Journal of Water Resources Research 水资源研究, 2017, 6(5), 426-434
Published Online October 2017 in Hans. http://www.hanspub.org/journal/jwrr
https://doi.org/10.12677/jwrr.2017.65050

Analysis of Correlation between River Flows Using Copula-Entropy Method

Kangdi Huang[1], Lu Chen[1*], Shenglian Guo[2], Jianzhong Zhou[1], Zhengying Yang[1]

[1]College of Hydropower & Information Engineering, Huazhong University of Science & Technology, Wuhan Hubei
[2]State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan Hubei
Email: *chl8505@126.com

Received: Jul. 26th, 2017; accepted: Aug. 9th, 2017; published: Aug. 17th, 2017

Abstract

Analysis of the dependence between the main stream and its upper tributaries is important for hydraulic design, flood prevention and risk control. In order to solve the disadvantages of the current hydrologic correlation analysis method, the method of copula entropy was introduced to estimate the dependence between Hydrological variables. The relationship between copula entropy and mutual information was discussed and the calculation procedures of copula entropy were deduced, and multiple integration and Monte Carlo methods were used to calculate the copula entropy. The upper Yangtze River was selected for case study. Results show that there is a significant difference in total correlation values, when different copula functions were used. The total correlation among the rivers is not high, and the one between Min and Tuo Rivers is the largest. There are some dependence among Jinsha, Min and Tuo Rivers, which constitutes a threat to flood control by the Three Gorges Dam (TGD). The flows of Jinsha, Jialing, Min and Tuo Rivers significantly influence the flood occurrence in the Yangtze River.

Keywords

Copula Entropy, Dependence, The Upper Yangtze River

基于Copula熵方法的河流之间的相关性研究

黄康迪[1]，陈　璐[1*]，郭生练[2]，周建中[1]，杨振莹[1]

[1]华中科技大学水电与数字化工程学院，湖北 武汉
[2]武汉大学水资源与水电工程科学国家重点实验室，湖北 武汉
Email: *chl8505@126.com

收稿日期：2017年7月26日；录用日期：2017年8月9日；发布日期：2017年8月17日

作者简介：黄康迪(1992-)，男，湖北孝感人，硕士，主要从事水文分析计算方面的研究。
*通讯作者。

文章引用: 黄康迪, 陈璐, 郭生练, 周建中, 杨振莹. 基于 Copula 熵方法的河流之间的相关性研究[J]. 水资源研究, 2017, 6(5): 426-434. DOI: 10.12677/jwrr.2017.65050

摘　要

研究长江干流及其支流之间的总相关性对于长江上游的水力设计、防洪和风险控制非常重要。针对现有的相关性计算方法的不足与缺陷，本文引入Copula熵方法，用来计算多变量之间的相关结构，并推导出Copula熵与互信息的关系与计算方法，采用多重积分法和蒙特卡罗法估计多变量之间的相关性特征。以长江上游的五条主要干支流：金沙江、岷江、沱江、嘉陵江和乌江为研究对象。计算结果表明，当使用不同的Copula函数时，总相关值存在显著差异。长江上游干支流河流之间的总相关性不大，其中岷江和沱江之间的相关性最大，金沙江，岷江和沱江之间也存在一定的相关性，因此，这几条河流有洪水遭遇的可能，对三峡大坝的防洪构成了威胁。

关键词

Copula熵，相关性，长江上游

1. 引言

洪水事件一般具有多方面因素的特征属性，而各个变量之间通常具有相关性，研究变量之间的相关性对正确认识水文事件的客观规律，合理地进行水资源开发利用和制定有效的防洪减灾措施，具有十分重要的意义。目前，表达水文变量之间相关性的方法一般采用 Pearson 线性相关方法。该方法有以下几个缺点：① 只适用于线性相关关系，② 变量必须服从多元正态分布的假设；③ 无法计算多元变量之间的相关性。但在实际应用中，并不是所有的水文变量之间都是线性相关且满足服从正态分布的假设。

