多调查期洪水频率计算及参数估计公式推导

黄华平,梁忠民

(河海大学水文水资源学院,江苏 南京 210098)

摘 要:基于实测极值流量系列进行水文频率分析时,为了提高系列的代表性及洪水估计的可靠性,常将历史洪水信息融合到实测系列构成不连续样本。考虑到目前给出的不连续样本经验频率计算方法多数是针对单个调查期,本文提出了适用于多个调查期不连续样本的经验频率计算方法,同时推导了均值(EX)和变差系数(C_v)的矩法计算公式,并将该方法应用于某水文站年洪峰系列的频率分析中。

关键词:洪水频率;调查考证期;不连续样本;统一处理法;分别处理法;矩法

中图分类号:P333 **文献标识码**:A **文章编号**:1000-0852(2016)03-0001-05

1 前言

设计洪水是确定水利工程建设规模及制定管理运行策略的重要依据。国内外大都以概率论与数理统计为理论基础,根据水文极值系列进行概率预估。然而目前的水文极值样本系列一般较短,代表性不高,用长度只有几十年的极值样本估计百年一遇、千年一遇的设计洪水将产生较大误差,故而提高洪水频率计算精度是推求设计洪水的关键。目前,将历史洪水信息融合到连续的观测系列是一种提高系列代表性及设计可靠性的重要手段。

国内外很多学者对融入历史洪水的不连续样本进行了相关研究,费永法[1]采用试验方法研究历史特大洪水对设计洪水频率曲线参数及设计值的影响,得出实测洪水系列中添加适量的历史洪水将大幅提高设计洪水的稳定性和精度的结论。王善序[2]使用次序统计量的方法对具有历史洪水不连续系列经验频率进行了相关研究;胡进宝等[3]进行了多级历史洪水频率及参数公式的研究,推导了两级历史洪水的经验频率与参数计算公式;丁晶、宋德敦等[4-5]提出了P-Ⅲ分布在具有历史洪水时的样本概率权重矩;陈元芳等[6]进行了具有历史洪水时P-Ⅲ分布线性矩法的研究;Martins等[7]研究了历史洪水在POT方法和AMS方法中的应用。然而,这些研究多数是集中在单调查期不连续样本的情况,而关于多调查期不连续系列经验频率计算及参数估计的研究鲜见报道。

本文基于单调查期不连续样本经验频率计算的概念,采用归纳法推导得出多调查期不连续样本的经验频率计算公式;并假设各调查期去除特大值后的均值及均方差与实测系列去除特大值后相应值相等,推导出多调查期不连续样本均值与C_v的矩法计算公式,并提供上述公式的应用示例。

2 理论方法

2.1 经验频率计算

2.1.1 单调查期洪水样本经验频率计算

在实测序列中加入历史洪水能相对提高洪水频率计算精度,进而提高设计洪水的精度。目前对于单调查期不连续洪水样本经验频率的计算通常有两种方法:

(1)统一处理法

$$P_M=\frac{M}{N+1}\qquad M=1,2,\cdots,a \tag{1}$$

$$P_m=P_a+(1-P_a)\frac{m}{n+1}\qquad m=1,2,\cdots,n \tag{2}$$

或者 $$P_m=P_a+(1-P_a)\frac{m-l}{n-l+1}\qquad m=l+1,l+2,\cdots\cdots,n \tag{3}$$

式中:N为调查期长度;n为实测系列年数;a为N年

收稿日期:2015-07-01
基金项目:国家科技支撑计划课题(2013BAB06B01)
作者简介:黄华平(1993-),男,江西抚州人,硕士研究生,主要从事水文水资源方面研究。E-mail:huanghp93@163.com
通讯作者:梁忠民(1962-),男,辽宁丹东人,教授,博导,主要从事水文水资源方面研究。E-mail:zmliang@hhu.edu.cn

内的特大值个数；l 为特大值中来自实测系列的个数；M 为特大洪水在调查期内的排位；m 为实测洪水在实测期内的排位；P_a 为调查期内排位为 a 的特大洪水的经验频率；P_M 为调查期内排位为 M 的特大洪水的经验频率；P_m 为实测期内排位为 m 的实测洪水的经验频率。

