



Article

# Stochastic Flood Simulation Method Combining Flood Intensity and Morphological Indicators

Xiaodi Fu <sup>1,2,3</sup>, Xiaoyan He <sup>1,2,3,\*</sup> and Liuqian Ding <sup>1,2</sup><sup>1</sup> State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, Beijing 100038, China; fuxiaodi2006@163.com (X.F.); dinglq@iwhr.com (L.D.)<sup>2</sup> China Institute of Water Resources and Hydropower Research, Beijing 100038, China<sup>3</sup> Research Center on Flood & Drought Disaster Prevention and Reduction, Ministry of Water Resources, Beijing 100038, China

\* Correspondence: hexy@iwhr.com

**Abstract:** The existing flood stochastic simulation methods are mostly applied to the stochastic simulation of flood intensity characteristics, with less consideration for the randomness of the flood hydrograph shape and its correlation with intensity characteristics. In view of this, this paper proposes a flood stochastic simulation method that combines intensity and morphological indicators. Using the Foziling and Xianghongdian reservoirs in the Pi River basin in China as examples, this method utilizes a three-dimensional asymmetric Archimedean M6 Copula to construct stochastic simulation models for peak flow, flood volume, and flood duration. Based on K-means clustering, a multivariate Gaussian Copula is employed to construct a dimensionless flood hydrograph stochastic simulation model. Furthermore, separate two-dimensional symmetric Copula stochastic simulation models are established to capture the correlations between flood intensity characteristics and shape variables such as peak shape coefficient, peak occurrence time, rising inflection point angle, and coefficient of variation. By evaluating the fit between the simulated flood characteristics and the dimensionless flood hydrograph, a complete flood hydrograph is synthesized, which can be applied in flood control dispatch simulations and other related fields. The feasibility and practicality of the proposed model are analyzed and demonstrated. The results indicate that the simulated floods closely resemble natural floods, making the simulation outcomes crucial for reservoir scheduling, risk assessment, and decision-making processes.



Citation: Fu, X.; He, X.; Ding, L.

Stochastic Flood Simulation Method Combining Flood Intensity and Morphological Indicators.

*Sustainability* **2023**, *15*, 14032.<https://doi.org/10.3390/su151814032>

Academic Editor: Andrzej Walega

Received: 10 August 2023

Revised: 4 September 2023

Accepted: 19 September 2023

Published: 21 September 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The variation of flood processes is influenced by numerous factors, making it extremely complex and characterized by evident randomness [1]. Stochastic flood simulation involves generating a large number of flood hydrographs based on the statistical characteristics and stochastic patterns derived from historical flood observations. This method can not only be used to forecast future hydrological conditions, but also to provide fundamental data for flood control scheduling simulation calculations and the development of scheduling strategies. Therefore, stochastic flood simulation holds significant importance for formulating reservoir scheduling plans and making decisions related to risk assessment in flood control [2].

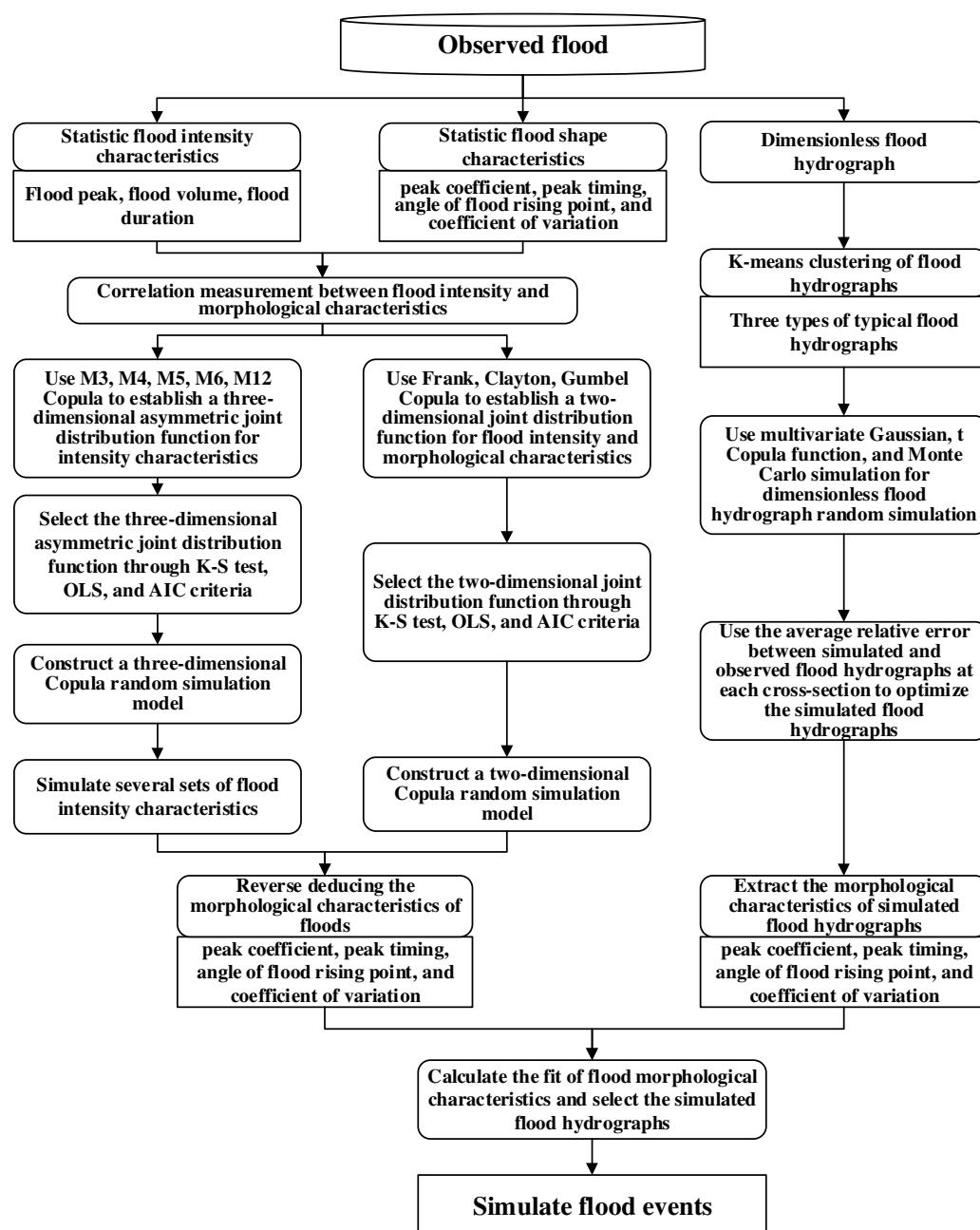
Flood events are hydrological stochastic events involving multiple variables and types. They include intensity characteristics such as flood peak, flood volume, and flood duration, as well as shape characteristics that represent flood hydrographs. Additionally, there are certain correlations among these characteristics [3]. Currently, the most widely used flood stochastic simulation models include regression-based models, set-based models, non-parametric methods, nonlinear methods, and wavelet analysis theory. However, these

models treat the flood process as a whole hydrological sequence and neglect the crucial role of flood characteristics. As a result, the accuracy of simulating flood characteristics is not high, and they are constrained by the type of marginal distributions, making it difficult to address complex multivariate joint simulation problems. Copula functions have also found extensive application in the field of hydrological stochastic simulations [4–9]. The Copula function can simulate flood characteristics, with a focus on considering the interdependencies among these characteristics. Furthermore, it offers diverse and flexible marginal distributions, significantly enhancing the accuracy and adaptability of flood stochastic simulations and finding numerous applications in practical production. For instance, Xiao and Guo [10] utilized the Gumbel Copula function to establish a two-dimensional stochastic simulation model for flood peak and flood volume, with better simulation results than traditional models. Three-dimensional flood characteristics can more effectively assess flood attributes. Gao [11] constructed a three-dimensional joint distribution model for flood peak, flood volume, and flood duration, carrying out a three-dimensional stochastic simulation for flood characteristics. Currently, most three-dimensional Copula functions employed are single-parameter functions [12]. However, for high-dimensional stochastic variables with different correlations, single-parameter Copula functions cannot accurately reflect complex asymmetric correlation structures. Asymmetric Copula functions possess more flexible parameters and forms, making them more suitable for fitting high-dimensional stochastic variables [13,14]. Ref. [15] created a three-variable asymmetric Archimedean Copula joint distribution model between flood peak and flood volume during specific time intervals, verifying the feasibility and practicality of asymmetric Copula simulation for high-dimensional stochastic variables.

Flood hydrographs possess both stochastic and correlated attributes, with their shapes varying significantly [16]. They represent a random process, and at the same time, there is a certain correlation between the flood process hydrograph and intensity features such as flood peak and flood volume. Traditional methods for simulating flood hydrographs involve using fixed single or a few typical flood hydrographs, and then simulating the flood process through equal-frequency or scaling calculations. However, these methods have significant limitations in practical application, as they fail to fully capture the stochastic of flood hydrographs and their correlations with flood intensity characteristics. Consequently, simulated flood hydrographs tend to exhibit overly uniform shapes, and may lead to unrealistic scenarios for certain combinations of peak and volume. To address these issues, Gao and Yan [17] incorporated the stochastic simulation of flood hydrograph shapes into the three-dimensional joint distribution model of flood characteristics. They employed the Monte Carlo method, combined with logarithmic, normal, and orthogonal transformations, for simulating dimensionless flood hydrographs.

Considering this, this article proposes a flood stochastic simulation method that takes into account the stochastic of flood hydrographs and the correlation between the hydrograph shape and flood intensity features. This approach builds upon the stochastic simulation model of flood characteristics, and further investigates the stochastic nature of flood hydrographs, analyzing the correlation between flood characteristics and hydrograph shapes, thereby achieving an organic integration of flood intensity characteristics with potential flood hydrographs. Firstly, a three-dimensional joint distribution function is built for flood peak, flood volume, and flood duration, and several sets of flood intensity characteristics are randomly simulated. Secondly, representative flood hydrographs are determined through cluster analysis using measured flood data. Considering the dependency between flood volumes at different time intervals, a Copula function is utilized to establish a multivariate joint distribution function for flood volumes at various time intervals. Several sets of flood volumes are then randomly simulated, based on the joint distribution function, and compared with the Monte Carlo method [18], for a dimensionless flood process hydrograph simulation. Furthermore, various characteristic parameters related to flood hydrograph shapes, such as the peak coefficient, peak timing, angle of flood rising point, and coefficient of variation, are calculated and analyzed for their correlation

with flood intensity characteristics, including flood peak, flood volume, and flood duration. The goodness of fit between each set of simulated flood intensity characteristics and dimensionless flood hydrographs is determined to identify representative flood processes. Finally, the representative dimensionless flood hydrographs are amplified according to the corresponding flood peak, flood volume, and flood duration, to obtain complete flood processes. The specific technical approach is shown in Figure 1. Considering the frequent occurrence of floods and the high flood control pressure in the Pihe River Basin, especially in the Fuziling and Xianghongdian reservoirs, the proposed method is applied to a flood stochastic simulation and simulation dispatching calculations during flood seasons. A comparison is made with observed floods to verify the applicability and superiority of this approach. This study aims to lay the foundation for the formulation of flood control scheduling schemes during the flood season for reservoir operations.



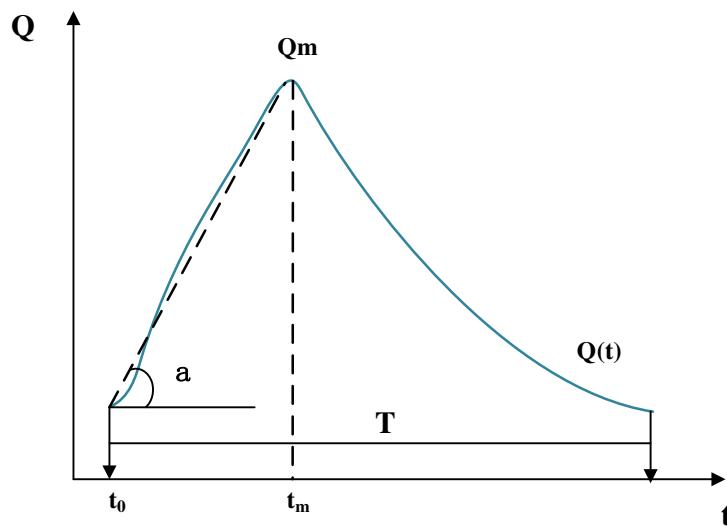
**Figure 1.** Technology Roadmap.

## 2. Materials and Methods

### 2.1. Three-Dimensional Copula Stochastic Simulation of Flood Characteristics

#### 2.1.1. Flood Characteristic Variables

Based on the basic characteristics of runoff conditions and hydrological information forecasting standards [19,20], flood process variations are generally described using two indicators: intensity and shape. The intensity indicators involved in this article mainly include flood peak, flood volume, and flood duration, while the shape indicators primarily consist of peak shape coefficient, peak timing, angle of flood rising point, and coefficient of variation. This section mainly introduces the flood shape indicators, as shown in Figure 2.