近年来基于熵理论的相关性分析方法备受推崇，许多研究者已经成功的利用互信息(Mutual Information)表征随机变量概率分布之间的相关性，该方法的优点是：① 它是一种非参数方法；② 不对边缘分布的函数形式作出假设，变量可以服从任何分布；③ 它可以用于多变量之间的相关性计算。该方法是一种实现多变量之间的相关性计算的有效途径，并在水文学领域中得到了广泛的应用。陈璐等[1]利用 Copula 熵方法对神经网络的径流预报模型中的预报因子进行了选择；李帆等[2]利用 Copula 熵方法对三变量洪水频率进行了分析；韩敏等[3]基于 Copula 熵方法对互信息进行了估计。本文以长江上游干支流为研究对象，采用信息论和 Copula 函数相结合的非线性技术，即 Copula 函数的熵，对长江上游干支流之间的相关性进行了分析，为长江上游的洪水遭遇以及三峡的防洪风险分析提供了依据。

2. Copula 熵

2.1. Copula 熵的定义

令 $x \in R_d$ 为 d 维随机变量，其边缘分布函数为 $F_i(X)$，$u_i = F_i(X)$，$i = 1,2,\cdots,d$。其中，u_i 为服从均匀分布的随机变量，将 Copula 函数的熵定义为 Copula 熵，可以表示为：

$$H_c(u_1,u_2,\cdots,u_d) = -\int_0^1 \cdots \int_0^1 c(u_1,u_2,\cdots,u_d)\log\big(c(u_1,u_2,\cdots,u_d)\big)\mathrm{d}u_1\mathrm{d}u_2\cdots\mathrm{d}u_d \tag{1}$$

式中：$c(u_1,u_2,\cdots,u_d)$ 为 Copula 函数的概率密度函数，可表示为 $\dfrac{\partial C(u_1,u_2,\cdots,u_d)}{\partial u_1 \partial u_2 \cdots \partial u_d}$。下面探讨联合熵和 Copula 熵

之间的关系，变量 X 的联合概率密度函数可以定义为[4]：

$$f(x_1,x_2,\cdots,x_d)=c(u_1,u_2,\cdots,u_d)\prod_{i=1}^{d}f(x_i) \tag{2}$$

结合等式(1)，多维联合熵可表示为：

$$\begin{aligned}
H(X_1,X_2,\cdots,X_d)&=-\int_0^\infty\cdots\int_0^\infty f(x_1,x_2,\cdots,x_d)\log\left[f(x_1,x_2,\cdots,x_d)\right]\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_d\\
&=-\int_0^\infty\cdots\int_0^\infty c(u_1,u_2,\cdots,u_d)\prod_{i=1}^{d}f(x_i)\log\left[c(u_1,u_2,\cdots,u_d)\prod_{i=1}^{d}f(x_i)\right]\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_d\\
&=-\int_0^\infty\cdots\int_0^\infty c(u_1,u_2,\cdots,u_d)\prod_{i=1}^{d}f(x_i)\left\{\log\left[c(u_1,u_2,\cdots,u_d)\right]+\sum_{i=1}^{d}\log\left[f(x_i)\right]\right\}\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_d\\
&=-\int_0^\infty\cdots\int_0^\infty c(u_1,u_2,\cdots,u_d)\prod_{i=1}^{d}f(x_i)\cdot\log\left[c(u_1,u_2,\cdots,u_d)\right]\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_d\\
&\quad-\int_0^\infty\cdots\int_0^\infty c(u_1,u_2,\cdots,u_d)\prod_{i=1}^{d}f(x_i)\cdot\sum_{i=1}^{d}\log\left[f(x_i)\right]\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_d\\
&=A+B
\end{aligned} \tag{3}$$

其中 A 可表示为：

$$\begin{aligned}
A&=-\int_0^\infty\cdots\int_0^\infty c(u_1,u_2,\cdots,u_d)\prod_{i=1}^{d}f(x_i)\cdot\sum_{i=1}^{d}\log\left[f(x_i)\right]\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_d\\
&=-\int_0^\infty\cdots\int_0^\infty f(x_1,x_2,\cdots,x_d)\cdot\sum_{i=1}^{d}\log\left[f(x_i)\right]\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_d\\
&=-\int_0^\infty\cdots\int_0^\infty f(x_1,x_2,\cdots,x_d)\cdot\log\left[f(x_1)\right]\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_d\\
&\quad-\cdots-\int_0^\infty\cdots\int_0^\infty f(x_1,x_2,\cdots,x_d)\cdot\log\left[f(x_d)\right]\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_d\\
&=-\int_0^\infty\log\left[f(x_1)\right]\cdot\left[\int_0^\infty\cdots\int_0^\infty f(x_1,x_2,\cdots,x_d)\mathrm{d}x_2\cdots\mathrm{d}x_d\right]\mathrm{d}x_1\\
&\quad-\cdots-\int_0^\infty\log\left[f(x_d)\right]\cdot\left[\int_0^\infty\cdots\int_0^\infty f(x_1,x_2,\cdots,x_d)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_{d-1}\right]\mathrm{d}x_d\\
&=-\sum_{i=1}^{d}\int_0^\infty f(x_i)\log\left[f(x_i)\right]\mathrm{d}x_i=\sum_{i=1}^{d}H(X_i)
\end{aligned} \tag{4}$$

