International Journal of Computational Intelligence Systems, Vol. 7, Supplement 1 (2014), 18-28

Bivariate analysis of typical hydrological series of the yellow river

Xin Tong

Key Laboratory of Surficial Geochemistry, Ministry of Education, Department of Hydrosciences, School of Earth Sciences and Engineering, State Key Laboratory of Pollution Control and Resource Reuse, Nanjing University, Nanjing 210093, P.R. China

Email: tongxin000111@163.com

Dong Wang

Key Laboratory of Surficial Geochemistry, Ministry of Education, Department of Hydrosciences, School of Earth Sciences and Engineering, State Key Laboratory of Pollution Control and Resource Reuse, Nanjing University, Nanjing 210093, P.R. China

Corresponding author Email: wangdong@nju.edu.cn

www.nju.edu.cn

Jichun Wu

Key Laboratory of Surficial Geochemistry, Ministry of Education, Department of Hydrosciences, School of Earth Sciences and Engineering, State Key Laboratory of Pollution Control and Resource Reuse, Nanjing University, Nanjing 210093, P.R. China

Yuanfang Chen

College of Water Resources and Environment, Hohai University

1 Xikang Road, Nanjing 210098, P.R. China

Xi Chen

State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, School of Hydrology and Water Resources, Hohai University, Nanjing 210098, P.R. China

Received 14 December 2012

Accepted 1 July 2013

Abstract

This paper uses Gumbel-Hougaard (G-H) copula, Clayton copula and Frank copula to construct joint distributions of hydrological variables of the two typical stations on the Yellow River Region, including the annual maximum flood magnitude (AMFM), the annual maximum flood occurrence date (AMFOD) and the annual runoffs (ARs). The results give the joint distribution between each pair of the variables. Also an isoline of the concurrence return periods between the AMFMs of the two stations was drawn up.

Keywords: copula; von Mises; the Yellow River; flood analysis; annual runoff

Co-published by Atlantis Press and Taylor & Francis

1. Introduction

Water resource is a kind of indispensible natural resource to human society. Large scale hydrological issues have drawn great attention by researchers all over the world because it may have some impacts on the living, economy and society. Costa, M. H. *et al* [1] studied the effect of changes in land cover on the discharge of the Tocantins River (175360 km^2) in Southeastern Amazonia and found that the change of vegetable cover did disturbed the hydrological response of the region. Lin, Y. and Wei, X. H. [2]collected lone-term hydrological and climate data and large scale cumulative forest harvesting in the Willow watershed of 2860 km^2 in Canada and investigated the impact of the harvesting on hydrology.

The Yellow River is the second longest river in china and it covers large area of 752443 km2 where about 110 million residents live. The water resource of the Yellow River is 2.6% of national total water resources, ranking the fourth in the national seven longest rivers. What is more, the per capital water resource is one third of the national average level. Thus, water resource in the Yellow River region is in shortage. The runoff of the Yellow River is not uniformly distributed through a year. About 60% of the runoff of the Yellow River happens during July to October (the flood season) and runoff from March to June only covers 10%-20% of the total runoff of a year. Most of the flood during the flood season comes from the middle region of the Yellow River where two types of flood combination can lead to devastating flood of the Huayuankou station and cause flood disaster to the downstream region of the Yellow River, namely, "up large" flood (it comes from the Hekou station to Longmen station or from longmen station to Sanmenxia station) and "down large" flood (it comes from sanmenxia station to Huayuankou station). Among the two types of flood, the "down large" is more threatening to the downstream region where 70 million people live because of the high flood peak discharge it brings and the short time it needs to reach peak. On the other hand, In the climate changing environment, water resources is gradually shrinking, Flood is a kind of water resource and the measures of making flood of Yellow River useful should be considered in the case that Yellow River is in shortage. So it is important to study the possible water resource—the flood. In order to know the water resource of Sanmenxia station and Huayuankou station well, some emphasis should be drawn to analyzing the characteristics of the flood and runoff of the Sanmenxia and Huayuankou station.

Some work has been done to analyze some typical great floods between Sanmenxia and Huayuankou stations. *Zhang et al.* [3]explored the relationship between the floods of July 1958 and August 1982 in Sanmenxia –Huayuankou reach and found that the temporal and spatial distribution of the rainstorm has great impact on the relation between rainstorm and flood. *Wang et al.* [4]studied the extraordinary rainstorm and modeled the eventual flood in 1761 on the Sanmenxia-Huayuankou reach of Yellow River. This paper, however, mainly considers the joint behavior between the flood and runoff of the two stations using copula method and von Mises distribution.

