

黔江洪水峰量组合概率研究

刘和昌¹, 金占伟²

(1.水利部珠江水利委员会技术咨询(广州)有限公司, 广东 广州 510611;

2.水利部珠江水利委员会珠江水利综合技术中心, 广东 广州 510611)

摘 要: 对黔江洪水的洪峰和洪量组合概率进行研究, 以为流域洪水防治提供参考。选取 Gumbel-Hougaard Copula 函数建立黔江武宣水文站洪峰和洪量的联合分布函数, 提出以概率密度函数最大推求多变量水文概率的峰量设计值, 并采用该方法对条件组合概率、同现概率、联合概率(或现概率)、二次概率的峰量值进行了计算。结果表明, 相同重现期下推求的峰量设计值从大到小依次为联合概率(或现概率)、单变量概率、条件组合概率、二次概率和同现概率, 这与各水文概率描述的危险域大小是相符的。基于概率密度函数最大推求的峰量组合值对应最可能情形, 可为流域洪水风险管理、水利工程设计提供新的选择。

关键词: 洪水峰量联合分布; Copula 函数; Kendall 测度; 重现期; 设计值

中图分类号: P333.9

文献标识码: A

文章编号: 1000-0852(2021)06-0007-06

0 引言

洪水事件包含洪峰、洪量等多个水文特征变量。传统水文分析多采用单变量法, 如用洪峰出现(或洪量)的概率表明一场洪水的发生概率, 或者采用多个水文变量分别描述水文事件发生的概率, 未考虑各水文变量间的相关性。Copula 函数能有效的描述特征变量间的非线性关系^[1], 将联合分布分解为相关性结构和变量的边缘分布, 为分析多变量水文概率提供了较好的数学工具。

在多变量水文概率研究方面, 肖义等^[2]采用 Copula 函数构建洪水的峰、量联合分布, 以同现概率推求峰、量设计值; 李天元等^[3]推荐采用两变量同频率组合作为峰量联合分布中的设计值; 刘学等^[4]采用同现概率推求设计潮位过程; 刘和昌等^[5]以洪峰(或洪量)为已知控制变量, 以条件概率推求洪量(或洪峰)的设计值, 提出峰量的条件最可能组合、条件期望组合; Salvadori 等^[6-9]提出了根据联合概率值定义危险区域的 Kendall 重现期计算方法, 把事件发生区域分为危险域、安全域和临界面三类, 并在海岸工程以及其他领域得到初步应用; 史黎翔等^[10]根据联合分布的危险区域, 基于 Copula 函数提出 Kendall 重现期(即二次概率对应的重现期)计算方

法与洪水设计值计算方法。总结前人成果, 多变量水文概率研究集中在设计洪水、设计潮位等方面, 经历了由同现概率到联合概率(或现概率), 再到条件组合概率, 最后到二次概率的过程, 丰富了水文统计的内涵。但多变量水文概率存在多组合问题, 以峰量组合为例, 在同一分布概率下, 洪峰、洪量存在无数种组合, 而已有成果多以洪峰、洪量同概率这一特定组合来推求设计值, 存在一定的人为性。为此, 本文采用 Copula 函数构建洪水峰量联合分布, 以黔江峰量组合概率研究为例, 从概率统计角度出发, 提出以多变量水文概率分布相应的概率密度函数最大这种峰量最可能组合情形, 采用多元函数极值法推求峰量设计值, 以减少峰量同概率选取的主观性。

1 研究方法

1.1 Copula 函数

Copula 函数是定义域为[0,1]的多维联合分布函数, 洪峰、洪量二维 Copula 函数可表示为^[11]:

$$F(q, w) = P(Q \leq q \cap W \leq w) \\ = C_{\theta}(F_Q(q), F_W(w)) = C_{\theta}(u, v) \quad (1)$$

