An Uncertainty Partition Approach for Inferring Interactive Hydrologic Risks

Yurui Fan¹, Kai Huang², Guohe Huang³, Yongping Li⁴, Feng Wang⁴

¹ Department of Civil and Environmental Engineering, Brunel University, London, Uxbridge, Middlesex, UB8 3PH, United Kingdom
² Faculty of Engineering and Applied Sciences, University of Regina, Regina, SK, Canada, S4S0A2
³ Institute for Energy, Environment and Sustainable Communities, University of Regina, Regina, Saskatchewan, Canada S4S 0A2
⁴ School of Environment, Beijing Normal University, Beijing 100875, China

Correspondence to: Yurui Fan (yurui.fan@brunel.ac.uk), Guohe Huang (huangg@uregina.ca)

Abstract:

Extensive uncertainties exist in hydrologic risk analysis. Particularly for interdependent hydrometeorological extremes, the random features in individual variables and their dependence structures may lead to bias and uncertainty in future risk inferences. In this study, a full-subsampling factorial copula (FSFC) approach is proposed to quantify parameter uncertainties and further reveal their contributions to predictive uncertainties in risk inferences. Specifically, a full-subsampling factorial analysis (FSFA) approach is developed to diminish the effect of the sample size and provide reliable characterization for parameters’ contributions to the resulting risk inferences. The proposed approach is applied to multivariate flood risk inference for Wei River basin to demonstrate the applicability of FSFC for tracking the major contributors to resulting uncertainty in a multivariate risk analysis framework. In detail, the multivariate risk model associated with flood peak and volume will be established and further introduced into the proposed full-subsampling factorial analysis framework to reveal the individual and interactive effects of parameter uncertainties on the predictive uncertainties in the resulting risk inferences. The results suggest that uncertainties in risk inferences would mainly be attributed to some parameters of the marginal distributions while the parameter of dependence structure...
(i.e. copula function) would not produce noticeable effects. Moreover, compared with traditional factorial analysis (FA), the proposed FSFA approach would produce more reliable visualization for parameters’ impacts on risk inferences, while the traditional FA would remarkably underestimate contribution of parameters’ interaction to the failure probability in AND, and at the same time, underestimate the contribution of parameters’ interaction to the failure probabilities in OR and Kendall.

1. Introduction

Many hydrological and climatological extremes are highly correlated among each other, and it is desired to explore their interdependence through multivariate approaches. Examples include sea level rise and fluvial flood (Moftakhari et al., 2017), drought and heat waves (Sun et al., 2019), soil moisture and precipitation (AghaKouchak, 2015). Moreover, even one specific hydrological extreme may have multiple attributes, such as the peak and volume for a flood, duration and severity for a drought, and duration and intensity of a storm (Karmakar and Simonovic, 2009; Kong et al., 2019). Traditional univariate approaches, mainly focusing on one variable or one attribute of hydrological extremes (e.g. flood peak), may not be sufficient to describe those hydrological extremes containing multivariate characteristics. Thus the univariate frequency/risk analysis methods may be unable to obtain reliable risk inferences for the failure probability or recurrence intervals of interdependent extreme events (Chebana and Ouarda, 2011; Requena et al., 2013; Salvadori et al., 2016; Sadegh et al., 2017).

Since the introduction of copula function into hydrology and geosciences by De Michele and Salvador (2003), the copula-based approaches have been widely used for multivariate hydrologic risk analysis. The copula functions are able to model correlated variables with complex or nonlinear dependence structures. Also, this kind of methods are easily to be implemented since the marginal distributions and
dependence model can be estimated in separate processes, which also give flexibility in selection of both marginal and dependence models. A great number of research has been developed for multivariate hydrologic simulation through copula functions, such as multivariate flood frequency analysis (Sraj et al., 2014; Xu et al., 2016; Fan et al., 2018); drought assessments (Song and Singh 2010; Kao and Govindaraju 2010; Ma et al. 2013); storm or rainfall dependence analysis (Zhang and Singh 2007; Vandenberghhe et al. 2010); streamflow simulation (Lee and Salas 2011; Kong et al., 2015) and other water and environmental engineering applications (Fan et al., 2017; Huang et al., 2017).

For both univariate and multivariate analysis for hydrometeorological risks, uncertainty would be one of the unavoidable issues which needs to be well addressed. The uncertainty in hydrometeorological risk inference mainly results from stochastic variability of hydrometeorological processes and incomplete knowledge of the watershed systems (Merz and Thieken, 2005). Amounts of studies have been proposed to address uncertainty in both univariate and multivariate hydrological risk analysis (e.g. Merz and Thieken, 2005; Serinaldi, 2013; Dung et al., 2015; Zhang et al., 2015; Sadegh et al., 2017; Fan et al., 2018). However, one critical issue in uncertainty quantification of hydrological inference is how to characterize the major sources for uncertain risk inference. Qi et al. (2016) employed a subsampling ANOVA approach (Bosshard et al., 2013), to quantify individual and interactive impacts of the uncertainties in data, probability distribution functions, and probability distribution parameters on the total cost for flood control in terms of flood peak flows. Even though the subsampling ANOVA approach is able to reduce the effect of the biased estimator on quantification of variance contribution resulting from traditional ANOVA approach, it would be noticed that merely subsampling one uncertainty parameter/factor (referred as single-subsampling ANOVA), as used in the studies by Bosshard et al. (2013) and Qi et al. (2016a), will lead to underestimation of individual contribution for the factor to be sampled and overestimation of contributions for those non-sampled factors. Moreover, few studies have been reported to characterize the individual and interactive effects of parameter
uncertainties in marginal and dependence models on the multivariate risk inferences.

Consequently, as an extension of previous research, this study aims to propose a Full-subsampling factorial copula (FSFC) approach for uncertainty quantification and partition in multivariate hydrologic risk inference. In detail, the parameter uncertainties are quantified through a Monte Carlo-based Bootstrap algorithm. The interactions of parameter uncertainties are explored through a multilevel factorial analysis approach. The contributions of parameter uncertainties are analyzed through a full-subsampling factorial analysis (FSFA) method, in which all uncertain factors will be subsampled to generate more reliable results. The applicability of the proposed FSFC approach will be demonstrated through case studies of flood risk analysis in the Wei River basin in China.

2. Methodology

Figure 1 illustrates the framework of the proposed FSFC approach. The framework consists of four modules: (i) selection of marginal distributions, (ii) identification of copulas, (iii) parameter uncertainty quantification, (iv) parameter interaction and sensitivity analysis. In FSFC, modules (i) and (ii) are proposed to construct the most appropriate copula-based hydrologic risk model. Module (iii) quantifies parameter uncertainties in marginal distributions and copulas. Modules (iv) would be the core part of our study to identify the main sources of uncertainties in multivariate risk inference by the proposed full-subsampling factorial analysis (FSFA) approach.
2.2. Copula-based Multivariate Risk Inference Framework

A copula function is a multivariate distribution function with uniform margins on the interval [0, 1]. Sklar’s Theorem states that any \(d\)-dimensional distribution function \(F\) can be formulated through a copula and its marginal distributions (Nelsen, 2006). In detail, a multivariate copula function can be expressed as:

\[
F(x_1, x_2, ..., x_d \mid \gamma_1, \gamma_2, ..., \gamma_d, \theta) = C(F_{X_1}(x_1 \mid \gamma_1), F_{X_2}(x_2 \mid \gamma_2), ..., F_{X_d}(x_d \mid \gamma_d) \mid \theta)
\]

(1)

where \(F_{X_1}(x_1 \mid \gamma_1), F_{X_2}(x_2 \mid \gamma_2), ..., F_{X_d}(x_d \mid \gamma_d)\) are marginal distributions of the random vector \((X_1, X_2, ..., X_d)\), with \(\gamma_1, \gamma_2, ..., \gamma_d\) respectively being the unknow parameters of marginal distributions. \(\theta\) is the parameter in the copula function describing dependence among the correlated variables. If these marginal distributions are continuous, then a single copula function \(C\) exists, which can be written as (Nelsen, 2006):

\[
C(u_1, u_2, ..., u_d \mid \theta) = F(F_{X_1}^{-1}(u_1 \mid \gamma_1), F_{X_2}^{-1}(u_2 \mid \gamma_2), ..., F_{X_d}^{-1}(u_d \mid \gamma_d))
\]

(2)

where \(u_1 = F_{X_1}(x_1 \mid \gamma_1), u_2 = F_{X_2}(x_2 \mid \gamma_2), ..., u_d = F_{X_d}(x_d \mid \gamma_d)\). More details on the theoretical background and properties of various copula families can be found in Nelsen (2006).

If appropriate copula functions are specified to reflect the joint probabilistic characteristics among for a multivariate extreme event, the conditional, primary and secondary return periods (RP) can be obtained. Consider one kind of hydrological extreme (denoted as \(X\)) with \(d\) attributes (i.e. \(X = (x_1, x_2, ..., x_d)\)), and for a specific extreme event \(X^*\) with its attributes being \(X^* = (x_1^*, x_2^*, ..., x_d^*)\), three categories of multivariate RP can be applied for revealing the potential risk of \(X^*\).

(i) “OR” case:
where \( \mu \) denotes the average time between two adjacent events under consideration.

(ii) “AND” case:

\[
T^{\text{AND}} = \{(x_1, x_2, ..., x_d) \in \mathbb{R}^d : x_1 > x_1^* \land x_2 > x_2^* \land ... \land x_d > x_d^* \}
\]

\[
= \frac{\mu}{C(F_1(x_1 | \gamma_1), ..., F_d(x_d | \gamma_d) | \theta)}
\]

where \( \hat{C} \) is the multivariate survival function of the \( X_i \)'s proposed by Salvadori et al. (2013; 2016), and \( \bar{F}_i(x_i | \gamma_i) = P(X > x_i) = 1 - F_i(x_i | \gamma_i) \). Following Salvadori et al. (2013; 2016), and the Inclusion-Exclusion principle proposed by Joe (2014), the multivariate survival function \( \hat{C} \) can be obtained by:

\[
\hat{C}(\mathbf{u}) = \bar{C}(1 - \mathbf{u})
\]

and

\[
\bar{C}(\mathbf{u}) = 1 - \sum_{i=1}^{d} u_i + \sum_{S \in \mathcal{P}} (-1)^{|S|} C_S(u_i : i \in S)
\]

(iii) “Kendall” case: The Kendall RP characterizes the hydrologic disasters exceeding a critical layer as defined by (Salvadori et al., 2011): \( L_t^F = \{x \in \mathbb{R}^d : F(x) = t\} \). The Kendall RP can be expressed as

\[
T^{\text{Kendall}} = \frac{\mu}{1 - K_C(t)}
\]

where \( K_C \) is the Kendall distribution function associated with \( C \), which can be expressed as:

\[
K_C(t) = \int_0^t \left( \sum_{i=1}^{d} u_i + \sum_{S \in \mathcal{P}} (-1)^{|S|} C_S(u_i : i \in S) \right) du
\]
In addition to the multivariate RP, Failure probability (FP) can be another index to provide more coherent, general and well devised tools for multivariate risk assessment and communication. In general, the failure probability \( p_M \) to indicate the occurrence of a critical event for at least one time in \( M \) years of design life can be defined as (Salvadori et al., 2016):

\[
p_M = 1 - \prod_{j=1}^{M} (1 - p_j) = 1 - (F(x_j))^M
\]

Similar to the multivariate RP concept, the failure probability in a multivariate context can also be characterized in “OR”, “AND”, and “Kendall” scenarios expressed by the following equations. For a given critical threshold \( x^* = \{x_1^*, x_2^*, ..., x_d^*\} \), the failure probabilities violating this critical value can be expressed as (Salvadori et al., 2016):

\[
p^{\text{OR}}_T = 1 - (C(F_1(x_1^* | \gamma_1), F_1(x_2^* | \gamma_2), ..., F_d(x_d^* | \gamma_d) | \theta))^T
\]

\[
p^{\text{AND}}_T = 1 - (1 - C(F_1(x_1^* | \gamma_1), F_2(x_2^* | \gamma_2), ..., F_d(x_d^* | \gamma_d) | \theta))^T
\]

\[
p^{\text{Kendall}}_T = 1 - (P(C(F_1(x_1^* | \gamma_1), F_2(x_2^* | \gamma_2), ..., F_d(x_d^* | \gamma_d) | \theta) \leq t))^T
\]

where \( p^{\text{OR}}_T \), \( p^{\text{AND}}_T \), and \( p^{\text{Kendall}}_T \) respectively denote the failure probability in “AND”, “OR” and “Kendall” cases. \( T \) indicates the service time of the facilities under consideration.

