

基于极值水文分区和层次贝叶斯的区域非一致性频率分析

曾 杭^{1, 2}, 周 洋^{1, 2}, 李伶俐³, 王国庆^{3, 4}

(1. 长沙理工大学 水利与海洋工程学院, 湖南 长沙 410114; 2. 长沙理工大学 洞庭湖水环境治理与生态修复湖南省重点实验室, 湖南 长沙 410114; 3. 南京水利科学研究院 水灾害防御全国重点实验室, 江苏 南京 210029; 4. 水利部应对气候变化研究中心, 江苏 南京 210029)

摘要: 非一致性水文频率分析是变化环境下工程水文设计研究的关键科学问题, 而现有方法的不确定性较高且未考虑站点序列间的空间相关性。因此本文提出基于极值水文分区和层次贝叶斯的区域非一致性水文频率分析方法, 首先采用变差函数 F -madogram 的围绕中心点划分(PAM)聚类算法对流域划分水文分区, 其次针对同一分区构建基于层次贝叶斯的区域非一致性频率分析模型, 具体包括无池化、部分池化和全池化3种模型。最后, 以湘江流域年最大24 h极值降雨为例开展实例研究。结果表明: 通过遵循极值理论的空间聚类算法, 将湘江流域36个雨量站划分为3个水文分区; 以第I水文分区为例, 与大部分站点呈显著正相关的气候驱动因子为FMA Nino12。相较单站非一致性模型, 部分池化模型模拟表现最优, 能较好地体现了气候驱动因子的区域同质响应特征并保留了站点个体差异, 各站区域参数的不确定性区间显著降低了10%~35%。随着重现期的不断增大, 各站非一致性重现期相较一致性情况不断减小, 如50年一遇的重现期, 部分站点减小比例接近60%, 表明在气候因子驱动下分区内极值降雨事件发生概率增大。本文方法丰富了非一致性水文频率分析的方法体系, 为科学确定流域设计暴雨并制订防灾策略提供依据。

关键词: 区域非一致性; 水文频率分析; 极值水文分区; 层次贝叶斯; 湘江流域

中图分类号: P333; TV11

文献标识码: A

doi: 10.3724/j.slx.20250304

1 研究背景

水文频率分析在水利工程规划设计、运行管理及市政工程设计中具有重要作用。目前, 国内外研究学者主要是基于物理成因一致的长期观测资料认识水文规律, 然而, 随着全球气候变化和人类活动加剧的影响, 流域的降雨特性、下垫面的产汇流过程、极值降雨/洪水序列等已发生变化, 引发了水文特征的时空不对应及水文序列的非一致性^[1-3], 进而导致基于一致性假设的水文频率计算方法的适用性受到严峻挑战。因此, 研究变化环境下的水文频率分析方法是进行流域水利工程设计及暴雨洪涝防治的迫切需要, 对水文学科的发展具有重要意义^[4]。

目前, 国内外水文学者已经普遍接受水文特征已发生变化的观点, 在非一致性水文频率分析方法上取得一系列代表性研究成果^[5-6]。关于单站单变量非一致性水文频率分析研究, 主要有3种计算途径: (1)基于还原/还原的方法^[7-8], 未考虑未来环境变化, 因此存在较大的外延风险; (2)条件概率分布和混合分布^[9-12], 有其特定物理意义的应用情景; (3)时变矩^[13-15], 此方法目前研究最多。对于单站多变量非一致性水文频率分析方法的研究也取得一定的进展, 主要是通过Copula函数构建两变量或三

收稿日期: 2025-05-27; 网络首发日期: 2026-05-09

网络首发地址: <https://link.cnki.net/urlid/11.1882.TV.20260508.1134.002>

基金项目: 国家自然科学基金创新研究群体项目(52121006); 湖南省自然科学基金面上项目(2024JJ5022); 湖南省教育厅科学研究优秀青年项目(24B0329); 湖南省水利科技项目(XSKJ2023059-06)

作者简介: 曾杭(1989—), 博士, 副教授, 主要从事水文学及水资源研究。E-mail: hzeng1989@csust.edu.cn

通信作者: 王国庆(1971—), 正高级工程师, 主要从事流域水文及生态水文过程模拟研究。E-mail: gqwang@nhri.cn

变量的联合分布,考虑边缘分布或变量间相关结构的非一致性进行研究。在气候或下垫面因子驱动下,基于时变矩的单站单变量非一致性水文频率分析,所需估计的参数较多,而我国水文测站实测水文系列样本长度有限,单站水文序列样本代表性不足,导致设计水文成果的不确定性显著提高,严重影响了水文设计值估计精度。因此,如何提高单站非一致性水文频率计算模型的计算精度,是亟待解决的科学问题。

区域水文频率分析是一种有效提高水文设计值估算精度的计算方法^[16-18]。国外在区域水文频率分析方面的研究相对较成熟,并得到广泛应用,此方法主要是通过有效利用水文分区内相似站点的水文信息,以弥补单站样本长度有限的不足,提高估计精度和降低成果不确定性。其中应用到工程实践最广泛的方法是标度洪水法(Index-Flood Method),此方法需满足的假定包括:同站点水文序列满足一致性假定;同一水文分区内,除不同站点间的洪水尺度系数不同外,所有站点水文序列的概率分布和参数完全一致;各站点的水文序列相互独立。其不足体现在^[16-18]:(1)对标度洪水的取值,一般采用站点水文序列的均值或中位数来估计,选取的物理意义不明确;(2)各站点水文序列之间不可能相互独立,即未考虑站点间空间相关性,且互相相关的站点所包含的相同水文信息被频率计算模型重复利用,会低估设计水文成果的不确定性^[19];(3)同一水文分区内各站点分布和参数一致的假定,实际中很难满足。随着全球气候变化的影响,同站点水文序列一致性假定遭到质疑,区域非一致性水文频率分析的研究逐渐受到关注,主要体现在:一方面,基于IFM原理,指定各站标度洪水值的同时,通过分析区域频率曲线的非一致性来研究^[18];另一方面,直接对同一水文分区内各站水文序列,建立区域非一致性频率模型^[20],如Wu等^[21]建立了基于层次贝叶斯的区域径流频率分析模型,但未考虑水文序列的非一致性和空间相关性;Sun等^[19]构建了区域非一致性频率分析模型,提出的两类模型参数分别反映区域信息和站点特性,但对两类模型参数没有清晰的定位;Renard等^[17]构建了区域非一致性模型,提出站点参数、区域参数和随机参数,由于模型结构非常复杂,导致难以应用;Chen等^[22]利用层次贝叶斯模型对淮河流域的径流和降雨进行了区域非一致性频率计算,但未进行水文分区,且未区分径流和降雨。

本文在湘江流域开展实例研究,以流域各站年最大24 h极值降雨序列为研究对象,主要开展了如下研究工作:(1)采用遵循极值理论的空间聚类方法,对极值水文序列进行水文分区;(2)同一水文分区内,在气候预测因子驱动下,对极值降雨构建清晰明确的区域非一致性频率分析模型,采用的层次贝叶斯框架结构,既充分利用分区内各站极值降雨信息,又考虑站点间序列相关性;(3)明确提出代表区域同质特征和站点特性的模型参数,以降低非一致性水文频率计算结果的不确定性。本文模型可进一步完善非一致性水文分析计算的方法体系,为流域防灾减灾和工程水文设计提供重要的科学理论依据。

2 研究方法

2.1 极值序列水文分区 在区域水文频率研究中,常将极值降雨/洪水频率分布曲线参数相同的区域称为同一水文分区,以表明该分区的降雨/洪水特性相同。常用的同一水文分区方法有等值线图、水文地理因子空间分布、聚类方法等^[23],其中聚类方法应用最广泛,如K均值聚类、模糊聚类、主成分聚类、高斯混合模型聚类等。以上聚类方法主要是以区域站点水文要素均值或标准差为标准来进行分类,而序列均值呈现高斯特征且以站点间欧氏距离量化空间相关性,对呈现高偏特征的水文极值序列并不适用^[24]。

为克服以上两点不足,本文采用基于变差函数F-madogram的围绕中心点划分(Partitioning Around Medoids, PAM)聚类算法^[24]对流域内各站点极值水文序列进行同一水文分区,其中变差函数F-madogram^[25]可以量化站点间极值序列的空间相关性,这种量化方法遵循极值理论体系。PAM算法^[26]可将站点序列划分到空间相关性最近的同一水文分区里,并且分区结果保留了极值高偏统计特征。假

