Flood Hydrograph Coincidence Analysis of the Upper Yangtze River and Dongting Lake, China

Chao Zhang

North China Electric Power University School of Water Resources and Hydropower Engineering

Changming Ji

North China Electric Power University School of Water Resources and Hydropower Engineering

Yi Wang (✉ hywy02@foxmail.com)

North China Electric Power University School of Water Resources and Hydropower Engineering

<https://orcid.org/0000-0001-7815-5254>

Qian Xiao

North China Electric Power University School of Water Resources and Hydropower Engineering

Research Article

Keywords: Flood hydrograph coincidence, Hydrological extremes, Joint distribution, Vine copula, Risk probability, Yangtze River Basin

Posted Date: March 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-277815/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Natural Hazards on September 5th, 2021. See the published version at <https://doi.org/10.1007/s11069-021-04993-2>.

Flood hydrograph coincidence analysis of the upper Yangtze River and Dongting Lake, China

Chao Zhang¹, Changming Ji¹, Yi Wang^{1,*}, Qian Xiao¹

Chao Zhang, Postgraduate Student, 916989850@qq.com, ORCID: 0000-0002-4443-6004

Changming Ji, Professor, cmji@ncepu.edu.cn

Yi Wang* (**The Corresponding Author**), Professor, hywy02@foxmail.com, ORCID: 0000-0001-7815-5254

Qian Xiao, Postgraduate Student, xiao_qian19@163.com

¹ School of Water Resources and Hydropower Engineering, North China Electric Power University, Beijing 102206, China

Abstract

In hydrological research, flood events can be analyzed by flood hydrograph coincidence. Existing flood hydrograph coincidence research mostly focuses on the analysis of the coincidence risk probability of the annual maximum flood event using the 15-day maximum annual flood volume; the actual duration of the flood hydrograph is neglected. The duration of the flood hydrograph is a key variable in (1) determining whether flood hydrograph coincidence occurs, and (2) accurately calculating the flood hydrograph coincidence risk probability. This paper creatively proposes a novel method to analyze the flood hydrograph coincidence risk probability by establishing a five-dimensional joint distribution of flood volumes, durations and interval time for two hydrologic stations. More specifically, using the annual maximum flood of the upper Yangtze River and input from Dongting Lake as an example, the Pearson Type III and the mixed von Mises distributions were used to establish the marginal distribution of flood volumes, flood duration and interval time. Subsequently, the five-dimensional joint distribution based on vine copula was established to analyze the flood hydrograph coincidence risk probability. The results were verified by comparison with a historical flood sequence. The flood hydrograph coincidence volume-risk probability curve was also obtained, providing theoretical support for flood control safety and risk management in the middle and lower Yangtze River. This study also demonstrates the significant

beneficial role of regulation by the Three Gorges Water Conservancy Project in mitigating flood risk of the Yangtze River.

Keywords: Flood hydrograph coincidence, Hydrological extremes, Joint distribution, Vine copula, Risk probability, Yangtze River Basin

Declarations

Funding

The study was supported by the Guangdong Foundation for Program of Science and Technology Research (2020B1111530001).

Conflicts of interest/Competing interests

- The authors have no relevant financial or non-financial interests to disclose.
- The authors have no conflicts of interest to declare that are relevant to the content of this article.
- All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.
- The authors have no financial or proprietary interests in any material discussed in this article.

Availability of data and material

The data that support the findings of this study are openly available in the Changjiang Water Resources Commission of The Ministry of Water Resources of China at <http://www.cjw.gov.cn/>.

Code availability

Not applicable

Author's contributions

All authors contributed to the study conception and design. Data collection and analysis were performed by Changming Ji and Yi Wang. The first draft of the manuscript was written by

65 Chao Zhang, the figures and tables were made by Qian Xiao. All authors commented on
66 previous versions of the manuscript. All authors read and approved the final manuscript.
67

1 Introduction

Floods are caused by the rapid increase of river water volume and water level in response to heavy or continuous rainfall, or snow- and ice-melt in a river basin. They can result in significant financial losses, the destruction of infrastructure, and the loss of human life, and as such are serious natural hazards (Brocca et al. 2011; Li et al. 2012; Stein et al. 2020). In the context of global climate change, extreme meteorological and hydrological events are becoming more frequent, and the risk of basin-wide flooding is increasing (Wang et al. 2019; Su 2020; Try et al. 2020; Yang et al. 2020). Therefore, more attention to basin-wide flooding mitigation is needed.

Flood characteristics at the basin exit are usually influenced by the coincidence of runoff from two or more subbasins in the catchment. For example, the peak or maximum flood volume of the main stream and one or more tributaries may reach the same river section at a given time (Chen et al. 2012; Feng et al. 2020). When runoff from two locations is coincident, the magnitude of the flood peak and flood volume increase primarily due to accumulation. Therefore, flood event coincidence analysis in a basin is of significance to the formulation of flood control strategies and the rational development and utilization of water resources, particularly in large river basins.

The traditional method to study flood event coincidence within a river basin is to conduct probabilistic statistical analysis on the historical flood event coincidence using synchronously collected flood data in the study area (Zhu et al. 2015; Gao et al. 2017; Zhang et al. 2018). However, the traditional approach only focuses on historical floods, and thus cannot provide the coincidence risk probability of extremely large floods (often missing from the historic record), such as floods with 100-year or even 1000-year return periods. However, extreme flood events are particularly important in the design of dams and other hydrologic systems as well as the mapping of flood-prone areas. Considering that flood event coincidence is a typical multivariate frequency combination problem, multivariate analysis can be applied (Prohaska et al. 2008). Prohaska et al. (2010, 2012), for example, used a two-dimensional normal distribution to study the flood event coincidence probability of the Danube River and its tributaries. However, flood sequences are usually characterized by a skewed distribution, rather than a perfectly normal distribution. For example, flood sequences in most parts of China generally follow a Pearson Type III (P-III) distribution (MWR 2006).

