water

MDPI

Article

Combined Exceedance Probability Assessment of Water Quality Indicators Based on Multivariate Joint Probability Distribution in Urban Rivers

Yang Liu, Yufei Cheng, Xi Zhang, Xitong Li and Shengle Cao *

School of Civil Engineering, Shandong University, Jinan 250061, China; fengyangjiao123@163.com (Y.L.); yufei_c2015@163.com (Y.C.); ZXSELZ@163.com (X.Z.); youngmxh@gmail.com (X.L.)

* Correspondence: cao_shengle@163.com; Tel.: +86-0531-88399636

Received: 7 May 2018; Accepted: 9 July 2018; Published: 25 July 2018

check for updates

Abstract: Discharge and water quality are two important attributes of rivers, although the joint response relationship between discharge and multiple water quality indicators is not clear. In this paper, the joint probability distributions are established by copula functions to reveal the statistical characteristics and occurrence probability of different combinations of discharge and multiple water quality indicators. Based on the data of discharge, ammonia nitrogen content index (NH_4^+) and permanganate index (COD_{Mn}) in the Xiaoqing River in Jinan, we first tested the joint change-point with the data from 1980–2016, before we focused on analyzing the data after the change-point and established the multivariate joint probability distributions. The results show that the Gaussian copula is more suitable for describing the joint distribution of discharge and water quality, while the year of 2005 is a joint change-point of water quantity and quality. Furthermore, it is more reasonable to use the trivariate joint probability distribution as compared to the bivariate distributions to reflect the exceedance probability of water quality combination events under different discharge conditions. The research results can provide technical support for the water quality management of urban rivers.

Keywords: Gaussian copula; joint probability distribution; multiple indicators; urban rivers; exceedance probability of water quality

1. Introduction

Water quantity and quality are two important attributes of rivers. In recent years, the water quality of rivers is deteriorating with the development of society and economy, with water dispatching having become one of the main technologies of water quality improvement [1–4]. However, due to the disturbance of urban rivers from human activities, the relationship between water quantity and water quality has become more complicated. The relationship between a single water quality indicator and the water quantity can hardly reflect the real situation of river pollution [5]. Thus, the methods for improvement of the river water quality should be based on a correct understanding of the joint response relationship between water quantity and multiple water quality indicators.

Many scholars have simulated and predicted the river water quality from the aspects of mechanism models and non-mechanism models. The mechanism models are based on the hydrodynamic model and water quality equations in order to simulate the diffusion and attenuation process of pollutants [6–8]. However, due to the wide variety of influencing factors and the cognitive limitations of the real water quality, this makes the model construction and parameter determination difficult [9–11]. Therefore, this indicates that the simulation effect of the nature river is better than that of the urban river, which is affected by human activities. The non-mechanism model can avoid this type of problems to some extent. Sun et al. [12] used the gray model group method and the exponent smoothing method to construct a probabilistic combination prediction framework of water quality to forecast the individual indicator of chemical oxygen demand in a complex environment. After this, some new models for water quality prediction have been proposed, which integrates different non-mechanism models, such as the neural networks model, fuzzy model, and wavelet model [13–15]. In addition, there are also some joint models that combine hydrodynamics and water quality simulation tests [16,17]. However, most non-mechanism models require that the variables obey the same marginal distribution or that these variables need to be transformed to the same distribution, which not only leads to the loss of the original data information, but it also does not conform to the fact that variables, such as discharge and water quality, obey different marginal distributions. Meanwhile, because less emphasis is placed on the joint distribution between the variables, non-mechanism models are more suitable to use a simulation of the discharge and individual water quality indicator when compared to a simulation of the discharge and multiple water quality indicators. Copula can construct the joint probability distribution of multiple variables [18] and it does not require each variable to follow the same marginal distribution. Thus, it can effectively preserve the nature of the original data. Copula has the advantages of great adaptability and flexibility, so it has been widely used in the field of hydrology, especially in determining flood risks [19–22] and drought characteristics [23–28]. However, copula has been used less frequently in the joint analysis of discharge and water quality, with most of the uses focusing on bivariate joint distribution. As such, there is a distinct lack of studies that use copula for a multivariable joint distribution.

Therefore, the purpose of this paper is to use copulas to establish the multivariate joint probability distribution of the discharge and water quality indicators, before analyzing the exceedance probability of multiple water quality indicators under different conditions of discharge.

2. Methods and Materials

2.1. Methodology

2.1.1. Establishment of Copula Function

Copula is a connection function, which is defined in R with marginal distribution functions $F_1, F_2, \cdots, F_n$ for n-dimensional random variables $X_1, X_2, \cdots, X_n$ [29]. If H is the n-dimensional distribution function of n-dimensional random variables, for any $x_i \in R(i = 1, 2, \cdots, n)$, there exists the following copula function:

$$\begin{aligned} H_{x_1,x_2,\ldots,x_d}(x_1, x_2, \ldots, x_d) &= P(X_1 \leq x_1, X_2 \leq x_2, \cdots, X_n \leq x_n) \\ &= C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)) \end{aligned} \tag{1}$$

where n is the number of variables; C is the associated dependence copula function; and, $F_i(x_i)$ is the marginal distribution functions.

Archimedean copulas [30] and Elliptical copulas [31] are commonly used in hydrology. The typical Archimedean copulas include Clayton, Gumbel-Hougaard (G-H), and Frank copula, which require the variables to fall within a symmetric structure. Elliptical copulas are neither limited by the correlation nor restricted by the symmetric structure between variables. However, the multiple hydrological variables may be asymmetric, with independent, positive, or negative relationships occurring among the variables. Thus, it is more reasonable to use the Elliptical copulas to describe the joint distribution of variables when compared to the Archimedean copulas. The Gaussian copula in Elliptical copula is expressed as:

$$\begin{aligned} C(u_1, u_2, \ldots, u_d; \textstyle\sum) &= \Phi_{\sum}(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \cdots, \Phi^{-1}(u_d)) \\ &= \int_{-\infty}^{\Phi^{-1}(u_1)} \cdots \int_{-\infty}^{\Phi^{-1}(u_d)} \frac{1}{(2\pi)^{\frac{d}{2}} \sum^{\frac{1}{2}}} \exp(-\tfrac{1}{2}\mathbf{w}^T \textstyle\sum^{-1} \mathbf{w}) dw \end{aligned} \tag{2}$$

where $\Phi^{-1}(.)$ is the inverse function of the standard normal distribution; $\Phi_{\sum}(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \cdots, \Phi^{-1}(u_d))$ is the multivariate normal distribution; $\sum = \begin{bmatrix} 1 & \cdots & \rho_{1d} \\ \vdots & \ddots & \vdots \\ \rho_{d1} & \cdots & 1 \end{bmatrix}$, $\rho_{ij} = \begin{cases} 1, & i = j \\ \rho_{ji}, & i = j \end{cases}$, $-1 < \rho_{ij} < 1$; d is the dimensions of the random variable; and, $\mathbf{w} = [w_1, w_2 \cdots w_d]^T$.