**Figure 2.** Schematic diagram of flood form characteristics.

The peak timing [21]  $T_m$  refers to the time when the flood peak appears, and it is generally taken as the initial moment when calculating the peak occurrence, starting from the moment the flood begins to rise.

$$T_m = T(Q_{t_m}) \quad (1)$$

The peak shape coefficient  $c$  refers to the ratio of the average flow before the peak to the peak flow during the flood.

$$c = \frac{\overline{Q}_{t_m}}{Q_{t_m}} \quad (2)$$

The angle of the flood rising point [22]  $\alpha$  is represented using the tangent value of the elevation angle,  $\tan\alpha$ . It is the ratio of the normalized peak flow to the pre-peak time.

$$\tan\alpha = \frac{Q_{t_m}'}{T_m'} \quad (3)$$

The coefficient of variation (CV) is the ratio of the standard deviation of the sub-flood process to the mean flow.

$$CV = \frac{\sigma[Q(t)]}{Q_{av}} \quad (4)$$

where  $t_m$  represents the time corresponding to the flood peak;  $Q_{t_m}$  denotes the flood peak.  $\overline{Q}_{t_m}$  stands for the average flow before the peak.  $Q_{t_m}'$  represents the normalized flow value.  $T_m'$  represents the normalized pre-peak time.  $Q(t)$  refers to the sub-flood process.  $\sigma[Q(t)]$  is the standard deviation of the sub-flood process.  $Q_{av}$  is the average flow of the sub-flood process.

For a flood event, the operators generally pay more attention to the rising stage, rather than the recession stage. The peak timing, peak shape coefficient, angle of flood rising point, and coefficient of variation are all important morphological indicators that characterize the rising characteristics. The peak timing reflects the time when the flood peak appears; the peak shape coefficient reflects the shape before the flood peak; the angle of flood rising point is a physical description of the flood hydrograph, indicating the overall shape of the flood as sharp and narrow or short and wide. The coefficient of variation reflects the intensity of changes in the flood fluctuation process. A larger value indicates a faster rise and fall of the flood and a more clustered process variation, making it prone to disasters in a short period.

### 2.1.2. Joint Distribution of Characteristics

Traditional joint distributions, such as multivariate normal, multivariate log-normal, etc., have certain limitations, as their marginal distributions must be the same. Copula is a multidimensional joint distribution function with a domain in [0, 1], representing a uniform distribution [23]. It can connect the marginal distributions of multiple random variables to obtain their joint distribution. Let  $X_1, X_2, \dots, X_n$  be  $n$  continuous random variables with marginal distribution functions  $F_1, F_2, \dots, F_n$ . According to Sklar's theorem, there exists an  $n$ -dimensional Copula function  $C$  that satisfies the following for any  $x \in R_n$ :

$$H(x_1, x_2, \dots, x_n) = C(F_1(x_1), F_2(x_2), \dots, F_n(x_n)) \quad (5)$$

According to different construction methods, Copula functions can generally be divided into three types: elliptical (multivariate Gaussian, multivariate Student t), quadratic, and Archimedean types (symmetric and asymmetric).

(1) Elliptical Copula is based on the elliptical distribution. The most commonly used elliptical Copula functions include Gaussian Copula [24] and Student t Copula [25].

#### ① Gaussian Copula

$$C(u_1, u_2, u_3; \Sigma) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \int_{-\infty}^{\Phi^{-1}(u_3)} \frac{1}{(2\pi)^{3/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2} \omega^T \Sigma^{-1} \omega\right) d\omega \quad (6)$$

where  $\Sigma = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix}$  is a symmetric and positive definite correlation matrix, with  $-1 \leq \rho_{ij} \leq 1$  ( $i, j = 1, 2, 3$ ).  $\Phi^{-1}$  represents the inverse of the univariate standard normal distribution function, and its bivariate margins are also characterized by the Gaussian Copula.

#### ② Student t Copula

$$\begin{aligned} C(u_1, u_2, u_3; \Sigma, v) &= T_{\Sigma, v}(T_v^{-1}(u_1), T_v^{-1}(u_2), T_v^{-1}(u_3)) \\ &= \int_{-\infty}^{T_v^{-1}(u_1)} \int_{-\infty}^{T_v^{-1}(u_2)} \int_{-\infty}^{T_v^{-1}(u_3)} \frac{\Gamma(\frac{v+3}{2})}{\Gamma(\frac{v}{2})} * \frac{1}{(\pi v)^{\frac{3}{2}} |\Sigma|^{\frac{1}{2}}} \left(1 + \frac{w^T \Sigma^{-1} w}{v}\right)^{\frac{v+3}{2}} dw \end{aligned} \quad (7)$$

where  $T_v^{-1}(.)$  represents the inverse function of the univariate Student t distribution.  $T_{\Sigma, v}(T_v^{-1}(u_1), T_v^{-1}(u_2), T_v^{-1}(u_3))$  denotes the multivariate Student t distribution function.

$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix}$  is a correlation matrix,  $\rho_{ij} = \begin{cases} 1; i = j \\ \rho_{ji}; i \neq j \end{cases} -1 \leq \rho_{ij} \leq 1$  ( $i, j = 1, 2, 3$ ).

(2) Archimedean Copula [26] functions are currently widely used Copula functions, known for their simplicity and ability to construct various forms of multivariate joint distribution functions with strong adaptability. They have extensive practical applications and are also the most commonly used functions in the field of hydrology. Archimedean Copula functions can be classified into symmetric and asymmetric types.

Taking the three-dimensional case as an example, the commonly used Copula functions [27] in symmetric Archimedean Copulas are shown in Table 1.

**Table 1.** Formula and parameter range of three-dimensional symmetric Archimedean Copula functions.

Copula	Function Expressions	Parameter Ranges
Gumbel	$C(u_1, u_2, u_3; \theta) = \exp\left\{-\left[(-\ln u_1)^\theta + (-\ln u_2)^\theta + (-\ln u_3)^\theta\right]^{1/\theta}\right\}$	$\theta \in [1, \infty)$
Frank	$C(u_1, u_2, u_3; \theta) = -\frac{1}{\theta} \ln\left\{1 + \frac{[\exp(-\theta u_1) - 1][\exp(-\theta u_2) - 1][\exp(-\theta u_3) - 1]}{[\exp(-\theta) - 1]^2}\right\}$	$\theta \in R, \theta \neq 0$
Clayton	$C(u_1, u_2, u_3; \theta) = \left[u_1^{-\theta} + u_2^{-\theta} + u_3^{-\theta} - 2\right]^{-1/\theta}$	$\theta \in (0, \infty)$
Ali-Mikhail-Haq	$C(u_1, u_2, u_3; \theta) = \frac{u_1 u_2 u_3}{[1 - \theta(1 - u_1)(1 - u_2)(1 - u_3)]^{-1/\theta}}$	$\theta \in [-1, 1]$

In the equation,  $u_1, u_2, u_3$  represent the marginal distribution functions, where  $u_1 = F_1(x)$ ,  $u_2 = F_2(x)$ ,  $u_3 = F_3(x)$ ;  $\theta$  is the parameter of the Copula function.

Asymmetric Archimedean Copula is a “fully nested” Copula proposed by Joe H, Nelsen RB, Embrechts P, Lingdskog F [28], and others, based on the study of two-dimensional Archimedean Copulas. Taking the three-dimensional case as an example, the expressions for five common asymmetric Archimedean Copulas are shown in Table 2 [29]:

**Table 2.** Formula of three-dimensional asymmetric Archimedean Copula function.

Copula	Function Expressions	Parameter Ranges
M3	$C(u_1, u_2, u_3; \theta_1, \theta_2) = -\frac{1}{\theta_1} \log\left\{1 - \left(1 - e^{-\theta_1}\right)^{-1} \left(1 - e^{-\theta_1 u_3}\right) \left(1 - \left[1 - \left(1 - e^{-\theta_2}\right)^{-1} \left(1 - e^{-\theta_2 u_1}\right) \left(1 - e^{-\theta_2 u_2}\right)\right]^{\frac{1}{\theta_2}}\right)\right\}$	$\theta_2 \geq \theta_1 \in [0, \infty]$
M4	$C(u_1, u_2, u_3; \theta_1, \theta_2) = \left[u_3^{-\theta_1} + \left(u_1^{-\theta_2} + u_2^{-\theta_2} - 1\right)^{\frac{1}{\theta_2}} - 1\right]^{-\frac{1}{\theta_1}}$	$\theta_2 \geq \theta_1 \in [0, \infty]$
M5	$C(u_1, u_2, u_3; \theta_1, \theta_2) = 1 - \left\{\left[(1 - u_1)^{\theta_2} \left(1 - (1 - u_2)^{\theta_2}\right) + (1 - u_2)^{\theta_2}\right]^{\frac{1}{\theta_2}} \left(1 - (1 - u_3)^{\theta_1}\right) + (1 - u_3)^{\theta_1}\right\}^{\frac{1}{\theta_1}}$	$\theta_2 \geq \theta_1 \in [1, \infty)$
M6	$C(u_1, u_2, u_3; \theta_1, \theta_2) = \exp\left\{-\left[\left((-ln u_1)^{\theta_2} + (-ln u_3)^{\theta_2}\right)^{\theta_1/\theta_2} + (-ln u_3)^{\theta_1}\right]^{1/\theta_1}\right\}$	$\theta_2 \geq \theta_1 \in [1, \infty)$
M12	$C(u_1, u_2, u_3; \theta_1, \theta_2) = \left\{1 + \left[\left(\frac{1}{u_3} - 1\right)^{\theta_1} + \left(\left(\frac{1}{u_1} - 1\right)^{\theta_2} + \left(\frac{1}{u_2} - 1\right)^{\theta_2}\right)^{\frac{1}{\theta_2}}\right]^{\frac{1}{\theta_1}}\right\}^{-1}$	$\theta_2 \geq \theta_1 \in [1, \infty)$

### 2.1.3. Construction of Copula Joint Distribution Models

The construction of the Copula joint distribution model mainly involves several steps, including the selection of flood characteristic indicators, determination of marginal distribution functions, correlation measurement, parameter estimation, testing, and goodness-of-fit evaluation.

After selecting the flood characteristics and individual marginal distributions, we focus on measuring the correlation between the characteristics. The pairwise correlation between characteristics determines the choice of joint distribution type, typically computed using Kendall and Spearman rank correlation coefficients. In commonly used Copula functions, parameter estimation methods include maximum likelihood, correlation-based indicators, and the method of moments. For two-dimensional functions, the correlation-based indicator method is often employed for indirect estimation, while, for high-dimensional functions, the maximum likelihood method is generally used to estimate parameters.

Finally, the Copula joint distribution function is tested and optimized. The Kolmogorov-Smirnov non-parametric test is used to verify whether the joint distribution of characteris-

tics represents the overall distribution type. For the various alternative Copula functions obtained through hypothesis testing, the Genest–Rivest plot [30] is used to visually compare the fit between empirical joint distribution function values and theoretical joint distribution function values. When the simulation results are similar, the Ordinary Least Squares (OLS) and Akaike Information Criterion (AIC) [31] are used to evaluate the fit discrepancies and select the optimal joint distribution model.

#### 2.1.4. Stochastic Simulation of Flood Characteristic Variables

After obtaining the  $n$ -dimensional joint distribution of flood characteristic variables  $X_1, X_2, \dots, X_n$  (where the joint distributions of dimensions  $1, 2, \dots, n - 1$  are also known), the steps for the stochastic simulation of each characteristic variable are as follows:

- (1) Generate  $n$  independent random numbers  $k_1, k_2, k_i, \dots, k_n$ , ( $1 \leq i \leq n$ ), following the uniform distribution on the interval  $[0, 1]$ .
- (2) Set  $k_1$  to be the probability of  $X_1$  not exceeding  $F(x_1)$ , i.e.,  $F(x_1) = k_1$ . Calculate  $x_1$  using  $x_1 = F^{-1}(k_1)$ .
- (3) Set  $k_2$  to be the conditional probability distribution value of  $X_2$  given  $X_1 = x_1$ , i.e.,  $F(x_2|x_1) = k_2$ . Then compute  $x_2$  using  $x_2 = F^{-1}(k_2|X_1 = x_1)$ .
- (4) For  $k_i$ , where  $i > 2$ , set it as the conditional probability distribution value of  $x_i$  given  $X_1 = x_1, X_2 = x_2, \dots, X_{i-1} = x_{i-1}$ , i.e.,  $F(x_i|x_{i-1}, \dots, x_2, x_1) = k_i$ . Calculate  $x_i$  using  $x_i = F^{-1}(k_i|X_1 = x_1, X_2 = x_2, \dots, X_{i-1} = x_{i-1})$ .
- (5) Repeat steps (2) to (4) until  $i = n$ , completing one random simulation.
- (6) Repeat steps (1) to (5) a total of  $H$  times to obtain  $H$  sets of correlated flood characteristic variables.

### 2.2. Classification and Stochastic Simulation of Flood Hydrographs

#### 2.2.1. Classification of Flood Hydrographs

During a flood event, the flow continuously changes over time, resulting in the randomness and diversity of flood hydrograph shapes. Additionally, the shapes of different flood hydrographs are influenced by peak, volume, and duration of the flood. Corresponding changes in reservoir operation and water resources management measures are made, based on different types of flood hydrographs, such as those with early, mid, or late peaks. Therefore, in-depth research on flood hydrograph types requires classifying the shapes of flood hydrographs and removing the influence of flood characteristic variables. This allows the flood intensity to be the sole factor affecting different flood types over time, achieved through nondimensionalization of flood hydrographs as shown in Equations (9) and (10).

In this study, the K-means clustering algorithm [32] is used to cluster the observed flood data, obtaining several representative flood hydrographs.