Copula functions are an effective method for constructing multivariable joint distributions. Due to their flexible structure (Favre et al. 2004), copulas have been widely used in hydrological multivariate analysis in recent years (Salvadori and Michele 2004, Zhang and Singh 2007, Karmakar and Simonovic 2009, Kao and Govindaraju 2010, Liu et al. 2011, Reddy and Ganguli 2012, Chen et al. 2013, Sraj et al. 2015). In analyzing flood event coincidence, Klein et al. (2010), Schulte and Schumann (2016) and Bing et al. (2018) used copulas to establish the joint distribution among different river flood peaks, and to calculate the probability of the simultaneous occurrence of different river flood peaks within the basin (i.e., the probability of the occurrence of a river flood peak under the condition of another given river flood peak). However, they only considered flood magnitude and neglected flood occurrence time. Further, Chen et al. (2012), Peng et al. (2018), Feng et al. (2020) and Zhang et al. (2020) used copulas to analyze the flood peak coincidence considering the flood magnitude and occurrence time simultaneously.

Even when the flood peaks are not directly coincident, but the flood hydrographs overlap, the impacts may be catastrophic. For example, in 2020 Shexian County in China was impacted by a flood with a 50-year return period; it caused US \$300 million in damages to the city and US \$100 million in damages to the countryside (Cheng and Zhang 2020). The peaks of the two “floods” that caused this damage occurred on July 3 and July 7, and thus were not perfectly coincident. However, the two flood hydrographs overlapped to a large degree. When the latter flood reached the area, the flood stage in the river was still receding from the previous flood. As a result, the second flood event led to higher rising water levels than from a single event, and therefore caused serious flood damage. As such, overlap in flood hydrographs should also be considered as a flood coincidence event.

Coincidence of different flood events can be analyzed by the coincidence of their flood hydrograph, and flood duration is an important factor. Chen et al. (2019) and Yan et al. (2013) established a three-dimensional joint distribution of the annual maximum 15-day flood volumes and their interval time to analyze flood hydrograph coincidence. While, the duration of annual maximum flood (AMF) in the real condition is not always 15 days, which will depends on the actual condition of the basin. So that the actual duration of the flood hydrograph will be neglected when the volume of AMF is replaced with the annual maximum 15-day flood volume. To address this problem, this paper creatively proposes a method to

analyze flood hydrograph coincidence by establishing a five-dimensional joint distribution of flood volumes, durations and interval time at two stations in the basin. The approach will depend on the nature of the river being studied, and can be applied to other basins with similar complex river systems and frequent flood hazards. Specifically, this study focuses on the analysis of flood hydrograph coincidence between the upper Yangtze River and input from Dongting Lake located downstream to establish a five-dimensional joint distribution (Fig. 1). The coincidence risk probability of the flood hydrograph of the upper Yangtze River and Dongting Lake was then calculated and analyzed through stochastic simulation, so that the curve of the flood hydrograph coincidence volume and flood event coincidence risk probability could be obtained. Moreover, the role of the Three Gorges Project on flood event coincidence risk prevention of the upper Yangtze River was assessed.

2 Study area

The middle and lower Yangtze River Basin is prone to frequent flood hazards. Floods along the main stream of the Yangtze River above Yichang exhibit large peak flows of long duration; the flood events generally last for more than 10 days. During the flood season, if the flood of the upper Yangtze River encounters a flood from Dongting Lake, which is in the south of the Jingjiang section of the river (Fig. 1), it will not only adversely affect the flood control of the middle Yangtze River, but also make flooding in the Jingjiang section more severe. The flood event coincidence risk probability of the upper Yangtze River and tributaries of the middle Yangtze River Basin is an important theoretical component of the Three Gorges Project to timely formulate reasonable flood control plans. Therefore, studying flood event coincidence risk probability of the upper Yangtze River and Dongting Lake is of great significance for flood control and disaster reduction in the middle and lower Yangtze River (Zhang et al. 2020).

3 Methodology

3.1 Utilized data and approach

In this study, the annual maximum floods (AMFs) from 1951 to 2016 at the Yichang Station (Station 1) and Chenglingji Station (Station 2) were collected to study the flood event coincidence risk probability of the upper Yangtze River and Dongting Lake (Fig. 1). The Yichang Station is located at the boundary between the upper and middle Yangtze River, and is a control station for water and sediment from the upper Yangtze River. It controls a

drainage area of $1 \times 10^6 \text{ km}^2$, with an average annual runoff of $4.29 \times 10^{11} \text{ m}^3$. The maximum annual runoff was $5.75 \times 10^{11} \text{ m}^3$ (in 1954), whereas the minimum annual runoff at the station was $2.85 \times 10^{11} \text{ m}^3$ (in 2006). Runoff during the flood season (May to October) accounts for 78% of the average annual runoff. The Chenglingji Station is an important control station for the discharge of water and sediment from Dongting Lake into the main stream of the Yangtze River (Fig. 1). It controls a drainage area of $2.59 \times 10^5 \text{ km}^2$, with an average annual runoff of $2.85 \times 10^{11} \text{ billion m}^3$. The maximum annual runoff was $5.27 \times 10^{11} \text{ m}^3$ (in 1954), whereas the minimum annual runoff was $1.48 \times 10^{11} \text{ m}^3$ (in 2011). Runoff during the flood season (May to October) accounts for 73% of the annual average runoff.

