Comparative Study of Flood Coincidence Risk in The Mainstream and Its Tributaries

Na Li
Wuhan University
Shenglian Guo (slguo@whu.edu.cn)
Wuhan University
Feng Xiong
Wuhan University
Jun Wang
Wuhan University

Research Article

Keywords: Flood coincidence, Copula function, Conditional probability, Comparative study, Fuhe River

Posted Date: June 25th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-644829/v1

Comparative study of flood coincidence risk in the mainstream and its tributaries

Na Li, Shenglian Guo*, Feng Xiong and Jun Wang

State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China

*slguo@whu.edu.cn

Abstract

The coincidence of floods in the mainstream and its tributaries may lead to a large flooding in the downstream confluence area, and the flood coincidence risk analysis is very important for flood prevention and disaster reduction. In this study, the multiple regression model was used to establish the functional relationship among flood magnitudes in the mainstream and its tributaries. The mixed von Mises distribution and Pearson Type III distribution were selected to fit the probability distribution of the annual maximum flood occurrence dates and magnitudes, respectively. The joint distributions of the annual maximum flood occurrence dates and magnitudes were established using copula function. Fuhe River in the Poyang Lake region was selected as a study case. The joint probability, co-occurrence probability and conditional probability of flood magnitudes were calculated and compared with the simulated results of the observed data. The results show that the selected marginal and joint distributions perform well in simulating the observed flood data. The coincidence probabilities of flood occurrence dates in the upper mainstream and its tributaries mainly occur from May to early July. Among the three coincidence probability calculation methods, the conditional probability is the most consistent with the flood coincidence risk in the mainstream and its tributaries, which is more reliable and rational in practice. The results reflect the actual flood coincidence situation in the Fuhe River basin and can provide technique support for flood control decision-making.

Keywords: Flood coincidence; Copula function; Conditional probability; Comparative study; Fuhe River

1 Introduction

Nowadays, flood problems have become more and more prominent, which account for a large part of all-natural hazards in the world (KvočKa et al. 2016). In addition to causing severe disasters to agriculture, floods can also bring loss to the industry, life and property (de Bruijn et al. 2015; Thieken et al. 2015). Affected by human activities, climate change, environmental degradation and El Niño, extreme hydrological events occur frequently worldwide (Hirabayashi et al. 2013; Alfieri et al. 2016; Zhang et al. 2016). Against this background, the frequency and intensity of floods continue to increase, and the resulting losses are also growing (Ceola et al. 2014; Daksiya et al. 2020). Generally, large floods are caused by the combination of floods in the mainstream and its tributaries (Ganguli and Reddy 2013). When floods occur simultaneously, the flood peaks and the flood volumes will superimpose into large floods, threating the safety of the downstream river (Chen et al. 2012; Wang 2016). Therefore, it is of great significance to study the flood coincidence laws in the mainstream and its tributaries, which can not only provide a theoretical basis for the formulation of flood control and dispatching plans in the basin, but also offer a reference for the construction of flood control facilities in the downstream.

For the analysis of flood coincidence, the traditional method is to statistically analyze flood coincidence events occurred based on the synchronized flood data over the years, so as to calculate the corresponding coincidence probabilities. However, the traditional hydrological statistical analysis method only focuses on historical data, and cannot quantitatively estimate the coincidence probabilities and return periods of design floods at specific frequencies. As cascade reservoirs continue to be built and put into operation, flood coincidence analysis is particularly important in joint flood control and dispatching work, but the traditional method cannot provide sufficient information. In fact, the essence of flood coincidence is a multivariable frequency combination event, which can be studied by the multivariable hydrological analysis method (Feng et al. 2020). At present, the commonly used multivariate hydrological analysis methods include the nonparametric method (Silverman 1986; Kim et al. 2006), the specific joint distribution method (Bacchi et al. 1994; Yue 2000a, 2002; Escalante 2007; Shimizu 2010), the multivariate Normal distribution method

(Goel et al. 1998; Yue 2000b; Prohaska and Ilic 2010), the FEI method of transforming the multi-dimensional joint distribution into one-dimensional distribution, and the empirical frequency method. However, the above methods have certain limitations and deficiencies. For example, the nonparametric method cannot give an analytical formula for the marginal distribution of variables, and the data prediction ability is weak. The specific joint distribution method requires the marginal distributions to be the same type (Shiau 2006). The multivariate Normal distribution method is prone to cause information distortion in the process of data conversion (Correia 1978). And the empirical frequency method does not have the ability to predict data extension.

Copula functions overcome the shortcomings of the traditional methods, and can connect arbitrary marginal distributions through correlation structures. As an effective method of constructing multivariate joint distribution, the advantages of copula functions are: (1) arbitrary marginal distribution types, (2) flexible and diverse structures, (3) simple solution method, (4) strong applicability and wide scalability. In recent years, copula functions have become a research hotspot in the field of hydrology, and been widely used in multivariate hydrological analysis. For example, they have been used for flood frequency analysis (Salvadori and Michele 2004; Zhang and Singh 2006; Reddy and Ganguli 2012; Li et al. 2013; Sraj et al. 2015; Zhong et al. 2018; Karahacane et al. 2020), rainfall frequency analysis (De Michele and Salvadori 2003; Kao and Govindaraju 2007, 2008; Ashkar and Aucoin 2011; Zhang and Singh 2012), rain and flood analysis (Xiao et al. 2009; Keef et al. 2009; Candela et al. 2014), and multivariate simulation (Aghakouchak et al. 2010a, b; Chen et al. 2015, 2016; Poduje and Haberlandt 2017).

Copula functions have also been applied in flood coincidence analysis. For example, Favre et al. (2004) used a copula function to construct the joint distribution of floods in a mainstream and its tributary, and calculated the flood coincidence probability. Wang et al. (2009) presented a Copula-based Flood Frequency (COFF) model with arbitrary marginal distributions to evaluate quantitatively flood risk at confluences. Klein et al. (2010) estimated coincidence probability of flood volumes at two reservoirs in a river basin using copula functions. Schulte and Schumann (2016) developed multivariate copula-approaches to analyze coincidence risk of flood peaks in adjoining catchment. Using Copula Monte Carlo (CMC) method, Peng et al. (2017) further estimated flood risk in the confluence flood control downstream of a reservoir. However, these researches only considered flood magnitudes and ignored flood occurrence time. In fact, flood coincidence means that the simultaneous occurrence of large floods in different rivers. It needs to meet two conditions: the flood occurrence time should be within a certain range, and the flood magnitudes should be above a certain level. Therefore, when analyzing flood coincidence risk, both factors of flood occurrence time and magnitudes should be taken into consideration. Recently, assuming that the flood occurrence dates and magnitudes were independent, Chen et al. (2012) selected the multi-dimensional asymmetric Archimedean copula functions to analyze the flood coincidence risk of the upstream Yangtze River and its tributaries. Peng et al. (2019) employed multivariate copulas to estimate flood coincidence probabilities, considering flood occurrence dates and magnitudes simultaneously. Huang et al. (2018) took flood magnitudes of two rivers and flood occurrence interval dates as three reference variables, and further explored the flood hydrograph coincidence risk using copulas.

The above researches revealed the characteristics of flood coincidence risk from different angles, and have made great progress in the flood coincidence analysis of mainstream and its tributaries. However, some researches only focused on the coincidence risk of flood magnitudes and neglected the flood occurrence time; other researches usually assumed that the flood magnitudes and occurrence time data series were independent and ignored the correlation of flood variables. In addition, most studies were limited to constructing distribution models to quantitatively evaluate the flood coincidence risk, and hardly compared with the actual situations, which cannot guarantee the rationality and reliability of the analysis results.

The objective of this study is therefore to analyze flood coincidence risk in the mainstream and its tributaries considering the link of the up-downstream flood variables. Based on copula function, the joint distributions of flood magnitudes and occurrence dates are constructed for the flood coincidence analysis.

Daily flow data series at three hydrologic stations in the Fuhe River are chosen as a case study. First, the multivariate regression model is used to simulate the functional relationship among the flood variables of the mainstream and its tributaries. Second, the mixed von Mises and Pearson Type III marginal distributions are used to describe the annual maximum flood occurrence dates and flood magnitudes, respectively. Third, the coincidence probabilities of flood occurrence dates and magnitudes are estimated. Finally, the estimated and simulated flood coincidence risks are compared and assessed.

2 Study area and data

The Fuhe River, located at the east of Jiangxi Province and feeding into the Poyang Lake, was selected as a case study. Fig. 1 depicts the distributions of the mainstream and tributaries of the Fuhe River and related hydrological stations. Affected by the subtropical humid monsoon climate, there are abundant rainfall in the area (Wang et al. 2013). Floods are mostly formed by heavy rainstorms, which temporal and spatial distributions are consistent with heavy rains. And the flood season is generally from April to early July. Flood events occur frequently in the Poyang Lake area, with an average of 4 years every 5 years. Among the most recent major flood events are those of 1954, 1983, 1995, 1998, 1999 and 2010. One of the main causes of flood disasters in the Poyang Lake area is that the flow from the five main rivers is too large and the water level is too high, leading the floods to overflow the levee or break it. On the other hand, the river networks are huge and floods are prone to encounter, which add to the severity of flood damage.

