

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

Na Li

Wuhan University

Shenglian Guo (✉ slguo@whu.edu.cn)

Wuhan University <https://orcid.org/0000-0002-8594-4988>

Feng Xiong

Changjiang Water Resources Commission

Jun Wang

Wuhan University

Yuzuo Xie

Wuhan University

Research Article

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

Posted Date: December 15th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-824942/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 Water Resources Management on January 6th, 2022. See the published version at <https://doi.org/10.1007/s11269-021-03050-8>.

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, respectively. 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 quantitatively estimated and compared with the predicted flood coincidence risks. The results show that the selected marginal and joint distributions can fit observed flood dataset very well. The coincidence probabilities of flood occurrence dates in the upper mainstream and its tributaries mainly occur from May to early July. It is found that the conditional probability is the most consistent with the predicted flood coincidence risks in the mainstream and its tributaries, and is more reliable and rational in practice.

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). Affected by human activities, climate change, environmental degradation and El Niño, extreme hydrological events occur frequently worldwide and the frequency and intensity of floods continue to increase (Alfieri et al. 2016). Generally, large floods can be caused by the combination of floods in the mainstream and its tributaries (Prohaska and Ilic 2010). When floods occur simultaneously, the flood peaks and volumes will superimpose into large floods, threatening 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 flood coincidence study, the traditional method only focuses on the statistical analysis of flood events based on the synchronized flood data, and can't quantitatively estimate the coincidence probabilities of design floods at specific site. Essentially, the flood coincidence is a multivariable frequency combination event, which can be studied by the multivariable hydrological analysis method (Feng et al. 2020). As an effective method, copula functions can connect arbitrary marginal distributions through correlation structures, and have been widely used in multivariate hydrological analysis. For example, in flood frequency analysis (Tsakiris et al. 2015; Zhong et al. 2018; Moftakhari et al. 2019; Karahacane et al. 2020; Dodangeh et al. 2020), rainfall frequency analysis (Ashkar and Aucoin 2011; Zhang and Singh 2012), drought frequency analysis (Montaseri et al. 2018), and multivariate simulation (Chen et al. 2016; Jane et al. 2020).

Copula functions have also been applied in flood coincidence analysis (Bing et al. 2018; Muthuvel and Mahesha 2021). For example, Schulte and Schumann (2016) developed multivariate copula-approaches to analyze coincidence risk of flood peaks in adjoining catchment. Bender et al. (2016) introduced a multivariate design framework by copula functions to evaluate the flood occurrence at the confluence. Peng et al. (2017) estimated flood risk in the confluence flood control downstream of a reservoir using Copula Monte Carlo method. Gilja et al. (2018) analyzed flood hazard at river confluences based on bivariate copula. However, all of 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, which 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, both factors of flood occurrence time and magnitudes should be taken into consideration. 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. Huang et al. (2018) took flood magnitudes of two rivers and flood occurrence interval dates as three reference variables to further explore the flood hydrograph coincidence risk using copulas. Peng et al. (2019) employed multivariate copulas to estimate flood coincidence probabilities and considered flood occurrence dates and magnitudes simultaneously.

All above work has revealed the characteristics of flood coincidence risk from different angles, and made some progress in the flood coincidence analysis. However, some researches only focused on the coincidence risk of flood magnitudes and neglected the flood occurrence time; other researches ignored the correlation of flood magnitudes and time. Furthermore, there are several flood coincidence risk estimation methods, which one is much rational and reliable in practice?

The objective of this study is to estimate flood coincidence risks quantitatively in the mainstream and its tributaries by considering the correlation between the upstream and downstream flood variables, and to evaluate the reliability and applicability of different coincidence probability estimation methods. The main steps of this study are as follow: First, the multivariate regression model is used to establish the functional relationship among the flood variables in the mainstream and its tributaries. Second, the mixed von Mises and Pearson Type III marginal distributions are selected to describe the annual maximum flood occurrence dates and flood magnitudes, respectively. Third, the joint distributions of flood magnitudes and occurrence dates are constructed based on copula function. Fourth, the coincidence probabilities of flood occurrence dates and magnitudes are quantitatively estimated. Finally, the flood coincidence probabilities obtained by different estimation methods are compared with the predicted values.

2 Methodology

2.1 Copula functions

Sklar's theorem assumes 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:

$$F(x, y) = C_\theta(F_X(x), F_Y(y)) = C_\theta(u, v) \quad (1)$$

where $C_\theta(\cdot)$ is a copula function; θ is a parameter of the copula function, which can capture the dependency between random variables; U and V are marginal distribution functions, satisfying $U = F_X(x) = P_X(X \leq x)$, and $V = F_Y(y) = P_Y(Y \leq y)$.

Copula function is a multi-dimensional joint distribution function uniformly distributed in the domain of $[0, 1]$, which can integrate marginal distributions and correlation structures to construct multi-dimensional joint distributions (Nelsen 2006). The Clayton (Clayton 1978), Gumbel-Hougaard (Hougaard 1986) and Frank copula (Frank 1979) only have one parameter to be estimated and are widely used in practice (Guo et al. 2018; Yin et al. 2018). The mathematical expressions of these copula functions are:

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

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

Frank copula:
$$C(u, v) = -\frac{1}{\theta} \ln \left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)} \right]; \theta \in (-\infty, \infty) \setminus \{0\} \quad (4)$$

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:

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

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.

2.2 Marginal distributions of flood occurrence dates and magnitudes

A mixed von Mises distributions is used to fit the flood occurrence dates series which often have the characteristics of periodicity and multi-peaked. The probability density function of the mixed von Mises

distribution can be written as:

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

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.

Pearson Type III (P3) distribution recommended by the Ministry of Water Resources, China as a uniform procedure for flood frequency analysis (MWR 2006), is used to fit the annual maximum flood magnitudes. The probability density function of P3 distribution is expressed as:

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} (x - \delta)^{\alpha-1} \exp[-\beta(x - \delta)]; \alpha > 0, \beta > 0, \delta \leq x \quad (7)$$

where α , β and δ are the shape, scale, and position parameters of the P3 distribution, respectively; $\Gamma(\cdot)$ is the gamma function.