Annual daily discharge data at the Yichang and Chenglingji Stations from 1951 to 2016 were collected by the Changjiang Water Resources Commission of The Ministry of Water Resources of China to study the flood risk probability of the upper Yangtze River and Dongting Lake. The water storage operation of the Three Gorges Project has changed the local hydrological cycle of the Three Gorges Reservoir, which has changed the amount, duration and occurrence time of the upstream flood. These alterations have also impacted the flood hydrograph coincidence risk probability of downstream reaches. Therefore, this study conducts risk probability analysis of flood hydrograph coincidence between the upper Yangtze River and Dongting Lake during two periods for comparative purposes: (1) from 1951 to 2002 (Period 1: before the construction of the Three Gorges Project), and (2) from 2003 to 2016 (Period 2: after the construction of the Three Gorges Project).

3.2 Definitions of AMF event characteristics

To analyze the AMF hydrograph coincidence risk probability, an AMF event must be defined by specific characteristics that allow the occurrence of an AMF event to be objectively recognized (Tosunoglu et al. 2020). The following criteria were used to define and characterize an AMF event.

- (1) The highest point in the annual daily discharge diagram was defined as the peak of the AMF event (Q);
- (2) The first point at which the daily increase in discharge exceeds a certain threshold (Q_t') on the rising limb of the hydrograph corresponding to the AMF peak is the starting point of the AMF event. Its corresponding time is the starting time of the AMF event (t');

- (3) The last point at which the daily decline in discharge exceeds a certain threshold (Q_t'') on the recessional limb of the hydrograph corresponding to the AMF peak is the ending point of the AMF event. Its corresponding time is the ending time of the AMF event (t'');
- (4) If the interval time between the AMF event peak and its adjacent flood event peak is no more than 7 days, the flood event corresponding to the adjacent flood peak is also considered as a part of the AMF event;
- (5) The volume of the AMF event (V) is the volume of runoff between t' and t'' ;
- (6) The duration of the AMF event (d) is the interval time between t' and t'' ;
- (7) The occurrence time of the AMF event (T) can be denoted by the starting time (t').

Based on the hydrological regime and typical historical flood hydrograph characteristics of the Yangtze River Basin, the two thresholds (Q_t' and Q_t'') were set at 2000 ($m^3 s^{-1} d^{-1}$) and 1000 ($m^3 s^{-1} d^{-1}$), respectively. The extracted AMF event characteristics were consistent with the historical data. The positions of the AMF event characteristics in an annual daily discharge diagram are shown in Fig. 2.

3.3 Flood hydrograph coincidence model

Recent research on flood hydrograph coincidence assumes that flood hydrograph coincidence occurs when the flood hydrographs from two stations overlap and the duration of the overlapping parts account for more than 1/2 of the flood duration of any station. The magnitudes of the two floods are represented by flood volumes (Yan et al. 2013; Huang et al. 2018). Assuming that there are two stations A and B, their flood volumes, flood durations and flood occurrence times are denoted by V_1 , V_2 , d_1 , d_2 , T_1 , and T_2 , whereas the interval time between the two stations is denoted by T_d where:

$$T_d = T_1 - T_2 . \quad (1)$$

The definition of flood hydrograph coincidence is schematically shown in Fig. 3 for two conditions:

- (1) When the flood duration of station A is longer than or equal to station B. In this case, if

$T_2 \in [T_1 - d_2 / 2, T_1 + d_1 - d_2 / 2]$, the duration of the overlapping part of the two floods accounts for more than 1/2 of the duration of the flood in station B, and a flood hydrograph coincidence occurs (Fig. 3(a));

- (2) When the flood duration of station A is shorter than station B. In this case, if

$T_2 \in [T_1 + d_1 / 2 - d_2, T_1 + d_1 / 2]$, the duration of the overlapping part of the two floods

accounts for more than 1/2 of the duration of the flood in station A, and a flood hydrograph coincidence occurs (Fig. 3(b)).

Put in the form of Equation (1), a flood hydrograph coincidence happens when:

$$\begin{cases} T_d \in \left[\frac{d_2}{2} - d_1, \frac{d_2}{2} \right], & (d_1 \geq d_2) \\ T_d \in \left[-\frac{d_1}{2}, d_2 - \frac{d_1}{2} \right], & (d_1 < d_2) \end{cases}, \quad (2)$$

Therefore, when studying the hydrograph coincidence risk probability of two floods greater than specified magnitudes, it is necessary to calculate the probability (P_{fc}) that the flood volumes of the two stations are greater than the design values, and their interval time meets the above definition of flood hydrograph coincidence. This probability is expressed as:

$$\begin{aligned} P_{fc} = & P(V_1 \geq V_1^{R_{A0}}, V_2 \geq V_2^{R_{B0}}, \underline{T_d} \leq T_d \leq \overline{T_d}) \\ & \begin{cases} \underline{T_d} = \frac{d_2}{2} - d_1, \overline{T_d} = \frac{d_2}{2}, & (d_1 \geq d_2) \\ \underline{T_d} = -\frac{d_1}{2}, \overline{T_d} = d_2 - \frac{d_1}{2}, & (d_1 < d_2) \end{cases}, \quad (3) \end{aligned}$$

where $V_1^{R_{A0}}$ and $V_2^{R_{B0}}$ are the design flood volumes of station A and B, respectively; and $\underline{T_d}$ and $\overline{T_d}$ are the upper and lower limits of the interval time (T_d), respectively. In light of Equation (3), it is necessary to establish the five-dimensional joint distribution of the variables, V_1 , V_2 , d_1 , d_2 , and T_d , for the flood hydrograph coincidence analysis. The upper and lower limits of T_d are related to d_1 and d_2 at the same time. The form is complex, and it is difficult to simplify directly. In this paper, the idea of stochastic simulation is adopted to calculate the probability (Roo et al. 1992). According to this approach, you get M sets of simulation values associated with the joint distribution, and the value of P_{fc}' is:

$$P_{fc}' = \frac{m}{M}, \quad (4)$$

where m is the number of the data satisfying Equation (3). We can assume that $P_{fc}' \approx P_{fc}$, when M is large enough (generally greater than 5000).

