## 考虑抽样不确定性的水文设计值估计

胡义明,梁忠民,王 军,杨好周

(河海大学水文水资源学院,江苏南京 210098)

**摘要:**水文频率分析计算过程中,水文极值样本系列容量一般都较小、代表性不高,使得水文设计值估计具有不确定性。利用 Bootstrap 方法,研究样本抽样不确定性对水文设计值的影响。与传统水文频率分析方法相比,基于 Bootstrap 方法不仅可提供设计值的点估计和区间估计,同时能够对设计值的不确定性进行定量评价。此外,基于 Bootstrap 技术,结合矩法、权函数法及线性矩法,设置3套方案,分析了该方法在不同参数估计方法间的有效性。以南通市 1970—2011 年共42 年的年降雨量数据资料为例,对所提方法进行实例应用分析,结果表明,从期望设计值、90% 置信区间及最终设计值角度而言,基于所提方法的设计成果受参数估计方法的选取影响不大,且可回避规 范中 *B* 值诺莫图通用性较差及误差显著问题。

关键词:水文频率分析;设计值;Bootstrap;样本不确定性 中图分类号:P333 文献标志码:A 文章编号:1001-6791(2013)05-0667-08

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。国内外目前一般大都以概率论与数 理统计为理论基础,根据水文极值系列进行概率预估,包括样本抽样、参数估计和线型选取等环节。然而, 目前的水文极值样本系列一般较短,代表性不高;加之参数估计方法和线型选择等不确定性因素,使得水文 设计值不可避免地存在不确定性。如何定量评价这种不确定性,并在工程设计中加以合理考虑,在工程水文 设计中得到了重视和研究。

目前,针对水文设计值不确定性研究主要集中在参数估计和线型选择两方面。如在围绕降低参数估计误 差方面,分别提出了改进矩法、概率权重矩、线性矩法等方法<sup>[1-5]</sup>;Wood和Rodriguze-iturbe<sup>[6-7]</sup>最早基于贝 叶斯理论分析了水文统计模型参数估计的不确定性及其对设计洪水的影响;还提出了一套同时考虑线型选择 和参数估计不确定性的理论框架,但由于计算条件限制,当时仅分析了几种简单线型(如正态分布、对数正 态分布)水文设计值的不确定性问题;Kuczera<sup>[8]</sup>提出了应用贝叶斯理论结合单站信息和地区信息推求设计洪 水,以降低估计的不确定性。在中国,关于水文设计值不确定性研究相对较晚。刘攀等<sup>[9]</sup>利用贝叶斯因子 法研究了水文频率分析中线型选择的不确定性问题;梁忠民等<sup>[10]</sup>和桑燕芳等<sup>[11]</sup>基于贝叶斯理论定量描述分 布参数的不确定性,进而分析了水文设计值的不确定性。梁忠民等<sup>[12]</sup>综合分析了国内外关于频率分析不确 定性的研究成果,并根据贝叶斯理论,对只考虑参数估计或线性选择不确定性以及同时考虑这两种不确定性 情况下的水文设计值问题进行研究,提出设计值估计及其不确定性评价方法<sup>[13]</sup>。然而,关于样本抽样不确 定性对水文设计值不确定性影响的研究鲜见报道。考虑到水文频率分析过程中,采用的水文极值样本系列通 常只有40~50年的长度,样本代表性不高,使得水文设计值估计具有不确定性。本文提出采用Bootstrap方 法开展样本抽样不确定性对水文设计值不确定性的影响研究,在提供设计值点估计和区间估计的同时,定量 评价设计值的不确定性;并将该方法应用于南通市年降雨系列的频率分析中。