2.1.2. Test of Joint Change-Point

The purpose of testing the change-point of the multivariable dependence is to make the data more representative in different periods. At the same time, we can utilize the collected data as much as possible to improve the accuracy of the simulation.

Assuming that there is one change-point in the copula function constructed based on the observed data $(x_1, y_1, z_1), \cdots (x_n, y_n, z_n)$, the original hypothesis and the alternative hypothesis can be, respectively, assumed as:

$$H_0 : \theta_1 = \theta_2 = \cdots = \theta_n, H_1 : \theta_1 = \cdots \theta_{n*} \neq \theta_{n^*+1} = \cdots \theta_n$$

If the original hypothesis is rejected, then k^* is the moment of change. If $k^* = k$, the log likelihood ratio statistic of copula can be established based on the maximum likelihood estimation method, which is shown in Equation (4):

$$\begin{aligned} -2\ln(\Lambda_k) &= 2[\sum_{i=1}^{k} C_{uvw}(\theta_k; U(x_i), V(y_i), W(z_i)) + \sum_{i=k+1}^{n} C_{uvw}(\theta_{k^*}; U(x_i), V(y_i), W(z_i)) \\ &+ \sum_{i=1}^{n} C_{uvw}(\theta_n; U(x_i), V(y_i), W(z_i))] \end{aligned} \tag{3}$$

where $\theta_k, \theta_{k^*}, \theta_n$ are the maximum likelihood estimations of the corresponding data; and, $C_{uvw}(U(x), V(y), W(z)) = \frac{\partial}{\partial u}\frac{\partial}{\partial v}\frac{\partial}{\partial w}C(u, v, w)$.

The statistic $Z_n = \max\limits_{1 \le k < n}(-2\ln(\Lambda_k))$ is constructed. If k^* is unknown, the original hypothesis will be rejected when Z_n is greater than the critical value, which essentially demonstrates that the copula function has change-points [32].

2.1.3. Exceedance Probability of Water Quality

In order to quantitatively analyze the response relationship between the discharge and water quality, according to the full probability formula and the copula functions [33,34], we can calculate the probability that the water quality indicators exceed specific values. However, discharge needs to be less than a specific value, which addresses the exceedance probability of water quality. Taking the joint probability distribution of two and three dimensions as an example, the expressions are as follows:

$$\begin{aligned} P_{U_1U_2} &= P(U_1 \le u_1, U_2 \ge u_2) &&= F_{U_1}(u_1) - F_{U_1U_2}(u_1, u_2) \\ & &&= F_{U_1}(u_1) - C(u_1, u_2) \end{aligned} \tag{4}$$

$$\begin{aligned} P_{U_1U_2U_3} &= P(U_1 \le u_1, U_2 \ge u_2, U_3 \ge u_3) \\ &= F_{U_1}(u_1) - F_{U_1U_2}(u_1, u_2) - F_{U_1U_3}(u_1, u_3) + F_{U_1U_2U_3}(u_1, u_2, u_3) \\ &= F_{U_1}(u_1) - C(u_1, u_2) - C(u_1, u_3) + C(u_1, u_2, u_3) \end{aligned} \tag{5}$$

where $P_{U_1U_2}$ is the value of the bivariate joint probability distribution; $P_{U_1U_2U_3}$ is the value of the trivariate joint probability distribution; U_1 is the discharge; U_2 and U_3 are the water quality indicators;

$F_{U_i}(u_i)$ is the marginal distribution function of U_1; and, $C(u_1, \cdots, u_i)$ is the copula function of pair $(U_1, \cdots, U_i)$.

2.2. Study Area and Data

Xiaoqing River, which is located in Jinan, is a manually excavated city river. Furthermore, it is the main flood discharge channel in Jinan. The Huangtaiqiao hydrologic station is located in the main stream of the river, which has monitored the discharge from the river since 1953 with complete water quality monitoring having begun in 1970. The distance between the Muli Gate to Huangtaiqiao hydrologic station is 21.0 km, with a drainage area of 321.0 km^2 and an average annual discharge of 10.4 m^3/s. The river basin is located in the main urban area of Jinan. Due to the rapid development of urbanization, the water pollution in this area is serious. In order to build the ecological civilization, there has been water diversion from the Eastern Route of South-to-North Water Transfer Project to improve the water quality. However, the combined effect of urbanization and water diversion result in the response relationship between the water quantity and quality in Xiaoqing River has been severely disturbed, which has seriously affected its water quality simulation. Therefore, the statistical analysis of the occurrence probability of the combined events of water quantity and water quality can provide a basis for the scientific dispatch. The drainage map of Xiaoqing River in Jinan can be seen in Figure 1.

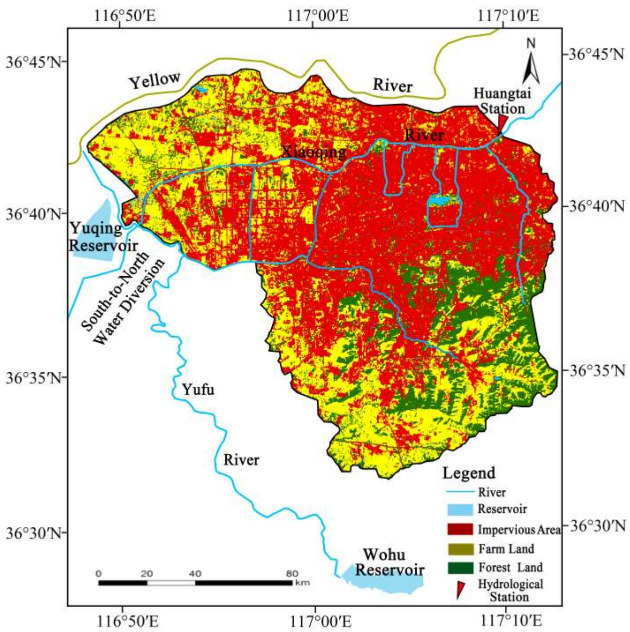

Figure 1. Location of Xiaoqing River in Jinan and drainage map.

This study selected the data of discharge and water quality of Huangtaiqiao station during 1980–2016, with the data of water quality having been recorded every other month and the discharge having been recorded at the same time when the water quality samples were taken (using the main pollutants, nitrogen content index (NH_4^+), which includes the ionised ammonia nitrogen and the free ammonia nitrogen, and permanganate index (COD_{Mn}) as the calculation indicators). To remove some unreasonable data, 161 sets of water quantity and water quality data were applied in this paper. The sample data series of NH_4^+, COD_{Mn}, and discharge of the Huangtai Station from 1980 to 2016 are shown in Figure 2, and the characteristics and trend results of the generalized basic data about the water quality are given in Table 1. The trends of the generalized basic data are tested by Mann-Kendall (M-K). It can be seen that the concentrations of NH_4^+ and COD_{Mn} are basically larger than the grade III value (Environmental Quality Standards for Surface Water (GB3838-2002)). Moreover, the COD_{Mn}

has a significant decreasing trend, and the discharge has a significant increasing trend, while the trend of NH_4^+ descending was not obvious.