Focusing on a bivariate case, the joint RP and the associate failure probability in “OR”, “AND”, and “Kendall” scenarios can be formulated as (Salvadori et al., 2007, 2011; Graler et al., 2013; Sraj et al., 2014; Serinaldi, 2015):

\[
T^{\text{OR}}_{u_1,u_2} = \frac{\mu}{1 - C_{u_1,u_2}(u_1, u_2 | \theta)}
\]
\[
T^{\text{AND}}_{u_1,u_2} = \frac{\mu}{1 - u_1 - u_2 + C_{U_1,U_2}(u_1,u_2 | \theta)}
\]

(12b)

\[
T^{\text{Kendall}}_{u_1,u_2} = \frac{\mu}{1 - P(C_{U_1,U_2}(u_1^*,u_2^*) \leq t)}
\]

(12c)

\[
p_T^{OR} = 1 - (C_{U_1,U_2}(u_1^*,u_2^* | \theta))^T
\]

(12d)

\[
p_T^{\text{AND}} = 1 - (u_1^* + u_2^* - \hat{C}_{U_1,U_2}(u_1^*,u_2^* | \theta))^T
\]

(12e)

\[
p_T^{\text{Kendall}} = 1 - (P(C_{U_1,U_2}(u_1^*,u_2^* | \theta) \leq t))^T
\]

(12f)

where \( u_1 = F_1(x_1 | \gamma_1), u_2 = F_2(x_2 | \gamma_2), u_1^* = F_1(x_1^* | \gamma_1), u_2^* = F_2(x_2^* | \gamma_2) \), \((x_1^*, x_2^*)\) defines the bivariate threshold.

### 2.3. Uncertainty in the Copula-base Risk 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) (Zhang et al., 2015). Several methods have been proposed to quantify parameter uncertainties in copula-based models. For instance, Dung et al. (2015) proposed bootstrap-based methods for quantifying the parameter uncertainties in bivariate copula models. Zhang et al. (2015) employed a Bayesian inference approach for evaluating uncertainties in copula-based hydrologic drought models, in which the Component-wise Hit-And-Run Metropolis algorithm is adopted to estimate the posterior probabilities of model parameters.
In this study, a bootstrap-based algorithm, will be applied to quantify parameter uncertainties in the copula-based multivariate risk model. The procedures for the bootstrap-based algorithm to derive probabilistic distributions of the parameters in both marginal and dependence models are presented as follows:

1. Predefine a large number of bootstrapping samplings $N_B$.

2. Implement the resampling with replacement over observed pairs $Z = (X, Y)$ to obtain $Z^* = (X^*, Y^*)$. $Z^*$ has the same size as $Z$.

3. Fit the chosen marginal distributions to $X^*$ and $Y^*$, and estimate the associated parameters $(\gamma_X, \gamma_Y)$.

4. Fit the chosen copula to $Z^*$, and estimate the parameter in the copula function $\theta$.

5. Repeat step 2–5 $N_B$ times, and obtain $N_B$ sets of $(\gamma_X, \gamma_Y, \theta)$. Moreover, reject those parameters that lead to bad fits for both marginal and copula models, the A-D test and the Cramer-von-Mises test are introduced in the bootstrap procedure to ensure that the obtained parameters can pass statistic tests for both the marginal distribution and copula models. Then the kernel method will be adopted to quantify the probabilistic features for $\gamma_X, \gamma_Y, \theta$.

6. In order to derive bivariate uncertainty bands for a predefined quantile curve $(QC)$ with certain joint RP in ‘AND’, ‘OR’ or ‘Kendall’ (denoted as $T^{AND}$, $T^{OR}$, $T^{Kendall}$), sample $N_B$ sets of $(\gamma_X, \gamma_Y, \theta)$ from the obtained $N_B$ samples.

7. Sample a large number ($N_s$) of $x_i y_j$ from their marginal distributions.

8. For each set of $(\gamma_X, \gamma_Y, \theta)$ from $N_B$, evaluate the joint RPs of $(x_i, y_j)$ ($i = 1, 2, \ldots, N_s$; $y = 1, 2, \ldots, N_s$), and store the pairs of $(x_i, y_j)$ approaching the predefined joint RPs.

9. Repeat step 8 for $N_B$, and for each predefined $QC$, and plot the bivariate uncertainty bands for each quantile $QC$.
2.4. Interactive and Sensitivity Analysis for Parameter 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 have been reported to analyze the effect of uncertainties in the copula model on evaluating the risk for a flood event. To address the above issue, a full-subsampling factorial analysis (FSFA) approach will be proposed to reveal the individual and interactive effects of parameter uncertainties on the predictive uncertainties of different risk inferences.

Consider a copula-based bivariate risk assessment model which has two marginal distributions ($A$ and $B$) and one copula ($C$). The parameters in the two marginal distributions are assumed to be respectively denoted as $\gamma^A$ with $a$ levels and $\gamma^B$ with $b$ levels, while the parameter in the copula is denoted with $\theta^C$ with $c$ levels. The three factor ANOVA model for such a factorial design in terms of the predictive risk (denoted as $R$) in response to the parameters $\gamma^A$, $\gamma^B$, $\theta^C$ and $n$ replicates, can be expressed as:

$$
R_{ijkl} = R_0 + R_{C} + R_{\gamma^A} + R_{\gamma^B} + R_{\theta^C} + R_{\gamma^A, \gamma^B} + R_{\gamma^A, \theta^C} + R_{\gamma^B, \theta^C} + R_{\gamma^A, \gamma^B, \theta^C} + \varepsilon_{ijkl}
$$

\[i = 1, 2, \ldots, c \]
\[j = 1, 2, \ldots, a \]
\[k = 1, 2, \ldots, b \]
\[l = 1, 2, \ldots, n \]

(13)

where $R_0$ denotes the overall mean effect; $R_{C}$, $R_{\gamma^A}$, $R_{\gamma^B}$ respectively indicate the effect for parameter $\theta^C$ in the copula at the $i$th level, parameter $\gamma^A$ in the first marginal distribution at the $j$th level, and parameter $\gamma^B$ in the first marginal distribution at the $k$th level; $R_{\gamma^A, \gamma^B}$, $R_{\gamma^A, \theta^C}$, $R_{\gamma^B, \theta^C}$ indicate interactions between factors $\theta^C$ and $\gamma^A$, $\theta^C$ and $\gamma^B$, as well as $\gamma^A$ and $\gamma^B$, respectively; $R_{\gamma^A, \gamma^B, \theta^C}$ denotes the interaction of factors $\theta^C$, $\gamma^A$ and $\gamma^B$; $\varepsilon_{ijkl}$ denotes the random error component.
Based on Equation (13), the total variability of the predictive risk can be decomposed into its component parts as follows (Montgomery, 2001):

\[ SS_T = SS_{\theta_c} + SS_{\gamma^A} + SS_{\gamma^B} + SS_I + SS_e \]  

(14a)

and

\[ SS_T = \sum_{i=1}^{c} \sum_{j=1}^{a} \sum_{k=1}^{b} \sum_{l=1}^{n} R_{ijkl}^2 - \frac{R_{\ldots}^2}{abc} \]  

(14b)

\[ SS_{\theta_c} = \frac{1}{abn} \sum_{i=1}^{c} R_{ji..}^2 - \frac{R_{\ldots}^2}{abc} \]  

(14c)

\[ SS_{\gamma^A} = \frac{1}{bcn} \sum_{j=1}^{a} R_{j..k}^2 - \frac{R_{\ldots}^2}{abc} \]  

(14d)

\[ SS_{\gamma^B} = \frac{1}{acn} \sum_{k=1}^{b} R_{..kj}^2 - \frac{R_{\ldots}^2}{abc} \]  

(14e)

\[ SS_I = \sum_{i=1}^{c} \sum_{j=1}^{a} \sum_{k=1}^{b} \sum_{l=1}^{n} R_{ijkl}^2 - \frac{1}{n} \sum_{i=1}^{c} \sum_{j=1}^{a} \sum_{k=1}^{b} R_{ijk.}^2 \]  

(14f)

\[ SS_e = SS_{\theta_c^\gamma^A} + SS_{\theta_c^\gamma^B} + SS_{\gamma^A^\gamma^B} + SS_{\theta_c^\gamma^A^\gamma^B} \]  

\[ = SS_T - SS_{\theta_c} - SS_{\gamma^A} - SS_{\gamma^B} - SS_e \]  

(14g)

where \( R_{ijkd} = \sum_{l=1}^{n} R_{ijkl} \), \( R_{ij..} = \sum_{k=1}^{b} \sum_{l=1}^{n} R_{ijkl} \), \( R_{j..k} = \sum_{i=1}^{c} \sum_{l=1}^{n} R_{ijkl} \), \( R_{..kj} = \sum_{i=1}^{c} \sum_{j=1}^{a} \sum_{l=1}^{n} R_{ijkl} \), \( R_{..\ldots} = \sum_{i=1}^{c} \sum_{j=1}^{a} \sum_{k=1}^{b} \sum_{l=1}^{n} R_{ijkl} \). Then the contributions of parameter uncertainties in marginal distributions and dependence structures can be calculated as:

(1) Contribution of parameters in marginal distributions A and B
\[ \eta_A = \frac{SS_{\gamma^A}}{SS_T} \]  
(15a)

\[ \eta_B = \frac{SS_{\gamma^B}}{SS_T} \]  
(15b)