设两个水文站 i 和 j 的极值水文序列分别为 M_i 和 M_j ，则两个极值序列空间相关性的变差函数 F -madogram 非参数估计量计算公式为

$$\hat{d}_{ij} = \frac{1}{2T} \sum_{t=1}^T \left| \hat{F}_i(M_i^{(t)}) - \hat{F}_j(M_j^{(t)}) \right| \quad (1)$$

式中： $\hat{d}_{ij}$ 为站点 i 和站点 j 极值序列空间相关性的估计量； $M_i^{(t)}$ 和 $M_j^{(t)}$ 分别为水文站编号为 i 和 j 的极值水文序列； T 为极值序列长度 ($t = 1, 2, \dots, T$)； $\hat{F}_i$ 和 $\hat{F}_j$ 均为经验分布函数，分别代表序列 $M_i^{(t)}$ 和 $M_j^{(t)}$ 的边际分布，其计算公式为

$$\hat{F}_i(u) = \frac{1}{T} \sum_{t=1}^T I_{\{M_i^{(t)} \leq u\}} \quad (2)$$

式中： $I_{\{M_i^{(t)} \leq u\}}$ 为事件 $\{M_i^{(t)} \leq u\}$ 的指示函数； u 为任一极值水文值。

基于变差函数 F -madogram 得到两个水文站极值序列之间的空间相关性，再采用 PAM 算法将流域内各站极值序列划分成多个水文分区。计算得到的同一水文分区，在算法中每一步不仅保留极大值序列高偏特征，而且保证分区里极值序列之间的总空间相关性最小。

为确定最优水文分区个数 K ，并判断流域内极值序列水文分区是否最优，Rousseeuw^[27] 提出采用“轮廓系数”来比较水文分区的紧密和分离程度。假设基于变差函数 F -madogram 的 PAM 聚类算法得到了 K 个分区，每个水文站点 i 有其分区聚类中心站点 k ，则水文站点 i 的轮廓系数计算公式为

$$s_i(K) = 1 - d_{ik}/\delta_{i,-k} \quad (3)$$

式中：流域被分为 K 个水文分区后， $s_i(K)$ 为水文站点 i 的轮廓系数； d_{ik} 为站点 i 到所属分区聚类中心站点 k 的空间相关性； $\delta_{i,-k}$ 为站点 i 与除 k 外的其他分区聚类中心站点的最近空间相关性。

为确定最优水文分区个数 K ，通过计算轮廓系数的均值来判断，计算公式为

$$\bar{s}(K) = \frac{1}{N} \sum_{i=1}^N s_i(K) \quad (4)$$

式中： $\bar{s}(K)$ 为流域被分成 K 个水文分区后的平均轮廓系数； N 为流域内水文站点总数。其中， K 个水文分区平均轮廓系数取值越接近于 1，表明此分区聚类的总空间相关性越近，水文分区效果越优；反之，取值越接近于 0，水文分区效果越不好。

2.2 区域非一致性水文频率分析方法 针对标度洪水法和单站非一致性模型的不足，本文对同一水文分区内各站极值水文序列，构建基于层次贝叶斯模型的区域非一致性水文频率分析模型 (Hierarchical Bayesian Model-Nonstationary Regional Frequency Analysis, HBM-NRFA)，建立分区极值序列的多变量联合分布。此模型不仅能充分利用同一水文分区内各站数据信息，降低水文样本长度限制的影响，从而降低计算成果的不确定性^[28]；而且其基于多层次结构的模拟，既能体现区域化同质特征，又能保持各站点的个体特点。

在层次贝叶斯框架结构中，多变量联合分布的函数形式需直接显式给出，因此不能通过 Copula 函数来构建。比如，若各序列的边缘分布假定为广义极值分布，且各站点的相关关系未知，则采用 Copula 函数建立的多变量联合分布形式无法显式应用^[17,29]。而多变量正态分布，是一个直接显式给出的函数形式^[19]，可直接用于层次贝叶斯结构，得到显式的多变量联合分布贝叶斯后验分布。Serago 等^[30] 曾总结所有适合洪水频率分析的概率分布，表明两变量对数正态分布 (LN2 分布) 和皮尔逊 III 型分布，是两个无显著偏差的分布线型；LN2 分布参数较少，所构建的非一致性模型更简化。

基于以上原因，本文采用两变量对数正态分布 (LN2 分布)，作为同一水文分区中各站极值水文序列所服从的边缘分布，即各极值序列取自然对数之后的新序列服从正态分布，则各站点新序列的联合分布为多变量正态分布，正适合应用于层次贝叶斯框架中。HBM-NRFA 模型的具体方法介绍如下。

同一水文分区内各站点原始极值水文变量为 X ，假设服从 LN2 分布，其概率密度函数为

$$f(x|\mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right] \quad x > 0, \sigma > 0 \quad (5)$$

式中 μ 和 σ 分别为位置和尺度参数。其中,取自然对数后的 $\ln x$ 服从正态分布,同一水文分区内各站点 $\ln x$ 新序列的联合分布为多变量正态分布。

各站点极值序列为 X_t^s ,服从LN2分布,其中 S 为同一水文分区内站点总数($s = 1, 2, \dots, S$), T 为样本长度($t = 1, 2, \dots, T$),则HBM-NRFA模型结构表示为

$$\text{第一层: } \ln(X_t^s) \sim N(\beta_t^s, \Sigma) \quad (6)$$

$$\beta_t^s = \mu_0^s + \sum_{i=1}^n \mu_i^s z_t^i \quad (7)$$

$$\text{第二层: } \mu_i^s \sim MVN(\mu_{\mu_i}, \Sigma_{\mu_i}) \quad (8)$$

式中各待估参数的贝叶斯先验分布分别为

$$\begin{cases} \mu_0^s \sim U(0, 10000), \mu_{\mu_i} \sim U(0, 10000) \\ \Sigma_{\mu_i} \sim IW(v_0, \Lambda_0), \Sigma \sim IW(v_1, \Lambda_1) \end{cases} \quad (9)$$

以上HBM-NRFA模型结构中,第一层为每个站点 s 的极值序列 $\ln(X_t^s)$ 服从正态分布,正态分布的位置参数为 β_t^s ,是各协变量 z_t^i ($i = 1, 2, \dots, n$, n 为协变量总数)的线性函数;各协变量线性回归系数 μ_i^s 为 $n \times S$ 的矩阵,被定义为区域参数,其取值代表各协变量(即影响因子)对各站极值序列的区域化同质响应程度; Σ 为 $S \times S$ 的多变量正态分布协变量矩阵,描述和代表各站点极值序列的相关关系。第二层参数 μ_i^s 代表极值序列被各协变量影响的响应程度,服从超参数为 μ_{μ_i} 和 Σ_{μ_i} 的多变量正态分布。在各待估参数的先验分布中, μ_0^s 和 μ_{μ_i} 是服从均匀分布的无上限正数;协方差矩阵 Σ_{μ_i} 和 Σ 的共轭先验分布,服从对称正定矩阵分别为 Λ_0 和 Λ_1 、自由度分别为 v_0 和 v_1 的逆威沙特分布(Inverse-Wishart distribution),即 $IW(v_0, \Lambda_0)$ 和 $IW(v_1, \Lambda_1)$ 分布^[22,31]。

同一水文分区内,各站极值序列信息的整合程度代表区域同质的响应特征,在HBM-NRFA模型中通过区域参数 μ_i^s 的取值来体现。基于区域参数 μ_i^s 取值的不同,构建的HBM-NRFA模型可分为3种情景^[28]:无池化、部分池化和全池化模型。

无池化模型:本质为单站非一致性频率计算模型,不考虑各站极值序列间的相关关系。各站点之间相互独立,没有相关性,各站极值序列各自独立进行单站非一致性频率计算。则HBM-NRFA模型中站点间的协方差矩阵 Σ 变为各站点的尺度参数 σ_s ,各站的区域参数 μ_i^s 各自单独进行参数估计,其HBM-NRFA模型如下

$$\ln(X_t^s) \sim N(\mu_0^s + \sum_{i=1}^n \mu_i^s z_t^i, \sigma_s) \quad (10)$$

全池化模型:只考虑同一水文分区各站极值序列受协变量影响的区域同质响应特征,不考虑各站极值序列的个体差异。则各站点区域参数 μ_i^s 均相等,其HBM-NRFA模型如下

$$\ln(X_t^s) \sim N(\mu_0^s + \sum_{i=1}^n \mu_i z_t^i, \Sigma) \quad (11)$$

部分池化模型:既考虑各站极值序列间的相关关系,以体现同一水文分区受协变量影响的区域同质响应特征,又保留站点各自的个体差异特征。式(6)—(9)描述的HBM-NRFA模型即为部分池化模型,各站点的协变量线性回归系数即各站区域参数 μ_i^s 取值均不同,但服从超参数为 μ_{μ_i} 和 Σ_{μ_i} 的同一正态分布,表明受协变量影响的线性回归系数趋向区域同质平均水平;模型里的多变量正态分布协方差矩阵即代表考虑各站极值序列间的相关关系。其中,无池化模型和全池化模型是部分池化模型的两种极端情景。