收稿日期: 2012-11-15; 网络出版时间: 2013-08-14

网络出版地址: http://www.cnki.net/kcms/detail/32.1309. P. 20130814.1037.003. html

**基金项目:**国家自然科学基金资助项目(51190095; 51079039)

作者简介:胡义明(1986-),男,江苏宿迁人,博士研究生,主要从事水文水资源方面研究。 E-mail: hymkyan@163.com

### **1** Bootstrap 方法概述

Bootstrap 方法是美国 Stanford 大学统计系教授 Efron 于 1979 年提出的一种统计推断方法<sup>[14]</sup>。该方法不需要对总体分布做假设,仅基于原始样本系列进行再抽样,并不断估计抽出样本的参数值,来推断未知总体的参数特性,并可定量描述参数估计的不确定性。

假设  $X_1$ ,  $X_2$ , …,  $X_n$  为独立同分布的随机变量,  $X_i \sim F(x)$ ;  $\theta = \theta(F)$  为总体分布的未知参数;  $X_1$ ,  $X_2$ , …,  $X_n$  的联合分布为  $F_n$ ;  $\hat{\theta} = \hat{\theta}(F_n)$  为总体参数  $\theta$  的估计。记

$$T_{n} = \hat{\theta}(F_{n}) - \theta(F) \tag{1}$$

从  $F_n$  中抽取相同容量的样本,获得再生样本  $X^* = (X_1^*, X_2^*, \dots, X_n^*)$ ,样本  $X_1^*, X_2^*, \dots, X_n^*$ 的联 合分布为  $F_n^*$ ,则由样本  $X^*$ 作出的  $\theta$  估计为  $\hat{\theta}(F_n^*)$ 。记

$$R_n^* = \hat{\theta}(F_n^*) - \theta(F_n) \tag{2}$$

 $R_n^*$  为  $T_n$  的自助统计量, 以  $R_n^*$  的分布去近似  $T_n$  的分布, 这就是自助法的中心思想。

从分布  $F_n$  中重复抽取相同容量的样本 C 组,以

$$X^{*(j)} = (X_1^{*(j)}, X_2^{*(j)}, \cdots, X_n^{*(j)}), j = 1, 2, \cdots, N$$
(3)

表示第j个再生样本。每个再生样本可分别求得 $R_n^*$ ,记为 $R_n^{*(j)}$ 

$$R_{n}^{*(j)} = \hat{\theta}(F_{n}^{*(j)}) - \theta(F_{n}), j = 1, 2, \cdots, N$$
(4)

对于每个  $R_n^{*(j)}$ , 可得  $\theta(F)$  的近似值, 记为  $\hat{\theta}^{(j)}(F)$ , 即

$$\theta^{(j)}(F) = \hat{\theta}(F_n) - T_n \approx \hat{\theta}(F_n) - R_n^{*(j)}, \ j = 1, 2, \ \cdots, \ N$$
(5)

通过 N 组样本,可获得总体参数 θ(F)的 N 个估计值,将其作为 θ(F)的样本,获得 θ(F)的抽样分布, 据此分布可求得参数的点/区间估计,并可定量评价设计值的不确定性。

#### 2 基于样本抽样不确定性的水文设计值推求

水文频率分析过程中,采用的水文极值样本系列一般较短,通常只有40~50年的资料系列,样本对总体的代表性存在偏差,使得基于给定样本系列推断出的总体参数存有较大不确定性,进而影响水文设计值的可靠性。为此,将借助上述的Bootstrap方法,对原始样本系列进行重抽样,获得总体参数的分布,进而研究给定频率条件下的水文设计值的分布,最终获得给定条件下水文设计值的点估计和区间估计,并定量评价设计值的不确定性。在此过程中,参数估计及线型选取的不确定性暂不考虑。

现以中国常用的 P-III型线型为例介绍该方法。设已知水文极值样本系列  $x_1, x_2, \dots, x_n; \bar{X}, C_v, C_s$ 分别为总体的均值、变差系数和偏态系数;  $\hat{\bar{X}}^{(j)}, \hat{C}_v^{(j)}, \hat{C}_s^{(j)}$ 分别为第 j个样本关于总体均值、变差系数和偏态系数;  $\hat{\bar{X}}^{(j)}, \hat{C}_v^{(j)}, \hat{C}_s^{(j)}$ 分别为第 j个样本关于总体均值、变差系数和偏态系数的估计值。

