

Article

# Development of a Maximum Entropy-Archimedean Copula-Based Bayesian Network Method for Streamflow Frequency Analysis—A Case Study of the Kaidu River Basin, China

Xiangming Kong <sup>1</sup>, Xueting Zeng <sup>2</sup>, Cong Chen <sup>3</sup>, Yurui Fan <sup>4,\*</sup>, Guohe Huang <sup>5,\*</sup>, Yongping Li <sup>5</sup> and Chunxiao Wang <sup>1</sup>

<sup>1</sup> School of Fundamental Science, Beijing Polytechnic, Beijing 100176, China; xiangming0913@126.com (X.K.); chunxiao20083366@163.com (C.W.)

<sup>2</sup> School of Labor Economics, Capital University of Economics and Business, Beijing 100070, China; zxt1231@sina.com

<sup>3</sup> Donlinks School of Economics and Management, University of Science and Technology Beijing, Beijing 100083, China; chencong@ustb.edu.cn

<sup>4</sup> Department of Civil and Environmental Engineering, Brunel University, London, Uxbridge UB8 3PH, UK

<sup>5</sup> Center for Energy, Environment and Ecology Research, Beijing Normal University, Beijing 100875, China; yongping.li33@gmail.com

\* Correspondence: yurui.fan@gmail.com (Y.F.); huangg@iseis.org (G.H.)

Received: 28 October 2018; Accepted: 17 December 2018; Published: 27 December 2018



**Abstract:** Frequency analysis of streamflow is critical for water-resources system planning, water conservancy projects and the mitigation of hydrological extremes events. In this study, a maximum entropy-Archimedean copula-based Bayesian network (MECBN) method has been proposed for frequency analysis of monthly streamflow in the Kaidu River Basin, which integrates the maximum entropy-Archimedean copula (MEAC) and Bayesian network methods into a general framework. MECBN is effective for representing the uncertainties that exist in model representation, preserving the distributional characteristics of streamflow records and addressing the correlation structure between streamflow pairs. Application to the Kaidu River Basin shows a good performance of MECBN in describing the historical data of this basin in China. The results indicate that the interactions between two adjacent monthly streamflow pairs are non-linear. There is upper tail dependence between monthly streamflow pairs. The dependence coefficients including Spearman's rho, Kendall's tau, and the upper tail dependence coefficient are in inverse proportion of monthly streamflow values in the Kaidu River Basin, due to the fact that other factors (i.e., rainfall, snow melting, evapotranspiration rate and requirement of water use) provide more contributions to the streamflow in the flooding season. These findings can be used for providing vital information in the prevention and control of hydrological extremes and to further water resources planning in Kaidu River Basin.

**Keywords:** maximum entropy; Archimedean Copula; Bayesian network; frequency analysis; Kaidu River Basin

## 1. Introduction

Many regions especially in developing countries are suffering from severe water stresses resulting from heterogenous precipitation, extreme hydrological events, water shortages, as well as numerous demands from socio-economic and natural systems [1–5]. However, extensive uncertainties in hydrological processes are among the major challenges to deriving reliable water system management

strategies. Uncertainties in hydrological processes have been addressed with a large number of research works [6–10]. For hydrological processes, the variation of natural circumstances, the lack of historical records, and the limitation of measurement may cause uncertainties in input information, i.e., randomness in the streamflow inputs [11]. Moreover, the interactive relationship among historical data records of streamflow may lead to uncertainties in hydrological predictions, e.g., sub-optimal parameter values and errors due to incomplete or biased model structures. All these uncertainties may affect the resulting results of frequency analysis for streamflow. Consequently, it is necessary to develop effective tools to identify and analyze these uncertainties in order to preserve the distributional characteristics of streamflow records, maintain the dependence structure of such records, and quantify the uncertainties existing in the hydrological processes.

In order to identify and analyze uncertainties in hydrological processes, a large number of methods have been proposed in recent decades [12–20]. For example, a regional frequency analysis of annual maximum streamflow for Gipuzkoa was developed by Erro and López [21]; the results showed that Gipuzkoa could be characterized by generalized logistic distribution. Schnier and Cai proposed a model tree ensembles (MTEs) method to predict streamflow frequency statistics, and they obtained the results that MTEs outperform global multiple-linear regression models in terms of predictions in watersheds which are ungauged [22]. A non-stationary flood frequency model was proposed by Zhang et al. [23] to analyze the annual peak streamflow data in the Pearl River Basin, China, which showed that the streamflow trends significantly impacted the estimation of flood frequencies and investigated the reason of changes in hydrological extremes. Among these methods, the Bayesian network was recently used since it can provide the opportunity to quantify the uncertainty of parameter and model presentation [24]. For instance, Nagarajan et al. utilized a scalable spatiotemporal method based on the Bayesian network to estimate streamflow [25]; the results indicated prediction accuracy of the developed method well. Mediero et al. applied the Bayesian network method to a homogeneous region in the Tagus Basin to estimate the flood quantile [26]; the results denoted as a probability distribution of discharges could supply information about the prediction uncertainty. The formal safety assessment and the Bayesian network technique was incorporated by Zhang et al. [27] to estimate the navigational risk of the Yangtze River, China; the advantage of the proposed model was that both probability and consequences of accidents were considered. D'Addabbo et al. [28] employed the Bayesian network to monitor flood events, which proved that the Bayesian network method can help gain insight into the complex phenomena related to floods.

As a probabilistic computational structure, the Bayesian network infers the joint probability distributions of the related random variables from observations. In addition, the Bayesian network allows taking different sources of estimation uncertainty into account. However, it is still a challenge as to whether the joint probability distributions in the Bayesian network can address the interactive and dependence structure between related random variables accurately. The copula method has been widely applied to complex hydrologic phenomena with strongly correlated variables. Using the copula method, one may successfully capture the non-linear dependence between random variables. Taking into account of the capability of reflecting correlation between random variables, the copula method was introduced into the Bayesian network framework. For example, Madargar and Moradkhani proposed copula-based modeling of Bayesian networks for seasonal drought forecasting [29]; then, they extended the work and proposed a probabilistic forecast model by using a statistical forecast model within Bayesian networks, which was applied for spatial variations of future droughts evaluation for the Gunnison River Basin in Gunnison, CO, USA (Madargar and Moradkhani) [30]. However, in terms of the copula method, the marginal distributions in the copula function were determined by parametric distributions, which may be unable to preserve the distributional characteristics of the historical records properly. In addition, the copula-based Bayesian network approach has not proved its capability and efficiency for frequency analysis of monthly streamflow in rivers of China.

Therefore, this study aims to advance a maximum entropy-Archimedean copula-based Bayesian network (MECBN) method for frequency analysis of monthly streamflow. The MECBN method is

then applied to a case study of the Kaidu River Basin in the Xinjiang Uyghur Autonomous Region, an arid region in northwest China. In the MECBN method, the maximum entropy-Archimedean copula approach is incorporated into a Bayesian network framework according to the dependences between adjacent monthly streamflows in the Kaidu River Basin. MECBN is used for: (i) assessing the uncertainties which exist in the model representation; (ii) addressing the correlation structure between streamflows in two adjacent months, and (iii) presenting the key statistical characters (i.e., mean, standard deviation, skewness and kurtosis) for the predictive streamflow.

## 2. Methodology

### 2.1. Maximum Entropy Method

The determination of marginal distribution for monthly streamflow is critical. The accuracy of marginal distribution significantly affects the follow-up study for interactions of adjacent monthly streamflows. As a method to generate distribution of random variable, the maximum entropy method has been widely used in hydrology. It can present the key statistical characters of monthly streamflow properly without any hypothesis [31]. The marginal probability density function (PDF) for random variable  $X$  can be obtained by maximizing the Shannon entropy with mean, standard deviation, skewness and kurtosis as constraints. The marginal cumulative density function (CDF) can be obtained by computing the integral for PDF. Then, the generated marginal PDF and CDF can be expressed as follows:

$$f(x) = \exp \left[ -\ln \left( \int_a^b \exp \left( -\sum_{i=1}^m \lambda_i h_i(x) \right) dx \right) - \sum_{i=1}^m \lambda_i h_i(x) \right] \tag{1}$$

$$E_X(x) = \int_a^x f(t) dt \tag{2}$$

where  $a$  and  $b$  are the lower and upper bounds of random variable  $X$ , respectively.  $\lambda_i (i = 1, 2, \dots, m)$  are the Lagrange multipliers which could be determined by the conjugate gradient (CG) method [31];  $h_i(x)$  is a known function of  $X$ , which can be specified as  $h_1 = x, h_2 = x^2, h_3 = x^3$  and  $h_4 = x^4$  for the constraints of mean, standard deviation, skewness and kurtosis Hao et al. [32].