2.3 Goodness-of-fit evaluation

The Root Mean Square Error (RMSE), Kolmogorov-Smirnov (K-S) and Chi-square (χ^2) test methods are selected to test the goodness-of-fit of the marginal and joint distributions. For the univariate series, the empirical probabilities of each flood variable are calculated by the Gringorten plotting position formula (Gringorten 1963), which is written as:

$$P_e(x_i) = \frac{i - 0.44}{n + 0.12} \quad (8)$$

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 specific formula as follows:

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

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 RMSE is selected to measure the difference between the theoretical and empirical probabilities and calculated by:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(P_{ei} - P_i \right)^2} \quad (10)$$

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.

For n observed data in an increasing order, the K-S test statistic is expressed as:

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

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

Chi-square (χ^2) test is to measure the degree of deviation between the observed and predicted values, which statistic is defined as:

$$\chi^2 = \frac{\left(M_1 - n p_1 \right)^2}{n p_1} + \frac{\left(M_2 - n p_2 \right)^2}{n p_2} + \dots + \frac{\left(M_k - n p_k \right)^2}{n p_k} \quad (12)$$

where $\{p_1, p_2, \dots, p_k\}$ are the hypothesized probabilities for k possible outcomes; and $\{M_1, M_2, \dots, M_k\}$ are the observed counts of each outcome to be compared for expected counts $n p_1, n p_2, \dots, n p_k$ in n independent trials.

2.4 Estimation of flood coincidence risks

2.4.1 Coincidence risk of 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:

$$\begin{aligned} P_k^i &= P(t_k < T_i \leq t_{k+1}, t_k < T_j \leq t_{k+1}) \\ &= F_T(t_{k+1}, t_k) + F_T(t_{k+1}, t_{k+1}) - F_T(t_k, t_{k+1}) - F_T(t_{k+1}, t_k) \end{aligned} \quad (13)$$

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.

2.4.2 Coincidence risk of flood magnitudes

The joint distributions of the annual maximum flood magnitudes in the mainstream and its tributaries are established firstly based on copula functions. Then, the joint probabilities, co-occurrence probabilities and conditional probabilities can be quantitatively estimated by following equations.

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 and can be expressed as:

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

where x and y are the flood magnitudes in i and j river, respectively; and $F\left(\{x,y\} \right)$ 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 and can be expressed as:

$$P\left(\{X>x \cap Y>y\} \right) = 1 - \{F_X\}\left(x \right) - \{F_Y\}\left(y \right) + F\left(\{x,y\} \right) = 1 - u - v + C\left(\{u,v\} \right) \quad (15)$$

For the conditional probabilities, it refers to the probabilities that when one of the variables is in a given range, the other variable falls within specified 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 and can be expressed as:

$$P\left(\left. \{Y>y\} \right| X>x \right) = \frac{P\left(\{X>x \cap Y>y\} \right)}{P\left(\{X>x\} \right)} = \frac{1 - \{F_X\}\left(x \right) - \{F_Y\}\left(y \right) + F\left(\{x,y\} \right)}{1 - \{F_X\}\left(x \right)} = \frac{1 - u - v + C\left(\{u,v\} \right)}{1 - u} \quad (16)$$

3 Study Area And Data

The Fuhe River, located at the east of Jiangxi Province, China, feeding into the Poyang Lake, was selected as a case study. Fig. 1 sketches the map of Fuhe River basin, the mainstream and tributaries as well as the hydrological stations. Affected by the subtropical humid monsoon climate, the flood season is generally from April to early July in this region.

The Fuhe River is 348 km long and has a drainage area of 16,493 km², and the control basin area of Liaojiawan and Loujiacun hydrological stations located in the upper mainstream and tributary is 8723 km² and 4969 km², respectively. The Lijiadu hydrological station with a catchment area of 15,812 km² is located in the down-mainstream, accounting for more than 95% of the entire drainage area. The daily average flow discharge data series at these three hydrological stations from 1953 to 2016 were collected from Jiangxi Province, and the annual maximum flood magnitudes and corresponding occurrence dates were sampled and shown in Fig. 2.

4 Results And Discussion

4.1 Flood correlation analysis

The Pearson correlation coefficient was used to describe the correlation between the annual maximum flood magnitude series of these three hydrological stations in the Fuhe River. Fig. 3(a) shows the correlation analysis of the annual maximum flood magnitudes between the Liaojiawan, Loujiacun and the Lijiadu station. The Pearson correlation coefficient between the Liaojiawan and Lijiadu stations is 0.93, and between the Loujiacun and Lijiadu stations is 0.90. Since the control basin area of the Loujiacun station at tributary is smaller than that of the Liaojiawan station at upper mainstream, its correlation with the Lijiadu station is relatively small.

Figure 3(b) shows the correlation analysis of the annual maximum flood magnitudes between the Liaojiawan and Loujiacun stations located in the upper reaches, and the Pearson correlation coefficient is equal to 0.80. Since their control basins are very close, the key factors for flood generation including climatic conditions, geographical environment and topographical position are very similar (Deidda et al. 2021), which implicitly indicates that the regularity of flood occurrence in the upper reaches is highly consistent.

The multiple regression model is employed to analyze the quantitative relationship among the annual maximum flood magnitude series. The functional relationship among the annual maximum flood magnitudes at these three hydrological stations is as follows:

$$Y = \hat{Y} + \varepsilon = 104.890 + 0.841X_1 + 1.082X_2 + \varepsilon \quad (17)$$

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

The relationship between the predicted and observed flood magnitudes at the Lijiadu station is plotted in Fig. 3(c). The scattered points are basically distributed along the 45° diagonal line, indicating the predicted and the observed flood values are very close. The simultaneous occurrence of the upstream floods is prone to cause large flooding in the downstream confluence area.

4.2 Estimation of marginal distribution parameters

The parameters of mixed von Mises distributions were estimated by the maximum likelihood method, and the K-S test and the RMSE were applied for the goodness-of-fit evaluation. Table 1 shows that the K-S test statistics values do not exceed critical values at a significance level of 0.05, implying the hypothesis that the annual maximum flood occurrence dates follow the mixed von Mises marginal distribution could be accepted. Meanwhile, the RMSE values between the theoretical and empirical probabilities are very small.

The parameters of the P3 distribution are estimated by the L-moment method. The estimated statistics, the Chi-square test and RMSE are also listed in Table 1. It is shown that the p -value (χ^2) are all

larger than 0.05, implying that the hypothesis test that the flood magnitudes obey the P3 marginal distribution is not rejected.