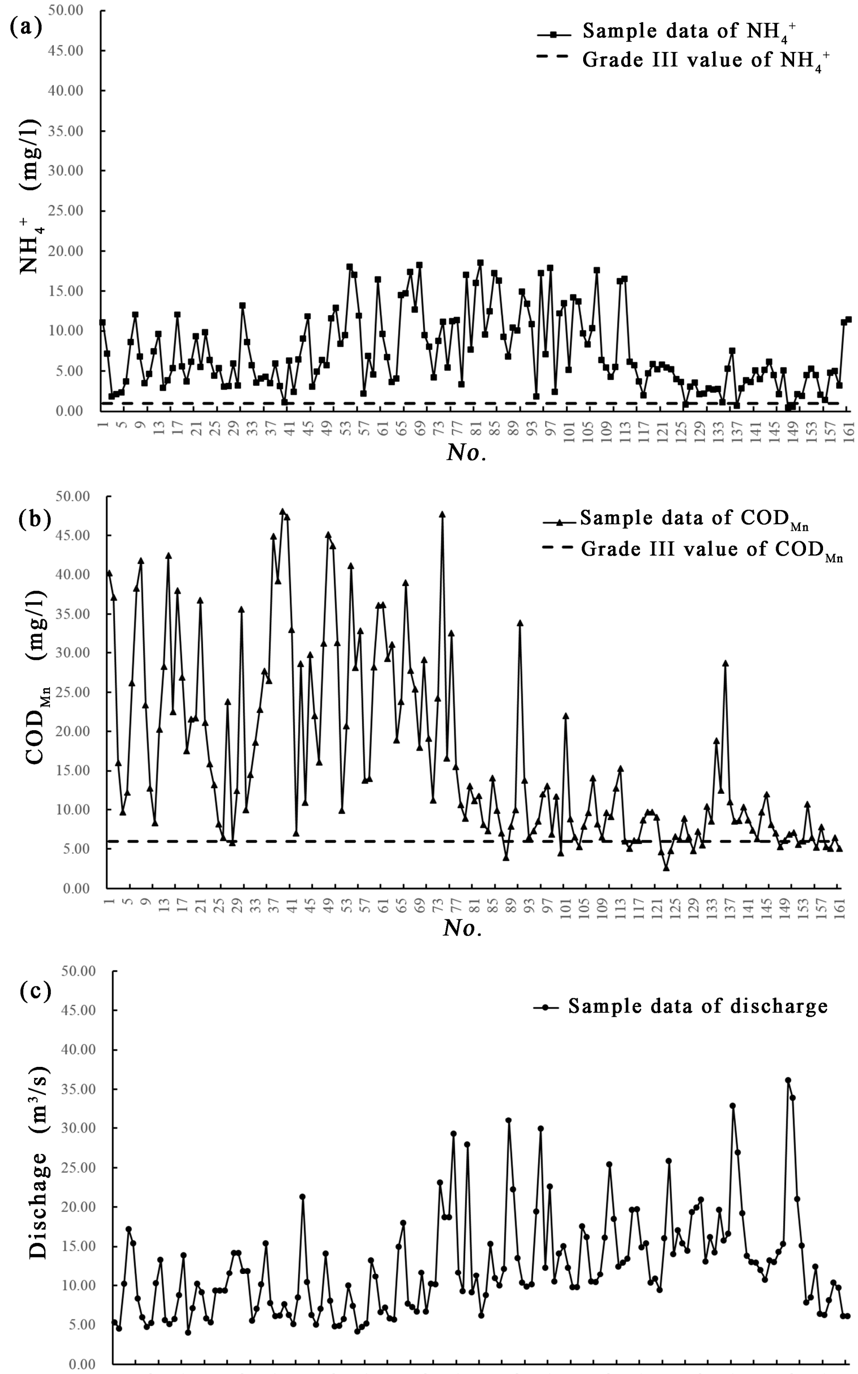

Figure 2. The data series of nitrogen content index (NH_4^+), permanganate index (COD_{Mn}), and discharge of the Huangtai Station from 1980 to 2016: (**a**) the sample data and grade III value of NH_4^+ ; (**b**) the sample data and grade III value of COD_{Mn}; and, (**c**) the sample data of discharge.

Table 1. The trend results of NH_4^+, COD_{Mn}, and discharge.

	NH_4^+ (mg/L)	COD_{Mn} (mg/L)	Discharge (m^3/s)	Water Temperature (°C)	pH
Minimum	0.43	2.60	3.99	3	5.50
Maximum	18.50	48.10	36.10	31	11.60
Mean	7.23	16.86	12.43	18.94	7.51
Trend	↓	$\downarrow^+$	$\uparrow^+$	$\uparrow^-$	↑

Note: ↓ means a decreasing trend; ↑ means a raising trend; $^+$ means significant trend; $^-$ means non-significant trend.

In order to better reflect the relationship between the discharge and the multiple water quality indicators in the Xiaoqing River Basin, which is a rapidly developing area, we carried out a test for the change-point of the long sequence data (1980–2016), before focusing on the analysis of the data after the change-point.

3. Results and Discussion

3.1. Marginal Distributions and Correlation Coefficient of Q, NH_4^+ and COD_{Mn}

The marginal distributions of the Q, NH_4^+, and COD_{Mn} of Xiaoqing River are established by the lognormal, gamma, and Pearson-III distributions to search for the proper cumulative distribution function (CDF). The proper CDF of the parameters was chosen using two goodness fit tests: the Kolmogorov–Smirov (K–S) test and root mean squared error (RMSE) test. The results of the proper marginal distribution of each variable are shown in Table 2. From Table 2, we can see that the marginal distribution of Q and NH_4^+ sequence is more consistent with a Gamma distribution, while COD_{Mn} is more consistent with a lognormal distribution.

Table 2. The parameters of marginal distributions of Q, NH_4^+, and COD_{Mn} and results of Kolmogorov–Smirov (K–S) and root mean squared error (RMSE).

	Lognormal			Gamma			Pearson III		
	(μ, σ)	KS	RMSE	(μ, σ)	KS	RMSE	(μ, σ, α_0)	KS	RMSE
Q	2.398, 0.494	0.057	0.028	4.255, 2.921	**0.051**	**0.021**	0.553, 23.406, −0.508	0.431	0.228
NH_4^+	1.738, 0.751	0.152	0.026	2.238, 3.230	**0.063**	**0.022**	2.367, 3.149, −0.222	0.066	0.028
COD_{Mn}	2.588, 0.690	**0.072**	**0.053**	2.266, 7.440	0.121	0.067	1.470, 10.291, 1.737	0.097	0.055

Note: The number in bold means that RSME is the minimum and the K–S is 0.10172 according to the 5% significance level when n = 161.