### 2.2. Archimedean Copula

After obtaining marginal distribution of monthly streamflow, joint distributions should be further constructed in order to further investigate interactions between adjacent monthly streamflows. The Archimedean copula is widely used in hydrologic frequency analysis. It can generate joint distribution through combining marginal distributions into a copula function. The superior property of the Archimedean copula is that it can be applied when the correlation amongst hydrologic variables is positive or negative [33–35]. In this study, one-parameter Archimedean copulas including the Gumbel–Hougaard, Clayton and Frank copulas are used to generate the joint distributions of adjacent monthly streamflows. The expressions for two random variables,  $X$  and  $Y$ , can be defined as follows [34]:

$$C(u_1, u_2) = \varphi^{-1}[\varphi(u_1), \varphi(u_2)], \quad 0 < u_1, u_2 < 1 \tag{3}$$

$$C_{\theta'}^{GH}(u_1, u_2) = \exp \left\{ - \left[ -(\log u_1)^{\theta'} + (-\log u_2)^{\theta'} \right]^{1/\theta'} \right\}, \quad \theta' \geq 1 \tag{4}$$

$$C_{\theta''}^{CT}(u_1, u_2) = \left( u_1^{-\theta''} + u_2^{-\theta''} - 1 \right)^{-1/\theta''}, \quad 0 < \theta'' < \infty \tag{5}$$

$$C_{\theta'''}^{FK}(u_1, u_2) = \frac{1}{\theta'''} \log \left( 1 + \frac{(e^{\theta''' u_1} - 1)(e^{\theta''' u_2} - 1)}{e^{\theta''' u_1} - 1} \right), \quad -\infty < \theta''' < \infty \tag{6}$$

where  $C(u_1, u_2)$  is the one-parameter Archimedean copula in general,  $\varphi(u)$  is the generating function of  $C(u_1, u_2)$ ;  $\varphi^{-1}(u)$  is the inverse function of  $\varphi(u)$ ;  $u_1 = F_X(x)$  and  $u_2 = F_Y(y)$  are the CDFs of  $X$  and  $Y$ , respectively.  $C_{\theta'}^{GH}(u_1, u_2)$ ,  $C_{\theta''}^{CT}(u_1, u_2)$  and  $C_{\theta'''}^{FK}(u_1, u_2)$  are the Gumbel–Hougaard, Clayton and Frank copulas, respectively.  $\theta'$ ,  $\theta''$  and  $\theta'''$  are the unknown parameters to be estimated for the Gumbel–Hougaard, Clayton and Frank copulas, respectively.

### 2.3. Maximum Entropy-Archimedean Copula Method

A maximum entropy-Archimedean copula (MEAC) method can be proposed through integrating techniques of maximum entropy and Archimedean copula into a general framework. In MEAC, the marginal distribution of monthly streamflow is constructed by the maximum entropy method and the joint distribution of two adjacent monthly streamflows is constructed by an Archimedean copula. The joint distribution of the MEAC method can be formulated as follows:

$$P(X \leq x, Y \leq y) = C_{\theta}(u_1, u_2) = C_{\theta}(F_X(x; \lambda), F_Y(y; \lambda')) \tag{7}$$

$$F_X(x; \lambda) = \int_{a_1}^x \exp \left[ -\ln \left( \int_{a_1}^{b_1} \exp \left( -\sum_{i=1}^m \lambda_i h_i(x) \right) dx \right) - \sum_{i=1}^m \lambda_i h_i(x) \right] dt \tag{8}$$

$$F_Y(y; \lambda') = \int_{a_2}^y \exp \left[ -\ln \left( \int_{a_2}^{b_2} \exp \left( -\sum_{i'=1}^{m'} \lambda_{i'} h_{i'}(y) \right) dy \right) - \sum_{i'=1}^{m'} \lambda_{i'} h_{i'}(y) \right] dt \tag{9}$$

where  $C_{\theta}(u_1, u_2)$  is the Archimedean copula with parameter  $\theta$ ;  $u_1 = F_X(x; \lambda)$  and  $u_2 = F_Y(y; \lambda')$  are the marginal distributions of random variables  $X$  and  $Y$ , respectively.  $h_i(x)$  and  $h_{i'}(y)$  are known functions of  $X$  and  $Y$ ;  $a_1$  and  $b_1$  are the lower and upper bounds of the random variable  $X$ , respectively;  $a_2$  and  $b_2$  are the lower and upper bounds of random variable  $Y$ , respectively;  $\lambda_i (i = 1, 2, \dots, m)$  and  $\lambda_{i'} (i' = 1, 2, \dots, m')$  are the Lagrange multipliers. The parameters of MEAC could be estimated by the inference functions for margins (IFM) method [31]. The parameter of the Archimedean copula could be estimated by Kendall's tau. The CG method could be used for the parameter estimation of marginal distribution.

### 2.4. Maximum Entropy-Archimedean Copula-Based Bayesian Network

Based on the MEAC method, the joint distribution of adjacent monthly streamflows could be generated. However, the frequency analysis of monthly streamflow also needs to construct the conditional dependency of adjacent monthly streamflows. Therefore, a Bayesian network is introduced in this study in order to obtain the conditional dependencies of monthly streamflow. As a probabilistic model, a Bayesian network can describe the conditional dependencies of a set of random variables via directed acyclic graphs [29,30]. The joint PDF of different random variables with a level of dependency  $(x_{t_1}, x_{t_2}, \dots, x_{t_n})$  within a Bayesian network can be defined as follows [30]:

$$f(x_{t_1}, x_{t_2}, \dots, x_{t_p}) = \prod_{t_j \in T} f(x_{t_j} | x_{t_1}, \dots, x_{t_{j-1}}) \tag{10}$$

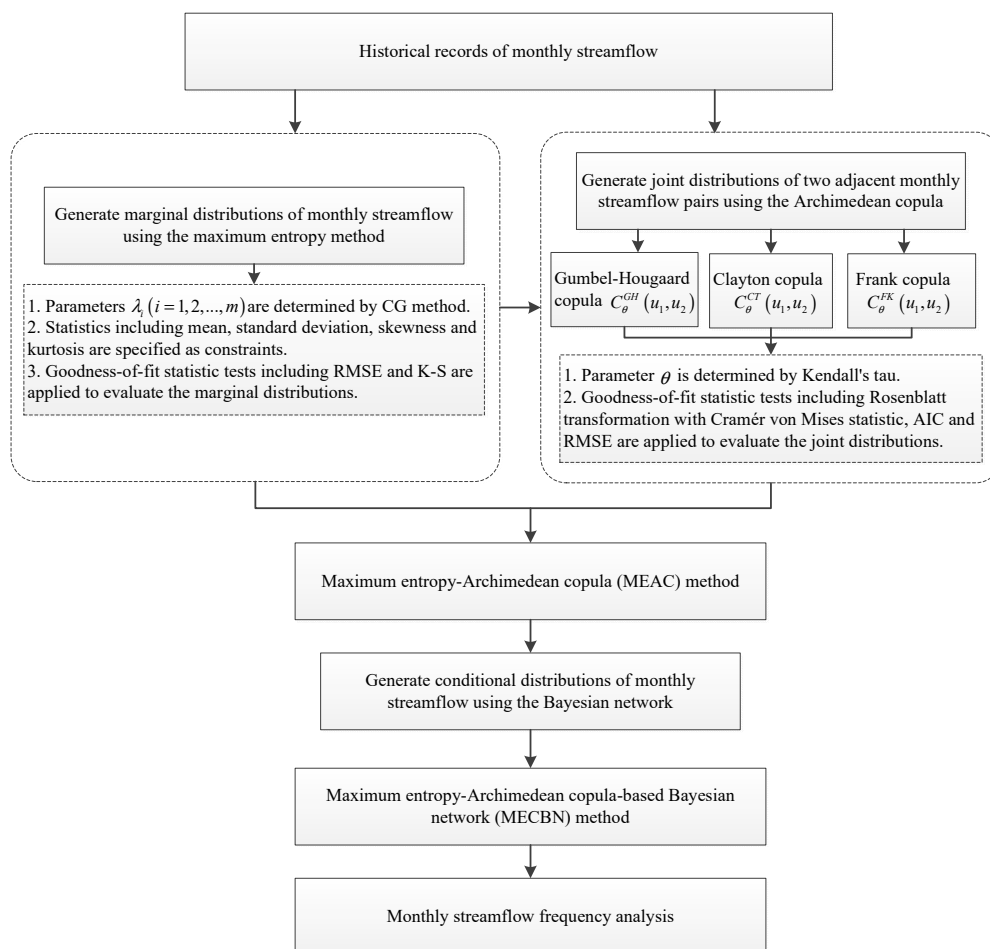
Based on Equation (10), the expression of conditional PDF of two adjacent monthly streamflows,  $X$  and  $Y$ , can be described as follows:

