

DOI: 10.13476/j.cnki.nsbdqk.2017.03.001

王军, 宁亚伟, 胡义明, 等. 混合分布在非一致性水 文频率分析中的应用[J]. 南水北调与水利科技, 2017, 15 (3):01-04, 24. WANG Jun, NING Ya wei, HU Yi ming, et al. Application of mixed probability distribution to nor stationary hydrological frequency analysis [J]. South to North Water Transfers and Water Science & Technology, 2017, 15(3):01-04, 24. (in Chinese)

混合分布在非一致性水文频率分析中的应用

王 军1,宁亚伟1,胡义明1,刘和昌2,梁忠民1,3

(1.河海大学 水文水资源学院, 南京 210098; 2. 水利部 珠江水利委员会技术咨询中心, 广州 510000; 3. 水安全与水科学协同创新中心, 南京 210098)

摘要: 传统水文频率分析方法的前提条件是水文极值系列需满足一致性要求, 然而由于气候变化和人类活动的影响, 使得水文资料的非一致性问题越来越突出, 导致现行水文频率分析方法的应用受到挑战。根据重建的淮河流域1470年—1999年共530年夏季面平均降雨量长系列数据, 进行分析, 通过假设检验表明系列在1534年、1724年和1923年发生跳跃性变异, 因而将系列划分为4个子平稳系列; 采用现行频率分析方法估计每个子系列的概率分布函数, 再根据混合分布概念由各子系列的概率分布构建一个综合的概率分布函数, 作为非一致性降雨系列总体分布的估计。在此基础上, 计算了给定设计标准下的设计值, 并与基于大样本(530年) 计算的经验设计值进行了对比,结果表明混合分布模型对观测系列具有较好的拟合效果。

关键词: 非一致性; 水文频率分析; 混合分布; 淮河流域

中图分类号: P333 文献标识码: A 文章编号: 1672-1683(2017)03 000 F 04

Application of mixed probability distribution to non stationary hydrological frequency analysis

WANG Jun¹, NING Yawei¹, HU Yi ming¹, LIU He chang², LIANG Zhong min^{1,3}

(1. College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China; 2. Technical Advisory Center of Pearl River Water Resources Commission of the Ministry of Water Resources, Guangzhou 510000, China;

3. National Cooperative Innovation Center for Water Safety & Hydro Science, Nanjing 210098, China)

Abstract: The traditional hydrological frequency analysis (THFA) method has been widely used to estimate the design flood for a given return period. It relies on the assumption that the series should be stationary. However, climate change and human activities caused the series to lose their stationarity, and thus limit the application of the THFA method. In this paper, we studied the 530 year summer precipitation series covering the period from 1470 to 1999. At the 0-01 significant level, it was found that the series had three change points: 1534, 1724, and 1923. Therefore, the whole series was divided into four sub-sample series. We estimated the probability distribution function (PDF) of each sub-sample using the THFA method. Then we obtained the mixed probability distribution function (MPDF) by weighting and synthesizing the four PDFs related to the four sub-samples. The MPDF was regarded as the PDF of the whole series and was used to estimate the design precipitation for a given return period. Besides, the 530 year large precipitation series was also used to estimate the empirical value of the design precipitation. Then we compared the observation sample based estimations and MPDF-based estimations of the design precipitation. The results showed that the MPDF fit the series well.

Key words: norr stationary; hydrological frequency analysis; mixed probability distribution; Huaihe River basin

收稿日期: 2016-12-20 修回日期: 2017-03-01 网络出版时间: 2017-05-06

网络出版地址: http://kns.cnki.net/kcms/detail/13.1334.TV.20170506.1644.016.html

基金项目: 国家重点研发计划项目"水资源高效开发利用"重点专项(2016YFC0402706; 2016YFC0402707; 2016YFC0402709)

Funds: National Key Research and Development Program of China "Water Resources Development and Utilization of High Efficiency" Key Projects (2016Y FC0402706; 2016Y FC0402707; 2016Y FC0402709)