式中: C 为 Copula 函数; θ 为 Copula 函数参数; $u = P(Q \leq q) =$

$F_Q(q)$ 为洪峰 Q 的边缘分布; $v=P(W \leq w)=F_W(w)$ 为洪量 W 的边缘分布。记 $c(u,v)$ 为 C 的密度函数, $f_Q(q)$ 和 $F_W(w)$ 分别为 $u=F_Q(q)$ 和 $v=F_W(w)$ 的概率密度函数, 则式(1)的概率密度函数为:

$$f(q,w)=c(u,v)f_Q(q)f_W(w) \quad (2)$$

目前, 水文事件联合分布函数多采用 Archimedean Copula 函数簇, Gumbel-Hougaard Copula 函数即为其一种。Gumbel-Hougaard Copula 函数适用于变量间存在正相关的情形, 并且能描述变量的上尾相关性^[12], 故本文选用 Gumbel-Hougaard Copula 函数。

1.2 重现期

重现期 T 是指特定水文事件连续两次发生的平均时间间隔长度。若洪水事件的发生概率为 P , 则其重现期为:

$$T=\frac{1}{P} \quad (3)$$

1.3 单变量概率

采用“ $Q>q$ ”表示洪峰大于 q 的事件, “ $W>w$ ”表示洪量大于 w 的事件, 则相应单变量事件发生的概率为:

$$P(Q>q)=1-u \quad (4)$$

$$P(W>w)=1-v \quad (5)$$

单变量概率分别从峰、量对洪水进行描述, 单变量概率分布区域示意图 1(a) 中的阴影部分, 当洪峰超过 q (或洪量超过 w) 时, 对象将遭受洪水的侵害, 较适于防洪安全主要受洪峰(或者洪量)单方面影响的概率分析。

1.4 条件组合概率

洪峰 Q 达到某设计标准设计值 q 时, 洪量 W 所对应的值并非唯一, 存在一个条件概率分布 $F_{W|Q}(w)$:

$$\begin{aligned} F_{W|Q}(w) &= P(W \leq w | Q=q) \\ &= \frac{\partial F(q,w)/\partial q}{dF_Q(q)/dq} = \frac{\partial C(u,v)}{\partial u} \end{aligned} \quad (6)$$

此时, $F_{W|Q}(w)$ 的密度函数 $F_{W|Q}(w)=c(u,v)f_W(w)$ 最大时, 所对应的 w 称为洪量条件最可能值。同理, 洪量 W 达到某设计值 w 时, 也存在一个条件概率分布 $F_{Q|W}(q)$, 当其密度函数最大时, 所对应的 q 称为洪峰条件最可能值^[9]。条件组合概率给出了洪峰达到 q 时(或者洪量达到 w) 时所对应洪量(或者洪峰)的概率组合情况, 条件概率的分布区域示意图 1(b) 中的阴影部分, 较适于防洪安全主要受洪峰或者洪量影响且需知晓峰量关联性的概率分析。

1.5 同现概率

洪峰“ $Q>q$ ”与洪量“ $W>w$ ”同时发生的概率称为同现概率, 表示为:

$$P(Q>q \cap W>w)=1-u-v+C_\theta(u,v) \quad (7)$$

同现概率的分布区域示意图 1(c) 中的阴影部分, 较适用于防洪安全受洪峰、洪量以同时出现的关联影响概率分析。但由同现概率变量值直接划分的危险区域却存在着一定的不合理性, 如对于图 1(c) 中的 A_1 点, 两变量均较大, 其联合概率大于给定的联合概率水平, 位于对象的危险区域内, 但却不在同现概率的危险区域内, 缩小了对象的危险区域。同一同现概率下存在着多种洪峰、洪量组合, 本次选取式(7)相应的概率密度函数最大时的这一种峰量组合。

1.6 联合概率(或现概率)

对于洪峰“ $Q>q$ ”或洪量“ $W>w$ ”发生的概率称, 李天元^[9]、史黎翔^[10]等为联合概率, 表示为:

$$P(Q>q \cup W>w)=1-C_\theta(u,v) \quad (8)$$

参照统计学将“ \cup ”称为“或”, 也可将式(8)称为或现概率。联合概率(或现概率)的分布区域见图 1(d) 中的阴影部分, 较适用于防洪安全受洪峰、洪量和峰量同时影响的概率分析。由联合概率(或现概率)变量值直接划分的危险区域亦存在着一定的不合理性, 如

