AMSE JOURNALS-2016-Series: Advances A; Vol. 53; N°2; pp 108-122
Submitted July 2016; Revised Oct. 15, 2016, Accepted Dec. 10, 2016

Analysis of the Relationship between Precipitation and Runoff based on Smoothing-Window-Based Dependence Structure Entropy

*Ru Hao, ** Xin Chang

* School of Technical secondary, Wulanchabu Vocational College; P.R. China
** Department of Mathematics, Jining Normal University, P.R. China
hr-831020@163.com

Abstract

The Copula function is used to construct multivariate distributions to the advantage of separating the effects of dependence from the effects of marginal distributions of correlated random variables. To analyze the change of bivariate structure, the paper proposed the bivariate analysis method based on smoothing-window-based dependence structure entropy by combining smoothing windows with dependence structure entropy. The measurement data of monthly precipitation time series and monthly runoff time series at Huaxian in Weihe river basin from 1951 to 2010 was taken as an example, according to which the dependence structure entropy in different smoothing windows was computed. The result shows that the year of 1972 witnesses the mutation of dependence structure of precipitation and runoff at Huaxian, and that precipitation and runoff obeys the bivariate T-Copula function distribution during the period 1951-1972, and the bivariate Gumbel-Copula function distribution during the period 1973-2010. With many advantages, the Copula-function-based bivariate simulation is conducive to further understanding of the dependence between precipitation and runoff, and provides a new perspective of analyzing the relationships between precipitation and runoff.

Keywords

Multivariate random variable, Copula function, Structure entropy, Smoothing window, Relationship between precipitation and runoff

1. Introduction

There are three frequently-used methods for multivariate analysis, namely normal distribution, jointly specific marginal distribution, and non-parameter approach. Despite the effectiveness in certain practical analysis, they still have some problems during research. For instance, in terms of normal distribution, normal conversion is required for massive data of non-normal distribution, during which data lose is inclined to happen; while jointly specified marginal distribution is largely effective if and only if there is one category of marginal distribution for targeted variables, which is difficult to meet in practical cases; despite a favorable reflection of measurement data, the multivariate joint distribution constructed by non-parameter approach cannot reveal the category of corresponding marginal distributions. In addition, the above methods fail to describe variables of negative correlation. To address these issues, researchers at home and abroad introduce into the multivariate domain the theory and method of copula function with flexible, varied forms and simple solution [1, 2]. With copula function, multivariate joint distribution can be constructed through marginal distribution and dependence [3].

The Copula function can reflect dependence structure of random variables by separating the effects of dependence from the effects of marginal distributions of correlated random variables. Specifically, joint distribution can be divided into dependence structure and marginal distribution for independent treatment, where the former one is described by copula function [4]. The advantage of the application of copula function is that marginal distributions of arbitrary types can be constructed into joint distribution, and that no data distortion or data loss occurs since the marginal distribution contains all univariate data [5]. By describing dependence structure of variables, copula function has been widely applied to the fields of hydrology, economic calculation, and data mining [6]. The paper first expounded the concepts of copula function and entropy. Then, the idea of dependence structure entropy was proposed by means of combining copula function with entropy. Next, the technology of smoothing window was introduced to dependence structure, whose production was finally applied to research into the dependence structure of precipitation and runoff, with the intention of providing a new way for analysis of dependence between precipitation and runoff.

2. Copula function

2.1 The theory of Copula

The theory of Copula can be traced back to Sklar's theorem. In 1999, Nelsen gave a precise definition of Copulas [7], whose fundamental thought was that a copula was a joint distribution function which joined or coupled N random variables to their uniform one-dimensional margins. The 2-dimensional random variables can be expressed as:

Let X and Y be random variables with marginal distribution functions $F_1(x)$ and $F_2(y)$, respectively, and joint distribution function $F_{X,Y}(x,y)$ that corresponds to the density function of $f_{X,Y}(x,y), f_1(x), f_2(y)$. Then there exists a copula C which satisfies

$$F_{X,Y}(x,y) = C(F_1(x), F_2(y))$$

(1)

$C(F_1(x), F_2(y))$ is a dependent distribution function, where C: $[0,1]^2 \rightarrow [0,1]$. Copulas can expand to n-dimensional (n > 2) joint distributions [8].