Copula is a relatively new way to build joint distribution of several variables independently of the marginal distribution. It is widely used in hydrologic frequency analysis. Zhang and Singh[5] derived bivariate rainfall frequency distribution using Archimedean copulas and proved that the advantage of copula method is the variables do not need to have the same marginal distributions. Xiao *et al.* [6] employed G-H copula to build joint distribution of flood peak and flood volume, from which synthetic flood hydrographs are constructed. Kao and Govindaraju[7] used copulas to capture the joint behaviors of drought information. Guo *et al.* [8] summarized the using of copulas in multivariate hydrological analysis and prospected the future applying of the method. Lee, T. and Salas, J. D.[9] introduced copula method to stochastic streamflow simulation. Chowdhary, H. et al. [10]discussed selection procedure of copulas and demonstrated their application in the bivariate flood frequency analysis. Some researchers have accomplished some work via copulas considering the runoff and drought of the Yellow River. Shiau *et al.*[11] built bivariate droughts distributions and investigated historical noticeable droughts of the Yellow River by means of Clayton copula. Fu *et al.* [12] used G-H copula to calculate the encounter risk between the abundant and low runoffs from Weihe River and Fenhe River which are the two largest tributary of the Yellow River.

The von Mises distribution is widely used in analysis of circular statistics in medical science, statistic analysis and some other research fields. Mooney *et al.* [13] introduced mixed von Mises distribution to illustrate the distribution of sudden infant death syndrome. Catar *et al.*

[14]explored the distribution of directional wind speed using a finite mixture of von Mises distributions. Recently some researchers have done work in the joint distribution of flood date and magnitude. Fang *et al.*[15] for the first time successfully fitted the flood occurrence date using von Mises method and construct joint distribution of flood occurrence date and flood magnitude of Qingjiang River using Gumbel Archimedean copula. Yan *et al*.[16] studied the joint distribution of flood occurrence date and magnitude of Qingjiang River and Changjiang River using mixed von Mises method and Clayton copula.

On the basis of the above treatises, this paper mainly discusses the encounter risk of the annual maximum flood magnitudes (AMFMs) and annual maximum flood occurrence dates (AMFODs) between Sanmenxia station and Huayuankou station and of each station using the Archimedean copula such as G-H copula, Clayton copula and Frank copula. Also the joint distribution of the annual runoffs (ARs) between the two stations will be established. First, the joint distributions will be constructed. Then the daily encounter risk of AMFMs between the two stations can be calculated, so as the encounter risk of a specific magnitude of flood peak and a given date of each station. Further, the paper also deals with the conditional joint probability between the AMFMs of the two stations and the conditional joint probability between the AMFOD and AMFM of each station. Also the joint return periods and concurrent return periods between the AMFMs of the two stations will be studied. Finally, an isoline of encounter return periods of equivalent frequency combination is made.

2. Methodology

2.1. Von Mises distribution

The AMFOD can be regarded as a circular vector which can be demonstrated by von Mises distribution. Von Mises distribution has been used in demonstrating the distribution of sudden infant death syndrome (see Ref. 13) and the distribution of directional wind speed (see Ref. 14). Some researchers have introduced von Mises distribution into hydrological variable analysis. Fang et al. [15]and Yan *et al.*[16] used this method to analyze the distribution of flood date. The probability density function (PDF) of a variable X fitting von Mises distribution is defined as:

$$f(x;\ \ \theta) = \frac{\exp[k\cos(x - \mu)]}{2\pi I_0(k)}. \quad (1)$$

Where, $\theta = (\mu, k)^T$, $0 < x \le 2\pi, 0 < \mu \le 2\pi$ and $k \ge 0$. μ is the mean position parameter, k is the concentration parameter. $I_0(k)$ is the modified Bessel function of order zero. The flood date can be transformed as follows:

$$x_i = D_i \frac{2\pi}{T}, 0 < x_i \le 2\pi. \quad (2)$$

Where, D_i is the i th day during the calculating period. T is the length of the calculating period. Thus, the date can be transformed to a series which can be illustrated by the von Mises distribution.

2.2. Marginal distributions

P-Ⅲ distribution is widely used in flood frequency analysis and the paper prefers the method to illustrate the distribution of the AMFM and AR of Sanmenxia station Huayuankou station. The PDF of a variable Y fitting P-Ⅲ distribution is given as:

$$f(y) = \frac{\beta^\alpha}{\Gamma(\alpha)}(y - a_0)^{\alpha-1}\exp^{-\beta(y-a_0)},\ \ y > a_0. \quad (3)$$

Where, α , β and a_0 are parameters of shape, concentration and mean position, respectively. These parameters are estimated by linear-moment (LM) method.