Given a sample set  $D = \{x_1, x_2, \dots, x_m\}$ , the K-means algorithm aims to minimize the squared error for the cluster partition  $C = \{C_1, C_2, \dots, C_k\}$ .

$$E = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|_2^2 \quad (8)$$

where  $\mu_i = \frac{1}{|C_i|} \sum_{x \in C_i} x$  represents the mean vector of cluster  $C_i$ . This expression partially characterizes the compactness of the samples around the cluster mean vector. A smaller value of  $E$  indicates a higher similarity among the samples within the cluster.

Using the K-means algorithm for flood hydrograph clustering, to remove the influence of flood characteristic variables, and ensure that the variation of flood intensity over time is the sole factor affecting different flood types, it is necessary to first normalize the flood hydrographs.

$$\tau = \frac{t}{T} \quad (9)$$

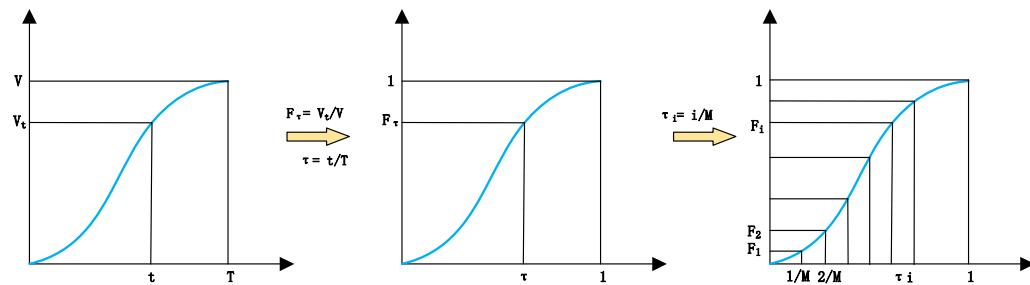
$$F_\tau = \frac{V_t}{V} \quad (10)$$

where  $T$  represents the flood duration;  $\tau$  denotes the non-dimensional time at time  $t$ , with  $\tau \in (0, 1]$ ;  $V_t$  is the accumulated flood volume at time  $t$ ;  $V$  is the total flood volume of a flood event;  $F_\tau$  represents the non-dimensional cumulative flood volume, which represents the accumulated percentage of flood volume over time, with  $F_\tau \in (0, 1]$ .

Furthermore, the non-dimensionalized cumulative flood volume curve for each flood event is partitioned into  $M$  equal time intervals, with non-dimensional time ( $\tau$ ) as follows:

$$\tau_i = \frac{i}{M}, i = 1, 2, \dots, M \quad (11)$$

Afterward, non-dimensional time  $\tau_i$  is used to interpolate the cumulative flood volume curve, resulting in the corresponding non-dimensional cumulative flood volume  $F_\tau$ . The detailed process is illustrated in Figure 3.



**Figure 3.** Dimensionalization and interpolation process of flood.

Finally, the non-dimensional flood volumes of each flood event in the  $M$  time intervals are input into the K-means algorithm for clustering, resulting in  $K$  representative types of typical non-dimensional flood hydrographs.

## 2.2.2. Non-Dimensional Flood Hydrograph Stochastic Simulation

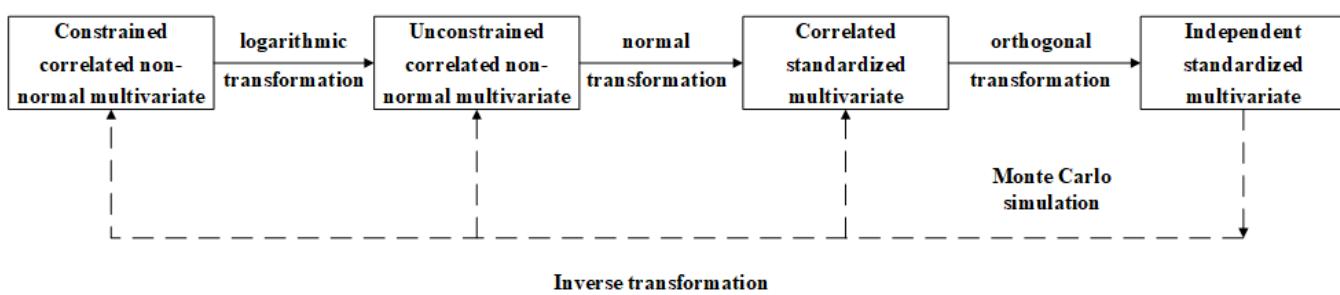
The core of simulating flood hydrographs lies in generating non-dimensional cumulative flood volume values, where  $0 \leq F_i \leq 1$  ( $i = 1, 2, \dots, K$ ). Essentially, this can be transformed into a problem of generating non-dimensional flood increments in each time interval.

$$P_1 = F_1 \quad (12)$$

$$P_i = F_i - F_{i-1} \quad (i = 2, 3, \dots, K) \quad (13)$$

The non-dimensional flood increment values, denoted as  $P_i$ , must satisfy the following constraints: ①  $P_1 + P_2 + \dots + P_K = 1$ ; ②  $0 \leq P_i \leq 1$  ( $i = 1, 2, \dots, K$ ). Considering that the  $P_i$  values in each time interval are mutually dependent non-normal variables, a multivariate joint distribution of non-dimensional flood increments in each time interval is established, based on the Copula functions. The joint distribution function is used for stochastic simulation of the flood increment values  $P_i$  in each time interval. In this study, elliptical Copula functions, which exhibit good performance in representing interdependence between multivariate variables, are employed to construct the joint distribution model for flood volume increments in each time interval and perform stochastic simulations.

In order to compare the stochastic simulation methods of the multivariate Copula-based flood volume increments, as described above, and following the methods from the literature, a Monte Carlo simulation is utilized to stochastically generate unconstrained independent normal multivariate variables for reverse calculation of non-dimensional flood volume increments in each time interval. The specific method is shown in Figure 4:



**Figure 4.** Conversion process of simulation problem.

First, a logarithmic transformation is applied to convert the correlated non-normal multivariate variables under constraints into correlated non-normal multivariate variables without constraints. Next, the Johnson system function is used to transform the correlated non-normal multivariate variables into correlated standard normal multivariate variables. Then, the Schmidt orthogonalization method is employed to obtain the orthogonal transformation matrix, which converts the correlated standard normal multivariate variables into independent standard normal multivariate variables. Finally, a Monte Carlo simulation is utilized to stochastically generate multidimensional normal random variables. Subsequently, the inverse transformation is performed to obtain the correlated non-normal random variables under the specified constraints.

### 2.3. Integration of Flood Characteristics and Flood Hydrographs

Flood shape characteristics have a certain correlation with flood peak, flood volume, and flood duration. Based on the flood shape characteristics at a certain flood peak, flood volume, and flood duration, suitable non-dimensional flood hydrographs are selected. This ensures that the simulated flood hydrographs adhere to the actual occurrence pattern of floods.

The morphological characteristics of the observed flood hydrographs, such as the peak shape coefficient, peak timing, angle of flood rising point, and coefficient of variation, are statistically calculated. Then, the correlation between these morphological characteristics and flood intensity characteristics is measured for pairwise combinations. The morphological characteristics that exhibit good correlation with flood intensity characteristics are selected to establish a joint distribution model. The fit between the morphological characteristics related to the simulated intensity characteristics and the corresponding values of each simulated flood process is calculated, and the representative flood process with the highest fit is chosen for magnification.

$$e = 1 - |T_h - T^k| - |c_h - c^k| - |Ta_h - Ta^k| - |CV_h - CV^k| \quad (14)$$

where  $e$  represents the goodness of fit;  $T_h$ ,  $c_h$ ,  $Ta_h$ , and  $CV_h$  are the non-dimensional peak timing, peak shape coefficient, angle of flood rising point, and coefficient of variation of flood characteristic variables for the  $h$ -th group of floods, respectively; and  $T^k$ ,  $c^k$ ,  $Ta^k$ , and  $CV^k$  are the non-dimensional peak timing, peak shape coefficient, angle of flood rising point, and coefficient of variation of the  $k$ -th representative flood hydrograph, respectively.

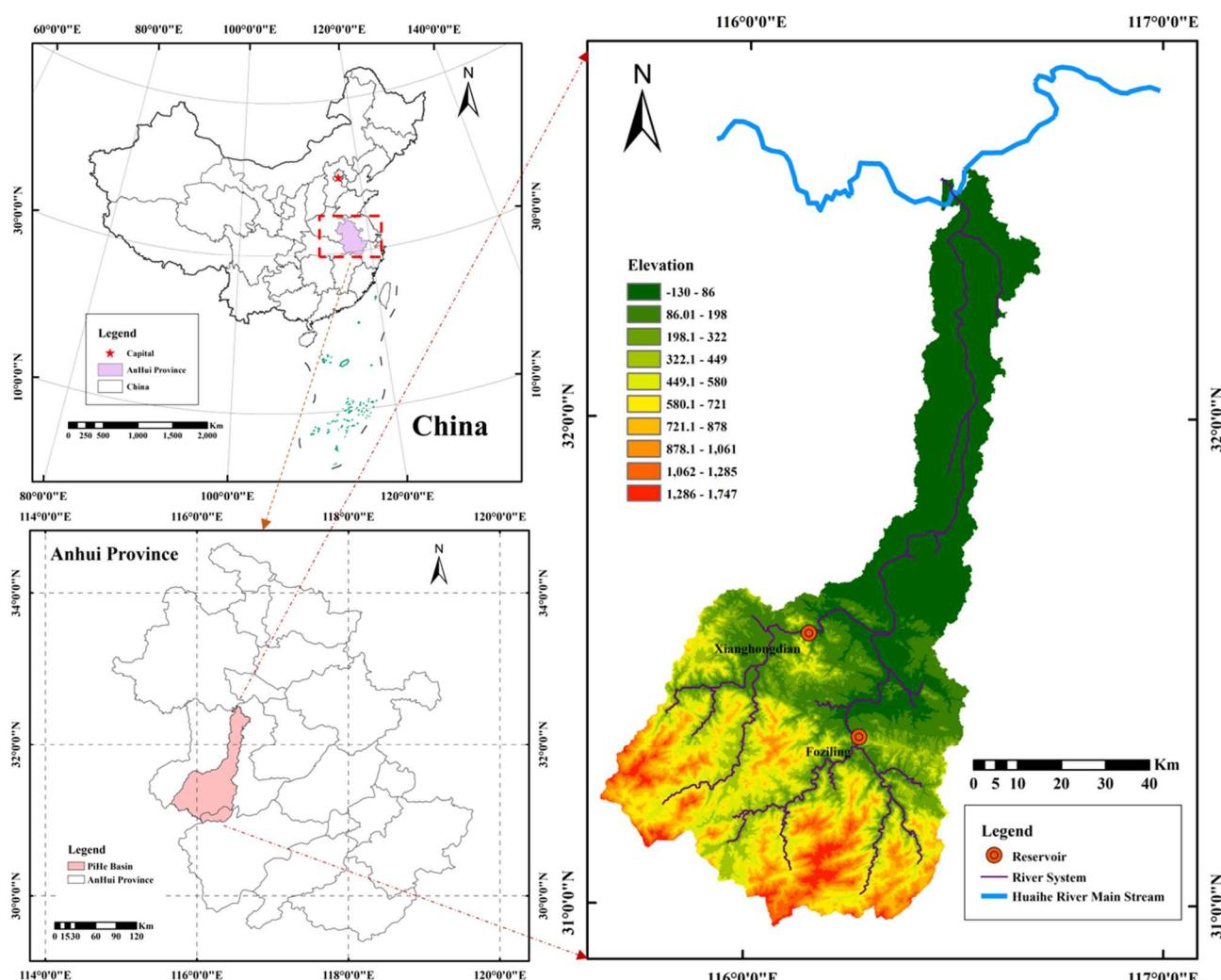
By calculating the goodness of fit, the representative flood hydrographs corresponding to each set of flood characteristic variables are determined. The corresponding type of flood hydrograph is then integrated with the flood peak, flood volume, and flood duration to generate a complete flood hydrograph.

## 3. Case Study

### 3.1. Study Area Overview

Fuziling and Xianghongdian Reservoir belong to the Huai River Basin, specifically within the Pi River system, located in the middle and upper reaches of East Pi River in

Huoshan County, and the upper reaches of West Pi River in Jinzhai County, respectively, both situated in Anhui Province. The geographical location of the study area is shown in Figure 5. Fuziling Reservoir has a controlled drainage area above the dam of  $1840 \text{ km}^2$  and is a large (2) type reservoir designed primarily for flood control and irrigation, with additional functions for power generation and water supply. Xianghongdian Reservoir has a controlled drainage area above the dam of  $1400 \text{ km}^2$  and is a large (1) type reservoir designed primarily for flood control and irrigation, with additional functions for power generation and water supply. Both reservoirs serve to protect downstream areas, including towns in Liuan City, the Hewu and Ningxi railways, G35, G42 expressways, G312 national road, and other essential infrastructures. They safeguard a population of approximately 1.3 million people and about 48,000 hectares of arable land. Xianghongdian Reservoir also plays a role in flood peak mitigation for the Huai River mainstream. Both reservoirs are situated in the subtropical continental monsoon zone, with mild and humid climates throughout the year. Frequent interactions between warm and cold air masses from the north and south, along with cyclone activities and the influence of land uplift from the Dabie Mountains and typhoon landfalls, often lead to concentrated rainfall events. Therefore, conducting flood stochastic simulation studies for Fuziling and Xianghongdian Reservoirs is of significant importance for their flood control and safety during the flood season. For this study, a total of 185 observed flood events from 1964 to 2020 for Fuziling Reservoir and 171 observed flood events from 1964 to 2020 for Xianghongdian Reservoir were selected for the extraction and analysis of flood characteristics.



