

线性矩法在长江中下游区域水文频率计算中的应用

陈元芳¹, 王庆荣¹, 沙志贵², 陈剑池²

(1. 河海大学水资源环境学院, 江苏 南京 210098 ; 2. 长江水利委员会水文局, 湖北 武汉 430010)

摘要 对长江中下游地区宜昌、沙市、螺山、汉口及大通 5 个水文站的年最大洪量序列, 采用美国学者 Hosking 等于 1997 年提出的以线性矩(L-moment)为基础的区域水文频率分析方法进行频率计算, 结果表明: 由此 5 个站点组成的区域可认为是水文频率计算的一致区, 可采用指标洪水法对各站进行频率计算, 区域统一的分布线型选择的结果是 Wakeby 分布, 而传统的极值分布、GLO、GPA、对数正态分布、P-Ⅲ分布等均不合适。
关键词 洪量序列 线性矩法 区域频率分析 指标洪水
中图分类号 P333.9 **文献标识码** A **文章编号** 1000-1980(2003)02-0207-05

我国在水文频率计算研究中, 一般仅侧重于单站情况下的线型及参数估计问题, 而对于如何利用地区的水文信息则研究得很少。实际频率计算时, 对于地区洪水信息的利用, 一般采用经验方法, 如通过利用上下游站的频率曲线目估对比作出经验判断, 必要时对频率曲线做一些调整。虽然该方法能在一定程度上利用地区水文信息, 但其经验性、任意性较大。此外, 目前我国绝大多数水文频率曲线的总体线型采用文献 [1] 所规定的 Pearson-Ⅲ(简称 P-Ⅲ) 型。这种线型对于一个具体样本是否合适, 一般也不做统计检验。而 Hosking 等^[2]在提出线性矩的基础上, 对于序列一致性、水文一致区的判别, 区域总体统一线型的选择及指标洪水法等进行研究, 形成了一套较为完善的区域水文频率分析算法。该算法曾应用于美国年暴雨量频率的分析与计算, 效果较为理想。该算法在利用地区水文信息、线型选择等方面, 比我国目前的做法更合理。关于线性矩法, 我国在此方面的研究起步较晚^[3], 而对于如何利用线性矩进行区域水文频率计算, 至今未见报道。本文采用文献 [2] 所建议的算法对长江中下游地区各主要测站的水文频率进行计算。

1 各站资料的收集与分析检验

据文献 [2], 区域频率分析步骤如下 (a) 各站资料的收集及分析检验 (b) 分区与一致区的鉴别 (c) 一致区统一分布线型的选择 (d) 各站点分布函数的确定。
本次收集了从宜昌至大通共 6 个站年最大 30 d、60 d 的洪量序列资料, 见表 1。

表 1 各站年最大 30 d、60 d 洪量
Table 1 Basic information of annual maximum 30 and 60 days' flood volume at six hydrological stations

站名	最大重现期 N/a	实测年数 n/a	历史洪水总次数 a	实测期中历史洪水次数 l	折算年数 n'/a
大通	150	45	2	1	98
汉口	134	45	4	2	98
螺山	134	45	4	2	98
皇庄	47	47	0	0	47
沙市	45	45	0	0	45
宜昌	122	122	0	0	122

注: 折算年数 n' 按文献 [1] 中建议的方法计算。
由于分洪、滞洪等人类活动影响, 序列不一致, 故宜昌、沙市、螺山、汉口及大通等站的年最大洪量未采用实测序列, 而是采用了河道洪水演算后的新序列。螺山站采用总入流。

1.1 独立性检验

对各站实测期序列计算了前 10 阶自相关系数,年最大 30 d 洪量的前 3 阶自相关系数结果见表 2。年最大 60 d 洪量与 30 d 洪量的结果是一致的,故未列出。从表 2 可知,各站均具有独立性。

表 2 最大 30 d 洪量自相关及趋势性检验结果

Table 2 Results of trend and independent tests for 30 days' maximum flood volume series

站名	r_1	r_2	r_3	u_1	u_2	实测年数 n/a
大通	0.17	0.01	0.05	1.25	-0.68	45
汉口	0.04	-0.07	0.15	0.43	1.09	45
螺山	0.02	0.09	0.10	1.08	0.54	45
沙市	0.11	0.10	0.00	-0.45	0.75	45
宜昌	0.00	0.17	0.07	-0.80	-1.45	122
皇庄	0.33	0.09	0.05	-2.58	-2.71	47