2.3. General theory about copulas

Copulas are functions that link multivariate probability distributions as the statement by Nelsen[17] and it can capture the dependence feature between multiple different marginals (see Ref. 18). The advantage of using copula is that copula can demonstrate the joint characteristics of multiple variables independent of their marginal distributions. In terms of bivariate situations, let $F_X(x) = u$ be the cumulative density function (CDF) of X and $F_Y(y) = v$ be the CDF of Y .Then the copula connects two marginal distributions to form a bivariate probability distribution $H(x, y)$ given by:

$$H(x, y) = C_\theta(F_X(x), F_Y(y)) = C_\theta(u, v). \quad (4)$$

Where, C is the function of the corresponding copula. θ is a parameter of the copula function. Details about copulas can be found in the book of Nelsen[17]. G-H copula, Clayton copula and Frank copula are three common used families from Archimedean copulas in hydrology as stated by Genest[19]. Table1 shows the relationships between Kendall's tau and the parameter θ and the corresponding functions of each of the three copulas.

Table1. Relationship between Kendall's tau and the parameter θ and the corresponding functions of three Archimedean copulas.

Family	Kendall's tau	Function of copula
G-H	$1-1/\theta$	$\exp\left\{-[(-\ln u)^{\theta}+(-\ln v)^{\theta}]^{1/\theta}\right\}$
Clayton	$\theta/(\theta+2)$	$(u^{-\theta}+v^{-\theta}-1)^{-1/\theta}$
Frank	$1-4/\theta+4D_1(\theta)/\theta$	$\frac{1}{\theta}\ln\left[1+\frac{[\exp(\theta u)-1][\exp(\theta v)-1]}{\exp(\theta)}\right]$

Note here $D_1(.)$ is the first Debye function. t is either u or v

2.4. Way of choosing proper copula and parameters

A key step of building joint distributions is choosing the best fitted copula and the corresponding parameters. This paper introduces two ways to decide the most suitable copula. One is the canonical maximum likelihood estimator (CMLE) which is one of the semiparametric (SP) methods and proved to be better than the maximum likelihood estimator (MLE) and the inference function from margins (IFM) method in most situations as mentioned by Kim[20]. Details about the CMLE can be found in the work done by Vandenberghe [21]. The other one is estimating the parameter θ via relationship between θ and Kendall's tau shown in Table 1.

Then, in order to test the goodness of fit of the copula methods, two criterions will be applied, including the AIC (see Ref. 5) value and ordinary least square (OLS) value:

$$OLS=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(p_{ei}-p_i)^2}. \tag{5}$$

$$AIC=n\log[\frac{1}{n}\sum_{i=1}^{n}(p_{ei}-p_i)^2]+2m. \tag{6}$$

Where, n is the length of the data, m is the number of the parameters, p_i is the theoretical joint probability, p_{ei} is the empirical joint probability given as:

$$p_{ei}=p(X\le x_i,Y\le y_i)=\frac{\sum_{j=1}^{n}No.of(x_j\le x_i,y_j\le y_i)-0.44}{n+0.12}. \tag{7}$$

Where the $No.of(x_j\le x_i,y_j\le y_i)$ means the number of the pairs of variables that satisfy $X\le x_i$ and $Y\le y_i$. The best choice of θ and copula are supposed to bring with the minimum AIC and OLS values.

3. Case Study

Annual maximum flood peaks and the related flood occurrence dates of 1958-1988 (no data of 1986) from Sanmenxia station and Huayuankou station are collected, so are the annual runoff series from the two stations during the same time domain. Using the von Mises distribution and P-Ⅲ distribution discussed above, the marginal distribution parameter and corresponding testing results are given in Table 2 and Table 3. The KSP in Table 3 is K-S (Kolmogorov Smirnov Test) test value calculated by Matlab and α is significance level. If $KSP>\alpha/2$ then the theoretical distribution can be accepted. The Tables show that the Bias and RMSE values are both small and all the corresponding KSP values are acceptable, which indicates that the von Mises and P-Ⅲdistribution of the two stations fit well.

After the marginal distribution is obtained, it comes to the choice of the parameter θ and the preference of copula through the following steps:

(i) Calculate the marginal empirical probability of AMFOD, AR, and AMFM using the equation:

$$F(x_i)=p(X\le x_i)=\frac{m(i)}{n+1}. \tag{8}$$

Where, $m(i)$ is the index of m th smallest observation in the data set of a variable arranged in ascending order.

Table2. Parameters of von Mises and P-III distribution of Sanmenxia and Huayuankou station.

		Von Mises		P-III		
		μ	k	C_v	C_s	mean(cms)
Annual maximum flood	Sanmenxia station	4.029	1.7907	0.36	1.33	5534.67
	Huayuankou station	3.932	1.6821	0.45	1.87	7799.67
		μ	k	C_v	C_s	mean(10^8cm)
Annual Runoff	Sanmenxia station	—	—	0.35	1.11	403.09
	Huayuankou station	—	—	0.38	1.04	439.61

Table3. Parameters testing results of von Mises and P-III distribution of Sanmenxia and Huayuankou station.