**Figure 5.** Geographical location of the study area.

### 3.2. Measurement of Flood Characteristics Correlation

The flood intensity characteristics selected in this study include flood peak, flood volume, and flood duration, while the shape characteristics consist of peak timing, peak shape coefficient, angle of flood rising point, and coefficient of variation. The Pearson correlation coefficients between each characteristic are presented in Table 3. Data marked with “\*\*\*” and “\*\*” indicate significant correlations at the 0.01 and 0.05 significance level respectively.

**Table 3.** Correlation coefficient of flood characteristic quantity.

Name of Reservoir	Characteristic	Flood Peak	Flood Volume	Flood Duration	Peak Shape Coefficient	Peak Timing	Angle of Flood Rising Point	Coefficient of Variation
Fuziling	flood peak	1	0.768 **	0.335 **	-0.492 **	0.203 **	0.364 **	0.663 **
	flood volume	0.768 **	1	0.614 **	-0.334 **	0.432 **	0.153 *	0.296 **
	flood duration	0.335 **	0.614 **	1	-0.312 **	0.597 **	0.156 *	0.175 *
	peak shape coefficient	-0.492 **	-0.334 **	-0.312 **	1	-0.227 **	-0.338 **	-0.685 **
	peak timing	0.203 **	0.432 **	0.597 **	-0.227 **	1	-0.223 **	-0.002
	angle of flood rising point	0.364 **	0.153 *	0.156 *	-0.338 **	-0.223 **	1	0.630 **
	coefficient of variation	0.663 **	0.296 **	0.175 *	-0.685 **	-0.002	0.630 **	1
Xianghongdian	flood peak	1	0.866 **	0.351 **	-0.292 **	0.188 *	0.351 **	0.483 **
	flood volume	0.866 **	1	0.528 **	-0.087	0.256 **	0.281 **	0.250 **
	flood duration	0.351 **	0.528 **	1	-0.201 **	0.504 **	0.236 **	0.257 **
	peak shape coefficient	-0.292 **	-0.087	-0.201 **	1	-0.230 **	-0.180 *	-0.588 **
	peak timing	0.188 **	0.256 **	0.504 **	-0.230 **	1	-0.350 **	0.087
	angle of flood rising point	0.351 **	0.281 **	0.236 **	-0.180 *	-0.350 **	1	0.631 **
	coefficient of variation	0.483 **	0.250 **	0.257 **	-0.588 **	0.087	0.631 **	1

From Table 3, it can be observed that both Fuziling and Xianghongdian Reservoirs exhibit significant positive correlations between flood peak, flood volume, and flood duration. For Fuziling Reservoir, there are significant positive correlations between flood peak and the coefficient of variation, as well as between flood duration and peak timing. However, there is a significant negative correlation between flood volume and the peak shape coefficient. As for Xianghongdian Reservoir, significant positive correlations are found between flood peak flow and the coefficient of variation, flood volume and angle of flood rising point, and flood duration and peak occurrence time. Based on these correlation patterns, three-dimensional Copula functions are used to establish joint distribution models for flood peak, flood volume, and flood duration for both Fuziling and Xianghongdian Reservoirs. Additionally, two-dimensional Copula functions are applied to establish joint distribution models for flood peak and coefficient of variation, flood duration and peak timing, as well as flood volume and peak shape coefficient for Fuziling Reservoir; and for flood peak and coefficient of variation, flood volume and angle of flood rising point, and flood duration and peak timing for Xianghongdian Reservoir.

### 3.3. Copula Simulation of Flood Characteristics

Based on the correlation analysis of flood characteristics, Copula functions are used to construct multivariate joint distribution functions. The commonly used distributions in hydrological frequency analysis, namely, normal distribution, log-logistic distribution [33], Weibull distribution, Generalized Extreme Value distribution (GEV) [34], and

gamma distribution, are fitted to the samples of flood peak, flood volume, duration, peak shape coefficient, peak timing, angle of flood rising point, and coefficient of variation for Fuziling and Xianghongdian Reservoirs. The marginal distributions of flood peak and flood duration for the Fuziling Reservoir are determined to be log-logistic distributions, while the total flood volume and angle of flood rising point are modeled as GEV (Generalized Extreme Value) distributions. The shape coefficient follows a Weibull distribution, and the coefficient of variation is represented by a gamma distribution. For the Xianghongdian Reservoir, the flood peak, total flood volume, angle of flood rising point, and coefficient of variation are modeled as log-logistic distributions, while flood duration and time of peak occurrence are modeled as GEV distributions. The distribution parameters are presented in Table 4.

**Table 4.** Edge distribution parameters.

Name of Reservoir	Characteristics	Marginal Distributions	Mu (a)	Sigma (b)	Morphological Parameters (k)
Foziling	flood peak	Log-Logistic	6.7773	0.5767	
	flood volume	GEV	0.5554	0.4457	0.5263
	flood duration	Log-Logistic	4.3966	0.2786	
	Peak timing	GEV	18.7545	12.4247	0.3440
	peak shape coefficient	Weibull	0.3887	2.1895	
	coefficient of variation	Gamma	5.7937	0.1532	
Xianghongdian	flood peak	Log-Logistic	6.9753	0.4746	
	flood volume	Log-Logistic	-0.3014	0.4878	
	flood duration	GEV	55.223	24.3223	0.0925
	Peak timing	GEV	18.5201	11.0115	0.2022
	angle of flood rising point	Log-Logistic	-1.0181	0.3163	
	coefficient of variation	Log-Logistic	-0.1648	0.1887	

Five types of non-symmetric Archimedean Copula functions (M3, M4, M5, M6, M12) are used to construct the joint distribution functions of flood peak, flood volume, and flood duration, which represent the three-dimensional flood intensity characteristics for both reservoirs. Additionally, three types of symmetric Archimedean Copula functions (Frank, Clayton, Gumbel) are used to construct the joint distribution functions between flood intensity and shape characteristics. The goodness-of-fit is evaluated through the Kolmogorov–Smirnov test, and the function types are further selected based on the OLS and AIC criteria. The Copula parameters and the results of goodness-of-fit evaluations are presented in Tables 5 and 6, respectively.

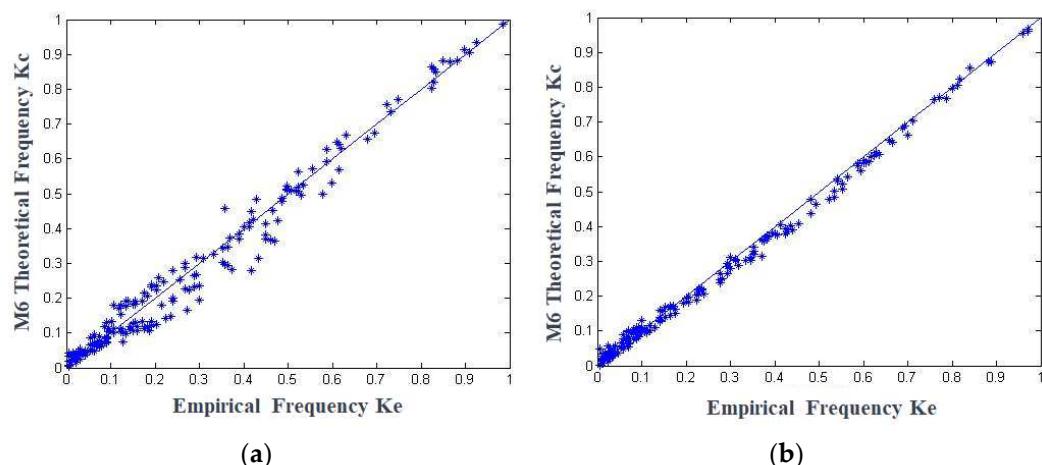
**Table 5.** Three-dimensional asymmetric Archimedean Copula parameters and goodness of fit evaluation results of flood peak, flood volume, and flood duration.

Name of Reservoir	Copula	$\theta_1$	$\theta_2$	H	P (%)	OLS	AIC
Foziling	M3	13.0846	9.4238	0	17.26	0.0964	-861.392
	M4	2.3086	3.4126	1	0.018	0.0769	-945.2933
	M5	1.5411	2.7317	0	17.26	0.0441	-1150.793
	M6	1.3963	2.5316	0	73.59	0.0405	-1181.948
	M12	1.0395	1.9459	0	13.49	0.0470	-1115.736
Xianghongdian	M3	14.6957	10.5305	0	17.76	0.0941	-801.3896
	M4	2.7373	3.3566	1	0.12	0.0713	-896.1163
	M5	1.5245	3.3884	0	13.86	0.0316	-1172.817
	M6	1.3904	2.939	0	85.75	0.0205	-1319.596
	M12	1.0348	2.1453	1	2.43	0.0336	-1151.915

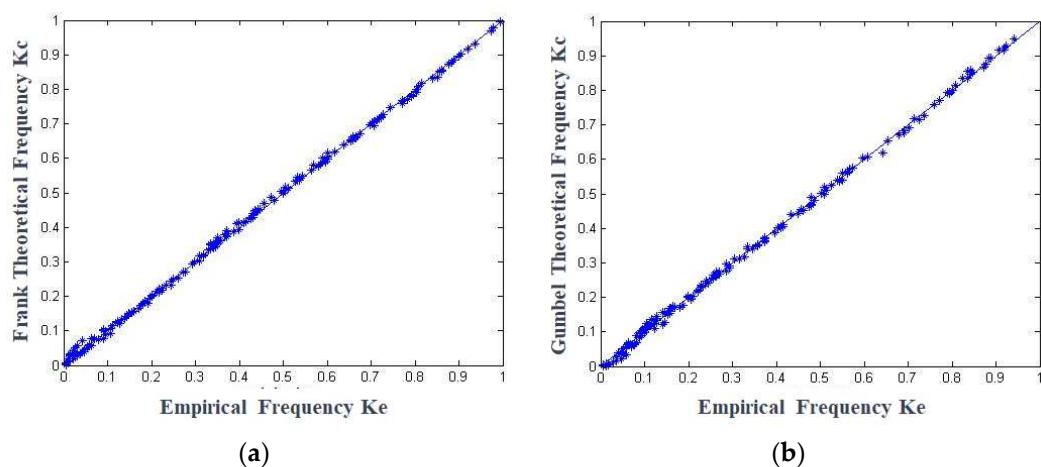
**Table 6.** Evaluation results of two-dimensional symmetrical Archimedean Copula parameters and goodness of fit.

Name of Reservoir	Characteristics	Copula	$\theta$	H	P (%)	OLS	AIC
Foziling	Flood peak and coefficient of variation	Frank	6.7103	0	99.93	0.0117	-1818.10
		Gumbel	2.0277	0	99.48	0.031	-1635.50
		Clayton	1.8761	0	89.56	0.0822	-1452.10
	Flood volume and peak shape coefficient	Frank	-2.4257	0	94.75	0.0178	-1739.50
		Frank	4.7534	0	65.79	0.0407	-1584.30
	Flood duration and peak timing	Gumbel	1.6766	0	97.98	0.0212	-1706.50
Xianghongdian	Flood peak and coefficient of variation	Clayton	1.1528	0	74.50	0.1244	-1374.40
		Frank	4.0522	0	77.93	0.0372	-1440.10
		Gumbel	1.5456	0	99.06	0.0142	-1604.40
	Flood volume and angle of flood rising point	Clayton	1.0926	0	42.45	0.0795	-1310.20
		Frank	1.3027	0	92.44	0.0467	-1401.10
		Gumbel	1.1992	0	92.44	0.0354	-1448.60
	Flood duration and peak timing	Clayton	0.3216	0	50.77	0.0687	-1335.10
		Frank	1.4608	0	98.99	0.0181	-1563.70
		Gumbel	1.1637	0	99.85	0.0144	-1602.50
		Clayton	0.4548	0	98.06	0.0248	-1509.70

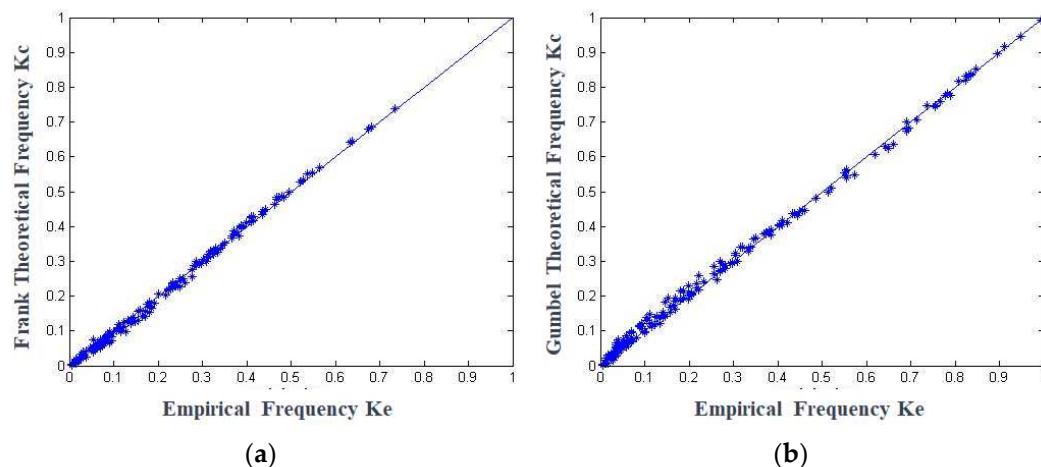
According to Table 5, the non-symmetric Archimedean M6 Copula provides the best fit for the joint distribution of flood peak, flood volume, and flood duration for both the Foziling and Xianghongdian reservoirs. As shown in Table 6, the Frank Copula provides the best fit for the joint distribution of flood peak and coefficient of variation, as well as the joint distribution of flood volume and peak shape coefficient for the Foziling reservoir. The Gumbel Copula provides the best fit for the joint distribution of flood duration and peak timing, as well as the joint distribution of flood peak and coefficient of variation, flood volume and angle of flood rising point, and flood duration and peak timing for the Xianghongdian reservoir. Therefore, the M6 Copula function is selected to fit the three-dimensional intensity characteristics for both reservoirs, while the Frank Copula and Gumbel Copula are selected to fit the two-dimensional intensity and shape characteristics for both reservoirs. The joint distribution Ke–Kc plots for each group of variables are shown in Figures 6–9. All data points in the figures are clustered around the 45-degree diagonal line, indicating a good fit of the joint distribution functions.