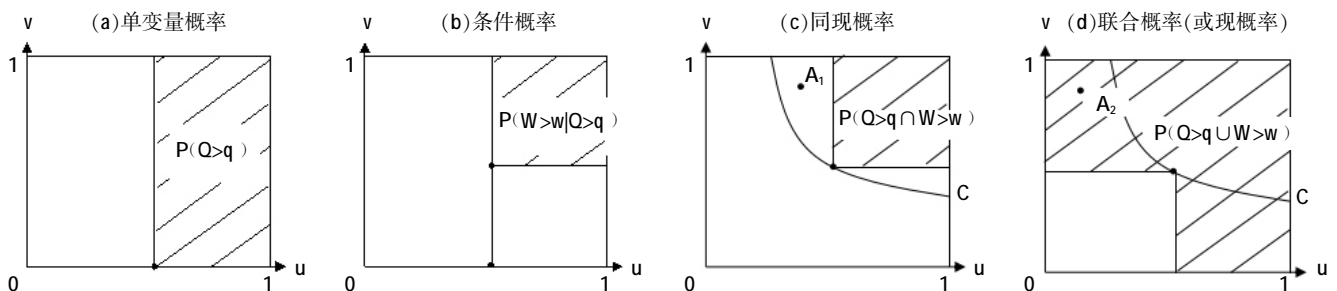


图1 概率分布区

Fig.1 Probability Area

对于图 1(d)中的 A_2 点,某一变量较大而另一变量较小,其联合概率小于给定的联合概率水平,不在对象的危险区域内,但却位于联合概率(或现概率)的危险区域内,从而扩大了对象的危险区域。同一联合概率(或现概率)亦存在着多种洪峰、洪量组合,本次选取式(8)相应的概率密度函数最大时的这一种峰量组合。

1.7 二次概率

针对同现概率缩小危险区域、联合概率(或现概率)扩大危险区域的问题,Salvadori 等^[6-9]提出了根据联合概率值定义危险区域的计算方法,将事件发生区域分为亚临界区域、临界面、超临界区域(见图 2),与洪水事件的安全域、临界面和危险相对应,其中 t_0 为临界面应的特定临界概率值,则事件发生在亚临界区域之内的概率采用 Kendall 测度 $K_c(t)$ ($t \in [0,1]$) 表示:

$$K_c(t) = P(C(u,v) < t_0) \tag{9}$$

由式(9)可知,Kendall 测度 K_c 其实是一种概率的概率,称为二次概率。对于 Gumbel-Hougaard Copula 函数,测度:

$$K_c(t) = t - \frac{\phi(t)}{\phi'(t)} = t - \frac{t \ln t}{\theta} \tag{10}$$

式中: $\phi(t) = [-\ln(t)]^\theta$ 为生成元, θ 为 Gumbel Copula 函数的参数, $t=C(u,v)$ 。

借助 Kendall 测度 K_c ,事件发生在临界面、超临界区域的概率为 $1-K_c$,二次概率的分布见图 2 中阴影部分,这与洪水事件危险区域契合较好,适用于防洪安全受洪峰、洪量危险区域影响的概率分析。由式(10)可知,一个测度 K_c 对应一个 t_0 值,而同一个 t_0 值下有多种洪峰、洪峰组合,本次选取 $t_0=C(u,v)$ 的概率密度函数最大时的这一种峰量组合。