	AMFM and AMFOD				AR	
	Sanmenxia station		Huayuankou station		Sanmenxia station	Huayuankou station
	P-III	Von Mises	P-III	Von Mises	P-III	P-III
RMSE	0.0718	3.5761	0.0901	5.2698	0.0425	0.0406
Bias	-0.0206	1.0362	0.0277	1.3847	-0.0164	-0.0124
KSP	0.9360	0.76	0.9360	0.9360	0.9970	0.9970

Note here KSP is the K-S test value and significance level $\alpha = 0.05$

(ii) Use CMLE to estimate θ ;

(iii) Calculate Kendall’s tau between the AMFOD and the AMFM of each station, between the AMFODs of the two stations, between the AMFMs of the two stations and between the ARs of the two stations respectively.

(iv) Estimate $\hat{\theta}$ via the relationship between Kendall’s tau and $\hat{\theta}$ as demonstrated in Table 1;

Using θ , $\hat{\theta}$ and the three types of copula to build joint distributions between pairs of variables and compare the results of Goodness of fit (GOF), then make a choice of θ and copula which brings the best GOF. The GOF results of different θ and copula are shown in Table 4. The best fitting situation are represented by the highlighted minimum OLS and AIC value. Thus it can be deduced that the G-H copula fits the bivariate distribution between AMFMs of the two stations and the joint distribution between the ARs of the two stations best; The Clayton copula is the best choice to construct joint distribution between AMFOD and AMFM of Sanmenxia station and the joint distribution between the AMFODs of the two stations. Frank copula performs best in building the joint distribution between the AMFOD and AMFM of Huayuankou station. Table 4 also shows that in most situations, the parameter estimated via Kendall’s tau performs better than that from CMLE, except for the joint distribution between the AMFOD and AMFM of Huayuankou station in which case the Kendall’s tau is negative.

We plot the best fit copula probabilities against the empirical probabilities to check the see the modeling effect more straightly. Comparison of empirical joint probability and theoretical joint probability are shown in Fig. 1-5 corresponding to the chosen copulas. It can be seen that all the copulas fit well with the empirical joint distributions. So we can implement the fitted copulas to analyze the characteristics of the variables of the two stations.

Table 4 GOF results of different θ and copulas.

Bivariate variables			CMLE	Via Kendall's tau
Between AMFMs of the two stations	OLS	**G-H**	0.0486	**0.0474**
		Clayton	0.0586	0.0594
		Frank	0.0540	0.0530
	AIC	**G-H**	-179.44	**-180.96**
		Clayton	-168.21	-167.45
		Frank	-173.08	-174.20
Between AMFM and AMFOD of Sanmenxia station	OLS	G-H	0.0808	0.0813
		Clayton	0.0814	**0.0804**
		Frank	0.0812	0.0809
	AIC	G-H	-148.93	-148.23
		Clayton	-148.49	**-149.23**
		Frank	-148.64	-148.89
Between AMFM and AMFOD of Huayuankou station	OLS	G-H	0.0381	—
		Clayton	0.0388	—
		Frank	**0.0353**	**0.0353**
	AIC	G-H	-194.00	—
		Clayton	-193.00	—
		Frank	**-198.60**	-198.57
between AMFODs of the two stations	OLS	G-H	0.0712	0.0719
		Clayton	0.0711	**0.0607**
		Frank	0.0649	0.0644
	AIC	G-H	-156.50	-155.90
		Clayton	-156.61	**-166.14**
		Frank	-162.07	-162.55
between ARs of the two stations	OLS	**G-H**	0.039	**0.0383**
		Clayton	0.0502	0.049
		Frank	0.0426	0.0431
	AIC	**G-H**	-192.72	**-193.69**
		Clayton	-177.50	-179.01
		Frank	-187.41	-186.72

4. Results and discussions

The paper implemented three common used Archimedean copulas, namely G-H copula, Clayton copula and Frank copula to analyze the joint behavior between several hydrological variables e.g. annual maximum flood peak and occurrence date of Sanmenxia station and Huayuankou station and compare the GOF of different copulas and estimators. Result shows that G-H copula is the most fitted model for the joint distribution between the AMFMs of the two stations and between ARs of the two stations. The Clayton copula performs best in constructing the joint distribution between AMFM and AMFOD of Sanmenxia station and the joint distribution between the AMFODs of the two stations. The Frank copula can illustrate the joint distribution between AMFM and AMFOD of Huayuankou station best. CDFs and contour isolines of the above joint distributions can be drawn up from which it is convenient to find the joint probability. Then the joint behaviors of the variables and encounter risks can be discussed.

