

P-III、对数 P-III和 GL 分布线型比较研究

刘 洁¹, 陈 华², 王金星³, 陈 星², 李天元²

(1. 成都大学城乡建设学院, 四川 成都 610106; 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 430072; 3. 水利部水文情报预报中心, 北京 100053)

摘 要:以淮河和珠江流域为研究对象, 收集了淮河 18 个和珠江 16 个水文站 1951~2010 年长系列日流量资料。通过年极大值取样, 应用线性矩法估计参数, 比较分析了 P-III、对数 P-III 和 GL 分布曲线在两个流域洪水频率分析计算情况。研究表明, 对于小概率事件, 在这两个流域对数 P-III 和 GL 分布曲线设计洪水值要大于 P-III 曲线分布计算的值, 而通过误差分析, P-III 曲线的误差整体上小于对数 P-III 曲线。

关键词:淮河; 珠江; P-III; 对数 P-III; GL

中图分类号: P333

文献标识码: A

文章编号: 1000-0852(2013)03-0001-04

1 引言

目前国外设计洪水的计算主要通过两种途径, 一是依据实测流量资料系列直接计算设计洪水 (或称数理统计法), 二是根据暴雨资料间接推算设计洪水 (或称水文气象法)。我国设计洪水的计算和分析, 主要采用的是第一种途径, 从历年实测流量资料中极值选择, 计算其经验频率后, 用 P-III 型适线的方法最后确定设计洪水参数值。由流量资料系列计算设计洪水, 国内已经在洪水频率及样本参数的估计、洪水频率线型的选择、不连续系列设计洪水计算方法等多方面进行了广泛的研究^[1-6], 其目的都是为了寻求更佳的概率模型来拟合样本资料, 以使估计的样本参数最能反映总体的内在分布特性, 使计算的设计洪水更加合理可靠。目前还无法得到哪一种洪水频率线型最适合国内设计洪水计算的研究结论。在样本参数估计上, 已有很多学者认为线性矩法是目前估计样本参数稳健性较好的方法^[7]。由流量推求设计洪水, 目前美国主要采用的是对数 P-III, 英国主要采用广义逻辑分布 (Generalised Logistic Distribution, 简记为 GL)^[8], 本文在收集淮河和珠江流域长系列水文资料的基础上, 用线性矩法估计参数, 对中英美三国目前采用的洪水频率计算线型进

行比较分析, 为洪水频率计算提供参考。

2 洪水频率理论分布线型

2.1 P-III 分布曲线

中国《水利水电工程设计洪水规范》规定用 P-III 型分布曲线作为水文随机变量的概率分布线型, 其概率密度函数为^[1]

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta(x-a_0)} (x-a_0)^{\alpha-1} \quad (1)$$

其中:

$$\left. \begin{aligned} \Gamma(\alpha) &= \int_0^\infty x^{\alpha-1} e^{-x} dx \\ a_0 &= \bar{x} \left(1 - \frac{2C_v}{C_s} \right) \\ \alpha &= \frac{4}{C_s^2} \\ \beta &= \frac{2}{\bar{x} C_v C_s} \end{aligned} \right\} \quad (2)$$

2.2 GL 分布曲线

英国 1999 年的《洪水估算手册》^[8]推荐采用广义逻辑分布 (GL) 作为频率曲线分布线型, GL 分布是一个三参数分布, 表达式如下:

收稿日期: 2012-01-29

基金项目: 国家重点基础研究发展规划 (973 计划) (2010CB428405); 国家自然科学基金重大项目课题 (51190094)

作者简介: 刘洁 (1983-), 女, 江西景德镇人, 讲师, 主要从事城市水文与市政工程研究。E-mail: lj2000liujie@tom.com

通讯作者: 陈华 (1977-), 女, 福建建瓯人, 副教授, 主要从事水文水资源研究。E-mail: chua@whu.edu.cn

$$Q(F)=\xi+\frac{\alpha}{k}\left\{1-\left(\frac{1-F}{F}\right)^k\right\} \quad k \neq 0 \quad (3)$$