$$\begin{aligned} f(y|x) &= \frac{f(x,y)}{f_X(x)} \\ &= \frac{c_{\theta}(u_1, u_2) f_X(x) f_Y(y)}{f_X(x)} \\ &= c_{\theta}(u_1, u_2) f_Y(y) \end{aligned} \tag{11}$$

where  $c_\theta(u_1, u_2)$  is the PDF of the Archimedean copula with parameter  $\theta$ , which describes the joint PDF of two adjacent monthly streamflows;  $u_1 = F_X(x; \lambda)$  and  $u_2 = F_Y(y; \lambda')$  are the marginal distributions of monthly streamflows whose marginal PDFs are  $f_X(x)$  and  $f_Y(y)$ , respectively. The PDF of copula  $c_\theta(u_1, u_2)$  can be defined as follows:

$$c_\theta(u_1, u_2) = \frac{C_\theta(u_1, u_2)}{\frac{\partial u_1, \partial u_2}{\partial u_1, \partial u_2}} = \frac{C_\theta(F_X(x; \lambda), F_Y(y; \lambda'))}{\frac{\partial u_1, \partial u_2}{\partial u_1, \partial u_2}} \tag{12}$$

The general framework of MECBN for modeling hydrological processes is shown in Figure 1. The MECBN method would be used for monthly streamflow frequency analysis. The application involves the construction of marginal, joint and conditional distributions, and the measurement of dependence. In detail, (a) the marginal distributions of monthly streamflows would be constructed by the maximum entropy method; (b) the joint distributions of two adjacent monthly streamflows would be estimated using Archimedean copula based on the generated marginal distributions; (c) the performance of marginal and joint distributions would be tested by the goodness-of-fit statistics; (d) the conditional distributions of two adjacent monthly streamflows would be generated by the Bayesian network combined with generated marginal and joint distributions; (e) the dependence between two adjacent monthly streamflows would be measured by methods including Spearman’s rho, Kendall’s tau and the tail dependence coefficient.



**Figure 1.** General framework of the maximum entropy-Archimedean copula-based Bayesian network (MECBN).

### 3. Study Area and Measures

#### 3.1. Study Area

The Kaidu River Basin (Figure 2), located in the middle reach of the Tarim River Basin, plays a key role in water supply to municipality, industry, agricultural, stockbreeding and forestry sectors in Xinjiang province of China. It lies between the latitudes of  $42^{\circ}14' N$  and  $43^{\circ}21' N$  and longitudes of  $82^{\circ}58' E$  and  $86^{\circ}05' E$  [36]. The Kaidu River Basin experiences a typical inner-continental climate, with an average annual temperature of  $4.16^{\circ}C$ . The spatial and temporal distribution of precipitation in the Kaidu River Basin is heterogeneous with an average rainfall of 273 mm/year Huang et al. [37]. It exhibits plentiful snowmelt water in spring, concentrated ice-melt water and rainfall recharge during the summer, significant temperature differences during the autumn, and frequent early frost in winter. More than 80% of the rainfall occurs between May and September [38]. The study site is Dashankou hydrological station, which is located at the Kaidu River Basin outlet. The Kaidu River Basin above Dashankou hydrological station has a catchment area of  $18,827 \text{ km}^2$  with the mean elevation of 3100 m. The data of streamflow records are obtained from the Dashankou hydrological station for the period of 1957–2009. The unit for streamflow is in cubic meter per second (CMS). The monthly variation of streamflow is shown in Figure 3, with a maximum value of 413, a minimum value of 33.2, a mean of 110.54, a standard deviation of 68.77, a skewness of 1.37, and a kurtosis of 5.07.

The Kaidu River Basin has important functions for agricultural irrigation and ecological sustainability. It also plays a critical role in protecting the Bosten Lake. However, the regional water-resources systems have been destroyed because of excessive exploitation and environmental degradation, leading to amounts of ecological and environmental problems. Moreover, the streamflow is replenished by snowmelt in spring, sometimes causing floods. Therefore, it is necessary to propose an effective method to analyze the inherent uncertain characteristics on streamflow of the Kaidu River Basin. In this paper, potential streamflow interactions are explored and formalized into a copula function combined with a Bayesian network framework. An application using such frequency analysis method could be developed as part of a strategy for water resources management and precaution of hydrological extremes.

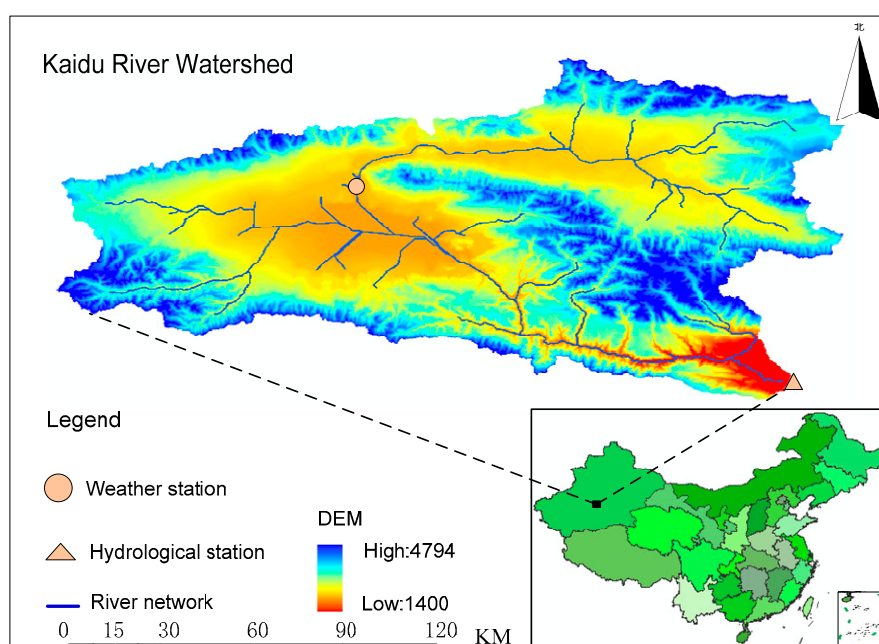


Figure 2. The Kaidu River Basin.