根据等式 $\mathrm{d}u=\mathrm{d}x\cdot f(x_i)$，

$$\begin{aligned}
B&=-\int_0^\infty\cdots\int_0^\infty c(u_1,u_2,\cdots,u_d)\prod_{i=1}^{d}f(x_i)\cdot\log\left[c(u_1,u_2,\cdots,u_d)\right]\mathrm{d}u_1\mathrm{d}u_2\cdots\mathrm{d}u_d\\
&=-\int_0^\infty\cdots\int_0^\infty c(u_1,u_2,\cdots,u_d)\cdot\log\left[c(u_1,u_2,\cdots,u_d)\right]\mathrm{d}u_1\mathrm{d}u_2\cdots\mathrm{d}u_d\\
&=H_C(\boldsymbol{u})
\end{aligned} \tag{5}$$

所以，联合熵可表示为：

$$H(X_1,X_2,\cdots,X_d)=\sum_{i=1}^{d}H(X_i)+H_C(u_1,u_2,\cdots,u_d) \tag{6}$$

由方程(6)可知，联合熵可分解为两部分：d 维边缘分布熵的和与 Copula 函数的熵。

2.2. Copula 熵的计算

2.2.1. 多重积分法

根据等式(1)，Copula 熵可以使用多重积分法求得。首先，需要估计 Copula 函数的参数，从而确定 Copula 的概率密度函数。多重积分法由 Berntsen 等提出(1991)，用于计算多重积分，将 Copula 概率密度函数进行积分，得到相应的 Copula 熵值[5]。

2.2.2. 蒙特卡罗法

对于变量较多的情况，多重积分法较难求得，可采用蒙特卡罗法计算 Copula 熵。对于在[0,1]中的多元向量，Copula 熵可以表示为：

$$H_C(u_1,u_2,\cdots,u_d) = -\int_{[0,1]^d} c(U)\ln c(U)\,\mathrm{d}U = -E\left[\ln c(U)\right] \tag{7}$$

Copula 熵等于 $\ln c(U)$的期望值，可以通过蒙特卡罗法求得，与多重积分法类似，首先需确定 Copula 函数的相关性系数和参数，从而确定 Copula 函数，然后计算$-\ln c(U)$的期望值。

3. 总相关

因为计算多变量联合概率时比较复杂，所以有必要计算多个变量之间的相关性。总相关性可以定义为[6]：

$$I(X_1,X_2,\cdots,X_d) = \sum_{i=1}^{d} H(X_i) - H(X_1,X_2,\cdots,X_d) \tag{8}$$

可得，对于 $d=2$ 的情况，总相关性等效于两变量公共的互信息。

根据方程(6)，方程(8)可以写为：

$$I(X_1,X_2,\cdots,X_d) = -H_C(\boldsymbol{u}) = -B \tag{9}$$

因多变量的边缘熵的和总是大于多变量的联合熵，因此相关性总为正值；当且仅当所有被选择的变量相互独立时，I 等于零。然而，如果两个变量有一些相关性，即使其余的变量是独立的，I 仍大于零。根据等式(9)，随机变量的总相关性等于它们的负 Copula 熵。基于公式(9)采用 Copula 熵方法可直接计算总相关，避免了通过公式(8)计算联合熵和边缘熵带来的误差累积效应，且基于 Copula 熵计算的总相关性仅依赖于由 Copula 参数所确定的 Copula 函数。因此，对于总相关性的估计仅需要确定 Copula 参数。