2.2 Commonly-used copula function

The commonly-used copula functions include Gauss-Copula, T-Copula and Archimedean Copula [9].

The bivariate Gauss-Copula distribution function is

$$C^{Ga}(u,v;\rho) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\right] dsdt$$

(2)

Whose density function is

$$c^{Ga}(u,v,\rho) = \frac{1}{\sqrt{1-\rho^2}} \exp\left[-\frac{\Phi^{-1}(u)^2 + \Phi^{-1}(v)^2 - 2\rho\Phi^{-1}(u)\Phi^{-1}(vu)}{2(1-\rho^2)}\right] \exp\left[-\frac{\Phi^{-1}(u)^2\Phi^{-1}(v)^2}{2}\right]$$

(3)

Where ρ denotes the linear dependence coefficient between variables, and Φ^{-1} represents the inverse of standard Gaussian distribution. The density function of bivariate Gauss-Copula has symmetric tails with tail dependence parameter about zero. It means that the asymmetric properties of tail dependence between random variables cannot be captured.

The bivariate T-Copula distribution function is

$$C^t(u,v;\rho,k) = \int_{-\infty}^{t_k^{-1}(u)} \int_{-\infty}^{t_k^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \left[1 + \frac{s^2 - 2\rho st + t^2}{k(1-\rho^2)}\right]^{-(k+2)/2} dsdt$$

(4)

Whose density function is

$$c^t(u,v;\rho,k)=\rho^{-\frac{1}{2}}\frac{\Gamma\left(\frac{k+2}{2}\right)\Gamma\left(\frac{k}{2}\right)\left[1+\frac{u^2+v^2-2\rho uv}{k(1-\rho^2)}\right]^{-\frac{k+2}{2}}}{\left[\Gamma\left(\frac{k+1}{2}\right)\right]^2\left(1+\frac{u^2}{k}\right)^{-\frac{k+2}{2}}\left(1+\frac{v^2}{k}\right)^{-\frac{k+2}{2}}}$$

(5)

Where ρ denotes the linear dependence coefficient between variables, k is the degree of freedom, and t^{-1} represents the inverse of T-distribution. The bivariate T- Copula density function has relatively heavy tails such that it is sensitive to changes of tail dependence between random variables. Thus the density function can better capture the asymmetric properties of tail dependence between random variables.

As families of Archimedean Copulas, Gumbel-Copula, Clayton-Copula and Frank-Copula have a broad range of application, because they are suitable for variables with either positive correlation or negative correlation. The corresponding three categories of distribution function are

$$C^{Gu}(u,v;\theta)=\exp\{-[(-\ln u)^{\theta}+(-\ln v)^{\theta}]^{\frac{1}{\theta}}\},\theta\in[1,\infty)$$

(6)

Whose density function is

$$c^{Gu}(u,v;\theta)=\frac{C^{Gu}(u,v;\theta)(\ln u\times\ln v)^{\frac{1}{\theta}-1}}{uv\left[(-\ln u)^{\frac{1}{\theta}}+(-\ln u)^{\frac{1}{\theta}}\right]^{2-\theta}}\left\{\left[(-\ln u)^{\frac{1}{\theta}}+(-\ln u)^{\frac{1}{\theta}}\right]^{\theta}+\frac{1}{\theta}-1\right\}$$

(7)

$$C^{Cl}(u,v;\theta)=\max[(u^{-\theta}+v^{-\theta}-1)^{\frac{1}{\theta}}],\theta\in[-1,\infty)$$

(8)

Whose density function is

$$c^{Cl}(u,v;\theta)=(1+\theta)(uv)^{-\theta-1}\left(u^{-\theta}+v^{-\theta}-1\right)^{-2-\frac{1}{\theta}},\theta\in[-1,\infty)$$