As one of the five main rivers in the Poyang Lake water system, the Fuhe River is 348 km long and has a drainage area of 164,93 km^2, which is the second largest river in Jiangxi Province. The Liaojiawan hydrological station with a control basin area of 8723 km^2, and the Loujiacun hydrological station with a control basin area of 4969 km^2, are located in the upper mainstream and tributary, respectively. The Lijiadu hydrological station with a catchment area of 15,812 km^2 is located in the down-mainstream, which accounts for more than 95% of the entire drainage area. In this study, the daily flow discharge data series of these three hydrological stations from 1953 to 2016 were collected and the annual maximum flood magnitudes and corresponding occurrence dates were sampled.

[Insert Figure 1 about here]

3 Methodology

3.1 Copula functions

Copula function is a multi-dimensional joint distribution function uniformly distributed in the domain of [0, 1]. It can connect arbitrary marginal distributions through correlation structures to construct multi-dimensional joint distributions (Joe 1997; Nelsen 2006). Based on Sklar's theorem (Schweizer and Sklar 1983), assuming that the marginal distribution functions of random variables *x* and *y*, are $F_X(x)$ and $F_Y(y)$ respectively, and $F(x, y)$ is their joint distribution, then the copula function can be written as:

$$F(x, y) = C_\theta\left(F_X(x), F_Y(y)\right) = C_\theta(u, v) \tag{1}$$

where $C_\theta(\cdot)$ is a copula function; θ is a parameter of the copula function to be estimated; u and v are marginal distribution functions, satisfying $u = F_X(x) = P_X(X \le x)$, and $v = F_Y(y) = P_Y(Y \le y)$.

Recently, the Archimedean copula functions have been widely used in hydrological frequency analysis (Grimaldi and Serinaldi 2006; Leonard et al. 2008; Guo et al. 2018; Yin et al. 2018). Among them, the Clayton copula (Clayton 1978), Gumbel-Hougaard copula (GH copula) (Hougaard 1986) and Frank copula (Frank 1979) are the most commonly used in practice. Because they just have one parameter and are easy to generate and solve. What's more, they can be used to describe hydrological variables with positive or negative correlations (Nelsen 2006). The mathematical expressions are shown below:

Clayton copula:
$$C(u, v) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-1/\theta}; \theta \in (0, \infty) \tag{2}$$

GH copula:
$$C(u, v) = \exp\left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right\}; \theta \in [1, \infty) \tag{3}$$

Frank copula：

$$C(u,v) = -\frac{1}{\theta}\ln\left[1+\frac{\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)}{\left(e^{-\theta}-1\right)}\right];\theta \in R \tag{4}$$

The parameter estimation methods of copula function mainly consist of the correlation index method, the line-of-fit method and the maximum likelihood method. In this study, the correlation index method was employed. Based on the relationship between the copula parameter θ and the two variables' Kendall correlation coefficient τ , the parameters of copula functions can be estimated by following formula:

$$\tau = \begin{cases} \theta/(\theta+2), & \text{Clayton} \\ 1-1/\theta, & \text{GH} \\ 1-4/\theta+4D_1(\theta)/\theta, & \text{Frank} \end{cases} \tag{5}$$

where τ is the Kendall correlation coefficient of two variables; θ is a parameter of the copula function; and $D_1(\cdot)$ is the first-order Debye function.

3.2 Marginal distributions of flood occurrence dates and magnitudes

The flood occurrence dates often have the characteristics of periodicity. The von Mises distribution has a good fitting effect for the distribution of periodic or seasonal variables with a single peak (Fisher 1993; Mardia and Jupp 2009). In general, the annual maximum flood is affected by many factors, so its occurrence dates series may be multi-peaked. In this situation, a mixed von Mises distribution which comprises m von Mises distributions, can be applied to describe the probability density function of multi-peaked variables. The probability density function of the mixed von Mises distribution can be written as:

$$f_X(x) = \sum_{i=1}^{m}\frac{p_i}{2\pi I_0(k_i)}\exp^{\left[k_i\cos(x-u_i)\right]};0\le x\le 2\pi, 0\le u_i\le 2\pi, k_i>0 \tag{6}$$

where p_i is the coefficient of the mixing proportion; k_i is the scale parameter; u_i is the position parameter; $I_0(k_i)$ is the 0-order modified Bessel function; and *m* is the order of the finite mixed von Mises distribution (*m*=2). The maximum likelihood estimate (MLE) method is frequently used to calculate the parameters in Eq. (6) (Michael and Stanislav 2013).

For the annual maximum flood series, many distributions including the Pearson Type III (P3) distribution, Log-Pearson Type III distribution, Gamma-type distribution, Generalized Extreme Value (GEV) distribution, and Lognormal distribution can be used to describe the probability density function. In China, the Pearson Type Ⅲ distribution has been recommended by the Chinese Ministry of Water Resources (MWR 2006) as a uniform procedure for flood frequency analysis. Therefore, assuming that the annual maximum flood magnitudes obey the P3 distribution, and its probability density function is expressed as:

$$f_X(x)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}(x-\delta)^{\alpha-1}\exp\left[-\beta(x-\delta)\right];\alpha>0,\beta>0,\delta\le x<\infty \tag{7}$$

where α, β and δ are the shape, scale and position parameters of the P3 distribution, respectively; and $\Gamma(\cdot)$ is the gamma function. The parameters of the P3 distribution can be estimated by the L-moment method (Hosking 1990).

3.3 Goodness-of-fit tests

The evaluation of goodness-of-fit is a very important step in the process of selecting the marginal distribution and the joint distribution. In order to judge whether the selected distribution is appropriate and whether it can accurately reflect the actual distribution of the sample, it is necessary to perform a fitting test and goodness evaluation. There are many methods for goodness-of-fit test in hydrological analysis, and the commonly used Root Mean Square Error (RMSE), Kolmogorov-Smirnov (K-S) and chi-square (χ^2) test methods were selected to evaluate the comprehensive performance of the distributions.

To test the goodness-of-fit of the marginal distributions and the joint distributions, the empirical probabilities of the samples should be obtained first. For the univariate series, the empirical probabilities of each flood variable generally can be obtained by the Weibull plotting position formula (Makkonen 2008), which can be written as:

$$P_e(x_i)=\frac{i}{n+1} \tag{8}$$

where x_i is the observed data; n is the sample length; and $P_e(x_i)$ is the empirical exceedance probability.

For the bivariate series, the empirical probabilities of joint distribution can be estimated using the Gringorten formula (Gringorten 1963), which has been widely applied in extreme flood events (Hirsch and Stedinger 1987; Yue 1999; Zhang and Singh 2007; Karmakar and Simonovic 2009; Xiong et al. 2019). The specific formula is as follows:

$$P_e(x_i, y_i) = \frac{\sum_{j=1}^{n}\left(X_j \le x_i, Y_j \le y_i\right) - 0.44}{n + 0.12} \tag{9}$$

where (x_i, y_i) is a combination of the observed data; n is the sample length; and $P_e(x_i, y_i)$ is the empirical joint distribution probability.

The Root Mean Square Error (RMSE) is often selected to measure the difference between the theoretical probabilities of the fitted distribution and the empirical probabilities of the observed data. RMSE can effectively evaluate the performance of the fitted distributions. The smaller the RMSE value, the better the fitting effect. The RMSE value can be obtained as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(P_{ei} - P_i\right)} \tag{10}$$

where n is the sample size; P_i is the theoretical probabilities obtained from the fitted distribution; P_{ei} is the empirical frequencies from the observed data.

The Kolmogorov-Smirnov (K-S) test is a goodness-of-fit test method that analyzes the distance between the empirical distributions and the theoretical distributions (Massey 1951; Weiss 1978; Razali and Wah 2011). It can judge whether the observed data of the sample obey the fitted distribution. For n observed data which are in an increasing order, the K-S test statistic is expressed as:

$$D_n = \sup_x \left|F_n(x) - F^*(x)\right| \tag{11}$$

where $F^*(x)$ is the theoretical distribution; $F_n(x)$ is the empirical distribution, $\sup_x$ is the maximum value of distances.

The chi-square (χ^2) test is to measure the degree of deviation between the observed values and the predicted values. The size of the chi-square value determines the degree of deviation. The chi-square test statistic is defined as:

$$\chi^2 = \frac{\left(M_1 - np_1\right)^2}{np_1} + \frac{\left(M_2 - np_2\right)^2}{np_2} + \cdots + \frac{\left(M_k - np_k\right)^2}{np_k} \tag{12}$$

where $p_1, p_2, \ldots, p_k$ are the hypothesized probabilities for *k* possible outcomes; and $M_1, M_2, \ldots, M_k$ are the observed counts of each outcome to be compared for expected counts $np_1, np_2, \ldots, np_k$ in *n* independent trails.

3.4 Estimation of flood coincidence risk

Flood coincidence means large floods in the mainstream and its tributaries occur simultaneously. In general, the probabilities are used to quantitatively describe the degree of flood coincidence. According to the definition of flood coincidence, it is obvious that flood events are mainly characterized by flood occurrence dates and flood magnitudes. Thus, both factors should be taken into consideration when evaluating flood coincidence. In this study, the flood occurrence dates and flood magnitudes were selected as reference variables.