4. 应用研究

4.1. 研究区域

长江上游汇集众多支流，主要是左岸的雅砻江、岷江、沱江和嘉陵江，右岸的乌江。雅砻江汇入金沙江，因此，本研究考虑了金沙江，而不考虑雅砻江。金沙江、岷江、沱江、嘉陵江、乌江从上游到下游的五个测站分别是屏山、高场、李家庄、北碚、武隆，采用各站 1951 年至 2007 年的日平均流量数据进行计算。采用年最大取样方法，应用最大熵原理方法获得边缘分布。Kendall 和 Pearson 相关系数的计算结果如表 1 所示。

假设边缘分布是正态分布，通过正态分布和 POME 方法估计河流的边缘概率密度函数，如图 1 所示，可以看出，对于岷江和乌江，POME 估计的分布所拟合经验分布的效果优于正态分布，两江数据表现出较高的峰度和偏度，所以在这种情况下，正态分布的假设是不合适的，然而，线性相关方法主要取决于边缘分布是正态分布的假设，因此，选择 POME 方法获得边缘分布是更加合理的。

Table 1. Dependence measures for the upper Yangtze River, China, based on annual maximum data
表 1. 上游长江年最大值的相关性

河流	金沙江	嘉陵江	岷江	沱江	乌江
金沙江	1.00	−0.08	0.11	0.13	0.12
嘉陵江	−0.12	1.00	0.03	0.18	−0.15
岷江	0.16	0.13	1.00	0.36	0.01
沱江	0.19	0.37	0.50	1.00	−0.05
乌江	0.16	−0.26	−0.02	−0.10	1.00

注：右上角元素是 Kendall 相关系数；左下角元素是 Pearson 相关系数。

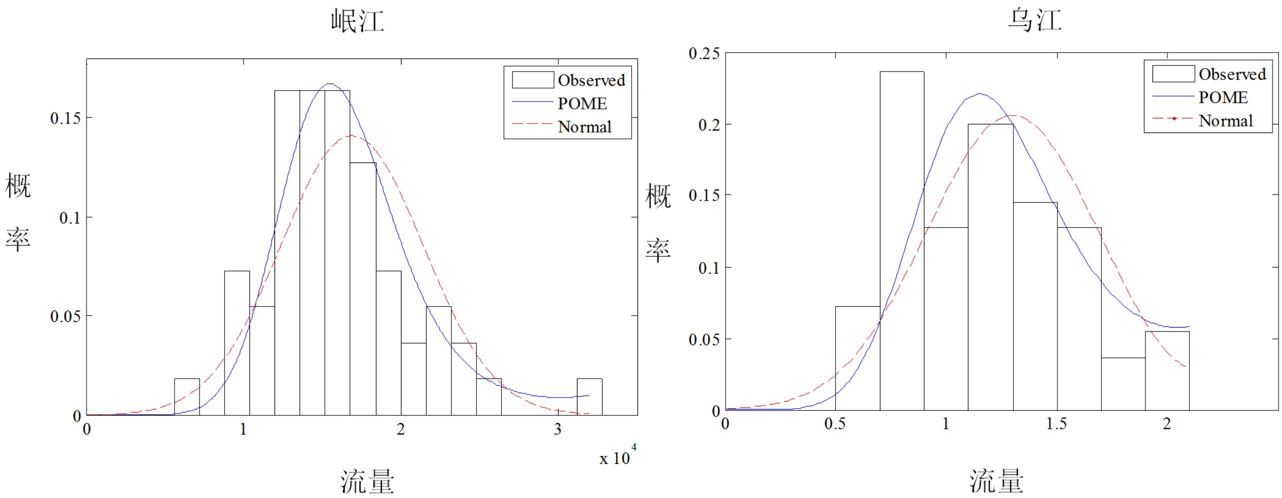

Figure 1. Results of fitting by POME method and normal distribution
图 1. 通过 POME 方法和正态分布拟合的结果

4.2. 拟合优度检验

本文利用 AIC 准则对所生成的联合分布进行拟合优度检验，结果可以判断 Copula 能否可以很好的体现出变量之间的相关性，其中最小 AIC 为 Copula 函数选取的重要依据。AIC 可以通过计算极大似然值或函数的均方误差来得到[7]，在一般情况下，AIC 可以表示为：

$$\mathrm{AIC} = 2k - 2\ln(L) \tag{10}$$

式中：k 为统计模型中的参数的数量，L 为估计模型的极大似然值。

4.3. 双变量模型

首先，确定长江上游任何两条河流的联合分布，根据该区域的 5 条河流，建立了 10 个双变量联合分布。Gumbel、Clayton、Frank、Normal 和 Student t Copula 分别用于模拟五个站之间的相关关系。利用伪极大似然法估计这些 Copula 的参数，AIC 值用于选择最合适的 Copula 函数。