总体来说,HBM-NRFA模型有四大核心优势:(1)各站点的参数取值不同,代表各站点极值序列的个体差异特征;(2)表征区域同质特征的区域参数 μ_i^s ,用取值相同(全池化)或在模型第二层中来自同一个正态分布(部分池化)来体现;(3)呈显式函数的多变量正态分布,其协方差矩阵代表各站极值序列之间的相关关系;(4)模型构建的多变量联合分布,整合了同一水文分区内全部的极值样本信息,能有效降低设计水文成果的不确定性。

2.3 模型参数估计和拟合优度 HBM-NRFA 模型所有参数包括超参数, 均采用贝叶斯推断理论进行估计。3 种情景下的模型参数先验分布如式(9)所示, 所有待估参数的后验分布均采用基于不回头 NUTS 采样器(No-U-Turn Sampler)的哈密顿蒙特卡洛算法进行估计^[32]。假设模型所有待估参数由 θ 表示, 其基于贝叶斯理论的联合后验分布为

$$f(\theta | \ln X) \propto f(\ln X | \theta) f(\theta) \propto \prod_{s=1}^S \prod_{t=1}^T N(x_t | \mu_0^s + \mu_i^s z_t^i, \Sigma) IW(\Sigma | v_1, \Lambda_1) N(\mu_0^s | 0, 10000) \times \\ N(\mu_i^s | \mu_{\mu_i}, \Sigma_{\mu_i}) N(\mu_{\mu_i} | 0, 10000) IW(\Sigma_{\mu_i} | v_0, \Lambda_0) \quad (12)$$

基于贝叶斯理论进行参数及不确定性区间的估计, 不同于赤池信息准则 AIC(Akaike Information Criterion)和贝叶斯信息准则 BIC(Bayesian Information Criterion), 本文采用离差信息准则 DIC(Deviation Information Criterion)^[33], 即充分考虑和利用贝叶斯推断的全部后验分布信息, 对模型进行拟合优度估计和比较, 模型的 DIC 值越小, 其模拟性能越优。DIC 计算公式详见文献[33]。

对 3 种情景下的 HBM-NRFA 模型, 本文采用期望超过次数法计算非一致性重现期, 分别与一致性重现期进行比较。Parey 等^[34]提出期望发生次数法(Expected Number of Exceedances, ENE)的思想, Salas 等^[35]将 ENE 思想扩展到非一致性情况, 方法介绍详见文献[35]。

3 实例应用

本文在长江中游地区的湘江流域开展实例研究。湘江流域面积为 9.46 万 km², 干流河长 856 km, 平均坡降 1.34‰。全流域属亚热带季风湿润气候, 雨量充沛, 洪涝灾害频发。

本文采用湘江流域 1959—2017 年 36 个雨量站的实测年最大 24 h 降雨序列为研究数据, 雨量站分布如图 1 所示, 降雨实测数据由湖南省水文水资源勘测中心提供。

湘江流域洪涝灾害频发的主要因素是极易形成的极端降雨, 且流域面积较大, 影响流域内各区域极端降雨形成的气候因素存在差异。为研究气候因素对同一水文分区内极值降雨的影响并开展预测, 本文采用常用的大气环流大尺度气候指数, 如厄尔尼诺-南方涛动指数(ENSO, 包括 Nino3.4、Nino12、Nino3、Nino4 和南方涛动指数 SOI)、太平洋十年涛动(PDO)、北大西洋涛动(NAO)和北极涛动(AO)^[36], 数据来源于荷兰皇家气象研究所网站(<https://climexp.knmi.nl/start.cgi>)^[37]。湘江流域每年 7、8 月份的降雨量最多, 年最大 24 h 降雨常出现在 4—8 月, 故本文考虑与极值降雨同一年 8 月(含)之前的气候预测因子, 同年 8 月之后的气候数据不考虑。所采用的气候预测因子时间范围为 1959—2017 年。

4 结果与讨论

4.1 流域极值降雨水文分区及气候驱动分析 基于变差函数 F -madogram 的 PAM 聚类算法, 对湘江流域 36 个雨量站的年最大 24 h 极值降雨序列进行水文分区。为防止出现小分区(如同一水文分区只有 2

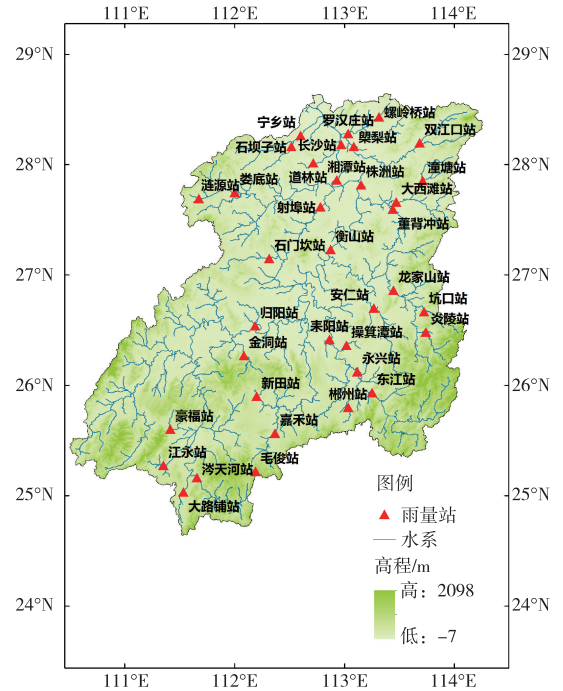


图 1 湘江流域雨量站的空间分布

Fig. 1 Spatial distribution of rainfall stations in Xiangjiang River Basin

个站点)的情况,将水文分区个数 K 的最大值取6。当 K 分别取2、3、4时,湘江流域极值降雨序列的水文分区结果如图2所示。流域各站极值降雨水文分区空间分布,除个别弱相关站点之外,不同水文分区的区分较显著。

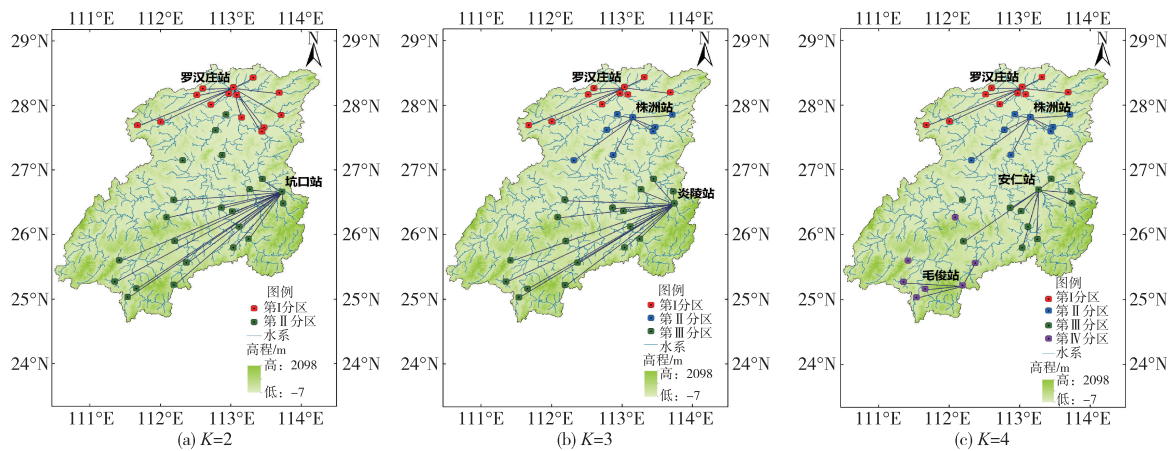


图2 基于 F -madogram的PAM算法水文分区结果

Fig. 2 Hydrologic regionalization of PAM algorithm with F -madogram

不同水文分区个数 K 对应的轮廓系数值和弱相关站点数如表1所示。综合水文分区空间分布和轮廓系数结果:当水文分区数为 $K=4$ 及以上时,其平均轮廓系数减小且出现弱相关站点,弱相关站点离水文分区中心距离远,相关性较弱;当水文分区数 $K=3$,平均轮廓系数减小,水文分区效果不好;当分区数 $K=2$ 时,轮廓系数稍小且弱相关站点数较多。因此综合而言,水文分区数 $K=3$ 时,湘江流域极值降雨水文分区空间分布最优,轮廓系数较大且无弱相关站点,各自水文分区内的极值降雨序列归属同一分区。