式中: ξ 为位置参数; α 为尺度参数; k 为形状参数; F 为不超过概率。由重现期与不超过概率的关系式 $T=\frac{1}{1-F}$, 又可将上式改写成如下形式:

$$Q_T=\xi+\frac{\alpha}{k}\{1-(T-1)^{-k}\} \quad (k \neq 0) \quad (4)$$

GL 分布的可能值域范围为

$$\begin{aligned} -\infty < Q < \xi + \frac{\alpha}{k} & \quad k > 0 \\ \xi + \frac{\alpha}{k} \leq Q < \infty & \quad k < 0 \end{aligned} \quad (5)$$

因此, 当 $k > 0$ 时, GL 有上限约束。当 $k < 0$ 时, 有下限约束。

2.3 对数 P-III 分布曲线

美国联邦政府建议的洪水频率分析方法详见公报 17B^[9]。其建议洪水流量频率分析统一采用对数 P-III 型分布, 即将洪峰流量取常用对数后再来拟合 P-III 型分布。

2.4 参数估计方法

L-矩是目前洪水频率分析参数估计的最新方法之一, 它是在常规矩的基础上发展起来的用于估计洪水频率曲线参数的新方法^[1]。L-矩法起源于“概率权重矩”, 是概率权重矩的线性组合。它的最大特点是对洪

水系列中的极大值和极小值远没有常规矩那么敏感, 因此根据 L-矩法估计的洪水频率曲线参数的值比常规矩要稳健的多^[1,7]。在本文采用 L-矩法估计三种线型的参数。

3 应用研究

3.1 研究区域

淮河流域地处我国东部, 介于长江和黄河两流域之间, 是中国南北气候过渡地带。年降水量南部大别山区最大达 1 300~1 400mm, 北部黄河沿岸仅为 600~700mm。每年 6~9 月份为汛期, 降水量占年总量的 60%~70%。珠江流域地处亚热带, 北回归线横贯流域的中部, 气候温和多雨, 多年平均温度在 14~22℃ 之间, 多年平均降雨量 1 200~2 200mm, 降雨量分布明显呈由东向西逐步减少, 降雨年内分配不均, 地区分布差异和年际变化大。每年汛期两个流域都将面临巨大的防汛抗洪压力, 选取这两个流域作为研究对象, 具有重要意义。选取断面控制站信息如表 1 所示, 测站径流资料为 1951~2010 年的日径流系列。

3.2 L-矩法估计参数

观测到变量 X 的 n 个样本值, 将其按从小到大的顺序进行排列。用 $x_{(j)}$ 表示样本序列中第 j 小的值, 对前四阶的样本 L-矩进行计算如下:

表1 流域及测站基本信息表
Table1 The information of the basin and hydrometry stations

流域	序号	河流	站名	流域面积/km ²	流域	序号	河流	站名	流域面积/km ²
淮河	1	淮河干流	长台关	3 090	珠江	1	西江	江边	25 116
	2	淮河干流	息县	10 190		2	西江	盘江桥	14 492
	3	淮河干流	王家坝	30 630		3	西江	涌尾	13 045
	4	淮河干流	润河集	40 360		4	西江	三岔	16 280
	5	淮河干流	鲁台子	88 630		5	西江	柳州	45 413
	6	淮河干流	吴家渡	121 330		6	西江	平乐	12 159
	7	竹竿河	竹竿铺	1 639		7	西江	百色	21 720
	8	洪河	遂平	1 760		8	西江	新和	5 791
	9	洪河	沙口	5 560		9	西江	金鸡	9 103
	10	洪河	班台	11 280		10	北江	长坝	6 794
	11	史河	蒋家集	5 930		11	北江	高道	9 007
	12	颍河	紫罗山	1 800		12	北江	石角	38 363
	13	颍河	漯河	12 150		13	东江	河源	15 750
	14	颍河	周口	25 800		14	东江	博罗	25 325
	15	洪泽湖	临涣集	2 470		15	漠阳江	双捷	4 345
	16	沂河	东里店	1 182		16	钦江	常乐	6 645
	17	沂河	临沂	10 315					
	18	浍河	大兴镇	5 108					

$$l_1=b_0; l_2=2b_1-b_0; l_3=6b_2-6b_1+b_0; l_4=20b_3-30b_2+12b_1-b_0 \quad (6)$$

$$\text{式中: } b_0=n^{-1} \sum_{j=1}^n x_{(j)}, b_1=n^{-1} \sum_{j=2}^n \frac{(j-1)}{(n-1)} x_{(j)}, b_2=n^{-1} \sum_{j=3}^n \frac{(j-1)(j-2)}{(n-1)(n-2)} x_{(j)},$$

$$b_3=n^{-1} \sum_{j=4}^n \frac{(j-1)(j-2)(j-3)}{(n-1)(n-2)(n-3)} x_{(j)}。 \text{ 样本系列的均值, L-}$$