3.4.1 Coincidence risk of flood occurrence dates

Considering flood occurrence dates, the coincidence of flood occurrence dates refers to that the annual maximum floods in the mainstream and its tributaries occur on the same day. Therefore, the coincidence probabilities of the annual maximum flood occurrence dates of two rivers on the *k*th day can be defined as:

$$\begin{aligned} P_k^t &= P\left(t_k < T_i \le t_{k+1}, t_k < T_j \le t_{k+1}\right) \\ &= F_T\left(t_k,\ t_k\right) + F_T\left(t_{k+1},\ t_{k+1}\right) - F_T\left(t_k,\ t_{k+1}\right) - F_T\left(t_{k+1},\ t_k\right) \end{aligned} \tag{13}$$

where *i* and *j* are the hydrological stations on the mainstream and its tributary; T_i represents the occurrence dates of the annual maximum flood, expressed as a certain day of the flood season; and t_k represents the *k*th day of the flood season.

3.4.2 Coincidence risk of flood magnitudes

Considering flood magnitudes coincidence, we can establish joint distributions of the annual maximum floods in the mainstream and its tributaries based on copula functions. Then, the flood coincidence risk can be quantitatively evaluated using the joint probabilities, co-occurrence probabilities and conditional probabilities.

For the joint probabilities of flood magnitudes coincidence, it refers to the probabilities that at least one of two rivers occurs floods surpassing certain values, expressed as the following equation:

$$P(X > x \cup Y > y) = 1 - F(x, y) = 1 - C(u, v) \tag{14}$$

where *x* and *y* are the flood magnitudes in *i* and *j* river, respectively; and $F(x, y)$ is the joint distribution function of flood magnitudes in two rivers.

For the co-occurrence probabilities of flood magnitudes coincidence, it refers to the probabilities that two rivers simultaneously occur floods surpassing certain values, expressed as the following equation:

$$\begin{aligned} P(X > x \cap Y > y) &= 1 - F_X(x) - F_Y(y) + F(x, y) \\ &= 1 - u - v + C(u, v) \end{aligned} \tag{15}$$

For the conditional probabilities, it refers to the probabilities that when given range for any one variable, the other variable fall into another range. In this study, under the condition that one river has occurred floods surpassing certain values, the probability that another river also occurs floods surpassing certain values, can be expressed as the following equation:

$$\begin{aligned} P(Y > y | X > x) &= \frac{P(X > x \cap Y > y)}{P(X > x)} \\ &= \frac{1 - F_X(x) - F_Y(y) + F(x, y)}{1 - F_X(x)} = \frac{1 - u - v + C(u, v)}{1 - u} \end{aligned} \tag{16}$$

4 Results and discussion

4.1 Flood correlation analysis

There are many methods to measure the correlation between variables in hydrological analysis, and the common Pearson correlation coefficient was used to describe the correlation between the annual maximum flood magnitudes series of these three hydrological stations in the Fuhe River. Fig. 2(a) shows the correlation analysis of the annual maximum flood magnitudes between the Liaojiawan, Loujiacun stations in the upstream and the Lijiadu station in the downstream, respectively. It can be seen that the Pearson correlation coefficient between the Liaojiawan and Lijiadu stations is 0.93, and between the Loujiacun and Lijiadu stations is 0.90. There are strong positive correlation in the annual maximum flood magnitudes between these stations. Since the control basin area of the Loujiacun station is smaller than that of the Liaojiawan station, its correlation with the Lijiadu station is relatively small.

Fig. 2(b) shows the correlation analysis of the annual maximum flood magnitudes between the Liaojiawan and Loujiacun stations which are both located in the upper reaches of the Fuhe River. It can be seen that the Pearson correlation coefficient is 0.80, and their control basins are very close in space. Therefore, the key factors for flood generation, such as climatic conditions, geographical environment and topographical position, are very similar. It also proves that the regularity of flood occurrence in the upper reaches is highly consistent.

In order to analyze the quantitative relationship among the annual maximum flood magnitudes series of the three stations, the multiple regression model was employed. The functional relationship among the annual maximum flood magnitudes of these three stations was constructed, and the result is as follows:

$$Y = \hat{Y} + \varepsilon = 104.890 + 0.841X_1 + 1.082X_2 + \varepsilon \tag{17}$$

where X_1 and X_2 represent the observed values of the annual maximum flood magnitudes at the Liaojiawan and Loujiacun stations, respectively; $\hat{Y}$ represents the simulated values of the annual maximum flood magnitudes at the Lijiadu station; ε is the random error of the model, obeying a Normal distribution with a mean value of 0, which can be expressed as $\varepsilon \sim N\left(0,\sigma_{\varepsilon}^{2}\right)$.

To evaluate the performance of the multiple regression equation, the relationship between the simulated values of the annual maximum flood magnitudes obtained by the multiple regression model and the observed values at the Lijiadu station, is plotted in Fig. 3. It can be seen that the scattered points are basically distributed along the 45° diagonal, indicating the simulated values and the observed values are very close. It further proves that the floods at the Lijiadu station are mainly formed by the superposition of the upstream floods, and the simultaneous occurrence of floods in the upstream are easy to cause large flooding in the downstream confluence area.

[Insert Figure 2 about here]

[Insert Figure 3 about here]

4.2 Estimation of marginal distributions

The mixed von Mises distribution was used to fit the occurrence dates series of the annual maximum flood at the three stations in the Fuhe River, and the parameters of the marginal distributions were estimated by the maximum likelihood method. Table 1 lists the values of the estimated parameters of the von Mises function. The empirical frequencies of the marginal distributions were obtained by Eq. (8), and the theoretical probabilities of the mixed von Mises distribution were calculated by Eq. (6). Fig. 4 shows the fitting relationship between the empirical frequency points and the theoretical probability curve for the annual maximum flood occurrence dates, in which the mixed von Mises distribution can fit the empirical distribution very well.

In order to evaluate the performance of the marginal distributions more comprehensively, the K-S test method was used for the fitting test, and the Root Mean Square Error method was applied for the goodness evaluation. The results of the goodness-of-fit evaluation are presented in Table 1. It is shown that at a significance level of 0.05, the values of K-S test statistics do not exceed their critical values, implying that all the marginal distributions of the annual maximum flood occurrence dates have passed the hypothesis test. Meanwhile, the RMSE values between the theoretical frequencies and the empirical probabilities are small enough. Therefore, the mixed von Mises distribution performs well in simulating the annual maximum flood occurrence dates in the Fuhe River.

The P3 distribution was applied to describe the annual maximum flood magnitudes series of the three stations, and the parameters were estimated by the L-moment method. The values of the estimated parameters of the P3 distribution are presented in Table 2. The empirical frequencies and the theoretical probabilities were calculated by Eqs. (8) and (7), respectively. The fitting curves of the marginal distributions are shown in Fig. 4. It can be seen that graphically, the P3 distribution fits the empirical distributions well. In addition, the chi-square test and the Root Mean Square Error method were selected for the fitting test and goodness evaluation. Table 2 lists the relative evaluation index results. It is shown that the p-values of the chi-square test of the three stations are all larger than the critical values at the 0.05 level of significance, implying that the hypothesis test that the flood magnitudes obey the P3 distribution is not rejected. So, the P3 distribution performs well in quantifying the marginal distributions of the annual maximum flood magnitudes.

[Insert Table 1 about here]

[Insert Table 2 about here]

[Insert Figure 4 about here]

4.3 Estimation of joint distributions

The dependence of the annual maximum floods among these three stations have been verified. Thus, the copula functions can be used for the flood coincidence analysis. Based on the Clayton, GH and Frank copula functions, the bivariate joint distribution of the annual maximum flood occurrence dates at the Liaojiawan and Loujiacun stations was constructed. Similarly, the two-dimensional joint distributions of the annual maximum flood magnitudes between the Liaojiawan, Loujiacun and Lijiadu stations were also constructed, respectively. The Kendall rank correlation coefficient was employed to reckon the parameters of the

Archimedean copula functions. Table 3 presents the values of the estimated parameters. Using the Eq. (9), the empirical frequencies of the joint distributions were calculated. In order to select the most appropriate copula function among the three candidate copulas, the Root Mean Square Error method was applied to test the goodness of fitting. The RMSE values between the empirical frequencies and the theoretical probabilities are listed in Table 3. The results show that for the annual maximum flood occurrence dates and flood magnitudes, the RMSE values of the Clayton copula function are the lowest. It indicates that the Clayton copula function is the most appropriate copula for modeling the joint distributions of flood variables at these stations of the Fuhe River. Therefore, the Clayton copula function is selected to establish the joint probability distributions of the annual maximum flood occurrence dates and magnitudes, respectively. The theoretical and the observed nonexceedance joint probabilities are exhibited in Fig. 5, in which the x-axis is sorted in ascending order of the theoretical nonexceedance joint probabilities. It can be seen that the theoretical frequency curves can fit the observed values well.