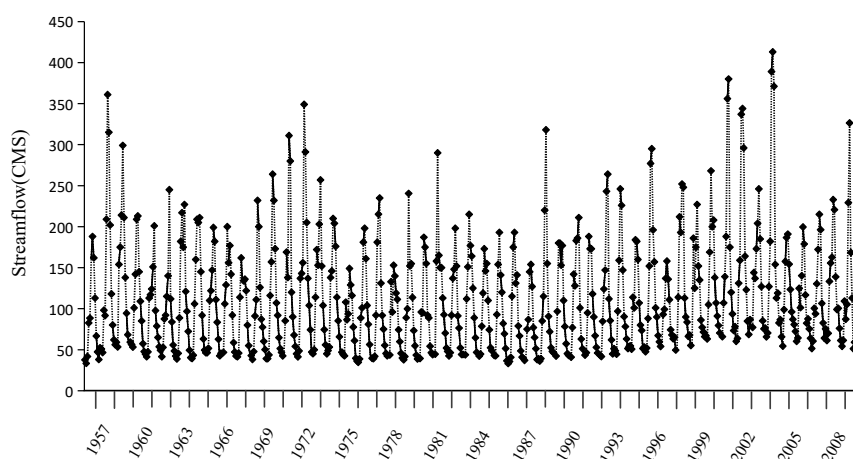


Figure 3. Trend of monthly streamflow from 1957 to 2009.

### 3.2. Dependence Measures

In streamflow analysis, one may be interested in dependence structure between adjacent streamflows, especially for the extreme behavior of the streamflows for risk analysis. Therefore, the rank-based coefficients of correlation including Spearman’s rho and Kendall’s tau, which are well-known non-parametric measures of dependence, are applied to examine the dependence structure of adjacent monthly streamflows. The tail dependence coefficient is used to analyze the extreme behavior of adjacent monthly streamflows [39]. The simplification of Spearman’s rho can be expressed as:

$$\hat{\rho}_s = 1 - \frac{6}{n(n^2 - 1)} \sum_{i=1}^n (R_i - Q_i)^2 \tag{13}$$

where  $\hat{\rho}_s \in [-1, 1]$ ,  $(X_i, Y_i)$ ,  $i = 1, 2, \dots, n$  are the sample of  $(X, Y)$ ;  $R_i$  stands for the rank of  $X_i$  among  $(X_1, X_2, \dots, X_n)$ , and  $Q_i$  stands for the rank of  $Y_i$  among  $(Y_1, Y_2, \dots, Y_n)$ . For Kendall’s tau, it can be expressed as:

$$\hat{\tau} = \frac{c - d}{c + d} = \frac{c - d}{C_n^2} \tag{14}$$

where  $c$  and  $d$  are the number of concordant of discordant pairs, respectively;  $(X_i, Y_i)$  and  $(X_j, Y_j)$  are said to be concordant when  $(X_i - X_j)(Y_i - Y_j) > 0$  and discordant when  $(X_i - X_j)(Y_i - Y_j) < 0$ . The lower and upper tail dependence coefficients are introduced as follows:

$$\lambda^{lo} = \lim P \left[ Y < G^{-1}(u) \mid X < F^{-1}(u) \right] \tag{15}$$

$$\lambda^{up} = \lim P \left[ Y > G^{-1}(u) \mid X > F^{-1}(u) \right] \tag{16}$$

where  $\lambda^{lo}$  and  $\lambda^{up}$  are the lower and upper dependence coefficients, respectively;  $F(x)$  and  $G(x)$  are the marginal distributions of random variable  $X$  and  $Y$ , respectively.

### 3.3. Goodness-of-Fit (GOF)

The goodness-of-fit (GOF) statistic tests for marginal and joint distributions can be performed separately. In this study, the GOF tests for marginal distributions include the root mean square error (RMSE) and the Kolmogorov–Smirnov (K–S) test. The GOF tests for joint distributions consist of the

Rosenblatt transformation with Cramér von Mises statistic, the Akaike information criterion (AIC), and RMSE. The expression of RMSE is Willmott and Matsuura [40]:

$$RMSE = \sqrt{\frac{\sum_{k=1}^N (x_k^{est} - x_k^{obs})^2}{N}} \quad (17)$$

where  $x_k^{est}$  is the estimated value from the fitted distribution;  $x_k^{obs}$  is the corresponding observed value;  $N$  is the sample size.

As a non-parametric probability distribution free test, the K–S test quantifies the largest vertical difference between the specified and empirical distributions. The statistic of the K–S test is defined as follows:

$$T = \sup_x |F^*(x) - F_n(x)| \quad (18)$$

where data points  $x$  are in an increasing order;  $F^*(x)$  and  $F_n(x)$  are the specified distribution and empirical distribution, respectively; If  $T$  exceeds the  $1 - \alpha$  quantile, then we reject the null hypothesis  $H_0$  (The sample data follow the hypothesized distribution);  $\alpha$  is the level of significance.

The GOF tests based on the Rosenblatt transformation with Cramér von Mises statistic, Akaike information criterion (AIC) and RMSE are employed in this study. The empirical distribution for the Rosenblatt transformation and the corresponding Cramér von Mises statistic can be expressed as follows [31]:

$$D_n(\mathbf{v}) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(\mathbf{E}_i \leq \mathbf{v}), \quad \mathbf{v} \in [0, 1] \quad (19)$$

$$\begin{aligned} S_n^{(B)} &= n \int_0^1 \{D_n(\mathbf{v}) - C_{\perp}(\mathbf{v})\}^2 d\mathbf{v} \\ &= \frac{n}{3^2} - \frac{1}{2} \sum_{i=1}^n (1 - E_{i1}^2)(1 - E_{i2}^2) + \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n (1 - E_{i1} \vee E_{j1})(1 - E_{i2} \vee E_{j2}) \end{aligned} \quad (20)$$

where  $\mathbf{E}_i$  stand for the pseudo-observations from the independence copula  $C_{\perp}$ ;  $\mathbf{v}$  is the vector of marginal distributions;  $n$  is the sample size;  $E_{il} \vee E_{jl} = \max(E_{il}, E_{jl})$ , ( $l = 1, 2$ ). The corresponding  $p$ -value of the Cramér von Mises test statistic exceeds the  $1 - \alpha$  quantile, then we reject the null hypothesis  $H_0$  (the hypothesized copula function is the fitted one);  $\alpha$  is the level of significance. Then, AIC can be used for identifying the most appropriate probability distribution, which can be expressed below:

$$AIC = N \log \left[ \frac{1}{N} \sum_{k=1}^N (x_k^{est} - x_k^{obs})^2 \right] + 2(\text{no. of fitted parameters}) \quad (21)$$

The RMSE value for joint distribution can also be determined using Equation (17). The best fitted copula function is the one that has the minimum values of AIC and RMSE.

## 4. Results and Discussion

### 4.1. Marginal Distributions

Firstly, marginal distributions of monthly streamflows are constructed. The Lagrange multipliers for the marginal distributions are estimated by the CG method. The results of estimation are shown in Table 1. Based on the Lagrange multipliers, the PDFs and CDFs of monthly streamflows in the Kaidu River Basin could be determined based on Equations (6) and (7). In order to evaluate the marginal distributions of monthly streamflow in Kaidu River Basin generated by the MECBN method, GOF tests including RMSE and K–S test are applied in this study. The results are shown in Table 2.



**Table 1.** Estimation of Lagrange multipliers for maximum entropy (ME)-based marginal distribution of monthly streamflow.