The Kendall correlation coefficient τ is used to analyze the correlation of Q, NH_4^+, and COD_{Mn}, respectively, which is shown in Table 3. We can see that the pair of (Q, NH_4^+) and the pair of (Q, COD_{Mn}) are negatively correlated, while the pair (NH_4^+, COD_{Mn}) has a positive correlation. This provides the conditions for the application of Gaussian Copula.

Table 3. The correlation coefficient τ of Q, NH_4^+, and $COD_{Mn.}$

Pair	Q, NH_4^+	Q, COD_{Mn}	NH_4^+, COD_{Mn}
τ	−0.2712	−0.3992	0.1933

3.2. Establishment of Joint Probability Distribution

3.2.1. Establishment of Multivariate Joint Probability Distribution of Q, NH_4^+, and COD_{Mn} with the Data from 1980–2016

The Frank, GH, and Gaussian copulas are used to construct the bivariate and trivariate joint probability distributions with the data of water quantity and water quality of Xiaoqing River in

Jinan that were obtained during 1980–2016. The proper distribution was chosen using K–S and the RMSE tests, with the parameters of copulas and the results of the goodness fit being shown in Table 4. Figure 3 compares the sample empirical distribution with the joint distribution of the pairs of (Q, NH_4^+), (Q, COD_{Mn}), and (Q, NH_4^+, COD_{Mn}). From these figures, we can see that the correlation coefficients R^2 of Gaussian copula are higher. This illustrates that the Gaussian copula is more suitable in establishing both the bivariate joint probability distributions and the trivariate joint probability distribution.

Table 4. The parameters of copulas of Q, NH_4^+, and COD_{Mn}, and results of K–S and RMSE.

	Pairs	Parameters	K–S	RMSE
Frank	Q, NH_4^+	$\theta = -0.6114$	0.1032	0.0541
	Q, COD_{Mn}	$\theta = -0.5890$	0.0867	0.0353
	Q, NH_4^+, COD_{Mn}	$\theta = -0.9315$	0.1078	0.0331
GH	Q, NH_4^+	$\theta = 0.7866$	0.0842	0.0251
	Q, COD_{Mn}	$\theta = 0.7147$	0.0892	0.0337
	Q, NH_4^+, COD_{Mn}	$\theta = 0.9236$	0.1003	0.0343
Gaussian	Q, NH_4^+	$\rho = -0.4036$	0.0812	0.0251
	Q, COD_{Mn}	$\rho = -0.5319$	0.0752	0.0323
		$\rho_{12} = -0.4036$		
	Q, NH_4^+, COD_{Mn}	$\rho_{13} = -0.5319$	0.0906	0.0222
		$\rho_{23} = 0.2645$		

Note: ρ_{12} is the correlation coefficient of (Q, NH_4^+); ρ_{13} is the correlation coefficient of (Q, COD_{Mn}); and ρ_{23} is the correlation coefficient of (NH_4^+, COD_{Mn}). The K–S is 0.1626 according to the 5% significance level when n = 70.

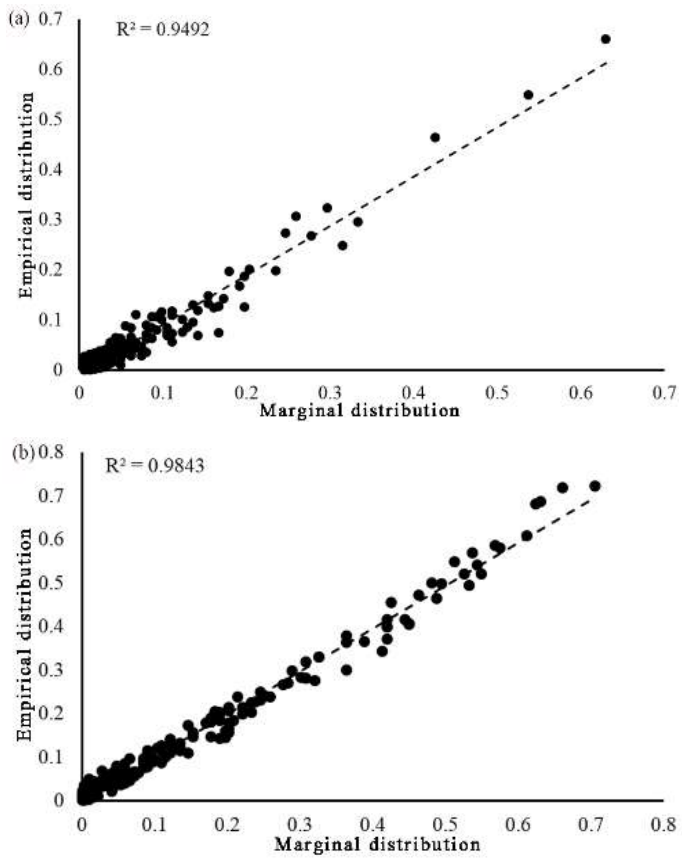

Figure 3. *Cont.*

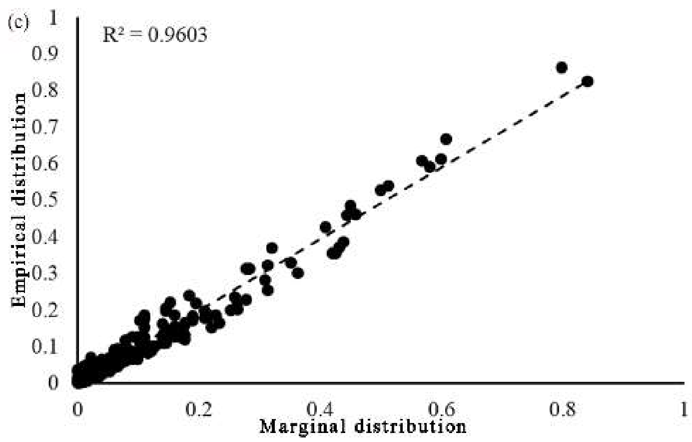

Figure 3. Comparison of the sample empirical distribution and marginal distribution: (**a**) pair of (Q, NH_4^+); (**b**) pair of (Q, COD_{Mn}); and, (**c**) pair of (Q, NH_4^+, COD_{Mn}).

3.2.2. Test of Joint Change-Point of Q, NH_4^+, and COD_{Mn}

According to the optimal fitting copula functions of Q, NH_4^+, and COD_{Mn} using the data from 1980–2016, the joint change-points can be identified. When the number of data is 161, the critical value of Z_n is 13.35 [35]. The results are shown in Figure 4, which show that Z_n is greater than its critical value in No. 94 of the data, which is July 2005. Therefore, this is the change-point of the combination of (Q, NH_4^+, COD_{Mn}). Furthermore, the result also matches the actual situation: the continuous drought for several years before 2005 in Jinan resulted in the management department transferring about 1×10^8 m^3 of water into the Xiaoqing River, which might cause a change in the relationship between the water quantity and quality.