[Insert Table 3 about here]

[Insert Figure 5 about here]

4.4 Analysis of flood coincidence risk

4.4.1 Coincidence risk of flood occurrence dates

With Eq. (13), the coincidence probabilities of the occurrence dates of the annual maximum flood at the Liaojiawan and Loujiacun stations were calculated and the results are presented in Fig. 6. The curve demonstrates that the coincidence probabilities of flood occurrence dates in the mainstream and tributaries of the upper Fuhe River present the characteristics of multiple-peak. Before March and after August, the coincidence probabilities are very small, which are basically close to zero. During the period from mid-March to late April, there are a relatively stable coincidence risk, which are almost below 0.015%. Early May to early July is a higher coincidence period of the flood occurrence dates, including two peaks. The first coincidence peak occurs on May 12, with an associated probability of 0.026%. The second peak with an associated probability of 0.057% occurs on June 21, which is the largest coincidence risk. According to the analysis of climate, it can be known that the Fuhe River is located in the subtropical monsoon climate zone, and the flood season is from April to July. During this period, the precipitation is frequent and the precipitation intensity is high, which is easy to cause floods. Moreover, floods have obvious characteristics that change with the seasons. The annual flow is mainly distributed in the flood season, and most of the large floods occur in individual months of the flood season, such as May, June and July. From the beginning to the end of the flood season, the flood intensity changes from weak to strong, and then from strong to weak. Therefore, May to early July are the overlapping periods of floods in the mainstream and tributaries. As shown in Fig. 6, the higher coincidence probabilities of the occurrence dates of the annual maximum flood occur in May to early July, which proves that the calculated results are in accordance with the actual situation. Adding together all the daily coincidence probabilities of the flood occurrence dates, expressed as $P^t = \sum_{k=1}^{n} P_k^t$, we can obtain the coincidence probability of the annual maximum flood occurrence dates in the mainstream and tributaries, and the result is 2.87%.

[Insert Figure 6 about here]

4.4.2 Coincidence risk of flood magnitudes

The coincidence probabilities of the annual maximum flood magnitudes for different design floods in the mainstream and its tributaries were estimated. With Eqs. (14) and (15), the joint probabilities and the co-occurrence probabilities of *T*-year design floods at the Liaojiawan and Loujiacun stations were calculated, respectively. With Eq. (16), the conditional probabilities of the occurrence of *T*-year design floods at the Lijiadu station, when given the occurrence of floods at the Liaojiawan station were obtained. In the same manner, the conditional probabilities of *T*-year design floods occurring at the Lijiadu station, given floods at the Loujiacun station were also calculated. Tables 4 and 5 display the flood coincidence probabilities including the joint probabilities, co-occurrence probabilities, and conditional probabilities for 5, 10, 20, 50, and 100-year design floods in the Fuhe River basin.

It can be seen that the joint probabilities of the occurrence of 5, 10, 20, 50 and 100-year design floods at the Liaojiawan and Loujiacun stations are 30.48%, 17.14%, 9.20%, 3.86%, and 1.96%, respectively. The co-occurrence probabilities for different design floods with the return periods of 5, 10, 20, 50 and 100 years at the Liaojiawan and Loujiacun stations are 9.52%, 2.86%, 0.80%, 0.14% and 0.04%, respectively. Obviously, the joint probabilities are greater than the co-occurrence probabilities for different design floods. The results show that as the return periods increase, the joint probabilities and co-occurrence probabilities of flood magnitudes coincidence decrease. That is, small and medium floods are more likely to occur simultaneously, which conforms to the general law in practice.

Given the occurrence of 5, 10, 20, 50, and 100-year design floods at the Liaojiawan station, the conditional probabilities of the same flood magnitudes occurring at the Lijiadu station are 16.41%, 5.01%, 1.46%, 0.26%, and 0.07%, respectively. Similarly, given the occurrence of 5, 10, 20, 50, and 100-year design floods at the Loujiacun station, the conditional probability of the same flood magnitudes occurring at the Lijiadu station are 15.21%, 4.47%, 1.26%, 0.22%, and 0.06%, respectively. It can be seen that floods with lower return periods result in higher coincidence probabilities. The conditional probabilities between the Liaojiawan and Lijiadu stations are higher than that between the Loujiacun and Lijiadu stations. With reference to the previous analysis, the drainage area controlled by the Liaojiawan station is larger than that of the Loujiacun station, the correlation of flood variables between the Liaojiawan and Lijiadu stations is stronger. Therefore, the floods at the Liaojiawan station have more significant impact on the downstream floods. From the point of view, the calculated results are reasonable.

[Insert Table 4 about here]

[Insert Table 5 about here]

4.5 Comparison of simulation results

Fig. 2 provides the evidence that the flood variables at the Liaojiawan and Loujiacun stations are highly correlated and consistent. Therefore, when floods simultaneously occur in the upper stream, the flood frequencies in the mainstream and its tributaries are more likely to be the same. In this study, assuming that floods with the same frequencies simultaneously occur at the Liaojiawan and Loujiacun stations. Then based on the constructed multiple regression model, expressed as Eq. (17), the design flood at the Lijiadu station can be obtained. Flood magnitudes obey the P3 distribution, so the corresponding design frequencies can be inferred using Eq. (7). And the return periods are calculated as well. Table 6 lists: (1) the design flood values at the Liaojiawan and Loujiacun stations under the return periods of 5, 10, 20, 50, and 100 years; (2) the simulated flood values at the Lijiadu station, which are obtained by the regression model after the same frequency floods occurring in the two upstream stations; (3) the corresponding frequencies and return periods of the simulated floods at the Lijiadu station. The results show that with the rise of flood magnitudes at the Liaojiawan and Loujiacun stations, the floods at the Lijiadu station are increasing. The flood return periods of the Lijiadu station are generally longer than those of the Liaojiawan and Loujiacun stations, especially when the return periods are larger. For example, when 10-year design floods simultaneously occur at the Liaojiawan and Loujiacun stations, an 11-year flood may occur at the Lijiadu station. However, when 100-year floods simultaneously occur at the Liaojiawan and Loujiacun stations, there may be a 137-year flood at the Lijiadu station.

In order to verify the rationality and feasibility of the above three flood coincidence probability calculation methods, the theoretically calculated coincidence probabilities were compared with the design frequencies of the simulated flood at the Lijiadu station. Table 7 lists five conditions under different return periods, including: (1) the design frequencies of the simulated floods at the Lijiadu station when floods with the same frequency simultaneously occurring at the Liaojiawan and Loujiacun stations, (2) the joint probabilities and the co-occurrence probabilities of the same frequency floods occurring at the Liaojiawan and Loujiacun stations, (3) the conditional probabilities of the same frequency floods occurring at the Lijiadu station, given the occurrence of T-year floods at the Liaojiawan station, named P_1^c; (4) the conditional probabilities of the same frequency floods occurring at the Lijiadu station, given the occurrence of *T*-year floods at the Loujiacun station, named P_2^c. Meanwhile, these five flood probabilities are plotted in Fig. 7.

The results show that compared with the design frequencies of the simulated flood at the Lijiadu station, the joint probabilities of floods simultaneously occurring in the upper mainstream and its tributaries are generally greater, while the co-occurrence probabilities and both conditional probabilities are less. Among the four flood coincidence probabilities, the conditional probabilities (P_1^c) are the closest to the design frequencies of the simulated floods, especially when the return periods are low. For example, the joint probability and the co-occurrence probability of 10-year floods at the Liaojiawan and Loujiacun stations are 17.14% and 2.86%, respectively. At this time, the design frequency of the simulated flood at the Lijiadu station is 9.41%. The conditional probability of a 10-year flood occurring at the Lijiadu station given a 10-year flood at the Liaojiawan station is 5.01%. The conditional probability of a 10-year flood occurring at the Lijiadu station given a 10-year flood at the Loujiacun station is 4.47%. Combined with the previous analysis, we can see that flood with the same frequency simultaneously occurring at the Liaojiawan and Loujiacun stations can lead to large floods at the Lijiadu station. The corresponding design frequencies of the simulated floods are more consistent with the conditional probabilities that the same frequency floods occurring at the Lijiadu station when given *T*-year floods at Liaojiawan station. In practice, the floods at the Lijiadu station are basically formed by the superposition of the floods in the upper mainstream and its tributaries. More importantly, the floods at the Liaojiawan station account for a larger proportion in that of the Lijiadu station. Therefore, the results conform to the general law.

[Insert Table 6 about here]

[Insert Table 7 about here]

[Insert Figure 7 about here]

5 Conclusions

Flood coincidence risk analysis plays an important role in reservoir operation and flood management. In this study, the coincidence probabilities of flood magnitudes and occurrence dates in the mainstream and tributaries of the Fuhe River were estimated with copula functions. The 66-year daily flow discharge data series were considered to construct the marginal distributions of flood magnitudes and occurrence dates. Based on copula functions, the joint distributions of the annual maximum flood magnitudes and occurrence dates in the mainstream and its tributaries were established, respectively. The coincidence risk of flood occurrence dates was calculated, and the coincidence analysis by joint probabilities, co-occurrence probabilities and conditional probabilities for different flood magnitudes were compared with the simulation results. The main conclusions of this study were summarized as follows:

(1) There is a strong consistency and significantly correlation between the floods at the upstream and downstream with the Pearson coefficients reaching 0.90. The floods at the Lijiadu station are mainly formed by the superimposition of the upstream floods. The mixed von Mises distribution and P3 distribution perform well in quantifying the marginal distributions of occurrence dates and flood magnitudes, and the Clayton copula is the best one for simulating the joint distributions of flood variables.

(2) The coincidence events of the annual maximum flood occurrence dates at the Liaojiawan and Loujiacun stations mainly occur from May to early July. There are two flood coincidence peaks occurring on May 12 and June 21, which coincidence probabilities are reaching 0.026% and 0.057%, respectively. The coincidence risk of the annual maximum flood occurrence dates in the upper mainstream and its tributaries throughout a year is 2.87%.