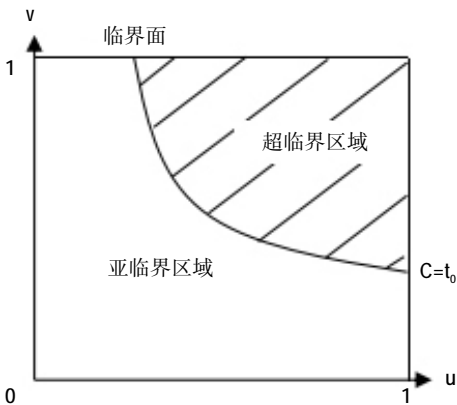


图 2 超临界洪水事件概率分布区域
Fig.2 Probability area of supercritical flood events

2 实例应用

2.1 流域概况

西江是珠江流域的干流,自古以来洪涝灾害频发,黔江是西江干流的中游段,黔江洪水峰量组合研究对掌握西江洪水规律、减少洪水风险具有重要意义。黔江武宣水文站位于广西壮族自治区武宣县城南门外窑步码头,设立于 1934 年 12 月,集水面积 196 655km²,是黔江的控制性水文站,下游为浔江防洪保护区,区内梧州市河西区远期 2025 年规划防洪标准为 100 年一遇。收集武宣水文站 1952~2005 年洪水资料,结合洪水特性和工程设计实践,选取对工程、下游防洪安全影响较大的年最大洪峰流量、7d 最大洪量两个特征变量进行峰量组合概率研究。经调查,黔江的历史特大洪水为 1902 年和 1924 年,洪峰流量分别为 51 000、48 900m³/s,7d 最大洪量分别为 270、260×10⁸m³;1902 年洪峰为 1833 年以来第一位,历史调查期为 2005-1833+1=173 年;1902 年 7d 最大洪量为 1902 年以来第一位,1924 年洪峰、1924 年 7d 最大洪量均为 1902 年以来第二位,历史调查期为 2005-1902+1=104 年。

2.2 边缘分布与联合分布

采用 P-III 型曲线,根据武宣水文站 1952~2005 年历年最大洪峰流量、7d 最大洪量系列,考虑历史洪水,确定洪峰、7d 洪量的边缘分布,结果见表 1。由武宣水文站历年洪峰与 7d 洪量系列,可计算洪峰与 7d 洪量的线性相关系数为 0.952,Kendall 秩相关系数为 0.817,点绘的峰量相关关系见图 3,据此可知洪峰与 7d 洪量相关性较强。由 Kendall 秩相关系数可推求 Gumbel-Hougaard Copula 函数参数 θ 为 5.45,由此构建武宣水文站洪峰与 7d 洪量的联合分布函数,联合分布的经验概率 P-理论概率 P 拟合情况见图 4,可知 Gumbel-Hougaard Copula 分布函数能较好的拟合洪峰与 7d 洪量联合分布情况。

表1 武宣水文站洪峰与洪量P-III型曲线统计参数
Table1 Statistical parameters of P-III curve of flood peak and flood volume at Wuxuan hydrological station

项目	均值	C_v	C_s/C_v
$Q_m/(m^3/s)$	26900	0.32	3.5
$W_{7d}/(10^8m^3)$	129	0.39	3.0

2.3 频率分析

在不同设计频率下,采用前述单变量概率、条件组合概率、同现概率、联合概率(或现概率)、二次概率五种分布概率,以相应的概率密度函数最大这种峰量组合情形,采用多元函数极值法推求峰量设计组合值,计算的武宣水文站洪峰、7d 洪量组合值见表 2;在不同洪峰、7d 洪量设计值下,5 种方法计算的洪水重现期见表 3。

由表 2 可知,同一重现期下,不同方法推求的武宣水文站峰量设计值从大到小依次为联合概率 (或现概率)、单变量概率、条件组合概率、二次概率、同现概率,最大与最小洪峰设计值相差 2.39%~4.88%,最大与最小洪量设计值相差 2.61%~5.66%。与能较好拟合洪峰、洪

量危险区域的二次概率法相比,传统单变量法设计的峰量值偏于安全,其中,100 年一遇防洪标准单变量洪峰设计值比二次概率偏大 0.9%,洪量设计值偏大 1.1%。

单变量概率与多变量概率相比,峰量设计值存在差异主要是因为武宣水文站峰量是相关关系,而不是完全相关,发生概率为 P 的洪峰时,洪量发生的概率不是 P,同一重现期下单变量概率与多变量概率的峰量设计值不一样;条件组合概率在峰(或量)已知的情況下推求量(或峰)的设计值,在同一重现期下,峰(或量)值与单变量概率的峰(或量)值一致。同现概率、联合概率(或现概率)和二次概率 3 种多变量概率,峰量相关系数和采用的 Copula 函数均一样,峰量设计值存在差异主要是因为各概率定义的危险区域不一样;在