采用 Bootstrap 方法,从原始序列中重复抽样,获得 N 组相同容量的样本  $X^{(j)} = (X_1^{(j)}, X_2^{(j)}, \dots, X_n^{(j)}),$  $j = 1, 2, \dots, N;$  采用稳健的线性矩法<sup>[4]</sup>估计每一组样本的 $\hat{X}^{(j)}, \hat{C}_{v}^{(j)}, \hat{C}_{s}^{(j)}, j = 1, 2 \dots, N,$ 可分别获得 总体  $\bar{X}, C_v, C_s$  的 N 个估计值,并将此 N 个估计值的分布  $F_{E}(\bar{X}), F_{E}(C_v), F_{E}(C_s)$ 作为总体均值、变差 系数和偏态系数的分布。

从分布  $F_{\rm E}(\bar{X})$ ,  $F_{\rm E}(C_{\rm v})$ ,  $F_{\rm E}(C_{\rm s})$  中抽取 N 组参数  $\theta^{(i)} = (\bar{X}^{(i)}, C_{\rm v}^{(i)}, C_{\rm s}^{(i)})$ , i = 1, 2, ..., N, 据每一组 参数,可获得某一给定频率 p 条件下的设计值  $x_p^{(i)}$ , i = 1, 2, ..., N, 将 N 个设计值的分布  $G_{\rm E}(x_p)$ 作为设计 值  $x_p$  的分布,进而获得频率 p 条件下水文设计值的点估计和区间估计。但考虑到给定样本条件下 X,  $C_v$ ,  $C_s$  的相依性及为了简化计算,本文采用如下策略进行给定频率 p 条件下水文设计值的推求。步骤如下:

(1)首先依据原始样本系列 x<sub>1</sub>, x<sub>2</sub>, …, x<sub>n</sub>,做有放回抽样,每次生成一个[1,n]区间上均匀分布的随机数 τ,则获得的新样本点为 x<sub>r</sub>,重复此过程 n 次,可得到容量为 n 的新样本。重复上述过程 N 次,即可获得 N 组容量为 n 的新样本,X<sup>(j)</sup> = (X<sup>(j)</sup><sub>1</sub>, X<sup>(j)</sup><sub>2</sub>, …, X<sup>(j)</sup><sub>n</sub>), j=1, 2, …, N。

(2) 若原始序列中存有特大值,抽样方法同样采用(1)中的策略。当新获得的样本中无特大值时,作常规连续样本处理;若新获得的样本中存有特大值,作不连续样本处理。

(3) 采用稳健的线性矩法估计(1)或(2)中*N*组样本的均值、离差系数和偏态系数,获得*N*组参数 $\theta^{i} = (\bar{X}^{(i)}, C_{v}^{(i)}, C_{s}^{(i)}), i = 1, 2, \dots, N_{o}$ 

(4) 基于(3)中的第*i* 组参数  $\theta^{(i)}$ ,可获得给定频率 *p* 条件下的水文设计值  $x_p^{(i)}$ 。

(5) 将 N 个给定频率 p 条件下的水文设计值  $x_p^{(i)}$ , i = 1, 2, ..., N 的经验分布作为  $x_p$  的分布,进而推断  $x_p$  的点估计和区间估计,并定量分析设计值  $x_p$  的不确定性。

3 实例应用

以南通市 1970—2011 年共 42 年的年降雨量数据资料为例,对本文所提方法进行实例应用分析。图 1 为 年降雨时间序列图,由图 1 可知,年降雨系列无明显变异且无特大值。另外,在 5% 的显著性水平下, Mann-Kendall 法检验结果同样表明该降雨量系列无明显变异(图 2)。