注: u_1, u_2 分别为 Kendall 及游程检验统计计算值,临界值 $u_{\alpha/2} = 1.96$ ($\alpha = 0.05$)。

1.2 一致性检验

一致性检验采用 Kendall 及游程检验法^[4],检验资料为各站实测期样本。最大 30 d 洪量的检验结果列于表 2。从表 2 可知,除皇庄站外,其余各站在 $\alpha = 0.05$ 时均可接受为无趋势,即具有一致性。但皇庄站在 $\alpha = 0.01$ 时也可认为具有一致性。最大 60 d 洪量检验结果与最大 30 d 洪量检验结果相同。

1.3 不和谐性检验 (discordancy measure test)

这是 Hosking 等^[2]提出的一种方法,可用于检验资料系列是否存在奇异点。

设一个区域有 N 个站点,每个站点均可计算出样本线性矩系数 (L-moment ratios) $\{t^{(i)}_1, t^{(i)}_3, t^{(i)}_4, i = 1, 2, \dots, N\}$ 。令

$$u_i = [t^{(i)}_1, t^{(i)}_3, t^{(i)}_4]^T \tag{1}$$

$$\overline{u} = \frac{1}{N} \sum_{i=1}^N u_i \quad A = \sum_{i=1}^N (u_i - \overline{u})(u_i - \overline{u})^T \tag{2}$$

$$D_i = \frac{1}{3} N (u_i - \overline{u})^T A^{-1} (u_i - \overline{u}) \tag{3}$$

计算出区域内各站 D_i ($i = 1, 2, \dots, N$) 后便可进行检验。当 D_i 超过临界值时,说明很可能存在奇异点或有趋势性。其临界值与站点数 N 有关,见表 3。

表 3 不和谐系数 D_i 临界值

Table 3 Critical values of discordancy coefficient D_i

站 数	5	6	7	8	9	10
D_i 临界值	1.333	1.648	1.917	2.140	2.329	2.491

由样本计算的各站 D_i (各站 D_i 下标值按表 2 中排列次序)结果见表 4。从表 4 可知,在不包括皇庄站的 5 个站点中,所有 D_i 均小于临界值,故可认为是无趋势或无奇异点存在,而包括皇庄站的 6 个站点作为一个区域计算时,皇庄站 D_i 比其临界值略大,说明皇庄站可能存在趋势性或奇异点,这与上述一致性检验结论一致。但从皇庄站洪量随时间变化过程来看,皇庄站这种趋势是不明显的,故本文仍认为皇庄可作为无趋势站进行多站综合分析。

表 4 各站不和谐系数 D_i 计算结果

Table 4 Calculated results of discordancy coefficient D_i for some stations

洪量天数/a	站 数	D_1	D_2	D_3	D_4	D_5	D_6
30	5 (不包括皇庄站)	1.32	1.00	1.08	0.77	0.83	
30	6 (包括皇庄站)	1.35	1.09	0.31	0.80	0.78	1.66*
60	5 (不包括皇庄站)	0.77	1.32	1.10	0.67	1.13	
60	6 (包括皇庄站)	0.78	0.20	1.09	0.87	1.39	1.67*

注:*表示超过临界值。

总之,该地区 5 或 6 个站点水文资料能通过可靠性、一致性及独立性检验。

2 分区与一致区的鉴别

笔者未对长江流域或更大范围内的各站点用聚类方法进行分区,而是直接对长江中下游地区选定了 6 个站所组成的区域作频率计算一致区的检验.检验方法与结论如下所述.

设研究区共有 N 个站,第 i 个站样本长度为 n_i ,其线性矩系数为 $(t^{(i)},t_3^{(i)},t_4^{(i)},i=1,2,\dots,N)$.区域平均的线性矩系数可采用如下公式进行计算:

$$t^R=\sum_{i=1}^Nn_it^{(i)}\bigg/\sum_{i=1}^Nn_i\qquad t_3^R=\sum_{i=1}^Nn_it_3^{(i)}\bigg/\sum_{i=1}^Nn_i\qquad t_4^R=\sum_{i=1}^Nn_it_4^{(i)}\bigg/\sum_{i=1}^Nn_i\qquad (4)$$

以样本长度为权重的区域样本线性矩系数 t 的标准差为

$$V=\left[\sum_{i=1}^Nn_i(t^{(i)}-t^R)^2\bigg/\sum_{i=1}^Nn_i\right]^{\frac{1}{2}}\qquad (5)$$

以 Kappa 分布(4 个参数的一种分布)为总体,并令其总体参数为区域平均的线性矩系数,即 $\lambda_{10}=1,t_0=t^R,t_{30}=t_3^R,t_{40}=t_4^R$.在给定 $\lambda_{10},t_0,t_{30},t_{40}$ 后,便可按文献[2]所介绍的方法求 Kappa 分布密度参数.这样就可利用统计试验方法生成 N_s 个各站之间独立同分布的简单样本,各站样本长度与实测系列一致,对每个样本,可按式(4)(5)计算每个模拟样本对应的 $V_i(i=1,2,\dots,N_s)$ 值.根据 N_s 个样本,便可计算 V 统计量的均值与均方差.

