

(12)

60-62

洪水

频率分析

分布

评价

计算

第15卷第1期

水利水电科技进步

1995年2月

洪水频率分析中各种分布的评价

赵贵德 编译

(北京市水文总站 北京 100038)

TV 122
P 333.2

近20年来,研究者们提出了众多概率分布(PDs)和参数估算方法,最优分布线型的选择一直为人们所关注。这里所介绍的是另一个方面,即各种概率分布和参数计算方法之间相互比较的问题^[1]。为此,编制了一个计算机软件包。它包括9种分布:①二参数对数正态分布(LN2);②三参数对数正态分布(LN3);③耿贝尔(Gumbel)普通极值分布(GEV);④皮尔逊Ⅲ型分布(P3);⑤对数皮尔逊Ⅲ型分布(LP3);⑥对数-逻辑分布(LL);⑦Wakeby分布;⑧二步指数转换正态化分布(TSPT);⑨LN3为零偏差(LN3-CS₂=0)分布。概率分布中最常用的参数计算方法是最大似然法(ML)和概率-加权矩法(PWM)。混合矩法和熵法也应用于LP3分布中。对于LN3分布计算其位置参数、使变换变量 x (这里 $x=\ln(Q-C)$; Q 为年最大洪峰值, C 为位置参数)的分布的偏态系数准确地等于零($CS_2=0$)。当相同分布用不同的参数计算方法时,这里也作为一种“不同的模型”,这样总共有19种模型,即:①LN2-ML;②Gumbel-ML;③Gumbel-PWM;④Gumbel-熵;⑤TSPT;⑥Log-logistic-ML(LL-ML);⑦LL-PWM;⑧P3-ML;⑨P3-PWM;⑩LP3-ML;⑪LP3-PWM;⑫LP3-MM1;⑬LP3-熵;⑭LN3-ML;⑮LN3-Kappenman;⑯LN3-CS₂=0;⑰GEV-ML;⑱GEV-PWM;⑲Wakeby-PWM。

三种传统的拟合优度检验(GOF): χ^2 检验, K-S检验和CvM检验也包括在此软件包内。文献[1]首先对19个模型进行拟合优度的评价,即直接对13条天然河流长度在30年以上的实测洪峰流量系列加以检验。其次对其中6条河随机生成了长达100 000年的系列,采用的总体分布为GEV和Wakeby。再者,对从总体分布中计算出的重现期分别为100、1000和10 000年的设计洪水值,基于样本容量分别为30和100的1000组随机生成样本以及样本容量为500的200组样本的19种模型的相应估计值进行了对比。并对相对误差和平方误差直方图进行了讨论。

关于概率分布,洪水频率分析中各种概率分布的有关详细资料及诸多参数的计算方法,其参考书很多,这里只简述一下有关重点。

GEV参数的计算,特别是对数-逻辑分布的参数计算(应用ML法)是非常困难的,而且耗时太多。采用ML和熵法估计耿贝尔参数,要求求解两个联立非线性方程,直接采用牛顿迭代算法总能求出结果,且工作量也不太大。在计算中,用PWM法所求出的参数值作为ML法参数估计的初值,随后再作为熵法的初值。

用ML法所计算和P3和LP3参数,并不像LL计算的参数那么麻烦,用MM1法计算的LP3参数类似于Arora和Singh于

1989 年建议的方法。LP3 参数也可用熵法解出, 只是这个方法要求用迭代法运算, 它类似于 ML 法。熵法总是产生两套参数, 采用的一套其偏态系数与 ML 法有相同的符号, 而另一套则舍弃不用。

