

文章编号: 1008-2786-(2015)4-425-09

DOI: 10.16089/j.cnki.1008-2786.000053

Bayes 统计模型在出山月均径流极小值研究中的应用

刘友存¹, 霍雪丽^{1,2}, 郝永红¹, 崔玉环³, 韩添丁⁴, 沈永平⁴, 王 建⁴

(1. 天津师范大学 天津市水资源与水环境重点实验室, 天津 300387; 2. 天津师范大学 城市与环境科学学院, 天津 300387;

3. 安徽农业大学 理学院, 安徽 合肥 230036; 4. 中国科学院 寒区旱区环境与工程研究所, 甘肃 兰州 730000)

摘 要: 数理统计方法在解决全球气候变化引起的洪水、干旱等极端水文事件中获得了越来越广泛的应用。选取乌鲁木齐河 1958—2006 年枯水期的月平均出山径流资料, 采用广义 Pareto 极值分布 (GPD) 模型, 并运用 Bayes 统计模型估计 GPD 的参数, 最后对乌鲁木齐河枯水期月均出山径流极小值变化进行了估算。研究表明: 1. 参数的初始值、先验分布的均值分别取其极大似然估计值, 先验分布的标准差取较小值, 随机游走项分布的标准差取较大值, 这种方法能使 Markov 链快速收敛; 2. 基于 Bayes 参数估计值的 GPD 在拟合月均径流量的极小值时具有很高的精确度, 与传统的极大似然估计方法相比, Bayes 统计模型的推断效果较好; 3. 乌鲁木齐河重现期为 10 a、25 a、50 a 和 100 a 的枯水期月均径流极小值分别约为 $0.60 \text{ m}^3/\text{s}$ 、 $0.44 \text{ m}^3/\text{s}$ 、 $0.32 \text{ m}^3/\text{s}$ 和 $0.20 \text{ m}^3/\text{s}$; 4. 100 a 重现水平的 95% 置信区间的下限为 $-0.238 \text{ m}^3/\text{s}$, 说明当乌鲁木齐河在枯水期遇上百年一遇的极小值时, 有可能出现断流的情况。

关键词: 径流极小值; 广义 Pareto 分布; Markov Chain Monte Carlo (MCMC) 方法; 乌鲁木齐河

中图分类号: P333.3

文献标志码: A

IPCC 第五次报告指出, 上个世纪以来, 全球变暖引起的水循环加剧, 导致极端水文事件明显增加^[1-4], 给环境和社会经济发展造成了重大损失。这些极端事件将在本世纪变得更加频繁、普遍甚至剧烈^[4-7]。区域乃至全球尺度的气候变化对洪水、干旱等极端事件的影响引起了广大学者的极大关注^[6,8-9]。许多学者应用数理统计方法对流域的径流极值变化进行了分析。Wilby 等 2006 年对英国泰晤士河^[10]、徐若兰等 2010 年对汉江流域上游^[11]、Xia 等 2012 年对淮河流域^[8]、周旭东等 2013 年对黄河源区^[12]、Liu 等 2015 年分别应用 GPD 模型和 ARMA 模型对乌鲁木齐河上游^[13-14] 这些典型流域的径流极值特征应用数理统计方法进行了分析; 此外, Xu 等 2008 年利用统计降尺度法与 SWAT

模型耦合技术^[15]、Kumar 等 2010 年运用空间离散和模型预测的方法^[16]、Müller 等 2011 年采用降尺度模型^[17]、蓝永超等 2011 年采用线性趋势法及多元线性与非线性回归方法^[18]、Paquet 等 2013 年应用半连续降水径流模型^[19] 对极端水文事件的变化趋势进行了研究、王跃峰等 2014 年运用于多时间尺度 SPI 对闽江流域的干湿变化与洪旱事件进行了有效识别^[20]。这些研究取得了较好的成果, 然而有关气候变化下水文极端事件的模拟和预估亟需进一步加强^[21]。

Bayes 统计推断较传统的统计方法能更好得处理复杂非线性问题的求解^[22-23]。然而参数后验分布总结和探索方面的困难, 阻碍了 Bayes 统计推断在水文上的广泛应用。近些年来, MCMC 方法算法

收稿日期(Received date): 2014-03-15; 修回日期(Accepted): 2015-05-12。

基金项目(Foundation item): 国家自然科学基金(41471001、41272245、41401022、41001006); 中国博士后科学基金项目(20100480444)。[The National Nature Science Foundation of China (41471001、41272245、41401022、41001006); China Postdoctoral Science Foundation (20100480444)。]

作者简介(Biography): 刘友存(1977-), 男, 河北迁安人, 理学博士学位, 主要从事寒旱区水文水资源、水文气象等方面的研究工作。[Liu Youcun, (1977-), male, Born in Qian'an of Hebei Province, Phd, mainly engaged in hydrology and water resources in cold and arid regions, and hydrometeorology.] Email: liuyoucun@gmail.com; yliu@ipgp.fr

的进步使其能够对 Bayes 公式中参数的后验分布进行直接抽样,这一问题得到了解决^[24]。Haario 等人 2001 年介绍了 MCMC 方法实际计算中易于应用的一种适应性 Metropolis 算法^[25]。Bates 等人 2001 年用 Bayes 方法对澳大利亚东南两个集水区的 8 参数概念降雨径流模型进行了参数估计和推断^[26]。Blasone 等人 2008 年运用 SCEM-UA 算法对参数的先验空间进行采样,从而提升了通用似然不确定性估计 (GLUE) 方法的计算效率,并通过三种不同的概念集水区模型证实了这种改进的 GLUE 方法的优越性^[27]。卫晓婧等人 2009 年提出了基于 MCMC 算法的改进 GLUE 方法,并将其应用在汉江玉带河流域的水文模型中^[28]。尽管 Bayes 统计推断在水文上已有应用,但是在径流极值估算尤其是西北内陆径流极值重现水平的估算方面还不多见^[29]。本文研究目的是探讨 Bayes 统计模型在出山径流极值研究中的适用性,重点分析乌鲁木齐河(乌河)上游枯水期径流极小值的时空变化规律和概率统计特征,预估未来乌河流域径流极小值的变化,为开展气候变化对西北内陆河流域水资源影响的研究提供基础和支撑。