(9)

$$C^{Fr}(u,v;\theta)=-\frac{1}{\theta}\ln[1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{(e^{-\theta}-1)}],\theta\in(-\infty,\infty)\setminus\{0\}$$

(10)

Whose density function is

$$c^{Fr}(u,v;\theta)=\frac{-\theta(e^{-\theta}-1)e^{-\theta(u+v)}}{\left[\left(e^{-\theta}-1\right)+\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)\right]^2},\theta\in(-\infty,\infty)\setminus\{0\}$$

(11)

There are no symmetric tails for the density function of bivariate Gumbel-Copula and for that of bivariate Clayton-Copula, thus they are able to capture the asymmetric properties of tail dependence between random variables. With a high upper tail and a less high lower tail, the density function of Gumbel-Copula is sensitive to changes of the upper tail, and hence is applicable to conditions with upper tail dependence and lower tail approximation independence. As a contrast, the density function of the bivariate Clayton-Copula that is characterized by a high lower tail and a less high upper tail is sensitive to changes of the lower tail, and hence is applicable to conditions with lower tail dependence and upper tail independence approximation. The measure of the tail dependence coefficient for the bivariate Frank-Copula equals zero, which means that the distributed tail variables are approximated and independent between each other, and that it fits the conditions with upper tail dependence and lower tail independence approximation [10,11].

2.3 Parameter estimation of copula function

Maximum likelihood estimation is one of the most frequently-used approaches of parameter estimation [12]. If considered generally, let the marginal distribution functions of random variables X and Y be $F_1(x,\theta_1)$ and $F_2(y,\theta_2)$, respectively. Then the corresponding marginal density functions are respective $f_1(x,\theta_1)$ and $f_2(y,\theta_2)$, where θ_1 and θ_2 are unknown parameters for marginal distribution. Let copula distribution function be $C(u,v;\alpha)$ and its density function be $c(u,v;\alpha)=\frac{\partial^2 C(u,v;\alpha)}{\partial u \partial v}$, where α is an unknown parameter in the copula distribution function. Then the joint distribution function for X and Y is

$$F(x,y;\theta_1,\theta_2,\alpha)=C[F_1(x,\theta_1),F_2(y,\theta_2);\alpha]$$

(12)

The density function corresponding to X and Y is

$$f(x,y;\theta_1,\theta_2,\alpha)=\frac{\partial^2 F}{\partial x \partial y}=c[F_1(x,\theta_1),F_2(y,\theta_2);\alpha]f_1(x,\theta_1)f_2(y,\theta_2)$$

(13)

Then, the likelihood function of the sample $(X_i,Y_i),\ (i=1,2,\cdots,n)$ is

$$L(\theta_1,\theta_2,\alpha)=\prod_{i=1}^{n} f(x,y;\theta_1,\theta_2,\alpha)=\prod_{i=1}^{n} c[F_1(x,\theta_1),F_2(y,\theta_2);\alpha]f_1(x,\theta_1)f_2(y,\theta_2)$$

The log-likelihood function is

$$\ln L(\theta_1,\theta_2,\alpha)=\sum_{i=1}^{n}\ln c[F_1(x,\theta_1),F_2(y,\theta_2);\alpha]+\sum_{i=1}^{n}\ln f_1(x,\theta_1)+\sum_{i=1}^{n}\ln f_2(y,\theta_2)$$

The maximum likelihood estimation of all parameters can be obtained by taking the derivative of the above log-likelihood function.

$$\hat{\theta}_1 = \arg\max \sum_{i=1}^{n} \ln f_1(x_i;\theta_1)$$

$$\hat{\theta}_2 = \arg\max \sum_{i=1}^{n} \ln f_2(y_i;\theta_2)$$

$$\hat{\alpha} = \arg\max \sum_{i=1}^{n} \ln c[F_1(x_i;\hat{\theta}_1)F_2(x_i;\hat{\theta}_2);\alpha]$$

2.4 Evaluation on the Copula function

To assess the quality of the proposed model, the experience copula function (E-Copula, for short) is introduced to the research [13]. Let $(x_i, y_i)(i = 1,2,\cdots,n)$ be the sample of the overall (X,Y), and the experience distribution functions for X,Y be $F_n(x)$ and $G_n(y)$, respectively. The E-Copula of the sample is defined as

$$\hat{C}_n(u,v) = \frac{1}{n}\sum_{i=1}^{n} I_{[F_n(x_i)\le u]} I_{[G_n(y_i)\le v]},\ u,v \in [0,1]$$

(14)

Where $I_{[\bullet]}$ is the indicative function when $F_n(x_i) \le u$, $I_{[F_n(x_i)\le u]} = 1$, or else $I_{[F_n(x_i)\le u]} = 0$.