表1 不同分区个数 K 的轮廓系数及弱相关站点数

Table 1 Silhouette coefficients and numbers of weak correlated stations with different K

K 取值	最大轮廓系数	最小轮廓系数	平均轮廓系数	弱相关站点数
2	0.273	0.059	0.147	5
3	0.291	0.035	0.153	0
4	0.291	0.045	0.144	3
5	0.291	0.017	0.128	2
6	0.291	0.017	0.128	2

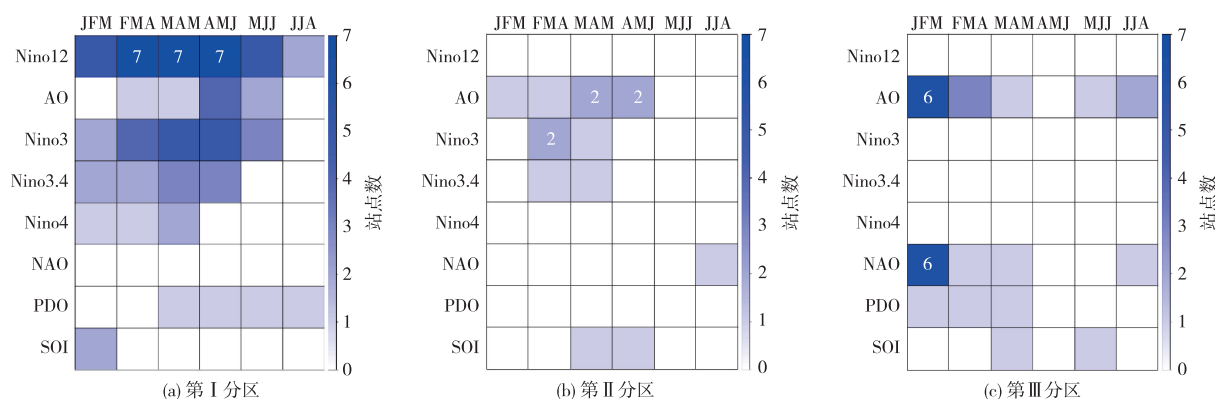
因此,将湘江流域极值降雨序列最终分成3个水文分区,其空间分布如图2(b)所示,定为第I、II和III分区,分别包含的站点数为10、8和18个,各水文分区中心代表站为罗汉庄站、株洲站和炎陵站。

为了更好应用HBM-NRFA模型来预测和量化气候驱动因子对湘江流域同一水文分区极值降雨频率的影响,本文研究了大气环流气候指数与湘江流域极值降雨之间的遥相关关系。采用Spearman秩相关检验法,分析和总结了与各水文分区内大部分站点极值降雨相关性显著的气候指数,结果如图3所示。

由图3(a)可得,第I分区的10个雨量站中,FMA Nino12、MAM Nino12和AMJ Nino12与分区内7个站点极值降雨呈显著正相关关系,而与其他气候指数相关性显著的站点数较少;其中FMA Nino12与第I分区中心代表站极值降雨序列相关性最显著,因此选取FMA Nino12为第I分区的气候驱动因子。同时这也说明第I分区大部分站点的极值降雨主要受FMA Nino12的影响,当太平洋东部赤道附近的海面温度距平指数升高,东太平洋大范围变暖,西南湿暖气流与北方冷空气在长江中下游地区湘江流域

交汇时,会产生持续性强降水。由图3(b)(c)可得,第Ⅱ、Ⅲ分区内与气候指数相关性显著的站点数不到1/2。因此,为较好地应用HBM-NRFA模型,以湘江流域第I水文分区为例进行验证。

基于湘江流域第I水文分区内10个雨量站年最大24 h极值降雨序列,谢平等^[38]认为因年最大24 h极值降雨序列是水文年极值序列,已消除年际周期对序列的影响。因此,本文采用Mann-Kendall秩相关检验法和非参数Pettitt检验法^[39],分别检验极值序列的趋势性和突变点,结果表明除槩梨站、道林站、宁乡站和罗汉庄站呈不显著的上升趋势外,其余6个站点极值降雨序列均呈现显著的上升趋势($\alpha = 0.05$);除槩梨站、道林站和宁乡站无突变点外,其他7个站点极值降雨序列最显著突变年份均集中在1990年左右($\alpha = 0.05$)。



注: JFM、FMA、MAM、AMJ、MJJ和JJA分别表示连续三个月份的平均值,如FMA(February March April)表示2、3和4月的平均值。

图3 湘江流域各水文分区与气候指数显著相关的站点个数

Fig. 3 Numbers of stations in different hydrologic regions of Xiangjiang River Basin whose precipitation are significantly correlated with climate indices

4.2 区域非一致性降雨频率模型构建 应用HBM-NRFA模型来预测和量化气候驱动因子FMA Nino12对湘江流域第I分区内极值降雨频率的影响,并与单站非一致性频率模型(Nonstationary Local Frequency Analysis, NLFA)即无池化模型的计算成果比较,以验证考虑各站极值降雨间相关性的区域非一致性模型能否提高计算精度。

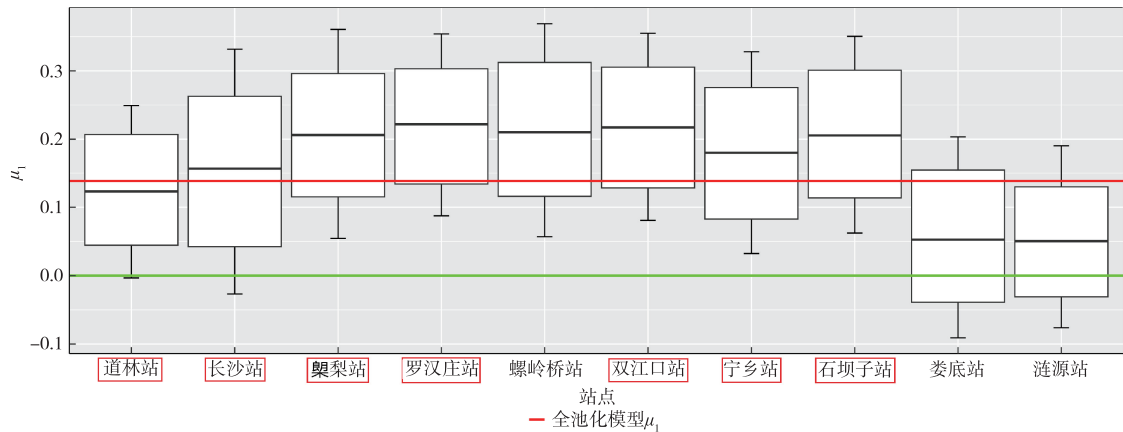
将无池化、部分池化和全池化3种情景的区域非一致性频率分析模型,分别用NLFA、HBM-NRFA_p和HBM-NRFA_f表示。与分区内大部分站点极值降雨显著相关的气候驱动因子为FMA Nino12,假设 $\ln(X_t^*)$ 为站点 s 的第 t 年取对数后的极值降雨,则HBM-NRFA模型中无池化、部分池化和全池化3种情景的构建,如表2所示。

无池化模型不考虑站点间相关性,采用NLFA模型进行计算。图4为无池化情景模型各站点区域参数 μ_1^* 后验分布的箱型图,本文采用箱型图中90%的后验分布值是否大于0,来表征气候驱动因子对极值降雨的影响是否显著。本文取参数 μ_1^* 箱型图中90%的后验分布值大于0表示影响显著,由图4可见8个站点的参数 μ_1^* 后验分布均显著大于0,另两个站点约70%后验分布值显著大于0,与图3(a)中与气候驱动因子FMA Nino12呈显著正相关的7个站点结果基本一致,且此7个站点参数 μ_1^* 的取值(箱型图中黑色横线)相对其他站点来说也较大,表明基于层次贝叶斯的NLFA模型,能较好地模拟气候因子驱动下的极值降雨频率。