First, the conditional joint probability can be inferred from the marginal distributions and joint distribution as follows:

$$P(X>x|Y>y)=\frac{P(X>x,Y>y)}{P(Y>y)}=\frac{1-F(x)-F(y)+H(x,y)}{1-F(y)}. \qquad (9)$$

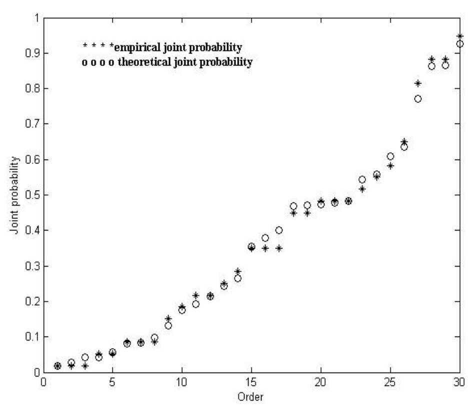

Fig. 1. Comparison of empirical and theoretical joint probability of the AMFMs of Sanmenxia and Huayuankou stations

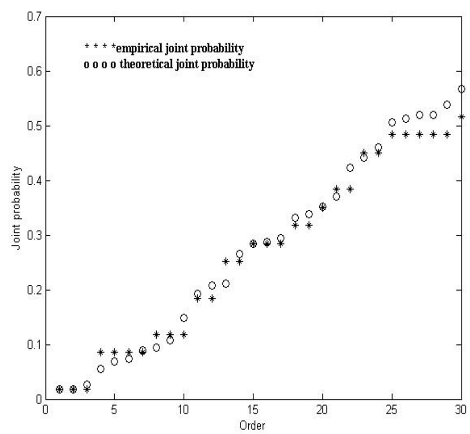

Fig. 2. Comparison of empirical and theoretical joint probability of the AMFOD and AMFM of Sanmenxia station

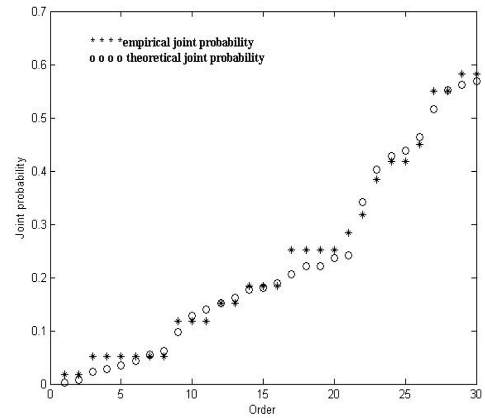

Fig. 3. Comparison of empirical and theoretical joint probability of the AMFOD and AMFM of Huayuankou station

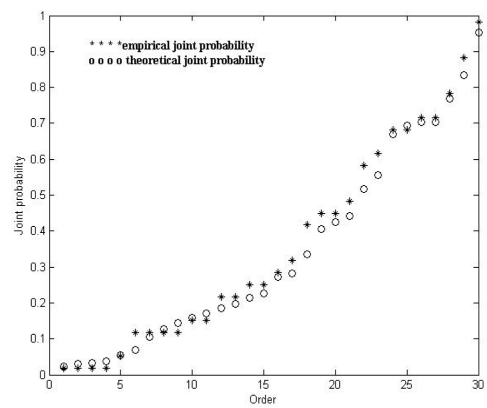

Fig. 4. Comparison of empirical and theoretical joint probability of the AMFODs of Sanmenxia and Huayuankou stations

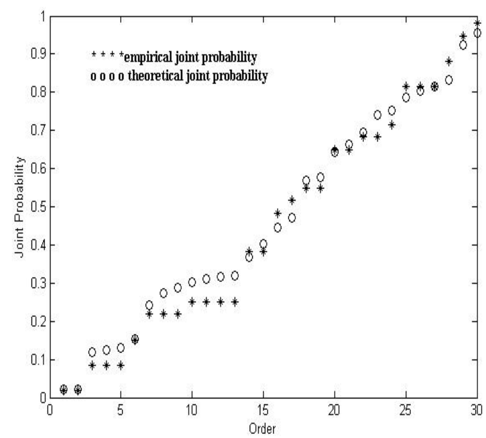

Fig. 5. Comparison of empirical and theoretical joint probability of the ARs of Sanmenxia and Huayuankou stations

The situation is the same when reversing X and Y . If we take the AMFMs of the two stations as X and Y , Eq. (9) can be used to analyze the probability of happening flood of or exceeding a given magnitude on one of the stations when the other station suffers a flood of or exceeds some magnitude. Another application of Eq. (9) is that if we set AMFM as X and AMFOD as Y or vise versa, we are able to estimate the probability of happening flood of or exceeding a given magnitude after some given date on each of the stations or the probability of a flood happening after a given date when the flood is of or exceeding some magnitude.