变差系数, L-偏态系数和 L-峰度系数则可分别定义如下: 均值($l_1=b_0$)、L-变差系数($t_2=l_2/l_1$)、L-偏态系数($t_3=l_3/l_2$)、L-峰度系数($t_4=l_4/l_2$)。

L-矩与 P-III型曲线参数的关系如下:

$$\bar{x}=l_1; C_v=l_2\pi^{\frac{1}{2}}\alpha^{-\frac{1}{2}}\Gamma(\alpha)/\Gamma(\alpha+0.5)/l_1; C_s=2\alpha^{-\frac{1}{2}}\text{sign}(\tau_3) \quad (7)$$

$$\text{式中: } \tau_3=6I_{1/3}(\alpha, 2\alpha)-3, I_x(p, q)=\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\int_0^x t^{p-1}(1-t)^{q-1}dt, \text{ 代表不完全贝塔函数比率, 其中 } \text{sign}(\tau_3) \text{ 是}$$

符号函数, 它表示 ξ, C_s 的符号与 τ_3 的符号相同。

L-矩与 GL 分布参数关系如下:

参数 k 和 β 可由样本 L-矩比值 t_2 和 t_3 计算而得

$$k=-t_3; \quad \beta=\frac{t_2k\sin\pi k}{\pi k(k+t_2)-t_2\sin\pi k} \quad (8)$$

3.3 对比结果分析

利用线性矩法估计每个站的 GL 分布、P-III 分布和对数 P-III 分布的参数, 通过适线法得到各个测站相应的理论频率曲线。由于站点比较多, 这里只给出了珠江博罗站和淮河吴家渡站两个站点的理论频率曲线图, 如图 1 和图 2 所示。从图 1 和图 2 可以看出, 线型尾端和经验点据匹配效果挺好, 其它站点也类似。通过选定不同频率, 计算得到三种线型的设计洪水值, 并将 P-III 线型和另外两种线型进行比较统计, 统计结果如表 2 和表 3 所示。

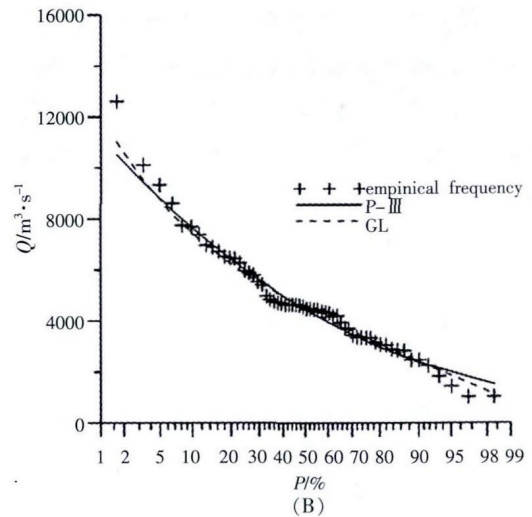
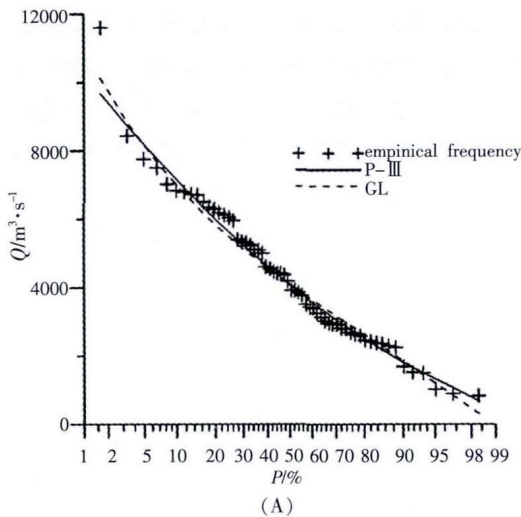