(3) The joint probability and co-occurrence probability of 50-year design floods at the Liaojiawan and Loujiacun stations are 3.86% and 0.14%. If a 50-year flood occurs at the Liaojiawan or Loujiacun station, the corresponding probability of a flood with the same frequency occurring at the Lijiadu station is 0.26% or 0.22%, respectively.

(4) Floods with the same frequencies simultaneously occurring at the Liaojiawan and Loujiacun stations are likely to superimpose into large floods at the Lijiadu station. Among the three coincidence

probability calculation methods, the conditional probability is the most consistent with the flood coincidence risk in the mainstream and its tributaries, which is more reliable and rational in practice.

In this study, the copula-based quantitative analysis of flood coincidence risk contributes us to better understand the spatiotemporal characteristics of floods in the Fuhe River. By calculating the flood coincidence probabilities of flood magnitudes and occurrence dates, we can verify the feasibility of the calculation method of flood coincidence probability considering the connection of the mainstream and its tributaries, so as to have a comprehensive knowledge in the flood coincidence laws. The results provide a scientific basis and effective support for improving the flood control capacity and ensuring the safety of flood control targets in the Poyang Lake region.

Acknowledgements

This study was financially supported by the National Natural Science Foundation of China (51879192) and China Three Gorges Corporation (0799254).

Author contributions statement

Shenglian Guo proposed the study, Feng Xiong and Jun Wang conducted the experiments, Na Li analyzed the results. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests. The corresponding author is responsible for submitting a competing interest statement on behalf of all authors of the paper.

References

Aghakouchak A, Bárdossy A, Habib E (2010a) Conditional simulation of remotely sensed rainfall data using a non-Gaussian v-transformed copula. Adv Water Resour 33(6): 624–634

Aghakouchak A, Bárdossy A, Habib E (2010b) Copula-based uncertainty modelling: application to multisensor precipitation estimates. Hydrol Process 24(15): 2111–2124

Alfieri L, Feyen L, Di Baldassarre G (2016) Increasing flood risk under climate change: a pan-European assessment of the benefits of four adaptation strategies. Clim Change 136: 507-521. https://doi.org/10.1007/s10584-016-1641-1

Ashkar F, Aucoin F (2011) A broader look at bivariate distributions applicable in hydrology. J Hydrol 405: 451–461

Bacchi B, Becciu G, Kottegoda NT (1994) Bivariate exponential model applied to intensities and durations of extreme rainfall. J Hydrol 155: 225–236

Candela A, Brigandì G, Aronica GT (2014) Estimation of synthetic flood design hydrographs using a distributed rainfall–runoff model coupled with a copula-based single storm rainfall generator. Nat Hazards Earth Syst Sci 14: 1819-1833

Ceola S, Laio F, Montanari A (2014) Satellite nighttime lights reveal increasing human exposure to floods worldwide. Geophys Res Lett 41(20): 7184–7190. https://doi.org/10.1002/2014GL061859

Chen L, Singh VP, Guo SL et al (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

Chen L, Singh VP, Guo SL et al (2015) Copula-based method for multisite monthly and daily streamflow simulation. J Hydrol 528: 369–384

Chen L, Singh VP, Lu WW et al (2016) Streamflow forecast uncertainty evolution and its effect on real-time reservoir operation. J Hydrol 540: 712–726

Clayton DG (1978) A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika 65(1): 141–151. https://doi.org/10.1093 /biomet/65.1.141

Correia FN (1978) Multivariate partial duration series in flood risk analysis. Springer, Dordrecht

Daksiya V, Mandapaka PV, Lo EYM (2020) Effect of climate change and urbanization on flood protection decision-making. J Flood Risk Manag 14: e12681. https://doi.org/10.1111/jfr3.12681

de Bruijn KM, Klijn F, van de Pas B, Slager CTJ (2015) Flood fatality hazard and flood damage hazard: combining multiple hazard characteristics into meaningful maps for spatial planning. Nat Hazards Earth Syst Sci 15: 1297-1309

De Michele C, Salvadori G (2003) A generalized Pareto intensity duration model of storm rainfall exploiting 2-copulas. J Geophys Res Atmos 108(D2): 1–11. https://doi.org/10.1029/2002JD002534

Escalante SC (2007) Application of bivariate extreme value distribution to flood frequency analysis: A case study of Northwestern Mexico. Nat Hazards 42: 37–46. https://doi.org/10.1007/s11069-006-9044-7

Favre AC, Adlouni SE, Perreault L et al (2004) Multivariate hydrological frequency analysis using copulas. Water Resour Res 40(1): W01101. https://doi.org/10.1029/2003WR002456

Feng Y, Shi P, Qu SM et al (2020) Nonstationary flood coincidence risk analysis using time-varying Copula functions. Sci Rep 10: 3395. https://doi.org/10.1038/s41598-020-60264-3

Fisher NI (1993) Statistical analysis of circular data. Cambridge University Press, Cambridge

Frank MJ (1979) On the simultaneous associatively of F(x,y) and x+y-F(x,y). Aequationes Math 19(1): 194–226. https://doi.org/10.1007/BF02189866

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

Goel NK, Seth SM, Chandra S (1998) Multivariate modeling of flood flows. J Hydrol Eng 124(2): 146-155

Grimaldi S, Serinaldi F (2006) Design hyetographs analysis with 3-copula function. Hydrol Sci J 51(2): 223–238. https://doi.org/10.1623/hysj.51.2.223

Gringorten II (1963) A plotting rule for extreme probability paper. J Geophys Res 68(3): 813–814. https://doi.org/10.1029/JZ068i003p00813

Guo SL, Muhammad R, Liu ZJ et al (2018) Design flood estimation methods for cascade reservoirs based on copulas. Water 10(5): 560

Hirabayashi Y, Mahendran R, Koirala S et al (2013) Global flood risk under climate change. Nat Clim Change 3(9): 816-821. https://doi.org/10.1038/nclimate1911

Hirsch RM, Stedinger JR (1987) Plotting positions for historical floods and their precision. Water Resour Res 23(4): 715–727. https://doi.org/10.1029/WR023i004p00715

Hosking JRM (1990) L-moments: analysis and estimation of distributions using linear combinations of order statistics. J R Stat Soc 52(2): 105–124

Hougaard P (1986) A class of multivariate failure time distributions. Biometrika 73(3): 671–678. https://doi.org/10.2307/2336531

Huang KD, Chen L, Zhou JZ et al (2018) Flood hydrograph coincidence analysis for mainstream and its tributaries. J Hydrol 565: 341-353

Joe H (1997) Multivariate models and dependence concepts. Chapman & Hall, London

Kao SC, Govindaraju RS (2007) A bivariate frequency analysis of extreme rainfall with implications for design. J Geophys Res Atmos 112(D1): 13119

Kao SC, Govindaraju RS (2008) Trivariate statistical analysis of extreme rainfall events via the Plackett family of copulas. Water Resour Res 44(2): W02415. https://doi.org/10.1029/ 2007WR006261

Karahacane H, Meddi M, Chebana F, Saaed HA (2020) Complete multivariate flood frequency analysis, applied to northern Algeria. J Flood Risk Manag 13: e12619

Karmakar S, Simonovic SP (2009) Bivariate flood frequency analysis. Part 2: a copula-based approach with mixed marginal distributions. J Flood Risk Manag 2(1): 32-44. https://doi.org/10.1111/j.1753-318X.2009.01020.x

Keef C, Svensson C, Tawn JA (2009) Spatial dependence in extreme river flows and precipitation for Great Britain. J Hydrol 378(3–4): 240–252. https://doi.org/10.1016/j.jhydrol.2009.09.026

Kim T, Valdés JB, Yoo C (2006) Nonparametric approach for bivariate drought characterization using Palmer drought index. J Hydrol Eng 11(2): 134-143. https://doi.org/10.1061/(ASCE)1084-0699(2006)11:2(134)

Klein B, Pahlow M, Hundecha Y, Schumann A (2010) Probability analysis of hydrological loads for the design of flood control systems using copulas. J Hydrol Eng 15(5): 360–369

KvočKa D, Falconer RA, Bray M (2016) Flood hazard assessment for extreme flood events. Nat Hazards 84(3): 1569-1599

Leonard M, Metcalfe A, Lambert M (2008) Frequency analysis of rainfall and streamflow extremes accounting for seasonal and climatic partitions. J Hydrol 348(1–2): 135–147

Li TY, Guo SL, Chen L, Guo JL (2013) Bivariate flood frequency analysis with historical information based on copula. J Hydrol Eng 18(8): 1018–1030. https://doi.org/10.1061/ (ASCE)HE.1943-5584.0000684

Makkonen L (2008) Bringing closure to the plotting position controversy. Commun Stat-Theory M 37(3): 460–467

Mardia K V, Jupp PE (2009) Directional statistics. Wiley, New York

Massey FJ (1951) The Kolmogorov–Smirnov test for goodness of fit. J Am Stat Assoc 46(253): 68–78

Michael Z, Stanislav U (2013) Statistical decision problems. Springer, New York

MWR (Ministry of Water Resources) (2006) Regulation for Calculating Design Flood of Water Resources and Hydropower Projects. Chinese Shuili Shuidian Press, Beijing (in Chinese)

Nelsen RB (2006) An introduction to copulas, 2nd edn. Springer, New York

Peng Y, Chen K, Yan HX, Yu XL (2017) Improving flood-risk analysis for confluence flooding control downstream using Copula Monte Carlo method. J Hydrol Eng 22(8): 04017018. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001526