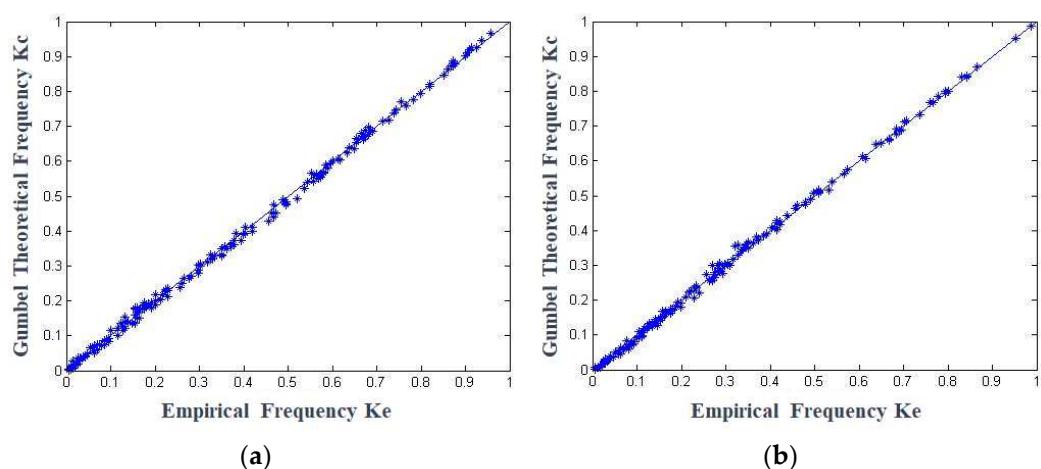
**Figure 6.** Ke–Kc diagram of three-dimensional joint distribution of flood peak, flood volume and flood duration: (a) Foziling Reservoir; (b) Xianghongdian Reservoir.



**Figure 7.** Ke–Kc diagram of two-dimensional joint distribution of peak discharge and coefficient of variation: (a) Foziling Reservoir; (b) Xianghongdian Reservoir.

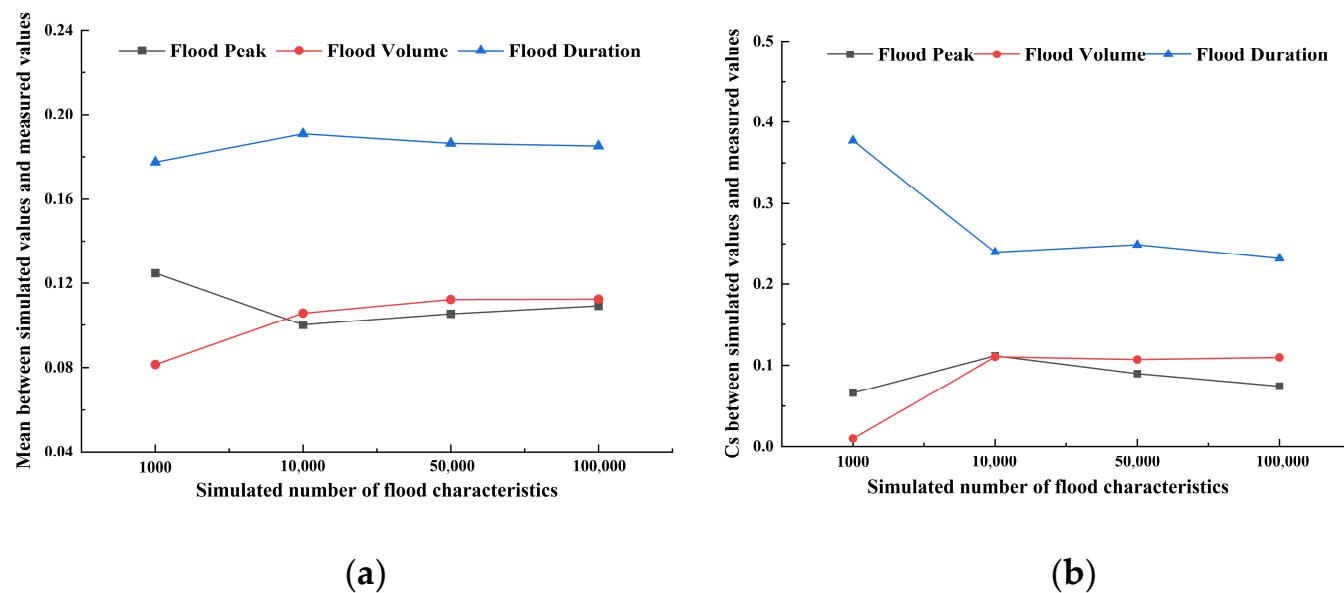


**Figure 8.** Ke-Kc diagram of two-dimensional joint distribution of site flood volume and peak type coefficient, site flood volume, and rising tide elevation: (a) Foziling Reservoir; (b) Xianghongdian Reservoir.



**Figure 9.** Two-dimensional joint distribution of flood duration and peak time, flood duration and peak time Ke-Kc diagram: (a) Foziling Reservoir; (b) Xianghongdian Reservoir.

Based on the joint distribution functions of the three-dimensional and two-dimensional variables, and in combination with the simulation method for flood characteristics described in Section 2.1.4, 10,000 sets of flood characteristics were simulated (by comparing the mean,  $C_V$ , and  $C_S$  errors of the simulated characteristics for 0.1 million, 1 million, 5 million, and 10 million simulations, it was found that the errors tend to stabilize at 10,000 simulations, as shown in Figure 10). Each set includes seven characteristics: flood peak, flood volume, flood duration, peak shape coefficient, peak timing, angle of flood rising point, and coefficient of variation. The statistical parameters of each simulated characteristic are shown in Table 7. Comparing them with the observed characteristics, it can be seen that the main statistical parameters are very close, passing the applicability test, and can be used for subsequent flood process magnification.



**Figure 10.** Comparison of relative error between simulated value and observed value of different number of flood characteristic quantities: (a) mean between simulated and observed value; (b)  $C_S$  between simulated and observed value.

**Table 7.** Statistical parameters of observed and simulated flood characteristics.

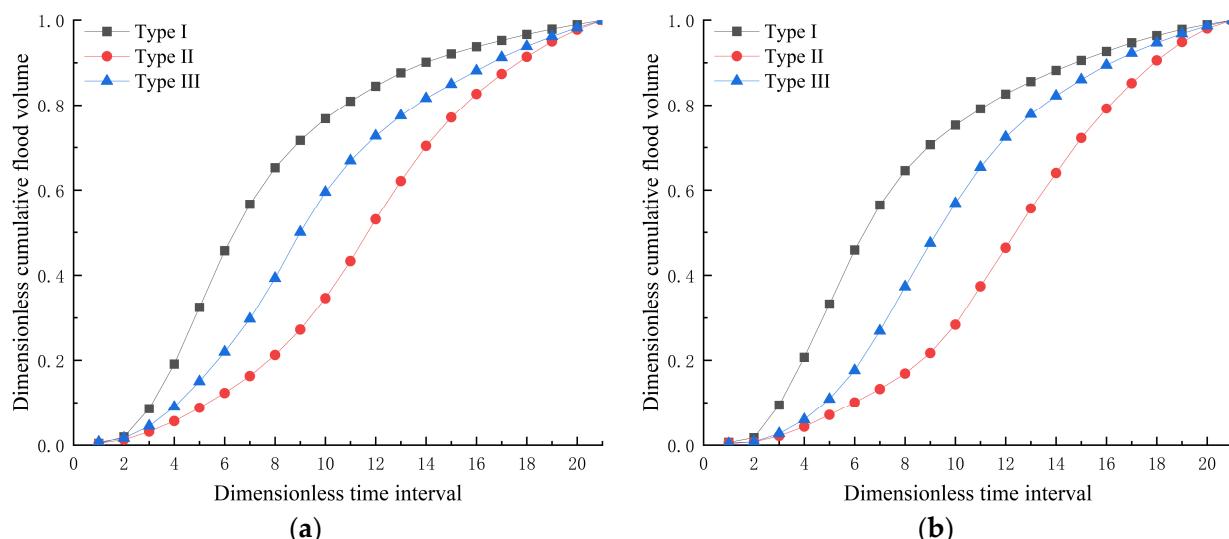
Characteristics	Flood Statistic	Foziling		Xianghongdian	
		Observed Flood	Simulated Flood	Observed Flood	Simulated Flood
Flood peak	Mean value	1410	1551	1508	1569
	$C_V$	1.08	0.99	0.97	0.92
	$C_S$	2.95	2.62	2.75	2.6
Flood volume	Mean value	1.15	1.27	1.06	1.10
	$C_V$	1.04	0.96	1.02	0.96
	$C_S$	2.59	2.30	3.06	2.97
Flood duration	Mean value	91.1	90.84	71.68	73.85
	$C_V$	0.5	0.51	0.48	0.49
	$C_S$	1.405	1.07	1.2	1.08
Peak shape coefficient	Mean value	0.34	0.40	/	
	$C_V$	0.48	0.42		
	$C_S$	0.28	0.28		
Peak timing	Mean value	31.05	19.05	27.33	23.91
	$C_V$	0.75	0.71	0.64	0.68
	$C_S$	1.56	2.65	1.69	2.66

**Table 7.** Cont.

Characteristics	Flood Statistic	Foziling		Xianghongdian	
		Observed Flood	Simulated Flood	Observed Flood	Simulated Flood
Angle of flood rising point	Mean value			0.43	0.37
	$C_V$	/		0.65	0.63
	$C_S$			2.18	3.52
Coefficient of variation	Mean value	0.89	0.69	0.89	0.77
	$C_V$	0.42	0.42	0.33	0.31
	$C_S$	1.03	1.16	0.71	0.93

### 3.4. Flood Hydrograph Classification

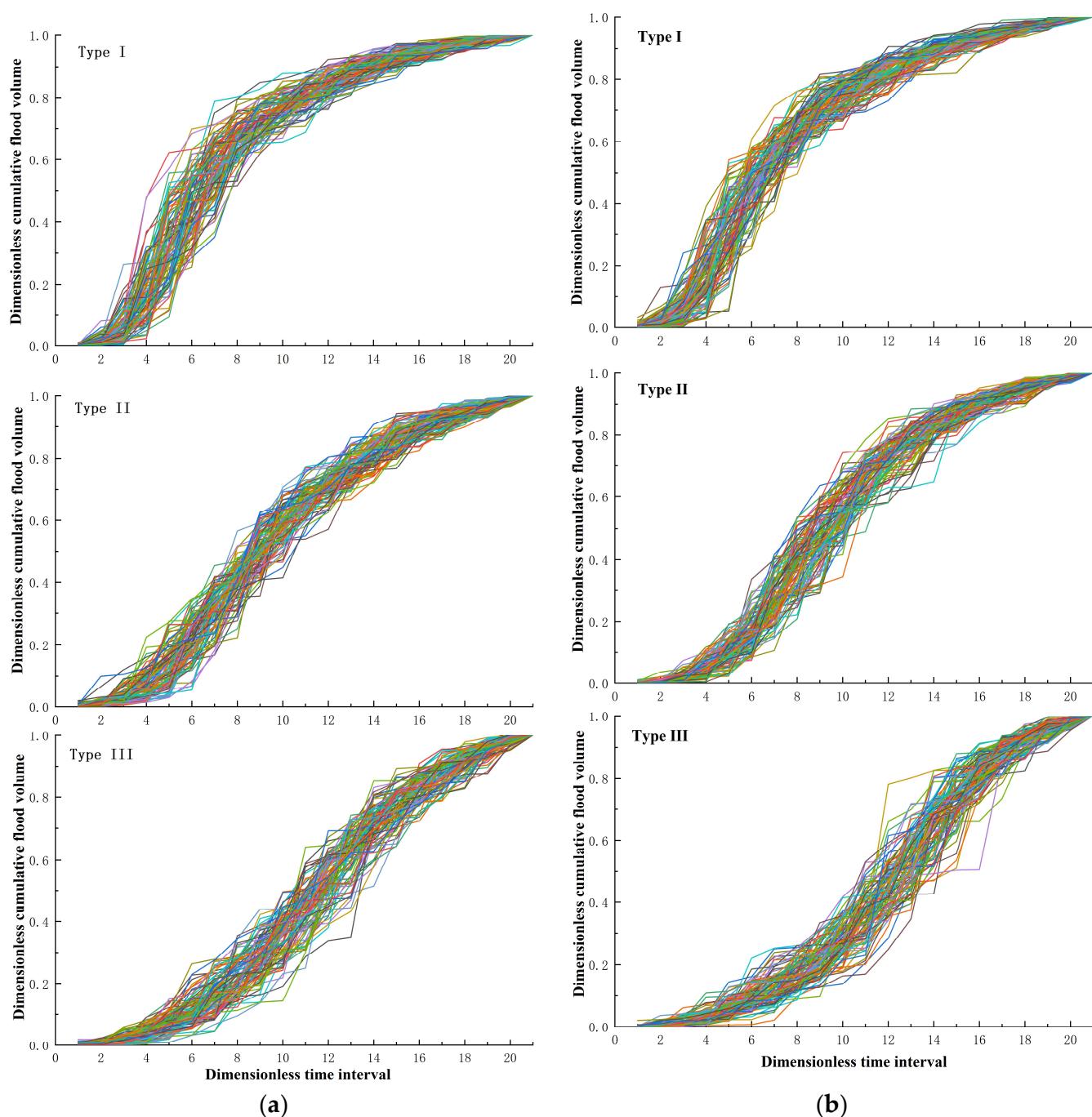
The 185-flood hydrograph from the Foziling Reservoir and the 171-flood hydrograph from the Xianghongdian Reservoir have been normalized and divided into 21 segments based on the characteristics of the watershed flood periods. The K-means clustering method was then applied to classify the flood process lines. Eventually, both reservoirs' flood hydrographs were divided into three classes, as shown in Figure 11. Class I represents floods where the peak occurs in the first half of the flood process line, Class II represents floods where the peak occurs in the middle of the process line, and Class III represents floods where the peak occurs in the latter half of the hydrograph. Based on the classification results, it is observed that the peaks of floods in both the Foziling Reservoir and the Xianghongdian Reservoir generally appear in the middle of the flood process.