公式(2)与公式(3)计算结果一般相差不大，我国现行 SL44-2006《水利水电工程设计洪水计算规范》[8]中，推荐使用式(3)。

(2)分别处理法

前 a 项特大洪水仍采用公式(1)进行计算，实测系列中 n-l 项的经验公式按照式(4)进行计算：

$$P_m=\frac{m}{n+1} \qquad m=l+1, l+2, \cdots\cdots, n \tag{4}$$

在进行历史洪水的调查过程中，有时将上溯到千年前，将这些历史洪水融入实测序列中将组成具有多个调查期的不连续洪水样本。上述公式仅仅只能处理单调查期不连续样本的问题，而多调查期不连续洪水样本的处理方法却鲜见报道。因此，本文将依据单调查期不连续样本经验频率计算公式，采用推广与归纳的方法，就多调查期不连续洪水样本经验频率计算展开讨论。

2.1.2 多调查期洪水样本经验频率计算

假定一组多调查期洪水样本存在 m 个调查期，调查期长度由长到短排列分别为 N_1, N_2, ……, N_{m-1}, N_m，实测期长度为 n，样本个数为 Z，如图 1 所示。

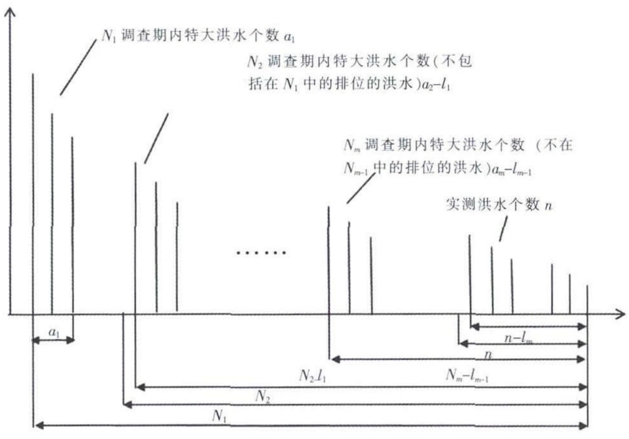

图 1 多个调查期的不连续洪水样本组成及参数示意图

Fig.1 The parameters and composition of the non-simple sample in the investigation period

图中 $a_i(i=1,2\cdots\cdots m)$为第 i 个调查期内的特大洪水个数(包括历史洪水与实测特大值洪水)；$l_i(i=1,2\cdots\cdots m)$为第 i 个调查期中来自第 i-1 个调查期内的特大洪水个数；l_m 为第 m 个调查期中来自实测期的特大洪水个数。(注：l_0 没有实际意义，如在下式中出现，则记为 0 代入计算)

(1)统一处理法

使用公式(1)计算可得第一个调查期内末位洪水的经验频率：

$$\frac{a_1}{N_1+1} \tag{5}$$

将公式(5)记为 P_a 代入公式(3)进行计算可得第二个调查期内末位洪水的经验频率：

$$\frac{a_1}{N_1+1}+\frac{a_2-l_1}{N_1+1}\cdot\frac{N_1-a_1+1}{N_2-l_1+1} \tag{6}$$

同理，将公式(6)记为 P_a 代入公式(3)进行计算可得第三个调查期内末位洪水的经验频率：

$$\frac{a_1}{N_1+1}+\frac{a_2-l_1}{N_1+1}\cdot\frac{N_1-a_1+1}{N_2-l_1+1}+\frac{a_2-l_2}{N_1+1}\cdot\frac{N_1-a_1+1}{N_2-l_1+1}\cdot\frac{N_2-a_2+1}{N_3-l_2+1} \tag{7}$$

以此类推，可得第 $s(2\leqslant s\leqslant m)$个调查期内末位洪水的经验频率：

$$\frac{a_1}{N_1+1}+\sum_{i=2}^{s}\left(\frac{a_i-l_{i-1}}{N_1+1}\cdot\prod_{j=2}^{i}\frac{N_{j-1}-a_{j-1}+1}{N_j-l_{j-1}+1}\right) \tag{8}$$

由数学归纳法：

记第 m-1 个调查期内末位洪水的经验频率为：

$$\frac{a_1}{N_1+1}+\sum_{i=2}^{m-1}\left(\frac{a_i-l_{i-1}}{N_1+1}\cdot\prod_{j=2}^{i}\frac{N_{j-1}-a_{j-1}+1}{N_j-l_{j-1}+1}\right) \tag{9}$$

将公式(9)记为 P_a 代入公式(3)进行计算，易化简得第 m 个调查期内末位洪水的经验频率为：

$$\frac{a_1}{N_1+1}+\sum_{i=2}^{m}\left(\frac{a_i-l_{i-1}}{N_1+1}\cdot\prod_{j=2}^{i}\frac{N_{j-1}-a_{j-1}+1}{N_j-l_{j-1}+1}\right) \tag{10}$$