3.4 Marginal distribution model

The Chinese Ministry of Water Resources (MWR 2006) advocates the P-III distribution as the unified model of flood frequency analysis. The duration and interval time of the two stations' AMFs can be regarded as vectors with periodic changes. Thus, they can be

described by the mixed von Mises distribution (Chen et al. 2012; Yan et al. 2013; Huang et al. 2018; Peng et al. 2019).

The P-III curve is an asymmetric and unimodally skewed curve. Its probability density function is expressed as:

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} (x - a_0)^{\alpha-1} e^{-\beta(x-a_0)}, \quad (5)$$

where $\Gamma(\alpha)$ is the gamma function of α , and α, β, a_0 are parameters of the P-III curve that can be calculated as:

$$\begin{cases} \alpha = \frac{4}{C_s^2} \\ \beta = \frac{2}{\bar{x} C_v C_s} \\ a_0 = \bar{x} \left(1 - 2 \frac{C_v}{C_s} \right) \end{cases} \quad (6)$$

where $\bar{x}$ is the average value; C_v is the coefficient of variation; and C_s is the coefficient of skewness. The curve-fitting method is often used to estimate the parameters of the P-III curve (Geyer 1940).

The von Mises distribution is known as the normal distribution on a circle; it is an important model to describe directional data (Fisher 1993). The probability density function of the unimodal von Mises distribution is:

$$f(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp\{\kappa \cos(\theta - \mu)\}, 0 \leq \theta \leq 2\pi, \quad (7)$$

where μ is the mean direction; κ is the concentration parameter; and $I_0(\kappa)$ is the 0-order modified Bessel function of the first kind calculated by:

$$I_0(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} \exp(\kappa \cos \theta) d\theta. \quad (8)$$

The probability density function of the mixed von Mises distribution is:

$$f(\theta; G) = \sum_{i=1}^p \alpha_i f(\theta; \mu_i, \kappa_i), i = 1, 2, \dots, p, \quad (9)$$

where $G = (\alpha_1, \alpha_2, \dots, \alpha_p; \mu_1, \mu_2, \dots, \mu_p; \kappa_1, \kappa_2, \dots, \kappa_p)$; α_p is the mean direction; and κ is the mixing proportion. The maximum likelihood method is often used to estimate the mixed von Mises distribution parameters (Spurr 1991).

3.5 Copula functions

Copula a tool that connects the joint distribution with its marginal distributions (Sklar 1959). The traditional copula functions used in the construction of a multidimensional joint distribution are mainly multidimensional elliptic copula and Archimedean copula, but their structures are relatively fixed, and they require the same correlation structure among variables. Therefore, they cannot accurately describe the dependencies of higher-dimensional variables (Aas et al. 2009; Schepsmeier and Czado 2016; Yu et al. 2019; Jane et al. 2020; Tosunoglu et al. 2020; Wu et al. 2020;).

The vine copula was first proposed by Joe (1996). In comparison to two previous forms, the structure of vine copula has been greatly improved. Vine copula allow a combination of $d(d-1)/2$ specified copulas, and its principle is to decompose a multivariate probability density function into $d(d-1)/2$ two-dimensional copula density functions. There are three main types of vine copula: C-vine, D-vine and R-vine. Among them, the structure of the R-vine copula is relatively flexible and there are many forms. C-vine and D-vine are special forms of R-vine with fixed structures (Bedford and Cooke 2001; Cooke 2002; Kurowicka and Cooke 2006; Aas et al. 2009); these are schematically shown in Fig. 4.

3.6 Goodness-of-fit of distributions

In order to quantitatively evaluate the fitting error and select the appropriate copula function, the mean square error (MSE) and the Akaike information criterion (AIC) were used. These indices can be expressed as:

$$\begin{aligned} MSE &= \frac{1}{n-1} \sum_{i=1}^n (Pe_i - P_i)^2; \\ AIC &= n \log(MSE) + 2m, \end{aligned} \tag{10}$$

where m is the number of model parameters, n is the number of samples, P_i represents the copula value of consecutive sample observations, and Pe_i represents the corresponding multivariate empirical probability. AIC is a measure of the quality of the statistical model's fit to the data. For a particular copula function, the smaller the AIC value of the objective function, the better the copula function simulation.

4 Results and analysis

4.1 Marginal distributions

Advantages of using a copula function to establish the multidimensional joint distribution is that the marginal distribution can work in many forms, and the marginal distribution and joint distribution can be considered separately. The flood hydrograph coincidence model proposed in this paper includes five variables: V_1 , V_2 , d_1 , d_2 , and T_d . The P-III distribution was used to establish the marginal distributions of AMF volumes at the Yichang and Chenglingji stations, whereas the mixed von Mises distribution was used to establish the marginal distributions of the AMF durations and their interval time at the Yichang and Chenglingji stations. The curve-fitting method and maximum likelihood method were used to estimate the parameters of the P-III distribution and mixed von Mises distribution, respectively. The distributions were subsequently analyzed using distribution fitting tests.

Results of the estimated parameters, along with the results of the Kolmogorov-Smirnov (K-S) and χ^2 tests of the P-III distributions of the AMF volumes, are shown in Table 1. When the significance level is 5%, the χ^2 test results and K-S test results are both less than the critical values, and the distributions pass the tests. The AMF volumes could be effectively described by the P-III distribution. The frequency curves provided in Fig. 5 show that the empirical data points fit the theoretical curves well.

Table 2 provides the results of the estimated parameters, the K-S test results and the χ^2 test results of the mixed von Mises distributions of the AMF durations and interval times. When the significance level is 5%, the χ^2 test results and K-S test results are both less than the critical values, and the distributions pass the tests. We therefore assume that the AMF durations and interval times could be effectively described by the mixed von Mises distribution. The probability density fitting diagrams are shown in Fig. 6 and Fig. 7 and illustrate that the theoretical curves fit the empirical bar charts well.