对于所有HBM-NRFA模型(包括无池化模型),本文采用PP图(Probability-Probability Plot)评估模型拟合是否合理可靠^[39],即模型的理论频率和经验频率是否接近。依据宋松柏等^[11]提出的计算原理,对存在突变点的极值序列,基于突变年份,将原序列分为两个子序列,假设 m_i 为第 i 个序列 X_i 中大于等于 x 出现的项数,采用期望公式,其经验频率计算公式为: $F(x) = m_1/(n+1) + m_2/(n+1)$;对无突变点,假设 m 为序列 X 中大于等于 x 出现的项数,经验频率计算公式为: $F(x) = m/(n+1)$ 。第I分区内10个站点无池化、部分池化和全池化3种模型的PP图结果表明,各模型的理论频率接近于经验频

表2 HBM-NRFA 模型构建
Table 2 Constructions of HBM-NRFA models

模型	模型结构形式和参数先验分布
NLFA (无池化)	$\ln(X_t^s) \sim N(\mu_0^s + \mu_1^s \times Z, \sigma_s^2); \mu_0^s \sim U(0, 10000), \mu_1^s \sim U(0, 10000)$
HBM-NRFA _p (部分池化)	$\ln(X_t^s) \sim N(\mu_0^s + \mu_1^s \times Z, \Sigma); \mu_0^s \sim U(0, 10000), \mu_1^s \sim N(\mu_{\mu_1}, \sigma_{\mu_1}), \mu_{\mu_1} \sim U(0, 10000), \Sigma_{\mu_1} \sim IW(v_0, \Lambda_0), \Sigma \sim IW(v_1, \Lambda_1)$
HBM-NRFA _F (全池化)	$\ln(X_t^s) \sim N(\mu_0^s + \mu_1 \times Z, \Sigma); \mu_0^s \sim U(0, 10000), \mu_1 \sim U(0, 10000), \Sigma \sim IW(v_1, \Lambda_1)$



注：红色框的站点表示其极值降雨序列与气候驱动因子呈显著正相关关系。

图4 无池化模型(单站NLFA模型)各站参数 μ_1 后验分布的箱型图

Fig. 4 Boxplots of the posterior distribution of coefficients μ_1 for all stations in the no pooling model (local model)

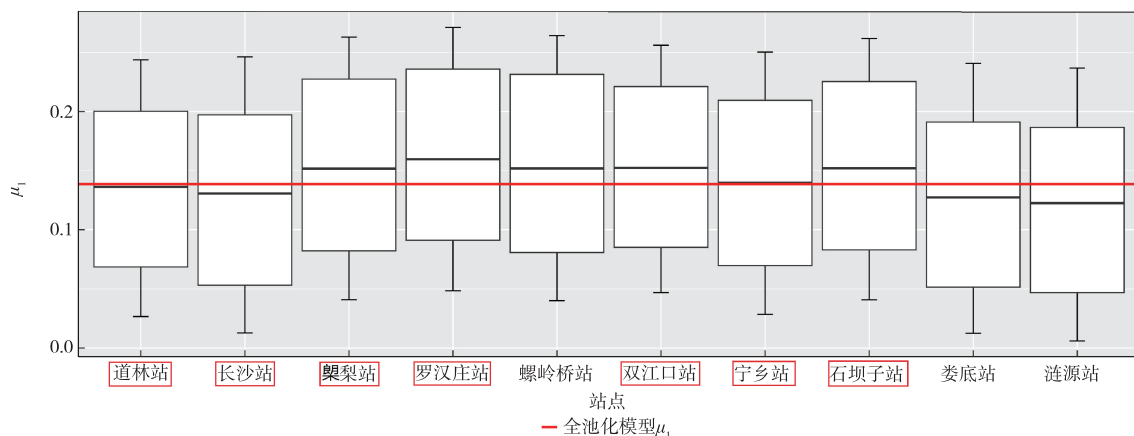
率，散点图均接近于对角线(即1:1线)，即各模型拟合气候因子驱动下的实测极值降雨序列合理可靠^[40-41]。

本文采用DIC值和AIC值来比较各模型拟合极值降雨的效果优劣，结果如表3所示。无池化模型(即单站非一致性模型)的DIC和AIC值最大，因待估参数最多，在模型复杂度上有较大的惩罚项。部分池化和全池化模型的DIC和AIC值显著降低，且部分池化模型的DIC值最小，表明既考虑气候因子驱动区域同质特征，又保留各站极值降雨个体差异特征的部分池化模型最优。

表3 各HBM-NRFA模型的模拟效果比较
Table 3 Model performance comparison with different HBM-NRFA models

	NLFA(无池化)	HBM-NRFA _p (部分池化)	HBM-NRFA _F (全池化)
DIC	373.7	118.5	119.4
AIC	578.2	300.1	305.1

4.3 区域非一致性与单站非一致性模型比较 基于HBM-NRFA模型，得到无池化(单站)、部分池化和全池化等3种情景下表征气候因子FMA Nino12影响程度的区域参数 μ_1^s ，分别如图4—5所示。无池化模型(如图4箱型图中位线)各站区域参数取值范围从最低0.054到最高0.221，而部分池化模型(如图5)区域参数取值范围缩窄至0.12到0.16，且接近于全池化模型参数取值0.138。图5为部分池化与全池化模型区域参数的后验分布比较，部分池化模型各站区域参数的取值，更直观地表现出各站极值降雨对气候驱动因子响应的区域一致性特点；此模型因整合了所有站点的极值降雨信息，将相关性显著和不显著站点对气候因子的响应，趋向于来自同一正态分布的平均气候响应水平且还保持着各站差异。部分池化模型区域参数的取值范围变窄，间接地强化了与气候因子相关性不显著站点的极值降雨也受到同一气候因子的显著影响。

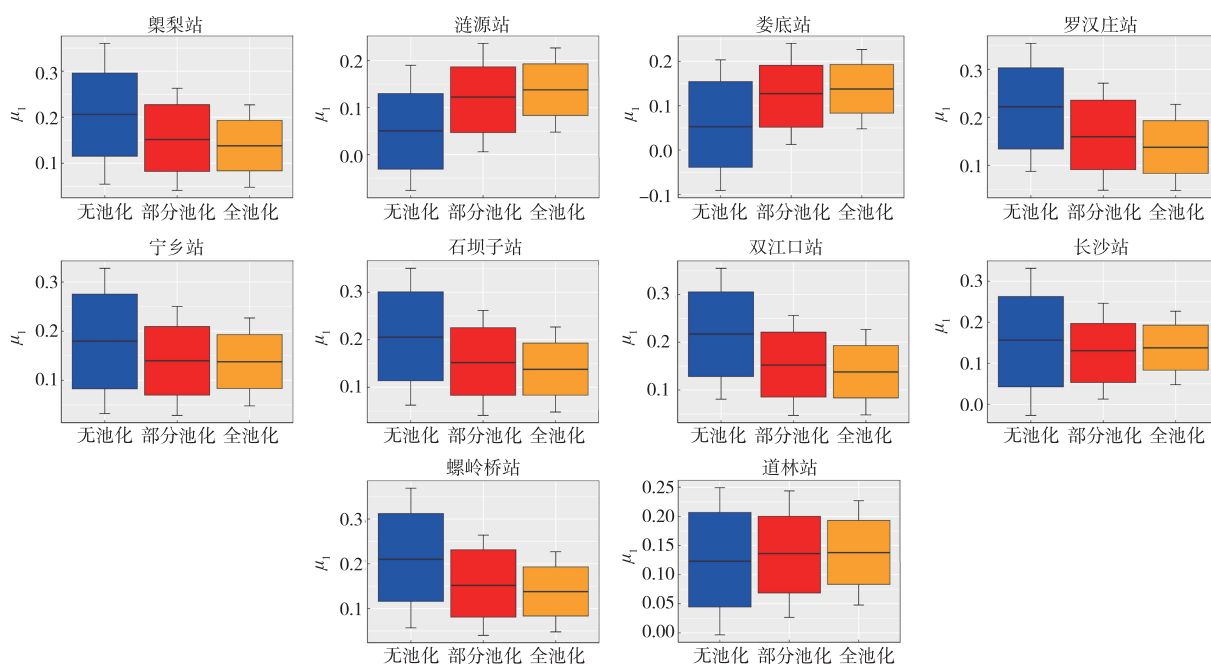


注：红色框的站点表示其极值降雨序列与气候因子呈显著正相关关系。

图5 部分池化模型各站参数 μ_1 后验分布的箱型图

Fig. 5 Boxplots of the posterior distribution of coefficients μ_1 for all stations in the partial pooling model