Month	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$
1	-6.78	16.35	-1.14	-7.60
2	-6.34	13.51	0.39	-5.47
3	-4.27	12.13	-0.62	-4.89
4	-8.62	19.32	2.04	-10.77
5	-3.83	10.90	5.20	-8.47
6	-13.78	26.80	5.40	-17.66
7	-1.12	7.93	-1.07	-3.26
8	-0.28	6.29	-0.53	-3.07
9	-7.05	14.75	1.53	-8.55
10	-6.57	12.78	1.53	-5.80
11	-3.29	9.95	0.27	-3.98
12	-4.73	10.39	-0.99	-3.82

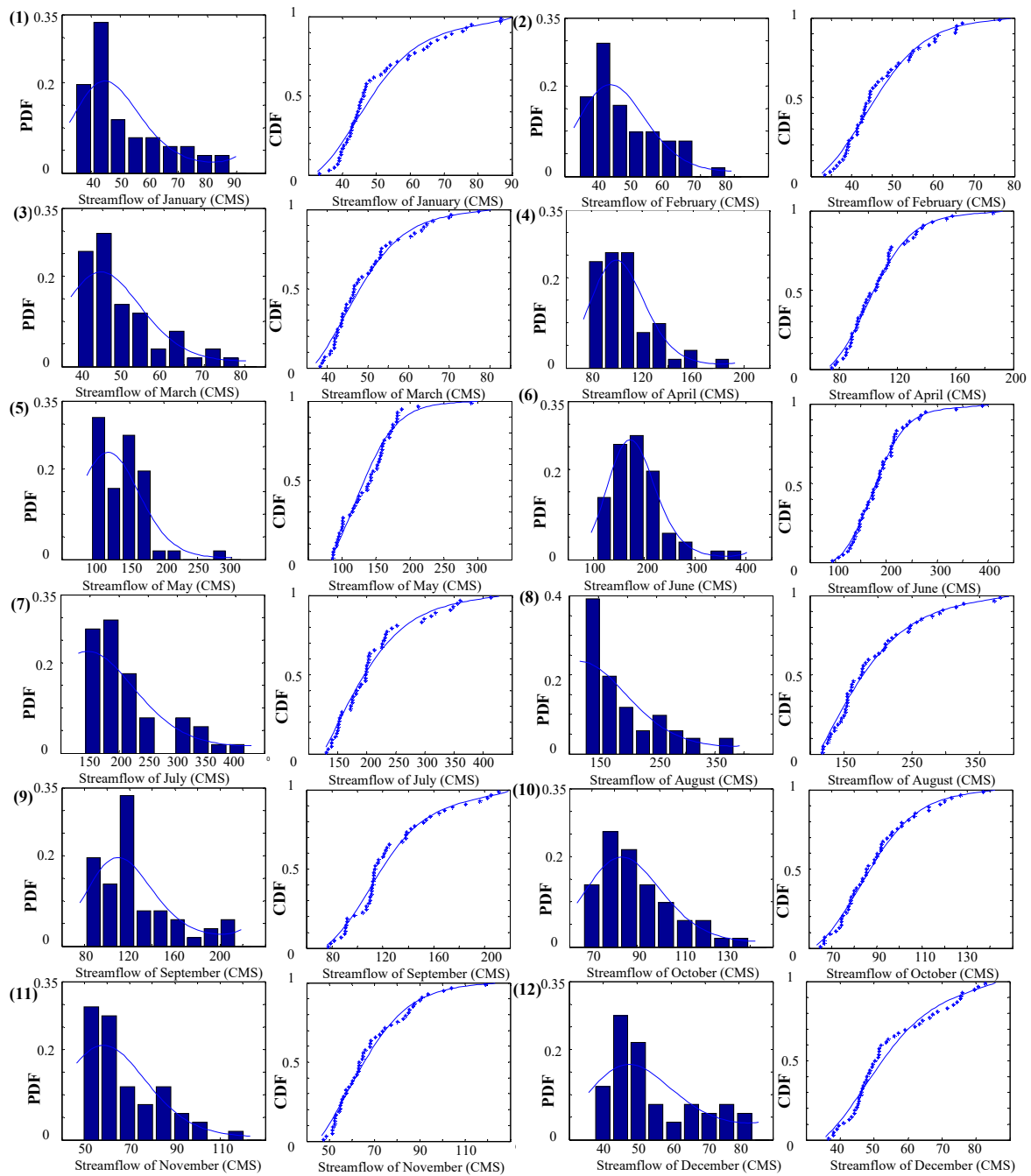
**Table 2.** The root means square error (RMSE) value and Kolmogorov–Smirnov (K–S) test for marginal distribution of monthly streamflow.

Month	RMSE	K–S test	
		$T$	$p$ -Value
1	2.28	0.135	0.145
2	1.46	0.099	0.347
3	1.20	0.081	0.490
4	2.31	0.074	0.546
5	8.62	0.100	0.340
6	5.06	0.055	0.706
7	10.28	0.086	0.447
8	5.41	0.058	0.684
9	4.44	0.096	0.371
10	1.52	0.054	0.722
11	1.82	0.087	0.436
12	2.06	0.107	0.290

As a measure which is regularly employed in model evaluation studies, RMSE indicates the deviation between generated monthly streamflow and observed monthly streamflow. Most RMSE values in Table 2 are small. The maximum value of RMSE is 10.28. This means that the maximum deviation between generated monthly streamflow and observed monthly streamflow occurs in July, while the minimum RMSE can be observed in March with a value of 1.20. The K–S test can quantify the largest vertical difference between the hypothesized and empirical distributions [31,41,42]. Therefore, it can be used to verify if the sample data of monthly streamflow follow the marginal distribution generated by the MECBN method. As presented in Table 2, all the  $p$ -Values calculated from K–S test are higher than the significant level  $\alpha = 0.05$ , where  $T$  is the K–S test statistic. The results indicate that observed monthly streamflows in the Kaidu River Basin could be appropriately represented by the marginal distributions generated by MECBN.

Figure 4 shows the comparison of theoretical and empirical marginal PDF and CDF for monthly streamflow in Kaidu River Basin. Numbers (1) to (12) stand for monthly streamflow from January to December. It could be concluded that the marginal PDF and CDF generated by MECBN are able to capture the shape of empirical histograms and empirical CDF. Therefore, MECBN is effective for generating marginal distributions of monthly streamflows of the Kaidu River Basin. In addition, Figure 4 indicates that most streamflow records in April, May, June, July, August and September are greater than 100 CMS. The streamflow values from April to September, corresponding to the 0.5 value of CDF, are greater than 100 CMS. That is mainly because the distribution of precipitation is uneven in Kaidu River Basin. More than 80% of the total precipitation in a year falls from May to September, and

the rest falls from October to April. However, there is a low rainfall but a high streamflow in April. This is due to the fact that snow melting is the main sources in April in the Kaidu River Basin.



**Figure 4.** Comparison of theoretical and empirical marginal probability density function (PDF) and cumulative density function (CDF). The values are calculated monthly from 1957 to 2009. The abscissa represents streamflow values. The corresponding size varies from month to month. PDF describes the output of variables, and indicates the frequency at a given streamflow value. CDF describes the integrate of PDF, and indicates the probability of a given streamflow value. Lines in PDFs and CDFs show the performance of generated values. Bars and dots depict the empirical values in PDFs and CDFs, respectively.

#### 4.2. Joint Distributions

In this study, Archimedean copulas including Gumbel–Hougaard, Clayton and Frank copulas would be used to generate the joint probability distributions for adjacent monthly streamflows. In order to evaluate the joint distributions generated by MECBN, the GOF tests of the Rosenblatt transformation with the Cramér von Mises statistic, AIC and RMSE would be employed. As shown in Table 3. All the  $p$ -Values of adjacent monthly streamflow pairs are higher than significant level  $\alpha = 0.05$  for the Gumbel–Hougaard and Frank copulas. When employing Clayton copula to generate joint distributions, the  $p$ -Values of months 4–5, 6–7 and 9–10 are lower than the significant level. Results in Table 3 mean that both Gumbel–Hougaard and Frank copulas could be used to generate joint distributions of two adjacent monthly streamflows. In order to choose the most appropriate copula for monthly streamflow pairs in the Kaidu River Basin, the AIC and RMSE values are applied to test goodness of fitness for sample data to the joint distribution obtained by the Gumbel–Hougaard and Frank copulas. Results are given in Table 4. The Gumbel–Hougaard copula would be chosen as the most appropriate copula with the minimum values of AIC and RMSE for monthly streamflow pairs.