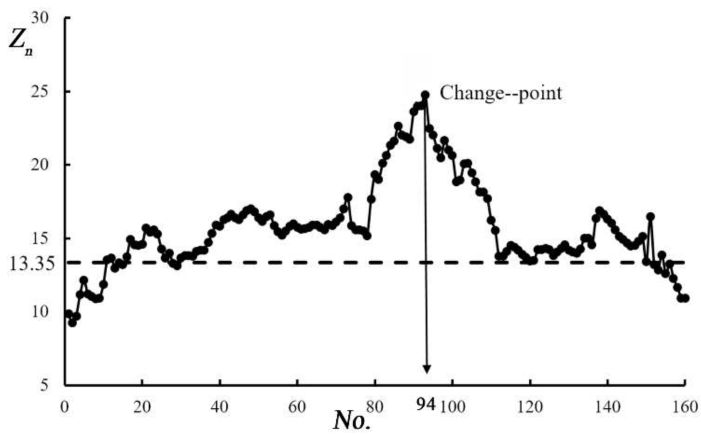

Figure 4. Analysis of the joint change-points of the combination of (Q, NH_4^+, COD_{Mn}).

This indicates that the dependency relationship of the data was inconsistent between 1980 and 2016, so a change-point test was needed to identify the data sequences that are consistent with the current situation.

3.2.3. Establishment of Multivariate Joint Probability Distribution after the Joint Change-Point

The data of Q, NH_4^+, and COD_{Mn} after the joint change-point is more representative of the relationship between water quantity and quality in recent years. Therefore, we re-established the multivariate joint probability distribution using the data after the change-point. The parameters of the copulas and the results of the goodness fit are shown in Table 5. From Table 5, we can see that the joint probability distributions are still consistent with the Gaussian copula function. Meanwhile, their corresponding contour plots and contour surface are presented in Figure 5.

Table 5. The parameters of copulas of Q, NH_4^+, and COD_{Mn} and the results of K–S and RMSE.

	Pairs	Parameters	K–S	RMSE
Frank	Q, NH_4^+	$\theta = -0.5869$	0.1667	0.0654
	Q, COD_{Mn}	$\theta = -0.5971$	0.1546	0.0621
	Q, NH_4^+, COD_{Mn}	$\theta = -0.2998$	0.1816	0.0651
GH	Q, NH_4^+	$\theta = 0.8295$	0.2136	0.0711
	Q, COD_{Mn}	$\theta = 0.9509$	0.2095	0.0604
	Q, NH_4^+, COD_{Mn}	$\theta = 2.8152$	0.1791	0.0743
Gaussian	Q, NH_4^+	$\rho = -0.4073$	0.1431	0.0462
	Q, COD_{Mn}	$\rho = -0.4037$	0.1315	0.0493
	Q, NH_4^+, COD_{Mn}	$\rho_{12} = -0.4073$ $\rho_{13} = -0.4037$ $\rho_{23} = 0.4151$	0.1163	0.0524

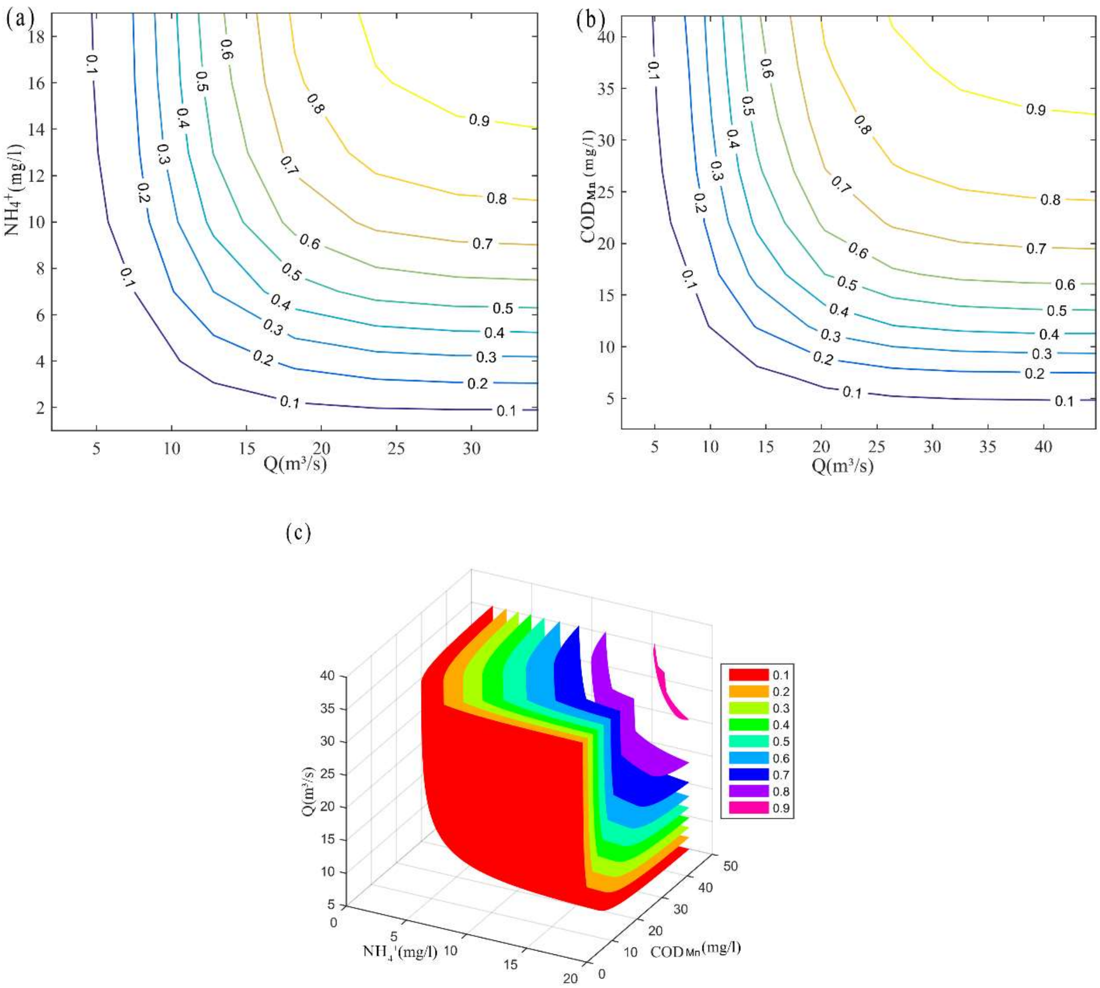

Figure 5. The contour plots and contour surface: (**a**) bivariate joint probability of (Q, NH_4^+); (**b**) bivariate joint probability of (Q, COD_{Mn}); and, (**c**) trivariate joint probability of the combination of (Q, NH_4^+, COD_{Mn}).

