Multivariate hydrologic design methods under nonstationary conditions and application to engineering practice

Cong Jiang1, Lihua Xiong2, Lei Yan3, Jianfan Dong4, and Chong-Yu Xu2,5

1School of Environmental Studies, China University of Geosciences, Wuhan 430074, China
2State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China
3College of Water Conservancy and Hydropower, Hebei University of Engineering, Handan 056002, China
4Guangxi Water Resources Management Center, Nanning 530023, China
5Department of Geosciences, University of Oslo, P.O. Box 1047 Blindern, 0316 Oslo, Norway

Correspondence: Cong Jiang (jiangcong@cug.edu.cn)

Received: 27 May 2018 – Discussion started: 12 July 2018
Revised: 14 February 2019 – Accepted: 9 March 2019 – Published: 22 March 2019

Abstract. Multivariate hydrologic design under stationary conditions is traditionally performed through the use of the design criterion of the return period, which is theoretically equal to the average inter-arrival time of flood events divided by the exceedance probability of the design flood event. Under nonstationary conditions, the exceedance probability of a given multivariate flood event varies over time. This suggests that the traditional return-period concept cannot apply to engineering practice under nonstationary conditions, since by such a definition, a given multivariate flood event would correspond to a time-varying return period. In this paper, average annual reliability (AAR) was employed as the criterion for multivariate design rather than the return period to ensure that a given multivariate flood event corresponded to a unique design level under nonstationary conditions. The multivariate hydrologic design conditioned on the given AAR was estimated from the nonstationary multivariate flood distribution constructed by a dynamic C-vine copula, allowing for time-varying marginal distributions and a time-varying dependence structure. Both the most-likely design event and confidence interval for the multivariate hydrologic design conditioned on the given AAR were identified to provide supporting information for designers. The multivariate flood series from the Xijiang River, China, were chosen as a case study. The results indicated that both the marginal distributions and dependence structure of the multivariate flood series were found to affect the outcome of the multivariate hydrologic design.

1 Introduction

A complete flood event or a flood hydrograph contains multiple feature variables, such as flood peak and flood volume, which can be associated with the safety of hydraulic structures (Salvadori et al., 2004, 2007, 2011; Xiao et al., 2009; Xiong et al., 2015; Loveridge et al., 2017; Shafaei et al., 2017). For example, the water level of a reservoir is controlled not only by flood peak flow but also by flood volume (Salvadori et al., 2011). Therefore, multivariate hydrologic design, which takes into account multiple flood characteristics as well as their dependence, provides a more rational design strategy for hydraulic structures compared to univariate hydrologic design (Zheng et al., 2013, 2014; Balistrocchi and Bacchi, 2017).

Multivariate hydrologic design under stationary conditions has been widely investigated, and the design criterion is usually quantified by the return period, similar to univariate hydrologic design. Under the definition of the average recurrence interval between flood events equaling or exceeding a given threshold (Chow, 1964), the return period of a given flood event under stationary conditions theoretically equals the average inter-arrival time between flood events divided by the exceedance probability (Salvadori et al., 2011). On the other hand, the exceedance probability of a univari-
ate flood event is usually uniquely defined without ambiguity, whereas the exceedance probability of a multivariate flood event could have multiple definitions (Salvadori and De Michele, 2004; Salvadori et al., 2011; Vandenbergh et al., 2011). To date, at least five kinds of different exceedance probabilities for a multivariate flood event have been defined: (1) the OR case in which at least one of the flood features exceeds the prescribed threshold, (2) the AND case in which all flood features exceed the prescribed thresholds, (3) the Kendall case in which the univariate representation transformed from Kendall's distribution function exceeds the prescribed threshold, (4) the survival Kendall case in which the univariate representation transformed from survival Kendall’s distribution function exceeds the prescribed threshold, and (5) the structural case in which the univariate representation transformed from a structure function exceeds the prescribed threshold (Favre et al., 2004; Salvadori and De Michele, 2004, 2010; Salvadori et al., 2007, 2013, 2015, 2016; Vandenbergh et al., 2011; Requena et al., 2013; Zheng et al., 2014).

Due to climate change as well as certain anthropogenic driving forces (Milly et al., 2008), the nonstationarities of both univariate and multivariate flood series have been widely reported (Xiong and Guo, 2004; Villarini et al., 2009; Vogel et al., 2011; López and Francés, 2013; Bender et al., 2014; Xiong et al., 2015; Blöschl et al., 2017; Kundzewicz et al., 2018). The multivariate flood distribution exhibits more complex nonstationarity behaviours than the univariate distribution, including nonstationarities of individual margins and the dependence structure between the margins (Quessy et al., 2013; Bender et al., 2014; Xiong et al., 2015; Kwon et al., 2016; Sarhadi et al., 2016; Qi and Liu, 2017; Vezzoli et al., 2017; Bracken et al., 2018; Salvadori et al., 2018). Both nonstationarities of the margins and dependence structure could impact the multivariate hydrodynamic design. Under nonstationary conditions, the exceedance probability $p$ of a given flood event varies from year to year; thus, the return period, calculated as the average inter-arrival time between two successive flood events divided by $p$, is no longer a constant (Salas and Obeysekera, 2014; Jiang et al., 2015a; Kwon et al., 2016; Sarhadi et al., 2016; Yan et al., 2017). As a result, a given flood event would correspond to a time-varying and non-unique return period. Consequently, the traditional return-period-based method for estimating hydrologic design may no longer be applicable to engineering practice under nonstationary conditions (Salas and Obeysekera, 2014).

Although increasing attention has been focused on the hydrologic designs under nonstationary conditions in recent years, the focus has mainly been on univariate designs (Obeysekera and Salas, 2014, 2016; Read and Vogel, 2016). To overcome the limitation of the traditional return period under nonstationary conditions, the concept of the return period has been revisited. Salas and Obeysekera (2014) extended two concepts of the return period into a nonstationary framework, defined as the expected waiting time (EWT) for an exceedance to occur (Olsen et al., 1998), and the time period that results in the expected number of exceedances (ENE) equal to 1 over this period (Parey et al., 2010).

Risk and reliability are both important measurements for assessing hydrologic designs (Vogel, 1987; Read and Vogel, 2015). Besides redefinitions of the return period, some risk-based or reliability-based metrics have been introduced as the hydrologic design criteria under nonstationary conditions (Rosner et al., 2014). Rootzén and Katz (2013) proposed the concept of the design life level (DLL) to quantify hydrologic risk in a nonstationary climate during the entire design life period of hydraulic structures. Read and Vogel (2015) introduced the concept of average annual reliability (AAR) to estimate the hydrologic designs under nonstationary conditions. Liang et al. (2016) defined the equivalent reliability (ER) to estimate the design flood under nonstationary conditions by linking the DLL to the return period. Salvadori et al. (2018) associated hydrologic designs with both given life times and failure probabilities to calculate bivariate design values under nonstationarity. These design criteria assess the risk or reliability of hydraulic structures associated with the flood distribution during the entire design life period, rather than for a single year. For a given design life period, these criteria can always yield a unique risk or reliability; therefore, they are applicable to the hydrologic designs under both stationary and nonstationary conditions.

Under the multivariate framework, a given design level would correspond to an infinite number of possible hydrologic design events (Hawkes, 2008; Kew et al., 2013; Zheng et al., 2015, 2017); however, these design events are generally not equivalent because their joint probability density values (i.e. likelihood) usually differ (Salvadori et al., 2011; Volpi and Fiori, 2012; Li et al., 2017; Yin et al., 2017). In engineering practice, it should be necessary to determine a typical design event as representative of a specific design level. For example, in Chinese engineering practice, a unique design flood hydrograph corresponding to a given design level is usually required to determine the scale of hydraulic structures (Yin et al., 2017). The most-likely design event, which theoretically has the largest joint probability density (likelihood) among all possible design events (Salvadori et al., 2011), appears to be the best representative candidate. Besides the most-likely design event, it is also necessary to identify the confidence interval for an infinite possible number of design events to provide a finite design range for designers (Volpi and Fiori, 2012; Yin et al., 2017). The most-likely design event and confidence interval for the bivariate hydrologic design under stationary conditions have been identified (Salvadori et al., 2011; Volpi and Fiori, 2012; Li et al., 2017; Yin et al., 2017; Salvadori et al., 2018); however, very few studies have focused on hydrologic designs with higher dimensions under nonstationary conditions.