所选的 Copula 和它们的参数如表 2。一般来说，阿基米德 Copula 比椭圆 Copula 能更好的拟合联合分布函数。由 Genest 等人定义的 Cramer-von Mises 函数 S_n (2009)用于拟合优度检验，计算其相应的 P 值，如表 2 所示，结果表明所选择的双变量 Copula 可用于模拟站点之间的相关关系。

分别用多重积分法和蒙特卡罗法计算 Copula 熵，对于第一种方法，采用 Berntsen 等人在 1991 年提出的多重积分法进行计算；对于第二种方法，产生 10000 对 $\boldsymbol{u}$，并计算 $\ln\left[c(\boldsymbol{u})\right]$ 的平均值，这两种方法的计算结果如表 3

Table 2. Selections of copulas and determination of the parameters
表 2. Copula 的选择和参数的确定

参数 \ Copula	金沙江	嘉陵江	岷江	沱江	乌江
金沙江	—	Frank	Clayton	Clayton	Clayton
嘉陵江	−0.74(0.71)	—	Gumbel	Gumbel	Normal
岷江	0.41(0.60)	1.09(0.10)	—	Clayton	t
沱江	0.43(0.49)	1.23(0.94)	1.21(0.38)	—	Frank
乌江	0.30(0.39)	−0.25(0.91)	0.04, 2.0(0.81)	−0.45(0.94)	—

注：右上角元素是选择的 Copula；左下角元素是相对应的参数；括号中的值为 P 值。

Table 3. Total correlation values of two tributaries in upstream Yangtze River
表 3. 长江上游两条支流的总相关值

总相关性	金沙江	嘉陵江	岷江	沱江	乌江
金沙江	1	0.008	0.053	0.057	0.032
嘉陵江	0.009	1	0.013	0.056	0.032
岷江	0.049	0.014	1	0.245	0.083
沱江	0.055	0.054	0.235	1	0.003
乌江	0.033	0.033	0.085	0.004	1

注：右上角元素是多重积分法的值；左下角元素是蒙特卡罗法的值。

所示，结果表明两种计算结果相似。从表 3 中还可看出，总相关值并不大。与 Pearson 和 Kendall 相关性系数相比，Copula 熵计算出的总相关性都为大于 0 的值，且一般小于前两种方法估计的值，根据总相关性的定义，Copula 熵计算出的值更加可靠。对于 Normal Copula，总相关可以表示为[8]：

$$I_N = -\frac{1}{2}\log\left(1-\rho^2\right) \tag{11}$$

式中：ρ 是 Pearson 相关系数。

为了评估所提出方法的合理性，采用等式(11)和提出的两种方法分别计算 Normal Copula 的总相关性。对于双变量的情况，可以采用 Pearson 相关系数来估计 Normal Copula 的参数，这三种方法的计算结果如表 4 所示，结果表明通过这三种方法计算的总相关性非常接近。因此，对特定情况的分析，可以证明所提出的方法是合理的，并且可以用于相关性的分析计算。

4.4. 多变量模型

阿基米德类的三维 Gumbel、Clayton 和 Frank Copula 以及椭圆类的 Normal Copula 和 t Copula 可用于模拟三个站之间的相关性特征。由于不对称的阿基米德 Copula 只能模拟正相关，对有负相关的情况，只能采用 Normal Copula 和 t Copula 建立联合分布。利用多重积分法计算得到的总相关性结果如表 5 所示，表中是具有代表性的组合情况，结果表明，当选择不同的 Copula 函数时，总相关值存在差异，因此，选择合适的 Copula 函数对于估计变量之间的总相关性是十分重要的。

建立四条河流的联合分布，由于复杂的相关性结构，采用椭圆 Copula，即 Normal Copula 和 t Copula。利用伪极大似然法估计参数，估计的参数如表 6 所示。从表 6 可以看出，金沙江、岷江、沱江和嘉陵江的最大总相关性为 0.36。由于金沙江、岷江、嘉陵江、沱江河流存在较大相关性，因此，这几天河流有洪水遭遇的可能，

Table 4. Comparisons of total correlation of Gaussian correlated variables in different methods
表 4. 不同方法中相关变量之间的总相关性的比较