The mixed von Mises distributions and empirical probabilities were calculated by Eq. (6) and Eq. (8), while the P3 distributions and empirical probabilities were calculated by Eq. (7) and Eq. (8), respectively. Fig. 4 shows the empirical frequency points can fit the theoretical probability curve for the annual maximum flood occurrence dates or magnitudes very well.

Table 1
Estimated parameters and test results of marginal distributions

Station	Mixed von Mises distribution				P3 distribution					
	u_j	k_j	p_j	K-S	RMSE	Mean (m ³ /s)	C_v	C_s	$\{\chi^2\}$	RMSE
	1.82	6.11	0.13							
Liaojiawan	2.28	39.34	0.14	0.061	0.022	2862	0.51	0.77	0.249	0.047
	2.98	10.86	0.73	(0.167)						
	2.58	3.01	0.68							
Loujiacun	2.98	32.51	0.32	0.053	0.019	1904	0.43	0.78	0.249	0.025
	2.75	8.04	0.00	(0.167)						
	1.83	1.75	0.15							
Lijiadu	2.82	8.98	0.54	0.085	0.033	4571	0.45	0.63	0.247	0.055
	2.96	3.42	0.31	(0.167)						

4.3 Selection of copula function

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, respectively. Similarly, the bivariate 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 copula functions, and the empirical frequencies of the joint distributions were calculated by Eq. (9). Table 2 shows that for the flood occurrence dates and magnitudes, the RMSE values of the Clayton copula function are the lowest, indicating the Clayton copula function is the most appropriate copula for modeling the joint distributions of flood variables at these stations. Therefore, the Clayton copula function is selected to establish the joint probability distributions of flood dates and magnitudes, respectively. The theoretical and the observed nonexceedance joint probabilities were plotted in Fig. 5, on which the x-axis is sorted in

ascending order of the theoretical nonexceedance joint probabilities. Obviously, the theoretical frequency curves can fit the observed flood dataset well.

Table 2

Estimated parameters and test results of joint distributions for 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

4.4 Analysis of flood coincidence risks

4.4.1 Coincidence risk of flood occurrence dates

The coincidence probabilities of the flood occurrence dates at the Liaojiawan and Loujiacun stations were calculated by Eq. (13) and plotted in Fig. 6. It is shown that the coincidence probabilities of flood occurrence dates present the characteristics of multiple-peak. Before March or after August, the coincidence probabilities are basically close to zero. Early May to early July is a higher coincidence period, including two peaks, which occur on May 12 and June 21 with probabilities of 0.026% and 0.057%, respectively. Adding together all the daily coincidence probabilities, we can get that the annual coincidence risk of the flood occurrence dates in the mainstream and tributaries is equal to 2.87%.

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. The joint probabilities and co-occurrence probabilities of T -year design floods were calculated by Eq. (14) and Eq. (15), respectively. Given the occurrence of floods at the Liaojiawan or Loujiacun station, the conditional probabilities of T -year design floods at the Lijiadu station were calculated by Eq. (16). Table 3 displays the flood coincidence probabilities including the joint probabilities, co-occurrence probabilities, and conditional probabilities for different design floods.

The joint probabilities, co-occurrence probabilities for 5, 10, 20, 50 and 100-year design floods are 30.48%, 17.14%, 9.20%, 3.86%, 1.96% and 9.52%, 2.86%, 0.80%, 0.14%, 0.04%, respectively. Obviously, the joint probabilities are greater than the co-occurrence probabilities. Moreover, both the joint probabilities and co-

occurrence probabilities decrease as the return periods increase. That is, small and medium floods are more likely to occur simultaneously, which conforms to the general observations in practice.

The conditional probabilities of the same flood magnitudes occurring at the Lijiadu station are 16.41%, 5.01%, 1.46%, 0.26%, 0.07% for given 5, 10, 20, 50, and 100-year design floods at the Liaojiawan station, 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. The floods at the Liaojiawan station in the upstream have more significant impact on the downstream floods.

Table 3

Joint probabilities, co-occurrence probabilities and conditional probabilities of flood magnitudes for different return periods at Liaojiawan, Loujiacun and Lijiadu stations (%)

Coincidence probabilities	T (year)	5	10	20	50	100
Joint probabilities between Liaojiawan and Loujiacun	5	30.48	24.85	22.33	20.91	20.45
	10	24.85	17.14	13.5	11.38	10.69
	20	22.33	13.5	9.2	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
Co-occurrence probabilities between Liaojiawan and Loujiacun	5	9.52	5.15	2.67	1.09	0.55
	10	5.15	2.86	1.5	0.62	0.31
	20	2.67	1.5	0.8	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
Conditional probabilities between Liaojiawan and Lijiadu	5	16.41	9.23	4.83	1.98	0.99
	10	8.2	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	0.8	0.53	0.31	0.13	0.07
Conditional probabilities between Loujiacun and Lijiadu	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

4.5 Comparison of different estimation methods

In this study, we assumed that floods with the same frequencies simultaneously occur at the Liaojiawan and Loujiacun stations. The design flood at the Lijiadu station can be predicted by Eq. (17). Since the flood magnitude series obey the P3 distribution, the corresponding design frequencies and the return periods are calculated by Eq. (7). For given return periods of 5, 10, 20, 50, and 100 years, the design floods at the Liaojiawan and Loujiacun stations and the predicted floods at the Lijiadu station were listed in

Table 4. The corresponding frequencies and return periods of the predicted floods at the Lijiadu station were also listed in Table 4. The results show that the floods at the Lijiadu station are increasing as the rise of flood magnitudes at the Liaojiawan and Loujiacun stations. The return periods of the Lijiadu station are generally longer than those of the Liaojiawan and Loujiacun stations. For example, when 100-year design floods simultaneously occur at the Liaojiawan and Loujiacun stations, a 137-year design flood may occur at the Lijiadu station.

To compare above three flood coincidence probability estimation methods, the theoretical coincidence probabilities were calculated and compared with the design frequencies of the predicted flood at the Lijiadu station. Under the different return periods, the design frequencies of the predicted floods at the Lijiadu station, the joint probabilities and the co-occurrence probabilities when floods with the same frequency occur at the Liaojiawan and Loujiacun stations were estimated. The conditional probabilities P_{1}^{c} (or P_{2}^{c}) of the same frequency floods occurring at the Lijiadu station for given the occurrence of T -year floods at the Liaojiawan (or Loujiacun) station were also calculated. All of these estimated flood probabilities are listed in Table 4.