4.2 Joint distributions

Vine copula parameter estimation and preferential selection were carried out using a routine in the R software package. The vine copula parameter estimation and preferential selection

results were used for establishing the five-dimensional joint distributions in this study. The parameters of the R software package were set as follows:

- (1) The optional two-dimensional copula functions were all types of copula in the package, which included 37 two-dimensional copula functions in total;
- (2) The alternative joint distribution structures were C-vine, D-vine and R-vine copula;
- (3) AIC was the evaluation standard to select suitable two-dimensional copula functions;
- (4) The confidence level of independent hypothesis testing was set at 0.05;
- (5) The correlation between variables was expressed by the Kendall correlation coefficient.

Table 3 shows the vine copula structures of the five-dimensional joint distributions of AMF hydrograph coincidence for the two study periods. After the establishment of the joint distributions, the accuracy and reliability of the joint distributions were tested. In this paper, 5000 sets of five-dimensional flood variables were simulated, and the scatter diagrams between the two variables of the simulated sequence and the historical sequence were compared (Figs. 8, 9). The comparison shows that the simulated sequence scatter diagrams not only cover almost all the historical sequence scatter diagrams, but also fully retain the shape and trend characteristics of the historical sequence scatter diagrams. Therefore, the simulated sequences retained the natural characteristics of the historical sequences, and the vine copula structure established above fully reflect the structural characteristics among the variables of the historical sequences. The established five-dimensional joint distributions in this study are therefore considered accurate and reliable and can be used for the risk probability analysis of AMF hydrograph coincidence.

4.3 Risk probability analysis of AMF hydrograph coincidence

As mentioned above, flood hydrograph coincidence refers to the overlap between flood hydrographs, and the duration of the overlapping parts account for more than 1/2 of the flood duration of any station. Equation (3) specifies the method to calculate the flood hydrograph coincidence risk probability. The stochastic simulation method mentioned above was used to calculate AMF hydrograph coincidence probability at the Yichang and Chenglingji stations. P'_{fc} was calculated using the five joint distributions based on vine copula and the stochastic simulation of 100000 sets of values of the five-dimensional flood variables. We assumed that $P'_{fc} \approx P_{fc}$ due to the large number of simulated data.

Fig. 10 shows the hydrograph coincidence risk probability of the AMFs greater than specified magnitudes of the simulated and historical sequences (in terms of the return period, $R_1 \geq R_0$; $R_2 \geq R_0$). During Period 1, the AMF hydrograph coincidence risk probabilities for each return period combination of the simulated and historical sequence were very close; the absolute value of errors were all less than 0.025 (Fig. 10(c)). During Period 2, the AMF hydrograph coincidence risk probabilities with a return period of 1-year were very similar between the simulated sequence and the historical sequence. Due to the short historical timeframe, there was no AMF hydrograph coincidence with a return period combination of more than 2-years, and the risk probability values corresponding to the simulated sequence were very small (less than 1/14). Therefore, the AMF hydrograph coincidence risk probability calculated by the simulated sequence was basically the same as that of the historical sequence; thus, this method can be used to calculate the AMF hydrograph coincidence risk probability with extremely large magnitudes.

In the AMF hydrograph coincidence risk probability analysis, six (6) return periods of the AMF were chosen at each station, and the hydrograph coincidence risk probability with 36 combinations ($R_1 \geq R_{01}$, $R_2 \geq R_{02}$) was calculated (Table 4), along with the AMF hydrograph coincidence volume (Table 5). Subsequently, the calculation results of the AMF hydrograph coincidence risk probability and volume were used to develop the hydrograph coincidence volume-risk probability curves (Fig. 11).

The calculation results of the hydrograph coincidence risk probability in Table 4 show that the AMF hydrograph coincidence risk probability of the upper Yangtze River and Dongting Lake decreases with an increase in the flood magnitude. For example, from 1951-2002, floods of the upper Yangtze River and Dongting Lake with a 5-year, 10-year, 50-year, 100-year, 500-year and 1000-year return period combination exhibited hydrograph coincidence risk probabilities of 1.41×10^{-2} , 5.26×10^{-3} , 5.80×10^{-4} , 2.70×10^{-4} , 4.00×10^{-5} and 3.00×10^{-5} , respectively; probabilities decreased with longer return periods. From 2003-2016, floods of the upper Yangtze River and Dongting Lake with a 5-year, 10-year, 50-year and 100-year return period combination possessed hydrograph coincidence risk probabilities of 5.70×10^{-3} , 1.81×10^{-3} , 7.00×10^{-5} and 2.00×10^{-5} , respectively, again exhibiting a decreasing trend. Interestingly, by comparing the hydrograph coincidence risk probabilities between the two periods (Table 4), the AMF hydrograph coincidence risk probabilities of the upper Yangtze River and Dongting Lake were greatly reduced by the operation of the Three

Gorges Project. For example, during 1951-2002, floods of the upper Yangtze River and Dongting Lake with a 5-year, 10-year, 50-year and 100-year return period combination had hydrograph coincidence risk probabilities of 1.41×10^{-2} , 5.26×10^{-3} , 5.80×10^{-4} and 2.70×10^{-4} , respectively, whereas during the 2003-2016 period, risk values decreased to 5.70×10^{-3} , 1.81×10^{-3} , 7.00×10^{-5} and 2.00×10^{-5} , respectively.