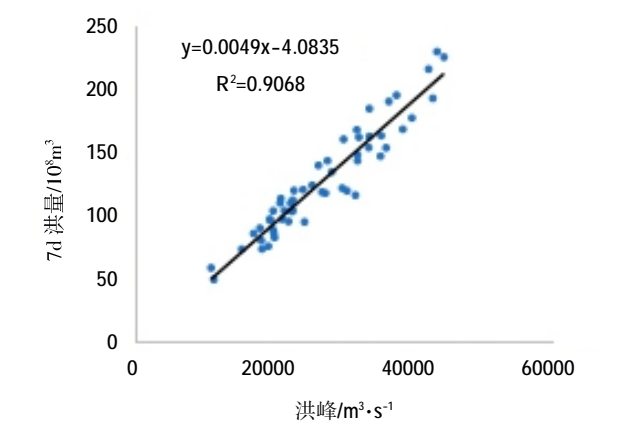


图 3 洪峰与洪量关系
Fig.3 Correlation between peak and volume

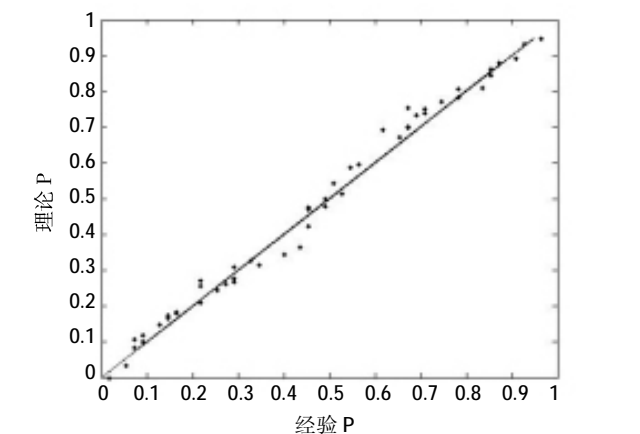


图 4 洪峰与洪量联合分布经验 P-理论 P 图
Fig.4 P-P graph of joint distribution

表2 不同频率分析方法计算的峰量设计值
Table2 Design values calculated by different frequency analysis methods

方法	项目	各重现期峰量设计值					
		1000a	500a	100a	50a	20a	10a
单变量概率	$Q_m/(m^3/s)$	67200	57800	53600	49200	43200	38400
	$W_{7d}/(10^8m^3)$	369	312	287	261	225	197
条件组合概率 1	$Q_m/(m^3/s)$	67200	57800	53600	49200	43200	38400
	最可能 $W_{7d}/(10^8m^3)$	367	311	285	259	223	195
条件组合概率 2	$W_{7d}/(10^8m^3)$	369	312	287	261	225	197
	最可能 $Q_m/(m^3/s)$	67145	57633	53450	49093	43046	38331
同现概率	$Q_m/(m^3/s)$	66522	56937	52673	48296	42275	37465
	$W_{7d}/(10^8m^3)$	364	307	281	255	219	191
联合概率(或现概率)	$Q_m/(m^3/s)$	68110	58586	54356	50017	44056	39292
	$W_{7d}/(10^8m^3)$	374	316	291	265	230	202
二次概率	$Q_m/(m^3/s)$	66881	57357	53108	48745	42743	37949
	$W_{7d}/(10^8m^3)$	366	309	284	258	222	194

同一重现期下,联合概率(或现概率)扩大危险区域范围,推求的峰量设计值最大;同现概率缩小了危险区域范围,推求的设计值最小;二次概率描述的超临界区域能较好的契合洪水危险区域范围,即不扩大,也不缩小,峰量设计值亦介于联合概率(或现概率)和同现概率之间。此外,比较图 1~2 可知,单变量概率危险区域范围略小于联合概率(或现概率)的,条件组合概率危险区域范围则与同现概率相靠近。由此可知,同一重现期下,推求的峰量值大小与各水文概率所描述的洪水危险区域大小是相符的。