图 1 P-III 和 GL 分布频率曲线拟合结果: (A) 淮河吴家渡站, (B) 珠江博罗站

Fig.1 The fitting results of the frequency curves of P-III and GL: (A) Wujiadu station (Huaihe River), (B) Boluo station (Pearl River)

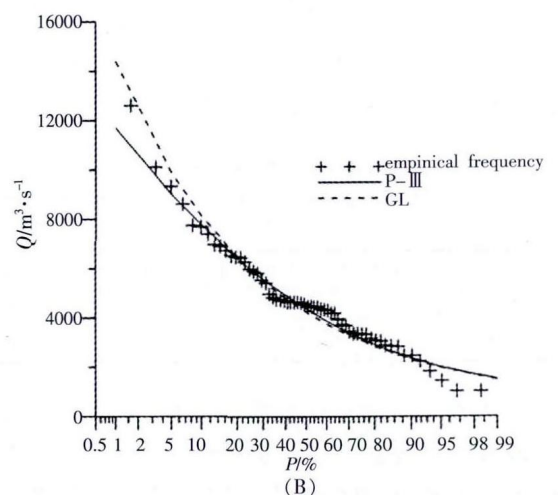
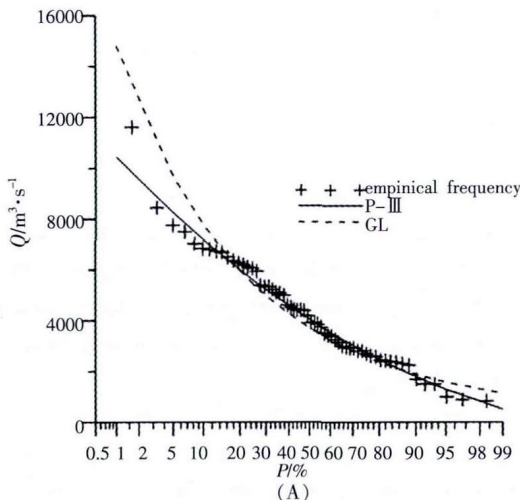


图 2 P-III 和对数 P-III (LOG P-III) 分布频率曲线拟合结果: (A) 淮河吴家渡站, (B) 珠江博罗站

Fig.2 The fitting results of the frequency curves of P-III and Log P-III: (A) Wujiadu station (Huaihe River), (B) Boluo station (Pearl River)

表2 P-III和GL分布线型洪水频率值计算比较

Table2 The comparison between the frequency curves of P-III and GL

	淮河(共 18 站)		珠江(共 16 站)	
	P-III < GL	P-III > GL	P-III < GL	P-III > GL
100 年一遇	18	0	16	0
20 年一遇	8	10	8	8
10 年一遇	0	18	0	16
5 年一遇	0	18	0	16

所有资料系列完整,代表性好,选样方法是采用年最大值选样。从图 1 和图 2 中,通过比较发现,在小概率事件上,GL 分布和对数 P-III 分布的设计洪水值要比 P-III 分布计算得到的大。从表 2 和表 3 中,通过进一步比较三种线型计算的不同频率的设计洪水值,能进一步说明从图 1 和图 2 中得到的结论。对于 100 年一遇的洪水重现期,两个流域所有站点的 GL 和对数 P-III 分布均大于 P-III 分布设计洪水值。对于 20 和 10 年一遇的设计洪水值,两个流域大部分站点的对数 P-III 分布计算值要大于 P-III 分布计算值。然而对于 GL 分布,重现期为 10 年一遇和 5 年一遇时,两个流域大部分站点的设计洪水计算值要小于 P-III 分布计算值。在工程设计上,关注更多的是稀遇洪水,而在小概率事件上,两个流域所有测站的计算结果表明 GL 分布和对数 P-III 分布的设计洪水值要大于 P-III 分布计算的设计洪水值。

