng, G.H. (2021). Offshore wind can power Canada. *Energy*, 121422. https://doi.org/10.1016/j.energy.2021.121422

De Michele, C. and Salvadori, G. (2003). A Generalized Pareto intensity-duration model of storm rainfall exploiting 2-copulas. *J. Geophys. Res. Atmos*, 108(D2), 4067. https://doi:10.1029/2002JD002534

Dung, N.V., Merz, B., Bardossy, A. and Apel, H. (2015). Handling uncertainty in bivariate quantile estimation – An application to flood hazard analysis in the Mekong Delta. *J. Hydrol*, 527, 704-717. https://doi.org/10.1016/j.jhydrol.2015.05.033

Fan, Y.R., Huang, W., Huang, G.H., Li, Z., Li, Y.P., Wang, X.Q., Cheng, G.H. and Jin, L. (2015). A stepwise-cluster forecasting approach for monthly streamflows based on climate teleconnections. *Stoch. Environ. Res. Risk Assess*, 29(6), 1557-1569. https://doi.org/10.1007/s00477-015-1048-y

Fan, Y.R., Huang, G.H., Baetz, B.W., Li, Y.P., Huang, K., Li, Z., Chen, X. and Xiong, L.H. (2016a). Parameter uncertainty and temporal dynamics of sensitivity for hydrologic models: A hybrid sequential data assimilation and probabilistic collocation method. *Environ. Modell. Softw*, 86, 30-49. https://doi.org/10.1016/j.envsoft.2016.09.012

Fan, Y.R., Huang, G.H., Li, Y.P., Wang, X.Q. and Li, Z. (2016b). Probabilistic prediction for monthly streamflow through coupling stepwise cluster analysis and quantile regression methods. *Water Resour. Manag*, 30(14), 5313-5331. https://doi.org/10.1007/s11269-016-1489-1

Fan, Y.R., Huang, G.H., Zhang, Y. and Li, Y.P. (2018). Uncertainty quantification for multivariate eco-hydrological risk in the Xiangxi River within the Three Gorges Reservoir Area in China. *Engineering*, 4(5), 617-626. https://doi.org/10.1016/j.eng.2018.06.006

Fan, Y.R., Huang, G.H., Li, Y.P., Baetz, B.W. and Huang, K. (2020a). Uncertainty characterization and partition in multivariate risk inference: a factorial bayesian copula framework. *Environ. Res*, 183, 109215. https://doi.org/10.1016/j.envres.2020.109215

Fan, Y.R., Huang, G.H., Huang, K. and Wang, F. (2020b). An uncertainty partition approach for inferring interactive hydrologic risks. *Hydrol. Earth Syst. Sci*, 24, 4601-4624. https://doi.org/10.5194/hess-2019-434

Fan, Y.R., Yu, L. and Shi, X. (2021). Uncertainty quantification and partition for multivariate risk inferences through a Factorial Multimodel Bayesian Copula (FMBC) system. *J. Hydrol*, 598, 126406. https://doi.org/10.1016/j.jhydrol.2021.126406

Ganguli, P. and Reddy, M.J. (2013). Probabilistic assessment of flood risks using trivariate copulas. *Theor. Appl. Climatol*, 111(1-2), 341-360. https://doi.org/10.1007/s00704-012-0664-4

Gebregiorgis, A.S. and Hossain, F. (2012). Hydrological risk assessment of old dams: case study on Wilson dam of Tennessee River basin. *J. Hydrol. Eng*, 17, 201-212. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000410

Genest, C. and Favre, A.C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *J. Hydrol. Eng*, 12(4), 347-368. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(347)

Genest, C., Rémillard, B. and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insur. Math. Econ*, 44, 199-213. https://doi.org/10.1016/j.insmatheco.2007.10.005

Graler, B., Van Den Berg, M.J., Vandenberghe, S., Petroselli, A., Grimaldi, S., Baets, B.D. and Verhoest, N.E.C. (2013). Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation. *Hydrol. Earth Syst. Sci*, 17, 1281-1296. https://doi.org/10.5194/hess-17-1281-2013, 2013

Han, J.C., Huang, G.H., Zhang, H., Li, Z. and Li, Y.P. (2014). Bayesian uncertainty analysis in hydrological modeling associated with watershed subdivision level: a case study of SLURP model applied to the Xiangxi River watershed, China. *Stoch. Environ. Res. Risk Assess*, 28(4), 973-989. https://doi.org/10.1007/s00477-013-0792-0