ρ	$-\frac{1}{2}\ln\left(1-\rho^2\right)$	方法	
		多重积分法	蒙特卡罗法
−0.9	0.8304	0.8305	0.8395
−0.7	0.3367	0.3276	0.3373
−0.5	0.1438	0.1438	0.1306
−0.3	0.0472	0.0472	0.0474
−0.1	0.0050	0.0050	0.0055
0	0.0000	0.0000	0.0000
0.1	0.0050	0.0050	0.0053
0.3	0.0472	0.0472	0.0448
0.5	0.1438	0.1438	0.1473
0.7	0.3367	0.3276	0.3276
0.9	0.8304	0.8305	0.8305

Table 5. Total correlation analysis of trivariate joint distribution
表 5. 三维联合分布的总相关性分析

编号	河流			Copula	θ1 ρ1	θ2 ρ1	ρ3	υ	AIC	总相关
1	金沙江	嘉陵江	岷江	Normal	−0.12	0.23	0.12		1.58	0.05
				t	−0.13	0.23	0.1	16.97	1.53	0.05
2	金沙江	岷江	沱江	Gumbel	1.17	1.17			0.01	0.1
					1.06	1.53			−2.14	0.19
				Frank	1.41	1.41			−0.57	0.076
					0.68	3.75			−2.34	0.17
				Clay	0.52	0.52			−2.36	0.19
					0.32	1.19			−3.16	0.28
				Normal	0.24	0.28	0.59		−1.94	0.26
				t	0.22	0.23	0.57	7.7	−1.99	0.25
3	金沙江	沱江	乌江	Normal	0.28	0.22	−0.009		0.87	0.07
				t	0.26	0.22	−0.02	65.51	0.90	0.06
4	嘉陵江	岷江	沱江	Gumbel	1.24	1.24			−1.65	0.16
					1.15	1.53			−2.49	0.22
				Frank	1.73	1.73			−1.37	0.11
					0.99	3.73			−2.44	0.18
				Clay	0.48	0.48			−1.86	0.17
					0.24	1.21			−3.01	0.27
				Normal	0.14	0.34	0.58		−2.10	0.27
				t	0.07	0.32	0.56	5.42	−2.22	0.29
5	嘉陵江	沱江	乌江	Normal	0.35	−0.25	−0.03		0.10	0.1
				t	0.33	−0.24	−0.03	26.73	0.09	0.1

注：v 表示这些模型的最大似然值，下同。

Table 6. Total correlation analysis of four-dimensional joint distribution
表 6. 四维联合分布的总相关性分析

编号	河流				Copula	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6	υ	AIC	总相关性
1	金沙江	嘉陵江	岷江	沱江	Normal	−0.11	0.24	0.27	0.13	0.34	0.58		0.48	0.33
					t	−0.13	0.24	0.25	0.09	0.33	0.57	7.94	1.39	0.36
2	金沙江	嘉陵江	岷江	乌江	Normal	−0.13	0.23	0.22	0.12	−0.25	0.04		3.09	0.10
					t	−0.13	0.22	0.17	0.08	−0.24	0.05	11.24	3.91	0.10
3	金沙江	嘉陵江	沱江	乌江	Normal	−0.12	0.27	0.22	0.34	−0.25	−0.02		1.87	0.18
					t	−0.12	0.26	0.21	0.33	−0.25	−0.03	57.63	2.88	0.19
4	嘉陵江	岷江	沱江	乌江	Normal	0.13	0.34	−0.25	0.58	0.03	−0.03		0.68	0.31
					t	0.06	0.29	−0.22	0.56	0.04	−0.03	5.49	1.47	0.29
5	金沙江	岷江	沱江	乌江	Normal	0.24	0.28	0.22	0.59	0.04	−0.01		0.87	0.29
					t	0.21	0.21	0.17	0.58	0.08	−0.04	5.81	1.68	0.31

Table 7. Total correlation analysis of five-dimensional joint distribution
表 7. 五维联合分布的总相关性分析

Copula	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6	ρ_7	ρ_8	ρ_9	ρ_{10}	υ	AIC	总相关性
Normal	−0.11	0.24	0.27	0.22	0.13	0.33	−0.25	0.59	0.03	−0.02		4.21	0.39
t	−0.13	0.23	0.25	0.17	0.09	0.31	−0.24	0.57	0.06	−0.03	8.86	5.11	0.39