The Euclidean distance between the bivariate copula function and the E-Copula function is defined as

$$d = \sqrt{\sum_{i=1}^{n} \left| C(u_i,v_i) - \hat{C}_n(u_i,v_i) \right|^2}$$

(15)

Euclidean distance reflects the situation that the bivariate copula function fits the original data. The smaller the Euclidean distance is, the higher the model's goodness-of-fit is.

3. Information entropy

3.1 Differential entropy

The differential entropy $H(X)$ [14] is defined as

$$H_X(X) = -\int f_X(x)\log f_X(x)dx$$

(16)

Where the base of the logarithm is one of the values among 2， e and 10.

3.2 Conditional entropy

Let $f_{X,Y}(x|y)$ be the conditional probability, and the definition of conditional entropy is given by

$$H_{X|Y}(X|Y) = -\int\int f_{X,Y}(x,y)\log f_{X|Y}(x|y)dxdy$$

(17)

3.3 Joint entropy

Joint entropy measures the uncertainty that a pair of random variables X and Y occurs simultaneously, whose definition is

$$H(X,Y) = -\int\int f_{X,Y}(x,y)\log f_{X,Y}(x,y)dxdy$$

(18)

The relationship among joint entropy, conditional entropy, and differential entropy is

$$\begin{aligned} H(X,Y) &= -\int\int f_{X,Y}(x,y)\log f_{X,Y}(x,y)dxdy \\ &= -\iint f_{X,Y}(x,y)\log f_Y(y)f_{X|Y}(x|y)dxdy \\ &= -\iint f_{X,Y}(x,y)[\log f_Y(y)+\log f_{X|Y}(x|y)]dxdy \\ &= -\iint f_{X,Y}(x,y)\log f_Y(y)dxdy + \iint f_{X,Y}(x,y)\log f_{X|Y}(x|y)dxdy \\ &= -\int f_Y(y)\log f_Y(y)dy + H_{X|Y}(X|Y) \\ &= H(Y) + H_{X|Y}(X|Y) \end{aligned}$$

And similarly for $H(X,Y) = H(X) + H_{Y|X}(Y|X)$.

4. Smoothing-window-based dependence structure entropy

4.1 Dependence structure entropy

Let $u = F_1(x)$, $v = F_2(y)$, and then $F_{X,Y}(x,y) = C(F_1(x),F_2(y)) = C(u,v)$, where $c(u,v)$ is the dependence structure entropy.

The dependence structure entropy of random variables X and Y is defined as

$$H_C(U,V) = -\iint c(u,v)\log c(u,v)dudv$$

(19)

$u = F_1(x)$, $v = F_2(y)$, $du = dF_1(x) = f_1(x)dx$, $dv = dF_2(y) = f_2(y)dy$

Substitute $x = F_1^{-1}(u), y = F_2^{-1}(v)$ into the above equations, and it is obtained that

$dx = dF_1^{-1}(u) = \frac{1}{f_1(x)}du$, $dy = dF_2^{-1}(v) = \frac{1}{f_2(y)}dy$

Below is the deduction of the relationship among joint entropy, dependence structure entropy, and differential entropy.