作者简介: 王 军(1982), 安徽亳州人, 副教授, 博士, 主要从事水文水资源方面研究。 E mail: wangjun_h hu@ 163. com

在各类水利工程的规划、设计、施工以及运行管 理中. 均涉及到给定设计标准下水文设计值的推求 问题。水文频率分析方法作为推求水文设计值的一 种标准工具已得到广泛应用。现行水文频率分析方 法的应用前提条件是水文极值系列必须满足一致性 要求,即水文极值系列的统计规律在过去、现在或未 来是一致的。但随着全球气候变化和人类活动影响 的加剧,降雨及其时空分配过程和流域的下垫面产 汇流条件发生了变化,导致水文极值系列的非一致 性问题越来越突出, 给现行频率计算方法的应用带 来挑战[12]。目前, 国内外进行非一致性水文频率计 算主要基于两个途径[3]:一是通过还原(还现)方法 将水文极值系列还原(还现)到过去(现在)某一状 态, 使其满足一致性要求; 二是直接基于非平稳极值 系列进行水文频率分析,如时变矩模型[46]和混合分 布模型[710]等。时变矩模型是假定分布函数中的参 数随时间(或与时间相关联的某一变量,即协变量) 之间呈现某种趋势性变化,通过建立分布参数与时 间(或协变量)之间的函数关系,实现对分布参数的 估计,从而确定未来任意时刻水文极值变量的分布 函数。混合分布模型认为整个非一致性的样本系列 可以分成若干个一致性的子系列, 先估计各子系列 的分布函数(子分布),再对各子分布进行加权综合, 以获得一个可综合考虑各子系列特征的分布函数, 用以描述极值系列的总体分布特征。

本文根据重构的淮河流域 1470 年- 1999 年共530 年的夏季长降雨样本系列,通过统计假设检验分析系列的不一致性特征(跳跃性特征),采用混合分布模型对该系列进行频率分析,比较基于混合分布模型计算的降雨量设计值和重构的大样本计算的经验设计值间的差异,分析了混合分布模型的适用性。

1 研究方法

1.1 滑动秩和检验法

滑动秩和检验法¹¹¹ 是针对传统秩和检验法只能用于判断给定变异点是否显著而不能分析变异点位置这一弊端而提出的, 其通过对水文极值系列进行逐点分析, 然后选取使得检验统计量最大且变异显著的点作为变异点。即通过应用滑动秩和检验法, 在获得变异点位置的同时, 也可对变异点的显著性进行检验。

假设水文极值系列为 $x_1, x_2, x_3, ..., x_n$, 首先将系列 $x_1, x_2, x_3, ..., x_n$ 从小到大进行排序, 排序后的系列记为 $x_{(1)}^*, x_{(2)}^*, x_{(3)}^*, ..., x_{(n)}^*$; 计算原始系列 x_i ,

当 n_1 , $n_2 > 10$ 时, 统计量 W 近似服从正态分布, 即:

$$W \sim N \left(\frac{n_1(n_1 + n_2 + 1)}{2}, \frac{n_1 n_2(n_1 + n_2 + 1)}{12} \right)$$
 (1)
式(1)可以进一步写为

$$U = \frac{W - \frac{n_1(n_1 + n_2 + 1)}{2}}{\frac{n_1 n_2(n_1 + n_2 + 1)}{12}} \sim N(0, 1)$$
 (2)

需要注意的是, 当 $n_1 > 10$ 且 $n_2 < 10$ 或 $n_1 < 10$ 且 $n_2 > 10$ 时, 变异点检验结果将可能存在较大不确定性。需要结合多种方法的检验结果及实际情况进行综合判定。

1.2 混合分布模型

气候变化及人类活动的影响,使得不同时期的水文事件的形成过程或形成机制存在差异,也即不同一时期之间水文极值事件的统计规律不一致,而在同一时期内水文极值事件的统计规律差异不显著。为此,可采用不同的分布函数对不同时期的水文极值系列的统计规律进行描述,再对各时期分布函数进行综合得到一个统一的分布函数,作为水文极值总体的估计,这是混合分布模型的基本思想。混合分布模型在非一致性水文频率分析中得到了较多研究⁽⁷⁻¹⁰⁾,其可表示为

 $F(x) = \alpha_i F_1(x) + \alpha_i F_2(x) + \dots + \alpha_k F_k(x)$ (3) 式中: k 为子样本系列的个数; $F_i(x)$ 为第 i 个子系列的分布函数; α_i 为第 i 个子系列分布函数对应的权重。且, $\sum_{i=1}^{k} \alpha_i = 1$ 。