Therefore, the objective of the present study was to address the issue of multivariate hydrologic design applying to engineering practice under nonstationary conditions, which
is achieved through the following steps. First, the nonstationary multivariate flood distribution was constructed using a dynamic canonical vine (C-vine) copula (Aas et al., 2009), which was able to capture the nonstationarities of both marginal distributions and the dependence structure. The design criterion for the multivariate flood event was then quantified according to AAR rather than the traditional return period, since a given multivariate flood event would correspond to a unique AAR under both stationary and nonstationary conditions (Read and Vogel, 2015; Yan et al., 2017). The multivariate hydrologic design for any given AAR was estimated from the nonstationary multivariate flood distribution.

The aforementioned methods for the multivariate hydrologic design under nonstationary conditions were applied to the Xijiang River, China. The four-dimensional (4-D) multivariate flood series, including the annual maximum daily discharge, annual maximum 3-day flood volume, annual maximum 7-day flood volume and annual maximum 15-day flood volume of the Xijiang River were chosen as the case study data because they constitute the variables used for deriving the design flood hydrograph for hydraulic structures. It has been found that the natural flood processes of this river have been significantly altered by urbanization and reservoir regulation (Xu et al., 2014), but these two factors have not yet been taken into account in multivariate hydrologic design.

The next section of the present paper describes the study area and data. Section 3 presents the methods developed in this paper. The results of the case study are provided in Sect. 4. Finally, the conclusion and remarks are provided in Sect. 5.

2 Study area and data

The multivariate flood series of the Xijiang River, South China (see Fig. 1), were selected as a case study to illustrate the multivariate hydrologic design methods under nonstationary conditions. The drainage area of the Xijiang River basin (XRB) is 353 120 km², with a river length of 2214 km. The basin falls within a humid subtropical monsoon climate region, with the flood season extending from May to October; therefore, floods have always been a serious natural hazard within the basin.

The calculation of design floods in China involving the derivation of flood hydrographs for hydraulic structures requires not only flood peak but also flood volumes with different durations, such as 3, 7, 15 and 30 days (Ministry of Water Resources of People’s Republic of China, 1996; Xiao et al., 2009; Xiong et al., 2015; Li et al., 2017). For a large catchment such as the XRB, the duration of a flood process is usually longer than 10 days. Therefore, the annual maximum daily discharge ($Q_1$), annual maximum 3-day flood volume ($V_3$), annual maximum 7-day flood volume ($V_7$) and annual maximum 15-day flood volume ($V_{15}$) of the Xijiang River were defined as the multivariate flood series ($Q_1, V_3, V_7, V_{15}$) for deriving design flood hydrographs. The flood data were from 1951 to 2012 and observed at the Dahuangjiangkou gauge located at the main stream of the Xijiang River, draining a total catchment area of 294669 km², approximately 83% of the total area of the XRB.

Rapid urbanization over recent decades has resulted in increasing river regulation projects built in the XRB, such as artificial levees for protecting urban areas from river flooding. As a result, flood flow has become increasingly constrained to the channel rather than overflowing to the floodplain, resulting in an increase in the observed river flood flow (Xu et al., 2014). For the purpose of flood control and hydropower generation, it is hard to find a river which is not impacted by reservoirs, particularly in rapidly developing China. Reservoir regulation has become an increasingly significant factor affecting flood processes of the XRB and should be seriously considered within downstream flood risk analysis and hydrologic design, particularly after 2006, when two reservoirs with considerable flood control capacities were put into operation. These are the Longtan and Baise reservoirs, with flood control capacities of $5 \times 10^6$ m³ and $1.64 \times 10^6$ m³ and catchment areas of 98500 and 9600 km², respectively. Climate change will likely result in flood nonstationarity by altering climatic conditions of the basin. Climatic conditions dominating flood processes in the XRB, such as extreme precipitation, appear to have been stationary over the past decades (Yang et al., 2010). Therefore, the current study introduced only urbanization and reservoir regulation as the potential driving forces of flood nonstationarity and ignored the effect of climate change.

The effect of urbanization on flood processes was quantified using the urban population (Pop). Given the unavailability of urban population data at the basin scale and the fact that the vast majority of cities in the XRB are distributed in Guangxi province, we used urban population data for Guangxi province to represent those of the XRB. The annual urban population data during the observation period were obtained from the China Compendium of Statistics 1949–2008 (Department of Comprehensive Statistics of National Bureau of Statistics, 2010) and the website of the National Bureau of Statistics of China (http://www.stats.gov.cn/tjsj/ndsj/, last access: 20 March 2019). The present study assumed the design life period for hydraulic structures to be from 2013 to 2100. The urban population over the design life period was estimated based on the predicted growth rate of China’s urban population reported by He (2014). The reservoir index (RI), which depends on the catchment area and flood controlling capacities of reservoirs, was used to quantify the effects of reservoir regulation on flood processes (López and Francés, 2013). As shown in Table 1, two reservoirs with flood control functions have been completed during the observation period from 1951 to 2012, and a further two are planned for operation during the design life period. Figure 2 illustrates the
evolution of the urban population and reservoir index during both the observation and design life periods.

3 Methods

The present study included the following methods: (1) the nonstationary multivariate flood distribution based on a dynamic C-vine copula, allowing for both time-varying marginal distributions and a time-varying dependence structure, and (2) estimation of the multivariate hydrologic design associated with AAR under nonstationary conditions.

To correspond to the case study in this paper, the multivariate flood series consisting of the annual maximum daily discharge ($Q_1$), annual maximum 3-day flood volume ($V_3$), annual maximum 7-day flood volume ($V_7$) and annual maximum 15-day flood volume ($V_{15}$) were chosen to illustrate the multivariate design methods under nonstationary conditions. It is worth noting that the proposed methods can be extended to other multivariate flood series, such as those consisting of flood peak, flood volume and flood duration.

3.1 Probability distribution of the nonstationary multivariate flood series

According to Sklar’s theorem (Sklar, 1959), the probability distribution of the 4-D flood series $(Q_1, V_3, V_7, V_{15})$ at time $t$ measured by years ($t = 1, 2, \ldots, n$, and $n$ is the length of the flood series) can be formulated through a copula $C(\cdot)$ as follows:

$$
F(q_1, t, v_3, t, v_7, t, v_{15}, t | \theta_c, t) = C\left[F_1(q_1, t | \theta_1, t), F_3(v_3, t | \theta_3, t), F_7(v_7, t | \theta_7, t), F_{15}(v_{15}, t | \theta_{15}, t | \theta_c, t)\right],
$$

where $F_1(q_1, t | \theta_1, t), F_3(v_3, t | \theta_3, t), F_7(v_7, t | \theta_7, t)$ and $F_{15}(v_{15}, t | \theta_{15}, t)$ denote the marginal distributions for $Q_1$, $V_3$, $V_7$ and $V_{15}$, respectively; $\theta_1, t, \theta_3, t, \theta_7, t$ and $\theta_{15}, t$ are the marginal probabilities of $Q_1$, $V_3$, $V_7$ and $V_{15}$, respectively; and $\theta_c, t$ stands for the copula param-
The time-varying moment model that expresses the distribution of the multivariate flood series \((Q_1, V_3, V_7, V_{15})\) can be separated into two modules, including the marginal distributions, i.e., \(f_1(q_{1,t}|\theta_{1,t})\), \(f_3(v_{3,t}|\theta_{3,t})\), \(f_7(v_{7,t}|\theta_{7,t})\) and \(f_{15}(v_{15,t}|\theta_{15,t})\), as well as the dependence structure expressed by the copula density function \(c(u_{1,t}, u_{3,t}, u_{7,t}, u_{15,t}|\theta_{c,t})\). Under nonstationary conditions, both the margins and dependence structure of \((Q_1, V_3, V_7, V_{15})\) can vary over time \(t\).

### 3.1.1 Nonstationary marginal distributions based on the time-varying moment model

The time-varying moment model that expresses the distribution parameters or moments as functions of time or some other explanatory variable or variables have been widely employed to capture the nonstationarities of univariate flood series (Strupczewski et al., 2001; Villarini et al., 2009). In this study, the nonstationary marginal distributions of the multivariate flood series \((Q_1, V_3, V_7, V_{15})\) were constructed by the time-varying moment model.