**Table 3.** Comparison of Cramér von Mises statistic for joint distribution of different streamflow pairs.

Month	Gumbel-Hougaard		Frank		Clayton	
	$S_n^{(B)}$	$p$ -Value	$S_n^{(B)}$	$p$ -Value	$S_n^{(B)}$	$p$ -Value
1–2	34.03	0.383	32.30	0.782	35.79	0.053
2–3	33.89	0.368	32.31	0.762	35.28	0.063
3–4	32.62	0.552	32.46	0.667	32.40	0.063
4–5	34.31	0.303	34.22	0.408	37.56	0.023
5–6	33.83	0.437	33.32	0.612	33.43	0.068
6–7	33.15	0.542	32.93	0.662	33.96	0.048
7–8	33.53	0.482	33.00	0.672	33.20	0.093
8–9	33.27	0.507	32.38	0.752	33.46	0.088
9–10	34.70	0.288	32.53	0.697	36.36	0.043
10–11	33.39	0.527	32.47	0.722	34.08	0.078
11–12	32.49	0.637	32.16	0.752	33.63	0.083
12–1	33.39	0.437	32.34	0.792	34.15	0.063

**Table 4.** Comparison of Akaike information criterion (AIC) and RMSE values for joint distribution of different streamflow pairs.

Month	Gumbel–Hougaard		Frank	
	AIC	RMSE	AIC	RMSE
1–2	−311.52	0.0462	−202.87	0.1342
2–3	−321.77	0.0418	−223.97	0.1091
3–4	−347.22	0.0329	−284.88	0.0601
4–5	−327.10	0.0397	−326.69	0.0399
5–6	−341.56	0.0345	−325.32	0.0404
6–7	−360.95	0.0285	−320.81	0.0422
7–8	−341.83	0.0344	−277.97	0.0643
8–9	−321.79	0.0418	−224.64	0.1084
9–10	−338.00	0.0357	−257.98	0.0782
10–11	−358.19	0.0293	−239.11	0.0941
11–12	−299.66	0.0520	−215.47	0.1186
12–1	−294.01	0.0549	−229.85	0.1030

Based on the analysis above, marginal distributions of monthly streamflow were generated and the Gumbel–Hougaard copula was chosen to describe the dependence structures for adjacent monthly streamflows in the Kaidu River Basin. Therefore, joint PDF, joint CDF and the corresponding counters are obtained by the MECBN method. Dependence measures (i.e., Table 5) indicate that the adjacent monthly streamflow pairs in the Kaidu River Basin are highly correlated with significant upper tail

dependence. The upper tail dependence between two adjacent monthly streamflows indicates that streamflow of one month in the Kaidu River Basin depends on the streamflow of its adjacent month, and also denotes the probability for one monthly streamflow exceeding a high threshold under the situation that its adjacent monthly streamflow exceeds that threshold. It is important to determine this in order to provide vital information for the prevention and control of hydrological extremes in the Kaidu River Basin.

**Table 5.** Results of dependence measures.

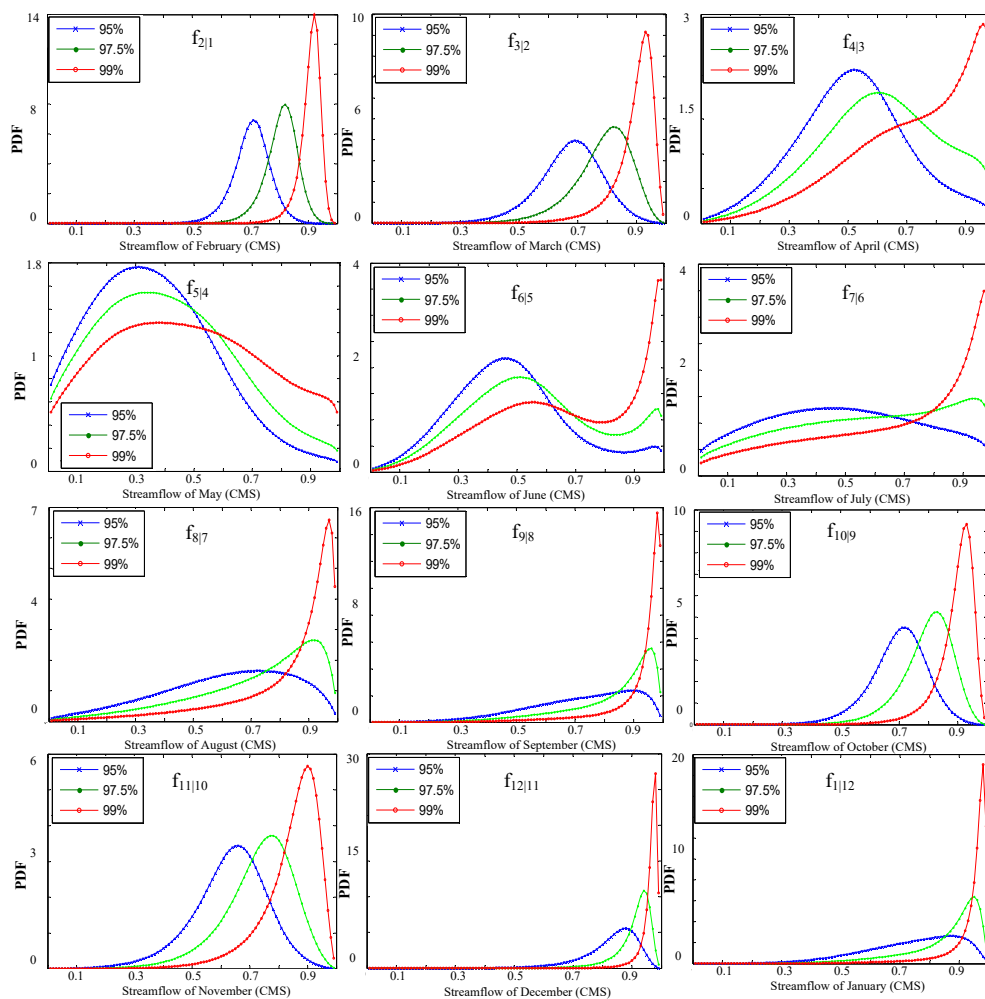
Month	Spearman's Rho $\hat{\rho}_s$	Kendall's Tau $\hat{\tau}$	Upper Tail Dependence Coefficient $\lambda_U^{GH}$
1–2	0.931	0.784	0.838
2–3	0.856	0.678	0.750
3–4	0.541	0.392	0.476
4–5	0.292	0.180	0.235
5–6	0.435	0.312	0.389
6–7	0.410	0.280	0.353
7–8	0.595	0.435	0.521
8–9	0.716	0.527	0.612
9–10	0.853	0.678	0.750
10–11	0.753	0.603	0.683
11–12	0.856	0.691	0.761
12–1	0.775	0.597	0.678

#### 4.3. Conditional Distributions

Once the marginal and joint distributions are fitted to streamflow records in the Kaidu River Basin, the conditional distributions of adjacent monthly streamflows can be derived according to Equation (11). The streamflow status is classified based on the quantile. In this study, 95, 97.5 and 99 percentile values reflect qualitative information on probability distributions of the scaled monthly streamflow of the Kaidu River Basin. They are used as streamflow classification schemes at high level to explore the conditional distribution. As the streamflow has been scaled into  $[0,1]$ , 95% means that the streamflow is 0.95. The relationship of conditional distribution of two adjacent months has been transformed to the exploration of the distribution of the streamflow in each month given the streamflow status (0.95, 0.975, 0.99) in the last month demonstrated in Figure 5.