(2) Contribution of the parameter in the dependence structure

\[ \eta_C = \frac{SS_{\gamma^C}}{SS_T} \]  
(15c)

(3) Contribution of internal variability

\[ \eta_e = \frac{SS_e}{SS_T} \]  
(15d)

(4) Contribution of parameter interactions

\[ \eta_i = 1 - \eta_A - \eta_B - \eta_C - \eta_e \]  
(15e)

However, one major issue for ANOVA approach is that the biased variance estimator in ANOVA would underestimate the variance in small sample size scenarios (Bosshard et al., 2013). Thus the sample size may significantly affect the resulting variance contributions expressed in Equations (15a) – (15e). A subsampling approach has been advanced by Bosshard et al. (2013) to diminish the effect of the sample size in ANOVA and has been employed for uncertainty partition in flood design and hydrological simulation (Qi et al., 2016a, b). In such a subsampling scheme, one factor (denoted as \( X \)) with \( T \) levels (these levels can be different values for numerical parameters, or different types for non-numerical factor (e.g. model type)), would choose two levels in each iteration. For \( T \) possible levels of \( X \), we can obtain a total of \( C_T^2 \) possible pairs for \( X \), expressed as a \( 2 \times C_T^2 \) matrix as follows:

\[ g(h, j) = \begin{pmatrix} X_1 & X_1 & \cdots & X_1 & X_2 & X_2 & \cdots & X_{T-2} & X_{T-2} & X_{T-1} \\ X_2 & X_3 & \cdots & X_T & X_3 & X_4 & \cdots & X_{T-1} & X_T & X_T \end{pmatrix} \]  
(16)

However, such a subsampling approach mainly applied to subsample merely one factor or one
parameter (here we refer to this method as single-subsampling ANOVA) in previous studies (Bosshard et al. 2013; Qi et al., 2016a, b). However, one critical issue for the single-subsampling ANOVA is that it will lead to underestimation of individual contribution for the factor to be sampled and overestimation of contributions for those non-sampled factors. Consequently, in this study, we will propose a FSFA approach to subsample all the factors to be addressed, and then quantify the contribution of each factor to the response variation. In the FSFA approach, all factors under consideration will be subsampled, and the corresponding sum of squares will be obtained. The contribution of one factor would be characterized by the mean value of its contribution in each iteration. In detail, for the three factor ANOVA model expressed by Equation (13), the subsampling schemes for the three parameters can be formulated as:

\[
g_{\theta^C}(h_C, j_C) = \left( \theta_1^C \theta_2^C \ldots \theta_{c-1}^C \theta_{c-2}^C \theta_{c-2}^C \theta_{c-2}^C \theta_{c-2}^C \theta_{c-2}^C \right) \quad (17a)
\]

\[
g_{\gamma^A}(h_A, j_A) = \left( \gamma_1^A \gamma_2^A \ldots \gamma_{a-2}^A \gamma_{a-2}^A \gamma_{a-2}^A \gamma_{a-2}^A \gamma_{a-2}^A \right) \quad (17b)
\]

\[
g_{\gamma^B}(h_B, j_B) = \left( \gamma_1^B \gamma_2^B \ldots \gamma_{b-2}^B \gamma_{b-2}^B \gamma_{b-2}^B \gamma_{b-2}^B \gamma_{b-2}^B \right) \quad (17c)
\]

Consequently, there are a total number of \( C_e^2 C_a^2 C_b^2 \) iterations in FSFA for the three-factor model expressed as Equation (13). For each iteration, the sums of squares can be reformulated as:

\[
SS_{\theta^C} = \frac{1}{8n} \sum_{h_c=1}^{2} \sum_{h_a=1}^{2} \sum_{h_b=1}^{2} \sum_{i=1}^{n} R^2 g_{\theta^C}(h_c; j_c) g_{\gamma^A}(h_a; j_a) g_{\gamma^B}(h_b; j_b) - \frac{R^2 g_{\theta^C}(\alpha, j_c) g_{\gamma^A}(\alpha, j_a) g_{\gamma^B}(\alpha, j_b)}{8n} \quad (18a)
\]

\[
SS_{\gamma^A} = \frac{1}{4n} \sum_{h_c=1}^{2} R^2 g_{\theta^C}(h_c; j_c) g_{\gamma^A}(\alpha, j_a) g_{\gamma^B}(\alpha, j_b) - \frac{R^2 g_{\theta^C}(\alpha, j_c) g_{\gamma^A}(\alpha, j_a) g_{\gamma^B}(\alpha, j_b)}{8n} \quad (18b)
\]
\[
SS^j_{y^A} = \frac{1}{4n} \sum_{h_c=1}^{2} R^2_{g_{c}^j(h_c \cdot J_c)g_{a}^j(h_a \cdot J_a)g_{b}^j(h_b \cdot J_b)} - \frac{R^2_{g_{c}^j(\alpha \cdot J_c)g_{a}^j(\alpha \cdot J_a)g_{b}^j(\alpha \cdot J_b)}}{8n}
\]

\[
SS^j_{y^B} = \frac{1}{4n} \sum_{h_b=1}^{2} R^2_{g_{c}^j(h_c \cdot J_c)g_{a}^j(h_a \cdot J_a)g_{b}^j(h_b \cdot J_b)} - \frac{R^2_{g_{c}^j(\alpha \cdot J_c)g_{a}^j(\alpha \cdot J_a)g_{b}^j(\alpha \cdot J_b)}}{8n}
\]

\[
SS^j_{\epsilon} = \sum_{h_c=1}^{2} \sum_{h_a=1}^{2} \sum_{h_b=1}^{2} \sum_{l=1}^{n} R^2_{g_{c}^j(h_c \cdot J_c)g_{a}^j(h_a \cdot J_a)g_{b}^j(h_b \cdot J_b)} - \frac{1}{n} \sum_{h_c=1}^{2} \sum_{h_a=1}^{2} \sum_{h_b=1}^{2} \sum_{l=1}^{n} R^2_{g_{c}^j(h_c \cdot J_c)g_{a}^j(h_a \cdot J_a)g_{b}^j(h_b \cdot J_b)}
\]

290 \[SS^j_f = SS^j_y - SS^j_{y^A} - SS^j_{y^B} - SS^j_{\epsilon} \] (18f)

where

\[
R_{g_{c}^j(h_c \cdot J_c)g_{a}^j(h_a \cdot J_a)g_{b}^j(h_b \cdot J_b)} = \sum_{l=1}^{n} R_{g_{c}^j(h_c \cdot J_c)g_{a}^j(h_a \cdot J_a)g_{b}^j(h_b \cdot J_b)}
\]

\[
R_{g_{c}^j(\alpha \cdot J_c)g_{a}^j(\alpha \cdot J_a)g_{b}^j(\alpha \cdot J_b)} = \sum_{h_c=1}^{2} \sum_{h_a=1}^{2} \sum_{h_b=1}^{2} \sum_{l=1}^{n} R_{g_{c}^j(h_c \cdot J_c)g_{a}^j(h_a \cdot J_a)g_{b}^j(h_b \cdot J_b)}
\]

295 \[
R_{g_{c}^j(\alpha \cdot J_c)g_{a}^j(\alpha \cdot J_a)g_{b}^j(h_b \cdot J_b)} = \sum_{h_c=1}^{2} \sum_{h_a=1}^{2} \sum_{h_b=1}^{2} \sum_{l=1}^{n} R_{g_{c}^j(h_c \cdot J_c)g_{a}^j(h_a \cdot J_a)g_{b}^j(h_b \cdot J_b)}
\]

Also, for each iteration, the corresponding contributions for each factor can be obtained as:

300 \[\eta^j_A = SS^j_{y^A} / SS^j_f \] (19a)

\[\eta^j_B = SS^j_{y^B} / SS^j_f \] (19b)

\[\eta^j_C = SS^j_{\epsilon} / SS^j_f \] (19c)

\[\eta^j_e = SS^j_f / SS^j_f \] (19d)

\[\eta^j_f = 1 - \eta^j_A - \eta^j_B - \eta^j_C - \eta^j_e \] (19e)
Finally, the individual and interactive contributions for those factors can be obtained by averaging the corresponding contributions in all iterations, expressed as:

\[
\eta_A = \frac{1}{J} \sum_{j=1}^{J} SS_{j \lambda} / SS_T
\]  
(20a)

\[
\eta_B = \frac{1}{J} \sum_{j=1}^{J} SS_{j \mu} / SS_T
\]  
(20b)

\[
\eta_C = \frac{1}{J} \sum_{j=1}^{J} SS_{j \phi} / SS_T
\]  
(20c)

\[
\eta_e = \frac{1}{J} \sum_{j=1}^{J} SS_{j \epsilon} / SS_T
\]  
(20d)

\[
\eta_i = \frac{1}{J} \sum_{j=1}^{J} \eta_i^j
\]  
(20e)

where \( J = C_c^2 C_a^2 C_b^2 \)

### 3. Applications

The proposed FSFC approach can be applied for various multivariate risk inference problems. In this study, we will apply FSFC for multivariate flood risk inference at the Wei River basin in China. The Weihe River plays a key role in the economic development of western China, and thus is known regionally as the ‘Mother River’ of the Guanzhong Plain of the southern part of the loess plateau (Song et al. 2007; Zuo et al. 2014; Du et al. 2015, Xu et al., 2016). It originates from the Niaoshu Mountain at an elevation of 3485 m above mean sea level in Weiyuan County of Gansu Province (Du et al. 2015). The Weihe River basin is characterized by a semi-arid and sub-humid continental monsoon climate, resulting in significant temporal-spatial variations in precipitation, with an annual precipitation of 559 mm (Xu et al., 2016). Furthermore, there is a strong decreasing gradient from south to north, in which
the southern region experiences a sub-humid climate with annual precipitation ranging from 800 to 1000 mm, whereas the northern region has a semi-arid climate with annual precipitation ranging from 400 to 700 mm (Xu et al., 2016). Over the entire basin, the mean temperature ranges from 6 to 14 °C, the annual potential evapotranspiration fluctuates from 660 to 1,600, and the annual actual evapotranspiration is about 500 mm (Du et al. 2015).

Observed daily streamflow data at Xianyang and Zhangjiashan gauging stations are applied for hydrologic risk analysis. Figure 1 show the locations of these three gauging stations. Based on the daily stream flow data, the flood peak applied is defined as the maximum daily flow over a period and the associated flood volume is considered as the cumulative flow during the flood period. In this study, the flood characteristics are obtained based on an annual scale. This means that one flood event is identified in each year. The detailed method to identify the flood peak and the associated flood volume can be found in Yue (2000, 2001). Table 1 shows some descriptive statistical values for the considered variables (peak discharge, Q; hydrograph volume, V), in which 47 and 55 flood events are characterized at the Xianyang and Zhangjiashan station, respectively.