Based on cause–effect analysis, the flood processes of the XRB were found to mainly be impacted by urbanization and reservoir operation. The reservoir index RI and urban population Pop were therefore used as potential covariates for marginal distribution parameters, including the location parameter \(\mu\), scale parameter \(\sigma\) and shape parameter \(\nu\) (if any). In this study, both linear and exponential functions were considered to build the relationships between distribution parameters and covariates (Strupczewski et al., 2001; Vogel et al., 2011; Salas and Obeysekera, 2014; Jiang et al., 2015; Sarhadi et al., 2016; Read and Vogel, 2016; Yan et al., 2017). Taking the location parameter for illustration, the candidate functions of \(\mu\) were generally formulated as follows.

**Linear:**

\[
\mu_t = \alpha_0 + \alpha_1 \text{Pop}_t + \alpha_2 \text{RI}_t.
\]

**Exponential:**

\[
\mu_t = \exp(\alpha_0 + \alpha_1 \text{Pop}_t + \alpha_2 \text{RI}_t).
\]

Here \(\alpha_0, \alpha_1\) and \(\alpha_2\) are model parameters estimated using the maximum likelihood estimate (MLE) method (Strupczewski et al., 2001). As above, the linear expression in Eq. (3) gives an additive model which suggests that the effects of the covariates RI and Pop on \(\mu\) are independent, while the exponential expression defines a multiplicative model which is able to take into account the possible interaction between the covariates RI and Pop. It is important to note that Eq. (3) defines four specific nonstationary models: the first one is the most complex nonstationary model where it is assumed that both RI and Pop are the driving factors of marginal distributions, the second and third models illustrate that the marginal nonstationarity is linked only to RI and Pop, respectively, and the final one represents the simplest and stationary model, which does not contain any covariates.

Five probability distributions widely used in flood frequency analysis, namely Pearson type III (PIII), generalized
3.1.2 Nonstationary dependence structure based on the dynamic C-vine copula

After estimating the marginal distributions, the nonstationary dependence structure of \((Q_1, V_3, V_7, V_{15})\) as formulated by the copula density function \(c(u_{1,t}, u_{3,t}, u_{7,t}, u_{15,t} | \theta_{c,t})\) was constructed. Given that most applied copula functions are for bivariate random variables, \(c(u_{1,t}, u_{3,t}, u_{7,t}, u_{15,t} | \theta_{c,t})\) cannot be directly expressed as a specific copula function. The pair copula method has been proven to be powerful for the construction of the distribution of multivariate random variables through the decomposition of the multivariate probability density into a series of bivariate copulas (Aas et al., 2009; Xiong et al., 2015; Shafaei et al., 2017). Therefore, this study constructed the dependence structure of \((Q_1, V_3, V_7, V_{15})\) using the pair copula method.

Numerous pair copula decomposition forms for a multivariate distribution are available, among which two kinds of decompositions with regular vine structures prevail in practice, namely the canonical vine (C vine) and the drawable vine (D vine; Aas et al., 2009). It is known that flood peak (e.g. \(Q_1\)) is the dominant feature quantifying a flood event as well as being the key factor in hydrologic design (Ministry of Water Resources of People’s Republic of China, 1996). The C vine is more suitable when there is a key variable governing multivariate dependence (Aas et al., 2009). In this case, the C vine was employed to construct the joint distribution of \((Q_1, V_3, V_7, V_{15})\) with \(Q_1\) elected as the key variable. Thus, the density function \(c(u_{1,t}, u_{3,t}, u_{7,t}, u_{15,t} | \theta_{c,t})\) can be decomposed into six bivariate pair copulas as follows:

\[
c(u_{1,t}, u_{3,t}, u_{7,t}, u_{15,t} | \theta_{c,t}) = c_{15}(u_{1,t}, u_{3,t} | \theta_{13,t}) \\
\cdot c_{17}(u_{1,t}, u_{7,t} | \theta_{17,t}) \\
\cdot c_{3711}(F(u_{3,t},u_{1,t}), F(u_{7,t},u_{1,t}) | \theta_{3711}) \\
\cdot c_{3151}(F(u_{3,t},u_{15,t}), F(u_{7,t},u_{15,t}) | \theta_{3151}) \\
\cdot c_{171513}(F(u_{1,t},u_{15,t},u_{3,t}) | \theta_{171513}) \\
\cdot c_{1513}(F(u_{15,t},u_{1,t},u_{3,t}) | \theta_{1513}),
\]

(4)

where \(\theta_{c,t} = (\theta_{13,t}, \theta_{17,t}, \theta_{1513}, \theta_{171513}, \theta_{3711}, \theta_{3151})\) is the parameter vector in the C-vine copula, and

\[
F(u_{3,t},u_{1,t}) = \frac{\partial C_{15}(u_{1,t}, u_{3,t} | \theta_{13,t})}{\partial u_{1,t}},
\]

\[
F(u_{7,t},u_{1,t}) = \frac{\partial C_{17}(u_{1,t}, u_{7,t} | \theta_{17,t})}{\partial u_{1,t}},
\]

\[
F(u_{15,t},u_{1,t}) = \frac{\partial C_{1513}(u_{1,t}, u_{15,t} | \theta_{1513})}{\partial u_{1,t}},
\]

\[
F(u_{7,t},u_{15,t}) = \frac{\partial c_{3711}(F(u_{3,t},u_{1,t}), F(u_{7,t},u_{1,t}) | \theta_{3711})}{\partial F(u_{3,t},u_{1,t})},
\]

\[
F(u_{15,t},u_{1,t}) = \frac{\partial c_{3151}(F(u_{3,t},u_{15,t}), F(u_{15,t},u_{1,t}) | \theta_{3151})}{\partial F(u_{3,t},u_{15,t})}.
\]

Figure 3 shows the schematic decomposition of the 4-D C-vine copula as expressed by Eq. (4). It is evident that the hierarchical structure of the 4-D C-vine copula contains three trees and six edges. The first tree (T1) includes three bivariate pair copulas, i.e. \(c_{13}(\cdot | \theta_{13,t})\), \(c_{17}(\cdot | \theta_{17,t})\) and \(c_{1513}(\cdot | \theta_{1513})\), which directly act on the marginal probabilities and describe the bivariate dependencies between the key variable \(Q_1\) and the other three variables, i.e. \(V_3, V_7\) and \(V_{15}\). The second tree (T2) includes two bivariate pair copulas \(c_{3711}(\cdot | \theta_{3711})\) and \(c_{3151}(\cdot | \theta_{3151})\), which act on the conditional distribution functions with \(u_{1,t}\) as the conditioning variable. Finally, the third tree (T3) includes only one bivariate pair copula \(c_{171513}(\cdot | \theta_{171513})\) acting on conditional distribution functions with both \(u_{1,t}\) and \(u_{3,t}\) as the conditioning variables.

In flood frequency analysis, the upper tail of the flood distribution deserves more attention because it allows the quantification of risks of the more serious flood events. The Gumbel–Hougaard copula, an extreme-value copula widely used in hydrology, accounts for the upper tail dependence and is well-suited to the dependence structure of a multivariate flood distribution (Salvadori et al., 2007; Zhang and Singh, 2007; Xiong et al., 2015). Consequently, the present study employed the bivariate Gumbel–Hougaard copula to construct the dynamic C-vine copula formulated by Eq. (4). The bivariate Gumbel–Hougaard copula is expressed as follows:

\[
C(u, v) = \exp \left\{ - \left[ (\ln u)^{\theta_c} + (\ln v)^{\theta_c} \right]^{1/\theta_c} \right\}, \\
\theta_c \in [1, \infty),
\]

(6)

where \(u\) and \(v\) are the bivariate marginal probabilities and \(\theta_c\) is the single parameter measuring the dependence strength.

Similar to the nonstationary marginal distributions, the nonstationarity of the dependence structure of \((Q_1, V_3, V_7, V_{15})\) was characterized by the time variations of the copula parameters in T1, T2 and T3. Both
linear and exponential functions were considered to characterize the time-varying copula parameters and generally formulated as follows.