1 研究区概况

乌河发源于东天山中段的天格尔 II 峰北坡的 1 号冰川末端,流域范围在 $86^{\circ}45' \sim 87^{\circ}56'E$ 、 $43^{\circ}00' \sim 44^{\circ}07'N$, 全长 214 km,流域总面积 $4\ 684\text{ km}^2$ ^[18,30]。其中,出山口(英雄桥水文站)以上河流长约 63 km,流域面积 924 km^2 ,平均海拔 3 083 m,山区多年平均出山径流量约为 $2.43 \times 10^8\text{ m}^3$ ^[31]。流域内(出山口以上区域)多年平均气温约为 2°C ^[14],年平均气温呈现上升趋势,约为 $0.22^{\circ}\text{C}/(10\text{ a})$;年降水量为 400~500 mm,集中在 6—8 月,占 60~80%。

2 研究方法

2.1 数据来源

本文选取了英雄桥水文站 1958—2006 年的月平均径流资料,该资料经过新疆水文水资源局整编。英雄桥水文站位于 $44^{\circ}22'01"N$, $87^{\circ}12'16"E$,海拔 1 920 m。2007 年开始在英雄桥水文站以上 5 km 处修建了大西沟水库,随之径流受到人为调节。因

此,本文选取的资料时间范围为 1958—2006 年。其中,丰水期(5—9 月)的多年平均径流占全年多年平均径流的 84.72%,枯水期(1—4 月和 10—12 月)的多年平均径流占全年多年平均径流的 15.28%。

2.2 广义 Pareto 极值分布 (GPD) 模型

假设 $\{X_1, \dots, X_n\}$ 是独立同分布的随机变量序列。令 $M_n = \max\{X_1, \dots, X_n\}$, 如果存在规范化数列 $\{a_n > 0\}$ 、 $\{b_n\}$, 使得对足够大的 n , $(M_n - b_n)/a_n$ 服从广义极值分布,则对于足够大的阈值 u , 在 $X > u$ 的条件下, $Y = X - u$ 的分布近似为广义 Pareto 分布 (GPD)

$$G(y) = 1 - (1 + \frac{\xi y}{\sigma})^{-1/\xi} \quad y > 0 \text{ 且 } (1 + \frac{\xi y}{\sigma}) > 0 \quad (1)$$

式中 $\xi \in R$ 为形状参数, $\sigma > 0$ 为尺度参数。当 $\xi \geq 0$ 时,径流量没有上限值。尺度参数 σ 表示超过阈值 u 的径流量的波动程度。如果 σ 大,表明超阈值的径流量变化大,反之,则超阈值的径流量的变化较小。

2.3 阈值选取

阈值选取通常有两种方法:平均剩余寿命图和判断阈值改变引起的参数估计量的变化。

2.4 参数估计

GPD 模型的参数估计采用基于 Bayes 统计模型的 MCMC 采样方法。该方法直接产生参数的模拟序列。

2.4.1 Bayes 统计模型

Bayes 公式的函数密度形式如下

$$\pi(\theta|\mathbf{x}) = \frac{f(\mathbf{x}|\theta)\pi(\theta)}{\int_{\Theta} f(\mathbf{x}|\theta)\pi(\theta)d\theta} \quad (2)$$

式中 $\pi(\theta|\mathbf{x})$ 为参数 θ 的后验密度函数; $f(\mathbf{x}|\theta)$ 为样本密度函数 $f(\mathbf{x}|\theta) = \prod_{i=1}^n f(x_i|\theta)$; $\pi(\theta)$ 为参数 θ 的先验密度函数; Θ 为 θ 的参数空间。

使用 Bayes 公式 (2) 的主要困难是关于公式 (2) 中分母的积分计算。MCMC 方法提供了解决这一问题的有效途径。

2.4.2 MCMC 方法

本文采用 Metropolis-Hastings 随机游走采样算法来产生模拟序列 $\theta_0, \theta_1, \theta_2, \dots$, 其中 θ_0 为任意初始值。这一模拟序列称为参数 θ 的 Markov 链。Metropolis-Hastings 随机游走采样算法如下:

先由提案分布 $q(\cdot|\theta_i)$ 产生一个建议值 θ^* , 并以概率 α_i 接受它为下一个 θ_{i+1} , 即

$$\theta^* = \theta_i + \varepsilon(\theta_i) \quad (3)$$

$$\theta_{i+1} = \begin{cases} \theta^* & \mu \leq \alpha_i \\ \theta_i & \mu > \alpha_i \end{cases} \quad (4)$$

$$\alpha_i = \min \left\{ 1, \frac{f(\mathbf{x}|\theta^*) \pi(\theta^*)}{f(\mathbf{x}|\theta_i) \pi(\theta_i)} \right\} \quad (5)$$

式中 $\varepsilon(\theta_i)$ 是 θ_i 的随机游走项,服从 $N(0, \omega(\theta))$, $\omega(\theta)$ 为方差。 u 是 $[0, 1]$ 均匀分布的随机数。

实际证明,对不同的初值 θ_0 ,经过一段时间的迭代后,可以认为此时的 θ_i 的边缘分布就是 $\pi(\theta|\mathbf{x})$,此时 Markov 链收敛。因此只需抛掉收敛之前的 $\theta_1, \dots, \theta_k$,用以后的 $\theta_{k+1}, \dots, \theta_n$ 作为后验分布 $\pi(\theta|\mathbf{x})$ 的样本即可。其中 $\theta_1, \dots, \theta_k$ 称为 Markov 链的预烧期, k 称为预烧长度, n 称为 Markov 链的运行长度。