比较单站非一致性和区域非一致性模型，区域参数 μ_1^* 后验分布的箱型图如图6所示，对于参数不确定性区间，相较无池化(单站非一致)模型，部分池化和全池化模型区域参数90%的贝叶斯不确定性区间显著降低。因单站极值降雨的低信噪比及样本长度有限，单站非一致性存在最大的不确定性；相反，区域非一致性模型整合了同一水文分区的降雨信息，并考虑各站极值降雨间的相关关系，在空间上加强信噪比获得更精确的设计成果；相较部分池化，全池化模型区域参数的不确定性区间有轻微的降低，这是因为综合了各站信息，且各站区域参数取值一样导致参数较少，进而降低不确定性。



注：每个箱型表示了后验分布25%、50%和75%的分位数值。

图6 无池化、部分池化和全池化模型各站参数 μ_1 后验分布的箱型图

Fig. 6 Boxplots of the posterior distribution of coefficients μ_1 for all stations in no pooling, partial pooling and full pooling models

总体而言，区域非一致性部分池化模型区域参数不确定性(后验分布箱型图宽度)的降低，表明基于多层次结构的区域非一致性模型比单站非一致性模型模拟的各站极值降雨频率更优；既考虑各站极值降雨的相关关系进而优化站点重合信息，又保留各站极值降雨响应气候的个体特点，显著提高非一致性极值降雨频率计算精度，因此能更好地描述气候驱动因子对湘江流域下游极值降雨频率的影响。

4.4 非一致性极值暴雨分位数和重现期比较 基于HBM-NRFA模型构建的无池化(单站非一致性)、部分池化和全池化模型,分析3种情景下各站极值降雨分位数随时间和气候驱动因子的变化,以湘江流域第I水文分区中心代表站罗汉庄站和宁乡站为例,计算结果如图7所示。

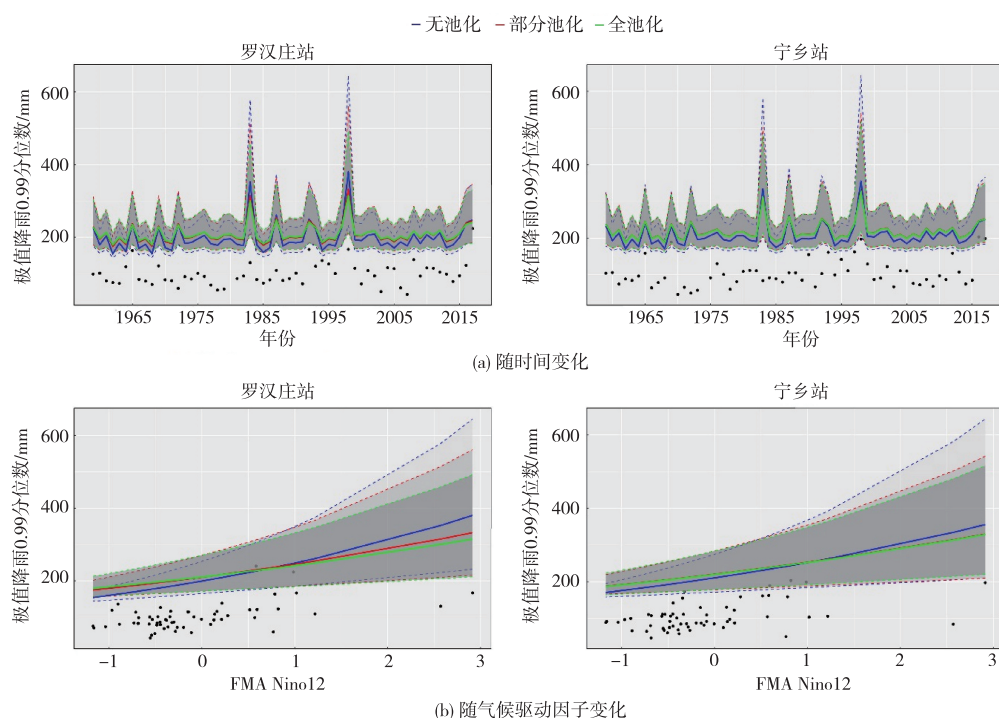


图7 不同模型下0.99极值降雨分位数随时间和气候驱动因子的变化及不确定性区间

Fig. 7 Extreme precipitation with 0.99 quantiles and uncertainties with respect to the year and climate drivers in no pooling, partial pooling and full pooling models

由图7(a)可知,因两个站点部分池化和全池化模型区域参数 μ_i^* 不确定性区间的降低,进而得到极值降雨分位数值不确定性区间的降低;所有的非一致性模型,均能很好地描述和模拟受气候因子驱动下0.99极值降雨分位数(即100年一遇设计暴雨)随年份的变化趋势;从各模型90%的不确定性区间比较,对较大的极值降雨事件,区域非一致性模型(即部分池化和全池化模型)的不确定性区间相较单站非一致性模型要低;对较小的极值降雨,单站非一致性模型的不确定性区间相对要小一些。

由图7(b)可知,随着气候驱动因子FMA Nino12每增加一个单位,两个站点0.99极值降雨分位数会增加近50 mm,且不确定性区间增加速率会更大,尤其是单站非一致性模型,随着气候因子的增加,设计暴雨的不确定区间在不断增大,不利于得到精确的设计成果。区域非一致性模型则不同,随着气候因子的增长,设计暴雨成果也不断增大,但不确定性区间有所降低。总体而言,相较单站非一致性模型,区域非一致性模型模拟气候驱动因子对极值降雨频率影响的表现更优,设计成果的计算精度显著提高。进一步分析表明,随着同年2—4月FMA Nino12事件的发生,海温升高,热带地区蒸发量增加,大气中水汽含量增多,同时可能引发大气环流调整冷空气南下,与西太平洋副高西伸的暖湿气流交汇,以致湘江流域下游区域对流活动增强,导致极值降雨事件频发。

从工程水文实践角度分析,受气候因子驱动的非一致性模型计算得到的重现期,会随着时间变化,很难或不便为工程师们直接利用。因此,本文采用ENE计算单站和区域非一致性模型的重现期,并以一致性重现期作为基准比较。同样以罗汉庄站和宁乡站为例,单站非一致性(无池化)、区域非一致性(部分池化和全池化)模型得到的重现期和传统一致性重现期比较,如图8所示。

两个站点的计算成果相似,以近20年一遇的一致性重现期为分界,20年一遇以下区间单站和区域非一致性重现期都比一致性重现期要大;20年一遇以上区间单站和区域非一致性重现期比一致性重