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

Poduje ACC, Haberlandt U (2017) Short time step continuous rainfall modeling and simulation of extreme events. J Hydrol 552: 182–197

Prohaska S, Ilic A (2010) Coincidence of flood flow of the Danube River and its tributaries: Hydrological processes of the Danube River Basin. Springer, Dordrecht

Razali NM, Wah YB (2011) Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. J Stat Model Anal 2(1): 21–33

Reddy MJ, Ganguli P (2012) Bivariate flood frequency analysis of Upper Godavari River flows using Archimedean copulas. Water Resour Manag 26(14): 3995–4018

Salvadori G, Michele CD (2004) Frequency analysis via copulas: theoretical aspects and applications to hydrological events. Water Resour Res 40(12): 229–244

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

Schweizer B, Sklar A (1983) Probabilistic metric spaces. Elsevier Science, New York

Shiau JT (2006) Fitting drought and severity with two-dimensional copulas. Water Resour Manag 20(5): 795- 815

Shimizu K (2010) A bivariate mixed Lognormal distribution with an analysis of rainfall data. J Appl Meteorol Clim 32(2): 161-171

Silverman BW (1986) Density Estimation for Statistics and Data Analysis. Chapman and Hall, New York

Sraj M, Bezak N, Brilly M (2015) Bivariate flood frequency analysis using the copula function: a case study of the Litija station on the Sava River. Hydrol Process 29: 235–248. https://doi.org/10.1002/hyp.10145

Thieken AH, Apel H, Merz B (2015) Assessing the probability of large-scale flood loss events: a case study for the river Rhine, Germany. J Flood Risk Manag 8: 247-262. https://doi.org/10.1111/jfr3.12091

Wang B, Xiang BQ, Lee JY (2013) Subtropical high predictability establishes a promising way for monsoon and tropical storm predictions. P Natl Acad Sci USA 110(8): 2718-2722

Wang C (2016) A joint probability approach for coincidental flood frequency analysis at ungauged basin confluences. Nat Hazards 82(3): 1727-1741. https://doi.org/10.1007/s11069-016-2265-5

Wang C, Chang NB, Yeh GT (2009) Copula-based flood frequency (COFF) analysis at the confluences of river systems. Hydrol Process 23(10): 1471–1486. https://doi.org/10.1002/hyp.7273

Weiss MS (1978) Modification of the Kolmogorov–Smirnov statistic for use with correlated data. J Am Stat Assoc 73(364): 872–875

Xiao Y, Guo SL, Liu P et al (2009) Design flood hydrograph based on multi-characteristic synthesis index method. J Hydrol Eng 14(12): 1359–1364. https://doi.org/10.1061/(ASCE)1084-0699(2009)14:12(1359)

Xiong F, Guo SL, Liu P et al (2019) A general framework of design flood estimation for cascade reservoirs in operation period. J Hydrol 577: 124003

Yin JB, Guo SL, Liu ZJ et al (2018) Uncertainty analysis of bivariate design flood estimation and its impacts on reservoir routing. Water Resour Manag 32(5): 1795-1809

Yue S (1999) Applying the bivariate normal distribution to flood frequency analysis. Water Int 24(3): 248–254. https://doi.org/10.1080/02508069908692168

Yue S (2000a) The bivariate lognormal distribution to model a multivariate flood episode. Hydrol Process

14 (14): 2575- 2588

Yue S (2000b) Joint probability distribution of annual maximum storm peaks and amounts as represented by daily rainfalls. Hydrol Sci J 45(2): 315-326

Yue S (2002) The bivariate lognormal distribution for describing joint statistical properties of a multivariate storm event. Environmetrics 13: 811-819

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

Zhang L, Singh VP (2007) Trivariate flood frequency analysis using the Gumbel-Hougaard copula. J Hydrol Eng 12(4): 431–439. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(431)

Zhang L, Singh VP (2012) Bivariate rainfall and runoff analysis using entropy and copula theories. Entropy 14(9): 1784–1812. https://doi.org/10.3390/e14091784.

Zhang WJ, Jin F, Stuecker MF et al (2016) Unraveling El Niño's impact on the East Asian Monsoon and Yangtze River summer flooding. Geophys Res Lett 43: 11375–11382.https://doi.org/10.1002/2016GL071190

Zhong YX, Guo SL, Liu ZJ et al (2018) Quantifying differences between reservoir inflows and dam site floods using frequency and risk analysis methods. Stoch Environ Res Risk Assess 32(2): 419-433

Lists of Tables

Table 1 Parameters and test results of marginal distributions of flood occurrence dates

Table 2 Parameters and test results of marginal distributions of flood magnitudes

Table 3 Parameters and test results of joint distributions of flood occurrence dates and magnitudes

Table 4 Joint probabilities and co-occurrence probabilities of flood magnitudes at Liaojiawan and Loujiacun stations (%)

Table 5 Conditional probabilities of flood magnitudes at three stations (%)

Table 6 Statistical analysis of design floods at three stations in Fuhe River

Table 7 Comparative analysis of flood coincidence probabilities under different conditions

Table 1 Parameters and test results of marginal distributions of flood occurrence dates

Station	u_i	k_i	p_i	K-S	RMSE
	1.82	6.11	0.13		
Liaojiawan	2.28	39.34	0.14	0.061(0.167)	0.022
	2.98	10.86	0.73		
	2.58	3.01	0.68		
Loujiacun	2.98	32.51	0.32	0.053(0.167)	0.019
	2.75	8.04	0.00		
	1.83	1.75	0.15		
Lijiadu	2.82	8.98	0.54	0.085(0.167)	0.033
	2.96	3.42	0.31		

Table 2 Parameters and test results of marginal distributions of flood magnitudes

Station	Mean (m^3/s)	*Cv*	*Cs*	*p*-value (χ^2)	RMSE
Liaojiawan	2861.69	0.51	0.77	0.249 (0.05)	0.047
Loujiacun	1903.75	0.43	0.78	0.249 (0.05)	0.025
Lijiadu	4570.62	0.45	0.63	0.247 (0.05)	0.055

Table 3 Parameters and test results of joint distributions of flood occurrence dates and magnitudes

Joint distribution	Clayton		GH		Frank	
	θ	RMSE	θ	RMSE	θ	RMSE
Occurrence dates between Liaojiawan and Loujiacun	3.69	0.028	2.84	0.030	9.38	0.029
Flood magnitudes between Liaojiawan and Loujiacun	2.59	0.033	2.30	0.043	7.05	0.037
Flood magnitudes between Liaojiawan and Lijiadu	6.25	0.039	4.12	0.048	14.64	0.043
Flood magnitudes between Loujiacun and Lijiadu	4.99	0.034	3.49	0.045	12.07	0.041

Table 4 Joint probabilities and co-occurrence probabilities of flood magnitudes at Liaojiawan and Loujiacun stations (%)

Station	Loujiacun						Coincidence probabilities
	T (year)	5	10	20	50	100	
Liaojiawan	5	30.48	24.85	22.33	20.91	20.45	Joint probabilities
	10	24.85	17.14	13.50	11.38	10.69	
	20	22.33	13.50	9.20	6.67	5.83	
	50	20.91	11.38	6.67	3.86	2.93	
	100	20.45	10.69	5.83	2.93	1.96	
	5	9.52	5.15	2.67	1.09	0.55	Co-occurrence probabilities
	10	5.15	2.86	1.50	0.62	0.31	
	20	2.67	1.50	0.80	0.33	0.17	
	50	1.09	0.62	0.33	0.14	0.07	
	100	0.55	0.31	0.17	0.07	0.04	

Table 5 Conditional probabilities of flood magnitudes at three stations (%)

Station	Lijiadu					
	T (year)	5	10	20	50	100
Liaojiawan	5	16.41	9.23	4.83	1.98	0.99
	10	8.20	5.01	2.74	1.15	0.58
	20	4.06	2.59	1.46	0.62	0.32
	50	1.61	1.06	0.61	0.26	0.14
	100	0.80	0.53	0.31	0.13	0.07
Loujiacun	5	15.21	8.48	4.43	1.82	0.91
	10	7.54	4.47	2.42	1.01	0.51
	20	3.73	2.29	1.26	0.54	0.27
	50	1.48	0.93	0.52	0.22	0.11
	100	0.74	0.47	0.26	0.11	0.06

Table 6 Statistical analysis of design floods at three stations in Fuhe River

Return period (year)	5	10	20	50	100
Design flood at Liaojiawan (m^3/s)	4007	4816	5543	6430	7061
Design flood at Loujiacun (m^3/s)	2544	2997	3406	3904	4258
Simulated flood at Lijiadu (m^3/s)	6227	7397	8451	9735	10649
Simulated flood frequency at Lijiadu (%)	19.79	9.41	4.41	1.59	0.73
Return period at Lijiadu (year)	5	11	23	63	137

Table 7 Comparative analysis of flood coincidence probabilities under different conditions

Return period (year)	5	10	20	50	100
Simulated flood frequency (%)	19.79	9.41	4.41	1.59	0.73
Joint probability (%)	30.48	17.14	9.20	3.86	1.96
Co-occurrence probability (%)	9.52	2.86	0.80	0.14	0.04
Conditional probability P_1^c (%)	16.41	5.01	1.46	0.26	0.07
Conditional probability P_2^c (%)	15.21	4.47	1.26	0.22	0.06