$$\mu_V=\sum_{i=1}^{N_s}V_i\qquad \sigma_V=\sqrt{\frac{1}{N_s}\sum_{i=1}^{N_s}(V_i-\mu_V)^2}\qquad (6)$$

表征区域内非一致性度量指标(Heterogeneity measure) H 统计量为

$$H=\frac{V-\mu_V}{\sigma_V}\qquad (7)$$

当 $H<1$ 时,该区域为可接受的一致区;当 $1\leq H\leq 2$ 时,该区域为可能的非一致区;当 $H>2$ 时,该区域为非一致区.Hosking 等[2]曾利用统计试验方法证明 H 统计量具有较强一致区鉴别能力,故本文采用该法进行一致区判别.为简化起见,Kappa 分布随机数生成时,各站的样本长度取为折算长度 n' ,生成的是简单样本.统计试验组数 $N_s=500$.

不包括皇庄站及包括皇庄站作为一致区检验时的检验结果如表 5 所示.

表 5 一致区检验结果

Table 5 Results of homogeneous region test for flood frequency analysis(H statistics)

洪量天数/ a	站数	t^R	t_3^R	t_4^R	V	μ_V	σ_V	H
30	5(不包括皇庄站)	0.089 3	-0.098 0	0.214 8	0.005 8	0.006 5	0.002 4	-0.31
30	6(包括皇庄站)	0.105 1	0.105 8	0.207 9	0.049 6	0.008 4	0.002 8	14.37
60	5(不包括皇庄站)	0.085 2	0.119 5	0.213 9	0.007 3	0.006 5	0.002 4	0.31
60	6(包括皇庄站)	0.101 3	0.126 3	0.205 6	0.051 2	0.008 5	0.003 0	14.41

从表 5 可以看出,包括皇庄的 6 个站组成的区域不是一致区,因为不管是 30 d 还是 60 d 洪量,其统计量 H 总是远大于 2,而不包括皇庄的 5 个站点,因所计算的 H 值分别为 0.31,-0.31,即满足 $H<1$ 的情况,故可认为由大通、汉口、螺山、沙市、宜昌所组成的中下游地区为水文频率计算的一致区.

3 一致区内统一分布线型的选择

3.1 原理

假设一致区内有 N 个站,每个站样本长度为 n_i ,由实测样本估算的样本线性矩系数为 $(t^{(i)},t_3^{(i)},t_4^{(i)}(i=1,2,\dots,N)$ 样本线性矩的区域平均值为 t^R,t_3^R,t_4^R .

对某一指定分布 DIST(3 个参数以上的分布),当其二、三阶总体线性矩系数可取为区域平均 t^R 及 t_3^R 时,所对应的四阶分布总体线性矩系数 t_4^{DIST} 则未必与 t_4^R 相同,因为 t_4^R 是由样本确定的,而 t_4^{DIST} 则不仅与 t^R,t_3^R 有关,而且与指定分布 DIST 形式有关.显然,当某一指定分布 DIST 所对应的 t_4^{DIST} 越接近于 t_4^R 时,指定分布

DIST 与实测资料拟合得越好.

考虑到 t_3 一般可认为是无偏的,但 t_4 在样本长度较短(或总体 $t_{30} \geq 0.4$)时可能会有一些偏差,故 Hosking 给出了考虑 t_4 偏差的 L-moment 拟优检验统计量.

3.2 检验方法

检验统计量

$$Z^{\text{DIST}} = (t_4^{\text{DIST}} - t_4^R + B_4)/\sigma_4 \tag{8}$$

$$B_4 = \frac{1}{N_s} \sum_{m=1}^{N_s} (t_4^{(m)} - t_4^R) \quad \sigma_4 = \{ (N_s - 1)^{-1} [\sum_{m=1}^{N_s} (t_4^{(m)} - t_4^R)^2 - N_s B_4^2] \}^{\frac{1}{2}} \tag{9}$$

式中 $t_4^{(m)}$ ——假定检验区域分布线型为 4 个参数的 Kappa 分布,总体参数为 $1, t^R, t_3^R, t_4^R$,由所生成的第 m 个样本(N 个站中各站长度与其折算长度 n_i' 一致,且各站之间独立)所计算的区域平均四阶线性矩系数即为 $t_4^{(m)}$; N_s ——所生成区域样本总个数.

考虑的总体线型有:Generalized Logistic (GLO), Generalized Extreme-value (GEV), Generalized Pareto (GPA), Generalized Lognormal (GLN), Pearson-III.