4. Results Analysis

4.1. Model Evaluation and Selection

There are a number of potential probabilistic models for modelling individual flood variables and their dependence structures. In this study, five alternative distributions, including gamma, generalized extreme value (GEV), lognormal (LN), Pearson type III (P III), and log-Pearson type III (LP III)
distributions, are employed to describe the probabilistic features of the chosen flood variables (i.e. peak and volume). Moreover, goodness-of-fit tests are performed through the indices of Kolmogorov-Smirnov test (K-S test), root mean square error (RMSE) and Akaike Information Criterion (AIC), to screen the performance of those potential models. The results are presented in Table 2. The results indicate that all five parametric distributions can produce satisfactory results, with all p-values larger than 0.05. However, it can be concluded that the GEV and lognormal approaches show best performance for respectively modelling flood peak and volume at both gauging stations.

In addition, a total number of six copulas, including Gaussian, Student t, Clayton, Gumbel, Frank and Joe copulas, are considered as the candidate models for quantifying the dependence structures for flood peak-volume at Xianyang and Zhangjiashan gauging stations. Also, the goodness-of-fit statistics is performed based on the Cramér von Mises statistic proposed by Genest et al. (2009). An e indices of RMSE and AIC would be further employed to evaluate the performance of the obtained copulas and identify the most appropriate ones. Table 3 shows statistical test results for the selected copulas. The results show that, for the Zhangjiashan station, all candidate copulas except the Joe copula performed well, while all six copulas would be able to provide satisfactory risk inferences at the Xianyang station. Moreover, based on the values of RMSE and AIC, the Gumbel copula was chosen to model the dependence of flood peak and volume at Zhangjiashan station, while the Joe copulas performed best at the Xianyang station. Consequently, the Gumbel and Joe copula would be chosen in this study to further characterize the uncertainty in model parameters and the resulting risks at Zhangjiashan and Xianyang station, respectively.
4.2. Uncertainty in Model Parameters and Risk Inferences

Based on the results in Tables 2 and 3, the multivariate risk inference model would be established, in which the GEV and lognormal distributions would respectively be adopted to model the individual flood variables at both gauging stations, while in comparison, the Gumbel and Joe copulas would respectively be employed for Zhangjiashan and Xianyang stations. Afterward, uncertainties would be characterized based on the bootstrap algorithm illustrated in Section 2.3. In current study, a total number of 5000 samples would be chosen in order to generally visualize the uncertainty features in model parameters. The probabilistic features for obtained parameters values (i.e. shape, scale and location for GEV, meanlog, sdlog for LN, and theta for copula) for each sample scenario would be described by the kernel method. Figure 3 exhibit the probabilistic distributions for the six unknown parameters in the established multivariate risk inference model. Extensive uncertainties exist in the parameters for both the marginal distribution and dependence model. As presented in Figure 3, each parameter, except the meanlog in the LN distribution, exhibit noticeable uncertainty. Moreover, most of the parameter uncertainties are approximately normally distributed.
It is quite apparent that different parameter values in the copula model would lead to different risk inference results. Consequently, parameter uncertainties in the marginal distributions and copula function would definitely result in uncertainties in multivariate risk inferences. Based on the copula model, some multivariate risk indices can be easily obtained, such as the joint return period in OR, AND and Kendall, as expressed in Equations (12a) – (12c). However, due to parameter uncertainties, these risk indices may also exhibit some degrees of uncertainty. Figures 4 – 6 describe uncertainties for the joint RP in AND, OR and Kendall at the two stations. In general, the predictive RP in AND exhibit most significant uncertainty, followed by the predictive RP in OR and Kendall. However, for moderate or large flood events, considerable uncertainties can be observed in the inferences for all the three joint RPs. Specifically, noticeable uncertainties exist in the predictive joint RP of AND even for a minor flood event with a 5-year joint RP. For some large flood events with a joint RP around 100 years, the predictive RP in AND shows remarkable uncertainty, ranging from less than 50 years to larger than 200 years. For the joint RP in OR and Kendall, slight uncertainty may exist for small flood events (e.g. 2-year or 5-year joint RP). Nevertheless, apparently uncertainties can be observed in the predictive joint RP even for moderate flood events. As shown in Figure 5, considerable uncertainties may appear in the predictive joint RP of OR even for a flood with an actual joint RP of 20 years, while prediction of the Kendall RP for a 20-year (in Kendall RP) flood event may range from 10 to 50 years, as presented in Figure 6.

Place Figures 4-6 here

4.3 Individual and Interactive Effects of Parameter Uncertainties

It has been observed that parameter uncertainties in the copula-based multivariate risk model would
lead to significantly imprecise risk predictions. However, one critical issue to be addressed is how the parameter uncertainties and their interaction would influence the risk inference. Consequently, a multilevel factorial analysis, based on Equations (13) and (14) would be proposed to primarily visualize the individual and interactive effects of parameter uncertainties in the marginal and dependence models on the resulting risk inferences. In this study, a total number of 6 parameters (i.e. three from GEV, two from LN, and one from copula) would be addresses, and based on probabilistic features of these parameters, three quantile levels (i.e. 0.1, 0.5 and 0.9) would be chosen to characterize the resulting risk inferences under different parameter values. This would finally form a $3^6$ factorial design, which has six factors with each having three levels. The failure probability denoted as Equations (11) would be considered as the responses in this factorial design.

The main and interactive effects of parameters uncertainties on the failure probabilities in AND are visualized in Figure 7. It is noticeable that at the two gauge stations, parameters uncertainties pose similar main and interactive effects on the failure probabilities in AND, which indicates that parameters’ effects (individual and interactive) on the failure probability in AND are independent with the location of gauge stations. More specifically, variations in the shape parameter in GEV and $sdlog$ parameter in LN would lead to more changes in the corresponding responses (i.e. failure probability in AND) than the variations in other parameters. Also, as shown in Figure 7, the parameter in the copula function (i.e. $Cop_{theta}$), describing dependence of the two flood variables, would not have an effect on the resulting risk as visible as the effects from the parameters (except the location parameter in GEV) in the marginal distributions. In terms of parameter interactions, the significance of interactive effects for different parameters is various. The interactive curves for some parameters (e.g. GEV_shape and GEV_location) are nearly parallel at the three levels, indicating an insignificant interaction for these two parameters on the inferred risk. In comparison, there are also some interactive curves intersecting among each other (e.g. GEV_shape and LN_meanlog), implying a significant interaction among these two parameters.
Table 4 provides the results from an ANOVA table for the failure probability in AND. It is quite interesting that: i) even though the effect from the parameter in the copula function is not as visible as the effects from the parameters (except the location parameter in GEV) in the marginal distributions (as shown in Figure 7), such an effect is still statistically significant; ii) the effect from the location parameter of GEV is statistically insignificant, which also lead to insignificant interactive effects between the location parameter and other parameters; iii) the interactions between the parameter in copula and the parameters in marginal distributions would be more likely statistically insignificant; iv) the statistical significance (significant or not) for individual and interactive effects from parameters is almost the same between these two gauge stations. All these conclusions obtained from Table 4 are consistent with the implications described in Figure 7.

In terms of the failure probabilities in OR and Kendall, as presented in Figures 8 and 9, these have similar pattern with the failure probability in AND (presented in Figure 7). The individual/main effects from the marginal distributions (except the location parameter in GEV) are generally more visible than the parameter in copula. Also, some interactive curves, especially the curves between GEV_location and others, are parallel, showing insignificant interaction between those parameters. More detailed characterization of the main and interactive effects for the failure probabilities in OR and Kendall is described in the ANOVA tables in Tables 5 and 6. These two tables show some slight differences from the conclusions given by Table 4. The location parameter in GEV also has statistically significant effect on the results failure probabilities in OR and Kendall, which also leads to some significant interactions between this parameter and other model parameters. For the failure probability in Kendall,
the parameter in copula would have more interactions with other parameters in marginal distributions than the interactions in the failure probability in AND and OR. As presented in Table 6, the parameter in copula would have statistically significant effect on the inferred failure probability in Kendall with other parameters except the location parameter in GEV. These results are also implied in the main effects plots and full interactions plot matrices in Figures 8 and 9.

Based on the three-level factorial analysis, it can be generally concluded that the parameters in marginal distributions (except the location parameter in GEV) would have more individual effects on joint risk inference than the parameter in copula. The risk indices (i.e. AND, OR, or Kendall) would not have significantly influence individual effects of model parameters. However, for the interactive effects among model parameters, they may exhibit slightly different patterns. Specifically, the parameter in copula would have more significant interactions with parameters in marginal distributions on the failure risk in Kendall than the other two risk indices. Moreover, the individual and interactive effects from model parameters on risk inferences would not influence by the location of gauge stations.

4.4. Contribution Partition of Uncertainty Sources

As a result of parameter uncertainties, the predictive failure probabilities exhibit noticeable uncertainties, as shown in Figures 4-6. The three-level factorial analysis based on Equations (11) is able to provide a primary description and visualization related to the individual and interactive effects of parameter uncertainty on the inferred failure probabilities. However, two critical issues be answered...
are: (i) how much would parameter uncertainties contribute to the variation of the inferred risk values? and (ii) do these contributions change significantly for failure probabilities with different service time scenarios? To address these two issues and get reliable results, a full-subsampling factorial approach (FSFA) has been proposed, which would be formulated as Equations (16) – (20). Also, similar with the three-level factorial analysis, three quantile levels would be selected at 0.1, 0.5, and 0.9. Based on FSFA, each parameter at its three quantile values (0.1, 0.5, 0.9) would be further subsampled into three scenarios of two quantile values (i.e. (0.1, 0.5), (0.1, 0.9), and (0.5, 0.9)). For this study, we have a total number of 6 parameters with each choose its three quantile values at 0.1, 0.5, and 0.9, which would lead to a total number of 729 (i.e. 3\(^6\)) two-level factorial designs.

Figure 10 shows the detailed contributions of the model parameters on uncertainty in predictive failure probabilities of AND at the two gauge stations. It can be observed that, even though some discrepancies exist at Zhangjiashan and Xingshan stations, the detailed contributions for each parameter and their interaction show quite similar features between these two stations. In detail, uncertainty in the shape parameter in GEV has the most significant impact on the failure probability in AND, followed by sdlog in LN, parameter interaction, meanlog in LN, and scale parameter in GEV. Moreover, the uncertainty in the parameter in copula would not lead to significant variation in the resulting failure probability predictions in AND, which would merely make a contribution less than 0.5%. Such conclusions are also generally consistent to the ANOVA results presented in Figure 7 and Table 4. Furthermore, as the increase in service time, the contributions of each parameter and their interactions would not vary significantly. Some individual contributions from parameter uncertainties would slightly increase while other individual contributions may slightly decrease. However, the effect from parameter interactions would generally increase as the increase of service time. In comparison, the enhancement in design standard for hydraulic infrastructure would lead to more chance for decreasing in individual effects and, at the same time, increasing in parameter interactions. For instance, as the flood design standard
increases from 200-year to 500-year for a hydraulic facility with 30-year service time near the Zhangjishan station, the interactive effect of model parameters would increase from 15.14% to 18.09%.