在我国,水文极值系列的频率分析一般采用 P- 0 型分布函数,为此,将采用 P- 0 型分布函数对 各个子系列的分布特征进行描述。其概率密度函数 表达式为

$$f_{i}(x) = \frac{\beta_{i}^{Y_{i}}}{\Gamma(Y_{i})} (x - Y_{i})^{\alpha_{1} - 1} e^{-\beta_{1}(x - \frac{\mu}{i})}$$
(4)

式中: V_i 、 β_i 和 μ_i (i= 1, 2, ..., k) 分别为第 i 个子系列的概率密度函数 f_i (x) 的形状、尺度和位置参数。

关于混合分布模型参数的估计方法较多,如全局优化算法^[10]、非线性优化算法^[12]、极人似然法^[13]等。但当子分布较多时,上述估计方法的应用较为复杂,且参数估计不确定性较大,如当子分布为3个时,需要估计的参数有12个之多,应用上述方法存在较大困难。为此,本文采用简化方法对各个字分布函数的权重进行估计,即根据各子系列容量占整个系列容量的比例作为各子系列对应的权重系数,即认为在不同时期,导致水文极值系列发生变异的主影响因素及其影响强度是不同的,子系列容量的长短在一定程度上可以反应这些因素的影响程度,因此可用子系列容量占整个系列容量的比例作为其分布函数的权重。

1.3 设计值的经验估计和基于混合分布模型的理论估计

由于实测样本系列为 1470 年 - 1999 共 530 年, 样本容量较大, 为此, 可基于此大容量样本计算 给定超过概率下的分位点, 作为分位点的经验估计 值。对于给定的观测值, 其对应的经验超过频率可 采用下式计算:

$$P(x > x_i) = \frac{m_i}{n+1}$$
 (i= 1, 2, ..., n) (5)

式中: m_i 为系列中数值大于 x_i 的个数;n 为系列长度。通过式(5)可以反求给定超过概率的分位点经验估计值。

根据式(3) 构建的混合分布模型, 可计算给定超过概率 P(或重现期 T) 对应的分位点值 X_T , 水文变量 $x > X_T$ 的重现期 T 计算公式如下:

$$T = \frac{1}{P(x > x_i)} = \frac{1}{1 - P(x \le x_i)} = \frac{1}{1 - F(x_i)}$$
 (6)
设计值 X_T 可采用下式估计:

$$X_T = F^{-1}(1-P) = F^{-1}(1-\frac{1}{T})$$
 (7)

式中, $F^{-1}(\acute{0})$ 是分布函数 $F(\acute{0})$ 的反函数。

因为 530 年的样本系列容量较大,通过对比分析混合分布的理论估计与基于大样本的经验估计,一定程度上可以检验混合分布的拟合情况。

2 实例分析

本研究使用的淮河流域 1470 年— 1999 年降雨数据,来自于 Feng^[14]等基于树轮、历史记录等信息重建的亚洲及太平洋地区近600 年0.5° 网格夏季降

雨数据集。根据重建的淮河流域 1470 年- 1999 共530 年 0.5° 网格逐年的夏季平均降雨,通过加权平均法求得整个淮河流域夏季的面平均降雨系列。

2.1 变异点诊断

图 1 给出了降雨系列的时间系列以及 10 年和 20 年滑动平均过程。

在显著性水平 α= 0.01 下,采用滑动秩和检验 法对降雨系列的变异点进行检验,诊断出第一个可能变异点为 1534 年,将整个系列在 1534 年划分为前后两段,即"1470年—1534年"和"1535年—1999年"两个子系列;同样在 0.01 显著性水平下,分别对"1470年—1534年"和"1535年—1999年"两个子系列的变异点进行诊断,结果表明 1470年—1534年子系列无显著跳跃点存在,而 1535年—1999年这段系列,在 1724年发生显著变异。为此,在 1724年处,将"1535年—1999年"子系列划分为两段,即"1535年—1724年"和"1725年—1999年"两个子系列;同样在 0 01 显著性水平下对各个子系列进行逐步检验。最终确定 1534年、1724年和 1923年为系列的三个变异点。