Huang, K., Dai, L.M., Yao, M., Fan, Y.R. and Kong, X.M. (2017). Modelling dependence between traffic noise and traffic flow through an entropy-copula method. *J. Environ. Inform*, 29(2), 134-151. https://doi:10.3808/jei.201500302

Kao, S.C. and Govindaraju, R.S. (2010). A copula-based joint deficit index for droughts. *J. Hydrol*, 380, 121-134. https://doi.org/10.1016/j.jhydrol.2009.10.029

Karmakar, S. and Simonovic, S.P. (2009). Bivariate flood frequency analysis. Part 2: a copula-based approach with mixed marginal distributions. *J. Flood Risk Manag*, 2, 32-44. https://doi.org/10.1111/j.1753-318X.2009.01020.x

Kidson, R. and Richards, K.S. (2005). Flood frequency analysis: assumption and alternatives. *Prog. Phys. Geogr*, 29(3), 392-410. https://doi.org/10.1191/0309133305pp454ra

Li, Y.F., Li, Y.P., Huang, G.H. and Chen, X. (2010). Energy and environmental systems planning under uncertainty—an inexact fuzzy-stochastic programming approach. *Appl. Energy*, 87(10), 3189-3211. https://doi.org/10.1016/j.apenergy.2010.02.030

Li, Y.P., Huang, G.H., Yang, Z.F. and Nie, S.L. (2008). IFMP: Interval-fuzzy multistage programming for water resources management under uncertainty. *Resour. Conserv. Recy.*, 52(5), 800-812. https://doi.org/10.1016/j.resconrec.2007.11.007

Li, Z., Huang, G.H., Fan, Y.R. and Xu, J.L. (2015). Hydrologic risk analysis for nonstationary streamflow records under uncertainty. *J. Environ. Inform*, 26(1), 41-51. doi:10.3808/jei.201500321

Merz, B. and Thieken, A.H. (2005). Separating natural and epistemic uncertainty in flood frequency analysis. *J. Hydrol*, 309(1-4), 114-132. https://doi.org/10.1016/j.jhydrol.2004.11.015

Montgomery, D. C. (2001). *Design and analysis of experiments* (5th ed.). New York: John Wiley & Sons Inc.

Nelsen, R.B. (2006). *An Introduction to Copulas*. Springer: New York. ISBN: 0-387-28659-4

Renard, B., Sun, X. and Lang, M. (2013). *Bayesian Methods for Non-stationary Extreme Value Analysis.* In: AghaKouchak A., Easterling D., Hsu K., Schubert S., Sorooshian S. (eds) *Extremes in a Changing Climate*. Water Science and Technology Library, vol 65. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-4479-0_3

Salvadori, G., Michele, C.D. and Durante, F. (2011). On the return period and design in a multivariate framework. *Hydrol. Earth Syst. Sci*, 15, 3293-3305. https://doi.org/10.5194/hess-15-3293-2011, 2011

Salvadori, G., Michele, C.D., Kottegoda, N.T. and Rosso, R. (2007). *Extremes in Nature: an Approach using Copula*. Springer: Dordrencht; 292. ISBN: 978-1-4020-4414-4

Serinaldi, F. (2013). An uncertain journey around the tails of multivariate hydrological distributions. *Water Resour. Res*, 49, 6527-6547. https://doi.org/10.1002/wrcr.20531

Song, S. and Singh, V.P. (2010). Meta-elliptical copulas for drought frequency analysis of periodic hydrologic data. *Stoch. Environ. Res. Risk Assess*, 24(3), 425-444. https://doi.org/10.1007/s00477-009-0331-1

Song, X., Zhang, J., Zhang, C., Xuan, Y., Ye, M. and Xu, C. (2015). Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications. *J. Hydrol*, 523, 739-757. https://doi.org/10.1016/j.jhydrol.2015.02.013

Sraj, M., Bezak, N. and Brilly, M. (2015). Bivariate flood frequency analysis using the copula function: a case study of the Litija station on the Sava River. *Hydrol. Process*, 29(2), 225-238. https://doi.org/10.1002/hyp.10145