3.3. Analysis of Multivariate Joint Probability Distribution of Q, NH_4^+ and COD_{Mn} after the Joint Change-Point

According to the real situation of the river water quality and the existing water quality standards "Environmental Quality Standards for Surface Water" (GB3838-2002), we conducted a cluster analysis of the water quality to obtain the new classes of water quality, which are shown in Table 6.

Table 6. Classes of water quality for NH_4^+ and COD_{Mn}.

	Cluster Center				Boundary Values				
	Center 1	**Center 2**	**Center 3**	**Center 4**	**Class 1**	**Class 2**	**Class 3**	**Class 4**	**Class 5**
NH_4^+	4.2	6.3	8.5	11.2	<4.2	4.2~6.3	6.3~8.5	8.5~11.2	>11.2
COD_{Mn}	8.0	12.0	24.0	35.0	<8.0	8.0~12.0	12.0~24.0	24.0~35.0	>35.0

Using the constructed joint probability distribution of Q, NH_4^+, and COD_{Mn}, the probability of the water quality exceeding a specific value under the different conditions of discharge can be analyzed by Equations (3) and (4). $P_{U_1U_2}$ and $P_{U_1U_3}$ are the bivariate joint probability distributions of the pair of (Q, NH_4^+) and the pair of (Q, COD_{Mn}), respectively, which represent the exceedance probability of NH_4^+ or COD_{Mn} in different discharge situations. $P_{U_1U_2U_3}$ is the trivariate joint probability distribution of the pair of (Q, NH_4^+, COD_{Mn}), which represents the combined exceedance probability of (NH_4^+, COD_{Mn}) in different discharge situations. The contour plots of $P_{U_1U_2}$ and $P_{U_1U_3}$ are given in Figure 6a,b, respectively, while the contour surface of $P_{U_1U_2U_3}$ is shown in Figure 6c.

From Figure 6a,b, we can see that when $P_{U_1U_2}$ or $P_{U_1U_3}$ is in the range of 0.1–0.4, the contour plot is relatively sparse, especially in the upper right panel of the figures. This indicates that when the concentration of NH_4^+ or COD_{Mn} is high, the joint probability is insensitive to a change in discharge. For the combination of NH_4^+ and COD_{Mn}, Figure 6c shows that, when $P_{U_1U_2U_3}$ is between 0.1–0.3, the contour surface is relatively sparse. This is mainly due to the large amount of pollutants in the water body of Xiaoqing River at this moment. At the same time, the tributaries also carry pollutants so the joint probability is insensitive to a change in the discharge.

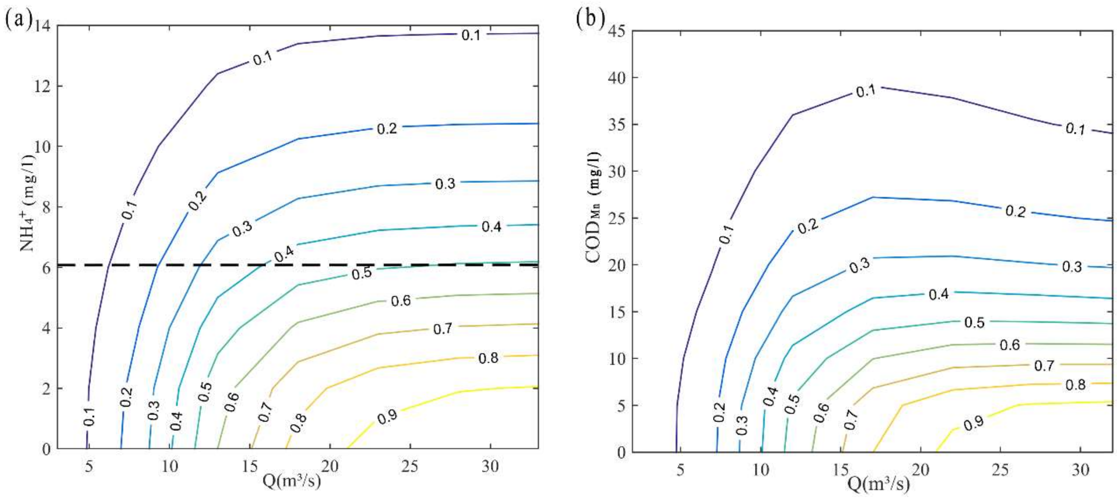

Figure 6. *Cont.*

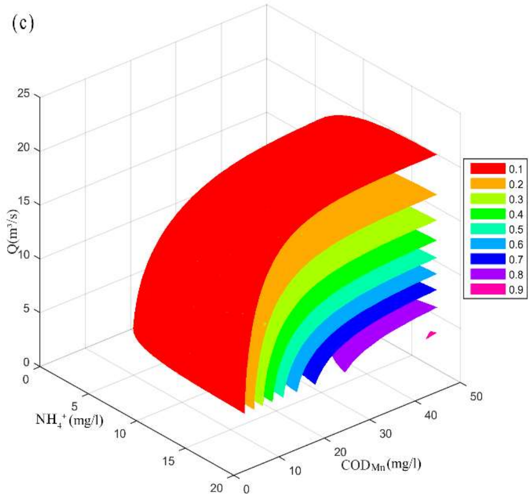

Figure 6. The contour plots and contour surface: (**a**) $P_{U_1U_2}$; (**b**) $P_{U_1U_3}$; and, (**c**) $P_{U_1U_2U_3}$.

The typical values of the joint probability of Q, NH_4^+, and COD_{Mn} are given in Tables 7–9, which describe the probability of the water quality indicators exceeding the specific values with different combinations of discharge and water quality indicators. For example, under the condition that only NH_4^+ is considered, $P_{U_1U_2} = P(U_1 \leq 12, U_2 \geq 6.3) = 0.1985$. Under the condition that only COD_{Mn} is considered, $P_{U_1U_3} = P(U_1 \leq 12, U_3 \geq 12.0) = 0.2135$. Finally, under the condition that NH_4^+ and COD_{Mn} are synergistically considered, $P_{U_1U_2U_3} = P(U_1 \leq 12, U_2 \geq 6.3, U_3 \geq 12.0) = 0.2493$. The results show that the trivariate joint probability of the combination of (Q, NH_4^+, COD_{Mn}) is larger than the bivariate joint probability of the pairs of (Q, NH_4^+) and (Q, COD_{Mn}). This indicates that the probability of exceeding the standards of multiple indicators is greater than that of exceeding individual indicators, which is more consistent with the actual situation. Therefore, only considering the discharge and individual water quality indicator cannot reasonably reflect the relationship between the discharge and water quality. The trivariate joint probability distribution considering Q, NH_4^+, and COD_{Mn} can more comprehensively express the characteristics of discharge and multiple water quality indicators.

Table 7. The typical values of exceedance probability of NH_4^+.