Compared with the design frequencies of the predicted flood at the Lijiadu station, the joint probabilities of floods simultaneously occurring in the upper mainstream and its tributaries are generally larger, while the co-occurrence probabilities or conditional probabilities are less. Among these flood probabilities, the conditional probabilities (P_{1}^{c}) are the closest to the design frequencies, especially with small return periods. For example, the joint and co-occurrence probability of 10-year floods at the Liaojiawan and Loujiacun stations are 17.14% and 2.86%, respectively, while the design frequency of the predicted 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 predicted 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. Generally, the floods at the Lijiadu station are basically formed by the superposition of the floods in the upper reaches. More importantly, the floods at the Liaojiawan station in the upstream account for a larger proportion of floods at the Lijiadu station.

Table 4
Comparison results including design floods and flood coincidence probabilities estimated by different methods

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
Predict flood at Lijiadu (m ³ /s)	6227	7397	8451	9735	10649
Return period at Lijiadu (year)	5	11	23	63	137
Predict 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

5 Conclusions

Flood coincidence risk analysis plays an important role in reservoir operation and flood management. In this study, the marginal and the joint distributions of annual maximum flood magnitude and occurrence date series were established, respectively. The coincidence risk of flood occurrence dates, and the joint, co-occurrence and conditional probabilities for different flood magnitudes were calculated and compared with the predicted flood coincidence risks. The main conclusions were summarized as follows:

- (1) There is a strong consistency and significant correlation between the floods at the upstream and downstream with the Pearson coefficient equal to 0.90. The floods at the Lijiadu station are mainly formed by the superimposition of the upstream floods. The mixed von Mises (or P3) marginal distribution can fit flood occurrence dates (or magnitudes) very well, and the Clayton copula is the best one for constructing 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. Two flood coincidence peaks occur on May 12 and June 21, and the annual coincidence risk in Fuhe River 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%. Given the occurrence of 50-year flood 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 these three coincidence probability estimation 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 reveals the spatiotemporal characteristics of floods in the Fuhe River. The feasibility and rationality of the flood coincidence probability estimation methods have been verified by considering the connection between the mainstream and its tributaries. This comparative study may provide more intuitive understanding in different flood coincidence probabilities. To improve the accuracy of flood coincidence risk estimation, other factors including climate, environment, and topography in addition to flood magnitudes and time should be considered as the variables in future studies. A comprehensive and systematic flood coincidence assessment method will provide a scientific basis and effective support for crucial decision-making and flood management.

Declarations

Acknowledgements

This study was financially supported by the National Key Research and Development Program of China (2021YFC3200301), and the National Natural Science Foundation of China (51879192).

Author contributions statement

Na Li calculated and analyzed the results, Shenglian Guo proposed the study and revised the manuscript, Feng Xiong, Jun Wang and Yuzuo Xie conducted the experiments.

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.

Availability of data and materials

The data used in this paper can be requested by contacting the corresponding author.

References

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
- Bender J, Wahl T, Müller A et al (2016) A multivariate design framework for river confluences. *Hydrolog Sci J* 61: 3,471-482. <https://doi.org/10.1080/02626667.2015.1052816>
- Bing JP, Deng PX, Zhang X et al (2018) Flood coincidence analysis of Poyang Lake and Yangtze River: risk and influencing factors. *Stoch Environ Res Risk Assess* 32: 879–891. <https://doi.org/10.1007/s00477-018-1514-4>
- 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](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000504)
- 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>
- Deidda C, Rahimi L, De Michele C (2021) Causes of dependence between extreme floods. *Environ Res Lett.* 10.1088/1748-9326/ac07d5.
- Dodangeh E, Singh VP, Pham BT et al (2020) Flood frequency analysis of interconnected rivers by copulas. *Water Resour Manag* 34: 3533–3549. <https://doi.org/10.1007/s11269-020-02634-0>
- 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>
- Frank MJ (1979) On the simultaneous associativity of $F(x,y)$ and $x+y-F(x,y)$. *Aequationes Math* 19(1): 194–226. <https://doi.org/10.1007/BF02189866>
- Gilja G, Ocvirk E, Kuspilić N (2018) Joint probability analysis of flood hazard at river confluences using bivariate copulas. *GRAĐEVINAR* 70: 267-275
- 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
- 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
- Jane R, Cadavid L, Obeysekera J et al (2020) Multivariate statistical modelling of the drivers of compound flood events in south Florida. *Nat Hazards Earth Syst Sci* 20: 2681-2699
- Karahacane H, Meddi M, Chebana F, Saaed HA (2020) Complete multivariate flood frequency analysis, applied to northern Algeria. *J Flood Risk Manag* 13: e12619
- Kvočka D, Falconer RA, Bray M (2016) Flood hazard assessment for extreme flood events. *Nat Hazards* 84(3): 1569-1599
- Moftakhari H, Schubert JE, AghaKouchak A et al (2019) Linking statistical and hydrodynamic modeling for compound flood hazard assessment in tidal channels and estuaries. *Adv Water Resour* 128: 28-38
- Montaseri M, Amirataee B, Rezaie H (2018) New approach in bivariate drought duration and severity analysis. *J Hydrol* 559: 166–181.
- Muthuvel D, Mahesha A (2021) Copula-based frequency and coincidence risk analysis of floods in tropical-seasonal rivers. *J Hydrol Eng* 26(5). <https://orcid.org/0000-0002-5903-7276>
- 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 ed. 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](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](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001744)
- Prohaska S, Ilic A (2010) Coincidence of flood flow of the Danube river and its tributaries. Springer Dordrecht. https://doi.org/10.1007/978-90-481-3423-6_6
- Schulte M, Schumann AH (2016) Evaluation of flood coincidence and retention measures by copulas. *Wasserwirtschaft* 106(2–3): 81–87
- Tsakiris G, Kordalis N, Tsakiris V (2015) Flood double frequency analysis: 2D-Archimedean copulas vs bivariate probability distributions. *Environ Process* 2: 705–716. <https://doi.org/10.1007/s40710-015-0078-2>

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>

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

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>.