Lists of Figures

Fig. 1 Location of hydrological stations in the Fuhe River basin

Fig. 2 Correlation analysis of annual maximum flood variables in Fuhe River

Fig. 3 Multiple regression model fitting results of flood variables at Lijiadu station

Fig. 4 Fitted empirical and theoretical probabilities of marginal distributions

Fig. 5 Fitted empirical and theoretical probabilities of joint distributions

Fig. 6 Daily coincidence probabilities of occurrence dates of annual maximum flood at Liaojiawan and Loujiacun stations

Fig. 7 Flood coincidence probabilities under different conditions

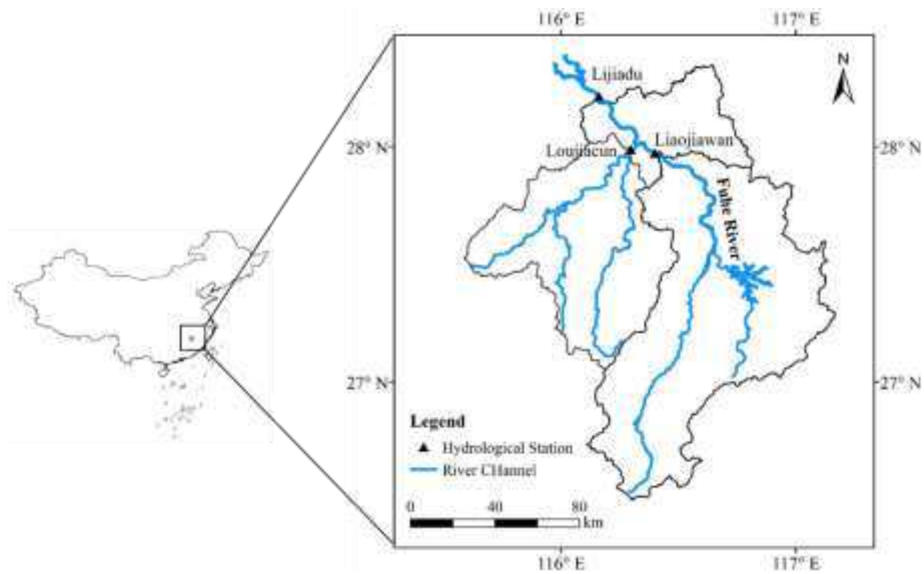
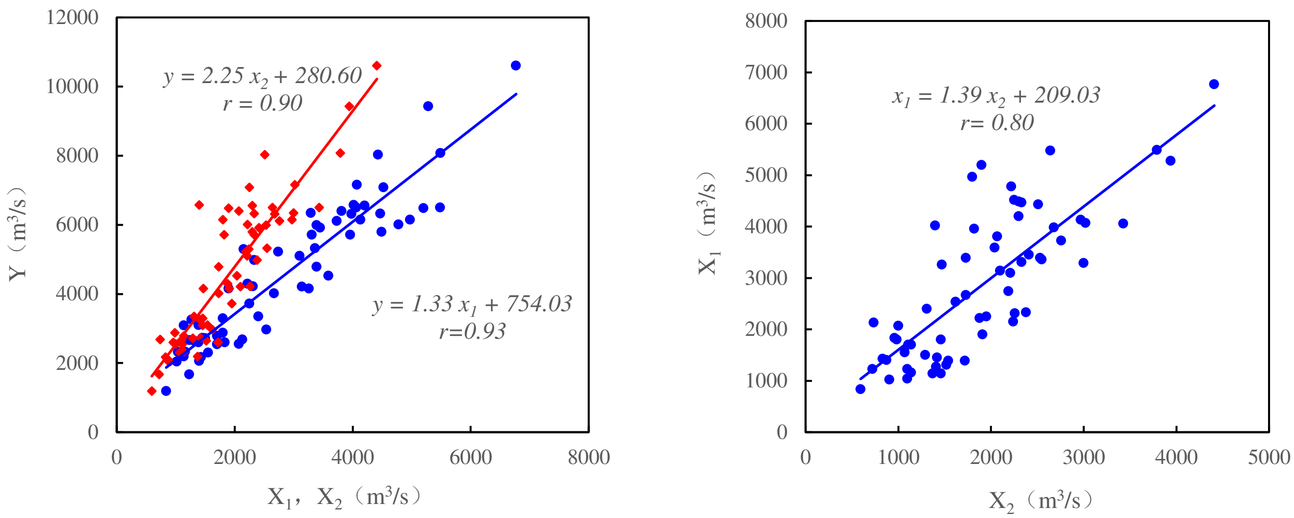


Fig. 1 Location of hydrological stations in the Fuhe River basin



(a) $Y \sim X_1$ and Y : X_2 (b) $X_1 \sim X_2$

695 **Fig. 2** Correlation analysis of annual maximum flood variables in Fuhe River

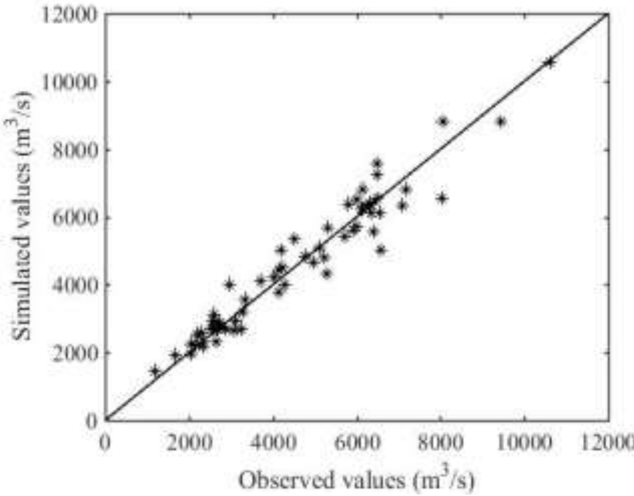
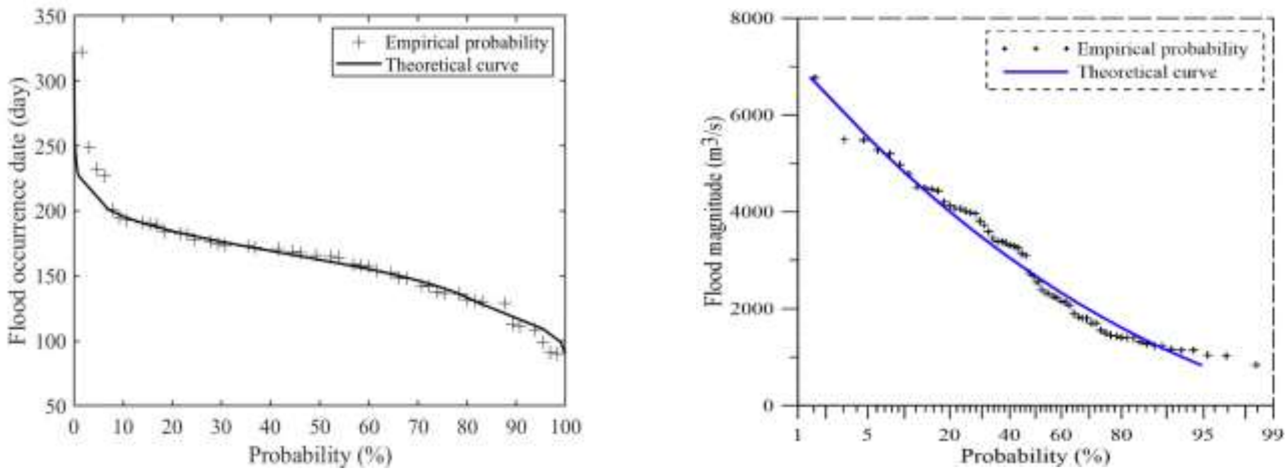
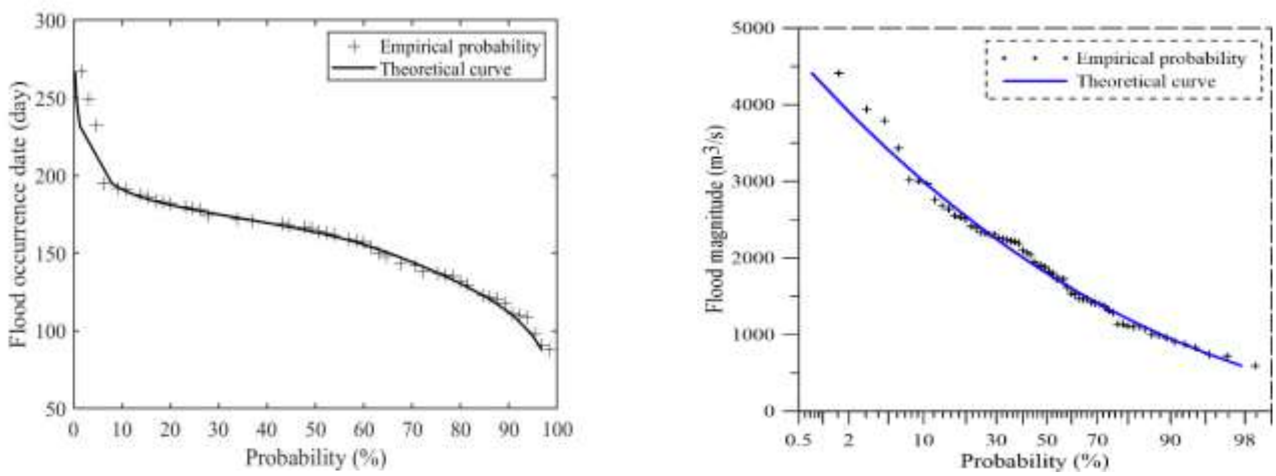