The flood hydrograph coincidence volume-risk probability curves in Fig. 11 show that the AMF hydrograph coincidence risk probability of the upper Yangtze River and Dongting Lake decreased with an increase in the hydrograph coincidence volume. For example, during the 1951-2002 period, the hydrograph coincidence probabilities corresponding to hydrograph coincidence volumes of $2.00 \times 10^{11} \text{ m}^3$, $4.00 \times 10^{11} \text{ m}^3$ and $6.00 \times 10^{11} \text{ m}^3$ were 0.213, 0.123, and 0.049, respectively. During the 2003-2016 period, the hydrograph coincidence probabilities corresponding to hydrograph coincidence volumes of $2.00 \times 10^{11} \text{ m}^3$, $4.00 \times 10^{11} \text{ m}^3$ and $6.00 \times 10^{11} \text{ m}^3$ were 0.072, 0.028 and 0.005, respectively. At the same time, a comparison of the flood hydrograph coincidence volume-risk probability curves in Fig. 11 show that the AMF hydrograph coincidence risk probabilities of the upper Yangtze River and Dongting Lake were greatly reduced by operation of the Three Gorges Project. For example, from 1951-2002, the hydrograph coincidence probabilities corresponding to hydrograph coincidence volumes of $2.00 \times 10^{11} \text{ m}^3$, $4.00 \times 10^{11} \text{ m}^3$ and $6.00 \times 10^{11} \text{ m}^3$ were 0.213, 0.123, and 0.049, respectively, whereas during the 2003-2016 period, these values decreased to 0.072, 0.028 and 0.005, respectively.

5 Discussion

This study proposed a creative methodology to analyze the AMF hydrograph coincidence risk probability of the upper Yangtze River and its downstream input from Dongting Lake. The main considerations and findings are as follows:

- (1) In view of the fact that flood duration is usually neglected in the analysis of the flood hydrograph coincidence in a basin, this study proposed a method for flood hydrograph coincidence risk probability analysis by establishing a five-dimensional joint distribution of the flood volumes, durations and interval time at two stations. This model takes the flood factors into account and was able to present the actual flood coincidence more comprehensively. Therefore, the results of risk probability analysis can describe the hydrological process of flood event coincidence more accurately and therefore more

effectively meet the needs of actual flood control planning. The approach can be applied to other basins with similar complex river systems.

- (2) In recent studies of flood event coincidence, traditional copulas were used to establish the joint distribution of flood peaks among stations, and to calculate the risk probability of flood peak coincidence. However, when there are certain differences in correlation structures among variables, the use of traditional copulas to establish high-dimensional joint distributions has limitations. In this paper, a vine copula was used to establish the five-dimensional joint distributions of the flood variables. It was then shown using historical sequences that the joint distributions established in this study were reliable. Therefore, when the correlation structures among hydrological variables under study are not identical and the joint distribution dimension is high, the vine copula function is recommended.
- (3) The results of this study show that the regulation of the Three Gorges Project during the flood season reduced the flood hydrograph coincidence risk probability and alleviated the flood control pressure in the middle and lower Yangtze River. At the same time, the flood hydrograph coincidence model developed in this study can more comprehensively analyze the flood event coincidence risk probability of the main stem of the Yangtze River and its tributaries, which provides a useful reference for the Three Gorges Project operation in term of flood mitigation.

6 Conclusions

This study proposed a method to analyze the hydrograph coincidence risk probability of the upper Yangtze River and input from Dongting Lake using a five-dimensional joint distribution of flood volumes, durations and interval time for two hydrologic stations. The risk probabilities and volumes of flood hydrograph coincidence with different return period combinations were obtained. Flood duration was used in this method based on the flood hydrograph, and result verification showed that the risk analysis was accurate and consistent with actual situations. Finally, by analyzing the flood hydrograph coincidence risk probability of the upper Yangtze River and Dongting Lake before and after operation of the Three Gorges Project, the positive effect of the Three Gorges Project on flood risk mitigation was demonstrated. In future work, a flood disaster model will be set up and combined with this study to further quantify flood risk, thereby providing a more intuitive theoretical basis for regional flood control and disaster reduction.

Acknowledgments

The study was supported by the Guangdong Foundation for Program of Science and Technology Research (2020B1111530001).

We thank LetPub (www.letpub.com) for its linguistic assistance and scientific consultation during the preparation of this manuscript.

References

- Aas K, Czado C, Frigessi A, Bakken H (2006) Pair-copula constructions of multiple dependence. *Insur Math Econ* 44(2):182-198.
<https://doi.org/10.1016/j.insmatheco.2007.02.001>
- Bedford T, Cooke RM (2001) Probability density decomposition for conditionally dependent random variables modeled by vines. *Ann Math Artif Intel* 32(1):245-268.
<https://doi.org/10.1023/a:1016725902970>
- Bing JP, Deng PX, Zhang X, Lv SY, Marco M, Xiao Y (2018) Flood coincidence analysis of Poyang Lake and Yangtze River: risk and influencing factors. *Stoch Env Res Risk A* 32(4):879-891. <https://doi.org/10.1007/s00477-018-1514-4>
- Brocca L, Melone F, Moramarco T (2011) Distributed rainfall-runoff modelling for flood frequency estimation and flood forecasting. *Hydrol Process* 25(18):2801-2813.
<https://doi.org/10.1002/hyp.8042>
- Chen L, Singh VP, Guo SL, Hao ZC, Li TY (2012) Flood coincidence risk analysis using multivariate copula functions. *J Hydrol Eng* 17(6):742-755.
[https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000504](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000504)
- Chen YD, Zhang Q, Xiao M, Singh VP (2013) Evaluation of risk of hydrological droughts by the trivariate Plackett copula in the East River Basin (China). *Nat Hazards* 68(2):529-547. <https://doi.org/10.1007/s11069-013-0628-8>
- Cheng Z, Zhang CX (2020) Analysis of the rainstorm flood process of "2020.7.7" in She County, Anhui Province. *China Flood & Drought Management* 30:236-240 (in Chinese)
<https://doi.org/10.16867/j.issn.1673-9264.2020252>
- Cooke BRM (2002) Vines: a new graphical model for dependent random variables. *Ann Stat* 30(4):1031-1068. <https://doi.org/10.2307/1558694>