3.4 误差分析

为比较各种分布线型的拟合结果与实测值的接近程度,对拟合结果进行误差分析。采用离差平方和最小准则:经验点距与同频率曲线纵坐标之差的平方和较小的拟合结果为优。对 3 种线型,比较各站点的离差平方和: $S(Q)=\sum_{i=1}^n [x_i - f(P_i, Q)]^2$ 的大小, x_i 为经验频率计算值, $f(P_i, Q)$ 为理论频率计算值。统计比较了淮河和珠江所有站点的离差平方和,结果如表 4 所示。

表4 应用离差平方和对各种分布线型计算值误差分析
Table4 The comparison of the errors between the calculated values from frequency curves

	测站数	$S(P-III)<S(GL)$	$S(P-III)<S(LogP-III)$
淮河	18	11	11
珠江	16	6	10

从表 4 可以看出在淮河流域,较多站点的 P-III 分布拟合结果误差要小于另外两种分布,而在珠江流域,较多站点的 P-III 分布拟合结果误差小于对数 P-III 分布,大于 GL 分布。

表3 P-III和对数P-III分布线型洪水频率值计算比较

Table3 The comparison between the frequency curves of P-III and Log P-III

	淮河(共 18 站)		珠江(共 16 站)	
	P-III<对数 P-III	P-III>对数 P-III	P-III<对数 P-III	P-III>对数 P-III
100 年一遇	18	0	15	1
20 年一遇	17	1	14	2
10 年一遇	15	3	14	2
5 年一遇	9	9	7	9

4 结语

设计洪水计算一直是水文学研究上的热点和难点问题。本文比较分析了三种目前在中国、英国和美国主要使用的理论频率曲线线型,并在淮河和珠江流域共 34 个站点进行应用比较。通过比较发现,在小概率事件上,GL 分布和对数 P-III 分布的洪水频率计算值比 P-III 分布计算值大。但是由于小概率事件受限于样本序列的长度,很难进行更为有效的科学验证,所以不能从以上的分析断定哪种分布更为合理。通过进一步的误差分析,在淮河和珠江流域 P-III 分布拟合结果误差要小于对数 P-III 分布,可以判断在两个流域洪水频率计算中,P-III 分布要优于对数 P-III 分布。对于 GL 分布,淮河较多站点的 P-III 分布拟合结果误差小于 GL 分布,而珠江流域较多站点误差要大于 GL 分布。

参考文献:

- [1] 郭生练. 设计洪水研究进展与评价 [M]. 北京: 中国水利水电出版社, 2005. (GUO Shenglian. Advancement and Assessment of Design Flood [M]. Beijing: China WaterPower Press, 2005. (in Chinese))
- [2] 马秀峰. 计算水文频率参数的权函数法 [J]. 水文, 1985, (3): 1-8. (MA Xiufeng. The weight function method in calculating parameters of hydrological frequency curve [J]. Journal of China Hydrology, 1985, (3): 1-8. (in Chinese))
- [3] 金光炎. 水文分析中的经验频率 [J]. 水文, 1994, 14 (1): 1-8. (JIN Guangyan. Empirical frequency calculation in hydrological analysis [J]. Journal of China Hydrology, 1994, (1): 1-8. (in Chinese))
- [4] 刘光文. 皮尔逊 III 型分布参数估计 [J]. 水文, 1990, (4): 1-15. (LIU Guangwen. Parameter estimation of P-III frequency curve [J]. Journal of China Hydrology, 1990, (4): 1-15. (in Chinese))
- [5] 费永法. 浅析历史特大洪水在设计洪水中的作用 [J]. 水文, 1998, (6): 7-12. (FEI Yongfa. Analysis of effects of historical extraordinary flood events on design flood [J]. Journal of China Hydrology, 1998, (6): 7-12. (in Chinese))
- [6] 熊立华, 郭生练. L-矩在区域洪水频率分析中的应用 [J]. 水力发电, 29 (3): 6-8. (XIONG Lihua, GUO Shenglian. Application of L-moments in the regional flood frequency analysis [J]. Water Power, 29(3): 6-8. (in Chinese))

(下转第 17 页)