Sun, C.X., Huang, G.H., Fan, Y.R., Zhou, X., Lu, C. and Wang, X.Q. (2021). Vine copula based jointly downscaling approach for precipitation of multi-RCMs: A case study in the Loess Plateau, China. *Water Resour. Res*, 57(1), 2020WR027698. https://doi.org/10.1029/2020WR027698

Sun, C.X., Huang, G.H., Fan, Y.R., Zhou, X., Lu, C. and Wang, X.Q. (2019). Drought occurring with hot extremes: changes under future climate change on Loess Plateau, China. *Earths Future*, 7, 587-604. https://doi.org/10.1029/2018EF001103

Vandenberghe, S., Verhoest, N.E.C. and De Baets, B. (2010). Fitting bivariate copulas to the dependence structure between storm characteristics: a detailed analysis based on 105 year 10 min rainfall. *Water Resour. Res*, 46(1). https://doi.org/10.1029/2009WR007857

Wang, S. and Huang, G.H. (2015). A multi-level Taguchi-factorial two-stage stochastic programming approach for characterization of parameter uncertainties and their interactions: An application to water resources management. *Eur. J. Oper. Res*, 240, 572-581. https://doi.org/10.1016/j.ejor.2014.07.011

Wu, Z.Y., Lu, G.H., Liu, Z.Y., Wang, J.X. and Xiao, H. (2013). Trends of extreme flood events in the Pearl River basin during 1951-2010. *Adv. Clim. Chang. Res*, 4(2), 110-116. https://doi.org/10.3724/SP.J.1248.2013.110

Xu, H., Taylor, R.G., Kinston, D.G., Jiang, T., Thompson, J.R. and Todd, M.C. (2010). Hydrological modeling of River Xiangxi using SWAT2005: A comparison of model parameterizations using station and gridded meteorological observations. *Quat. Int*, 226, 54-59. https://doi.org/10.1016/j.quaint.2009.11.037

Xu, Y., Huang, G.H. and Fan, Y.R. (2017). Multivariate flood risk analysis for Wei River. *Stoch. Environ. Res. Risk Assess*, 31 (1), 225-242. http://dx.doi.org/10.1007/s00477-015-1196-0

Yen, B.C. (1970). Risk analysis in design of engineering projects. *Journal of Hydrologic Engineering*, 96(4), 959-966. https://doi.org/10.1061/JYCEAJ.0002476

Yue, S. (2000). The bivariate lognormal distribution to model a multivariate flood episode. *Hydrol. Process*, 14(14), 2575-2588. https://doi.org/10.1002/1099-1085(20001015)14:14<2575::AID-HYP115>3.0.CO;2-L

Yue, S. (2001). A bivariate gamma distribution for use in multivariate flood frequency analysis. *Hydrol. Process*, 15(6), 1033-1045. https://doi.org/10.1002/hyp.259

Zhang, C., Chu, J. and Fu, G. (2013). Sobol's sensitivity analysis for a distributed hydrological model of Yichun River Basin, China. *J. Hydrol*, 480, 58-68. https://doi.org/10.1016/j.jhydrol.2012.12.005

Zhang, L. and Singh, V.P. (2006). Bivariate flood frequency analysis using the copula method. *J. Hydrol. Eng*, 11(2), 150-164. https://doi.org/10.1061/(ASCE)1084-0699(2006)11:2(150)

Zhang, L. and Singh, V.P. (2007). Bivariate rainfall frequency distributions using Archimedean copulas. *J. Hydrol*, 332(1-2), 93-109. https://doi.org/10.1016/j.jhydrol.2006.06.033

Zhang, Q., Xiao, M.Z. and Singh, V.P. (2015). Uncertainty evaluation of copula analysis of hydrological droughts in the East River basin, China. *Glob. Planet. Change*, 129, 1-9. https://doi.org/10.1016/j.gloplacha.2015.03.001

Zhai, M.Y., Huang, G.H., Liu, L.R., Zheng, B.Y. and Guan, Y.R. (2020). Inter-regional carbon flows embodied in electricity transmission: network simulation for energy-carbon nexus. *Renew. Sustain. Energy Rev.*, 118, 109511. https://doi.org/10.1016/j.rser.2019.109511

Zhou, X., Huang, G.H., Wang, X.Q., Fan, Y.R. and Cheng, G.H. (2018). A coupled dynamical-copula downscaling approach for temperature projections over the Canadian Prairies, *Climate Dynamics*, 51, 2413-2431. https://doi.org/10.1007/s00382-017-4020-3