509 Favre AC, Adlouni SE, Perreault L, Thiémonge N, Bobée B (2004) Multivariate hydrological
510 frequency analysis using copulas. *Water Resour Res* 40(1):290-294.
511 <https://doi.org/10.1029/2003WR002456>

512 Feng Y, Shi P, Qu SM, Mou SY, Chen C, Dong FC (2020) Nonstationary flood coincidence
513 risk analysis using time-varying copula functions. *Scientific Reports* 10(1):3395.
514 <https://doi.org/10.1038/s41598-020-60264-3>

515 Fisher NI (1993) *Statistical Analysis of Circular Data*. Cambridge University Press
516 Cambridge, UK

517 Gao HQ, Dai J, Shen Y, Fan HX (2017) Analysis of the flood composition and flood
518 occurrence in Xijiang River. *Pearl River* 38(7):18-21 (in Chinese)
519 <https://doi.org/10.3969/j.issn.1001-9235.2017.7.004>

520 Geyer JC (1940) New curve - fitting method for analysis of flood - records. *Eos*
521 *Transactions, American Geophysical Union* 21(2):660-668.
522 <https://doi.org/10.1029/TR021i002p00660>

523 Huang KD, Chen L, Zhou JZ, Zhang JH, Singh VP (2018) Flood hydrograph coincidence
524 analysis for mainstream and its tributaries. *J Hydrol* 565:341-353.
525 <https://doi.org/10.1016/j.jhydrol.2018.08.007>

526 Jane R, Cadavid L, Obeysekera J, Wahl T (2020) Multivariate statistical modelling of the
527 drivers of compound flood events in south Florida. *Natural Hazards and Earth System*
528 *Sciences* 20(10):2681-2699. <https://doi.org/10.5194/nhess-20-2681-2020>

529 Joe H (1996) Families of m-variate distributions with given margins and $m(m+1)/2$ bivariate
530 dependence parameters. *Distributions with Fixed Marginals and Related Topics*
531 (28):120-141

532 Kao SC, Govindaraju RS (2010) A copula-based joint deficit index for droughts. *J Hydrol*
533 380(1-2):121-134. <https://doi.org/10.1016/j.jhydrol.2009.10.029>

534 Karmakar S, Simonovic SP (2009) Bivariate flood frequency analysis Part 2: a copula-based
535 approach with mixed marginal distributions. *Journal of Flood Risk Management*
536 2(1):32-44. <https://doi.org/10.1111/j.1753-318X.2009.01020.x>

537 Klein B, Pahlow M, Hundecha Y, Schumann A (2010) Probability analysis of hydrological
538 loads for the design of flood control systems using copulas. *J Hydrol Eng* 15(5):360-
539 369. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000204](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000204)

540 Cooke R, Kurowicka D (2006) *Uncertainty Analysis with High Dimensional Dependence*
541 *Modelling*. John Wiley & Sons Ltd, New York

542 Li K, Wu S, Dai E, Xu Z (2012) Flood loss analysis and quantitative risk assessment in
543 China. *Nat Hazards* 63(2):737-760. <https://doi.org/10.1007/s11069-012-0180-y>

544 Liu CL, Zhang Q, Singh VP, Cui Y (2011) Copula-based evaluations of drought variations in
545 Guangdong, South China. *Nat Hazards* 59(3):1533-1546.
546 <https://doi.org/10.1007/s11069-011-9850-4>

547 Ministry of Water Resources (MWR) (2006) Regulation for Calculating Design Flood of
548 Water Resources and Hydropower Projects. Beijing: Water Resources and Hydropower
549 Press (in Chinese)

550 Peng Y, Shi YL, Yan HX, Chen K, Zhang JP (2019) Coincidence risk analysis of floods
551 using multivariate copulas: case study of Jinsha River and Min River, China. *J Hydrol*
552 *Eng* 24(2):05018030. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001744](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001744)

553 Prohaska S, Ilic A (2010) Coincidence of flood flow of the Danube River and its tributaries.
554 *Hydrological Processes of the Danube River Basin*:175-226

555 Prohaska S, Ilic A, Majkic B (2008) Multiple-coincidence of flood waves on the main river
556 and its tributaries. *IOP Conference* 4:012013. [https://doi.org/10.1088/1755-](https://doi.org/10.1088/1755-1307/4/1/012013)
557 [1307/4/1/012013](https://doi.org/10.1088/1755-1307/4/1/012013)

558 Prohaska S, Ilic A, Tripkovic V (2012) Methodology for assessing multiple-coincidence of
559 flood wave peaks in complex river systems. *Water Resour Manag* 2(1):45-60

560 Reddy MJ, Ganguli P (2012) Bivariate flood frequency analysis of upper Godavari River
561 flows using Archimedean copulas. *Water Resour Manag* 26(14):3995-4018.
562 <https://doi.org/10.1007/s11269-012-0124-z>

563 Roo APJD, Hazelhoff L, Heuvelink GBM (1992) Estimating the effects of spatial variability
564 of infiltration on the output of a distributed runoff and soil erosion model using Monte
565 Carlo methods. *Hydrol Process* 6(2):127-143. <https://doi.org/10.1002/hyp.3360060202>

566 Salvadori G, Michele CD (2004) Frequency analysis via copulas: Theoretical aspects and
567 applications to hydrological events. *Water Resour Res* 40(12):W12511.
568 <https://doi.org/10.1029/2004WR003133>

569 Schepsmeier U, Czado C (2016) Dependence modelling with regular vine copula models: a
570 case-study for car crash simulation data. *J Roy Stat Soc C-App* 65(3):415-429.
571 <https://doi.org/10.1111/rssc.12125>

572 Schulte M, Schumann A (2016) Evaluation of flood coincidence and retention measures by
573 copulas. *Wasserwirtschaft* 106(2–3):81–87