2.4.3 Markov 链收敛性诊断

本文选取定性和定量的方法来诊断 Markov 链的收敛性^[32]。

定性的方法包括模拟抽样路径图和自相关性图。通过观察图中统计量的特征,判断 Markov 链是否达到收敛。定量的方法包括建议值 θ^* 的接受率和 Raftery-Lewis 方法。建议值 θ^* 的接受率 $r = \frac{n_a}{n}$ 表示在 Markov 链的整个运行过程中,建议值 θ^* 被接受的次数 n_a 与运行长度 n 的比值。如果 r 在 $0.2 \sim 0.5$,则认为 Markov 链收敛。Raftery-Lewis 方法给出 Markov 链的预烧期和依赖因子 l 。如果 $l \leq 5$,则认为 Markov 链到达收敛。

2.5 重现水平预测

由公式(1)可以得到随机变量 X 的 N 年重现水平

$$x_N = u_0 + \frac{\hat{\sigma}}{\hat{\xi}} [N n_x \zeta_{u_0}^{\hat{\xi}} - 1] \quad (6)$$

式中 u_0 为选定的阈值, n_x 表示随机变量每年的观测次数, ζ_{u_0} 为超阈值的观测值数占观测值总数的比例, $\hat{\sigma}$, $\hat{\xi}$ 分别为尺度参数和形状参数基于 Bayes 统计模型的 MCMC 采样值。重现水平 x_N 表示平均 N 年出现一次的径流量,选取不同的重现期 N 值,由公式(6),可以得到对应的重现水平 x_N 。

3 结果与讨论

乌河作为乌鲁木齐市重要的水源地,其水量约占乌鲁木齐市地区总地表水量的 40%。由于乌河枯水期的径流量较小,其变化受到了人们广泛的关

注。在当前气候变化的背景下,研究乌河径流极小值的变化对当地居民的生产生活具有十分重要的意义。本文选取乌河上游英雄桥水文站 49 a (1958—2006 年) 枯水期 (1—4 月和 10—12 月) 的月均径流数据,如图 1 所示。在这 49 a 中,枯水期最小的月均径流量是 $0.32 \text{ m}^3/\text{s}$ 。考虑到 GPD 求极大值的特点,取枯水期月均径流序列的相反数,得到序列 $X_i, i=1, \dots, 343$ 。月均径流量经过取反,最小值变成了最大值,由此,我们可以运用 GPD 求枯水期月均径流量的极小值。

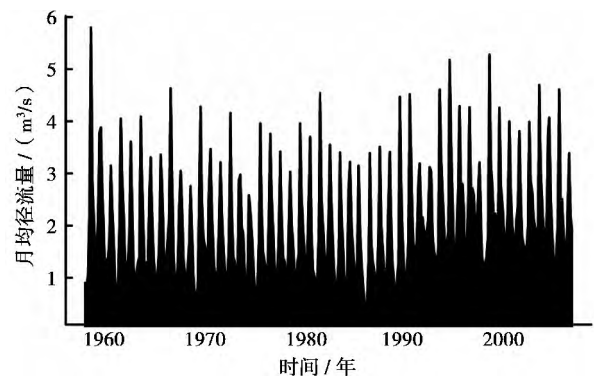


图1 1958—2006 年枯水期月均径流量变化图

Fig.1 Plot of monthly average runoff during dry periods from 1958 to 2006

3.1 阈值确定

首先通过平均剩余寿命图和判断阈值变化引起的修正的尺度参数 σ^* 和形状参数 ξ 的估计量的改变来选取合理的阈值 u_0 。平均剩余寿命图 (未给出) 显示,当阈值在 $[-1.3, -1.03]$ 内变化时,样本平均超出量与阈值之间表现出明显的线性关系。当阈值在 $[-1.06, -0.63]$ 内时,修正的尺度参数 σ^* 的估计量以 0.22 为中心上下波动,形状参数 ξ 的估计量以 0 为中心上下波动。因此,在 $[-1.06, -1.03]$ 之间选取合理的阈值 u_0 。根据超阈值样本个数尽量多的原则,选阈值 $u_0 = -1.05$ 。在序列 $X_i, i=1, \dots, 343$ 中,有 53 个值超出阈值 u_0 ,超阈值的比例 $\zeta_{u_0} = 0.155$ 。记 $x_i, i=1, \dots, 53$ 为超阈值序列。

确定阈值后,假定超阈值 $x_i + 1.05$ 服从 GPD,并用基于 Bayes 统计模型 MCMC 方法估计 GPD 模型中的尺度参数 σ 和形状参数 ξ 。

3.2 参数估计

3.2.1 参数 Markov 链的生成

考虑到公式(1)中的尺度参数 $\sigma > 0$,设 $\eta =$

$\ln \sigma$ 。 η ξ 的先验分布都取正态分布。它们提案分布中的随机项 $\varepsilon(\cdot)$ 同样取为正态分布。为了使得参数的 Markov 链快速收敛,令 η ξ 的初始值、先验分布的均值分别取它们的极大似然估计值,先验分布的标准差取较小值,随机游走项分布的标准差取较大值。上述结果见表 1。

令

$$\boldsymbol{\theta} = (\eta \ \xi) = (\boldsymbol{\theta}[1] \ \boldsymbol{\theta}[2]) \quad (7)$$

由于尺度参数 σ 和形状参数 ξ 相互独立,则参数向量 $\boldsymbol{\theta}$ 的先验分布为其各分量先验分布的乘积,即

$$\pi(\boldsymbol{\theta}) = \prod_{j=1}^2 \frac{1}{\sigma_{\boldsymbol{\theta}[j]} \sqrt{2\pi}} \exp \left\{ -\frac{(\boldsymbol{\theta}[j] - \mu_{\boldsymbol{\theta}[j]})^2}{2 \sigma_{\boldsymbol{\theta}[j]}^2} \right\} \quad (8)$$