In terms of the failure probability in OR, the individual and interactive effects of model parameters on predictive risk uncertainties show similar patterns with the parameters’ effects on the failure probability in AND. As shown in Figure 11, the shape parameter in the GEV distribution and the sdlog in the LN distribution are the two major sources for uncertainties in failure probabilities in OR. However, compared with the failure probability in AND, parameter interaction has a less effect on the resulting uncertainty of risk inference in OR. As shown in Figures 10 and 11, the effect of parameter interaction on the risk in AND ranges between 13.96% and 20.05%, while in comparison, the parameters’ interactive effect on the risk in OR varies within [10.25%, 11.57%]. Apparently, it can also be observed that some external factors such as the design standard and service time of hydraulic infrastructures have less influence on the parameters’ interaction on risk in OR than the risk in AND. However, the first contributor (i.e. shape parameter in GEV) would have a more contribution on the predictive uncertainty in the failure probability in OR as the increase in the design standard, while in comparison, this contributor would have a less contribution on the risk in AND. For instance, as the design return period of flood (i.e. design standard) increases from 200 to 500 years and the service time of the hydraulic facility is 30 years, the contribution of the shape parameter in GEV would increase from 47.62% to 50.64% for the failure probability in OR at the Xianyang station, while the parameter’s contribution on the failure probability in AND decreases from 49.26% to 45.77%.
For the failure probability in Kendall, the contributions of model parameters and their interaction are presented in Figure 12. Similar to the failure probabilities in AND and OR, the shape parameter in the GEV distribution and the sdlog parameter in the LN distribution are the two major contributors, which can account for nearly 70% or more in the predictive uncertainty of the failure probability in Kendall. Meanwhile, the scale parameter in GEV, meanlog in LN, and parameters’ interaction also have noticeable effects on the risk in Kendall, ranging from 4.72% (scale parameter in GEV) to 12.64% (mean log in LN). Conversely, the location parameter in GEV and dependence parameter in copula merely have quite minor individual effects. However, it is noticeable that, although the dependence parameter has a minor effect (0.78%, 1.03%) on the risk in Kendall, such an effect is much higher than the effect on the risk in AND (less than 0.23%) and the risk in OR (less than 0.06%).

Even though the prediction equations for the failure probabilities in AND, OR and Kendall, as presented in Equations (12) are different, the impacts of parameter uncertainties show quite similar features, in which the shape parameter in GEV and the sdlog in LN are the two major contributors to the predictive uncertainties in risk inferences. Nearly 70% and more variability in the uncertainties in risk inferences can be accounted by the uncertainties in the shape parameter in GEV and sdlog parameter LN. Also some external factors such as flood design and facility service time may have different
influence for parameters’ effects on different risk indices, such influence is not significant and would not lead to remarkable changes in parameters’ contribution to risk inferences. Parameters’ interaction has a more effect on risk inference in AND than the other two risk indices (i.e. OR, Kendall), while the contribution from the dependence parameter, even though not noteworthy, has a more effect on the risk inference in Kendall.

5. Discussion

5.1. Contribution Partition of Uncertainty Sources through Different Approaches

In current study, the individual and interactive contributions of parameter uncertainties are quantified through the developed FSFA approach, in which each parameter has three levels (i.e. 0.1, 0.5, 0.9 quantiles) to be subsampled. In fact, the parameters’ contribution can also be characterized by traditional factorial analysis (FA) approach based on Equations (15) as well as the FSFA approach with more factor levels (e.g. 4 or 5 levels for each parameter).

Figure 13 shows the comparison of parameter contributions to predictive uncertainty for failure probabilities in AND at Zhangjiashan station for three and four parameter levels scenarios for the design standard of 200-year. The results of Figure 13(b) are obtained through the FSFA approach with each parameter having four levels to be its quantiles at 0.1, 0.35, 0.6, 0.85. Also, Table 7 presents the parameter contributions to predictive uncertainty in failure probabilities obtained by traditional FA approach for Zhangjiashan stations with the design standard of 200-year and service time of 30-year.

Place Figure 13 and Table 7 here
It can be observed that for different subsampling scenarios, the resulting contributions may be different. However, such a difference would be tolerant since (1) the variations of parameters’ contributions are relatively small and mainly happen for the first two contributors, (2) the total contribution of the first two contributors does not change remarkably (around 70% in total), (3) the contributions of other factors especially the parameters’ interaction do not vary significantly, and (4) the rank of the contributions from different sources does not change for the two subsampling scenarios. In comparison, as presented in Table 7, the contribution partition of parameter uncertainties obtained through traditional FA shows totally different patterns for different risk inferences. Specifically, traditional FA approach would significantly overestimate parameter interactive effect on risk inference in AND, while at the same time, underestimate the interactive effect on risk inference in OR and Kendall. Consequently, the contribution rank of parameter uncertainties from traditional FA is different from the results obtained through the developed FSFA approach.

As shown in Figure 13, the proposed FSFA approach may lead to slightly different results for different subsampling schemes (four or five levels). However, an increase in parameter level would highly increase computational demand. For instance, if each parameter has four levels, FSFA approach would lead to a total number of 46,656 (i.e. $6^6$) two-level factorial designs. Moreover, the subsampling scheme for factors with five levels would lead to a total number of one million (i.e. $6^{10}$) two-level factorial designs. Consequently, the three-level subsampling scheme would generally be recommended and also can generate acceptable results.

5.2. Correlation among Parameters’ Contributions

The proposed FSFA approach would generally produce a great number of two-level factorial designs. For one specific factor (e.g. GEV_shape), it would have two levels (lower and upper levels) for all
factorial designs. However, the detailed value for the lower or upper level may be different in different factorial designs. This may finally lead to different contributions for this factor. Figure 14 presents the variations of parameters’ contributions to the prediction of failure probabilities in AND, OR, and Kendall. We already concluded that the shape parameter in GEV (i.e. GEV_shape) and the sdlog in LN (i.e. LN_sdlog) distribution would generally have the most significant contributions to predictive uncertainties in risk inferences. However, as shown in Figure 14, the detailed contributions for these two parameters would vary remarkably for different level values in different factorial designs. In comparison, the contributions from other parameters and their interaction have less fluctuation than the individual contributions of GEV_shape and LN_sdlog. For instance, although the meanlog in LN (i.e. LN_meanlog), with an average contribution more than 10%, may have some chances to pose a predominant contribution more than 50%, most of its contribution are positively distributed within [0, 25%]. Also, even though the parameters’ interaction has a noteworthy average contribution larger than 10%, all the detailed contributions in different factorial designs are located within [0, 25%].

--------------------------------------------
Place Figure 14 here
--------------------------------------------

It has been observed that the parameters’ contribution may vary significantly due to the differences in factor values in different factorial designs. One potential issue to be addressed is that how those individual and interactive contributions correlate with each other. Figure 15 presents the Pearson’s correlation among individual and interactive contributions of model parameters to different risk inferences (i.e. failure probabilities in AND, OR, and Kendall). It is noticeable that the parameters in the LN distribution (i.e. LN_sdlog, LN_meanlog) are generally negatively correlated with the parameters in the GEV distribution (i.e. GEV_shape, GEV_scale, and GEV_location). Also, for one marginal
distribution (LN or GEV), its parameters are positively correlated. This implies that an increase in the contribution of one parameter would lead to a contribution increase for parameters within the same distribution and at the same time result in a contribution decrease for all parameters in the other distribution. Moreover, if statistically significant, the contribution of the dependent parameter (i.e. parameter in copula) generally has positive correlation with the contributions from other parameters except GEV_shape and parameters’ interaction. Also, the contributions from parameters’ interaction are generally negatively correlated with the individual contributions from other parameters if such correlation is statistically significant.

The proposed FSFA approach can generally characterize how parameter uncertainties would influence the predictive uncertainties in risk inferences. A large number of two-level factorial designs would be produced due to different subsampling procedures and then generate different partition results for parameters’ contributions. However, for different risk inferences (i.e. failure probabilities in AND, OR, and Kendall), these partition results have similar variation features and also show similar correlation plots.

6. Conclusions

Uncertainty quantification is an essential issue for both univariate and multivariate hydrological risk analysis. A number of research works have been posed to reveal uncertain features in multivariate hydrological risk inference. However, it is required to know the major sources/contributors for
predictive uncertainties in multivariate risk inferences. In this study, a full-subsampling factorial copula approach (FSFC) has been proposed for uncertainty quantification and partition in multivariate hydrologic risk inference. In FSFC, a copula-based multivariate risk model has been developed and the bootstrap method is adopted to quantify the probabilistic features for the parameters in both marginal distributions and the dependence model. A full-subsampling factorial analysis (FSFA) approach is finally developed to diminish the effect of the sample size in traditional ANOVA and provided reliable contribution partition for parameter uncertainties in different risk inferences.

This study is the first attempt to characterize parameter uncertainties in a copula-based multivariate hydrological risk model and further reveal their contributions to predictive uncertainties for different risk inferences. As an improvement of ANOVA, the developed FSFA method can mitigate the effect of bias variance estimation in ANOVA and generate reliable results. Moreover, another noteworthy feature for the FSFA approach is that it cannot only characterize the impacts for continuous factors (e.g. model parameters in this study), but also reveal the impacts of discrete or non-numeric factors. Such a feature can allow the proposed FSFA approach to be employed to further explore the impacts of non-numeric factors (e.g. model structures, sample size) in hydrologic systems analysis.

**Code and data availability:** The flooding data for the studied catchments as well as the associated code for this study can be gathered upon email request to the corresponding authors.

**Author contributions.** YRF, KH, GHH and YPL designed the research. YRF and FW carried out the research, developed the model code and performed the simulations. YRF prepared the manuscript with contributions from all co-authors.

**Competing interests.** The authors declare that they have no conflict of interest.
Financial support. This work was jointly funded by the National Key Research and Development Plan (2016YFC0502800), the National Natural Science Foundation of China (51520105013), and the Natural Sciences and Engineering Research Council of Canada.

References

https://doi.org/10.5194/hess-2019-434
Preprint. Discussion started: 11 October 2019
© Author(s) 2019. CC BY 4.0 License.


Yue S.: The bivariate lognormal distribution to model a multivariate flood episode. Hydrol. Processes, 14(14), 2575-2588, 2000


Captions of Tables

Table 1. Flood characteristics for different stations

Table 2. Statistical test results for marginal distribution estimation: LN means lognormal distribution, P III means Pearson Type III distribution, and LP III means log-Pearson Type III distribution. K-S test denotes the Kolmogorov–Smirnov test.