基于以上结论，对于该不连续洪水样本中按从大到小排列的第 k 项洪水，如果在第 1 个调查期内排位，则相应的经验频率为：

$$P(k)=\frac{k}{N_1+1} \tag{11}$$

如果在第 $s(2\leqslant s\leqslant m)$个调查期内排位，则相应的经验频率为：

$$P(k)=P_{\text{前末}}+(1-P_{\text{前末}})\cdot\frac{k-\sum_{i=1}^{s-1}(a_i-l_{i-1})}{N_s-l_{s-1}+1} \tag{12}$$

其中 $P_{\text{前末}}=\frac{a_1}{N_1+1}+\sum_{i=2}^{m}\left(\frac{a_i-l_{i-1}}{N_1+1}\cdot\prod_{j=2}^{i}\frac{N_{j-1}-a_{j-1}+1}{N_j-l_{j-1}+1}\right)$

如果在实测期内排位，则相应的经验频率为：

$$P(k)=P_{前末}+(1-P_{前末})\cdot\frac{(k-\sum_{i=1}^{m}(a_i-l_{i-1})}{n-l_m+1} \tag{13}$$

其中：$P_{前末}=\frac{a_1}{N_1+1}+\sum_{i=2}^{m}\left(\frac{a_i-l_{i-1}}{N_1+1}\cdot\prod_{j=2}^{i}\frac{N_{j-1}-a_{j-1}+1}{N_j-l_{j-1}+1}\right)$

（2）分别处理法

基于分别处理法的多调查期洪水样本经验频率计算与单调查期洪水样本类似，第一个调查期内的前 a_1 项特大洪水的经验频率仍然采用式（1）计算；第 i（i=2,3……m）个调查期前 l_{i-1} 个特大洪水的序位保持“空位”，从 $l_{i-1}+1$ 开始计算其他样本的经验频率。因此，对于该不连续洪水样本中按从大到小排列的第 k 项洪水，如果在第 1 个调查期内排位，经验频率计算公式同公式（11），如果在第 s（$2\leqslant s\leqslant m$）个调查期内排位，则相应的经验频率为：

$$P(k)=\frac{k-\sum_{i=1}^{s-1}(a_i-l_{i-1})+l_{s-1}}{N_s+1} \tag{14}$$

如果在实测期内排位，则相应的经验频率为：

$$P(k)=\frac{k-\sum_{i=1}^{m}(a_i-l_{i-1})+l_m}{n+1} \tag{15}$$

目前在我国，统一处理与分别处理法都在使用，关于统一处理法与分别处理法的对比问题已有共识，即两种方法的结果一般差别不大。但是使用分别处理法公式计算不连续系列的经验频率时，可能会出现相邻两个调查期内特大洪水“重叠”的现象，即前一个调查期内末位的几项洪水的经验频率可能比后一个调查期内首几项洪水更大，造成不合理的现象。因此目前更倾向于使用统一处理法进行经验频率的计算。

2.2　参数估计

对于不连续洪水样本均值与 C_v 的矩估计，目前大部分采用克里茨基–闵光里所推荐的公式：

$$\overline{Q}=\frac{1}{N}\left[\sum_{i=1}^{a}Q_i+\frac{N-a}{n-l}\sum_{j=l+1}^{n}Q_j\right] \tag{16}$$

$$C_v=\frac{1}{\overline{Q}}\sqrt{\frac{1}{N-1}\left[\sum_{i=1}^{a}(Q_i-\overline{Q})^2+\frac{N-a}{n-l}\sum_{j=l+1}^{n}(Q_j-\overline{Q})^2\right]} \tag{17}$$

式中：Q_i 为特大洪水；Q_j 为一般洪水；l 为实测系列内特大洪水的个数；a 为特大洪水的总个数；N 为调查期长度；n 为实测期长度。

对于单调查期不连续样本，公式（16）、（17）假定 $\overline{Q}_{N-a}=\overline{Q}_{n-l}$ 和 $\overline{\sigma}_{N-a}=\overline{\sigma}_{n-l}$，即去除特大值后的 $N-a$ 年的均值和均方差与 $n-l$ 年的相等，将该假设推广到多调查期不连续样本中，令 $\overline{Q}_{N_k-Rl}=\overline{Q}_{n-l}$ 和 $\overline{\sigma}_{N_k-Rl}=\overline{\sigma}_{n-l}$（$k$=1,2,……,$m$），式中 $Rl=a_1+\sum_{i=2}^{k}(a_i-l_{i-1})$，即各调查期去除历史洪水与实测特大值后 N_k-Rl 年（k=1,2,……,m）的均值和均方差与 $n-l$ 年的对应值相等。利用矩法，可推导得：