式中 $\sigma_{\boldsymbol{\theta}[j]}^2$ $\mu_{\boldsymbol{\theta}[j]}$ 分别表示参数向量 $\boldsymbol{\theta}$ 的第 j 个分量先验分布的方差和均值 $j=1, 2$ 。

超阈值样本 $x_i, i=1, \dots, 53$ 的密度函数:

$$f(\mathbf{x}|\boldsymbol{\theta}) = \prod_{i=1}^{53} f(x_i|\boldsymbol{\theta}) = \frac{1}{\exp(\eta)} \left(1 + \xi \frac{(x_i - u_0)}{\exp(\eta)} \right)^{-\left(\frac{1}{\xi} + 1\right)}, \quad \left(1 + \xi \frac{(x_i - u_0)}{\exp(\eta)} \right) > 0 \quad (9)$$

接着由公式 (3) ~ (5) 可以确定参数向量 $\boldsymbol{\theta}$ 在第 $i+1$ 步的值 $\boldsymbol{\theta}_{i+1}$ 。最后通过 $\sigma = \exp(\eta)$ 变换,可

以得到尺度参数 σ 在每一步的迭代值。

取运行长度 n 为 5 000 按上述步骤,我们可以得到 GPD 模型中尺度参数 σ 和形状参数的 Markov 链。

表 1 MCMC 抽样过程中参数的初始值、先验分布和建议分布

Table 1 Initial value, prior distribution and proposal distribution used in MCMC sampling

参数	初始值	先验分布 (正态分布)		随机游走项分布 (正态分布)	
		均值	标准差	均值	标准差
η	-1.61	-1.61	0.008	0	0.5
ξ	-0.11	-0.11	0.005	0	0.55

3.2.2 参数 Markov 链的收敛性诊断

首先通过抽样路径图和自相关性图来诊断两参数的 Markov 链是否收敛。尺度参数 σ 和形状参数 ξ 的 Markov 链的抽样路径图,分别如图 2a 和图 2b 所示。

由图 2 可以看出,尺度参数 σ 和形状参数 ξ 的采样值在各自的 Markov 链的运行过程中,分别以 0.202 和 -0.109 为中心上下强烈得摆动,并且遍及参数空间。同时,做出两参数采样值的自相关性图,如图 3 所示。

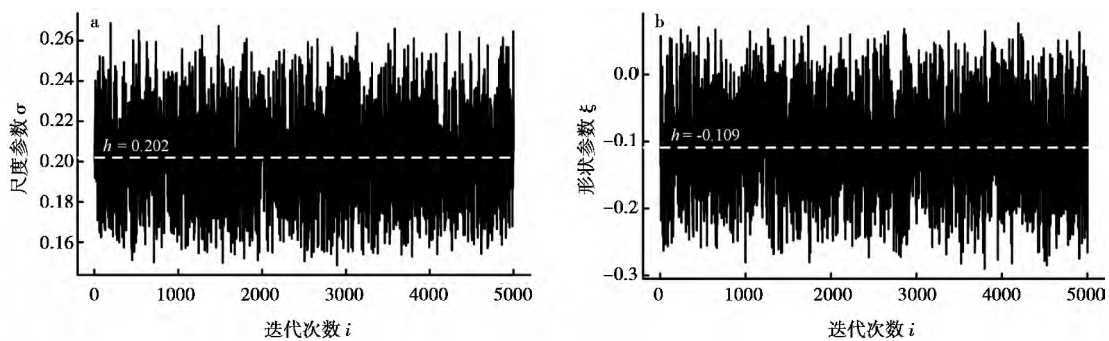


图 2 尺度参数 σ 和形状参数 ξ 的 MCMC 抽样路径图

Fig. 2 Sampling trace plots of scale parameter σ and shape parameter ξ achieved by MCMC method

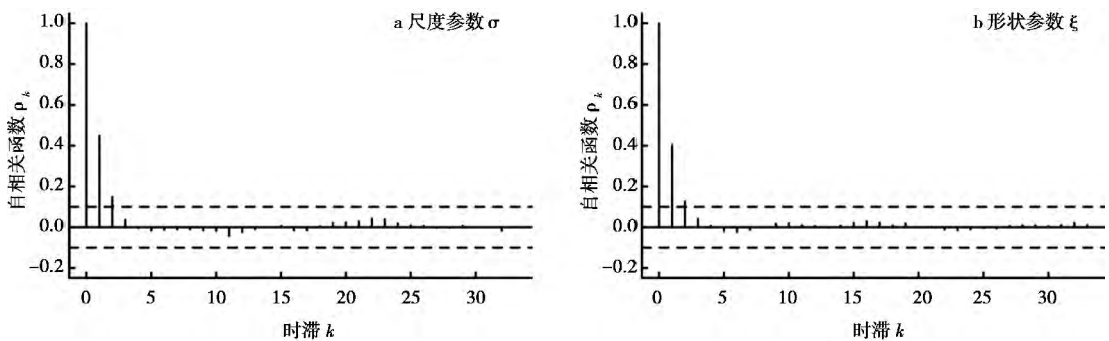


图 3 尺度参数的 σ 和形状参数 ξ MCMC 采样值的自相关性图

Fig. 3 Autocorrelation plots of MCMC samplings of scale parameter σ and shape parameter ξ

在图 3 中可以看到,随着时滞 k 的增加,尺度参数 σ 和形状参数 ξ 的 MCMC 采样值的自相关函数 ρ_k 迅速衰减。当 $k \geq 3$ 时,图 3a 和图 3b 中的自相关函数 $-0.1 < \rho_k < 0.1$ (虚线标出)。此外,在尺度参数 σ 和形状参数 ξ 的 Markov 链的运行过程中,建议值 $\theta^* = (\eta^*, \xi^*)$ 的接受率 $r = 0.497$,在 $0.2 \sim 0.5$ 。同时,采用 Raftery - Lewis 对两参数的 Markov 链进行检验,结果见表 2。表 2 中显示尺度参数 σ 的 Markov 链的依赖因子是 1.81,形状参数 ξ 的 Markov 链的依赖因子是 2.03,均 < 5 。此外,两参数 Markov 链的预烧长度与总运行长度(5 000)相比,非常短,说明设置的初值,算法能够使 Markov 链迅速收敛。