因为所采用的 13 条河其年洪峰系列的直方图都是向右偏, 又因这些直方图中的三个在右边的尾端展示出第二峰轻微的不规则性, 所以实际转换概率导致其变量值为对称性直方图, 对此也进行了检验。首先应用了二步幂变换 (TSPT)。其次, 尽管它不是很普遍地被采用, 但 TSPT 的根据是著名的 Box-Cox 转换, 在 TSPT 中, 将原始 Q 系列转换成 x 变量系列, 使 x 系列的偏差系数和峰度系数 (CS_x 和 CE_x) 都完全等于零, 其理由是, 当这两个系数等于零时, 其正态分布应与 x 系列拟合得很好。借助于 TSPT 计算 $Q_T (-) T$ 相关关系的必要公式已由 Gupta 等人于 1987 年给出。通过采用迭代法求解方程组 $CS_x=0$ 和 $CE_x=0$, 可得变换所要求的两个参数值。

与 ML 法一起, 另外两个方法也用于 LN3 分布的参数估计: 一个是 Kappenman 法; 另一个是确定位置系数法, 以使变换变量 x ($x=\ln(Q-c)$) 的偏态系数精确等于零。这最后一种方法在此之前并未采用, 希望这样确定 LN3 分布应该拟合得更好, 因为这表明 x 变量遵守标准的二参数正态分布, 其 CS 应等于零。恰如其总体分布的 CS , 这里 LN3- $CS_x=0$ 模型被认为是正态化的另一种转换。

应用有关资料和 19 种洪水频率分析方法, 再用三种传统的拟合优度检验, 可得出如下结论:

a. Wakeby-PWM 模型最好, TSPT 次之, 以后的次序是: LL-PWM, GEV-PWM, LL-ML, GEV-ML, LP3-熵, LP3-ML, P3-PWM。

b. 耿贝尔分布的箱形图分布最窄, 但多

次低估了右端尾部值。三个耿贝尔模型 (Gumbel-ML; Gumbel-PWM; Gumbel-entropy) 的右端值 Q_{Ts} 彼此很靠近, 当以 Wakeby 为总体分布时, 耿贝尔模型大多是低估了右端尾部值。当 GEV-Ⅰ型 (作为期望值) 作为基本分布模型时, 耿贝尔也总是估计偏低, 而且 GEV-Ⅲ型 (再作为期望值) 作为总体分布时, 耿贝尔又估计偏高。

c. 类似地, 大多数场合下 LN2 分布的估计值比正确预测值偏低。

d. TSPT 分布的 RE_s (相对误差) 分布很宽。偏离零位很远, 此模型计算困难, 稀遇洪水的计算很难进行, 这可能是由于这些系列有高偏态系数, 受其截断或累积误差影响所致。因此, 尽管它在传统的检验中较好, 但 TSPT 仍不值得推荐。

e. GEV-PWM 模型的估值与 100, 1000, 10 000 年重现期的总体值 Q_{Ts} 相比, 有稍稍偏大的趋势。GEV-PWM 分布的偏差始终紧靠零。而其它模型的偏差则或正或负波动很大, 因此 GEV-PWM 模型看起来在较保守的设计中值得推荐。

f. 所有 P3 和 LP3 模型与它们曾估算偏高相比, 都低估了高重现期基本 Q_{Ts} (Q_{Ts} (某重现期流量))。只有 LP3-PWM 例外。LP3-PWM 模型与其它几个 LP3 模型相比, 所得出的极右端 Q_T 值一般较高。总的说来, 在 $\ln(Q)$ 系列中, 所采用偏态系数愈大, LP3 模型的右端, Q_{Ts} 值愈高。一个有负值 CS_x (x 系列的偏态系数) 的 LP3 分布, 如 -0.7 或 -0.9 就会产生一个频率曲线, 其右尾部以低速增加。因此, 保守而小心的设计总是避免使用具有大负值 CS 的 LP3 模型 (可应用于低等设计)。

g. 应用 ML 法求解 LL 概率分布参数非常困难, 且耗时多。此模型大大高估了总体分布的右尾 Q_{Ts} 值。当偏态系数较大时, 此法甚至不能求出一个收敛结果。可以认为