采用第 2 节所述方法,对原始样本进行重复抽样,考虑到 P-III分布中,当  $C_s/C_v < 2$  时,其下限为负数,不符合水文特征,在抽样时,选取满足  $C_s/C_v \ge 2$  的样本。文献 [15] 指出,如要相对误差小于 5%,抽样组数在50~200 即可,本文设定抽样组数 N = 1000。采用稳健的线性矩法估计每组样本的参数  $\theta^{(i)} = (\bar{X}^{(i)}, C_v^{(i)}, C_s^{(i)}), i = 1, 2, ..., N;$ 结合 P-III分布,求得给定参数  $\theta^{(i)}, i = 1, 2, ..., N,$ 频率 p 的年降 雨设计值  $x_p^{(i)}, i = 1, 2..., N_o$ .通过统计分析频率 p 条件下的 N 个估计值,进而推求  $x_p$  的点估计和区间估计,并定量分析设计值  $x_p$  的不确定性。

图 3(a)~图 3(f)分别给出了频率 0.1%、0.5%、1%、2%、5% 和 10% 条件下设计值的频率直方图。可以看出,设计值的频率直方图与正态分布函数拟合较好。

图 4(a) ~图 4(f)提供了频率 0.1%、0.5%、1%、2%、5% 和 10% 条件下设计值的正态概率图。可以 看出,各频率条件下,散点基本上均匀分布在红色直线(45°直线)附近,表明各频率条件下设计值的分布近 似服从正态分布。另外,对上述给定频率的设计值进行了 χ<sup>2</sup>(卡方) 拟合优度检验(显著性水平 α = 0.05), 检验结果同样表明各频率条件下设计值的分布近似服从正态分布。





表1列出了频率0.1%、0.5%、1%、2%、5%和10%条件下,采用本文所提方法计算获得的设计值分布、期望值、90%置信区间及设计值的标准差。

表 1	给定频率下设计值估计结果
-----	--------------

Table 1 Estim	ation of design	values	corresponding	to	given	probability
---------------	-----------------	--------	---------------	----	-------	-------------

频率	本文方法(线性矩法 + Bootsrtap 抽样技术)									
/%	设计值分布	期望设计值/mm	90% 置信区间估计/mm	区间宽度/mm	标准差					
0.1	$X_{0.1\%} \sim N(1974, 121)$	1 974	[ 1 774 ,2 173 ]	399	121					
0.5	$X_{0.5\%} \sim N(1786,93)$	1 786	[1633,1938]	305	93					
1	$X_{1\%} \sim N(1\ 701\ ,81)$	1 701	[1567,1834]	267	81					
2	$X_{2\%} \sim N(1\ 611\ ,70)$	1 611	[1496,1726]	230	70					
5	$X_{5\%} \sim N(1\ 485\ ,57)$	1 485	[ 1 391 ,1 578 ]	187	57					
10	$X_{10\%} \sim N(1380,48)$	1 380	[1 300,1 459]	159	48					

从表1可以看出,随设计频率的减小(即设计标准的提高),设计值分布的标准差在增大,且90%置信区间的宽度也在增加,表明设计值的不确定性随着设计频率的减小而增加。

在工程水文设计中,考虑工程安全,通常在设计值上再加一个安全修正值<sup>[10]</sup>,以确定最终设计值。安 全修正值根据设计值的标准差进行折算。

$$\Delta X_p = \alpha \sigma_{x_p} \tag{6}$$

标准差一般采用下式近似估算:

$$\sigma_{z} = \bar{x}C_{y}B/\sqrt{n} \tag{7}$$

式中  $\alpha$  为可靠性系数,通常取 0.7;  $\sigma_{x_p}$ 为设计值的标准差;  $\bar{x}$  为均值;  $C_v$  为偏差系数; n 为样本容量; B 为  $C_s$  和设计频率 p 的函数,通过查 B 值诺莫图获得。