鉴于上述四种诊断方法的结果,我们认为两参数预烧期后的 Markov 链收敛。

表 2 Raftery - Lewis 方法对参数 Markov 链收敛性的检验结果
Table 2 Test result of convergence of parameters' Markov chains
by Raftery - Lewis method

参数	预烧长度	依赖因子 l
σ	6	1.81
ξ	7	2.03

3.2.3 确定参数的估计值

基于预烧期后的参数模拟值,我们对尺度参数 σ 和形状参数 ξ 做出有关的统计推断。分别计算两参数预烧期后的 Markov 链的均值和方差作为参数

的估计值和标准差,可以得到 GPD 的参数估计值 $\hat{\sigma}$ 和 $\hat{\xi}$ 分别为 0.202 和 -0.109 ,它们的标准差分别为 0.023 和 0.069。尺度参数的估计值 $\hat{\sigma}$ 为 0.202,标准差为 0.023,说明 < 1.05 的月均径流量变化不大。

3.3 GPD 模型检验

模型检验图包括概率图和分位数图,如图 4 所示。

在概率图(4a)和分位数图(4b)中,空心点和实心点分别表示基于极大似然参数估计值和 Bayes 参数估计值得到的概率点和分位数点。从概率图(4a)中可以看出,空心点和实心点都接近于斜率为 1 的直线。经计算,空心点和实心点与理论概率值的相关系数均为 0.9928,因此两者基本没有差别。在分位数图(4b)中,空心点和实心点在直线的中下部表现一致,但在直线上部,空心点的 5 个最高点中,只有一个点落在了斜率为 1 的直线上,而实心点的 5 个最高点中有 3 个落在斜率为 1 的直线上,且实心点的最高点比空心点的最高点更接近于直线。经计算,实心点与理论分位数的相关系数为 0.9925,高于空心点与理论分位数的相关系数为 0.9910,说明基于 Bayes 参数估计值的 GPD 模型在拟合高分位数时比基于极大似然参数估计值的 GPD 模型更为准确。

3.4 重现水平估计

在运行长度为 5 000 的尺度参数 σ 和形状参数 ξ 的 Markov 链中,由于形状参数 ξ 的 Markov 链的预

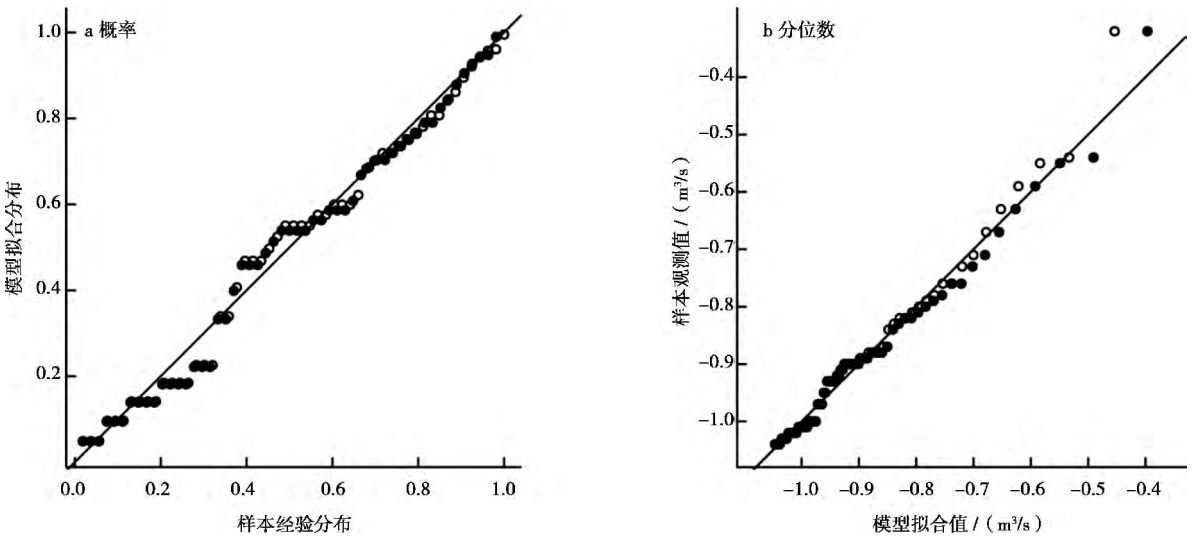


图 4 分别基于极大似然参数估计值和 Bayes 参数估计值的 GPD 模型检验图: 概率图和分位数图

Fig. 4 Model diagnosis plots of GPD based on parameter estimates achieved by maximum likelihood and Bayesian method respectively: probability plot and quantile plot

烧长度为7(表2)我们取迭代次数从8到5 000的尺度参数 σ 和形状参数 ξ 的样本值根据公式(6),可以得到重现期为10 a、25 a、50 a和100 a的重现水平,其密度图如图5所示。

图5显示,随着重现期的增加,密度函数最大值对应的径流量逐渐变小,说明枯水期月均径流量的最小值在不断减少。此外,从图(5d)中可以看到,100 a重现水平即枯水期月均径流最小值的估计中存在 <0 的估计值,说明当乌河在枯水期遇上百年一遇的极小值时,有可能出现断流的情况。