且对三峡的防洪构成了威胁。

建立了五条河流的联合分布，使用椭圆 Copula，估计的参数如表 7 所示。计算的 Normal Copula 和 t copula 的总相关值为 0.39，大于四变量总相关值。

5. 结论

本研究分析了长江上游五条河流的相关性。首先在水文领域引入由 Copula 和熵理论构建的 Copula 熵。由于变量之间存在复杂的相关性结构，所以采用 Copula 熵计算总相关性，即首先估计 Copula 函数的参数从而确定 Copula 函数，最后计算总相关值。主要结论如下：

1) 应用阿基米德和椭圆 Copula 建立多元变量的联合分布。一般阿基米德 Copula 更适合于维数少的情况，对于维数多的情况，其相关性复杂，只能使用椭圆 Copula，对于所建立的联合分布的理论曲线都能很好地拟合经验频率。

2) 基于信息理论和 Copula 函数，引入 Copula 熵方法，这是一种非参数方法，可以表示线性和非线性相关性，且不对边缘分布做出假设，可用于更高的维数。此外，该方法仅需要通过计算 Copula 熵来直接地估计总相关性，而不需要计算边缘熵和联合熵的值，从而避免了偏差的累积效应。

3) 使用多重积分法和蒙特卡罗法计算总相关值，得到的计算结果相近。对于特殊情况，可以采用不同类型的 Copula 函数，当使用不同的 Copula 函数时，总相关值存在显著差异，因此，选择合适的 Copula 函数对于相关性的估计是很重要的。

4) 计算结果表明，河流之间的总相关性不是很大，这与研究区域的气候特征有关。由于岷江和沱江这两条河流之间的距离最短，且属于同一暴雨区，所以它们之间的总相关系数最大为 0.33。金沙江，岷江和沱江之间

也存在一定的相关性，在平水年，由于长江左岸和右岸一般不会同时发生降雨，所以岷江和乌江的相关性不能忽视。由于金沙江，岷江，嘉陵江，沱江河流存在较大相关性，所以这几条河流有洪水遭遇的可能，且对三峡大坝的防洪构成了威胁。

基金项目

国家自然科学基金面上项目(51679094)；国家自然科学基金重点资助项目(91547208)。中央高校基本科研业务费专项资金资助(2017KFYXJJ194, 2016YXZD048)。

参考文献 (References)

[1] 陈璐，郭生练. Copula 熵理论及其在水文相关性分析中的应用[J]. 水资源研究, 2013, 2(2): 103-108.
CHEN Lu, GUO Shenglian. Copula entropy and its application in hydrological correlation analysis. Journal of Resources Research, 2013, 2(2): 103-108. (in Chinese)

[2] 李帆，郑骞，张磊. Copula 熵方法及其在三变量洪水频率计算中的应用[J]. 河海大学学报, 2016, 44(5): 443-448.
LI Fan, ZHENG Qian and ZHANG Lei. Copula entropy method and its application to trivariate flood frequency calculation. Journal of Hohai University, 2016, 44(5): 443-448. (in Chinese)

[3] 韩敏，刘晓欣. 基于 Copula 熵的互信息估计方法[J]. 控制理论与应用, 2013, 30(7): 875-879.
HAN Min, LIU Xiaoxin. Mutual information estimation based on Copula entropy. Control Theory &Application, 2013, 30(7): 875-879. (in Chinese)

[4] GRIMALDI, S., SERINALDI, F. Asymmetric copula in multivariate flood frequency analysis. Advances in Water Resources, 2006, 29(8): 1115-1167.

[5] BERNTSON, J., ESPELID, T. O. and GENZ, A. An adaptive algorithm for the approximate calculation of multiple integrals. ACM Transactions on Mathematical Software, 1991, 17(4): 437-451.

[6] MCGILL, W. J. Multivariate information transmission. Transactions of the Ire Professional Group on Information Theory, 2003, 4(4): 93-111.

[7] ZHANG, L., SINGH, V. P. Bivariate flood frequency analysis using the copula method. Journal of Hydrologic Engineering Impact Factor, 2006, 11(2): 150-164.

[8] CALSAVERINI, R. S., VICENTE, R. An information-theoretic approach to statistical dependence: Copula information. Europhysics Letters, 2009, 88(6): 3-12.