$$\overline{Q}=\frac{1}{N_1}\left[\sum_{i=1}^{Rl}Q_i+\frac{N_1-Rl}{n-l_m}\sum_{i=Rh+1}^{z}Q_i\right] \tag{18}$$

$$C_v=\frac{1}{\overline{Q}}\sqrt{\frac{1}{N_1-1}\left[\sum_{i=1}^{Rl}(Q_i-\overline{Q})^2+\frac{N_1-Rl}{n-l_m}\sum_{i=Rh+1}^{z}(Q_i-\overline{Q})^2\right]} \tag{19}$$

表1　历史洪水分级排位与经验频率计算表

Table1　The Sequenec and classification of historical flood and the results of the empirical frequency calculation

调查考证期或实测期	系列年数/a	年份	Q/m^3·s^{-1}	排位	总排位	统一处理法	分别处理法
调查考证期 N_1（1456~2009 年）	554	1597	11000	1	1	0.0018	0.0018
调查考证期 N_2（1723~2009 年）	287	1723	9500	1	2	0.0053	0.0035
		1996	9200	2	3	0.0087	0.0069
		1795	8700	3	4	0.0122	0.0104
		1852	8500	4	5	0.0157	0.0139
调查考证期 N_3（1842~2009 年）	168	1996	9200	1		已在 N_2 中排位	已在 N_2 中排位
		1852	8500	2		已在 N_2 中排位	已在 N_2 中排位
		1921	7400	3	6	0.0216	0.0178
		1998	6470	4	7	0.0275	0.0237
实测期 n（1978~2009 年）	32	1996	9200	1		已在 N_3 中排位	已在 N_3 中排位
		1998	6470	2		已在 N_3 中排位	已在 N_3 中排位
		1997		3	8	0.0588	0.0909
		……	600~5400	……	……	……	……
		1981	（不包括 1996 年和 1998 年洪峰流量）	32	37	0.9686	0.9697

式中：$Rl=a_1+\sum_{i=2}^{m}(a_i-l_{i-1})$；$Q_i$ 为该不连续洪水样本中由大向小排列的第 i 场洪水；Z 为该不连续洪水样本个数。

3 算例

某水文站 1978~2009 年共有 32 年连续的实测洪峰资料，实测最大洪峰为 9 200m³/s，发生在 1996 年，次大洪峰为 6 470m³/s，发生在 1998 年。根据水文站所处地区附近的州、县文献记载，历史洪水与实测特大值分级、排位情况表如表 1 所示。

根据历史洪水调查表与实测资料组成的不连续洪水样本，采用统一处理法与分别处理法分别对该样本进行处理，处理结果如表 1。

采用公式(18)、(19)估计的洪峰流量均值为 2 090m³/s，C_v 计算为 0.66。采用基于 OLS 准则的优化适线法对上述样本点据分别进行适线处理，适线结果如图 2 所示，适线参数如表 2 所示，二者的均值与公式(18)估计的均值没有差别，C_v 比公式(19)估计的 C_v 稍大，曲线整体拟合较好。