根据图5得到10 a、25 a、50 a和100 a重现期的重现水平及其相应的95%置信区间,如表3所示。

由表3可以看出,重现期为10 a、25 a、50 a和100 a的枯水期月均径流极小值分别约为 $0.60 \text{ m}^3/\text{s}$ 、 $0.44 \text{ m}^3/\text{s}$ 、 $0.32 \text{ m}^3/\text{s}$ 和 $0.20 \text{ m}^3/\text{s}$ 。值得注意的是,50 a重现水平的估算值为 $0.322 \text{ m}^3/\text{s}$,与过去49 a枯水期最小月均径流量 $0.32 \text{ m}^3/\text{s}$,仅仅相差 $0.002 \text{ m}^3/\text{s}$,说明基于Bayes统计模型的参数估计值拟合的GPD模型在估计乌河枯水期月均径流极小值方面具有很高的精确度。此外,100 a重现水平的估计值为 $0.199 \text{ m}^3/\text{s}$,这与刘友存等2013年对乌河^[14]百年一遇径流极小值的分析结果仅相差约 $0.02 \text{ m}^3/\text{s}$ 。进一步分析发现,随着重现期的增加,

重现水平的估计值在不断减小,而相应的95%置信区间的长度在增大,说明随着重现期的增加,相应的重现水平的变化范围在增大,其不确定性增加。

由于100 a重现水平估计值的95%置信区间下限为 $-0.238 \text{ m}^3/\text{s}$,明显 <0 ,说明1958年以来,当乌河在枯水期遇上百年一遇的极小值时,有可能出现断流的情况。

4 结 论

本文选择具有长期观测基础的乌河上游作为研究区,运用1958—2006年间的枯水期(1—4月和10—12月)的径流资料,讨论了Bayes统计模型在预测径流极值变化中的应用,并对乌河上游月均径流极小值的重现水平进行了预测,得到以下结论:

1. 通常情况下,Markov链很难快速收敛。本文中,令参数的初始值、先验分布的均值分别取其极大似然估计值,先验分布的标准差取较小值,随机游走项分布的标准差取较大值。结果表明这种做法能使Markov链快速收敛。

2. 通过比较基于Bayes参数估计值的GPD模型检验图和基于极大似然参数估计值的GPD模型检验图发现,两者在概率估计方面精度一样,但前者在拟合高分位数时比后者更为准确,说明基于Bayes

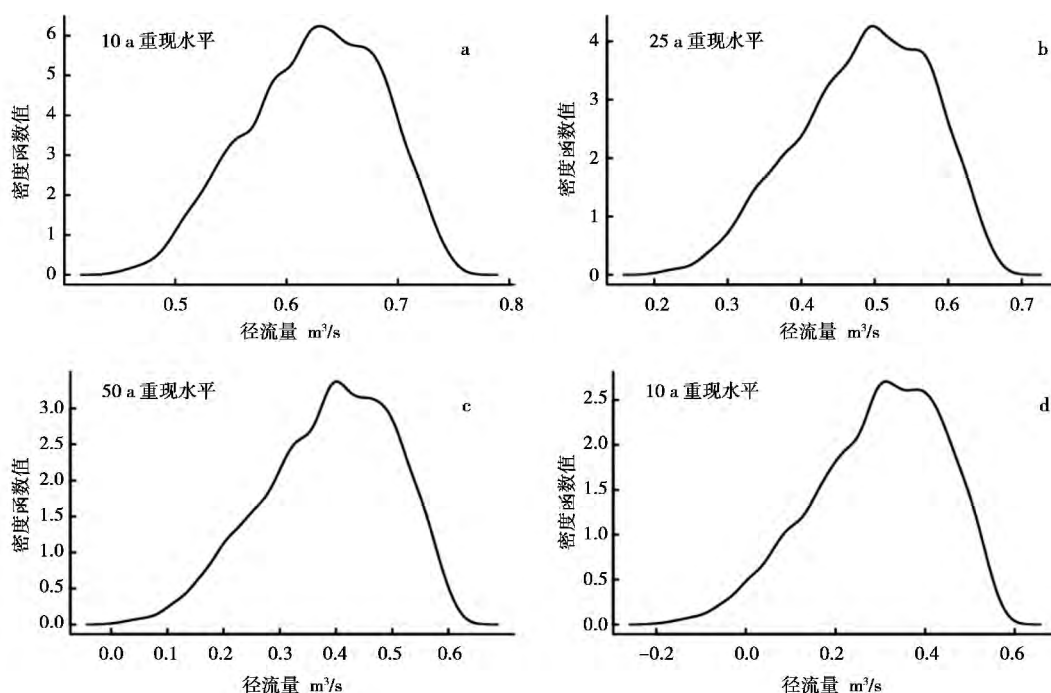


图5 乌河枯水期月均径流极小值10 a、25 a、50 a和100 a重现水平的密度函数图

Fig. 5 Density plots of 10 a, 25 a, 50 a, and 100 a return level of monthly average runoff minimum in dry periods

表 3 乌河枯水期月均径流极小值不同重现水平估算

Table 3 Different return level estimates of monthly average runoff minimum in dry periods

重现期 /a	95% 置信区间下限 /(m ³ /s)	估算值 /(m ³ /s)	95% 置信区间上限 /(m ³ /s)	置信区间长度 /(m ³ /s)
10	0.424	0.602	0.781	0.357
25	0.172	0.442	0.712	0.540
50	-0.026	0.322	0.671	0.697
100	-0.238	0.199	0.637	0.875

统计模型的统计推断有较好的效果。因此,使用 Bayes 估计方法,是对经典估计方法的一个较好的补充。

3. 基于 Bayes 统计模型,采用 MCMC 抽样方法,得到 GPD 模型的参数模拟序列,进而得到重现期分别为 10 a、25 a、50 a 和 100 a 的枯水期径流极小值的密度分布。结果表明,重现期为 10 a、25 a、50 a 和 100 a 的枯水期月均径流极小值分别约为 0.60 m³/s、0.44 m³/s、0.32 m³/s 和 0.20 m³/s。由于 100 a 重现水平估计值的 95% 置信区间下限值明显 < 0,说明当乌河在枯水期遇上百年一遇的极小值时,有可能出现断流的情况。

参考文献(References)

[1] IPCC. Working Group I Contribution to the IPCC Fifth Assessment Report (WGI AR5), Climate Change 2013: The Physical Science Basis: Summary for Policymakers [R/OL]. [2013-10-28]. http://www.climatechange2013.org/images/report/WG1AR5_SPM_FINAL.pdf