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

Figures

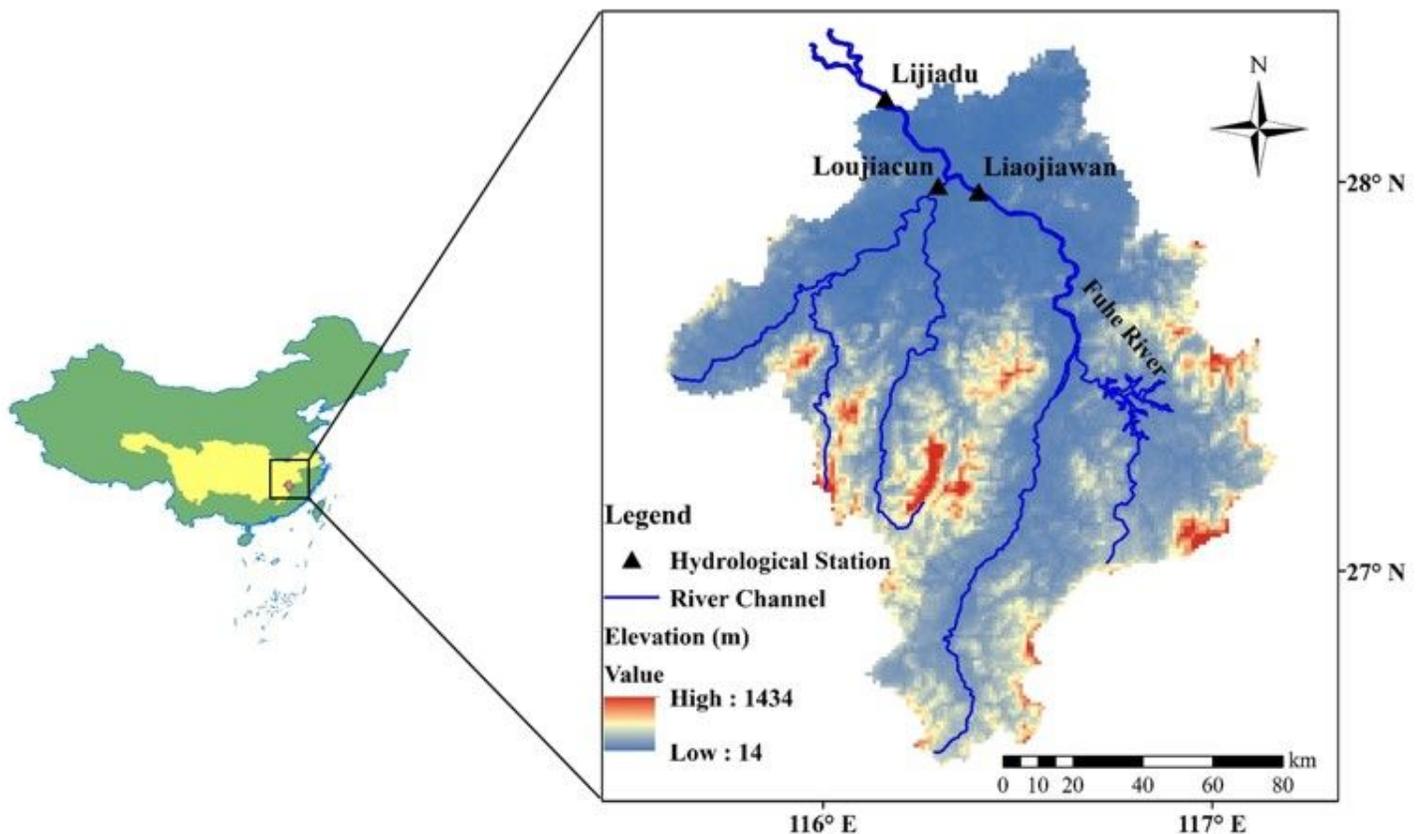
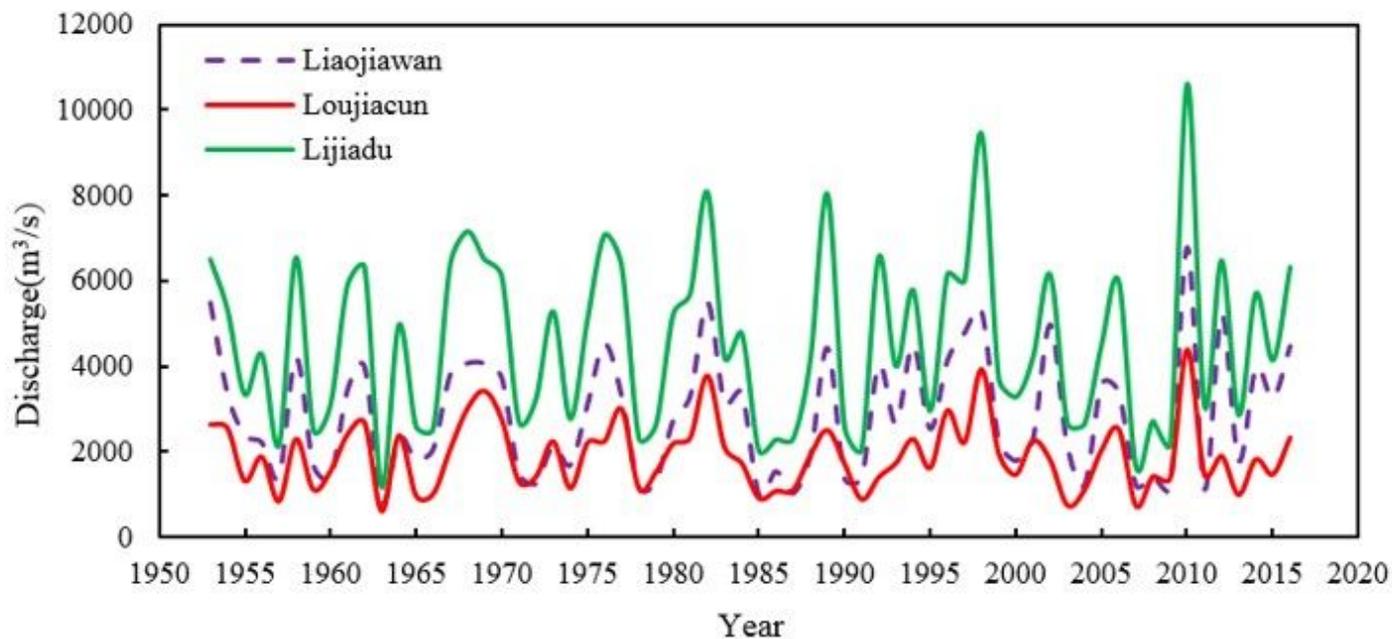
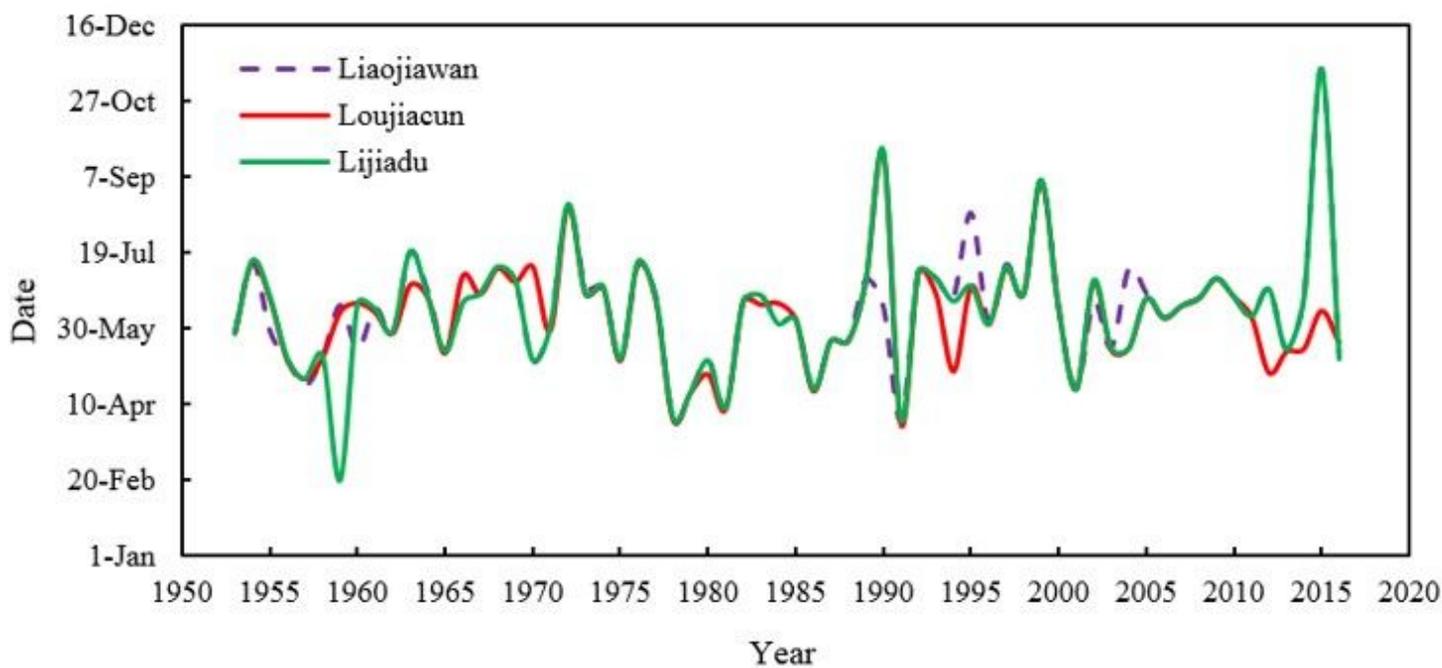