表2 优化适线参数

Table2 The parameters of the optimum curve-fitting method

	均值/m³·s⁻¹	C_v	C_s
统一处理法	2090	0.82	2.23
分别处理法	2090	0.83	1.94

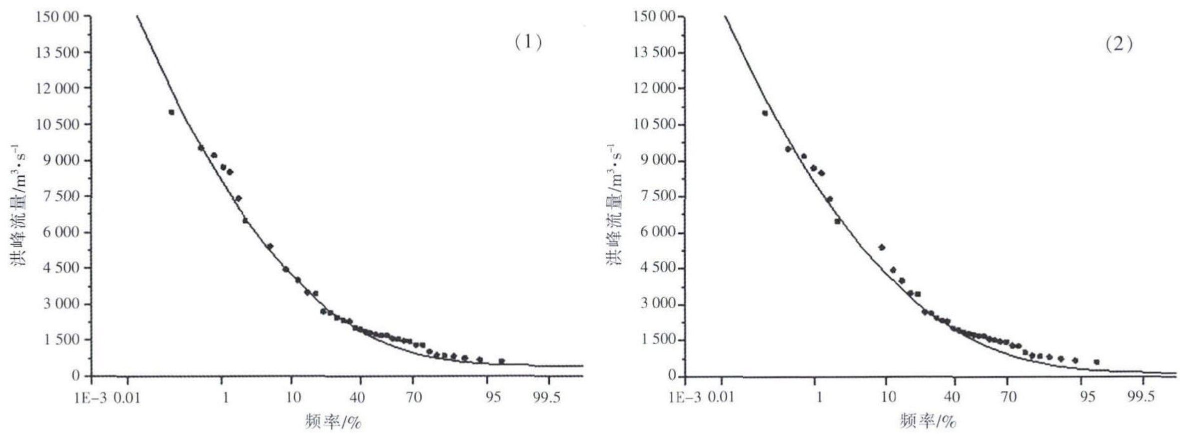

图 2 洪峰流量优化适线结果(1)统一处理法 (2)分别处理法

Fig.2 The result of optimum curve-fitting method (1) unified processing method (2) separately processing method

4 结论

(1)本文基于单调查期不连续洪水样本经验频率计算公式的概念，推导了多调查期不连续洪水样本的经验频率计算公式，同时推导了多调查期不连续样本的均值和 C_v 矩法的计算公式。

(2)基于 OLS 准则的优化适线得到的参数与矩法估计得到的参数进行比较，发现均值与 C_v 均差别不大，说明基于矩法估计多调查期不连续洪水样本参数的方法是可行的。

参考文献：

[1] 费永法. 历史特大洪水对设计洪水频率曲线参数及设计值的影响[J]. 水利发电学报，1999,(4):1-7.(FEI Yongfa. The effect of historical events on design flood[J]. Journal of Hydroelectric Engineering, 1999,(4):1-7. (in Chinese))

[2] 王善序. 具有历史洪水不连续系列经验频率的确定[J]. 人民长江, 1979,(3):39-50.(WANG Shanxu.Determination of empirical frequency of the non-simple sample with historical event[J]. Yangtze River, 1979,(3):39-50. (in Chinese))

[3] 胡进宝,张红月,黄振平. 多级历史洪水频率和参数公式的推导[J]. 人民黄河,2006,28(3):27-28.(HU Jinbao, ZHANG Hongyue, HUANG Zhenping. Derivation of the formula of frequency and parameter estimation for non-simple sample[J]. Yellow River, 2006,28(3):27-28. (in Chinese))

[4] 丁晶,杨荣富. 特大洪水时概率权重矩的计算及其应用[J].成都科技大学学报，1989,44 (2):43-54.(DING Jing, YANG Rongfu. The determination of probability weighted moments with the incorporation of extraordinary values into sample data and their application to estimating parameters for Pearson type three distribution [J]. Journal of Chengdu University of Science and Technology, 1989,44(2):43-54. (in Chinese))

[5] 宋德敦,丁晶. 概率权重矩法及其在 P-III型分布中的应用[J]. 水利

学报，1988,(3):1-11.(SONG Dedun, Ding Jing. The application of probability weighted moments to estimating the parameters for Pearson type-Ⅲ distribution [J]. Journal of Hydraulic Engineering, 1988,(3):1-11. (in Chinese))
[6] 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，2001,29(4):76-80. (CHEN Yuanfang, SHA Zhigui, CHEN Jianchi, et al. Study on L-moment estimation method for P-Ⅲ distribution with historical flood [J]. Journal of Hohai University(Natural Sciences) , 2001,29(4):76-80. (in Chinese))
[7] Martins E S, Stedinger J R. Historical information in a generalized maximum likelihood framework with partial duration and annual maximum series[J]. Water Resource Research, 2001,37(10):2559-2567.
[8] SL44-2006，水利水电工程设计洪水计算规范[S].(SL44-2006, Regulation for Calculation Design Flood of Water Resources and Hydropower Projects[S]. (in Chinese))

Formula Derivation for Frequency Analysis and Parameter Estimation of Non-simple Sample in Multi-investigation Periods

HUANG Huaping, LIANG Zhongmin

(*College of Hydrology and Water Resource, Hohai University, Nanjing* 210098, *China*)

Abstract: Adding historical floods into sequential series is a common method to improve representativity and reliability of flood series in hydrologic frequency analysis. However, the methods of empirical frequency calculation are mainly used for sequential series and discrete series with single investigation period at present. In this paper, the method of empirical frequency calculation and formula for estimating mean value based on moment method for discrete series with multi-investigation periods were presented, which have been applied in frequency analysis of the flood peak series at a station.

Key words: investigation period; non-simple sample; method of unified processing; method of separately processing; moment method