Linear: \[ \theta_{c,t} = \beta_0 + \beta_1 \text{Pop}_t + \beta_2 \text{RI}_t. \]

Exponential: \[ \theta_{c,t} = 1 + \exp(\beta_0 + \beta_1 \text{Pop}_t + \beta_2 \text{RI}_t). \] (7)

Here \( \beta_0, \beta_1 \) and \( \beta_2 \) are model parameters estimated using the MLE method (Aas et al., 2009). Here, the exponential expression in Eq. (7) was written as the sum of 1 and an exponential function of the covariates so that the domain range of the copula parameter \( \theta_c \) can be satisfied under any condition. To make it easy in parameter estimation, the model parameters for each pair copula were separately estimated. The model parameters for \( \theta_{13,t}, \theta_{17,t} \) and \( \theta_{115,t} \) in T1 were first estimated, and those for the remaining copula parameters \( \theta_{37|1,t}, \theta_{315|1,t} \) and \( \theta_{715|3,t} \) in T2 and T3 were then estimated in sequence. It is worth noting that these parameters can be also simultaneously estimated. These two methods could result in possible difference in parameter estimation.

The available GoF tests for vine copulas are very limited, with the probability integral transform (PIT) test appearing to be reliable (Aas et al., 2009). Under a null hypothesis of the multivariate flood variables \( (Q_1, V_3, V_7, V_{15}) \) following a given C-vine copula, the PIT converts the dependent flood variables into a new set of variables that are independent and uniformly distributed on \([0, 1]^n\). The GoF of vine copulas can be obtained through determining whether the resulting variables are independent and uniform in \([0, 1]\). For more details of the PIT test, readers are referred to Aas et al. (2009). The best nonstationary model for each bivariate pair copula in Eq. (4) was chosen from the nonstationary models generally expressed by Eq. (7) in terms of the AICc value.

### 3.2 Multivariate hydrologic design under nonstationary conditions

#### 3.2.1 Average annual reliability for multivariate flood events

The AAR introduced by Read and Vogel (2015) was calculated using the arithmetic-average method, thereby taking into account the reliability of each year with the same weighting factor. A safer design strategy should pay more attention to worse (i.e. lower) annual reliability; however, the arithmetic-average AAR is not capable of this function. The present study employed the geometric-average method to calculate AAR, which is dominated more by the minimum than arithmetic average and is theoretically able to yield safer design values. The geometric-average AAR is also equivalent to the metrics of the DLL (Rootzén and Katz, 2013) and ER (Liang et al., 2016; Yan et al., 2017).

Denoting \( (q_1, v_3, v_7, v_{15}) \) as a given multivariate flood event, its exceedance probability \( p_t \), which is the occurrence probability of a more dangerous multivariate event than \( (q_1, v_3, v_7, v_{15}) \) in a specific hazard scenario, would vary from year to year under nonstationary conditions. The AAR for \( (q_1, v_3, v_7, v_{15}) \) was calculated by the geometric-average method as follows:

\[
\text{AAR}(q_1, v_3, v_7, v_{15}) = \left[ \prod_{t=T_1}^{T_2} (1 - p_t) \right]^{-\frac{1}{T_2 - T_1 + 1}},
\] (8)

where \( T_1 \) and \( T_2 \) stand for the beginning year and ending year of the operation of an assumed hydraulic structure, respectively. \( T_2 - T_1 + 1 \) is the length of the design life period of the hydraulic structure, and \( 1 - p_t \) measures the annual reli-
ability of the given multivariate flood event \((q_1, v_3, v_7, v_{15})\) at time \(t\).

### 3.2.2 Exceedance probabilities of multivariate flood events

The present study characterized AAR by considering three widely used definitions of the exceedance probabilities of the multivariate flood event \((q_1, v_3, v_7, v_{15})\), i.e., the OR, AND and Kendall cases (Salvadori and De Michele, 2004, 2010; Favre et al., 2004; Salvadori et al., 2007, 2016; Vandenberghe et al., 2011). The OR case for \((q_1, v_3, v_7, v_{15})\) defines the case under which at least one of the flood features exceeds the prescribed threshold. The exceedance probability in the OR case at time \(t\) was denoted as \(p^{\text{OR}}_t\) and was calculated by the following:

\[
p^{\text{OR}}_t = P(\{Q_1 > q_1 \lor V_3 > v_3 \lor V_7 > v_7 \lor V_{15} > v_{15}\}) = 1 - F(q_1, v_3, v_7, v_{15}|\theta_t),
\]

where “\(\lor\)” stands for the OR operator and \(F(\cdot|\theta_t)\) is defined in Eq. (1).

The AND case for \((q_1, v_3, v_7, v_{15})\) defines the case under which all of the flood features exceed the prescribed thresholds, and the corresponding exceedance probability \(p^{\text{AND}}_t\) at time \(t\) was expressed as follows:

\[
p^{\text{AND}}_t = P(\{Q_1 > q_1 \land V_3 > v_3 \land V_7 > v_7 \land V_{15} > v_{15}\}) = \int \int \int \int \Omega^{\text{AND}} \cdot f(Q_1, v_3, v_7, v_{15}|\theta_t) \cdot dQ_1, dV_3, dV_7, dV_{15},
\]

where “\(\land\)” is the AND operator and \(f(\cdot|\theta_t)\) is defined in Eq. (2).

Under the Kendall case, the multivariate flood event \((q_1, v_3, v_7, v_{15})\) was first transformed into a univariate representation via Kendall’s distribution function \(K_C(\cdot)\) as follows:

\[
K_C(\rho_t) = P(C(\{U_1, U_3, U_7, U_{15}\}|\theta_{\rho}) \leq \rho_t) = P(F(Q_1, v_3, v_7, v_{15}|\theta_t) \leq \rho_t),
\]

where \(\rho_t = F(q_1, v_3, v_7, v_{15}|\theta_t)\) is the probability level corresponding to the given flood event \((q_1, v_3, v_7, v_{15})\). The exceedance probability \(p^{\text{KEN}}_t\) in the Kendall case at time \(t\) was expressed as follows:

\[
p^{\text{KEN}}_t = 1 - K_C(\rho_t).
\]

For general multivariate cases, the exceedance probabilities \(p^{\text{OR}}_t, p^{\text{AND}}_t\) and \(p^{\text{KEN}}_t\) could have no analytical solutions but can be numerically estimated through the Monte Carlo method (Niederreiter, 1978; Salvadori et al., 2011, 2013). The AAR in the OR, AND and Kendall cases can be calculated by replacing the exceedance probability \(p_t\) in Eq. (8) by \(p^{\text{OR}}_t, p^{\text{AND}}_t\) and \(p^{\text{KEN}}_t\), respectively.

### 3.2.3 Most-likely design event and confidence interval for multivariate hydrologic design

The methods identifying both the most-likely design event, denoted by \((z^{\text{MILE}}_{Q1}, z^{\text{MILE}}_{V3}, z^{\text{MILE}}_{V7}, z^{\text{MILE}}_{V15})\), and the confidence interval for the multivariate hydrologic design \((z_{Q1}, z_{V3}, z_{V7}, z_{V15})\) given AAR = \(\eta\) are introduced below. The average annual probability density, denoted by \(g(\cdot)\), of \((z_{Q1}, z_{V3}, z_{V7}, z_{V15})\) over the entire design life period from \(T_1\) to \(T_2\), was expressed as follows:

\[
g(z_{Q1}, z_{V3}, z_{V7}, z_{V15}) = \frac{1}{T_2 - T_1} \sum_{t=T_1}^{T_2} f(z_{Q1}, z_{V3}, z_{V7}, z_{V15}|\theta_t).
\]

The probability distribution function for AAR \(\leq \eta\) can be written as:

\[
\Phi(\eta) = \int \int \int \int \Omega^{\text{AAR}} g(q_1, v_3, v_7, v_{15})
\]

\[
dq_1dv_3dv_7dv_{15}.
\]

By denoting the density function of \(\Phi(\eta)\) as \(\phi(\eta)\), the probability density of \((z_{Q1}, z_{V3}, z_{V7}, z_{V15})\) conditioned on AAR = \(\eta\) can be expressed as follows:

\[
g(z_{Q1}, z_{V3}, z_{V7}, z_{V15}) = \frac{\Phi(\eta)}{\phi(\eta)}.
\]