$$
\begin{aligned}
H(X,Y) &= -\int\int f_{X,Y}(x,y)\log f_{X,Y}(x,y)dxdy \\
&= -\int_0^1\int_0^1 c(u,v)\log c(u,v)[f_1(F_1^{-1}(u))f_2(F_2^{-1}(v))]dudv \\
&= -\int_0^1\int_0^1 c(u,v)\log c(u,v)dudv - \int_0^1\int_0^1 c(u,v)\log f_1(F_1^{-1}(u))dudv - \int_0^1\int_0^1 c(u,v)\log f_2(F_2^{-1}(v))dudv \\
&= H_C(U,V) - \int_0^1[\int_0^1 c(u,v)dv]\log f_1(F_1^{-1}(u))du - \int_0^1[\int_0^1 c(u,v)du]\log f_2(F_2^{-1}(v))dv \\
&= H_C(U,V) - \iint f_{X,Y}(x,y)\log f_1(x)dxdy - \iint f_{X,Y}(x,y)\log f_2(y)dxdy \\
&= H_C(U,V) - \int f_1(x)\log f_1(x)dx - \int f_2(y)\log f_2(y)dy \\
&= H_C(U,V) + H(X) + H(Y)
\end{aligned}
$$

It has trivially shown that the joint entropy can be decomposed into and equals the sum of dependence structure entropy $H_C(U,V)$ as well as differential entropies $H(X)$ and $H(Y)$. If the random variables X and Y are completely independent from each other, then the measure of the dependence structure entropy equals zero, or $H_C(U,V)=0$. It indicates that the more independence approximation X and Y has, the more $H_C(U,V)$ approximates zero; and that the more closely the dependence between X and Y is, the larger the distance between $H_C(U,V)$ and zero is. In this way, dependence structure entropy can be used to discuss changes of dependence structures between X and Y.

4.2 Smoothing-window-based dependence structure entropy

When the dependence between a pair of random variables changes, their dependence structure varies noticeably. The paper introduces the concept of smoothing window into dependence structure entropy, and calls their combination as smoothing-window-based dependence structure entropy.

There are mainly two parameters for smoothing windows, one is window length W, and the other is smoothing step L. Below are the procedures to compute the smoothing-window-based dependence structure entropy of two-time series.

1) Select W. It should be guaranteed that the Ws for the two-time series are the same;

2) Select L. Keep W unchanged. Move the window at the distance of L in order from the first data in the time series, until all the data has been passed;

3) Compute the dependence structure entropy of both time series in their respective windows, and obtain the corresponding smoothing-window-based dependence structure entropy;

4) Draw out the sequence diagram of the dependence structure entropy. Identify preliminarily the structure change of the two time series;

5) With Euclidean distance as the evaluation criteria, determine the optimal family of copula functions before and after the change happens.

5. Case study

5.1 Data source

The dependence structures of time series data for precipitation and runoff at Xianyang hydrological station in the Weihe river basin were chosen for analysis in the paper. The Weihe River is the largest tributary of the Yellow River. Influenced by human activities and climate change, the precipitation and runoff volume in the basin witness a decreasing trend [10]. Changes perhaps have happened accordingly for the relationship of dependence structures between precipitation and runoff, whose binary distribution changes as a result. Xianyang hydrological station lies in the midstream and downstream reaches of the Weihe river basin, and is greatly affected by human activities in the upstream area. The paper conducted analysis of dependence structures of time series data for monthly precipitation and monthly runoff, respectively, at Xianyang hydrological station in the Weihe river basin, from January 1951 to December 2010. Both time series had a window length of 720. The data of runoff came from hydrological information on Yellow River from National Hydrological Yearbooks, and the meteorological data was derived from
http://www.cams.cma.gov.cn/cams_kxsy/qky_kxsy_index.htm.

5.2 The smoothing-window-based dependence structure entropy of Huaxian

During data analysis based on smoothing windows, the key to success is to find out effective window length and smoothing steps. Changes of dependence structures between two random variables cannot be observed in overlarge or undersized windows. To determine the effective window length, one in practical operation usually selects first a window of a relatively small (large) length, scales up (down) the length such that several window lengths in proportion can be used for analysis, and conducts dependence structure analysis on these windows. Interested in changes of the dependence structures of annual precipitation and annual runoff, respectively, we determine the smoothing step L for analysis as 12, and the window lengths as 12, 24, 60,120,240, respectively (Figure 1).