574 Sklar A (1959) Fonctions de repartition a n dimensions et leurs marges. *Publ Inst Stat Univ*
575 *Paris*:22shou9–231

576 Spurr BD, Koutbeiy MA (1991) A comparison of various methods for estimating the
577 parameters in mixtures of von Mises distribution. *Commun Stat-Simul C* 20(2-3):725-
578 741. <https://doi.org/10.1080/03610919108812980>

579 Sraj M, Bezak N, Brilly M (2015) Bivariate flood frequency analysis using the copula
580 function: a case study of the Litija station on the Sava River. *Hydrol Process* 29(2):225-
581 238. <https://doi.org/10.1002/lt.22388>

582 Stein L, Pianosi F, Woods R (2020) Event - based classification for global study of river
583 flood generating processes. *Hydrol Process* 34(7):1514-1529.
584 <https://doi.org/10.1002/hyp.13678>

585 Su Q (2020) Long-term flood risk assessment of watersheds under climate change based on
586 the game cross-efficiency DEA. *Nat Hazards* 104(3):2213-2237.
587 <https://doi.org/10.1007/s11069-020-04269-1>

588 Tosunoglu F, Faruk G, Spirli MN (2020) Multivariate modeling of flood characteristics using
589 vine copulas. *Environmental Earth Sciences* 79(19):1-21.
590 <https://doi.org/10.1007/s12665-020-09199-6>

591 Try S, Tanaka S, Tanaka K, Sayama T, Oeurng C (2020) Projection of extreme flood
592 inundation in the Mekong River basin under 4k increasing scenario using large
593 ensemble climate data. *Hydrol Process* 34(22):4350-4364.
594 <https://doi.org/10.1002/hyp.13859>

595 Wang HJ, Xiao WH, Wan YC, Zhao Y, Lu F, Yang MZ, Hou BD, Yang H (2019)
596 Assessment of the impact of climate change on hydropower potential in the Nanliujiang
597 River basin of China. *Energy* 167(JAN15):950-959.
598 <https://doi.org/10.1016/j.energy.2018.10.159>

599 Wu ZN, He CT, Wang HL, Zhang Q (2020) Reservoir inflow synchronization analysis for
600 four reservoirs on a mainstream and its tributaries in flood season based on a
601 multivariate copula model. *Water Resour Manag* 34(9):2753-2770.
602 <https://doi.org/10.1007/s11269-020-02572-x>

603 Yan BW, Guo SL, Yu W (2013) Coincidence risk of flood hydrographs between Yangtze
604 River and Qing River. *Journal of Hydroelectric Engineering* 32(1):50-53 (in Chinese)

605 Yang J, Wang Y, Yao J, Chang J, Xu G, Wang X, Hu H (2020) Coincidence probability
606 analysis of hydrologic low-flow under the changing environment in the Wei River
607 Basin. *Nat Hazards* 103(2):1711-1726. <https://doi.org/10.1007/s11069-020-04051-3>

608 Yu KX, Zhang X, Li P, Li ZB, Qin Y, Sun Q (2019) Probability prediction of peak break-up
609 water level through vine copulas. Hydrol Process 33(6):962-977.
610 <https://doi.org/10.1002/hyp.13377>

611 Zhang C, Peng Y, Ji CM, Shi YL (2020) Floods encountering risk analysis for upper Yangtze
612 River and Dongting Lake. Journal of Hydroelectric Engineering 39(8):55-68 (in
613 Chinese) <https://doi.org/10.11660/slfdb.20200806>

614 Zhang L, Singh VP (2007) Bivariate rainfall frequency distributions using Archimedean
615 copulas. J Hydrol 332(1-2):93-109. <https://doi.org/10.1016/j.jhydrol.2006.06.033>

616 Zhang XT, Shao J, Guo W (2018) Analysis of flood area composition and encounter law for
617 Yalong River and Chuanjiang River. Yangtze River 49(22):23-28 (in Chinese)
618 <https://doi.org/10.16232/j.cnki.1001-4179.2018.22.005>

619 Zhu LT, Liu YC, Yan FJ, Chen LF, Li GS (2015) Flood coincidence probability analysis for
620 Nansi Lake valley. Transactions of Oceanology and Limnology 37(1):149-154 (in
621 Chinese) <https://doi.org/10.13984/j.cnki.cn37-1141.2015.01.021>

Figures

Figure 1

Map showing the geographic setting of the Yangtze River and the location of the hydrological stations within the study area. (ArcGIS was used to create this artwork) Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its

authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

Figure 2

Schematic diagram illustrating how AMF characteristics were defined. (Excel XL Toolbox was used to create this artwork)

Figure 3

Schematic diagram of flood hydrograph coincidence. (Visio was used to create this artwork)

Figure 4

General structures of five-dimensional C-vine and D-vine. (Visio was used to create this artwork)

Figure 5

Developed frequency curves of AMF volumes for the Yichang and Chenglingji stations and their comparison to empirical data. (Excel XL Toolbox was used to create this artwork)

Figure 6

Developed frequency curves of AMF volumes for the Yichang and Chenglingji stations and their comparison to empirical data. (Excel XL Toolbox was used to create this artwork)

Figure 7

Probability densities of AMF interval times for the Yichang and Chenglingji stations and their comparison to empirical data. (Excel XL Toolbox was used to create this artwork)

Figure 8

Scatter diagrams of the simulated and the historical sequences during Period 1. (Excel XL Toolbox was used to create this artwork)

Figure 9

Scatter diagrams of the simulated and the historical sequences during Period 2. (Excel XL Toolbox was used to create this artwork)

Figure 10

Validation of AMF hydrograph coincidence risk probability. (Excel XL Toolbox was used to create this artwork)

Figure 11

The AMF hydrograph coincidence volume-risk probability curves. (Excel XL Toolbox was used to create this artwork)