- Chinese))
- [18] 丁永建,叶佑生,周文娟. 黑河流域过去 40a 来降水时空分布特征 [J]. 冰川冻土, 1999, 21 (1):42-48. (DING Yongjian, YE Baisheng, ZHOU Wenjuan. Temporal and spatial precipitation distribution in the Heihe catchment, northwest China, during the past [J]. Journal of Glaciology and Geocryology, 1999, 21(1):42-48. (in Chinese))
- [19] Richard R. Heim Jr, 周跃武, 冯建英. 美国 20 世纪干旱指数评述 [J]. 干旱气象, 2006, 24 (1):79-89. (Richard R. Heim Jr, ZHOU Yuewu, FENG Jianying, A Review of twentieth-century drought indices used in the United States [J]. Arid Meteorology, 2006, 24 (1):79-89. (in Chinese))
- [20] 焦菊英, 王万忠. 黄土高原降雨空间分布的不均匀性研究 [J]. 水文, 2001, 21 (2):20-24. (JIAO Juying, WANG Wanzhong. Non-uniformity of spatial distribution of rainfall on the loess plateau [J]. Journal of China Hydrology, 2001, 21(2):20-24. (in Chinese))

Study on Similarity Model of Precipitation Series Based on Precipitation Type Histogram

ZHANG Longfei¹, ZHU Yuelong¹, LI Shijin¹, GAO Xiangtao²

(1. School of Computer & Information, Hohai University, Nanjing 210098, China;

2. Hydrology and Water Resources Survey Bureau of Jiangsu Province, Nanjing 210029, China)

Abstract: Similarity model for precipitation stations based on analysis of the single precipitation type histogram was proposed to study the similarity between the precipitation stations. Firstly, the precipitation item was extracted from the precipitation time series, the features of single rainfall were chosen and extracted after data preprocessing. Then, the data standardization was conducted on the feature vectors. Further more, the standardized feature data were clustered using K-means clustering method based on the Davies-Bouldin index cluster validity index. Finally, the clustering results were analyzed using the single precipitation type histogram similarity model. The experimental results show that the similarity for the stations in the related regions is consistent with the underlying fact, which verifies that the proposed model is of great value for water resources management.

Key words: precipitation time series; similarity analysis; feature extraction; Davies-Bouldin index; K-means clustering; histogram of precipitation type

(上接第 4 页)

- [7] 陈元芳,沙志贵,顾圣华,等. 可考虑历史洪水对数正态分布线性矩法的研究[J]. 河海大学学报(自然科学版), 2003, (01):80-83. (CHEN Yuanfang, SHA Zhigui, GU Shenghua, et al, Study on L-moment estimation method with consideration of historical floods for log-normal distribution [J]. Journal of Hehai University (Natural Sciences), 2003, (01):80-83. (in Chinese))
- [8] Institute of Hydrology. Flood Estimation Handbook (Volumes 1-5) [M]. Institute of Hydrology, Wallingford, UK, 1999.
- [9] Maidment D R. Handbook of Hydrology [M]. McGraw-Hill, INC., 1992.
- [10] 周芬,郭生练. P-III 型分布参数估计方法的比较研究[J]. 水电能源科学, 2003, 21 (3):10-13. (ZHOU Fen, GUO Shenglian. Comparative study on parameter estimation methods of Pearson type III distribution [J]. Water Resources and Power, 2003, 21(3):10-13.(in Chinese))

Comparison among Theoretical Frequency Distributions of P-III, Log P-III and GL

LIU Jie¹, CHEN Hua², WANG Jinxing³, CHEN Xing², LI Tianyuan²

(1. School of Urban and Rural Construction, Chengdu University, Chengdu 610106, China;

2. State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China;

3. Bureau of Hydrology, MWR, Beijing 100053, China)

Abstract: The theoretical frequency distributions of P-III, Log P-III and GL (Generalized Logistic Distribution) were compared and analyzed by using annual maximum sampling and L-moment method in the river basins of the Huaihe and Pearl, where there are 18 and 16 hydrometry stations with data series (1951-2010) collected, respectively. The results show that the design flood from the curve of P-III is less than that from the other two curves and the errors of P-III are less than these of log P-III by analyzing the errors of the curve fitting in these two basins.

Key words: Huaihe River; Pearl River; P-III; Log P-III; GL