For instance, use Eq. (9) to calculate the conditional joint distribution between AMFMs of the two stations and between AMFM and AMFOD of each station. Here, take the former situation for example, the conditional probability of Huayuankou station is drawn up in Fig. 6 on the condition that the annual maximum flood magnitudes (AMFMs) of Sanmenxia station are respectively 11969 cubic meter per second (cms) and 8201 cms. It can be deduced from Fig. 6 that the flood peak values of marginal distribution are lower than estimated values from conditional joint distribution, which indicate that it is safer to use joint distribution than univariate distribution to analyze flood frequency.

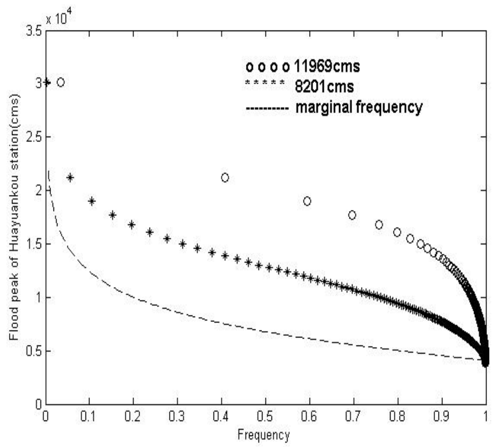

Fig. 6. Conditional frequency of Huayuankou station on the condition of different AMFM of Sanmenxia station

Further more, it is of interest to discuss the probability of a flood of or exceeding a given magnitude happening on a specific date. The following equation can be used to derive the risk of each station:

$$P_q^i = P(t_i \le T \le t_{i+1}, Q \ge q) \\ = F_t(t_{i+1}) - F_t(t_i) - H(t_i, q) + H(t_{i+1}, q). \tag{10}$$

Where, P_q^i is the encounter risk of AMFOD and AMFM. T is the happening date of the AMFM. Q is the magnitude of the AMFM. Similarly, daily encounter risk of AMFODs between the two stations can be inferred as:

$$P_i^i = P(t_i \le T_s \le t_{t+1}, t_i \le T_h \le t_{t+1}) \\ = H(t_i, t_i) + H(t_{i+1}, t_{i+1}) - H(t_i, t_{i+1}) + H(t_{i+1}, t_i). \tag{11}$$

Where, T_s and T_h is the AMFOD of Sanmenxia station and Huayuankou station, respectively.

Eqs. (10-11) are applied to illustrate the encounter risk of AMFODs of the two stations and the encounter risk of AMFM and AMFOD of each station. Using the first encounter risk, it is convenient to obtain the probability that the AMFMs of the two stations happening on the same date. The second risk stands for the probability that a flood of or exceeding some magnitude happening on a given date. Take the former encounter risk for example, Fig. 7 shows the daily risk of the Sanmenxia and Huayuankou station both suffering annual maximum flood. It can be referred from the figure that during the flood season, the risk first rise in June and arrives at peak. After that the encounter risk declines with the time. The greatest risk appears on about August 24th.

When analyzing the flood characteristics of the two stations, joint return period and concurrent return period are always should be considered. Joint return period stands for the return for either one of the two stations suffering some magnitudes of AMFMs. And the concurrent return period can represent the return for the two stations both suffering some magnitudes of AMFMs. They are given as:

$$T_j(x, y) = \frac{1}{P(X > x \bigcup Y > y)} = \frac{1}{1 - F(x, y)}. \tag{12}$$

$$T_c(x, y) = \frac{1}{P(X > x, Y > y)} \\ = \frac{1}{1 - F_X(x) - F_Y(y) + F(x, y)}. \tag{13}$$

Where, $T_j(x, y)$ and $T_c(x, y)$ represent for joint return periods and concurrent return periods.

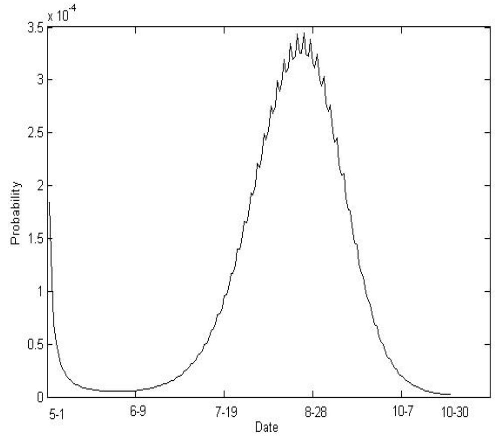

Fig. 7. Daily encounter risk between the AMFODs of Sanmenxia and Huayuankou station