**Figure 11.** Classification of dimensionless flood accumulation curve: (a) Foziling Reservoir; (b) Xianghongdian Reservoir.

### 3.5. Stochastic Simulation of Different Types of Flood Hydrograph

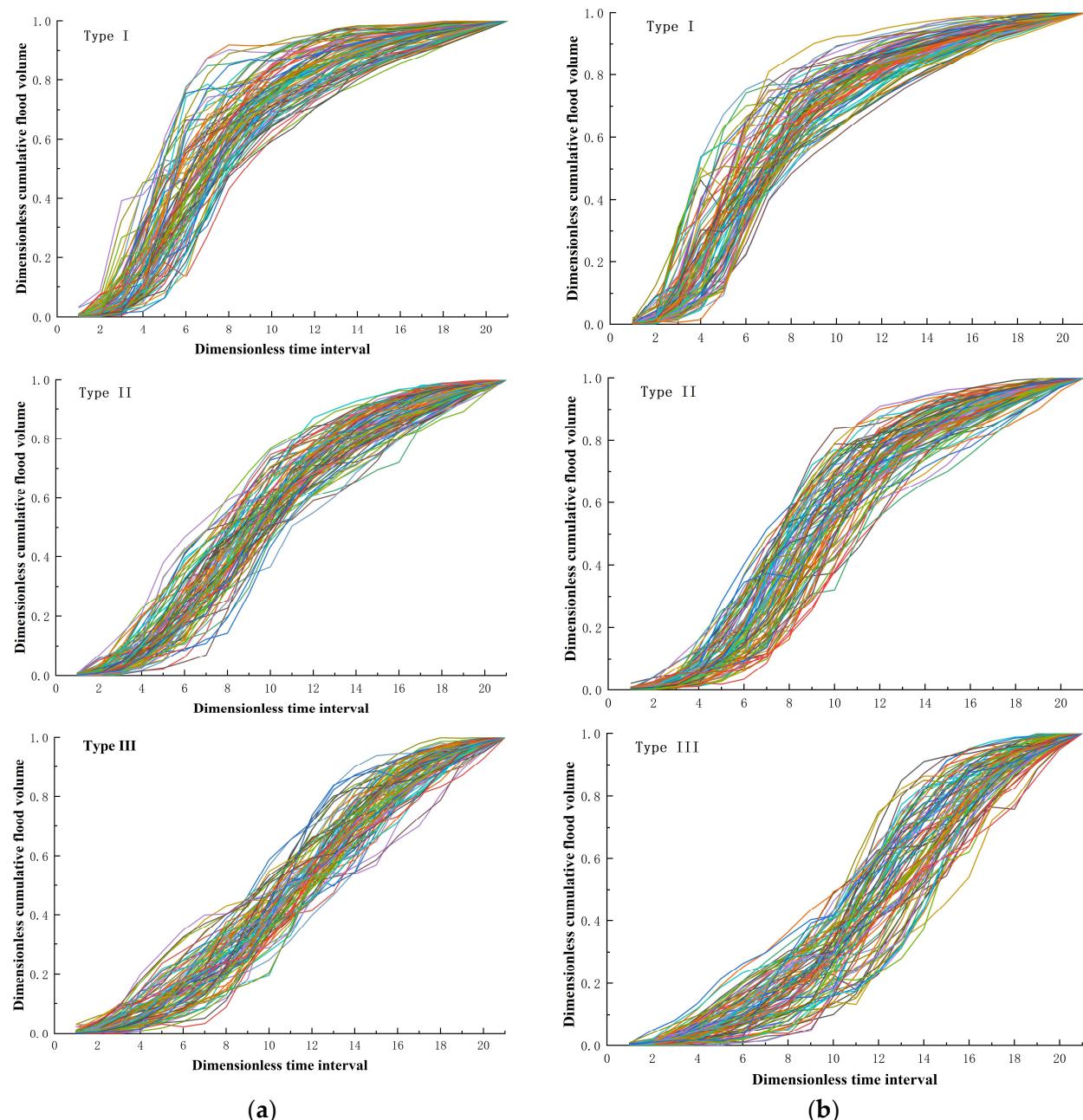
The joint distribution models of flood increments for each time period are constructed using multivariate Gaussian and multivariate t Copula functions, and stochastic simulations are performed. For illustration purposes, we present the results of 300 simulated dimensionless flood hydrograph, considering 100 each of Class I, II, and III flood types for both the Foziling Reservoir and Xianghongdian Reservoir. Figures 12a and 13a show the simulated flood hydrograph for the Foziling Reservoir, while Figures 12b and 13b display those for the Xianghongdian Reservoir. Additionally, for comparison, we use Monte Carlo simulation to randomly generate unconstrained independent normal multivariate inversely derived flood increments, as shown in Figure 11.



**Figure 12.** Monte Carlo method for simulating flood hydrographs of Class I, II, and III: (a) Foziling Reservoir; (b) Xianghongdian Reservoir.

Statistical comparisons are made for the means of the three flood simulation methods and the normalized observed mean values at each cross-section. The results of the comparison are presented in Table 8. From Figures 12–14 and Table 8, it can be observed that the simulation of the three flood process line types (Class I, II, and III) for both the Foziling Reservoir and Xianghongdian Reservoir are highly accurate. The relative errors of the means, except for Class I flood at the Foziling Reservoir, are all within 20%. This indicates that the simulated flood hydrograph maintains a similar distribution to the observed flood data at each cross-section, demonstrating the effectiveness of the random simulation of flood process lines. Furthermore, the multivariate Gaussian Copula simulation performs the best among the three methods, with relative errors within 5% for all flood types, except

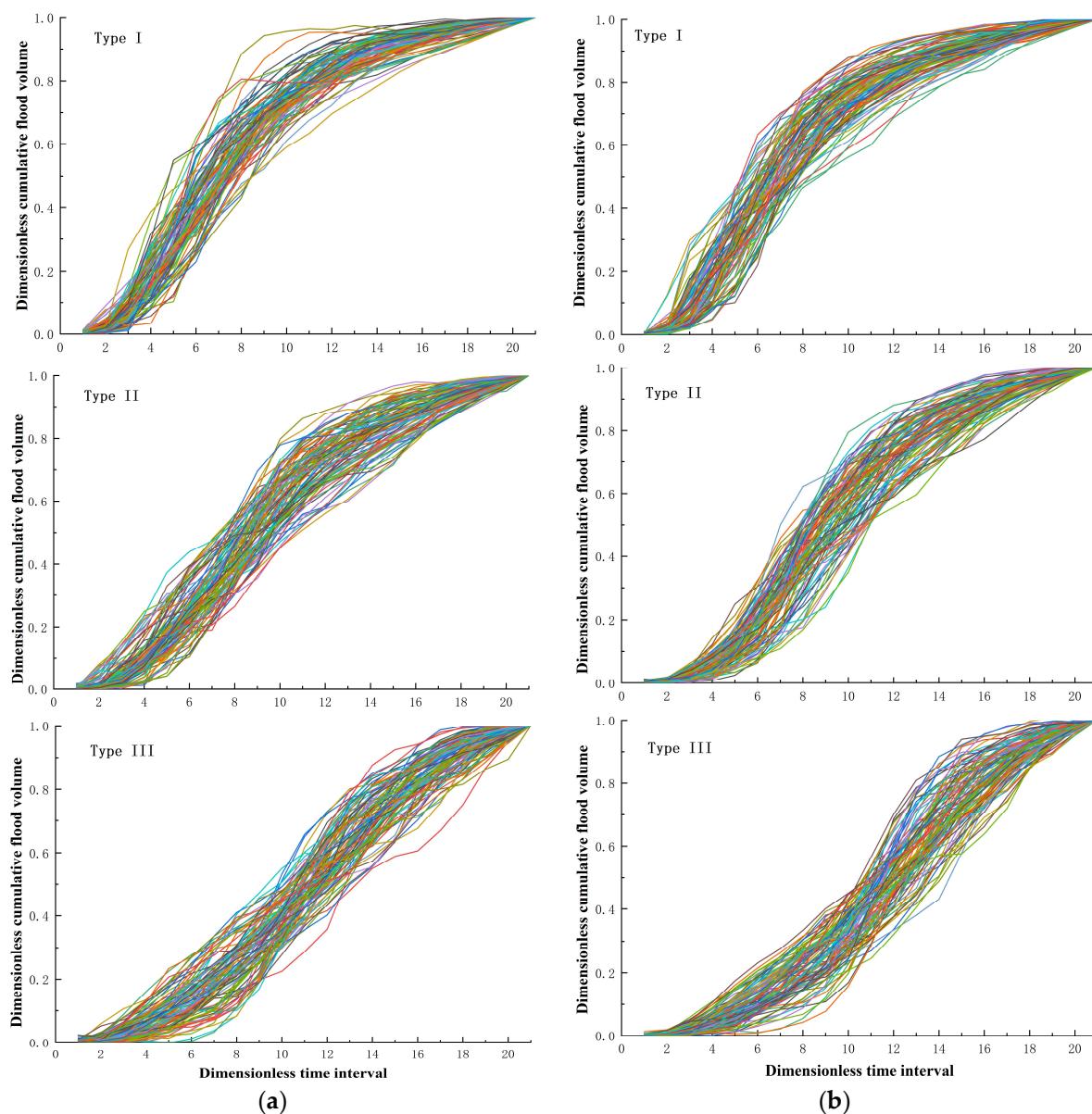
for Class III flood at the Xianghongdian Reservoir, which has a relative error of 6.87%. This is superior to the Monte Carlo random transformation and multivariate t Copula methods. Therefore, the flood hydrograph simulated using the multivariate Gaussian Copula method is chosen for fusion with flood characteristics.



**Figure 13.** Multivariate t-Copula simulation flood hydrograph of Class I, II, and III: (a) Foziling Reservoir; (b) Xianghongdian Reservoir.

**Table 8.** Relative errors between measured value and simulated mean value of each section.

Type of Flood Hydrograph	Foziling			Xianghongdian		
	Logarithmic Transformation	t	Gaussian	Logarithmic Transformation	t	Gaussian
I	46.70%	2.02%	2.41%	4.98%	1.57%	1.48%
II	7.99%	3.37%	1.87%	5.10%	2.25%	1.88%
III	10.03%	3.73%	3.17%	8.04%	7.11%	6.87%



**Figure 14.** Multivariate Gaussian\_Copula simulates flood hydrograph of Class I, II, and III: (a) Foziling Reservoir; (b) Xianghongdian Reservoir.