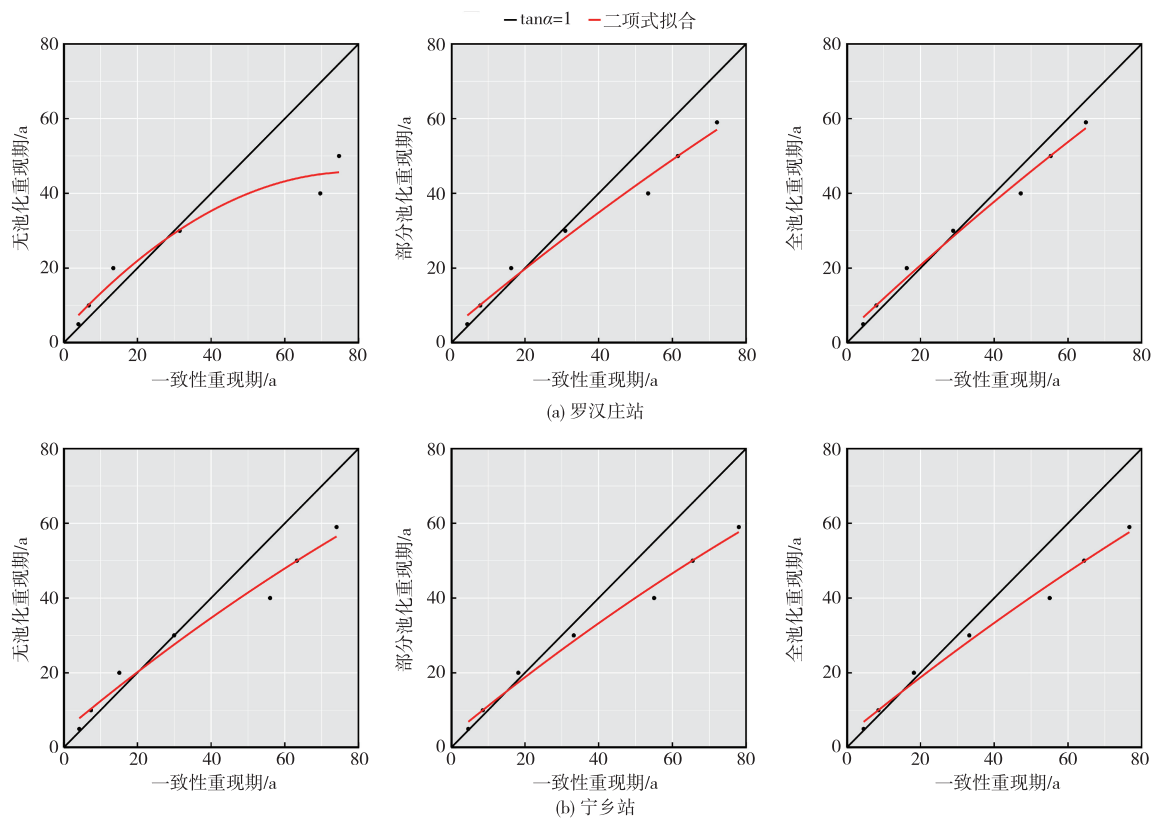


图8 非一致性模型与一致性模型极值降雨重现期比较

Fig. 8 Extreme precipitation return periods comparison between stationary and nonstationary models including no pooling, partial pooling and full pooling models

现期都要小；随着极值降雨的增大，相较一致性条件，受气候因子驱动的非一致性重现期越来越小，即极值降雨事件的发生概率不断增大。值得注意的是，相较20年一遇以上的一致性重现期，非一致性重现期的减小速率越来越大；不断减小的非一致性重现期，在一定程度上表明了湘江流域下游极值降雨事件发生概率受到气候驱动因子FMA Nino12强烈的影响，随着极值降雨越大，发生概率也越大。总体而言，相较一致性条件，非一致性模型计算得到的气候因子和极值降雨频率之间的定量关系能提供更多有效的信息，为工程师和决策者避免低估潜在暴雨风险、制定更好的应急措施提供科学的参考依据。

5 结论

基于遵循极值理论的极值空间聚类水文分区结果，本文提出一种基于层次贝叶斯的区域非一致性水文频率计算模型，以湘江流域为例开展了实例研究，并与单站非一致性模型、一致性模型计算成果进行比较。主要结论如下：

(1) 采用遵循极值理论的空间聚类算法，即基于变差函数 F -madogram的PAM聚类算法，既量化极值序列的空间相关性并保证同一分区内空间相关最近，又保留序列的高偏统计特征，是一种合理有效的、适合极值序列的水文分区方法。

(2) 对同一水文分区，基于层次贝叶斯的区域非一致性模型(HBM-NRFA)具有较灵活和方便的多层结构，其优势体现在以下3个方面：既考虑气候驱动因子协变量(即非一致性)，又考虑高维的各站极值序列相关关系；通过设置区域参数，整合各站极值序列的信息利用程度；同时贝叶斯模拟直接得到不确定性区间，信息的整合能显著降低计算成果不确定性。

(3)本文以湘江流域年最大24 h极值降雨为例进行研究,采用基于 F -madogram的PAM算法,获得3个水文分区,与第I分区大部分站点极值降雨呈显著正相关的气候驱动因子为FMA Nino12。相较一致性重现期,受气候因子影响的非一致性重现期不断减小,表明受同年2—4月FMA Nino12的影响,湘江流域下游地区极值降雨发生概率不断增大。

本文采用的基于层次贝叶斯框架的区域非一致性水文频率分析模型,在未来运用时仍存在一些可进一步深入研究之处。采用贝叶斯理论进行层次模型的参数估计时,对于高维变量序列的联合分布,需采用显式表达的多变量联合分布函数,但非高斯边缘分布以何种方式与层次贝叶斯结构相结合,是未来可研究探索的一个方向。

参 考 文 献:

- [1] MILLY P C D, BETANCOURT J, FALKENMARK M, et al. Stationarity is dead: Whither water management? [J]. Science, 2008, 319(5863): 573–574.
- [2] KHALIQ M N, OUARDA T B M J, ONDO J C, et al. Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review[J]. Journal of Hydrology, 2006, 329(3/4): 534–552.
- [3] YAN L, XIONG L H, GUO S L, et al. Comparison of four nonstationary hydrologic design methods for changing environment[J]. Journal of Hydrology, 2017, 551: 132–150.
- [4] 熊立华, 江聪, 杜涛, 等. 变化环境下非一致性水文频率分析研究综述[J]. 水资源研究, 2015, 4(4): 310–319. (XIONG L H, JIANG C, DU T, et al. Review on nonstationary hydrological frequency analysis under changing environments[J]. Journal of Water Resources Research, 2015, 4(4): 310–319. (in Chinese))
- [5] 谢雨祚, 郭生练, 熊立华, 等. 考虑历史洪水资料的三峡水库运行期非一致性设计洪水估算[J]. 水利学报, 2024, 55(6): 643–653. (XIE Y Z, GUO S L, XIONG L H, et al. Non-stationary design flood estimation of Three Gorges Reservoir in operation period considering historical information[J]. Journal of Hydraulic Engineering, 2024, 55(6): 643–653. (in Chinese))
- [6] 郭生练, 刘章君, 熊立华. 设计洪水计算方法研究进展与评价[J]. 水利学报, 2016, 47(3): 302–314. (GUO S L, LIU Z J, XIONG L H. Advances and assessment on design flood estimation methods[J]. Journal of Hydraulic Engineering, 2016, 47(3): 302–314. (in Chinese))
- [7] 谢平, 陈广才, 雷红富. 变化环境下基于趋势分析的水资源评价方法[J]. 水力发电学报, 2009, 28(2): 14–19. (XIE P, CHEN G C, LEI H F. The assessment method of water resources based on trend analysis of changing environments[J]. Journal of Hydroelectric Engineering, 2009, 28(2): 14–19. (in Chinese))
- [8] 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方法研究[J]. 水力发电学报, 2013, 32(5): 21–25. (HU Y M, LIANG Z M, YANG H Z, 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))
- [9] SINGH K P, SINCLAIR R A. Two-distribution method for flood frequency analysis[J]. Journal of the Hydraulics Division, 1972, 98(1): 29–44.
- [10] 冯平, 李新. 基于Copula函数的非一致性洪水峰量联合分析[J]. 水利学报, 2013, 44(10): 1137–1147. (FENG P, LI X. Bivariate frequency analysis of non-stationary flood time series based on Copula methods[J]. Journal of Hydraulic Engineering, 2013, 44(10): 1137–1147. (in Chinese))
- [11] 宋松柏, 李扬, 蔡明科. 具有跳跃变异的非一致分布水文序列频率计算方法[J]. 水利学报, 2012, 43(6): 734–739. (SONG S B, LI Y, CAI M K. Methods of frequency analysis for hydrologic data with jump up components [J]. Journal of Hydraulic Engineering, 2012, 43(6): 734–739. (in Chinese))
- [12] 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 P, ZENG H, LI X. Non-stationary flood-frequency analysis based on mixed distribution[J]. Journal of Tianjin University: Science and Technology, 2013, 46(4): 298–303. (in Chinese))
- [13] 董立俊, 董晓华, 马耀明, 等. 基于集成时变矩模型和动态蒙特卡洛方法的防洪风险分析[J]. 水利学报, 2024, 55(9): 1084–1097. (DONG L J, DONG X H, MA Y M, et al. Flood control risk analysis based on Integrated Time-Varying Moment model and Dynamic Monte Carlo method[J]. Journal of Hydraulic Engineering, 2024,