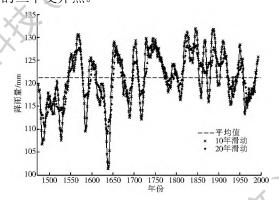


图 1 降雨时间系列的滑动平均结果 Fig. 1 Moving average results of precipitation time series

对于本文采用的滑动秩和检验法,显著性水平 a越小表示检验越严格,所以认为整个系列存在三个变异点在统计上具有较高的可信度。另外,杨传国等[15]通过对淮河 500 年来的降雨数据分析认为,流域存在 3 个主要的湿润期:明代末期 1550 年—1590 年、清朝早期 1730 年—1780 年、清朝晚期民国初 1820 年—1920 年。3 个主要的干旱期:明代晚期 1470 年—1545 年、明代末期清朝初期 1615 年—1660 年、近现代 1935 年—1980 年,即旱涝变异年份大致在 1545 年—1550 年、1590 年—1615 年、1660 年—1730 年、1920 年—1935 年之间。本研究识别出的变异点位置与旱涝变异年份比较吻合,因此,最终确定 1534 年、1724 年和 1923 年为变异年份,据此份将原系列划分为" 1470 年—1534 年"、

"1535年-1723年"、"1724年-1923年"和"1924 年- 1999年"四个子系列。

2.2 混合分布函数估计

采用 P- 0 型分布函数分别对四个子系列进行 拟合, 记为 $F_1(x)$ 、 $F_2(x)$ 、 $F_3(x)$ 和 $F_4(x)$; 采用经 验适线法进行分布参数估计,估计结果见表 1,频率 曲线 见图 2。

表 1 四个子系列统计参数

Tab. 1 Statistical parameters of the four subseries

子系列	均值 EX/mm	变差系数 C_v	偏态系数 C_s
1470年- 1534年	114. 2	0.08	0. 14
1535年- 1724年	120. 4	0.08	0. 25
1725年- 1923年	124. 9	0.07	0. 24
1924年- 1999年	119. 8	0.10	0.69

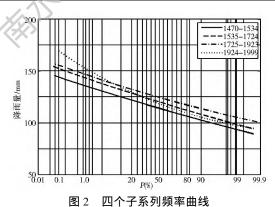


Fig. 2 Frequency curves of the four subseries

根据各个子样本系列容量占整个样本容量的比 值确定各个子分布函数的权重,依次计算的各子分 布函数的权重分别为 $\alpha_1 = 0.1226$, $\alpha_2 = 0.3585$, α= 0.375 5, α₄= 0.143 4, 混合分布模型可表示为

$$F(x) = 0.1226F_1(x) + 0.3585F_2(x) + 0.3755F_3(x) + 0.1434F_4(x)$$
 (8)

分别根据式(5)和式(7)、式(8)计算重现期为 5、10、20、50和100年时降雨量设计值的经验估计 值和混合分布模型计算的估计值,结果见表 2。表 2 中同时给出了经验估计值和模型计算值之间的相对 误差。由表 2 可知, 关于给定重现期的设计降雨量 估计,基于530年样本系列的经验估计值与模型计 算值差别很小,表明混合分布模型较好地反应了降 雨量系列分布特征。

表 2 各重现期下降雨量的估计值和经验值

Tab. 2 Estimated and empirical values of precipitation in different return periods

重现期/年	经验值	估计值	相对误差(%)
5	130. 1	129. 7	- 0.3
10	134. 3	134. 4	0. 1
20	138. 1	138. 4	0.3
50	141.6	143. 1	1. 1
100	145. 1	146. 3	0. 9

结语 3

气候变化和人类活动的影响导致不同时期的水 文事件的孕育背景存在差异, 进而导致水文极值系 列的非一致性问题越来越突出, 使得现行水文频率 分析方法对水文极值系列的一致性要求难以满足。 本文根据重构的淮河流域 1470 年- 1999 年共 530 年的夏季面平均降雨大样本系列, 对混合分布模型 在非一致水文频率分析中的适用性进行了应用研 究,主要结论如下。