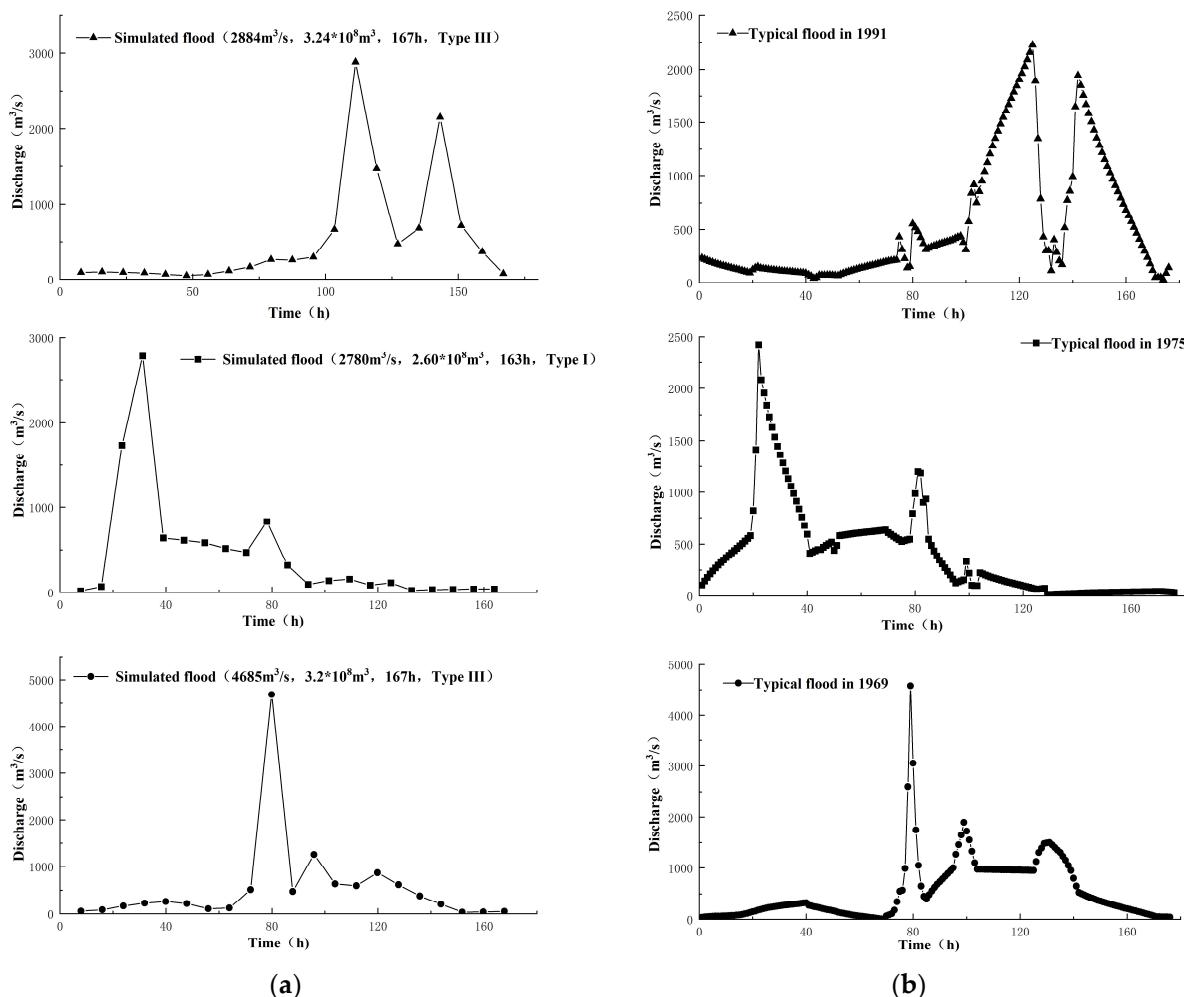
### 3.6. Fusion of Flood Characteristics with Different Types of Flood Hydrographs

By applying the flood hydrograph identification method, we analyzed the goodness of fit between the 10,000 sets of flood characteristics generated in Section 2.3 and the three representative types of dimensionless flood hydrograph obtained in Section 2.2.2. This analysis allows us to determine the flood hydrograph type corresponding to each set of flood characteristics. After the calculation, for the Foziling Reservoir, out of the 10,000 sets of flood characteristics, 2632 sets showed the best fit with Class I flood process, 3554 sets with Class II flood process, and the remaining 3814 sets with Class III flood process. For the Xianghongdian Reservoir, out of the 10,000 sets of flood characteristics, 2033 sets showed the best fit with Class I flood process, 4935 sets with Class II flood process, and the remaining 3032 sets with Class III flood process. The comparison between the occurrence frequencies of different types of simulated floods and observed floods is presented in Table 9. The results demonstrate that the frequencies of different types of flood processes obtained from the proposed random simulation method closely align with the frequencies observed in actual flood events, indicating the reliability of the simulation results.

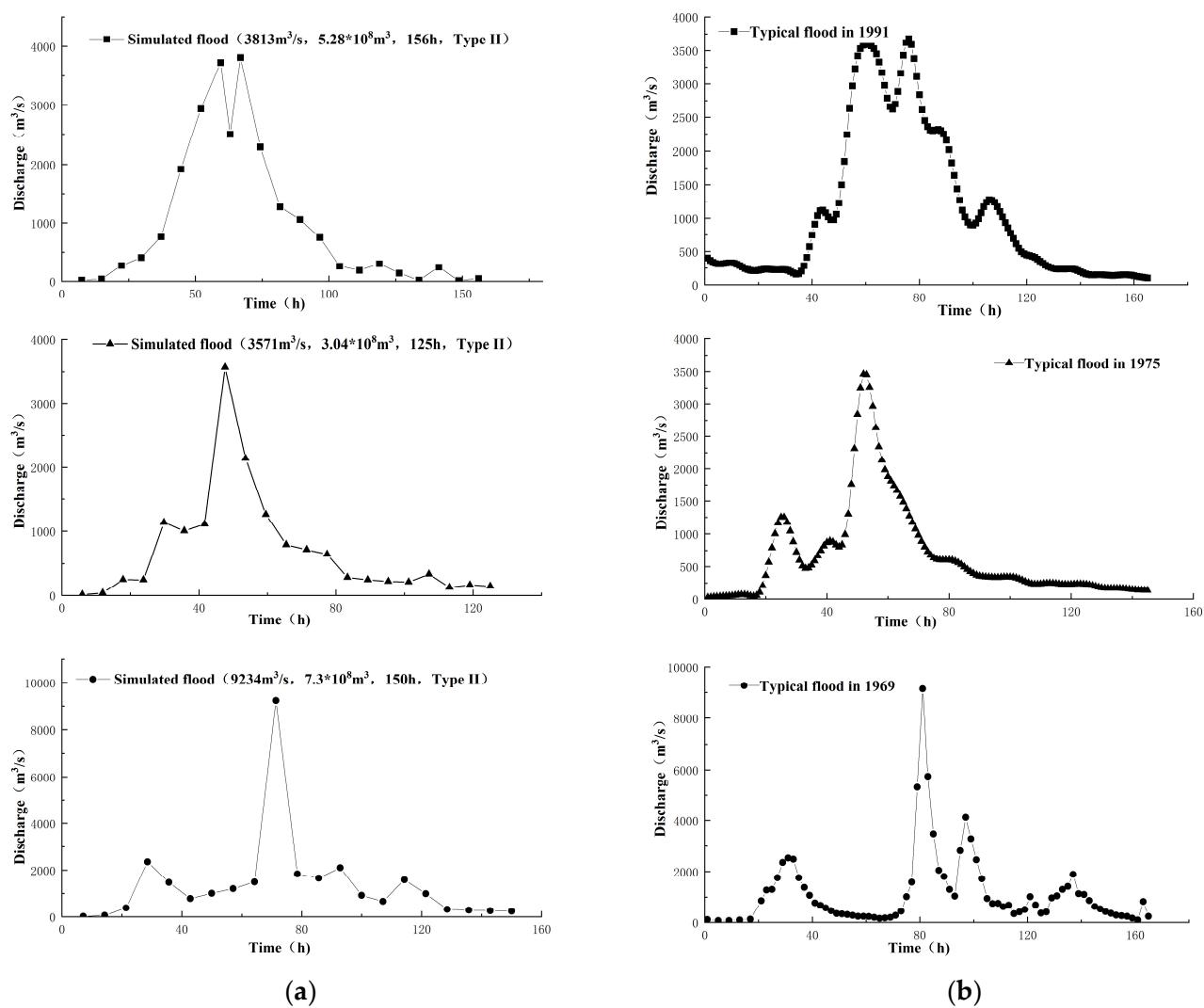
**Table 9.** Comparison between simulated frequency and measured frequency of different types of flood hydrographs.

Name of Reservoir	Type of Flood Hydrograph	Observed Value		Simulated Value	
		Events	Frequency	Events	Frequency
Foziling	I	55	29.10%	2632	26.32%
	II	67	35.45%	3554	35.54%
	III	67	35.45%	3814	38.14%
Xianghongdian	I	36	21.05%	2033	20.33%
	II	86	50.29%	4935	49.35%
	III	49	28.65%	3032	30.32%

In the stochastic simulation of 10,000 flood events, multiple flood events were obtained that bear similarities to the “1991”, “1975”, and “1969” typical floods. The comparison between the simulated floods and the typical flood events is illustrated in Figures 15 and 16, and the characteristics are summarized in Table 10. Both Foziling and Xianghongdian Reservoirs’ simulated flood events show a close resemblance in terms of flood intensity characteristics to the typical floods. Additionally, the type of flood process remains consistent, indicating that the stochastic simulation of floods, considering both intensity and morphology indicators, is capable of capturing historically typical flood events, demonstrating the representativeness and reliability of the simulation results.



**Figure 15.** Stochastic simulation diagram of typical floods of Foziling Reservoir in 1991, 1975, and 1969: (a) simulated flood; (b) typical flood.



**Figure 16.** Stochastic simulation diagram of typical floods of Xianghongdian Reservoir in 1991, 1975, and 1969: (a) simulated flood; (b) typical flood.

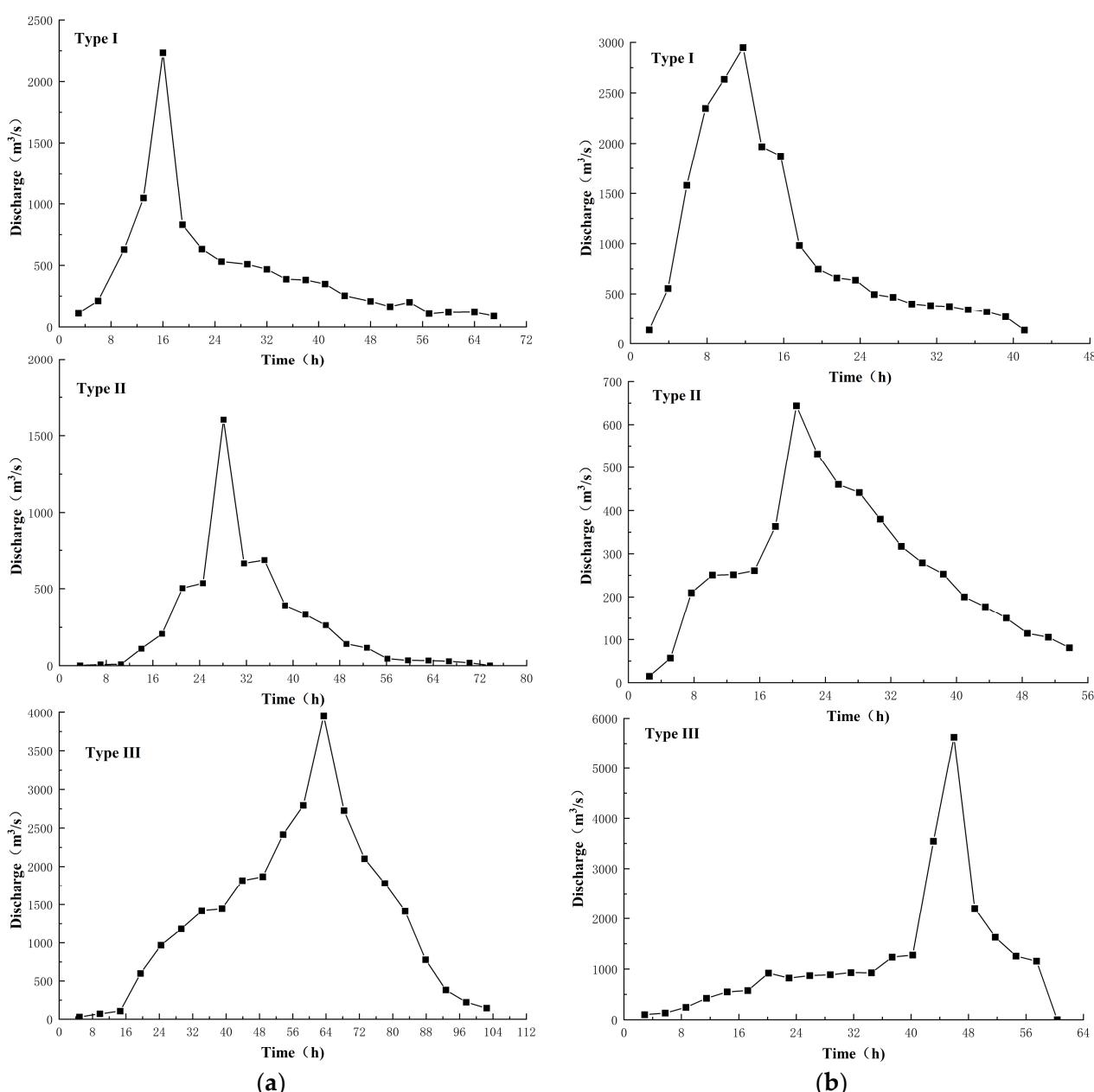
**Table 10.** Comparison of typical flood simulation and measured characteristic quantities of Foziling and Xianghongdian reservoirs.