[2] Liu Weilong, Zhao Hui, Wang Xiaodan, et al. Review and evaluation of the effect of the climate change on the high altitude wetland ecosystem in Tibet Plateau [J]. Mountain Research, 2014, 32(4): 481-487 [刘伟龙, 赵慧, 王小丹, 等. 气候变化下西藏高寒湿地生态系统研究的意义和特点 [J]. 山地学报, 2014, 32(4): 481-487]

[3] Huziy N O, Sushama L, Khaliq M N, et al. Analysis of streamflow characteristics over Northeastern Canada in a changing climate [J]. Climate Dynamics, 2013, 40: 1879-1901

[4] Dumas P, S. Hallegatte, P. Quintana-Segui, E. Martin. The influence of climate change on flood risks in France—first estimates and uncertainty analysis [J]. Natural Hazards and Earth System Sciences, 2013, 13(3), 809-821

[5] Peterson B J, Holmes R M, McClelland J W, et al. Increasing river discharge to the arctic ocean [J]. Science, 2002, 298: 2171-2173

[6] Douglas A Burns, Julian Klaus, Michael R McHale. Recent climate trends and implications for water resources in the Catskill Mountain

region, New York, USA. Journal of Hydrology, 2007, 336(1), 155-170

[7] Liu Y, Ye B, Metivier F, et al. Preliminary study on reaction of mountain runoff to climate change in Urumqi River, China [A]. In: International Symposium on Water Resource and Environmental Protection, vol. 3: 2350-2354. doi: 10.1109/ISWREP.2011.5893739

[8] Xia J, Du H, Zeng S, et al. Temporal and spatial variations and statistical models of extreme runoff in Huaihe River Basin during 1956—2010 [J]. Journal of Geographical Sciences, 2012, 22(6): 1045-1060

[9] Chen Yaning, Xu Changchun, Yang Yuhui, et al. Hydrology and water Resources variation and its responses to regional climate change in Xinjiang [J]. Acta Geographica Sinica, 2009, 64(11), 1331-1341 [陈亚宁, 徐长春, 杨余辉, 等. 新疆水文水资源变化及对区域气候变化的响应 [J]. 地理学报, 2009, 64(11), 1331-1341]

[10] Wilby R L, Harris I. A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK [J]. Water Resource Res, 2006, 42: W02419.

[11] Xu Ruolan, Chen Hua, Guo Jing. Impact of climate change on hydrological extreme events in upper reaches of the hanjiang river basin [J]. Journal of Beijing Normal University: Natural Science, 2010, 46(3): 383-386 [徐若兰, 陈华, 郭靖. 气候变化对汉江流域上游水文极值事件的影响 [J]. 北京师范大学学报: 自然科学版, 2010, 46(3): 383-386]

[12] Zhou Xudong, Yang Tao, Liang Huidi. Application of generalized extreme value distribution model to low water flow in headwater of Yellow River basin [J]. Water Resources and Power, 2013, 31(2): 12-14, 240 [周旭东, 杨涛, 梁慧迪. 广义极值分布模型在黄河源区枯季径流中的应用 [J]. 水电能源科学, 2013, 31(2): 12-14, 240]

[13] Liu Y, Huo X, Liu Y, et al. Analyzing monthly average streamflow extremes in the upper Ürümqi River based on a GPD model [J]. Environmental Earth Sciences, 2015 (in press) doi: 10.1007/s12665-015-4583-4

[14] Liu Y, Wu J, Liu Y, et al. Analyzing effects of climate change on streamflow in a glacier mountain catchment using an ARMA model [J]. Quaternary International, 2015, 358: 137-145. doi: 10.1016/j.quaint.2014.10.001

[15] Xu Z X, Zhao F F, Li J Y. Response of streamflow to climate change in the headwater catchment of the Yellow River basin [J]. Quaternary International, 2008, 208(1/2): 62-75

[16] Kumar R, Samaniego L, Attinger S. The effects of spatial discretization and model parameterization on the prediction of extreme runoff characteristics [J]. Journal of Hydrology, 2010, 392(1-2): 54-69

[17] Müller M, Kaspar M. Association between anomalies of moisture flux and extreme runoff events in the south-eastern Alps [J]. Natural Hazards and Earth System Sciences, 2011, 11(3): 915-920

[18] Lan Yongchao, Zhong Yingjun, Wu Sufen, et al. Sensitivity of mountain runoff of rivers originated from the south slope and the north slope of the Tianshan Mountain to climate change: Taking