Eqs. (12-13) are implemented to calculate the joint and concurrent return periods between the AMFMs of Sanmenxia station and Huayuankou station. Fig. 8 and Fig. 9 illustrate the CDF of the joint return periods and concurrent return periods between the annual maximum flood magnitudes (AMFMs) of the two stations, from which the return periods of any flood magnitude combination between the two stations can be inferred. Fig. 10 gives an isoline of the concurrent return periods. From Fig. 10, it is convenient to check out the concurrent return periods between the AMFMs of Sanmenxia and Huayuankou station or find theAMFM of one station when the other and their concurrent return period is given. For instance, when the flood peak of Sanmenxia station is 11777 m^3/s and that of Huayuankou is 13875 m^3/s, then it can be inferred that their concurrent return period is 100 year.

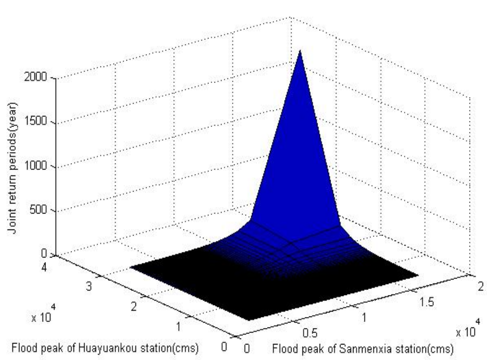

Fig. 8. Joint return periods of AMFMs of Sanmenxia and Huayuankou station

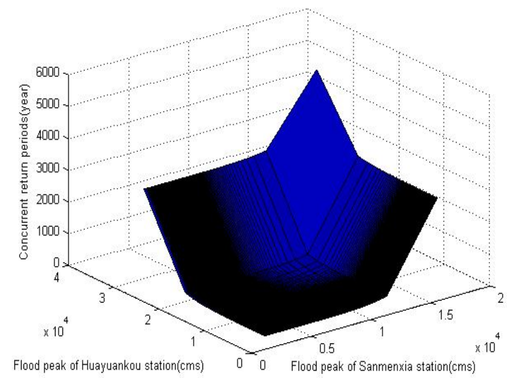

Fig. 9. Concurrent return periods between the annual maximum flood magnitudes of Sanmenxia and Huayuankou station

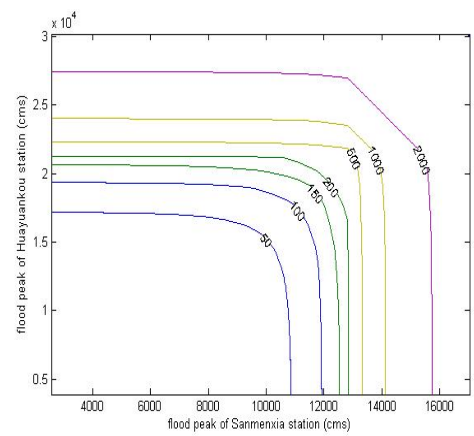

Fig. 10. isoline of concurrent return periods between the annual maximum flood magnitudes of Sanmenxia and Huayuankou station

5. Conclusions

This paper studied the characteristics of the flood and runoff of the Yellow River, China using 30-year long time series of annual runoff and annual maximum flood from two typical stations on the middle and lower reaches of the region. Von Mises distribution is applied to analyze the AMFOD distributions and P-III distribution is used for illustrating the AMFM and AR distribution. The paper presents the joint distribution between the AMFMs and the corresponding annual AMFODs of Sanmenxia and Huayuankou station. The joint characteristic of ARs of the two stations is also demonstrated. Some conclusions can be drawn as follows:

(i) It is safer to choose copula method than univariate method in flood frequency analysis.
(ii) The highest risk of Sanmenxia station and Huayuankou station both suffering the annual maximum flood peak appears on about August 24th.
(iii) This paper mainly discusses 3 common Archimedean copulas and do not consider the fitness of the other one parameter copulas and this needs further study.
(iv) The paper uses CMLE and estimating by Kendall's tau method to estimate the parameter of copula and the results show that parameters estimated via Kendall's tau produce a better GOF when the Kendall's tau is positive and parameters estimated by CMLE are fit better when the Kendall's tau is negative. But since the number of the flood data is relatively small, it may be necessary to have a deeper discuss about the parameter estimating methods according to Favre[22].

Acknowledgements

This study was supported by the National Natural Science Fund of China (No. 51190091, 41071018, 41030746), Program for New Century Excellent Talents in University (NCET-12-0262), China Doctoral Program of Higher Education (20120091110026), Qing Lan Project, the Skeleton Young Teachers Program and Excellent Disciplines Leaders in Midlife-Youth Program of Nanjing University.