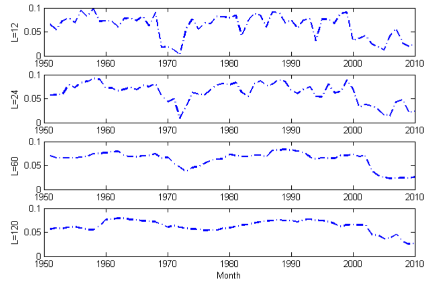

Fig. 1. The dependence structure entropy between precipitation and runoff at Huaxian (Ws (from top to bottom) are 12, 24, 60,120,240, respectively)

As can be seen from the above figure, when Ws are 12, 24, 60,120,240, respectively, there are two phases for changes of dependence structures of time series for precipitation and runoff. Specifically, the period from 1951 to 1972 sees a trend of slow increase first and then decrease; the curve hits the bottom in the year of 1972 as a turning point, and then fluctuates slowly. The changes of dependence structures of time series for precipitation and runoff in different windows are consistent. What's more, under the effect of human activities, the dependence between time series for precipitation and that for runoff tends to drop gradually. [16, 17] show that the year of 1972 marks the abrupt change of the relationship between precipitation and runoff in the Weihe river basin.

Table 1 is the parameters of bivariate Copula function for precipitation and runoff during the research period of 1951-2010.

Table 1. The parameters of Copula function

	1751.1-1972.12		1973.1-2010.12	
Gauss	ρ =0.8202	d =0.1641	ρ =0.6116	d =0.1531
T	ρ =0.8202 k =4	d =0.1042	ρ =0.6706 k =7	d =0.1652
Gumbel	θ=1.9283	d =0.1315	θ=1.7156	d =0.1153
Clayton	θ=1.7251	d =0.1129	θ=1.5931	d =0.1762
Frank	θ=5.7956	d =0.1273	θ=4.8158	d =0.1531
Optimal Copula	T-Copula		Gumbel -Copula	

Under the criteria of Euclidean distance, T-Copula fits the observation data at the first stage (1951-1971) best, while Gumbel–Copula has the optimal goodness-of-fit for the observation data at the second stage (1972-2010). In other words, the family of bivariate density function begins to change from T-Copula to Gumbel–Copula in 1972. In terms of the dependence structure entropy and the optimal copula distribution, the dependence structure of precipitation and runoff mutates in 1972. Figure 2-7 show the optimal density function, distribution function, and contour map at different stages.