As shown in Figure 5, if the given streamflow percentile value changes in the previous, the probability distribution of the streamflow in the current month varies, which also exhibits the conclusion that there exists upper tail dependence between every two adjacent monthly streamflows in the Kaidu River Basin. Moreover, the probability distributions under a given streamflow percentile value, and the fluctuation ranges under different streamflow percentile values vary considerably between different streamflow pairs. It indicates that the upper tail dependences between each streamflow pairs are different, which are coincide to the upper tail coefficients presented in Table 5. It is also interesting to find that the dependence coefficients including Spearman's rho, Kendall's tau and upper tail dependence coefficients are in inverse proportion to monthly streamflow values in the Kaidu River Basin. In detail, the streamflows in April, May, June, August and September are higher than those in other months. But the dependence coefficients between March and April, April and May, May and June, June and August, and August and September are lower than other monthly streamflow pairs. Among all the streamflow pairs, the dependence between January and February is most significant with the values of Spearman's rho, Kendall's tau and upper tail dependence coefficient being 0.931, 0.784, and 0.838, respectively. It leads to the results that the streamflow in February depends significantly on the streamflow in January and the PDF of streamflow in February varies greatly along with the changes of streamflow status in January. The minimum values of Spearman's rho, Kendall's tau and upper tail dependence coefficient are 0.292, 0.180 and 0.235, respectively. They reveal the interaction of

monthly streamflows between April and May. Obviously,  $f_{5|4}$  indicates the small change of PDF for streamflow in May when the scaled streamflow of April changes from 0.95 to 0.99. These results show that: (a) in general, high magnitudes of two adjacent streamflow corresponds to a low dependence coefficient among them, and vice versa; (b) the two minimum adjacent streamflows exhibit the highest dependence, but the two maximum streamflows do not show the minimum dependent coefficient, which indicates the interaction between streamflow and dependence coefficient is non-linear. That is mainly because (a) there are a variety of factors which affect the streamflow in the Kaidu River Basin, such as rainfall, snow melting, evapotranspiration rate, requirement of water use, and so on; (b) the conditions of streamflow are more complex under flooding season in the Kaidu River Basin, which leads to the situation that months with high streamflow values under flooding season have weak correlations with their adjacent months. In addition, the dependence coefficient between April and May is minimum mainly due to the types of streamflow contributors being different in April and May. It has been concluded before that snow melting is the main source in April in the Kaidu River Basin. However, both rainfall and snow melting make major contributions to the runoff in May. Otherwise, the conditions that the evapotranspiration rate increases with increasing temperature and the requirement of water increases due to the growth of the plants and the impact of human use may also affect the interaction of streamflow pairs between April and May.



**Figure 5.** Distribution of streamflow in each month given the streamflow status (0.95, 0.975, 0.99) in the previous month. The values are calculated monthly from 1957 to 2009. The abscissa represents streamflow values, which are scaled into [0,1].  $f_{K+1|K}$  indicates the conditional PDF of two adjacent monthly streamflows, K and K+1. 0.95, 0.975 and 0.99 mean the steamflow status of K's streamflow.

## 5. Conclusions

In this study, a maximum entropy-Archimedean copula-based Bayesian network (MECBN) method has been developed in order to analyze the interaction of monthly streamflows. MECBN is effective for representing the uncertainties existing in the model representation, preserving the distributional characteristics of streamflow records and addressing the correlation structure between streamflow pairs. This study is the first attempt at introducing a maximum entropy-copula into the Bayesian network modeling framework. Compared to the conventional Bayesian network approach, MECBN is more effective for addressing the correlation structure between streamflow pairs with arbitrary random characteristics. Moreover, the marginal distribution of monthly streamflow is approximated by the maximum entropy method. Such a method could preserve key statistics, and does not require specific assumptions (e.g., normal, gamma etc.) on the historical streamflow data. The limitations of MECBN is that the selection of copula functions needs more computation to ensure the validity of preserving the non-linear dependence structure. Moreover, the determination of parameters in the MECBN method is computationally cumbersome.

The proposed method has been applied to frequency analysis of monthly streamflow in Kaidu River Basin. The results indicate that (a) MECBN can preserve the key statistical characteristics (i.e., mean, standard deviation, skewness and kurtosis) for each monthly streamflow in the Kaidu River Basin; (b) the Gumbel–Hougaard copula among Archimedean copulas is the most appropriate for reflecting the interactions of monthly streamflow pairs, with the minimum values of AIC and RMSE; (c) the joint distributions, the conditional distributions and the dependence measures including Spearman’s Rho, Kendall’s Tau and the tail dependence coefficient reveal the upper tail dependence between monthly streamflow pairs in the Kaidu River Basin; (d) the goodness-of-fit statistical tests show a good performance of MECBN in describing the historical data of the Kaidu River Basin, China.

The frequency analysis of monthly streamflow in the Kaidu River Basin indicates that the interactions between two adjacent monthly streamflow pairs are non-linear. Compared with the other months, there is a low rainfall but a high streamflow in April due to the fact that snow melting is the main source for the streamflow in April in the Kaidu River Basin. The upper tail dependence between monthly streamflow pairs in the Kaidu River Basin indicates that streamflow of one month depends on the streamflow of its adjacent month. It also denotes a high probability for one monthly streamflow exceeding a high threshold under the situation that its adjacent monthly streamflow exceeds the threshold. The upper tail dependence varies from different streamflow status (95%, 97.5% and 99% quantiles of streamflow) and is described by the upper tail dependence coefficients. The dependence coefficients of Spearman’s rho, Kendall’s tau and the upper tail dependence coefficient are in inverse proportion to monthly streamflow values in the Kaidu River Basin. This is because a variety of factors (i.e., rainfall, snow melting, evapotranspiration rate and requirement of water use) would affect the streamflow in the Kaidu River Basin in the flooding season, leading to less impact from the streamflow in the antecedent month. These findings have great significance for the Kaidu River Basin, which is located in the arid region of China. They can be used for providing vital information for the prevention and control of hydrological extremes and further water resources planning in the Kaidu River Basin.

**Author Contributions:** Methodology and Writing–Original Draft, X.K.; Software and Modelling calculation, X.Z. and C.C.; Formal Analysis and Validation, Y.F.; Resources and Data Curation, G.H. and Y.L.; Writing–Review & Editing, C.W.

**Funding:** This research was funded by [the Training Programme Foundation for the Beijing Municipal Excellent Talents] grant number [2017000020124G179], [National Sciences Foundation] grant number [51520105013, 51679087], [National Natural Science Foundation of China] grant number [41701621], [the 111 Program] grant number [B14008].

**Acknowledgments:** The authors are grateful to the editors and the anonymous reviewers for their insightful comments and suggestions.