根据本文方法和传统方法,得到的最终设计结果见表2。

	Table 2   Calculation of safety factor										
ilst strategy		本文方法			传统方法						
<b></b>	$x_p/mm$	$\Delta x_p / \text{mm}$	$x_p + \Delta x_p / \text{mm}$	$x_p/mm$	$\Delta x_p / \text{mm}$	$x_p + \Delta x_p / \text{mm}$	相利差值/%				
0.1	1 973	85	2 058	1 816	136	1 952	5.2				
0.5	1 786	65	1 851	1 685	112	1 797	2.9				
1	1 701	57	1 758	1 623	83	1 706	3.0				
2	1 611	49	1 660	1 556	75	1 631	1.8				
5	1 485	40	1 525	1 458	51	1 509	1.1				
10	1 379	34	1 413	1 372	39	1 411	0.1				

表 2 安全修正值计算结果

表 2 中的数据表明,基于传统方法近似估算的修正值,较本文所提方法稍微偏小,相对误差值在 0.1%~5.2%之间,差别不大。但公式(7)中 B 值的大小与具体采用的参数估算公式有关,如采用矩法和线 性矩法时, B 值将不一样, B 值诺莫图的通用性较差,而目前规范中,仅给出了矩法下的 B 值诺莫图; 另 外,当 C<sub>s</sub> 值大于 2 及频率较小时,差异显著<sup>[16]</sup>。而本文方法算法简单,在获得频率 p 的期望设计值的同时,也可获得其标准差,可直接采用式(6)进行修正值的计算,较式(7)的估算方法更为方便有效。

为了说明本文方法设计成果受参数估计方法选取影响不大,以及在不同参数估计方法间的有效性,基于 上述所抽1000组样本,设计3套方案:① 矩法估计参数,结合 Bootstrap 抽样技术进行设计值计算;② 权 函数法估计参数<sup>[5]</sup>,结合 Bootstrap 抽样技术进行设计值计算;③ 线性矩法估计参数,结合 Bootstrap 抽样技 术进行设计值计算(表1)。

在方案①和方案②情况下(方案③成果已有),通过分析频率0.1%、0.5%、1%、2%、5%和10%条件下,设计值的频率直方图、正态概率图表明,方案①和方案②条件下各频率设计值分布近似服从正态分布。 具体计算成果见表3、表4。

		5		• –	
频率/%	设计值分布	期望设计值/mm	90% 置信区间估计/mm	区间宽度/mm	标准差
0.1	$X_{0.1\%} \sim N(1920,95)$	1 920	[1763,2076]	313	95
0.5	$X_{0.5\%} \sim N(1.745,75)$	1 745	[1622,1868]	246	75
1	$X_{1\%} \sim N(1.665, 67)$	1 665	[1555,1774]	219	67
2	$X_{2\%} \sim N(1581,59)$	1 581	[1484,1677]	193	59
5	$X_{5\%} \sim N(1461,49)$	1 461	[1 380,1 541]	161	49
10	$X_{10\%} \sim N(1361,42)$	1 361	[ 1 291 ,1 430 ]	139	42

表 3 方案①(矩法 + Bootsrtap 抽样技术)条件下给定频率设计值估计结果 Table 3 Estimation of design values corresponding to given probability under scheme ①

从表 3、表 4 可以看出,随着设计标准的提高,设计值分布的标准差及 90% 置信区间的宽度均在增加, 表明设计值的不确定性随着设计标准的提高而增加。该结论与表 1 的结论一致。