- (1) 混合分布模型认为非一致性水文极值的概 率分布函数, 可以通过综合样本系列中若干个一致 性子系列的分布函数估计得到,从而以一个统一的 概率分布函数描述非一致性水文极值的总体特征, 为解决变化环境下水文分析计算问题提供了一种新 的思路。
- (2) 对淮河流域 1470 年- 1999 年共 530 年的 降雨大样本分析表明,降雨系列在 1534、1724 和 1923年发生跳跃性变异, 据此对原系列划分了 4个 子系列,估计了各子系列的分布函数并推求了其混 合分布 作为流域未来降雨量的总体分布。
- (3) 对比分析混合分布模型估计的理论设计值 与基于大样本系列(530年)的经验设计值,两者相 对误差较小, 表明混合分布模型对样本系列具有较 好的拟合效果。

参考文献(References):

- [1] Petra S F, Felix N. More frequent flooding? Changes in flood frequency in Switzerland since 1850[J]. Journal of Hydrology, 2010, 381: 1-8.
- [2] 胡义明,梁忠民,杨好周,等.基于趋势分析的非一致性水文频 率分析研究[J]. 水力发电学报, 2013, 32(5): 21-25. (HU Yi ming, LIANG Zhong min, YANG Hao zhou, et al. . Study on frequency analysis method of non stationary observation series based on trend analysis[J]. Journal of Hydroelectric Engineering, 2013, 32(5): 21-25. (in Chinese))
- 梁忠民, 胡义明, 王军. 非一致性水文频率分析的研究进展[J]. 水科学进展, 2011, 22(6): 864 871. (LIANG Zhong min, HU Yiming, WANG Jun. Advances in hydrological frequency analysis of nonstationary time series[J]. Advances in Water Science, 2011, 22(6): 864 871. (in Chinese))
- [4] Strupczewski W G, Singh V P, Feluch W. Norr stationary apr proach to at site flood frequency modeling I. Maximum likelihood estimation[J]. Journal of Hydrology, 2001, 248: 123-142.
- Richard M V, Chad Y, Meghan W. Nonstationarity: flood magnification and recurrence reduction factors in the United States [J]. Journal of the American Water Resources Association, 2010, 47(3): 464-474.

(下转第24页)

- eling, 2014, 1: 238 242.
- [9] Burger C M, Kollet S, Schuacher J, et al. Introduction of a web service for cloud computing with the integrated hydrologics inr ulation platform ParFlow [J]. Computers & Geosciences, 2012, 48: 334-336.
- [10] Zhang G J, Yao Y Y, Zheng C M. HPC environment on Azure cloud for hydrological parameter estimation [C]. 17th IEEE International Conference on Computational Science and Engineering, 2013, 8: 299 304.
- [11] 朱仕杰. 基于 SOA 的水文模型集成研究[D].1 兰州: 兰州大学, 2009. (ZHU Shi jie. Research of Hydrological Model Integration Based on SOA [D]. Lanzhou: Lanzhou University, 2009. (in Chinese).
- [12] 崔云飞, 李艺, 李昀, 等. 基于 SOA 的云 计算体系结构研究 [J]. 装备指挥技术学院学报, 2011, 22(4): 78 81. (CUI Yurfei, Li Yi, LI Jun, et al. Cloud computing architecture research based on service oriented architecture [J]. Journal of the Academy of Equipment Command & Technology, 2013: 1231-1237. (in Chinese))
- [13] Zhang Q, Cheng L, Boutaba R. Cloud computing: state of the art and research challenges [J]. Journal of Internet Services and Applications, 2010, (1): 7-18.
- [14] Zhang X, Kunjithapatham A, Jeong S, Gibbs S. Towards an E-lastic Application Model for Augmenting the Computing Capabilities of Mobile Devices with Cloud Computing [J]. Mobile Networks and Application, 2011, (16): 270 284.
- [15] 梁爽. 基于 SOA 的云计算框架模型的研究与实现[J]. 计算机