由表 3 可知,同一峰量设计值下,不同方法推求的洪水重现期从大到小依次为同现概率、二次概率、条件组合概率、单变量概率、联合概率(或现概率),相同重现期下的峰量设计值越大,则相同峰量设计值下的重现期越小,这与联合概区域范围,出现概率小,重现期较大是相协调的。

3 结论

(1)黔江武宣水文站洪峰与 7d 洪量的 Kendeall 相关系数 0.817,洪水峰量存在强相关性,构建的洪峰与 7d 洪量 Gumbel-Hougaard Copula 分布函数,可较好的拟合洪峰与 7d 洪量的联合分布。二次概率能较好拟合洪水洪峰、洪量的危险区域,由其计算的武宣水文站峰量设计值小于传统单变量概率法推求的峰量设计值,传统单变量方法偏于安全。

(2)洪水峰量组合出现的不同遭遇概率可为流域防洪减灾风险管理提供多方参考。从最可能情形出发,提出以多变量水文概率分布相应的概率密度函数最大推求峰量组合值,减少了峰量同概率选取的主观性,可为工程设计提供选择。

(3)在已有研究成果基础上,本文比较分析了条件组合概率、同现概率、联合概率(或现概率)、二次概率

4 种常见多变量水文概率的概念、适用性以及危险区域的关系等,以便设计人员选用。

参考文献:

[1] Favre A C, Adlouni S E, Perrault L, et al. Multivariate hydrological frequency analysis using Copulas [J]. Water Resource Research, 2004,40(1).

[2] 肖义,郭生练,刘攀,等. 基于 Copula 函数的设计洪水过程线方法[J]. 武汉大学学报:工学版,2007,40(4):13-17. (XIAO Yi, GUO Shenglian, LIU Pan, et al. Derivation of design flood hydrograph based on Copula function [J]. Engineering Journal of Wuhan University, 2007,40(4):13-17. (in Chinese))

[3] 李天元, 郭生练, 刘章君, 等. 基于峰量联合分布推求设计洪水[J]. 水利学报, 2014,32(3):269-276. (LI Tianyuan, GUO Shenglian, LIU Zhangjun, et al. Design flood estimation based on bivariate joint distribution of flood peak and volume [J]. Journal of Hydraulic Engineering, 2014,32(3):269-276. (in Chinese))

[4] 刘学, 诸裕良, 孙波, 等. 基于 Copula 函数推求设计潮位过程线[J]. 水利学报, 2014,45(2):243-247. (LIU Xue, ZHU Yuliang, SUN Bo, et al. Derivation of design tide curve based on Copula function [J]. Journal of Hydraulic Engineering, 2014,45(2):243-247. (in Chinese))

[5] 刘和昌, 梁忠民, 姚轶, 等. 基于 Copula 函数的水文变量条件组合分析[J]. 水力发电, 2014,40(5):13-16. (LIU Hechang, LIANG Zhongmin, YAO Yi, et al. Analysis on conditional compositions of hydrological variants based on copula function [J]. Water Power, 2014,40(5):13-16. (in Chinese))

[6] Salvadori G, De Michele C, Durante F. On the return period and design in a multivariate framework [J]. Hydrology and Earth System Sciences, 2011,15(11):3293-3305.

[7] Salvadori G, De Michele C. Multivariate multiparameter extreme value models and return periods: A copula approach [J]. Water Resources Research, 2010,46(10).

[8] Salvadori G, De Michele C. On the use of Copulas in hydrology theory and practice [J]. Journal of Hydrologic Engineering, 2007,12(4):369-380.

[9] Salvadori G. Bivariate return periods via 2-Copulas [J]. Statistical Methodology, 2004,1(1):129-144.