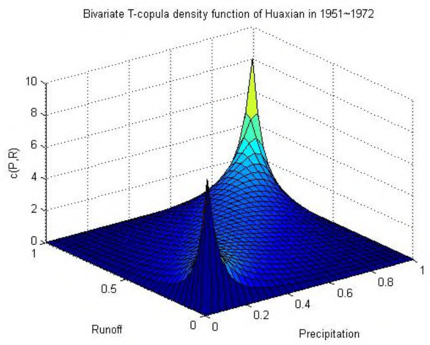

Fig. 2. Bivariate T-copula density function of Huaxian in 1951~1972

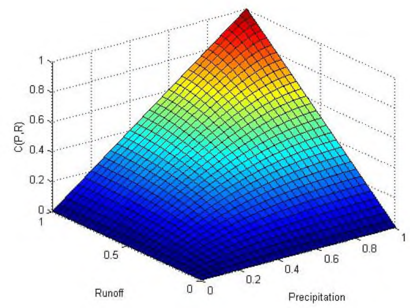

Fig. 3. Bivariate T-copula distribution function of Huaxian in 1951~1972

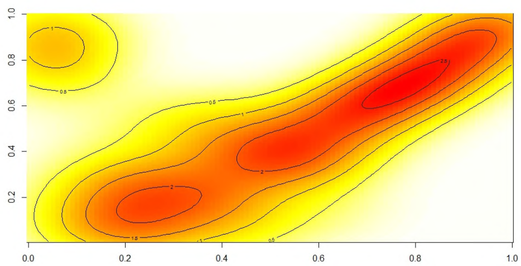

Fig. 4. Contour map of Huaxian in 1951~1972

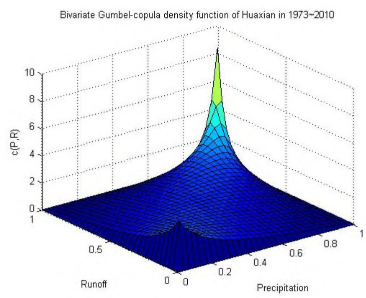

Fig. 5. Bivariate Gumbel-copula density function of Huaxian in 1973~2010

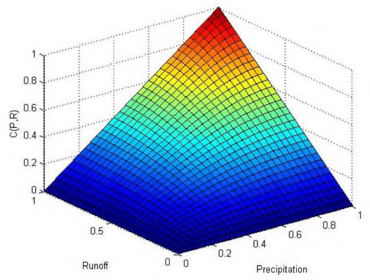

Fig. 6. Bivariate Gumbel-copula distribution function of Huaxian in 1973~2010

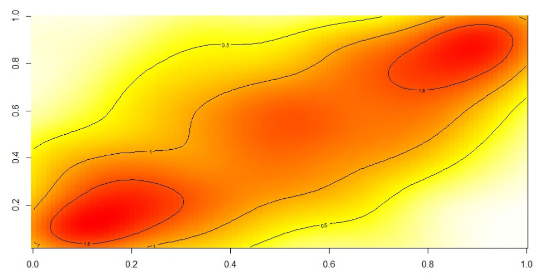

Fig. 7. Contour map of Huaxian in 1973~2010

5.3 Discussion

The density function for precipitation and runoff during the period of 1951-1971 has relatively heavy tails, which means that there is large mutual influence between precipitation and runoff. The high upper tail and low lower tail shown in the density function diagram indicates that the upper tail dependence overmatches lower tail dependence, namely that the interaction between maximum values of precipitation and those of runoff surpasses the interaction between minimum values of precipitation and those of runoff. The density function for precipitation and runoff during the period of 1972-2010 has a high upper tail and an approximating and independent lower tail. This means that there are mutual effects between maximum values of precipitation and those of runoff, and that minimum values of precipitation and those of runoff are basically independence from each other.

There may be two causes to the different distributions between precipitation and runoff at Huaxian from 1951 to 2010: 1. Climate change. The precipitation amount in Weihe river basin dropped remarkably in 1972 due to El Nino, leading to a great turn of the local precipitation; 2. Human activities. There are large-scale development of reservoirs and irrigation engineering in the Weihe river basin during the 1970-1980 period. For example, the main canal on the uplands that ring Baoji Gorge was established in 1971, and the year of 1970 witnessed the implementation of the Dongfanghong Chouwei Irrigation Project. As a result of the above hydraulic engineering, the runoff production rate declined in the area, which further reduced the actual runoff in 1972. All in all, the climate change as a natural cause and the human activities as a man-made reason together mutate the dependence structure of precipitation and runoff at Huaxian in Weihe river basin in 1972.

6. Conclusion

By introducing the idea of smoothing-window-based dependence structure entropy to the paper, the dependence structure of precipitation and runoff at Huaxian in the Weihe river basin was analyzed. The result shows that the dependence structure of precipitation and runoff mutates in 1972. The period from 1951 to 1972 sees a trend of slow increase first and then decrease; while the period from 1972 to 2010 witnesses another tendency of slow increase first and then decrease. In different smoothing windows with the fixed smoothing step of 12, the dependence between precipitation and runoff at Huaxian is enlarged first, then shrinks, then intensifies again, and finally declines. This phenomenon indicates that under the influence of climate change and human activities, the dependence between precipitation and runoff is weakened. As a crucial link of water circulation, precipitation and runoff in the Weihe river watershed is affected by climate change (temperature, evaporation, for example) and human activities (such as underlying surface change, and hydraulic engineering), and is characterized by high intricacy, non-steadiness, and non-linearity. Research into the dependence structure of precipitation and runoff is conducive to further understanding of the evolution law of time series for precipitation and runoff. It is also of great significance to efficient water usage, positive function and safe operation of hydraulic engineering, as well as social stability and progression.