Table 3. Performance for quantifying the joint distributions between flood peak and volume through different copulas: CvM is the Cramér von Mises statistic proposed by Genest et al. (2009), with p-value larger than 0.05 indicating satisfactory performance. Table 4. Statistical test results for marginal distribution estimation

Table 4. ANOVA table for failure probability in AND: A indicates the shape parameter in GEV, B indicates the scale parameter of GEV, C indicates the location parameter of GEV, D means the meanlog of LN, E means the sdlog of LN, and F mean the parameters (i.e. theta) in copula Comparison of RMSE and AIC values for joint distributions through different copulas

Table 5. ANOVA table for failure probability in OR: A indicates the shape parameter in GEV, B indicates the scale parameter of GEV, C indicates the location parameter of GEV, D means the meanlog of LN, E means the sdlog of LN, and F mean the parameters (i.e. theta) in copula

Table 6. ANOVA table for failure probability in Kendall: A indicates the shape parameter in GEV, B indicates the scale parameter of GEV, C indicates the location parameter of GEV, D means the meanlog of LN, E means the sdlog of LN, and F mean the parameters (i.e. theta) in copula

Table 7. Contributions of parameter uncertainties obtained by three level ANOVA to predictive failure probabilities for a design return period of 200-year and service time of 30-year

Captions of Figures

Figure 1: Framework of the proposed FSFC approach

Figure 2. The location of the studied watersheds. Wei River is the largest tributary of Yellow river, with a drainage area of 135,000 km². The historical flood data from Xianyang and Zhangjiashan stations on the Wei River are analyzed through the proposed FSFC approach. Figure 3. Probabilistic features for parameters in marginal distributions and copula: for both Xianyang and Zhangjiashan stations, the GEV (parameters include shape, scale and location) function would be employed to quantify the distribution of flood peak, while the lognormal distribution (parameters denoted as meanlog and sdlog) is applied for flood volume. The Gumbel and Joe copula (parameter denoted as theta) would be respectively adopted to model the dependence between flood peak and volume at Zhangjiashan and Xianyang stations.

Figure 4. Uncertainty quantification of the joint RP in “AND”: the red dash lines indicate the predictive means, the two blue dash lines respectively indicate the 5% and 95% quantiles, and the grey lines indicate the predictions under different parameter samples with the same joint RP of the red and blue
The cyan lines denote the predictions under different return periods with the model parameters being their mean values.

**Figure 5.** Uncertainty quantification of the joint RP in “OR”: the red dash lines indicate the predictive means, the two blue dash lines respectively indicate the 5% and 95% quantiles, and the grey lines indicate the predictions under different parameter samples with the same joint RP of the red and blue dash lines; The cyan lines denote the predictions under different return periods with the model parameters being their mean values.

**Figure 6.** Uncertainty quantification of the joint RP in “Kendall”: the red dash lines indicate the predictive means, the two blue dash lines respectively indicate the 5% and 95% quantiles, and the grey lines indicate the predictions under different parameter samples with the same joint RP of the red and blue dash lines; The cyan lines denote the predictions under different return periods with the model parameters being their mean values.

**Figure 7.** Main effects plot and full interactions plot matrix for parameters on the failure probability in AND at the two gauge stations.

**Figure 8.** Main effects plot and full interactions plot matrix for parameters on the failure probability in OR at the two gauge stations.

**Figure 9.** Main effects plot and full interactions plot matrix for parameters on the failure probability in Kendall at the two gauge stations.

**Figure 10.** Contributions of parameter uncertainties to predictive failure probabilities in AND under different design standards (i.e. return periods (RP)) and different service periods.

**Figure 11.** Contributions of parameter uncertainties to predictive failure probabilities in OR under different design standards (i.e. return periods (RP)) and different service periods.

**Figure 12.** Contributions of parameter uncertainties to predictive failure probabilities in Kendall under different design standards (i.e. return periods (RP)) and different service periods.

**Figure 13.** Comparison of parameter contributions to predictive uncertainty for failure probabilities under different levels of subsampling for Zhangjiashan station: three (i.e. 0.1, 0.5, 0.9) and four (i.e. 0.1, 0.35, 0.6, 0.85) level quantiles are adopted for subsampling and the design return period is 200 years.

**Figure 14.** Variation of parameters’ contributions for different risk inferences at the Zhangjiashan Station for a design standard of 200-year and a service time of 30-year.

**Figure 15.** Correlation for parameters’ contributions on risk inferences at Zhangjiashan station for a design standard of 200-year and a service time of 30-year: The cross sign indicates the correlation is statistically insignificant.
Table 1. Flood characteristics for different stations

<table>
<thead>
<tr>
<th>Station name</th>
<th>period</th>
<th>flood variable</th>
<th>Peak (m$^3$/s)</th>
<th>Volume (m$^3$/s day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xianyang</td>
<td>1960-2006</td>
<td>Minimum</td>
<td>139</td>
<td>317</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Median</td>
<td>1350</td>
<td>2491</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maximum</td>
<td>12380</td>
<td>17802</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minimum</td>
<td>217</td>
<td>303.7</td>
</tr>
<tr>
<td>Zhangjiashan</td>
<td>1958-2012</td>
<td>Median</td>
<td>775</td>
<td>1365.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maximum</td>
<td>3730</td>
<td>7576.1</td>
</tr>
</tbody>
</table>
Table 2. Statistical test results for marginal distribution estimation: LN means lognormal distribution, P III means Pearson Type III distribution, and LP III means log-Pearson Type III distribution. K-S test denotes the Kolmogorov–Smirnov test.

<table>
<thead>
<tr>
<th>Station name</th>
<th>Flooding variables</th>
<th>Marginal distribution</th>
<th>K-S test</th>
<th>RMSE</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$T$</td>
<td>P-value</td>
<td></td>
</tr>
<tr>
<td>Zhangjiashan</td>
<td>Peak</td>
<td>Gamma</td>
<td>0.0745</td>
<td>0.5471</td>
<td>0.0378</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GEV</td>
<td>0.0724</td>
<td>0.9151</td>
<td>0.0275</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LN</td>
<td>0.0805</td>
<td>0.8403</td>
<td>0.0283</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P III</td>
<td>0.0893</td>
<td>0.7386</td>
<td>0.0395</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LP III</td>
<td>0.0795</td>
<td>0.8508</td>
<td>0.0324</td>
</tr>
<tr>
<td></td>
<td>Volume</td>
<td>Gamma</td>
<td>0.1460</td>
<td>0.1735</td>
<td>0.0596</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GEV</td>
<td>0.1017</td>
<td>0.5839</td>
<td>0.0369</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LN</td>
<td>0.0904</td>
<td>0.7250</td>
<td>0.0361</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P III</td>
<td>0.1589</td>
<td>0.1112</td>
<td>0.0737</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LP III</td>
<td>0.0967</td>
<td>0.6468</td>
<td>0.0367</td>
</tr>
<tr>
<td>Xianyang</td>
<td>Peak</td>
<td>Gamma</td>
<td>0.1159</td>
<td>0.5533</td>
<td>0.0372</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GEV</td>
<td>0.0875</td>
<td>0.8645</td>
<td>0.0305</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LN</td>
<td>0.1051</td>
<td>0.6763</td>
<td>0.0436</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P III</td>
<td>0.1202</td>
<td>0.5051</td>
<td>0.0416</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LP III</td>
<td>0.1321</td>
<td>0.3848</td>
<td>0.0617</td>
</tr>
<tr>
<td></td>
<td>Volume</td>
<td>Gamma</td>
<td>0.1146</td>
<td>0.5305</td>
<td>0.0450</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GEV</td>
<td>0.0540</td>
<td>0.9980</td>
<td>0.0195</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LN</td>
<td>0.0670</td>
<td>0.9749</td>
<td>0.0192</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P III</td>
<td>0.1005</td>
<td>0.6913</td>
<td>0.0377</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LP III</td>
<td>0.0722</td>
<td>0.9522</td>
<td>0.0313</td>
</tr>
</tbody>
</table>
Table 3. Performance for quantifying the joint distributions between flood peak and volume through different copulas: CvM is the Cramér von Mises statistic proposed by Genest et al. (2009), with p-value larger than 0.05 indicating satisfactory performance.

<table>
<thead>
<tr>
<th>Location</th>
<th>Copula</th>
<th>RMSE</th>
<th>AIC</th>
<th>CvM</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhangjiashan</td>
<td>Gaussian</td>
<td>0.0669</td>
<td>-295.5144</td>
<td>7.9302</td>
<td>0.7770</td>
</tr>
<tr>
<td></td>
<td>Student t</td>
<td>0.0669</td>
<td>-293.5237</td>
<td>8.5203</td>
<td>0.5976</td>
</tr>
<tr>
<td></td>
<td>Gumbel</td>
<td>0.0637</td>
<td>-300.8577</td>
<td>7.9342</td>
<td>0.7580</td>
</tr>
<tr>
<td></td>
<td>Frank</td>
<td>0.0690</td>
<td>-292.0723</td>
<td>9.0704</td>
<td>0.4480</td>
</tr>
<tr>
<td></td>
<td>Joe</td>
<td>0.0606</td>
<td>-306.3185</td>
<td>11.0321</td>
<td>0.0290</td>
</tr>
<tr>
<td>Xinshan</td>
<td>Gaussian</td>
<td>0.0513</td>
<td>-277.1704</td>
<td>8.4731</td>
<td>0.2400</td>
</tr>
<tr>
<td></td>
<td>Student t</td>
<td>0.0510</td>
<td>-275.6834</td>
<td>8.2295</td>
<td>0.2885</td>
</tr>
<tr>
<td></td>
<td>Clayton</td>
<td>0.0618</td>
<td>-259.7391</td>
<td>8.2051</td>
<td>0.3240</td>
</tr>
<tr>
<td></td>
<td>Gumbel</td>
<td>0.0477</td>
<td>-283.9933</td>
<td>7.1344</td>
<td>0.6700</td>
</tr>
<tr>
<td></td>
<td>Frank</td>
<td>0.0562</td>
<td>-268.6861</td>
<td>8.2725</td>
<td>0.2940</td>
</tr>
<tr>
<td></td>
<td>Joe</td>
<td>0.0446</td>
<td>-290.2631</td>
<td>6.9905</td>
<td>0.6540</td>
</tr>
</tbody>
</table>
Table 4. ANOVA table for failure probability in AND: A indicates the shape parameter in GEV, B indicates the scale parameter of GEV, C indicates the location parameter of GEV, D means the mean log of LN, E means the sd log of LN, and F mean the parameters (i.e. theta) in copula