NH_4^+	**Q(m^3/s)**				
(mg/L)	**3**	**6**	**12**	**24**	**36**
4.2	0.0133	0.0653	0.2589	0.4079	0.4177
6.3	0.0120	0.0552	0.1985	0.2916	0.2964
8.5	—	0.0426	0.1379	0.1900	0.1921
11.2	—	0.0287	0.0827	0.1070	0.1078

Note: "—" denotes impossible events.

Table 8. The typical values of exceedance probability of COD_{Mn}.

COD_{Mn}	**Q(m^3/s)**				
(mg/L)	**3**	**6**	**12**	**24**	**36**
8.0	0.0135	0.0645	0.2638	0.4147	0.4231
12.0	0.0134	0.0587	0.2135	0.3043	0.3077
24.0	—	0.0341	0.0896	0.1077	0.1080
35.0	—	0.0184	0.0395	0.0443	0.0443

Note: "—" denotes impossible events.

Table 9. The typical values of combined exceedance probability of NH_4^+ and COD_{Mn}.

NH_4^+	COD_{Mn}	Q(m^3/s)				
(mg/L)	(mg/L)	3	6	12	24	36
4.2	8	0.0138	0.1056	0.3897	0.5596	0.5658
	12	0.0128	0.0964	0.3194	0.4244	0.4270
	24	0.0096	0.0566	0.1387	0.1607	0.1609
	35	0.0064	0.0308	0.0625	0.0685	0.0685
6.3	8	0.0120	0.0894	0.3011	0.4094	0.4125
	12	0.0116	0.0819	0.2493	0.3173	0.3187
	24	0.0088	0.0487	0.1116	0.1263	0.1264
	35	0.0059	0.0268	0.0514	0.0554	0.0554
8.5	8	0.0101	0.0691	0.2108	0.2725	0.2740
	12	0.0098	0.0635	0.1766	0.2159	0.2165
	24	—	0.0384	0.0818	0.0905	0.0906
	35	—	0.0214	0.0386	0.0411	0.0411
11.2	8	—	0.0467	0.1275	0.1569	0.1575
	12	—	0.0432	0.1082	0.1272	0.1274
	24	—	—	0.0522	0.0565	0.0566
	35	—	—	0.0254	0.0266	0.0266

Note: "—" denotes impossible events.

3.4. Discussion

There is not always a positive correlation between the discharge and water quality indicators, which means that the correlation between variables is asymmetric. In this paper, although Frank and G-H Copulas, which are symmetric copula functions, can fit the probability distribution between the discharge and water quality, the fitting accuracy, and the simulation of higher dimensions are not sufficient when compared to the asymmetric Gaussian copulas. Based on the statistical analysis of the historical data of discharge and water quality, this paper provides a method to quantitatively evaluate the exceedance probability of water quality by considering the joint probability of Q, NH_4^+, and COD_{Mn}. When compared with the previous non-mechanical methods, this method adopts a multi-dimensional joint distribution that is based on a copula to more accurately describe the response relationship between water quality and discharge under the condition of the coexistence of multiple pollutants in rivers.

The trivariate joint probability distribution of the combination of (Q, NH_4^+, COD_{Mn}) can reflect the joint probability for the different combinations of Q, NH_4^+, and COD_{Mn}. Thus, when the combination events (NH_4^+, COD_{Mn}) exceed their given design values in different discharge situations, the exceedance probability can be calculated. Generally, regardless of whether the pollutant is NH_4^+, COD_{Mn}, or a combination of NH_4^+ and COD_{Mn}, the joint probability tends to increase with an increase in the discharge and a decrease in the concentration of pollutants discharged into rivers. Thus, the exceedance probability of the water quality can be reduced by an increase in discharge (Q) and a decrease in concentration of pollutants (NH_4^+ and COD_{Mn}). However, with an increase in discharge, the counter lines or the counter surface tend to level, which means that there is a limit imposed on the improvement of water quality. Therefore, further research is needed to determine the limit value. In addition, we suggest that the administration should strengthen the monitoring of water quality at each entry section of Xiaoqing River so that we can obtain a more comprehensive relationship between the discharge and water quality.

4. Conclusions

In this paper, the joint probability distributions of discharge and the main pollutants of NH_4^+ and COD_{Mn} of Xiaoqing River in Jinan were constructed by copulas, before the exceedance probability of water quality combination events under different discharge conditions was analyzed. The specific conclusions are as follows:

(1) The Gaussian copula is more suitable for describing the multivariate joint probability distribution of discharge and water quality. As the relationship between the discharge and water quality indicators is not always positive, the Gaussian copula is more suitable than the Archimedean copulas in the simulation of trivariate or the above joint distribution.
(2) Based on the copula, the joint change-point can be identified. For urban rivers, the dependence of water quantity and quality is often affected by human activities, which leads to the emergence of change-point. Therefore, it is necessary to identify the mutation points.
(3) The trivariate joint probability distribution of the combination of (Q, NH_4^+, COD_{Mn}) is more suitable for estimating the various exceedance probability of the water quality effectively under different discharge conditions. Thus, when the combination events (NH_4^+, COD_{Mn}) exceed their given values in the specific discharge situation, the exceedance probability can be calculated.

Author Contributions: Y.L. and X.Z. collected and analyzed the data; Y.C. and X.L. performed the collation of figures and tables; S.C. contributed analysis tools; Y.L. carried out calculation and completed the manuscript.

Funding: This study was funded by the Special Scientific Research Fund of Public Welfare Profession of Ministry of Water Resources of China: Key Technology Research and Demonstration Project of Water Ecological Civilization Construction (No. 201401003) and the Shandong Provincial Water Conservancy Scientific Research and Technology Promotion Project (No. SDSLKY201302).