- 工程与应用, 2011, 47(35): 92 94. (LIANG Shuang. The design and realization of cloud computing framework model based on SOA [J]. Computer Engineering and Applications, 2011, 47(35): 92 94. (in Chinese))
- [16] 张潇元,刘利人,韩海雯. 基于 SOA 云架构的电子监察业务平台的设计与构建[J]. 计算机科学, 2014, 41(1A): 473 474. (ZHANG Xiao yuan, LIU Li ren, HAN Hai wen. Supervision collaboration platform based on SOA and cloud computing technologies [J]. Computer Science, 2014, 41(1A): 473 474. (in Chinese))
- [17] 朱志良, 苑海涛, 宋杰, 等. SOA 与云计算: 竞争还是融合[J]. 计算机科学, 2011, 38(12): 69. (ZHU Zhi liang, YUAN Hai tao, SONG Jie, et al. SOA and cloud computing: competition or integration [J]. Computer Science, 2011, 38(12): 69. (in Chinese))
- [18] 李新,程国栋,康尔泗,等. 数字黑河的思考与实践 3: 模型集成[J]. 地球科学进展, 2010, 25(8): 851 865. (LI Xin, CHENG Guordong, KANG Ersi, et al. Digital Heihe River basin 3: model integration [J]. Advances in Earth Science, 2010, 25(8): 851 865. (in Chinese))
- [19] 刘海燕, 刘晓民, 魏加华, 等. 组件式流域模拟模型集成技术进展及发展趋势[J]. 南水北调与水利科技, 2013, 11(1): 140-145. (LIU Hair yan, LIU Xiao min, WEI Jiar hua, et al. Progress and development trend of watershed simulation model integration based on component technology [J]. South to North Water Transfers and Water Science & Technology, 2013, 11(1): 140-145. (in Chinese))

(上接第4页)

- [6] Gabriele V, James A S, Francesco S, et al. Flood frequency analysis for nonstationary annual peak records in an urban drainage basin [J]. Advances in Water Resources 2009, 32: 1255-1266.
- [7] Waylen P, Woo M K. Prediction of annual floods generated by mixed processes [J]. Water Resources Research, 1982, 18 (4): 1283-1286.
- [8] Diehl T, Potter K W. Mixed flood distribution in Wisconsin [A]/ / SINGH V P. Hydrologic Frequency Modelling [C]. Netherlands: D. Reidel Publishing Company, 1987: 213-226.
- [9] 冯平,曾杭,李新. 混合分布在非一致性洪水频率分析的应用 [J]. 天津大学学报: 自然科学与工程技术版, 2013, 46(4): 298-303. (FENG Ping, ZENG Hang, LI Xin. Non stationary floodfrequency analysis based on mixed distribution[J]. Journal of Tianjin University: Science and Technology, 2013, 46(4): 298-303. (in Chinese))
- [10] 成静清, 宋松柏. 基于混合分布非一致性年径流序列频率参数的计算[J]. 西北农林科技大学学报: 自然科学版, 2010, 38(2): 229 234. (CHENG Jing qing; SONG Song bai. Calculation of hydrological frequency parameters of inconsistent annual runoff series based on mixed distribution [J]. Journal of

- Northwest A & F University: Nature Science Edition, 2010, 38(2): 229-234. (in Chinese)
- [11] 谢平, 陈广才, 雷红富. 水文变异综合诊断系统[J]. 水力发电学报, 2010, 29(1): 85-91. (XIE Ping, CHEN Guang-eri, LEI Hong fu. Hydrological alteration diagnosis system[J]. Journal of Hydroelectric Engineering, 2010, 29(1): 85-91. (in Chinese))
- [12] Singh K P. A versatile flood frequency methodology[J]. War ter International, 1987, 12 (3). 139 145.
- [13] Rossi J, Fiorentino M, Versace P. Two-component extreme value distribution for flood frequency analysis [J]. Water Resources Research, 1984, 20(7): 847-856.
- [14] Feng S, Hu Q, Wu Q, et al. A gridded reconstruction of warm season precipitation for Asia spanning the past half millennium [J]. Journal of Climate, 2013, 26(7): 2192-2204.
- [15] 杨传国, 陈喜, 张润润, 等. 淮河流域近 500 年洪旱事件演变特征分析[J]. 水科学进展, 2014, 25 (4): 503 510. (YANG Chuar guo, CHEN Xi, ZHANG Rur run, et al. Characteristics of flood and drought events of the last half millennium in Huaihe River basin[J]. Advances in Water Science, 2014, 25 (4): 503 510. (in Chinese))