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Zhangjiashan</th>
<th>Xinyang</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.37</td>
<td>0.59</td>
</tr>
<tr>
<td>B</td>
<td>0.018</td>
<td>0.013</td>
</tr>
<tr>
<td>C</td>
<td>8.313E-005</td>
<td>8.642E-005</td>
</tr>
<tr>
<td>D</td>
<td>0.059</td>
<td>0.082</td>
</tr>
<tr>
<td>E</td>
<td>0.18</td>
<td>0.31</td>
</tr>
<tr>
<td>F</td>
<td>9.379E-004</td>
<td>7.813E-004</td>
</tr>
<tr>
<td>AB</td>
<td>2.874E-003</td>
<td>8.730E-003</td>
</tr>
<tr>
<td>AC</td>
<td>1.179E-005</td>
<td>5.434E-005</td>
</tr>
<tr>
<td>AD</td>
<td>0.047</td>
<td>0.079</td>
</tr>
<tr>
<td>AE</td>
<td>0.14</td>
<td>0.28</td>
</tr>
<tr>
<td>AF</td>
<td>4.311E-004</td>
<td>4.687E-004</td>
</tr>
<tr>
<td>BC</td>
<td>2.905E-007</td>
<td>2.235E-008</td>
</tr>
<tr>
<td>BD</td>
<td>2.422E-003</td>
<td>2.465E-003</td>
</tr>
<tr>
<td>BE</td>
<td>6.956E-003</td>
<td>8.669E-003</td>
</tr>
<tr>
<td>BF</td>
<td>8.325E-006</td>
<td>2.555E-006</td>
</tr>
<tr>
<td>CD</td>
<td>1.143E-005</td>
<td>1.652E-005</td>
</tr>
<tr>
<td>CE</td>
<td>3.235E-005</td>
<td>5.669E-005</td>
</tr>
<tr>
<td>CF</td>
<td>3.820E-008</td>
<td>1.559E-008</td>
</tr>
<tr>
<td>DE</td>
<td>1.792E-003</td>
<td>6.919E-003</td>
</tr>
<tr>
<td>DF</td>
<td>9.625E-005</td>
<td>1.288E-004</td>
</tr>
<tr>
<td>EF</td>
<td>3.238E-004</td>
<td>4.540E-004</td>
</tr>
<tr>
<td>Error</td>
<td>0.016</td>
<td>0.038</td>
</tr>
<tr>
<td>Total SS</td>
<td>0.85</td>
<td>0.42</td>
</tr>
</tbody>
</table>

what do the red highlights mean? what are SS, DF and MS?
Table 5. ANOVA table for failure probability in OR: A indicates the shape parameter in GEV, B indicates the scale parameter of GEV, C indicates the location parameter of GEV, D means the meanlog of LN, E means the sdlog of LN, and F mean the parameters (i.e. theta) in copula

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Zhangjiashan</th>
<th></th>
<th></th>
<th>Xianyang</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SS</td>
<td>DF</td>
<td>MS</td>
<td>F-Value</td>
<td>P-value</td>
<td>SS</td>
<td>DF</td>
</tr>
<tr>
<td>A</td>
<td>2.04</td>
<td>2</td>
<td>1.02</td>
<td>39285.40</td>
<td>&lt; 0.0001</td>
<td>5.71</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>0.20</td>
<td>2</td>
<td>0.098</td>
<td>3784.17</td>
<td>&lt; 0.0001</td>
<td>0.26</td>
<td>2</td>
</tr>
<tr>
<td>C</td>
<td>9.466E-004</td>
<td>2</td>
<td>4.733E-004</td>
<td>18.22</td>
<td>&lt; 0.0001</td>
<td>1.811E-003</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>0.24</td>
<td>2</td>
<td>0.12</td>
<td>4679.22</td>
<td>&lt; 0.0001</td>
<td>0.30</td>
<td>2</td>
</tr>
<tr>
<td>E</td>
<td>0.60</td>
<td>2</td>
<td>0.30</td>
<td>11626.79</td>
<td>&lt; 0.0001</td>
<td>0.87</td>
<td>2</td>
</tr>
<tr>
<td>F</td>
<td>7.833E-004</td>
<td>2</td>
<td>3.916E-004</td>
<td>15.08</td>
<td>&lt; 0.0001</td>
<td>6.382E-004</td>
<td>2</td>
</tr>
<tr>
<td>AB</td>
<td>0.17</td>
<td>4</td>
<td>0.043</td>
<td>1666.34</td>
<td>&lt; 0.0001</td>
<td>0.27</td>
<td>4</td>
</tr>
<tr>
<td>AD</td>
<td>8.076E-004</td>
<td>4</td>
<td>2.019E-004</td>
<td>7.77</td>
<td>&lt; 0.0001</td>
<td>1.830E-003</td>
<td>4</td>
</tr>
<tr>
<td>AE</td>
<td>0.048</td>
<td>4</td>
<td>0.012</td>
<td>465.73</td>
<td>&lt; 0.0001</td>
<td>0.081</td>
<td>4</td>
</tr>
<tr>
<td>AF</td>
<td>0.15</td>
<td>4</td>
<td>0.037</td>
<td>1414.17</td>
<td>&lt; 0.0001</td>
<td>0.29</td>
<td>4</td>
</tr>
<tr>
<td>BC</td>
<td>3.442E-004</td>
<td>4</td>
<td>8.604E-005</td>
<td>3.31</td>
<td>0.0106</td>
<td>3.658E-004</td>
<td>4</td>
</tr>
<tr>
<td>BD</td>
<td>8.013E-005</td>
<td>4</td>
<td>2.003E-005</td>
<td>0.77</td>
<td>0.5442</td>
<td>1.226E-004</td>
<td>4</td>
</tr>
<tr>
<td>BE</td>
<td>2.528E-003</td>
<td>4</td>
<td>6.319E-004</td>
<td>24.33</td>
<td>&lt; 0.0001</td>
<td>2.534E-003</td>
<td>4</td>
</tr>
<tr>
<td>BF</td>
<td>7.212E-003</td>
<td>4</td>
<td>1.803E-003</td>
<td>69.40</td>
<td>&lt; 0.0001</td>
<td>8.837E-003</td>
<td>4</td>
</tr>
<tr>
<td>CF</td>
<td>5.294E-006</td>
<td>4</td>
<td>1.323E-006</td>
<td>0.051</td>
<td>0.9951</td>
<td>1.077E-006</td>
<td>4</td>
</tr>
<tr>
<td>DE</td>
<td>1.192E-005</td>
<td>4</td>
<td>2.981E-006</td>
<td>0.11</td>
<td>0.9773</td>
<td>1.697E-005</td>
<td>4</td>
</tr>
<tr>
<td>CE</td>
<td>3.353E-005</td>
<td>4</td>
<td>8.382E-006</td>
<td>0.32</td>
<td>0.8628</td>
<td>5.780E-005</td>
<td>4</td>
</tr>
<tr>
<td>EF</td>
<td>2.416E-008</td>
<td>4</td>
<td>6.040E-009</td>
<td>2.325E-004</td>
<td>1.0000</td>
<td>7.119E-009</td>
<td>4</td>
</tr>
<tr>
<td>DF</td>
<td>4.11</td>
<td>4</td>
<td>0.028</td>
<td>1069.79</td>
<td>&lt; 0.0001</td>
<td>0.17</td>
<td>4</td>
</tr>
<tr>
<td>DF</td>
<td>7.482E-005</td>
<td>4</td>
<td>1.870E-005</td>
<td>0.72</td>
<td>0.5784</td>
<td>9.919E-005</td>
<td>4</td>
</tr>
<tr>
<td>EF</td>
<td>2.568E-004</td>
<td>4</td>
<td>6.420E-005</td>
<td>2.47</td>
<td>0.0435</td>
<td>3.550E-004</td>
<td>4</td>
</tr>
<tr>
<td>Error</td>
<td>0.017</td>
<td>656</td>
<td>2.598E-005</td>
<td>0.040</td>
<td>656</td>
<td>6.071E-005</td>
<td></td>
</tr>
<tr>
<td>Total SS</td>
<td>3.60</td>
<td>728</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 6. ANOVA table for failure probability in Kendall: A indicates the shape parameter in GEV, B indicates the scale parameter of GEV, C indicates the location parameter of GEV, D means the meanlog of LN, E means the sdlog of LN, and F mean the parameters (i.e. theta) in copula parameters (i.e. theta) in copula

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Zhangjiashan</th>
<th></th>
<th>Xianyang</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SS</td>
<td>DF</td>
<td>MS</td>
<td>F-Value</td>
<td>P-value</td>
</tr>
<tr>
<td>A</td>
<td>0.97</td>
<td>2</td>
<td>0.48</td>
<td>33813.15</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>B</td>
<td>0.096</td>
<td>2</td>
<td>0.048</td>
<td>3349.45</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>C</td>
<td>4.627E-004</td>
<td>2</td>
<td>2.313E-004</td>
<td>16.19</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>D</td>
<td>0.11</td>
<td>2</td>
<td>0.057</td>
<td>3987.58</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>E</td>
<td>0.28</td>
<td>2</td>
<td>0.14</td>
<td>9809.61</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>F</td>
<td>0.013</td>
<td>2</td>
<td>6.451E-003</td>
<td>451.42</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>AB</td>
<td>0.087</td>
<td>4</td>
<td>0.022</td>
<td>1525.14</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>AC</td>
<td>0.090E-004</td>
<td>4</td>
<td>1.022E-004</td>
<td>7.15</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>AD</td>
<td>0.022</td>
<td>4</td>
<td>5.448E-003</td>
<td>381.22</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>AE</td>
<td>0.066</td>
<td>4</td>
<td>0.017</td>
<td>1156.06</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>AF</td>
<td>2.163E-003</td>
<td>4</td>
<td>5.407E-004</td>
<td>37.84</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>BC</td>
<td>4.233E-005</td>
<td>4</td>
<td>1.058E-005</td>
<td>0.74</td>
<td>0.5645</td>
</tr>
<tr>
<td>BD</td>
<td>1.147E-003</td>
<td>4</td>
<td>2.866E-004</td>
<td>20.06</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>BE</td>
<td>3.254E-003</td>
<td>4</td>
<td>8.135E-004</td>
<td>56.93</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>BF</td>
<td>2.479E-004</td>
<td>4</td>
<td>6.198E-005</td>
<td>4.34</td>
<td>0.0018</td>
</tr>
<tr>
<td>CD</td>
<td>5.400E-006</td>
<td>4</td>
<td>1.352E-006</td>
<td>0.994</td>
<td>0.3942</td>
</tr>
<tr>
<td>CE</td>
<td>1.513E-005</td>
<td>4</td>
<td>3.782E-006</td>
<td>0.26</td>
<td>0.9907</td>
</tr>
<tr>
<td>CF</td>
<td>1.190E-006</td>
<td>4</td>
<td>2.974E-007</td>
<td>0.021</td>
<td>0.9992</td>
</tr>
<tr>
<td>DE</td>
<td>0.054</td>
<td>4</td>
<td>0.014</td>
<td>950.06</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>DF</td>
<td>1.870E-004</td>
<td>4</td>
<td>4.676E-005</td>
<td>3.27</td>
<td>0.0114</td>
</tr>
<tr>
<td>EF</td>
<td>4.113E-004</td>
<td>4</td>
<td>1.028E-004</td>
<td>7.20</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Error</td>
<td>9.374E-003</td>
<td>656</td>
<td>1.429E-005</td>
<td>0.025</td>
<td>656</td>
</tr>
</tbody>
</table>
Table 7. Contributions of parameter uncertainties obtained by three level ANOVA to predictive failure probabilities for a design return period of 200-year and service time of 30-year

<table>
<thead>
<tr>
<th>Factor</th>
<th>FPand</th>
<th>FPor</th>
<th>FPkendall</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>43.53%</td>
<td>56.67%</td>
<td>56.40%</td>
</tr>
<tr>
<td>B</td>
<td>2.12%</td>
<td>5.56%</td>
<td>5.58%</td>
</tr>
<tr>
<td>C</td>
<td>0.01%</td>
<td>0.03%</td>
<td>0.03%</td>
</tr>
<tr>
<td>D</td>
<td>6.94%</td>
<td>6.67%</td>
<td>6.40%</td>
</tr>
<tr>
<td>E</td>
<td>21.18%</td>
<td>16.67%</td>
<td>16.28%</td>
</tr>
<tr>
<td>F</td>
<td>0.11%</td>
<td>0.02%</td>
<td>0.76%</td>
</tr>
<tr>
<td>Interaction</td>
<td>26.12%</td>
<td>4.72%</td>
<td>5.06%</td>
</tr>
</tbody>
</table>
Figure 1. Framework of the proposed FSFC approach.