LL-ML 模型在任何情况下都不宜应用。应用 PWM 法求得的 LL 模型参数, 计算虽容易, 但右端的 Q_{Ts} 仍偏高。LL-PWM 分布的偏高程度比 LL-ML 分布所求出的 Q_{Ts} 要小得多, 所以 LL-PWM 模型在 T' (重现期) 达到 100 年时还是可以应用的。

h. GEV-ML 与 GEV-PWM 很相似, GEV-ML 的似然值可得一个收敛解, 它比 LL-ML 模型所求结果高很多。

i. 所有三个传统的 GOF 检验结果, 它们之间有很好的 consistency。但这些 GOF 检验的结果与 Q_T 的比较分析并不一致。因此, 传统的 GOF 检验并不能真正的判别一个特别模型, 对严格的拟合优度而言, 不能只把注意力集中在研究概念上, 应该注意的是 100 或更多年重现期右端各种不同概率分布的性质。如考虑到 GEV 模型, 如果其结果形态是 III 型, 那么, 一个保守而慎重的分析者就应采用耿贝尔分布, 而不用 GEV-I 型。同样, 当 LP3 概率分布的偏态系数为负值时, 应该用 LN3, 而不用 LP3 模型。

j. LN3 模型稍稍低估了基本分布的右端 Q_{Ts} 值, 估计偏高的情况较少。而运算情况好于 LP3 分布。所有 LN3 模型彼此都非常接近。特别是 LN3- $CS_s=0$ 分布时, 得出的 Q_{Ts} 值相当接近于 LN3-ML 值。因此, $CS_s=0$ 法对这种特殊频率分布来说, 不会产生明显改善。当 ML 法不能求出一个稳定解

时, $CS_s=0$ 法可应用, 而不是 ML 法。LN3-Kappenman 法未形成一种可信服的改进。其它两个 LN3 模型也是如此, 在任何情况下, 应用这种特殊的方法计算位置参数都不容易。然而 RE 箱形图相互比较后, 明显发现 LN3 概率分布比 LN2 要好得多。

k. 从 Wakeby 分布与 18 个其它模型的箱形图上看, 未展示出它的优势, 但它是较好的分布之一。在作区域频率分析时, Wakeby 分布更为合适, 应用单站概念分布则欠妥。

l. 不管什么样的概率分布, ML 的计算工作量都比 PWM 法大得多。另外, PWM 模型看起来比 ML 模型有时差不多, 有时好一些。不管概率是什么分布类型, PWM 法好于 ML 法。

m. 用计算机计算各种模型都很方便, 但最后究竟应使用哪种模型应根据经验和对工程的判断而定, 洪水频率分析应该采用已证实模型的区域化参数予以计算。

参考文献

- 1 Tefaruk Haktanir, et al. Evaluation of various distributions for flood frequency analysis. Hydrological Sciences-Journal, 1993, 38 (1, 2): 15 ~ 31

(收稿日期: 1994-07-05)

· 新书介绍 ·

《第一届亚洲-大洋洲国际塑性力学会议论文集 (英文版)》——第一届亚洲-大洋洲国际塑性力学会议论文集 (英文版) / 王自强、徐秉业主编。——北京: 北京大学出版社, 1994, 4

第一届亚洲-大洋洲国际塑性力学会议于 1993 年 8 月 15~19 日在北京召开。文集汇集了会议论文 72 篇, 其中 24 篇为特邀报告。如国际塑性力学界的知名学者 E. H. Lee, E. Kremple, Q. S. Ngnyen 等在大会上所作的精采报告。文集内容包括: 分子动力学模拟; 相变、形状记忆合金; 单晶、多晶塑性; 损伤、断裂和局部化; 塑性动力响应; 循环塑性和高温塑性; 变形与断裂的位错结构; 岩石的本构及损伤演化; 应变与微结构引起的各向异性; 塑性工程理论及其应用等, 及时地概括了当前国际塑性力学发展主流, 反映了最新研究成果。文集可供高等院校的力学、土木材料等有关专业教师和研究生使用, 也可供科研机构从事力学、物理、材料科学的研究人员参考。

(余清芬 供稿)