Figure 1

Sketch map of the Fuhe River basin and the location of hydrological stations



(a) Annual maximum flood magnitudes



(b) Annual maximum flood occurrence dates

Figure 2

Annual maximum flood magnitude and corresponding occurrence date series at the Liaojiawan, Loujiacun and Lijiadu hydrological stations from 1953 to 2016

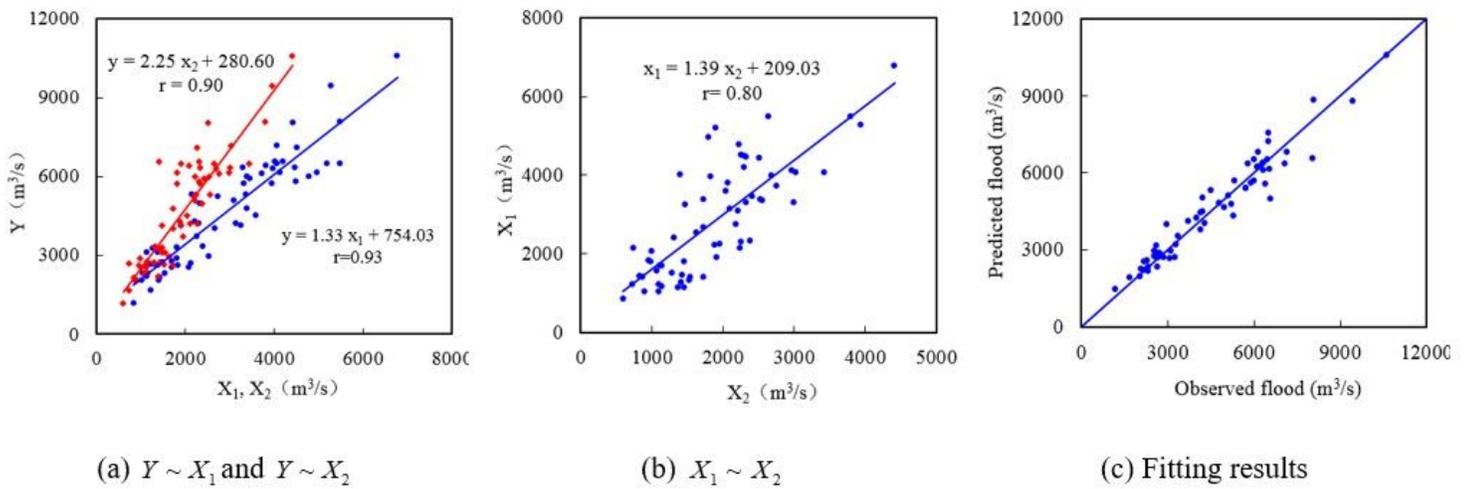
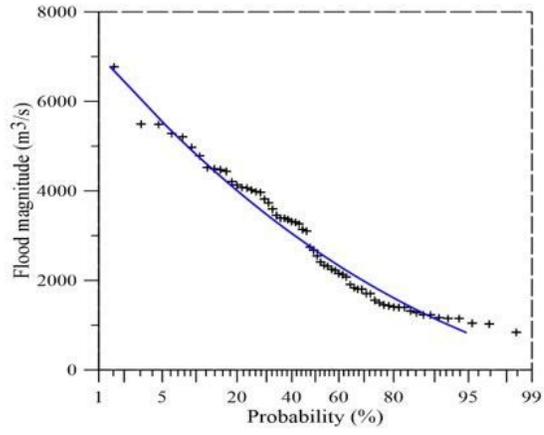
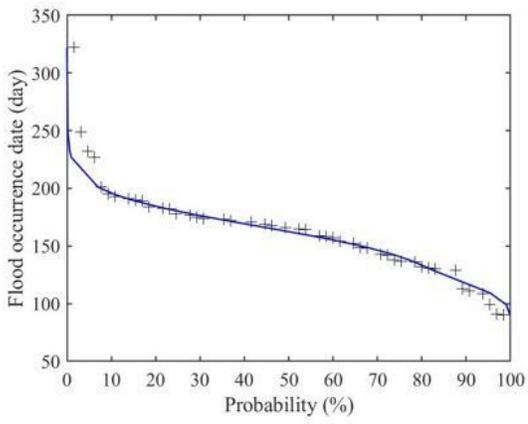
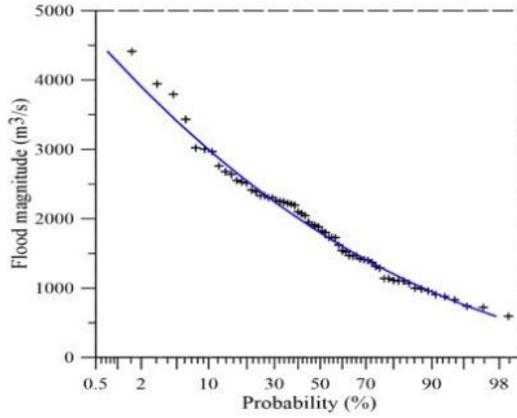
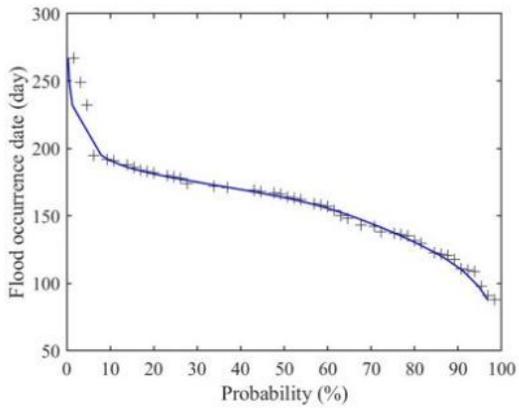