The most-likely design event conditioned on AAR = \(\eta\) was theoretically written as:

\[
(z^{\text{MILE}}_{Q1}, z^{\text{MILE}}_{V3}, z^{\text{MILE}}_{V7}, z^{\text{MILE}}_{V15}) = \arg \max_{|\text{AAR}=\eta} g(z_{Q1}, z_{V3}, z_{V7}, z_{V15}).
\]

Unfortunately, the analytical solutions of both the most-likely design event \((z^{\text{MILE}}_{Q1}, z^{\text{MILE}}_{V3}, z^{\text{MILE}}_{V7}, z^{\text{MILE}}_{V15})\) and confidence interval are unavailable but can be approximately estimated through the Monte Carlo simulation method. First, the design events with sample size \(N\) conditioned on AAR = \(\eta\) were generated. These design events were then sorted in descending order of their multivariate probability densities, denoted by:

\[
(z_1^{\text{MILE}}_{Q1}, z_1^{\text{MILE}}_{V3}, z_1^{\text{MILE}}_{V7}, z_1^{\text{MILE}}_{V15}), (z_2^{\text{MILE}}_{Q1}, z_2^{\text{MILE}}_{V3}, z_2^{\text{MILE}}_{V7}, z_2^{\text{MILE}}_{V15}), \ldots,
\]

\[
(z_N^{\text{MILE}}_{Q1}, z_N^{\text{MILE}}_{V3}, z_N^{\text{MILE}}_{V7}, z_N^{\text{MILE}}_{V15}), (z_{N-c}^{\text{MILE}}_{Q1}, z_{N-c}^{\text{MILE}}_{V3}, z_{N-c}^{\text{MILE}}_{V7}, z_{N-c}^{\text{MILE}}_{V15}), \ldots,
\]

\[
(z_{N}^{\text{MILE}}_{Q1}, z_{N}^{\text{MILE}}_{V3}, z_{N}^{\text{MILE}}_{V7}, z_{N}^{\text{MILE}}_{V15}).
\]

where \(N_c = N \cdot p_c\) and \(p_c\) is the critical probability level for the confidence interval. Thus, the approximate solution for \((z^{\text{MILE}}_{Q1}, z^{\text{MILE}}_{V3}, z^{\text{MILE}}_{V7}, z^{\text{MILE}}_{V15})\) is \((z_1^{\text{MILE}}_{Q1}, z_1^{\text{MILE}}_{V3}, z_1^{\text{MILE}}_{V7}, z_1^{\text{MILE}}_{V15})\). The lower
boundary for the confidence interval was expressed as follows:

\[
\begin{align*}
L_{Q_i}^t &= \min \left( z_{Q_1}^1, z_{Q_1}^2, \ldots, z_{Q_1}^{N_c} \right) \\
L_{V_3}^t &= \min \left( z_{V_3}^1, z_{V_3}^2, \ldots, z_{V_3}^{N_c} \right) \\
L_{V_7}^t &= \min \left( z_{V_7}^1, z_{V_7}^2, \ldots, z_{V_7}^{N_c} \right) \\
L_{V_{15}}^t &= \min \left( z_{V_{15}}^1, z_{V_{15}}^2, \ldots, z_{V_{15}}^{N_c} \right)
\end{align*}
\]

The upper boundary for the confidence interval was estimated by the following:

\[
\begin{align*}
U_{Q_i}^t &= \max \left( z_{Q_1}^1, z_{Q_1}^2, \ldots, z_{Q_1}^{N_c} \right) \\
U_{V_3}^t &= \max \left( z_{V_3}^1, z_{V_3}^2, \ldots, z_{V_3}^{N_c} \right) \\
U_{V_7}^t &= \max \left( z_{V_7}^1, z_{V_7}^2, \ldots, z_{V_7}^{N_c} \right) \\
U_{V_{15}}^t &= \max \left( z_{V_{15}}^1, z_{V_{15}}^2, \ldots, z_{V_{15}}^{N_c} \right)
\end{align*}
\]

3.2.4 Derivation of design flood hydrographs

In China, the design flood hydrographs for hydraulic structures are usually derived from the design flood events set against a benchmark flood hydrograph, which is chosen from the observed flood processes (Ministry of Water Resources of People’s Republic of China, 1996; Xiao et al., 2009, Yin et al., 2017). For example, suppose that a flood hydrograph consists of the features of annual maximum daily discharge, 3-day flood volume, 7-day flood volume and 15-day flood volume. The four features of the benchmark flood hydrograph are denoted by \(Q_1^B, V_3^B, V_7^B\) and \(V_{15}^B\), respectively. The design flood hydrograph corresponding to the multivariate hydrologic design realization \((z_{Q_1}, z_{V_3}, z_{V_7}, z_{V_{15}})\) can be derived by multiplying the benchmark flood hydrograph by different amplifiers, given as described below.

The amplifier \(K_1\) for the annual maximum daily discharge was calculated by

\[K_1 = \frac{z_{Q_1}}{Q_1^B} \] (20)

The amplifier \(K_{3-1}\) for the 3-day flood volume except for the annual maximum daily discharge was calculated by

\[K_{3-1} = \frac{z_{V_3} - V \left( z_{Q_1} \right)}{V_3^B - V \left( Q_1^B \right)} \] (21)

where \(V (\cdot)\) is the operator transforming daily discharge into flood volume. The amplifier \(K_{7-3}\) for the 7-day flood volume except for the 3-day flood volume was calculated by

\[K_{7-3} = \frac{z_{V_7} - z_{V_3}}{V_7^B - V_3^B} \] (22)

Finally, the amplifier \(K_{15-7}\) for the 15-day flood volume except for the 7-day flood volume was calculated by

\[K_{15-7} = \frac{z_{V_{15}} - z_{V_7}}{V_{15}^B - V_7^B} \] (23)

4 Results

4.1 Nonstationary analysis for marginal distributions

The time-varying moment model was employed to perform nonstationary analysis for each marginal distribution of the multivariate flood series \((Q_1, V_3, V_7, V_{15})\) of the Xijiang River. In general, the candidate distributions for all margins passed the GoF test at the 0.05 significance level. The chosen models featured with the smallest AICc values were shown in Table 2. The results indicated that the GEV distribution provided the best fit for the annual maximum daily discharge series \(Q_1\), whereas the Gamma distribution was chosen as the theoretical distribution for the flood volume series \(V_3, V_7\) and \(V_{15}\). All estimated model parameters were found to be statistically significant at the 0.05 level. The 95% uncertainty intervals for the estimated parameters were calculated by the parametric bootstrap method (Kyselý, 2009). In accordance with the modelling results, it can be seen that the location parameters \(\mu\) for all flood series were nonstationary, while the scale and shape parameters were stationary. Through an exponential function, the location parameters \(\mu\) referring to the means of the flood series were generally positively related to the urban population \(Pop\), whereas they were negatively related to the reservoir index \(RI\). This finding revealed the opposite roles played by urbanization and reservoir regulation on the flood processes of the XRB. In particular, more artificial levees were required to protect urban areas from flooding by constraining the flood flow to river channels, which resulted in increasing the river channel flood flow. The reservoirs played an active role in flood control by reducing the flood discharge downstream.

More specific to each margin of \((Q_1, V_3, V_7, V_{15})\), the location parameters \(\mu\) of the three short-duration flood series, i.e. \(Q_1, V_3, V_7\), were positively linked to Pop, whereas the RI was the driving factor reducing \(\mu\) for all flood series, including \(Q_1, V_3, V_7\) and \(V_{15}\). Owing to the difference in covariate selections, the short-duration flood series, including \(Q_1, V_3, V_7\), displayed asynchronous nonstationary behaviours with the long-duration flood series \(V_{15}\) occurring in the observation period of 1951–2012. As shown in Fig. 4, \(Q_1, V_3\) and \(V_7\) presented significantly increasing trends during 1951–2005, particularly since the 1980s, marking the beginning of a period of rapid urbanization in China. \(V_{15}\) tended to follow a stationary process during 1951–2005. After the two flood control reservoirs were put into operation in 2006, all flood series, including \(Q_1, V_3, V_7\) and \(V_{15}\), exhibited a sharp decline.

The predicted marginal distributions for \((Q_1, V_3, V_7, V_{15})\) during the design life period from 2013 to 2100 were estimated using the time-varying moment model by replacing the observed covariates for \(\mu\) with those predicted. Figure 4 also shows that the mean values of \(Q_1, V_3, V_7, V_{15}\) during the design life period increased with the growth of the urban population, following which they decreased sharply in
Table 2. Results of nonstationary analysis for the marginal distributions of \((Q_1, V_3, V_7, V_{15})\).