The concept of copula function and dependence structure entropy was brought into the paper. Compared to existing multivariate modeling, copula function can divide the bivariate joint distribution into dependence structure and marginal distribution for independent processing, and is free from limitations of variate distribution families as well. Thus, there is no data loss caused by the measure of assuming distribution types for this approach, and non-normal, asymmetric distribution data can be captured according to tail dependence parameters. In light of entropy, this simple, convenient, and noise-resistant nonlinear theory can be combined with the copula theory in a way that calculating the change of dependence structure of two random variables, so that this combination offsets the disadvantage of analysis by conventional linear methods. The proposed smoothing-window-based dependence structure entropy in the paper help reveal the tendency of precipitation time series and runoff time series, and also provides theoretical basis for planning and managing hydraulic engineering.

References

1. X. Chen, Y. Fan, Estimation of copula-based semiparametric time series models, 2006, Journal of Econometrics, 130(2), 307-335.
2. F. Abegaz, U.V. Naik-Nimbalkar, Modeling statistical dependence of Markov chains via copula models, 2008, Journal of Statistical Planning and Inference, 138(4), 1131-1146.
3. L. Zhang, V. P. Singh. Bivariate rainfall frequency distributions using Archimedean copulas, 2007, Journal of Hydrology, 332, 93-109.
4. J. T. Shiau, F. Song, S. Nadarajah, Assessment of hydrological droughts for the Yellow River, China, using copulas, 2007, Hydrological Processes, 21, 2157–2163.
5. A.C. Favre, S.E. Adlouni, L. Perrault, N. Thiémonge, B. Bobée, Multivariate hydrological frequency analysis using Copulas, 2004, Water resources research, 40.
6. A. Meucci, A New Breed of Copulas for Risk and Portfolio Management [J], 2011, Risk 24(9), 122–126.
7. R.B Nelsen, An Introduction to Copulas, Springer Series in Statistics. Springer, 1999, New York, pp. 216.
8. P.Embrechts, F. Lindskog, A. McNeil, Modeling Dependence with Copulas and Applications to Risk Management Handbook of Heavy Tailed Distributions in Finance, 2003, Elsevier Science, 2003, Publishers, New York.
9. S.L. Guo, B.W. Yan, Y. Xiao, et al. Multivariate Hydrological Analysis and Estimation [J], 2008, Journal of China Hydrology. 28(3), 1-7.
10. K. Aas, C. Czado, H. Bakken, Pair-copula constructions of multiple dependence, 2009, Insurance: Mathematics and Economics, 44(2), 182–198.doi:10.1016/j.insmatheco.2007.02.001
11. C. Donnelly, P. Embrechts, The devil is in the tails: actuarial mathematics and the subprime mortgage crisis, 2010, ASTIN Bulletin, 40(1), 1–33.
12. D. Brigo, A. Pallavicini, R. Torresetti, Credit Models and the Crisis: A Journey into CDOs, Copulas, Correlations and dynamic Models. Wiley and Sons. Jones, Sam, April 24, 2009

13. S. Wu, Construction of asymmetric copulas and its application in bivariate reliability modelling, 2014, European Journal of Operational Research 238, 476–485. doi:10.1016/j.ejor.2014.03.016
14. C.E. Shannon, Mathematical Theory of Communication, 1948, Bell System Technical Journal, vol.27 (7), 379-423.
15. Entropy inference and the James-Stein estimator, with application to nonlinear gene association networks [J]. J. Mach. Learn. Res. 10: 1469-1484.
16. Y.L. Li, Q.X. Chang, Detection of the abrupt changes in runoff based on the Morlet wavelet, 2012, Journal of Xi'an University of Technology, 28(3), 322-325.
17. Y.L. Li, Q.X. Chang, Z.L. Wang, H.L. Li, Detection of Abrupt Changes in Runoff in the Weihe River Basin, 2016, Advances in Meteorology, pp. 1-8. http://dx.doi.org/10.1155/2016/6962082