Figure 3

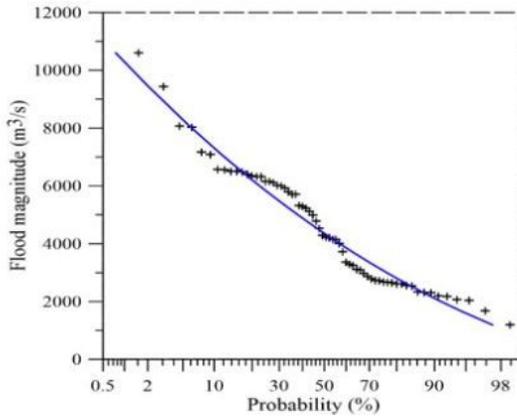
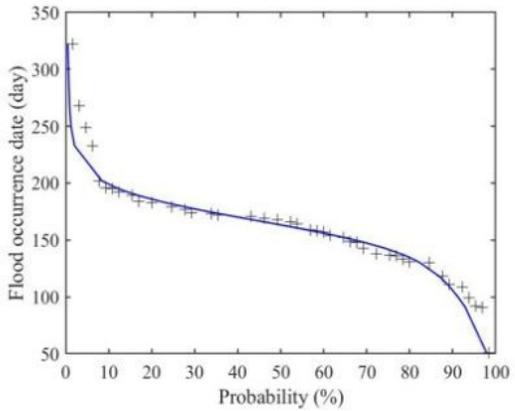
Correlation analysis of annual maximum flood magnitudes (a, b) and fitting results of observed and predicted flood magnitudes (c) at the Liaojiawan (X1), Loujiacun (X2) and Lijiadu (Y) stations



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



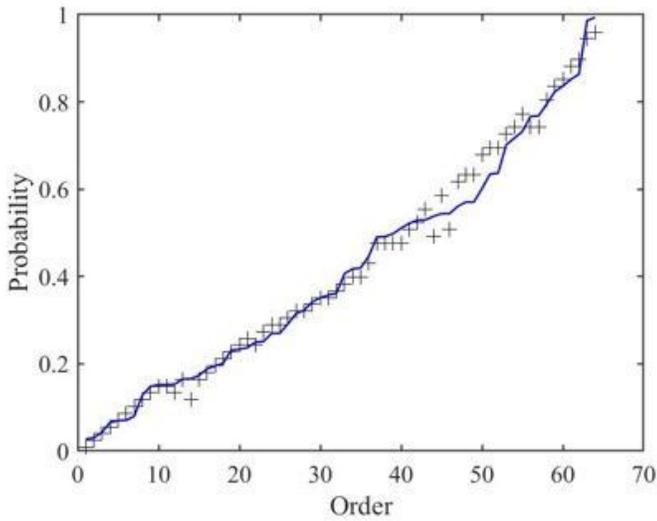
(b) Flood occurrence dates and flood magnitudes at Loujiacun station



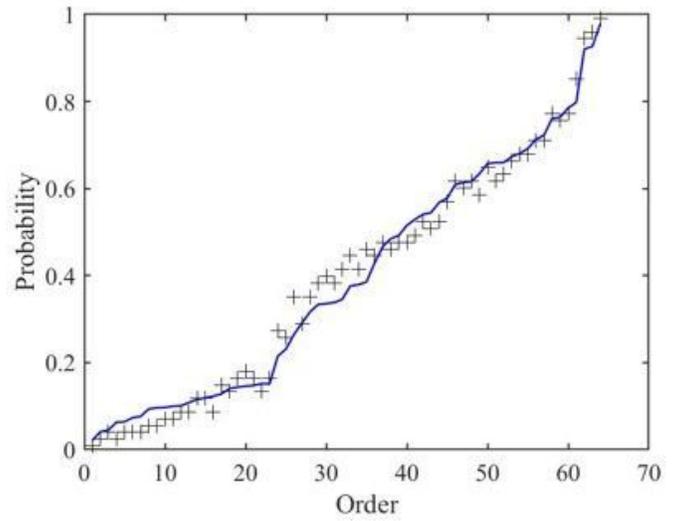
(c) Flood occurrence dates and flood magnitudes at Lijiadu station