<table>
<thead>
<tr>
<th>Flood series</th>
<th>Distribution</th>
<th>(\alpha_0)</th>
<th>(\mu)</th>
<th>(\sigma)</th>
<th>(\nu)</th>
<th>(p_{KS})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Q_1)</td>
<td>GEV</td>
<td>10.050(^c)</td>
<td>0.0212(^c)</td>
<td>-2.166(^b)</td>
<td>6892.085(^c)</td>
<td>-0.271(^b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[9.931, 10.182]</td>
<td>[0.005, 0.036]</td>
<td>[-4.006, -0.481]</td>
<td>[5313.291, 8176.206]</td>
<td>[-0.527, -0.092]</td>
</tr>
<tr>
<td>(V_3)</td>
<td>Gamma</td>
<td>1.866(^c)</td>
<td>0.0185(^b)</td>
<td>-2.094(^b)</td>
<td>0.261(^c)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1.751, 1.977]</td>
<td>[0.002, 0.034]</td>
<td>[-3.801, -0.403]</td>
<td>[0.209, 0.300]</td>
<td></td>
</tr>
<tr>
<td>(V_7)</td>
<td>Gamma</td>
<td>2.638(^c)</td>
<td>0.0119(^b)</td>
<td>-1.934(^b)</td>
<td>0.269(^c)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[2.522, 2.754]</td>
<td>[-0.005, 0.028]</td>
<td>[-3.713, -0.166]</td>
<td>[0.215, 0.308]</td>
<td></td>
</tr>
<tr>
<td>(V_{15})</td>
<td>Gamma</td>
<td>3.258(^c)</td>
<td>-</td>
<td>-1.525(^b)</td>
<td>0.265(^c)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[3.213, 3.354]</td>
<td>-</td>
<td>[-2.807, 0.155]</td>
<td>[0.215, 0.307]</td>
<td></td>
</tr>
</tbody>
</table>

The relationships between \(\mu\) and covariates were built by the exponential function in Eq. (3). \(\alpha_1\) and \(\alpha_2\) are the parameters related to urban population (Pop) and reservoir index (RI), respectively. The symbols \(^a\), \(^b\) and \(^c\) denote that the estimated model parameters are significant at the levels of 0.1, 0.05 and 0.1, respectively. The numbers in brackets are the 95\% uncertainty interval. \(p_{KS}\) stands for the \(p\) value of the KS test for marginal distributions.

![Figure 4](https://example.com/figure4.png)

**Figure 4.** Nonstationary marginal distributions during both the observation and design life periods.

2023 after a larger reservoir named Datengxia is expected to be put into operation. After 2023, with no more reservoirs planned, the predicted mean values of \(Q_1, V_3, V_7\) would be expected to reach their peaks in the mid-21st century followed by a slight declining trend because of a shrinking urban population. Since \(V_{15}\) was only related to RI, \(V_{15}\) would show an abrupt decline in 2023 due to the regulation of the Datengxia reservoir. In general, the predicted nonstationary marginal distributions for \(Q_1\) and \(V_3\) during 2013–2100 were roughly approximate to the marginal distributions under the assumption of stationarity, whereas the predicted nonstationary marginal distributions for \(V_7\) and \(V_{15}\) exhibited smaller mean values than those of the stationary distributions.
Table 3. Results of nonstationary analysis for the dependence structure of \((Q_1, V_3, V_7, V_{15})\).

<table>
<thead>
<tr>
<th>Copula parameter</th>
<th>Model parameters</th>
<th>(\beta_0)</th>
<th>(\beta_1)</th>
<th>(\beta_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\theta_{13})</td>
<td>(3.023^c)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>[2.816, 3.249]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\theta_{17})</td>
<td>(1.719^c)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>[1.483, 1.976]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\theta_{115})</td>
<td>(1.461^c)</td>
<td>(-0.111^b)</td>
<td>(9.426^b)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.958, 2.038]</td>
<td>([-0.021, -0.226])</td>
<td>[0.970, 20.416]</td>
<td></td>
</tr>
<tr>
<td>(\theta_{37})</td>
<td>(0.092^a)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>([-0.316, 0.473])</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\theta_{315})</td>
<td>(-1.444^b)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>([-3.036, -0.693])</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\theta_{115})</td>
<td>(-0.231^a)</td>
<td>([-0.728, 0.199])</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

The relationships between copula parameters and covariates were built by the exponential function in Eq. (7). \(\beta_1\) and \(\beta_2\) are the parameters related to urban population (Pop) and reservoir index (RI), respectively. The symbols \(^a\), \(^b\) and \(^c\) denote that the estimated model parameters are significant at the levels of 0.1, 0.05 and 0.01, respectively. The numbers in brackets are the 95% uncertainty interval.

Figure 5. Statistical correlations between flood peak and flood volumes. Three asterisks (***) indicate the statistical correlation at the 0.01 significance level.

4.2 Nonstationary dependence structure for \((Q_1, V_3, V_7, V_{15})\)

After estimating the nonstationary marginal distributions for \((Q_1, V_3, V_7, V_{15})\), the multivariate dependence structure was constructed by the dynamic C-vine copula with \(Q_1\) elected as the key variable. Figure 5 illustrates significant correlations between the flood peak \(Q_1\) and the flood volumes (i.e. \(V_3\), \(V_7\) and \(V_{15}\)). Table 3 shows the estimation results of the dynamic C-vine copula. The PIT test for the nonstationary dependence structure of \((Q_1, V_3, V_7, V_{15})\) suggested a satisfactory fitting effect, and most estimated parameters were statistically significant at the 0.05 level. The results indicated that the copula parameter \(\theta_{115}\) for pair \((Q_1, V_{15})\) was found to be nonstationary and expressed as an exponential function of both the urban population Pop and reservoir index RI, whereas other copula parameters indicated stationary dependences. It was seen that the margin of \(Q_1\) displayed asynchronous nonstationarity behaviours with \(V_{15}\) (see Table 2 and Fig. 4). Therefore, the dependence nonstationarity of the pair \((Q_1, V_{15})\) could possibly be attributed to the asynchronous marginal nonstationarities.

According to the regression function, \(\theta_{115}\) was negatively related to Pop, whereas it was positively related to RI. In other words, growing urbanization weakened the multivariate flood dependence, whereas reservoir regulation played an opposite role, enhancing the dependence. This finding indicated that human activities, including urbanization and reservoir regulation, not only changed the statistical characteristics of the marginal distributions of \((Q_1, V_3, V_7, V_{15})\) but also affected the dependence of \((Q_1, V_3, V_7, V_{15})\). Figure 6 shows the time variations of \(\theta_{115}\) during the observation pe-
Nonstationary copula parameter for pair \((Q_1, V_{15})\) during both the observation and design life periods.

The multivariate hydrologic designs, characterized by AAR associated with the OR, AND and Kendall exceedance probabilities, were estimated from the predicted nonstationary multivariate distribution for \((Q_1, V_3, V_7, V_{15})\) during the design life period from 2013 to 2100. The left columns in Figs. 7–10 show the most-likely design events and the 90 % confidence intervals conditioned on the AAR varying from 0.01 to 0.99. The multivariate hydrologic design events associated with both the OR and Kendall exceedance probabilities exhibited the lower boundaries, whereas the design events associated with the AND exceedance probability exhibited the upper boundaries.

The design flood hydrographs were derived from the multivariate hydrologic designs against the benchmark flood hydrograph observed in 1988. Figure 11 shows the design flood hydrographs by setting AAR equal to 0.90, 0.95 and 0.99. For any given multivariate flood event, the corresponding OR exceedance probability was larger than that of AND, with the Kendall exceedance probability somewhere in between (Vandenbergh et al., 2011). These differences among the OR, AND and Kendall exceedance probabilities indicate the different design strategies. It must be noted that the choice of design strategy in engineering practice is usually a priori and is dependent on the specific design requirements and mechanisms of failure for hydraulic structures (Serinaldi, 2015; Salvadore et al., 2016).