	Table 4 Estimation of design values corresponding to given probability under scheme (2)										
频率/%	设计值分布	期望设计值/mm	90%置信区间估计/mm	区间宽度/mm	标准差						
0.1	$X_{0.1\%} \sim N(1959, 121)$	1 959	[1760,2158]	398	121						
0.5	$X_{0.5\%} \sim N(1.771,92)$	1 771	[1620,1921]	301	92						
1	$X_{1\%} \sim N(1.685, 80)$	1 685	[1554,1816]	262	80						
2	$X_{2\%} \sim N(1595,69)$	1 595	[1482,1708]	226	69						
5	$X_{5\%} \sim N(1\ 468\ 55)$	1 468	[1377,1559]	182	55						
10	$X_{10\%} \sim N(1363,47)$	1 363	[1286,1440]	154	47						

表 4 方案② (权函数法 + Bootsrtap 抽样技术)条件下给定频率设计值估计结果 Table 4 Estimation of design values corresponding to given probability under scheme ②

通过对比分析方案①~方案③ 计算的安全修正值与传统方法计算值(表 5)可知,传统方法的计算结果 与3套方案的计算结果较为一致。

频率		$x_p + \Delta$	x <sub>p</sub> /mm					
/%	传统方法	方案①	方案②	方案③	方案①与传统方法	方案②与传统方法	方案③与传统方法	
0.1	1 952	1 987	2 044	2 058	1.8	4.7	5.4	
0.5	1 797	1 798	1 835	1 851	0.1	2.1	3.0	
1	1 706	1 712	1 741	1 758	0.4	2.1	3.0	
2	1 631	1 622	1 643	1 660	-0.6	0.7	1.8	
5	1 509	1 495	1 507	1 525	- 0. 9	- 0. 1	1.1	
10	1 411	1 390	1 396	1 413	- 1. 5	- 1. 1	0.1	

表 5 不同方案下最终设计值计算结果对比分析 Table 5 Comparative analysis of safety value under different schemes

表 6 对比分析了 3 套方案(基于 Bootstrap 抽样技术)条件下,期望设计值、90% 置信区间、区间宽 度及标准差间的差异。可以看出,3 套方案的期望设计值、90% 置信区间的上下限差别均不大,尤其 是方案②与方案③,其期望设计值、90% 区间的上下限、区间宽度及标准差指标几乎一致。方案①的 标准差和区间宽度较其他 2 套方案偏小,可能是由于采用矩法估计参数 *C*<sub>s</sub>时不确定性较大,而 Bootstrap 抽样技术,在一定程度上对其不确定性反向消减。但方案①的期望设计值及 90% 置信区间的上下 限和方案②、方案③差异微小;相对于方案③,期望设计值差幅仅为 - 2.7% ~ -1.3%,90% 置信区 间上下限差异分别为 - 0.8% ~ -0.6% 和 - 4.5% ~ -2.0%。就期望设计值和 90% 置信区间而言,3 套方案设计成果无明显差异。另外,表5提供的不同方案下的最终设计值(*x<sub>p</sub>* + Δ*x<sub>p</sub>*)也无明显差别。为 此,从期望设计值、90% 置信区间及最终设计值角度而言,可认为基于本文方法的设计成果受参数估 计方法的选取影响不大。

频率		方案①	与方案③相对	付差值/%		方案②与方案③相对差值/%				
/%	期望设计值	下限	上限	区间宽	标准差	期望设计值	下限	上限	区间宽	标准差
0.1	- 2. 7	-0.6	- 4. 5	- 21. 7	-21.7	- 0. 7	-0.7	- 0. 7	-0.4	-0.5
0.5	-2.3	-0.7	- 3.6	- 19. 2	- 19. 1	-0.8	-0.8	- 0. 9	- 1. 3	-1.4
1	-2.1	-0.8	- 3.2	- 17. 7	- 17.8	- 0. 9	-0.9	- 1.0	-1.7	-1.7
2	- 1.9	-0.8	- 2. 9	- 16. 3	- 16. 2	-1.0	- 0. 9	- 1. 1	-2.1	- 2. 1
5	- 1.6	-0.8	- 2. 3	- 13. 9	- 14. 0	- 1. 1	-1.0	- 1. 2	-2.5	- 2.6
10	- 1. 3	-0.7	- 2. 0	- 12. 5	- 12.4	-1.2	-1.1	- 1. 3	-2.7	- 2. 8