表3 不同频率分析方法计算的峰量重现期
Table3 Return periods calculated by different frequency analysis methods

方法	项目	各峰量设计值重现期/a					
		(67200,369)	(57800,312)	(53600,287)	(49200,261)	(43200,225)	(38400,197)
单变量概率	峰(量)	1000	200	100	50	30	20
条件组合概率	峰(量)	1000	200	100	50	30	20
同现概率	峰量	1250	233	115	57	35	23
联合概率(或现概率)	峰量	833	175	88	44	26	18
二次概率	峰量	1064	216	109	54	32	21

- [10] 史黎翔, 宋松柏. 基于 Copula 函数的两变量洪水重现期与设计值计算研究[J]. 水力发电学报, 2015,34(10):27-34. (SHI Lixiang, SONG Songbai. Calculation of return periods and design values of bivariate floods based on Copula function [J]. Journal of Hydroelectric Engineering, 2015,34(10):27-34. (in Chinese))
- [11] Nelson R B. An Introduction to Copulas [M]. Springer, New York, 1999.
- [12] Poulin A, Huard D, Faver A C, et al. Importance of tail dependence in bivariate frequency analysis [J]. Journal of Hydrologic Engineering, 2007,12(4):394-403.
- [13] Genest C, Rivest L P. Statistical inference procedures for bivariate archimedean copulas [J]. Journal of the American Statistical Association, 1993,88(423):1034-1043.

Study on Probability of Qianjiang Flood Peak Volume Combination

LIU Hechang¹, JIN Zhanwei²

(1. Technical Advisory Company of Pearl River Water Resources Commission of the Ministry of Water Resources, Guangzhou 510611, China; 2. Comprehensive Technology Center of Pearl River Water Resources Commission of the Ministry of Water Resources, Guangzhou 510611, China)

Abstract: The paper studied the probability of flood peak and flood volume combination of the Qianjiang River, providing reference for flood control. The Gumbel-Hougaard Copula function was selected to establish the joint distribution function of flood peak and flood volume at the Wuxuan Hydrological Station in Qianjiang River. It was proposed to calculate flood peak and flood volume when the probability density function value of joint distribution was maximum. Conditional combination probability, co-occurrence probability, joint probability, and quadratic probability were adopted to calculate flood peak and flood volume with the method. The results show that the values of flood peak and flood volume under the same return period are in order of joint probability, univariate probability, conditional combination probability, quadratic probability and co-occurrence probability. They are consistent with the sizes of danger zone described by various probabilities. The most likely combination of flood peak and flood volume provides a new option on the flood risk management and engineering design.

Key words: flood peak and flood volume joint distribution; copula function; Kendall measure; return period; design value

(上接第 108 页)

Research on Hydrochemical Characteristics and Recharge Sources in Yangzhuang Basin

FENG Yawei^{1,2,3}, CHEN Hongnian^{1,2,3}, ZHANG Cheng⁴, GUO Yongli⁴, SHI Fengfeng^{1,2}, TANG Wei⁴

(1. Shandong Lunan Geological Engineering Survey Institute (The second Geological Brigade of SDGM), Yanzhou 272100, China; 2. Key Laboratory of Karst Geology, SDGM, Yanzhou 272100, China; 3. School of Environmental Studies, China University of Geosciences, Wuhan 430074, China; 4. Key Laboratory of Karst Dynamics, MNR&GZAR / Institute of Karst Geology, CAGS, Guilin 541004, China)

Abstract: Based on the collection and utilization of previous research in Yangzhuang basin, this paper carried out the sampling and monitoring of Hydrochemistry and isotopes, and studied evolution of the hydrochemistry process in Yangzhuang basin, which is aiming at laying a foundation for the study of Karst Critical Zone. The results show that the main type of Karst water in the study area is $\text{HCO}_3\text{-Ca}$, and the main supply source is atmospheric precipitation. In the process of runoff generation, Karst water has strong interaction with limestone, and to some extent, it is affected by evaporation, cation exchange and mixing. Affected by human activities, the content of NO_3^- and SO_4^{2-} in some areas is relatively high.

Key words: Yangzhuang basin; hydrochemistry; hydrogen and oxygen isotopes; runoff recharge