We calculated the univariate hydrologic design events from the predicted marginal distributions to compare the design strategies under the multivariate framework with those under the univariate framework. Figures 7–10 show that the univariate hydrologic design events exactly constituted the lower boundaries of the multivariate hydrologic design events associated with the OR exceedance probability as well as the upper boundaries of the design events associated with the AND exceedance probability. Under a given AAR, the hydrologic designs under the univariate framework were generally smaller than the most-likely design events associated with the OR exceedance probability, whereas they were larger than those associated with the AND exceedance probability; they were most approximate to those associated with the Kendall exceedance probability. The comparisons of the flood hydrographs displayed in Fig. 11 reinforced these findings.

4.4 Impacts of multivariate nonstationarity behaviours on hydrologic design values

Section 4.1 and 4.2 show the marginal distribution and dependence structure of the multivariate flood distribution of \((Q_1, V_3, V_7, V_{15})\) to be nonstationary. We estimated the multivariate hydrologic design events under an assumption of stationarity to illustrate how these nonstationarities act on the
Table 4. Results of change-point detection for the marginal distributions and dependence of \((Q_1, V_3, V_7, V_{15})\).

<table>
<thead>
<tr>
<th>Flood series</th>
<th>Change point of margin</th>
<th>p value</th>
<th>Flood series</th>
<th>Change point of dependence</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Q_1)</td>
<td>1993</td>
<td>0.072</td>
<td>((Q_1, V_3))</td>
<td>1955</td>
<td>0.083</td>
</tr>
<tr>
<td>(V_3)</td>
<td>1993</td>
<td>0.186</td>
<td>((Q_1, V_7))</td>
<td>1955</td>
<td>0.537</td>
</tr>
<tr>
<td>(V_7)</td>
<td>1994</td>
<td>0.752</td>
<td>((Q_1, V_{15}))</td>
<td>1972</td>
<td>0.599</td>
</tr>
<tr>
<td>(V_{15})</td>
<td>1981</td>
<td>0.423</td>
<td>((Q_1, V_3, V_7, V_{15}))</td>
<td>1972</td>
<td>0.995</td>
</tr>
</tbody>
</table>

Figure 7. Design values of the annual maximum daily discharge for different average annual reliability (AAR) varying from 0.01 to 0.99 under three nonstationary conditions.
multivariate hydrologic designs, i.e. both marginal distributions and the dependence structure were treated as stationary (see the right columns in Figs. 7–10). Figure 4 suggests that both the predicted nonstationary marginal distributions for $Q_1$ and $V_3$ during the design life period were approximate to the stationary marginal distributions. Therefore, the nonstationary and stationary marginal distributions yielded similar design values for $z_{Q_1}$ and $z_{V_3}$ (see Figs. 7 and 8). The predicted nonstationary distributions for both $V_7$ and $V_{15}$ indicated smaller mean values compared to those of the stationary distributions (see Fig. 4); therefore, the corresponding hydrologic design values estimated from the nonstationary marginal distributions were generally smaller than those estimated from the stationary marginal distributions (see Figs. 9 and 10).

The nonstationary multivariate flood distribution during the design life period was also predicted to exhibit a weaker dependence structure than that of the stationary distribution (see Fig. 6). The dependence nonstationarity was expected to have a much subtler effect on the multivariate hydrologic

Figure 8. Design values of the 3-day flood volume for different average annual reliability (AAR) varying from 0.01 to 0.99 under three nonstationary conditions.
design compared to the marginal nonstationarities (Xiong et al., 2015). To illustrate the individual effect of the dependence nonstationarity on the multivariate hydrologic design, an artificial nonstationary condition for the multivariate flood distribution was set so that only the marginal nonstationarities were considered, whereas the dependence structure was treated as stationary. The results of the multivariate hydrologic design events are shown in the middle columns in Figs. 7–10. In general, the dependence nonstationarity had less of an effect on the multivariate hydrologic designs compared the marginal nonstationarities; however, some visible differences in both the 90% confidence intervals were still identified. In summary, the nonstationary and weaker dependence structure generally suggested wider confidence intervals for the multivariate hydrologic design values.

Figure 9. Design values of the 7-day flood volume for different average annual reliability (AAR) varying from 0.01 to 0.99 under three nonstationary conditions.
5 Conclusion and remarks

The statistical characteristics of both the marginal distributions and the dependence structure of multivariate flood variables can vary with time under nonstationary conditions. It is possible that the multivariate flood distribution estimated from the historical information will not reflect the statistical characteristics of flooding in the future. As a result, the stationary-based hydrologic design would not be able to deal with potential hydrologic risks of hydraulic structures. It is necessary for hydrologic designers to take into account the physical driving forces (such as human activities and climate change) responsible for the nonstationarities of multivariate flood variables.

The present study introduced possible methods for addressing multivariate hydrologic design for application in engineering practice under nonstationary conditions. A dynamic C-vine copula allowing for both time-varying marginal distributions and time-varying dependence structure was developed to capture the nonstationarities of a multi-
C. Jiang et al.: Multivariate hydrologic design methods under nonstationary conditions 1699

Figure 11. Design flood hydrographs associated with OR, AND and Kendall probabilities.

The multivariate hydrologic design under nonstationary conditions was estimated by specifying the design criterion by average annual reliability. The most-likely design event and confidence interval were identified as the outcome of the multivariate hydrologic design. Multivariate flood series \((Q_1, V_3, V_7, V_{15})\) from the Xijiang River were chosen as a case study, with the main findings given below.

Urbanization and reservoir regulation were found to be the driving forces responsible for the nonstationarities of both the marginal distributions and dependence structure of the multivariate flood series \((Q_1, V_3, V_7, V_{15})\). The growth of the urban population generally resulted in an increased mean value of the individual flood series, whereas it weakened the dependence of \((Q_1, V_3, V_7, V_{15})\). The increasing reservoir index had the opposite effect on the individual flood series as well as their dependence. Under a given average annual reliability, the OR exceedance probability yielded the largest design values, followed by the Kendall and the AND exceedance probabilities. Nonstationarities in both marginal distributions and dependence structure affected the outcome of the multivariate hydrologic design. It is the marginal non-
stationarities that played a dominant role in affecting the multivariate hydrologic design.

There are two remarks that can be made that are related to the practical implications of the hydrologic design methods developed in the current study that are detailed below.

The first remark relates to the length of observed flood data required for multivariate and nonstationary hydrologic designs. In theory, sufficiently long observed flood data (or other extreme-value data) are required to derive robust estimations of the distribution parameters and the correct hydrologic design values (Zheng et al., 2018). However, in reality, most data series are limited in length, thus forcing us to use what we have at hand to do research or design works without fulfilling the theoretical assumptions or requirements. Some recent studies suggested that univariate flood frequency analysis under stationary conditions usually requires flood data with a continuous period of at least 30 years (Ministry of Water Resources of People’s Republic of China, 1996; Engeland et al., 2018; Kobierska et al., 2018). However, determining a definitive answer to the length of observed flood data that is required for flood frequency analysis under multivariate and/or nonstationary settings poses a challenge, since this issue has not yet been fully addressed. However, it is certain that multivariate and nonstationary hydrologic designs naturally require datasets of longer length, since they usually contain more parameters to be estimated.

The second remark related to the trade-off between reducing estimation bias and increasing model uncertainty. Nonstationary models generally improve performance in fitting observation data by reducing estimation bias (Jiang et al., 2015b), but this is usually achieved at the expense of increasing model complexity, such as adding more model parameters and introducing more nonstationary covariates, which might induce additional sources of model uncertainty (Sernaldi and Kilsby, 2015; Read and Vogel, 2016). A careful balance between the model fitting effect and the model complexity should be maintained in practice when employing multivariate and nonstationary hydrologic design by keeping in mind the following two points: (1) the multivariate and nonstationary models should remain effective but should also be kept as simple as possible to avoid overfitting, and (2) to ensure a robust relationship between the distribution parameters and the explanatory covariates, the chosen covariates should be physically related to the flood processes and supported by a well-defined cause–effect analysis.

Data availability. All the data used in this study can be requested by contacting the corresponding author Cong Jiang at jiangcong@cug.edu.cn.