Figure 4

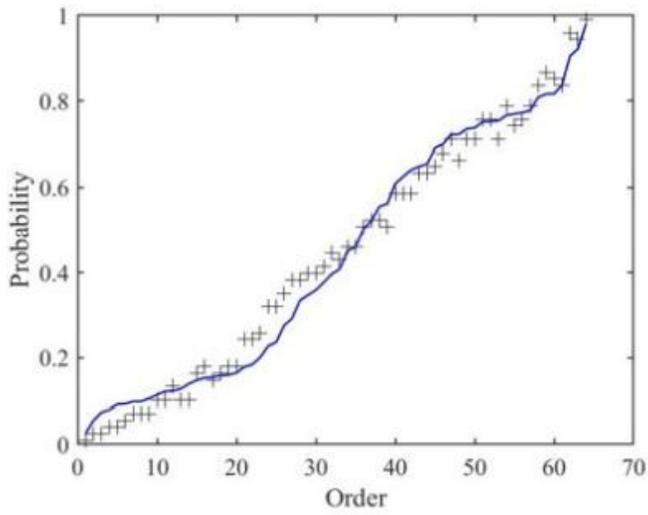
Fitted empirical points (+++) and theoretical probabilities curve (—) of marginal distributions



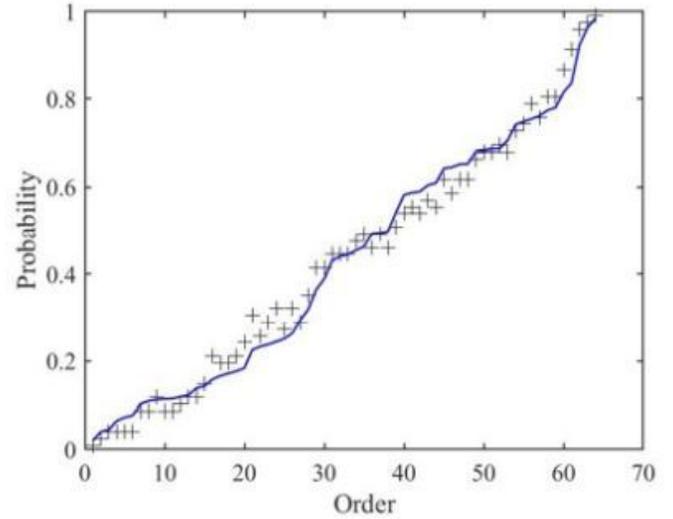
(a) Occurrence dates between Liaojiawan and Loujiacun stations



(b) Flood magnitudes between Liaojiawan and Loujiacun stations



(c) Flood magnitudes between Liaojiawan and Lijiadu stations



(d) Flood magnitudes between Loujiacun and Lijiadu stations

Figure 5

Fitted empirical points (+++) and theoretical probabilities curve (—) of joint distributions based on Clayton copula function

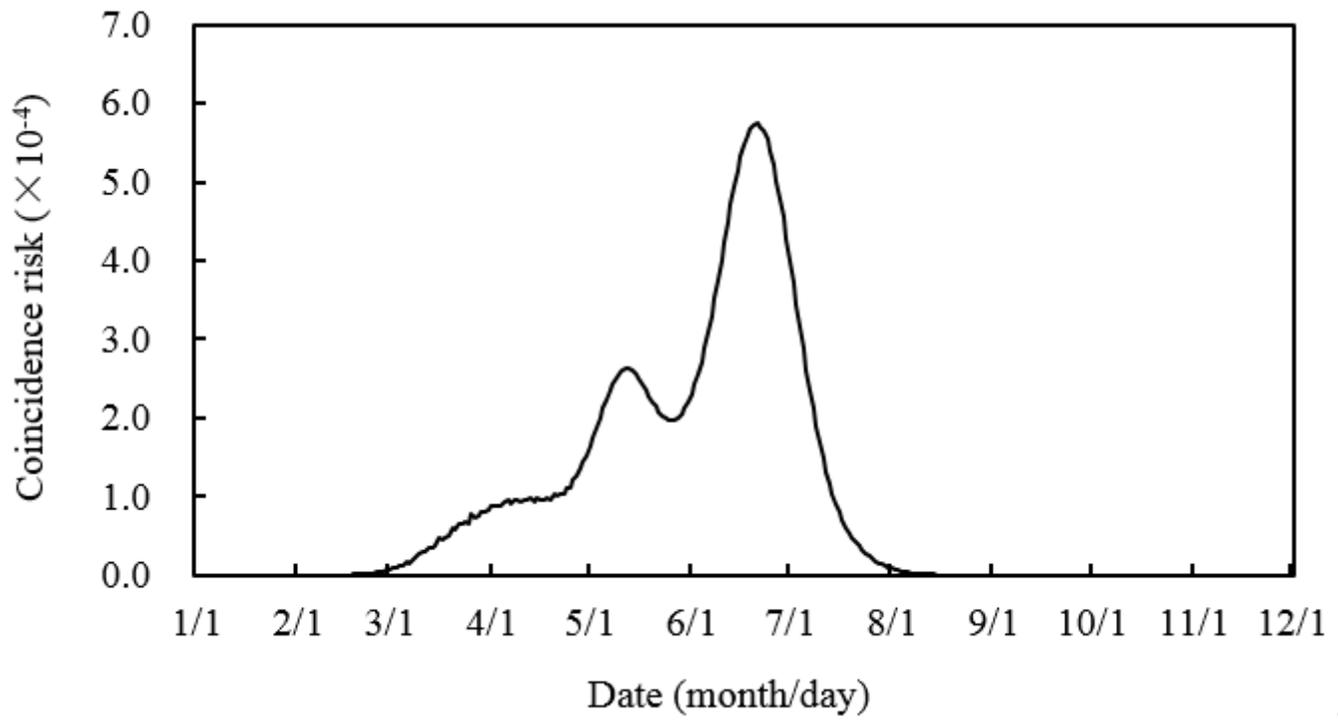


Figure 6

Daily coincidence probabilities of flood occurrence dates at Liaojiawan and Loujiacun stations