表 6 3 套方案设计成果对比分析 Table 6 Comparative analysis of safety value under schemes ① to schemes③

图 5 提供了采用方案③与经验适线法获得的频率曲线成果。可以看出,随着设计频率的减小,设计值的 不确定性增加(90%置信区间的宽度在增加);但绝大多数实测点据均置于 90% 的置信区间内。



图 5 本文方法与传统方法的频率分析结果

Fig. 5 Frequency analysis results of Bootstrap method and empirical curve-fitting method

#### 4 结 论

(1)本文提出采用 Bootstrap 重抽样技术,研究样本抽样不确定性对水文设计值的影响。该方法可获得 设计值的分布及统计参数(如标准差等),在提供设计值的点估计和区间估计的同时,可定量评价设计值的 不确定性。

(2)中国目前主要采用经验方法对设计值进行安全修正,且目前规范中,仅给出了矩法下的 B 值诺奠图。考虑到公式 B 值的大小与具体采用的参数估算方法有关,如采用矩法和线性矩法时, B 值将不一样, B 值诺莫图的通用性较差,且当 C<sub>s</sub> 值大于 2 及概率较小时,差异显著。本文基于 Bootstrap 技术,结合矩法、权函数法及线性矩法,设置 3 套方案,分析了该方法在不同参数估计方法间的有效性。结果表明,从期望设计值、90% 置信区间及最终设计值角度而言,基于本文所提方法的设计成果受参数估计方法选取影响不大,可回避规范中 B 值诺莫图通用性较差及误差显著问题。

(3)目前,本文仅讨论了特定容量下常规连续样本,关于不同调查期、不同实测样本容量及特大值组 合条件下的不连续样本的讨论,有待进一步研究。

#### 参考文献:

- [1] WALLIS J R, MATALS N C, SLACK J R. Just a moment! [J]. Water Resources Research, 1974, 10:211-219.
- [2] GREENWOOD J A, LANDWEHR J M, MATALS N C, et al. Probability weighted moments: Definition and relation to parameters of several distributions expressible in inverse form [J]. Water Resources Research, 1979, 15:1049-1054.
- [3] DING J, SONG D, YANG R. Further research on application of probability weighted moments in estimating the parameters of the Pearson type three distribution [J]. Water Resources Research, 1989, 110:239-257.
- [4] HOSKING J R M. L-moments: Analysis and estimation of distribution using linear combinations of order statistics [J]. Journal of the Royal Statistical Society: Series B, 1990, 52: 105-124.
- [5] 马秀峰. 计算水文频率参数的权函数法[J]. 水文, 1984(3):1-8. (MA Xiufeng. The weighted function method for hydrological frequency parameter estimation [J]. Journal of China Hydrology, 1984(3): 1-8. (in Chinese))
- [6] WOOD E F, RODRIGUZE-ITURBE I. Bayesian inference and decision making for extreme hydrologic events [J]. Water Resources Research, 1975, 11(4): 533-542.
- [7] WOOD E F, RODRIGUZE-ITURBE I. Bayesian approach to analyzing uncertainty among flood frequency models [J]. Water Resources Research, 1975, 11(6): 839-843.
- [8] KUCZERA G. Combining site-specific and regional information: An empirical Bayes approach [J]. Water Resources Research, 1982, 18(2): 306-314.
- [9] 刘攀, 郭生练, 田向荣,等. 基于贝叶斯理论的水文频率线型选择与综合[J]. 武汉大学学报:工学版, 2005, 38(5): 36-40. (LIU Pan, GUO Shenglian, TIAN Xiangrong, et al. Selecting and averaging of flood frequency models based on Bayesian theory