References

1. M. H. Costa, A. Botta and J. A. Cardille, Effects of large-scale changes in land cover on the discharge of the Tocantins River, Southeastern Amazonia. *J. Hydrol.***283** (2003) 206-217.
2. Y. Lin and X. H. Wei, The impact of large-scale forest harvesting on hydrology in the Willow watershed of Central British Columbia. *J. Hydrol.***359** (2008)141-149.
3. Z. H. Zhang, B. Y. Wang W. P. Li and Z. S. Liu, Analysis of relation between floods of July 1958 and August 1982 in the Yellow River. *Yellow River*. **22**(11) (2000) 28-30. (in Chinese)
4. B. Y. Wang, Y. F. Wang and H. R. Li, Study on the extraordinary rainstorm and eventual flood occurred in 1761 in Sanmenxia-Huayuankou reach of the Yellow River. *Yellow River*. **24**(10) (2002) 14-15. (in Chinese)
5. L. Zhang and V. P. Singh, Bivariate rainfall frequency distributions using Archimedean copulas. *J. Hydrol.* **332** (2007) 93-109.
6. Y. Xiao, S.L. Guo, L. H. Xiong, P. Liu and B. Fang, A new Random Simulation method for constructing synthetic flood hydrographs. *J. Sichuan Univ.(Eng. Sci. Ed.)*. **39**(2) (2007) 55-60. (in Chinese)
7. S.-C. Kao and R. S. Govindaraju, A copula-based joint deficit index for droughts. *J. Hydrol.* **380** (2010) 121-134.
8. S. L. Guo, B. W. Yan, Y. Xiao, B. Fang and N. Zhang, Multivariate hydrological analysis and estimation. *J. China Hydrol.* **28**(3) (2008) 1-7. (in Chinese)
9. T. Lee and J. D. Salas, Copula-based stochastic simulation of hydrological data applied to Nile River flows. *Hydrol. Res*. **42**(4) (2011) 318-330.
10. H. Chowdhary, L. A. Escobar and V. P. Singh, Identification of suitable copulas for bivariate frequency analysis of flood peak and flood volume data. *Hydrol. Res*. **42**(2-3) (2011) 192-216.
11. J.-T. Shiau, S. Feng and S.Nadarajah, Assessment of hydrological droughts for the Yellow River, China, using copulas. *Hydrol. Process*. **21** (2007) 2157-2163.
12. Y. Y. Fu, S. W. Yan and Z. H. Ding, Study on probability between abundant and low runoffs from Weihe River and Fenhe River based on Gumble-Hougaard copula. *Water Res. Hydropower Eng.* **41**(1) (2010) 15-17. (in Chinese)
13. J. A. Mooney, P. J. Heims and I. T. Jolliffe, Fitting mixture of von Mises distributions: a case study involving sudden infant death syndrome. *Comput. statistics & Data Analysis*. **41**(3-4) (2003) 505-513.
14. J. A.Carta, C. Bueno and P. Ramírez, Statistical modeling of directional wind speeds using mixtures of von Mises distributions: case study. *Energy Conversion & Management*. **49**(5) (2008) 897-907.
15. B. Fang, S. L. Guo, Y. Xiao, P. Liu and J. Wu, Annual maximum flood occurrence dates and magnitudes frequency analysis based on bivariate joint distribution. *Adv. Water Sci*. **19**(4) (2008) 505-511. (in Chinese)
16. B. W. Yan, S. L. Guo, L. Chen and P. Liu, (2010) Flood encountering risk analysis for the Yangtze River and Qingjiang River. *J. Hydraul. Eng*. **41**(5), 553-559. (in Chinese)
17. R. B. Nelsen *An introduction to copulas*. (Springer, New York,1999)
18. L. Zhang, and V. P. Singh, Bivariate flood frequency analysis using the copula method. *J. Hydrol. Eng.* **11**(2) (2006) 150– 164.
19. C. Genest and A.-C. Favre, Everything you always wanted to know about copula modeling but were afraid to ask. *J. Hydrol. Eng*. **12**(4) (2007) 347– 368.
20. G. Kim, M. J. Silvapulle and P. Silvapulle, Comparison of semiparametric and parametric methods for estimating copulas. *Comput. Stat. data Annual*. **51** (2007) 2836-2850.
21. S. Vandenberghe, N.E.C. Werhoest and B. De Baets, Fitting bivariate copulas to the dependence structure between storm characteristics: A detailed analysis based on 105 year 10 min rainfall. *Water Resour. Res.* **46** (2010)W01512.

22. A.-C. Favre, S. E. Adlouni, L. Perreault, N. Thiémonge and B. Bobée, Multivariate hydrological frequency analysis using copulas. *Water Resour. Res.* **40** (2004) W01101.