- mountain runoff of Urumqi River and Kaidu River for example [J]. *Journal of Mountain Science*, 2009, 27(6): 712–718 [蓝永超, 钟英君, 吴素芬, 等. 天山南、北坡河流出山径流对气候变化的敏感性分析——以开都河与乌鲁木齐河出山径流为例 [J]. *山地学报*, 2009, 27(6): 712–718]
- [19] Paquet E, Garavaglia F, Garçon R, et al. The SCHADEX method: a semi-continuous rainfall-runoff simulation for extreme flood estimation [J]. *Journal of Hydrology*, 2013, 495: 23–37
- [20] Wang Yuefeng, Chen Xingwei, Chen Ying. Based on multiple time scales of SPI dry /wet conditions variation and identification of flood and drought events in Minjiang River of Fujian [J]. *Mountain Research*, 2014, 32(1): 52–57 [王跃峰, 陈兴伟, 陈莹. 基于多时间尺度 SPI 的闽江流域干湿变化与洪旱事件识别 [J]. *山地学报*, 2014, 32(1): 52–57]
- [21] Xia Jun, Liu Chunzhen, Ren Guoyu. Opportunity and Challenge of the Climate Change Impact on the Water Resource of China. Advances in the study of projection of climate change impacts on hydrological extremes [J]. *Advances in Earth Science*, 2011, 26(1): 1–12 [夏军, 刘春蓁, 任国玉. 气候变化对我国水资源影响研究面临的机遇与挑战 [J]. *地球科学进展*, 2011, 26(1): 1–12]
- [22] Coles S. An Introduction to Statistical Modeling of Extreme Values [M]. London: Springer. 2001.
- [23] 史道济. 实用极值统计方法 [M]. 天津: 天津科技出版社, 2006. [Shi Daoji. Practical Methods of Extreme Value Statistics. Tianjin: Tianjin Science and Technology Press 2006.]
- [24] Marshall L, Nott D, Sharma A. A comparative study of Markov chain Monte Carlo methods for conceptual rainfall – runoff modeling [J]. *Water Resources Research*, 2004, 40(2), W02501.
- [25] Haario H, Sakaman E, Tamminen J. An adaptive Metropolis algorithm [J]. *Bernoulli*, 2001, 7(2): 223–242
- [26] Bates BC, Campbell EP. A Markov chain Monte – Carlo scheme for parameter estimation and inference in conceptual rainfall – runoff modeling [J]. *Water Resources Research*, 2001, 37(4): 937–947
- [27] Blasone RS, Vrugt JA, Madsen H, et al. Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling [J]. *Advances in Water Resources*, 2008, 31(4): 630–648
- [28] 卫晓靖, 熊立华, 万民, 等. 融合马尔可夫链 – 蒙特卡洛算法的改进通用似然不确定性估计方法在流域水文模型中的应用 [J]. *水利学报*, 2009, 40(4): 464–480
- [29] Xia Jun, Zhai Jinliang, Zhan chesheng. Some reflections on the research and of development water resources in China [J]. *Advances in Earth Science*, 2011, 26(9): 905–915 [夏军, 翟金良, 占车生. 我国水资源研究与发展的若干思考 [J]. *地球科学进展*, 2011, 26(9): 905–915]
- [30] Liu Y, Metivier F, Gaillardet J, et al. Erosion rates deduced from seasonal mass balance along the upper Urumqi River in Tianshan [J]. *Solid Earth*, 2011, 2(2): 283–301
- [31] Liu Y, Metivier F, Lajeunesse E, et al. Measuring bed load in gravel bed mountain rivers: averaging methods and sampling strategies [J]. *Geodinamica Acta*, 2008, 21(1–2): 81–92
- [32] Cowles M K and Carlin BP. Markov Chain Monte Carlo convergence diagnostics: a comparative review [J]. *Journal of the American Statistical Association*, 1996, 91(434): 883–904

A Bayesian Analysis of Monthly Average Runoff Minima in Mountain Areas

LIU Youcun¹, HUO Xueli^{1,2}, HAO Yonghong¹, CUI Yuhuan³, HAN Tianding⁴,
SHEN Yongping⁴, WANG Jian⁴

(1. Tianjin Key Laboratory of Water Resources and Environment, Tianjin Normal University, Tianjin 300387, China;

2. College of Urban and Environmental Science, Tianjin Normal University, Tianjin 300387, China;

3. School of Science, Anhui Agriculture University Anhui Hefei 230036, China;

4. Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Gansu Lanzhou 730000, China)

Abstract: Global warming has intensified hydrological extreme events and resulted in disasters around the world. For disaster management and adaption of extreme events, it is essential to improve the accuracy of extreme value statistical models. In this study, Bayes' Theorem is introduced to estimate parameters in the Generalized Pareto Distribution (GPD) model which is applied to simulate the distribution of monthly average runoff minima during dry periods in mountain areas of Ürümqi River. Bayes' Theorem treats parameters as random variables and provides machinery way to convert the prior distribution of parameters into a posterior distribution. Statistical inferences based on posterior distribution can provide a more comprehensive representation of the parameters. An improved

Markov Chain Monte Carlo (MCMC) method, which can solve high-dimensional integral computation in the Bayes equation, is used to generate parameter simulations from the posterior distribution. Model diagnosis plots are made to guarantee the fitted GPD model is appropriate. Then based on the GPD model with Bayesian parameter estimates, monthly average minima corresponding to different return periods can be calculated. The results show that the improved MCMC method is able to make Markov chains converge at a high speed. Compared with the GPD model based on maximum likelihood parameter estimates, the GPD model based on Bayesian parameter estimates obtain more accurate estimations of minimum monthly average runoff. Moreover, the monthly average runoff minima in dry periods corresponding to 10 a, 25 a, 50 a and 100 a return periods are $0.60 \text{ m}^3/\text{s}$, $0.44 \text{ m}^3/\text{s}$, $0.32 \text{ m}^3/\text{s}$ and $0.20 \text{ m}^3/\text{s}$ respectively. The lower boundary of 95% confidence interval of 100a return level is $-0.238 \text{ m}^3/\text{s}$, which implies that Ürümqi River is likely to cease when 100 a return level occurs in dry periods.

Key words: runoff minima; GPD model; MCMC method; Ürümqi River

浙江天台山丹霞地貌景观

天台山位于浙江省天台县城北,是浙江省东部名山。山体呈东北-西南走向,西南连仙霞岭,东北遥接舟山群岛,为曹娥江与甬江的分水岭,平均海拔 500 m 以上,主峰华顶山在天台县东北,海拔 1 098 m,由花岗岩构成。天台山自然景观得天独厚,风光奇丽秀美,人文积淀深邃厚实,悠久灿烂。不仅是一座少有的佛道双栖的名山,也是天然的植物和动物园,奇草异木、珍禽异兽极多。天台山主要包括国清景区、赤城景区、佛陇景区等 13 个主要景区,面积 187.1 km^2 ,各景区天然成趣,别具一格,各擅其胜,美不胜收。其中赤城景区是一处典型的丹霞地貌,山体由水平的中生代红色砂岩、砾岩层叠而成。赤城山孤峰卓立,土色皆赤,有“不与众山同一色,敢于平地拔千仞”之誉。

(然子桐)