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Jang, C.S.; Chen, S.K.; Kuo, Y.M. Establishing an irrigation management plan of sustainable groundwater based on spatial variability of water quality and quantity. *J. Hydrol.* **2012**, *414*, 201–210. [CrossRef]
2. Qian, L.; Liu, Y.; Chao, J.Y. The current situation and development trend of China joint regulating of water quality and water quantity. *Environ. Sci. Technol.* **2013**, *36*, 484–487. [CrossRef]
3. Shokri, A.; Haddad, O.B.; Mariño, M.A. Multi-objective quantity-quality reservoir operation in sudden pollution. *Water Resour. Manag.* **2014**, *28*, 567–586. [CrossRef]
4. Zhang, M.N.; Dolatshah, A.; Zhu, W.L.; Yu, G.L. Case Study on Water Quality Improvement in Xihu Lake through Diversion and Water Distribution. *Water* **2018**, *10*, 333. [CrossRef]
5. Zhang, X.; Ran, Q.X.; Xia, J.; Song, X.Y. Jointed distribution function of water quality and water quantity based on Copula. *J. Hydraul. Eng.* **2011**, *42*, 483–489. [CrossRef]
6. Cox, B.A. A review of currently available in-stream water-quality models and their applicability for simulating dissolved oxygen in lowland rivers. *Sci. Total. Environ.* **2003**, *314–316*, 335–377. [CrossRef]
7. Wang, S.W.; Qi, S.Q.; Yu, D.D.; Zhang, Y.W.; Wan, L.H. Forecast and evaluation of water environment quality based on WASP model: A case study on Harbin section of Songhua River. *J. Nat. Disast.* **2015**, *24*, 39–45. [CrossRef]
8. Yu, S.; He, L.; Lu, H.W. An environmental fairness based optimisation model for the decision-support of joint control over the water quantity and quality of a river basin. *J. Hydrol.* **2016**, *535*, 366–376. [CrossRef]
9. Phan, T.D.; Smart, J.C.R.; Capon, S.J.; Hadwen, W.L.; Sahin, O. Applications of Bayesian belief networks in water resource management. *Environ. Model. Softw.* **2016**, *85*, 98–111. [CrossRef]
10. Yang, B.; Lai, C.G.; Chen, X.H.; Wu, X.Q.; He, Y.H. Surface water quality evaluation based on a Game Theory-Based Cloud model. *Water* **2018**, *10*, 510. [CrossRef]
11. Jia, H.; Xu, T.; Liang, S.; Zhao, P.; Xu, C. Bayesian framework of parameter sensitivity, uncertainty, and identifiability analysis in complex water quality models. *Environ. Model. Softw.* **2018**, *104*, 13–26. [CrossRef]
12. Sun, Z.B.; Wang, B.L.; Ji, H.F.; Huang, Z.Y.; Li, H.Q. Water quality prediction based on probability-combination. *China Environ. Sci.* **2011**, *31*, 1657–1662.

13. Junyi, N.; Hu, H.; Na, C. Combined risk prediction in the water environment based on an MS-AR model and copula theory. *WST* **2013**, *67*, 1967. [CrossRef] [PubMed]
14. Parmar, K.S.; Bhardwaj, R. River water prediction modeling using neural networks, fuzzy and wavelet coupled model. *Water Resour. Manag.* **2015**, *29*, 17–33. [CrossRef]
15. Zhao, F.; Li, C.H.; Chen, L.B.; Zhang, Y. An integrated method for accounting for water environmental capacity of the River–Reservoir combination system. *Water* **2018**, *10*, 483. [CrossRef]
16. Wang, G.; Wang, S.; Kang, Q.; Duan, H.; Wang, X. An integrated model for simulating and diagnosing the water quality based on the system dynamics and Bayesian network. *WST* **2016**, *74*, 2639. [CrossRef] [PubMed]
17. Barzegar, R.; Adamowski, J.; Moghaddam, A.A. Application of wavelet-artificial intelligence hybrid models for water quality prediction: A case study in Aji-Chay River, Iran. *Stoch. Environ. Res. Risk A* **2016**, *30*, 1–23. [CrossRef]
18. Chowdhary, H.; Escobar, L.A.; Singh, V.P. Identification of suitable copulas for bivariate frequency analysis of flood peak and flood volume data. *Hydrol. Res.* **2011**, *42*, 193–216. [CrossRef]
19. Bezak, N.; Mikoš, M.; Šraj, M. Trivariate frequency analyses of peak discharge, hydrograph volume and suspended sediment concentration data using copulas. *Water Resour. Manag.* **2014**, *28*, 2195–2212. [CrossRef]
20. Fontanazza, C.M.; Freni, G.; La, L.G.; Notaro, V. Uncertainty evaluation of design rainfall for urban flood risk analysis. *WST* **2011**, *63*, 2641–2650. [CrossRef]
21. Luca, D.L.D.; Biondi, D. Bivariate return period for design hyetograph and relationship with T-Year design flood peak. *Water* **2018**, *9*, 673. [CrossRef]
22. Guo, S.L.; Muhammad, R.; Liu, Z.J.; Xiong, F.; Yin, J.B. Design flood estimation methods for Cascade Reservoirs based on Copulas. *Water* **2018**, *10*, 560. [CrossRef]
23. Bazrafshan, J.; Nadi, M.; Ghorbani, K. Comparison of empirical copula-based joint deficit index (JDI) and multivariate standardized precipitation index (MSPI) for drought monitoring in Iran. *Water Resour. Manag.* **2015**, *29*, 2027–2044. [CrossRef]
24. Salvadori, G.; Michele, C.D. Multivariate real-time assessment of droughts via copula-based multisite Hazard Trajectories and Fans. *J. Hydrol.* **2015**, *526*, 101–115. [CrossRef]
25. Katipoğlu, O.M.; Can, İ. Determining the lengths of dry periods in annual and monthly stream flows using runs analysis at Karasu River, in Turkey. *WST* **2017**, *18*. [CrossRef]
26. Li, Y.; Chen, C.; Sun, C. Drought severity and change in Xinjiang, China, over 1961–2013. *Hydrol. Res.* **2017**, *48*, 1343–1362. [CrossRef]
27. Muhammad, A.; Maeng, S.J.; San Kim, H.; Murtazaev, A. Copula-Based stochastic simulation for regional drought risk assessment in South Korea. *Water* **2018**, *10*, 359. [CrossRef]
28. Jaewon, K.; Kim, S.; Kim, G.; Singh, V.P.; Park, J.; Kim, H.S. Bivariate drought analysis using streamflow reconstruction with Tree Ring Indices in the Sacramento Basin, California, USA. *Water* **2016**, *8*, 122. [CrossRef]
29. Sklar, A. Fonctions de rèpartition à n dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris* **1959**, *8*, 229–231.
30. Wang, Y.; Li, C.Z.; Liu, J.; Yu, F.L.; Qiu, Q.T.; Tian, J.Y.; Zhang, M.J. Multivariate Analysis of Joint Probability of Different Rainfall Frequencies Based on Copulas. *Water* **2017**, *9*, 198. [CrossRef]
31. Genest, C.; Favre, A.C.; Béliveau, J.; Jacques, C. Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data. *Water Resour. Res.* **2007**, *43*, 223–236. [CrossRef]
32. Guo, A.; Qiang, H.; Wang, Y.; Li, Y.; Chang, J.; Mo, S. Detection of variations in precipitation-runoff relationship based on archimedean copula. *J. Hydroelectr. Eng.* **2015**, *34*, 7–13.
33. Fan, Y.R.; Huang, W.W.; Huang, G.H.; Li, Y.P.; Huang, K.; Li, Z. Hydrologic risk analysis in the Yangtze River basin through coupling Gaussian mixtures into copulas. *Adv. Water Resour.* **2016**, *88*, 170–185. [CrossRef]
34. Song, S.B.; Nie, R. Asymmetric Archimedean Copulas for multivariate hydrological drought frequency analysis. *J. Hydroelectr. Eng.* **2011**, *30*, 20–29.
35. Dias, A.D.C. Copula Inference for Finance and Insurance. Ph.D. Thesis, Swiss Federal Institute of Technology Zurich, Zurich, Swiss, 2004.