- 55(9): 1084–1097. (in Chinese))
- [14] ZENG H, SUN X, LALL U, et al. Nonstationary extreme flood/rainfall frequency analysis informed by large-scale oceanic fields for Xidayang Reservoir in North China[J]. *International Journal of Climatology*, 2017, 37(10): 3810–3820.
 - [15] 孙博凯, 郭生练, 钟斯睿, 等. 考虑长江上游水库群调蓄影响的寸滩站设计洪水计算分析[J]. *水利学报*, 2025, 56(8): 1002–1011. (SUN B K, GUO S L, ZHONG S R, et al. Analysis of design flood estimation at Cuntan station considering regulation impact of upstream reservoir groups on the Yangtze River[J]. *Journal of Hydraulic Engineering*, 2025, 56(8): 1002–1011. (in Chinese))
 - [16] 熊立华, 郭生练, 王才君. 国外区域洪水频率分析方法研究进展[J]. *水科学进展*, 2004, 15(2): 261–267. (XIONG L H, GUO S L, WANG C J. Advance in regional flood frequency analysis from abroad[J]. *Advances in Water Science*, 2004, 15(2): 261–267. (in Chinese))
 - [17] RENARD B. A Bayesian hierarchical approach to regional frequency analysis[J]. *Water Resources Research*, 2011 47(11): 602–610.
 - [18] GREGERSEN I B, MADSEN H, ROSBJERG D, et al. A regional and nonstationary model for partial duration series of extreme rainfall[J]. *Water Resources Research*, 2017, 53(4): 2659–2678.
 - [19] SUN X, LALL U, MERZ B, et al. Hierarchical bayesian clustering for nonstationary flood frequency analysis: Application to trends of annual maximum flow in Germany[J]. *Water Resources Research*, 2015, 51(8): 6586–6601.
 - [20] GREGERSEN I B, MADSEN H, ROSBJERG D, et al. A spatial and nonstationary model for the frequency of extreme rainfall events[J]. *Water Resources Research*, 2013, 49(1): 127–136.
 - [21] WU Y, LALL U, LIMA C H R, et al. Local and regional flood frequency analysis based on hierarchical Bayesian model: application to annual maximum streamflow for the Huaihe River basin[J]. *Hydrology and Earth System Sciences Discussions*, 2018: 1–21.
 - [22] CHEN X, HAO Z, DEVINENI N, et al. Climate information based streamflow and rainfall forecasts for Huai River basin using hierarchical Bayesian modeling[J]. *Hydrology and Earth System Sciences*, 2014, 18(4): 1539–1548.
 - [23] 孙周亮, 王婕, 刘艳丽, 等. 中国水文区划及水文分区方法综述[J]. *水利水电工程学报*, 2024(1): 77–86. (SUN Z L, WANG J, LIU Y L, et al. Evaluation of hydrological zoning technique and its utilization in China[J]. *Hydro-Science and Engineering*, 2024(1): 77–86. (in Chinese))
 - [24] BERNARD E, NAVEAU P, VRAC M, et al. Clustering of maxima: Spatial dependencies among heavy rainfall in France[J]. *Journal of Climate*, 2013, 26(20): 7929–7937.
 - [25] COOLEY D, NAVEAU P, PONCET P. *Variograms for Spatial Max-stable Random Fields*[M]. New York: Springer, 2006.
 - [26] KAUFMAN L R, ROUSSEEUW P J. Finding Groups in Data: An introduction to cluster analysis[J]. *Technometrics*, 1992, 34(1): 111–112.
 - [27] ROUSSEEUW P J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis[J]. *Journal of Computational and Applied Mathematics*, 1987(20): 53–65.
 - [28] DEVINENI N, LALL U, PEDERSON N, et al. A tree-ring-based reconstruction of Delaware River Basin streamflow using hierarchical bayesian regression[J]. *Journal of Climate*. 2013, 26(12): 4357–4374.
 - [29] SUN X, THYER M, RENARD B, et al. A general regional frequency analysis framework for quantifying local-scale climate effects: A case study of ENSO effects on Southeast Queensland rainfall[J]. *Journal of Hydrology*, 2014, 512(6): 53–68.
 - [30] SERAGO J M, VOGEL R M. Parsimonious nonstationary flood frequency analysis[J]. *Advances in Water Resources*, 2018, 112: 1–16.
 - [31] GELMAN A, HILL J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*[M]. Cambridge University Press, 2007.
 - [32] HOFFMAN M D, GELMAN A. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo[J]. *Journal of Machine Learning Research*, 2011: 1593–1623.
 - [33] SPIEGELHALTER D J, BEST N G, CARLIN B P, et al. Bayesian measures of model complexity and fit[J]. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2002, 64(4): 583–639.
 - [34] PAREY S, HOANG T T H, DACUNHA-CASTELLE D. Different ways to compute temperature return levels in the climate change context[J]. *Environmetrics*, 2010, 21(7/8): 698–718.

- [35] SALAS J D, OBEYSEKERA J. Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events[J]. *Journal of Hydrologic Engineering*, 2014, 19(3): 554–568.
- [36] CHAQDID A, TUEL A, FATIMY El A, et al. Extreme rainfall events in morocco: spatial dependence and climate drivers[J]. *Weather and Climate Extremes*, 2023 (40): 100556.
- [37] RAYNER N A, PARKER D E, HORTON E B, et al. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century[J]. *Journal of Geophysical Research: Atmospheres*, 2003, 108(D14): 1063–1082.
- [38] 谢平, 窦明, 朱勇, 等. 流域水文模型—气候变化和土地利用/覆被变化的水文水资源效应[M]. 北京: 科学出版社, 2010. (XIE P, DOU M, ZHU Y, et al. Watershed Hydrological Model—hydrological and Water Resources Effects of Climate Change and Land Use/Land Cover Change[M]. Beijing: Science Press, 2010. (in Chinese))
- [39] 曾杭. 非一致性洪水分析计算及对水利工程防洪影响研究[D]. 天津: 天津大学, 2015. (ZENG H. Non-stationary Flood Frequency Analysis and its Effect on Flood Control[D]. Tianjin: Tianjin University, 2015. (in Chinese))
- [40] SUN X. Analyse propabiliste régionale des précipitations: prise en compte de la variabilité et du changement climatique[D]. Grenoble: Université de Grenoble, 2013.
- [41] WILK M B, GNANADESIKAN R. Probability plotting methods for the analysis for the analysis of data[J]. *Biometrika*, 1968, 55(1): 1–17.

Regional non-stationary frequency analysis based on extreme hydrological regionalization and hierarchical Bayesian framework

ZENG Hang^{1,2}, ZHOU Yang^{1,2}, LI Lingjie³, WANG Guoqing^{3,4}

(1. School of Hydraulic and Ocean Engineering, Changsha University of Science & Technology, Changsha 410114, China;

2. Key Laboratory of Dongting Lake Aquatic Eco-Environmental Control and Restoration of Hunan Province, Changsha 410114, China;

3. The National Key Laboratory of Water Disaster Prevention, Nanjing Hydraulic Research Institute, Nanjing 210029, China;

4. Research Center for Climate Change, Ministry of Water Resources, Nanjing 210029, China)

Abstract: Non-stationary hydrological frequency analysis is a critical scientific research issue for engineering hydrologic design under changing environments. Existing methods exhibit high uncertainty and often fail to account for the spatial dependences among data series of different stations. This paper proposes a regional non-stationary hydrological frequency analysis methodology based on hierarchical Bayesian framework and extreme hydrological regionalization. Firstly, the partitioning around medoids (PAM) clustering algorithm based on the F -madogram variogram is utilized to partition the basin into hydrologically homogeneous regions. And then regional non-stationary hydrological frequency analysis models based on hierarchical Bayesian are constructed. They can be classified into no pooling, partial pooling and full pooling models. Finally, a case study is conducted using the annual maximum 24 hour extreme rainfall data from the Xiangjiang River Basin. Results indicate that based on the spatial clustering algorithm adapting with extreme value theory, the Xiangjiang River Basin with 36 rainfall gauges is divided into three hydrological regions. For Region I, FMA Nino12 is identified as a climatic driver significantly positively correlated with most stations. Compared with the at-site non-stationary model, the partial pooling model exhibits superior performance, effectively capturing the regionally homogeneous response to climatic drivers while preserving individual site characteristics. Simultaneously, the regional models exhibit a great benefit with the uncertainty for regional parameters at each station reduced by about 10% to 35%. Moreover, as the return period increases, the non-stationary return periods consistently decrease, compared with the stationary conditions. For instance, at the 50-year return period, the reduction rate approaches 60% for some stations, indicating an increased probability of extreme rainfall events within the region driven by climatic factors. This methodology enriches the methodological framework of non-stationary hydrological frequency analysis and can provide a scientific basis for determining the design rainstorms in basins and formulating disaster prevention strategies.

Keywords: regional non-stationarity; hydrological frequency analysis; extreme hydrological regionalization; hierarchical Bayesian framework; Xiangjiang River Basin

(责任编辑: 尹 婧)