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

## References

1. Hu, Q.; Huang, G.H.; Liu, Z.F.; Fan, Y.R.; Li, W. Inexact fuzzy two-stage programming for water resources management in an environment of fuzziness and randomness. *Stoch. Environ. Res. Risk Assess.* **2012**, *26*, 261–280. [[CrossRef](#)]
2. Huang, G.H.; Cohen, S.J.; Yin, Y.Y.; Bass, B. Incorporation of inexact dynamic optimization with fuzzy relation analysis for integrated climate change impact study. *J. Environ. Manag.* **1996**, *48*, 45–68. [[CrossRef](#)]
3. Cheng, G.H.; Huang, G.H.; Dong, C.; Baetz, B.W.; Li, Y.P. Interval recourse linear programming for resources and environmental systems management under uncertainty. *J. Environ. Inform.* **2017**, *30*, 119–136. [[CrossRef](#)]
4. Huang, Y.; Qin, X.S. A pseudospectral collocation approach for flood inundation modelling with random input fields. *J. Environ. Inform.* **2017**, *30*, 95–106. [[CrossRef](#)]
5. Li, Y.P.; Huang, G.H.; Nie, S.L.; Liu, L. Inexact multistage stochastic integer programming for water resources management under uncertainty. *J. Environ. Manag.* **2008**, *88*, 93–107. [[CrossRef](#)] [[PubMed](#)]
6. Fan, Y.R.; Huang, G.H.; Baetz, B.W.; Li, Y.P.; Huang, K. Development of a copula-based particle filter (CopPF) approach for hydrologic data assimilation under consideration of parameter interdependence. *Water Resour. Res.* **2017**, *53*, 4850–4875. [[CrossRef](#)]
7. Fan, Y.R.; Huang, W.W.; Huang, G.H.; Huang, K.; Zhou, X. A PCM-based stochastic hydrological model for uncertainty quantification in watershed systems. *Stoch. Environ. Res. Risk Assess.* **2015**, *29*, 915–927. [[CrossRef](#)]
8. Chen, B.; Li, P.; Wu, H.J.; Husain, T.; Khan, F. MCFP: A monte carlo simulation-based fuzzy programming approach for optimization under dual uncertainties of possibility and continuous probability. *J. Environ. Inform.* **2017**, *29*, 88–97. [[CrossRef](#)]
9. Han, J.C.; Huang, G.H.; Zhang, H.; Li, Z.; Li, Y.P. Bayesian uncertainty analysis in hydrological modeling associated with watershed subdivision level: A case study of SLURP model applied to the Xiangxi River watershed, China. *Stoch. Environ. Res. Risk Assess.* **2014**, *28*, 973–989. [[CrossRef](#)]
10. Pastori, M.; Udías, A.; Bouraoui, F.; Bidoglio, G. A multi-objective approach to evaluate the economic and environmental impacts of alternative water and nutrient management strategies in Africa. *J. Environ. Inform.* **2017**, *29*, 16–28. [[CrossRef](#)]
11. Jordan, Y.C.; Ghulam, A.; Chu, M.L. Assessing the impacts of future urban development patterns and climate changes on total suspended sediment loading in surface waters using Geoinformatics. *J. Environ. Inform.* **2014**, *24*, 65–79. [[CrossRef](#)]
12. Li, Z.; Huang, G.H.; Fan, Y.R.; Xu, J.L. Hydrologic Risk Analysis for Nonstationary Streamflow Records under Uncertainty. *J. Environ. Inform.* **2016**, *26*, 41–51.
13. Kong, X.M.; Huang, G.H.; Li, Y.P.; Fan, Y.R.; Zeng, X.T.; Zhu, Y. Inexact copula-based stochastic programming method for water resources management under multiple uncertainties. *J. Water Resour. Plan. Manag.* **2018**, *144*, 04018069. [[CrossRef](#)]
14. Fan, Y.R.; Huang, G.H.; Zhang, Y.; Li, Y.P. Uncertainty quantification for multivariate eco-hydrological risk in the Xiangxi River within the Three Gorges Reservoir Area in China. *Engineering* **2018**, *4*, 617–626. [[CrossRef](#)]
15. Fan, Y.R.; Huang, W.W.; Huang, G.H.; Huang, K.; Li, Y.P.; Kong, X.M. Bivariate hydrologic risk analysis based on a coupled entropy-copula method for the Xiangxi River in the Three Gorges Reservoir area, China. *Theor. Appl. Climatol.* **2016**, *125*, 381–397. [[CrossRef](#)]
16. 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](#)]
17. Asztalos, J.R.; Kim, Y. Lab-Scale Experiment and model study on enhanced digestion of wastewater sludge using bioelectrochemical systems. *J. Environ. Inform.* **2017**, *29*, 98–109. [[CrossRef](#)]
18. Kong, X.M.; Huang, G.H.; Fan, Y.R.; Li, Y.P.; Zeng, X.T.; Zhu, Y. Risk analysis for water resources management under dual uncertainties through factorial analysis and fuzzy random value-at-risk. *Stoch. Environ. Res. Risk Assess.* **2017**, *31*, 1–16. [[CrossRef](#)]
19. Kong, X.M.; Huang, G.H.; Fan, Y.R.; Li, Y.P. A duality theorem-based algorithm for inexact quadratic programming problems: Application to waste management under uncertainty. *Eng. Optim.* **2016**, *48*, 562–581. [[CrossRef](#)]
20. Lima, C.H.R.; Lall, U. Spatial scaling in a changing climate: A hierarchical Bayesian model for non-stationary multi-site annual maximum and monthly streamflow. *J. Hydrol.* **2010**, *383*, 307–318. [[CrossRef](#)]

21. Erro, J.; López, J.J. Regional frequency analysis of annual maximum streamflow in Gipuzkoa (Spain). *Geophys. Res. Abstr.* **2012**, *14*, 8274.
22. Schnier, S.; Cai, X.M. Prediction of regional streamflow frequency using model tree ensembles. *J. Hydrol.* **2014**, *517*, 298–309. [[CrossRef](#)]
23. Zhang, Q.; Gu, X.H.; Singh, V.P.; Xiao, M.Z.; Xu, C.Y. Flood frequency under the influence of trends in the Pearl River basin, China: Changing patterns, causes and implications. *Hydrol. Process.* **2015**, *29*, 1406–1417. [[CrossRef](#)]
24. Chan, T.U.; Hart, B.T.; Kennard, M.J.; Pusey, B.J.; Shenton, W.; Douglas, M.M.; Valentine, E.; Patel, S. Bayesian network models for environmental flow decision making in the Daly River, Northern Territory, Australia. *River Res. Appl.* **2012**, *28*, 283–301. [[CrossRef](#)]
25. Nagarajan, K.; Krekeler, C.; Slatton, K.C.; Graham, W.D. A scalable approach to fusing spatiotemporal data to estimate streamflow via a Bayesian network. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 3720–3732. [[CrossRef](#)]
26. Mediero, L.; Santillán, D.; Garrote, L. Flood quantile estimation at ungauged sites by Bayesian networks. *Geophys. Res. Abstr.* **2012**, *14*, 11998.
27. Zhang, D.; Yan, X.P.; Yang, Z.L.; Wall, A.; Wang, J. Incorporation of formal safety assessment and Bayesian network in navigational risk estimation of the Yangtze River. *Reliab. Eng. Syst. Saf.* **2013**, *118*, 93–105. [[CrossRef](#)]
28. D'Addabbo, A.; Refice, A.; Pasquariello, G. A Bayesian network approach to perform SAR/InSAR data fusion in a flood detection problem. *Proc. SPIE* **2014**, *9224*, 9244.
29. Madadgar, S.; Moradkhani, H. A Bayesian framework for probabilistic seasonal drought forecasting. *J. Hydrometeorol.* **2013**, *14*, 1685. [[CrossRef](#)]
30. Madadgar, S.; Moradkhani, H. Spatio-temporal drought forecasting within Bayesian networks. *J. Hydrol.* **2014**, *512*, 134–146. [[CrossRef](#)]
31. Kong, X.M.; Huang, G.H.; Fan, Y.R.; Li, Y.P. Maximum entropy-Gumbel-Hougaard copula method for simulation of monthly streamflow in Xiangxi River, China. *Stoch. Environ. Res. Risk Assess.* **2015**, *29*, 833–846. [[CrossRef](#)]
32. Hao, Z.; Singh, V.P. Entropy-based parameter estimation for extended three-parameter Burr III distribution for low-flow frequency analysis. *Trans. ASABE* **2009**, *52*, 1193–1202. [[CrossRef](#)]
33. Genest, C.; Favre, A.C. Everything you always wanted to know about copula modeling but were afraid to ask. *J. Hydrol. Eng.* **2007**, *12*, 347–368. [[CrossRef](#)]
34. Genest, C.; MacKay, J. The joy of copulas: Bivariate distributions with uniform marginal (Com: 87V41 P248). *Am. Stat.* **1986**, *40*, 280–283.
35. Nelsen, R.B. *An Introduction to Copulas*; Springer: New York, NY, USA, 1999.
36. Wang, C.X.; Li, Y.P.; Zhang, J.L.; Huang, G.H. Assessing parameter uncertainty in semi-distributed hydrological model based on type-2 fuzzy analysis—A case study of Kaidu River. *Hydrol. Res.* **2015**. [[CrossRef](#)]
37. Huang, Y.; Chen, X.; Li, Y.P.; Huang, G.H.; Liu, T. A fuzzy-based simulation method for modelling hydrological processes under uncertainty. *Hydrol. Process.* **2010**, *24*, 3718–3732. [[CrossRef](#)]
38. Fu, A.H.; Chen, Y.N.; Li, W.H.; Li, B.F.; Yang, Y.H.; Zhang, S.H. Spatial and temporal patterns of climate variations in the Kaidu River Basin of Xinjiang, Northwest China. *Quat. Int.* **2013**, *311*, 117–122. [[CrossRef](#)]
39. Poulin, A.; Huard, D.; Favre, A.C.; Pugin, S. Importance of tail dependence in bivariate frequency analysis. *J. Hydrol. Eng.* **2007**, *12*, 394–403. [[CrossRef](#)]
40. Willmott, C.J.; Matsuura, K. Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79–82. [[CrossRef](#)]
41. Massey, F.J. The Kolmogorov-Smirnov test for goodness of fit. *J. Am. Stat. Assoc.* **1951**, *46*, 68–78. [[CrossRef](#)]
42. Razali, N.M.; Wah, Y.B. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *J. Stat. Model. Anal.* **2011**, *2*, 21–33.