Name of Reservoir	Characteristics	Typical Floods			Simulated Floods		
		Year 1991	Year 1975	Year 1969	Year 1991	Year 1975	Year 1969
Foziling	Flood peak ( $\text{m}^3/\text{s}$ )	2223	2424	4580	2884	2780	4685
	Flood volume ( $10^8 \text{ m}^3$ )	3.39	2.57	3.47	3.24	2.6	3.2
	Flood duration (h)	176	176	176	167	164	167
	Type of flood hydrograph	III	I	III	III	I	III
Xianghongdian	Flood peak ( $\text{m}^3/\text{s}$ )	3680	3460	9148	3818	3571	9234
	Flood volume ( $10^8 \text{ m}^3$ )	6.43	3.48	6.55	5.28	3.04	7.3
	Flood duration (h)	165	145	161	156	125	150
	Type of flood hydrograph	II	II	II	II	II	II

Figure 17a shows the flood hydrograph of a single event for each of the three classes (I, II, III) at Foziling Reservoir. For Class I, the flood peak, flood volume, and flood duration are  $2234 \text{ m}^3/\text{s}$ , 120 million  $\text{m}^3$ , and 66 h, respectively. For Class II, the flood peak, flood volume, and flood duration are  $1648 \text{ m}^3/\text{s}$ , 60 million  $\text{m}^3$ , and 73 h, respectively. For Class III, the flood peak, flood volume, and flood duration are  $3951 \text{ m}^3/\text{s}$ , 479 million  $\text{m}^3$ , and

102 h, respectively. Figure 17b shows the flood hydrograph of a single event for each of the three classes (I, II, III) at Xianghongdian Reservoir. For Class I, the flood peak, flood volume, and flood duration are  $2948 \text{ m}^3/\text{s}$ , 143 million  $\text{m}^3$ , and 41 h, respectively. For Class II, the flood peak, flood volume, and flood duration are  $643 \text{ m}^3/\text{s}$ , 51 million  $\text{m}^3$ , and 53 h, respectively. For Class III, the flood peak, flood volume, and flood duration are  $5626 \text{ m}^3/\text{s}$ , 262 million  $\text{m}^3$ , and 60 h, respectively. These classes represent different scenarios: peak with a large volume and short duration, peak with a small volume and short duration, and peak with a large volume and long duration, for both Foziling and Xianghongdian Reservoirs.

After these steps, a flood hydrograph of any type under the joint distribution of flood peak, flood volume, and flood duration can be randomly simulated, considering different inflow possibilities. This provides a data foundation for flood control scheduling and risk assessment.



**Figure 17.** Example of flood hydrograph simulation results of Class I, II, and III: (a) Foziling Reservoir; (b) Xianghongdian Reservoir.

#### 4. Conclusions and Outlook

This paper is based on the Copula function to simulate flood characteristics and flood hydrograph. We specifically focused on the randomness of flood hydrograph and the correlation between their morphological features and intensity characteristics. This approach provides a new perspective for flood stochastic simulation, resulting in flood hydrographs that better match real-world scenarios. It offers crucial insights for flood control scheduling, risk assessment decisions, and serves as a valuable foundation for decision-making in these areas.

- (1) When establishing the stochastic simulation model for flood characteristic variables, significant consideration was given to the asymmetric correlation among high-dimensional flood characteristic variables. A non-symmetric Archimedean Copula was employed to construct the joint distribution. Compared to traditional symmetric methods, the simulated flood characteristic variables using this approach more closely resemble natural flood conditions.
- (2) Taking the inflow flood data of Fuziling and Xianghongdian Reservoirs as examples, the dimensionless flood process lines were clustered and analyzed. For different types of flood hydrograph, three methods, namely, multivariate Gaussian Copula, multivariate t Copula, and Monte Carlo simulation, were used to stochastically simulate the related cumulative flood volumes for each time interval. These methods enhanced the diversity and randomness of the hydrograph. A comparative analysis of the relative errors between the three simulation methods and the measured data showed that the multivariate Gaussian Copula method provided process lines that closely approximated the observed ones.
- (3) Emphasis was placed on the influence of flood intensity characteristics on the shape of hydrograph. Two-dimensional joint distributions between flood peak, flood volume, flood duration, and flood shape characteristics were established to achieve an organic fusion between flood hydrograph and characteristic variables. The results of practical calculations demonstrated that the simulated flood data closely matched the statistical characteristics and type proportions of the measured flood data, indicating the applicability and reliability of this method in flood random simulation.
- (4) Using Copula functions to randomly simulate multivariate flood characteristics and flood hydrographs requires a substantial amount of observed flood data for estimating model parameters. Insufficient flood data length and precision may impact the accuracy of the model. While three typical flood hydrographs obtained through clustering methods can, to some extent, enrich the diversity of flood hydrographs, they still do not fully represent the characteristics of rare flood hydrographs.
- (5) This paper generalizes the flood hydrographs of reservoirs into 21 intervals, with the option to increase the number of segments when the flood duration in the basin is longer. During the fusion of flood characteristic variables and flood hydrograph, the model employed flood regulation calculations to back-calculate and deduce the flood process, resulting in a sawtooth-shaped pattern. To address this issue, this study appropriately smoothed the flood hydrograph, while keeping the flood peak, flood volume, and flood duration unchanged. However, further improvements are required to enhance the fusion method.

**Author Contributions:** Conceptualization, X.F. and X.H.; methodology, X.F.; software, X.F.; validation, X.H. and L.D.; formal analysis, X.F.; investigation, X.H.; resources, X.H.; data curation, L.D.; writing—original draft preparation, X.F.; writing—review and editing, X.F. and X.H.; visualization, X.H.; supervision, L.D.; project administration, X.H.; funding acquisition, X.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The authors have no permission to submit the data, since it belongs to the Ministry.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Yang, H.P. Current Status and Development Direction of Hydrological and Water Resources Random Simulation. *Water Sci. Eng. Technol.* **2009**, *3*, 6–8. (In Chinese)
2. Zhou, H.C.; Dong, S.H.; Deng, C.L. Risk Analysis of Flood Control Scheduling Based on Stochastic Hydrological Processes. *J. Hydraul. Eng.* **2006**, *2*, 227–232. (In Chinese)
3. Xu, C.; Yin, J.; Guo, S. Deriving design flood hydrograph based on conditional distribution: A case study of Danjiangkou reservoir in Hanjiang basin. *Math. Probl. Eng.* **2016**, *2016*, 4319646. [[CrossRef](#)]
4. Nelsen, R.B. *An Introduction to Copulas*; Springer: New York, NY, USA, 2006.
5. Xie, H.; Huang, J.S. A review of bivariate hydrological frequency distribution. *Adv. Water Sci.* **2008**, *19*, 443–452. (In Chinese)
6. Salvadori, G.; De Michele, C. Estimating strategies for multiparameter multivariate extreme value copulas. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 141–150. [[CrossRef](#)]
7. Rauf, U.F.A. A Copula-Based Analysis of Flood Phenomena in Victoria, Australia. Ph.D. Thesis, RMIT University, Melbourne, Australia, 2014.
8. Requena, A.I.; Flores, I.; Mediero, L. Extension of observed flood series by combining a distributed hydro-meteorological model and a copula-based model. *Stoch. Environ. Res. Risk Assess.* **2016**, *30*, 1363–1378. [[CrossRef](#)]
9. Djibo, A.G.; Seidou, O.; Saley, H.M.; Philippon, N.; Sittichok, K.; Karambiri, H.; Paturel, J.E. A copula-based approach for assessing flood protection overtopping associated with a seasonal flood forecast in Niamey, West Africa. *J. Geogr. Environ. Earth Sci. Int.* **2018**, *15*, 1–22. [[CrossRef](#)]
10. Xiao, Y.; Guo, S.L.; Xiong, L.H.; Liu, P.; Fang, B. Study on a New Method for Stochastic Simulation of Flood Processes. *J. Sichuan Univ.* **2007**, *2*, 55–60. (In Chinese)
11. Gao, C.; Zhu, C.; Pan, S.L.; Liu, L.; Xu, Y.P. Stochastic Simulation of Different Types of Flood Hydrographs. *J. Basic Sci. Eng.* **2018**, *26*, 767–779. (In Chinese)
12. Zhang, L.; Singh, V.P. Trivariate flood frequency analysis using the Gumbel-Hougaard Copula. *J. Hydrol. Eng.* **2007**, *12*, 431–439. [[CrossRef](#)]
13. Grimaldi, S.; Seringesco, F. Asymmetric Copula in multivariate flood frequency analysis. *Adv. Water Resour.* **2006**, *29*, 150–164. [[CrossRef](#)]
14. Ganguli, P.; Reddy, J.M. Probabilistic assessment of flood risks using trivariate copulas. *Theor. Appl. Climatol.* **2013**, *111*, 341–360. [[CrossRef](#)]
15. Li, T.Y.; Guo, S.L.; Yan, B.W.; Chen, L. A Novel Method for Deriving Design Flood Hydrographs Based on Multivariate Joint Distribution. *J. Hydroelectr. Eng.* **2013**, *32*, 10–14+38.
16. Guo, S.L.; Liu, Z.J.; Xiong, L.H. Research Progress and Evaluation of Design Flood Calculation Methods. *J. Hydraul. Eng.* **2016**, *47*, 302–314. (In Chinese)
17. Yan, X.R.; Wang, L.P.; Zhang, Y.K.; Yu, H.J.; Li, N.N. Research on Flood Random Simulation Method Considering Peak Shape and Frequency. *J. Hydroelectr. Eng.* **2019**, *38*, 61–72. (In Chinese)
18. Al-Amri, N.S.; Ewea, H.A.; Elfeki, A.M. Stochastic Rational Method for Estimation of Flood Peak Uncertainty in Arid Basins: Comparison between Monte Carlo and First Order Second Moment Methods with a Case Study in Southwest Saudi Arabia. *Sustainability* **2023**, *15*, 4719. [[CrossRef](#)]
19. Zhai, X.Y.; Guo, L.; Zhang, Y.Y. Identification and Simulation of Mountainous Flood Types in China Based on Flood Behavioral Characteristic Index. *Sci. China: Earth Sci.* **2021**, *51*, 1092–1106. (In Chinese)
20. Poff, N.L.R.; Allan, J.D.; Bain, M.B.; Karr, J.R.; Prestegaard, K.L.; Richter, B.D.; Sparks, R.E.; Stromberg, J.C. The natural flow regime: A paradigm for river conservation and restoration. *Bioscience* **1997**, *47*, 769–784. [[CrossRef](#)]
21. Zhang, Y.Y.; Zhai, X.Y.; Shao, Q.X. Assessing temporal and spatial alterations of flow regimes in the regulated Huai River Basin, China. *J. Hydrol.* **2015**, *529*, 384–397. [[CrossRef](#)]
22. Wang, L.N.; Li, Y.; Chen, X.H.; Gu, J.G. Identification of Flood Morphology. *J. Hydroelectr. Eng.* **2013**, *32*, 20–24.
23. Schweizer, B. Introduction to Copulas. *J. Hydrol. Eng.* **2007**, *12*, 346. [[CrossRef](#)]
24. Zili, Z.; Christiana, C.; Peter, F. A Gaussian copula joint model for longitudinal and time-to-event data with random effects. *Comput. Stat. Data Anal.* **2023**, *181*, 107685.
25. Ma, M.; Song, S.; Ren, L.; Jiang, S.; Song, J. Multivariate drought characteristics using trivariate Gaussian and Student t copulas. *Hydrol. Process.* **2013**, *27*, 1175–1190. [[CrossRef](#)]
26. Mohamad, H.K.; Hamid, R.S.; Mohammad, H.G.; Maamoun, A. Comparison between bivariate and trivariate flood frequency analysis using the Archimedean copula functions, a case study of the Karun River in Iran. *Nat. Hazards* **2022**, *112*, 1589–1610.
27. Lei, H.; Yu, J.; Pan, H.; Li, J.; Leghari, S.J.; Shang, C.; Xiao, Z.; Jin, C.; Shi, L. A New Agricultural Drought Disaster Risk Assessment Framework: Coupled a Copula Function to Select Return Periods and the Jensen Model to Calculate Yield Loss. *Sustainability* **2023**, *15*, 3786. [[CrossRef](#)]

28. Zhang, L.; Singh, P.V. Trivariate Flood Frequency Analysis Using Discharge Time Series with Possible Different Lengths: Cuyahoga River Case Study. *J. Hydrol. Eng.* **2014**, *19*, 05014012. [[CrossRef](#)]
29. Song, S.; Nie, R. Joint probability study of multivariate hydrological drought based on asymmetric Archimedean copula. *J. Hydroelectr. Eng.* **2011**, *30*, 20–29. (In Chinese)
30. Genest, C.; Rivest, L.P. Statistical inference procedures for bivariate Archimedean Copulas. *J. Am. Stat. Assoc.* **1993**, *88*, 1034–1043. [[CrossRef](#)]
31. da Rocha Júnior, R.L.; dos Santos Silva, F.D.; Costa, R.L.; Gomes, H.B.; Pinto, D.D.C.; Herdies, D.L. Bivariate Assessment of Drought Return Periods and Frequency in Brazilian Northeast Using Joint Distribution by Copula Method. *Geosciences* **2020**, *10*, 135. [[CrossRef](#)]
32. Sohee, K.; Sunghee, C.; Yeun, J. Statistical Assessment on Student Engagement in Asynchronous Online Learning Using the k-Means Clustering Algorithm. *Sustainability* **2023**, *15*, 2049.
33. Sharda, V.N.; Das, P.K.; Ojasvi, P.R. Probability distribution of extreme rainfall series for conservation planning in a sub-humid climate. *Appl. Eng. Agric.* **2008**, *24*, 447–453. [[CrossRef](#)]
34. Shelton, S.; Pushpawela, B. Seasonal Dependence and Variability of Rainfall Extremes in a Tropical River Basin, South Asia. *Sustainability* **2023**, *15*, 5106. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.