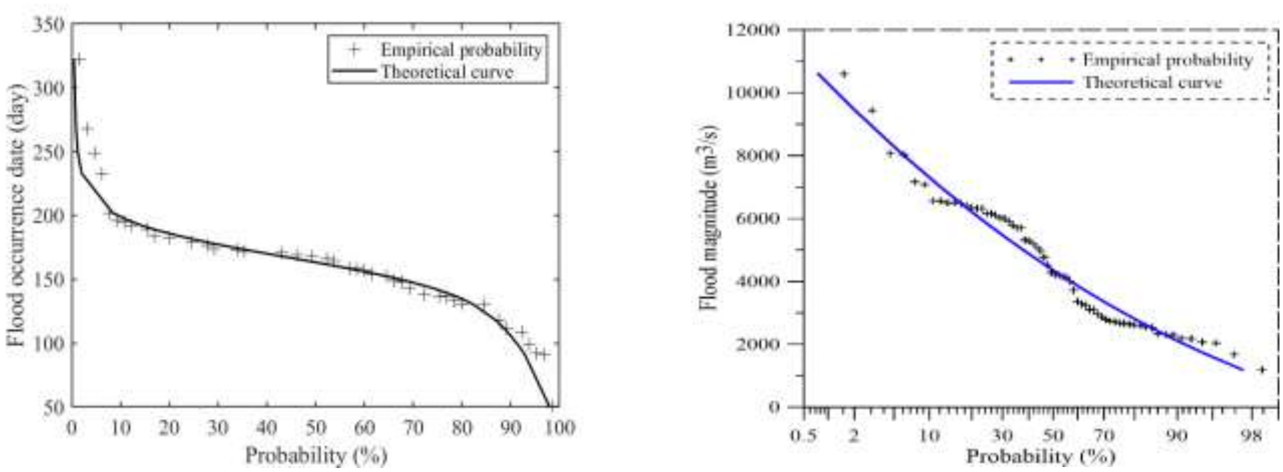
Fig. 3 Multiple regression model fitting results of flood variables at Lijiadu station



(a) Flood occurrence dates and flood magnitudes at Liaojiawan station

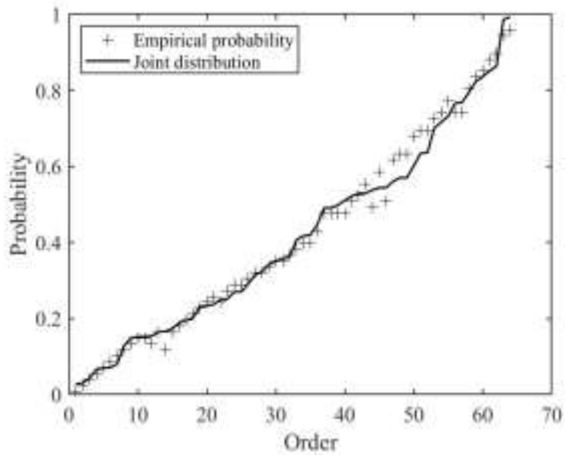


(b) Flood occurrence dates and flood magnitudes at Loujiacun station

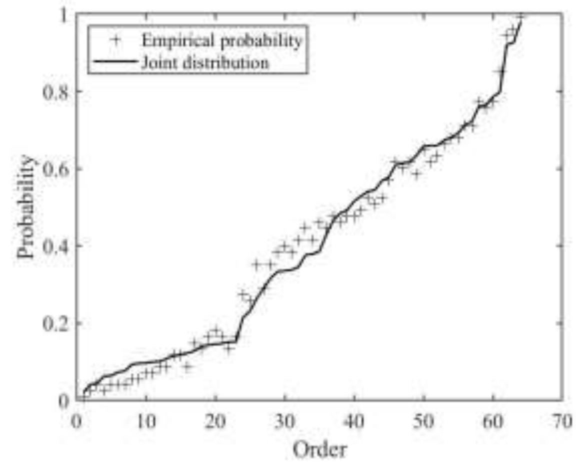


(c) Flood occurrence dates and flood magnitudes at Lijiadu station

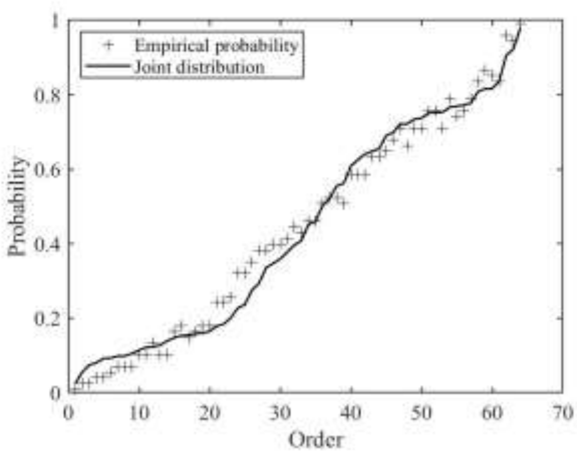
Fig. 4 Fitted empirical and theoretical probabilities of marginal distributions



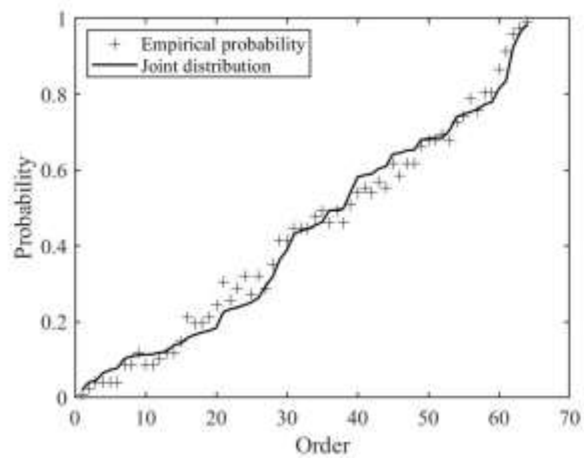
(a) Occurrence dates between Liaojiawan and Loujiacun stations



(b) Flood magnitudes between Liaojiawan and Loujiacun stations



(c) Flood magnitudes between Liaojiawan and Lijiadu stations



(d) Flood magnitudes between Loujiacun and Lijiadu stations

Fig. 5 Fitted empirical and theoretical probabilities of joint distributions

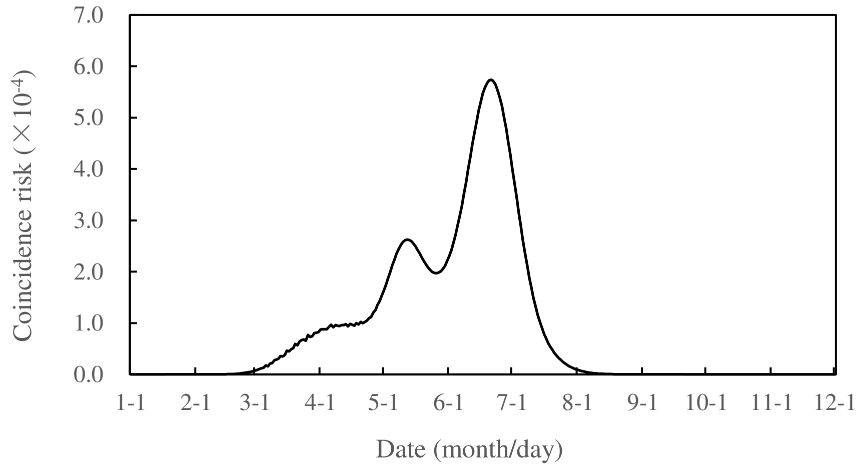


Fig. 6 Daily coincidence probabilities of occurrence dates of annual maximum flood at Liaojiawan and Loujiacun stations

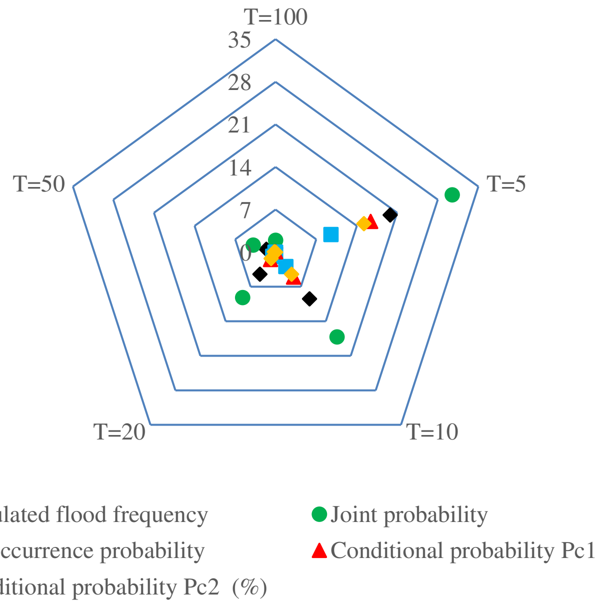


Fig. 7 Flood coincidence probabilities under different conditions