当计算的 $|Z| \leq 1.64$ 时,则接受所指定的分布 DIST 为区域统一的分布线型.

Hosking 等^[2]利用统计试验方法对利用上述方法选择分布线型的效果进行了检验,结果表明,假定总体分布与接受分布一致的方案占总方案的 90% 以上.

笔者对 5 个站所组成的一致区做了线型选择,结果见表 6. 由于以上 5 类分布线型所计算的 $|Z|$ 远大于 1.64,故最终选择一种具有 5 个参数能通过检验的 Wakeby 分布作为该地区的统一分布线型.笔者还曾用 P-III 型作为该地区样本总体线型对各站分别进行过频率计算,发现矩法估计参数所求设计值比绝对值准则适线法、线性矩法的结果大得多.

表 6 L-moment 法线型拟合计算结果
Table 6 Results of distribution selection test based on L-moment

分布类型	30 d 洪量			60 d 洪量		
	t_4^R	t_4^{DIST}	Z	t_4^R	t_4^{DIST}	Z
GLO	0.2148	0.0175	-4.09	0.2139	0.1790	-3.49
GEV	0.2148	0.1260	-6.55	0.2139	0.1320	-5.75
GLN	0.2148	0.1300	-6.33	0.2139	0.1340	-5.67
GPA	0.2148	0.1250	-6.58	0.2139	0.1270	-6.01
P-III	0.2148	0.0290	-11.44	0.2139	0.0370	-10.36

4 各站分布函数的确定

采用指标洪水法推求的各站分布函数为

$$X_i(p) = \mu_i q(p) \quad (0 \leq p \leq 1; i = 1, 2, \dots, N)$$

式中 μ_i ——指标洪水值,实际上可取各站样本平均值; $X_i(p)$ ——第 i 个站的分布函数; $q(p)$ ——区域综合的无尺度分布函数,即均值为 1 的分布函数.

各站最大 30 d、60 d 洪量不同频率下的设计值见表 7.

表 7 年最大 30 d、60 d 洪量不同频率设计值计算结果
Table 7 Design values of annual maximum 30 and 60 days' flood volume for different stations

频率 设计值	大通		汉口		螺山		宜昌	
	30 d	60 d	30 d	60 d	30 d	60 d	30 d	60 d
0.02	2091	3845	1748	3126	1629	2908	1317	2298
0.01	2303	4154	1897	3377	1768	3141	1430	2482
0.001	3049	5367	2511	4363	2340	4058	1892	3207
0.0001	4078	6926	3359	5631	3130	5238	2532	4139

5 结 语

本文利用 Hosking 所建议的区域水文频率计算方法对长江中下游地区洪水频率所进行的计算 ,为长江中下游地区各有关站点的洪水频率计算提供了较为可靠的设计成果 .对于其他地区或长江流域更大区域 ,也可开展这方面的研究 ,研究变量既可以是洪水也可以是暴雨等 .

参考文献 :

[1] SL 44—93 水利水电工程设计洪水计算规范[S].
[2] HOSKING J R M , WALLIS J R. Regional frequency analysis—an approach based on L-moment[M]. London :Cambridge University Press ,1997. 1—280.
[3] 陈元芳 ,沙志贵 ,陈剑池 ,等 .具有历史洪水时 P-Ⅲ 分布线性矩法的研究[J].河海大学学报(自然科学版) ,2001 29(4) :76—80.
[4] 杨位钦 ,顾岚 .时间序列分析与动态数据建模[M].北京 :北京工业学院出版社 ,1986. 150—154.

Application of L-moment based regional flood frequency analysis method to middle and lower Yangtze River

CHEN Yuan-fang¹ , WANG Qing-rong¹ , SHA Zhi-gui² , CHEN Jian-chi²

(1. College of Water Resources and Environment , Hohai Univ. , Nanjing 210098 , China ;
2. Hydrological Bureau of Yangtze River Water Resources Commission , Wuhan 430010 , China)

Abstract : The regional flood frequency analysis based on L-moment proposed by Hosking in 1997 was applied to the flood frequency analysis of annual maximum flood volume for five hydrological stations , including Yichang , Shashi , Luoshan , Hankou , and Datong on the middle and lower reaches of the Yangtze River. The calculated results showed that the area consisting of the five stations could be considered as a homogeneous region for frequency analysis , so the index-flood method was used to calculate the design value for each station , and the Wakeby distribution was recommended to be the population grow-curve (or distribution) , because the other distributions , including Pearson-III , GLO , GEV , GPA , and GLN , were excluded by the Z statistical test.

Key words : sequence of flood volume ; L-moment method ; regional frequency analysis ; index-flood