Flood variable 1

Marginal distribution

\[ F_1 = F(x, \theta_1) \]

Flood variable 2

Marginal distribution

\[ F_2 = F(y, \theta_2) \]

Joint distribution estimation

\[ F_{XY}(x, y) = C(F(x, \theta_1), (y, \theta_2), \theta_c) \]

Parameter uncertainty quantification through Bootstrap

\[ (\theta_1, \theta_2, \theta_c)^1 \sim [(X_1, Y_1)^1, (X_2, Y_2)^1, \ldots, (X_n, Y_n)^1] \]

\[ (\theta_1, \theta_2, \theta_c)^2 \sim [(X_1, Y_1)^2, (X_2, Y_2)^2, \ldots, (X_n, Y_n)^2] \]

\[ \ldots \]

\[ (\theta_1, \theta_2, \theta_c)^B \sim [(X_1, Y_1)^B, (X_2, Y_2)^B, \ldots, (X_n, Y_n)^B] \]

Multi-level factorial analysis

Full-subsampling ANOVA

Main and interactive effects of parameters

Uncertainty sources identification

Contributions of parameter uncertainties to risk inferences
Figure 2. The location of the studied watersheds. Wei River is the largest tributary of Yellow river, with a drainage area of 135,000 km². The historical flood data from Xianyang and Zhangjiashan stations on the Wei River are analyzed through the proposed FSFC approach.
Figure 3. Probabilistic features for parameters in marginal distributions and copula: for both Xianyang and Zhangjiashan stations, the GEV (parameters include shape, scale and location) function would be employed to quantify the distribution of flood peak, while the lognormal distribution (parameters denoted as meanlog and sdlog) is applied for flood volume. The Gumbel and Joe copula (parameter denoted as theta) would be respectively adopted to model the dependence between flood peak and volume at Zhangjiashan and Xianyang stations.
Figure 4. Uncertainty quantification of the joint RP in “AND”: the red dash lines indicate the predictive means, the two blue dash lines respectively indicate the 5% and 95% quantiles, and the grey lines indicate the predictions under different parameter samples with the same joint RP of the red and blue dash lines; The cyan lines denote the predictions under different return periods with the model parameters being their mean values.
Figure 5. Uncertainty quantification of the joint RP in “OR”: the red dash lines indicate the predictive means, the two blue dash lines respectively indicate the 5% and 95% quantiles, and the grey lines indicate the predictions under different parameter samples with the same joint RP of the red and blue dash lines; The cyan lines denote the predictions under different return periods with the model parameters being their mean values.
Figure 6. Uncertainty quantification of the joint RP in "Kendall": the red dash lines indicate the predictive means, the two blue dash lines respectively indicate the 5% and 95% quantiles, and the grey lines indicate the predictions under different parameter samples with the same joint RP of the red and blue dash lines; The cyan lines denote the predictions under different return periods with the model parameters being their mean values.
Figure 7. Main effects plot and full interactions plot matrix for parameters on the failure probability in AND at the two gauge stations.

In figures 7 through 9, please add some text in the caption to explain what the figures are showing the reader. In the images of 'interactive effect', why is the labelling put on the diagonal? are the labels describing vertical columns or horizontal rows? what are the green, red and black plots telling us? This layout poses an unnecessary puzzle for the reader.
Figure 8. Main effects plot and full interactions plot matrix for parameters on the failure probability in OR at the two gauge stations.
Figure 9. Main effects plot and full interactions plot matrix for parameters on the failure probability in Kendall at the two gauge stations.
Figure 10. Contributions of parameter uncertainties to predictive failure probabilities in AND under different design standards (i.e. return periods (RP)) and different service periods.
Table 1. Contributions of parameter uncertainties to predictive failure probabilities in OR under different design standards (i.e. return periods (RP)) and different service periods

<table>
<thead>
<tr>
<th></th>
<th>Zhangjiashan</th>
<th></th>
<th>Xianyang</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30 year</td>
<td>50 year</td>
<td>70 year</td>
<td>30 year</td>
<td>50 year</td>
<td>70 year</td>
</tr>
<tr>
<td>GEV_shape</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEV_scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEV_location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LN_meanlog</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LN_sdlog</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cop_theta</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Par_interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Legend

- GEV_shape: 45.60% 45.64% 45.52%
- GEV_scale: 6.94% 6.57% 6.23%
- GEV_location: 0.07% 0.07% 0.06%
- LN_meanlog: 12.60% 12.55% 12.50%
- LN_sdlog: 24.48% 24.75% 24.98%
- Cop_theta: 0.06% 0.06% 0.06%
- Par_interaction: 10.25% 10.37% 10.65%

Figure 11. Contributions of parameter uncertainties to predictive failure probabilities in OR under different design standards (i.e. return periods (RP)) and different service periods.
Figure 12. Contributions of parameter uncertainties to predictive failure probabilities in Kendall under different design standards (i.e. return periods (RP)) and different service periods.
Figure 13. Comparison of parameter contributions to predictive uncertainty for failure probabilities under different levels of subsampling for Zhangjiashan station: three (i.e. 0.1, 0.5, 0.9) and four (i.e. 0.1, 0.35, 0.6, 0.85) level quantiles are adopted for subsampling and the design return period is 200 years.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>3-Level</th>
<th>4-Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEV_shape</td>
<td>45.83%</td>
<td>51.52%</td>
</tr>
<tr>
<td>GEV_scale</td>
<td>4.93%</td>
<td>6.09%</td>
</tr>
<tr>
<td>GEV_location</td>
<td>0.05%</td>
<td>0.06%</td>
</tr>
<tr>
<td>LN_meanlog</td>
<td>10.22%</td>
<td>7.95%</td>
</tr>
<tr>
<td>LN_sdlog</td>
<td>23.61%</td>
<td>19.96%</td>
</tr>
<tr>
<td>Cop_theta</td>
<td>0.23%</td>
<td>0.27%</td>
</tr>
<tr>
<td>Par_interaction</td>
<td>15.14%</td>
<td>14.14%</td>
</tr>
</tbody>
</table>

GFp(kendall)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>3-Level</th>
<th>4-Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEV_shape</td>
<td>45.60%</td>
<td>39.28%</td>
</tr>
<tr>
<td>GEV_scale</td>
<td>6.94%</td>
<td>6.86%</td>
</tr>
<tr>
<td>GEV_location</td>
<td>0.07%</td>
<td>0.07%</td>
</tr>
<tr>
<td>LN_meanlog</td>
<td>12.60%</td>
<td>14.25%</td>
</tr>
<tr>
<td>LN_sdlog</td>
<td>24.48%</td>
<td>29.82%</td>
</tr>
<tr>
<td>Cop_theta</td>
<td>0.06%</td>
<td>0.08%</td>
</tr>
<tr>
<td>Par_interaction</td>
<td>10.25%</td>
<td>9.62%</td>
</tr>
</tbody>
</table>

GFp(For)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>3-Level</th>
<th>4-Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEV_shape</td>
<td>45.83%</td>
<td>38.87%</td>
</tr>
<tr>
<td>GEV_scale</td>
<td>4.93%</td>
<td>6.89%</td>
</tr>
<tr>
<td>GEV_location</td>
<td>0.05%</td>
<td>0.06%</td>
</tr>
<tr>
<td>LN_meanlog</td>
<td>10.22%</td>
<td>7.95%</td>
</tr>
<tr>
<td>LN_sdlog</td>
<td>23.61%</td>
<td>19.96%</td>
</tr>
<tr>
<td>Cop_theta</td>
<td>0.23%</td>
<td>0.27%</td>
</tr>
<tr>
<td>Par_interaction</td>
<td>15.14%</td>
<td>14.14%</td>
</tr>
</tbody>
</table>

GFp(For and kendall)

Legend

% 1% 10% 20% 30% 40% 50%
30 year 50 year 70 year
30 year 50 year 70 year
30 year 50 year 70 year

38.87% 38.94% 39.00%
6.89% 6.78% 6.68%
0.07% 0.07% 0.07%
14.13% 14.10% 14.06%
29.45% 29.55% 29.64%
0.92% 0.90% 0.88%
9.68% 9.66% 9.66%

39.28% 39.35% 39.38%
6.86% 6.70% 6.54%
0.07% 0.07% 0.07%
14.25% 14.19% 14.13%
29.82% 29.95% 30.05%
0.08% 0.08% 0.08%
9.62% 9.66% 9.75%

51.52% 51.93% 52.32%
6.09% 6.11% 6.13%
0.06% 0.06% 0.06%
7.95% 7.92% 7.89%
19.96% 19.92% 19.86%
0.27% 0.27% 0.27%
14.14% 13.80% 13.47%

51.60% 51.64% 51.52%
6.94% 6.57% 6.23%
0.07% 0.07% 0.06%
12.60% 12.55% 12.50%
24.48% 24.75% 24.98%
0.06% 0.06% 0.06%
10.25% 10.37% 10.65%

51.83% 52.37% 52.88%
4.93% 4.96% 5.00%
0.05% 0.05% 0.05%
10.21% 10.19% 10.19%
23.67% 23.70% 23.70%
0.23% 0.23% 0.22%
14.52% 13.96% 13.96%

51.60% 51.64% 51.52%
6.94% 6.57% 6.23%
0.07% 0.07% 0.06%
12.60% 12.55% 12.50%
24.48% 24.75% 24.98%
0.06% 0.06% 0.06%
10.25% 10.37% 10.65%

51.60% 51.64% 51.52%
6.94% 6.57% 6.23%
0.07% 0.07% 0.06%
12.60% 12.55% 12.50%
24.48% 24.75% 24.98%
0.06% 0.06% 0.06%
10.25% 10.37% 10.65%

51.83% 52.37% 52.88%
4.93% 4.96% 5.00%
0.05% 0.05% 0.05%
10.21% 10.19% 10.19%
23.67% 23.70% 23.70%
0.23% 0.23% 0.22%
14.52% 13.96% 13.96%
Figure 14. Variation of parameters’ contributions for different risk inferences at the Zhangjiashan Station for a design standard of 200-year and a service time of 30-year.
Figure 15. Correlation for parameters’ contributions on risk inferences at Zhangjiashan station for a design standard of 200-year and a service time of 30-year: The cross sign indicates the correlation is statistically insignificant.

(a) FPand
(b) FPor
(c) FPkendall

what are the units of this scale?

the lettering is very small and almost illegible. please enlarge the font and print in black