Hydrol. Earth Syst. Sci., 23, 1683–1704, 2019
Appendix A: Calculating multivariate exceedance probabilities

A1 OR exceedance probability (formulated by Eq. 9 in the paper)

Since the cumulative distribution function \( F(q_1, v_3, v_7, v_{15}|\theta_t) \) has no analytical expression, the OR exceedance probability \( p_t^{\text{or}} \) at time \( t \) is calculated by the Monte Carlo method as follows:

1. Calculate the marginal probabilities \((u_1, u_3, u_7, u_{15})\) of \((q_1, v_3, v_7, v_{15})\).
2. Generate \( m \) samples \((u_{1,i}, u_{3,i}, u_{7,i}, u_{15,i})\) \((i = 1, 2, \ldots, m)\) from the C-vine copula expressed by Eq. (4).
3. Calculate \( F(q_1, v_3, v_7, v_{15}|\theta_t) = \frac{1}{m+1} \sum_{i=1}^{m} \mathbf{1}(u_{1,i} \leq u_1, u_{3,i} \leq u_3, u_{7,i} \leq u_7, u_{15,i} \leq u_{15}) \).
4. Calculate \( p_t^{\text{or}} = 1 - F(q_1, v_3, v_7, v_{15}|\theta_t) \).

A2 AND exceedance probability (formulated by Eq. 10 in the paper)

The AND exceedance probability \( p_t^{\text{and}} \) at time \( t \) is calculated by the Monte Carlo method as follows:

1. Calculate the marginal probabilities \((u_1, u_3, u_7, u_{15})\) of \((q_1, v_3, v_7, v_{15})\).
2. Generate \( m \) samples \((u_{1,i}, u_{3,i}, u_{7,i}, u_{15,i})\) \((i = 1, 2, \ldots, m)\) from the C-vine copula expressed by Eq. (4).
3. Calculate \( p_t^{\text{and}} = \frac{1}{m+1} \sum_{i=1}^{m} \mathbf{1}(u_{1,i} \geq u_1, u_{3,i} \geq u_3, u_{7,i} \geq u_7, u_{15,i} \geq u_{15}) \).

A3 The Kendall exceedance probability (formulated by Eqs. 11 and 12 in the paper)

The Kendall exceedance probability \( p_t^{\text{ken}} \) at time \( t \) is calculated by the Monte Carlo method as follows:

1. Calculate the marginal probabilities \((u_1, u_3, u_7, u_{15})\) of \((q_1, v_3, v_7, v_{15})\).
2. Calculate \( \rho_t = F(q_1, v_3, v_7, v_{15}|\theta_t) \) (see calculation steps 2–3 for OR exceedance probability).
3. Generate \( m \) samples \((u_{1,i}, u_{3,i}, u_{7,i}, u_{15,i})\) \((i = 1, 2, \ldots, m)\) from the C-vine copula expressed by Eq. (4).
4. For \( j = 1, 2, \ldots, m \), calculate \( v_j = \frac{1}{m+1} \sum_{i=1}^{m} \mathbf{1}(u_{1,i} \leq u_1, u_{3,i} \leq u_3, u_{7,i} \leq u_7, u_{15,i} \leq u_{15,j}) \).
5. Calculate \( K_C(\rho_t) = \frac{1}{m} \sum_{i=1}^{m} \mathbf{1}(v_i \leq \rho_t) \).
6. Calculate \( p_t^{\text{ken}} = 1 - K_C(\rho_t) \).

Appendix B: Generating the multivariate design event samples (formulated by Eq. 17 in the paper)

To calculate the most-likely design event and confidence interval conditioned on \( \text{AAR} = \eta \), we need to generate numerous multivariate design event samples by the Monte Carlo method. Here, we give the details of generating the design event samples as follows:

1. Define the total number of design event samples \( N \) and the initial number of the design event sample \( i = 0 \).
2. Generate a random integer (denoted by \( t_r \)) among \((T_1, T_1+1, \ldots, T_2)\).
3. Generate a random sample \((z_{Q_1}, z_{V_3}, z_{V_7}, z_{V_{15}})\) following the multivariate distribution \( F(z_{Q_1}, z_{V_3}, z_{V_7}, z_{V_{15}}|\theta_t) \) with the distribution parameter vector \( \theta_t \).
4. Calculate the annual exceedance probability for each year throughout the period from \( T_1 \) to \( T_2 \).
5. Calculate AAR during the period from \( T_1 \) to \( T_2 \).
6. If \(|\text{AAR} - \eta| < \varepsilon\) (where \( \varepsilon \) is a very small value, such as 0.0001), \( i = i + 1 \).
7. If \( i < N \), repeat steps (2)–(6).
Author contributions. CJ and LX developed the main ideas. CJ and LY implemented the algorithms of the methods. CJ and JD collected the data used in the case study. CJ, LX and CYX prepared the paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This research was jointly financially supported by the National Natural Science Foundation of China (NSFC grants 51809243 and 51525902), the Fundamental Research Funds for the Central Universities (grant CUG170679), the Research Council of Norway (FRINATEK Project 274310) and the “111 Project” Fund of China (B18037), all of which are greatly appreciated. The authors express their gratitude to the reviewers for their constructive comments and suggestions that have helped improve the paper.

Review statement. This paper was edited by Carlo De Michele and reviewed by four anonymous referees.

References


Vogel, R. M.: Reliability indices for water supply systems, J. Wat.
Ter Res. Pl., 113, 563–579, https://doi.org/10.1061/(ASCE)0733-
Vogel, R. M., Yaindl, C., and Walter, M.: Nonstationarity: 
Flood magnification and recurrence reduction factors in the
United States, J. Am. Water Resour. As., 47, 464–474, 
Volpi, E. and Fiori, A.: Design event selection in bivariate hydro-
logical frequency analysis, Hydrolog. Sci. J., 57, 1506–1515, 
Xiao, Y., Guo, S., Liu, P., Yan, B., and Chen, L.: De-
sign flood hydrograph based on multicharacteristic syn-
thesis index method, J. Hydrol. Eng., 14, 1359–1364, 
https://doi.org/10.1061/(ASCE)1084-0699(2009)4:12(1359),
2009.
Xiong, L., and Guo, S.: Trend test and change-point detection 
for the annual discharge series of the Yangtze River at the 
Yichang hydrological station, Hydrolog. Sci. J., 49, 99–112, 
Xiong, L., Jiang, C., Xu, C.-Y., Yu, K.-X., and Guo, S.: 
A framework of changepoint detection for multivariate hydrological series, Water Resour. Res., 51, 8198–8217, 
Xu, B., Xie, P., Tan, Y., Li, X., and Liu, Y.: Analysis of flood re-
turning to main channel influence on the flood control ability of 
Xijiang River, Journal of Hydroelectric Engineering, 33, 65–72, 
Yan, L., Xiong, L., Guo, S., Xu, C.-Y., Xia, J., and Du, T.: 
Comparison of four nonstationary hydrologic design meth-
ods for changing environment, J. Hydrol., 551, 132–150, 
Yang, T., Shao, Q., Hao, Z., Chen, X., Zhang, Z., Xu, C.-
Y., and Sun, L.: Regional frequency analysis and spatio-
temporal pattern characterization of rainfall extremes in 
the Pearl River Basin, China, J. Hydrol., 380, 386–405, 
Yin, J., Guo, S., Liu, Z., Chen, K., Chang, F., and Xiong, F.: Bivari-
ate seasonal design flood estimation based on copulas, J. Hydrol.
Eng., 22, 05017028, https://doi.org/10.1061/(ASCE)HE.1943-
5584.0001594, 2017.
Zhang, L. and Singh, V. P.: Trivariate flood frequency analysis using 
the Gumbel–Hougaard copula, J. Hydrol. Eng., 12, 431–439, 
2007.
the dependence between extreme rainfall and storm 
surge in the coastal zone, J. Hydrol., 505, 172–187, 
dependence between extreme rainfall and storm surge to esti-
mate coastal flooding risk, Water Resour. Res., 50, 2050–2071, 
variable method to estimate coastal flood risk, J. Flood Risk 
Zheng, F., Tao, R., Maier, H. R., See, L., Savic, D., Zhang, T., 
Chen, O., Assumpção, T. H., Yang, P., Heidari, B., Rickermann, 
J., Minsker, B., Bi, W., Cai, X., Solomatine, D., and Popescu, I.: 
Crowdsourcing methods for data collection in geophysics: State 
of the art, issues, and future directions, Rev. Geophys., 56, 698– 