[J]. Engineering Journal of Wuhan University, 2005, 38(5): 36-40. (in Chinese))

- [10] 梁忠民, 戴荣, 雷杨, 等. 基于贝叶斯理论的水文频率分析方法[J]. 水力发电学报, 2009, 28(4): 22-26. (LIANG Zhongmin, DAI Rong, LEI Yang, et al. Hydrological frequency analysis based on Bayesian theory [J]. Journal of Hydroelectric Engineering, 2009, 28(4): 22-26. (in Chinese))
- [11] 桑燕芳, 王栋, 吴吉春. 基于贝叶斯理论的水文线型参数不确定性分析[J]. 水电能源科学, 2009, 27(6): 15-19. (SANG Yanfang, WANG Dong, WU Jichun. Parameter uncertainty analysis on hydrologic frequency models based on Bayesian theory[J]. Water Resources and Power, 2009, 27(6): 15-19. (in Chinese))
- [12] 梁忠民.戴荣.李彬权.基于贝叶斯理论的水文不确定性分析研究进展[J].水科学进展,2010,21(2):274-281.(LIANG Zhongmin, DAI Rong, LI Binquan. Areview of hydrological uncertainty analysis based on Bayesian theory[J]. Advances in Water Science, 2010,21(2):274-281.(in chinese)).
- [13] 梁忠民,李磊,王军,等.考虑参数和线型不确定性的水文设计值估计的贝叶斯方法[J].天津大学学报,2010,43(5): 379-384.(LIANG Zhongmin, LI Lei, WANG Jun, et al. Bayesian method for hydrological frequency analysis considering uncertainties of parameter and model[J]. Journal of Tianjin University, 2010,43(5): 379-384. (in Chinese))
- [14] EFRON B, TIBSHIRANI R. An introduction to the bootstrap[M]. New York: Chapman & Hall Ltd, 1993.
- [15] EFRON B. Better bootstrap confidence intervals[J]. J Amer Statist Assoc, 1987, 82: 171-200.
- [16] 水利部长江水利委员会,水利部南京水文水资源研究所.水利水电工程设计洪水计算手册[M].北京:中国水利水电出版 社,2001: 84-86. (Hydrological Bureau of Changjiang Water Resources Commission Ministry of Water Resources, Nanjing Institute of Hydrology and Water Resources Ministry of Water Resources. Handbook for calculating design flood of water resources and hydropower projects[M]. Beijing: China WaterPower Press, 2001: 84-86. (in Chinese))

# Determination of design hydrologic characteristics with sampling uncertainty considerations \*

HU Yiming, LIANG Zhongmin, WANG Jun, YANG Haozhou

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

**Abstract**: A common problem in hydrologic frequency analysis is that the sample size is often too small to adequately represent the statistical property of an entire population. Uncertainties can thus be introduced during the determination of design hydrologic characteristics. In this paper, a new method based on Bootstrap is used to estimate the influence of sample uncertainties on the design characteristics. Compared to conventional approaches, the new method provides not only point estimates and interval estimation for the design characteristics, but also a quantitative evaluation of uncertainties associated with hydrological frequency analysis. In addition, combing with the Bootstrap technique, three schemes are established for assessing the effectiveness of the new method when using different parameter estimation approaches. The schemes are the moment method, the weight function method and the linear-moment method. The Nantong annual rainfall data from 1970 to 2011 is used to validate the new method. The result shows that with the new method, the selection of different parameter estimation approaches will have an insignificant impact on the determination of design hydrologic characteristics. This is from the point of view of the expected value, 90% confidence intervals and the final design characteristics. The new method can also deal with the limitation of the *B* value.

Key words: hydrologic frequency analysis; design characteristics; Bootstrap; sampling uncertainty

<sup>\*</sup> The study is financially supported by the National Natural Science Foundation of China (No. 51190095; No. 51079039).