

考虑山洪灾害影响的贝叶斯径流模拟方法

梁冀雨, 刘曙光, 周正正, 钟桂辉, 方琦, 甄亿位

(同济大学 土木工程学院, 上海 200092)

摘要: 山洪灾害壅高河道水位带来的流量放大作用影响了贝叶斯径流模拟方法在山丘区流域的表现。针对这一问题, 以汶川地震后山洪灾害频发的都江堰白沙河流域为研究区域, 探究引发流域山洪灾害的临界雨量, 引入线性放大系数修正山洪灾害对流量的放大作用, 耦合一阶自回归残差模型与不同异方差转换函数构建了 6 种贝叶斯径流模拟方案, 利用 GR4J 模型模拟不同方案下流域出口日径流过程。研究结果表明, 临界雨量与线性放大系数能够有效识别、修正山洪灾害对径流的影响, 与未考虑山洪灾害影响的贝叶斯径流模拟方案相比, 白沙河流域径流模拟结果的准确性、可靠性、精确性指标分别平均提升 19%、32%、62%。该方法可在山洪灾害频发流域推广应用, 为山丘区流域径流预报提供更为准确的科学依据。

关键词: 山洪灾害; 贝叶斯理论; 异方差; 残差模型; 水文模拟

中图分类号: TV124

文献标志码: A

A Bayesian Runoff Simulation Method Considering the Influence of Flash Flood Disaster

LIANG Jiyu, LIU Shuguang, ZHOU Zhengzheng, ZHONG Guihui, FANG Qi, ZHEN Yiwei

(College of Civil Engineering, Tongji University, Shanghai 200092, China)

Abstract: The amplification of discharge caused by flash flood disasters has affected the performance of Bayesian runoff simulation in hilly watersheds. This paper, taking the Baisha River Basin in Dujiangyan with frequent flash flood disasters after the Wenchuan earthquake as the research area, explores the critical rainfall causing flash flood disaster in the basin, introduces the linear amplification factor to correct the amplification effect of flash flood disaster on flood flow, and constructs six

Bayesian flood simulation schemes by coupling the first-order autoregressive residual error model with different heteroscedastic transformation functions. The outlet daily runoff process of the basin is then simulated using the GR4J model. The results show that the critical rainfall and linear amplification factor can effectively identify and correct the impact of flash flood disasters on discharge. After considering the influence of flash flood disasters, the accuracy, reliability, and precision metrics of the daily runoff simulation results in the Baisha River Basin are increased by 19%, 32%, and 62%, respectively. This method can be widely used in the basin with frequent flash flood disasters and provide a more accurate scientific basis for runoff forecasting in hilly watersheds.

Key words: flash flood disaster; Bayesian theory; heteroscedasticity; residual error model; hydrological simulation

山洪灾害是山丘区由降雨、融雪等引起的洪水、泥石流河滑坡灾害的总称^[1], 其导致泥沙超量补给现象壅高河道水位产生流量高估现象^[2], 使流域出口断面实测径流大于响应流域降雨的清水流量, 破坏了流域水量平衡关系, 增加了山丘区流域径流模拟结果的输入不确定性。现有水文模型受限于其简化的物理过程、参数估算方法及资料观测条件, 在模拟复杂的山丘区小流域水文过程中也会产生来自模型输入、模型参数及模型结构的不确定性^[3-4]。在校准水文模型时, 上述所有来源的不确定性将集中传递给模型参数^[5-6], 降低了径流模拟结果的准确性并造成“异参同效”现象^[4]。

贝叶斯方法通过模拟径流的后验分布定量计算其不确定性, 提供了较确定性模拟方法更加合理的系统描述, 增加了径流模拟结果的可靠性^[7-8]。贝叶

收稿日期: 2021-05-10

基金项目: 国家重点研发计划(2018YFD1100401)

第一作者: 梁冀雨(1992—), 男, 博士生, 主要研究方向为水文水资源。E-mail: liangjiyu120@foxmail.com

通信作者: 刘曙光(1962—), 男, 教授, 博士生导师, 工学博士, 主要研究方向为水文水资源、流域与城市防洪减灾、洪水风险管理。E-mail: liusgliu@tongji.edu.cn



论文
拓展
介绍

斯方法通常假定观测径流可分解为由水文模型输出表示的确定性分量及代表残差的随机分量^[9]。其中,随机残差的统计模型确定了似然函数及径流预测分布的形式,直接影响着模型参数及贝叶斯径流预测分布估计的鲁棒性和准确性^[10]。模拟径流残差具有2个重要统计特征,即时间上的自相关性与异方差性(残差方差随流量增加)^[11]。大量的研究证明^[6, 11-14],湿润流域径流残差序列的自相关性可使用一阶自回归模型描述,而对于具有明显丰枯变化的流域,则可利用Copula函数建立更加灵活的残差自相关形式^[6, 11-14]。另一方面,对径流残差的异方差性的处理可分为2种思路:直接处理法和转换方法。直接处理法通过建立残差标准差与协变量(如流量)的函数关系,以此对残差进行标准化处理后,假定残差为零均值、同分布^[6, 14-17]。转换方法则是在对观测和模拟径流进行函数变换后,直接假定转换后的残差服从均值为零的常方差正态分布^[18]。

虽然贝叶斯径流模拟方法能够提升水文模型参数估计的鲁棒性并给出径流预测分布^[19],但是,目前的贝叶斯径流模拟方法皆没有考虑水文过程之外的不确定性来源^[20],特别是忽略了山丘区流域山洪灾害引起的流量放大作用,使其在山丘区流域应用时仍存在以下不足:①假定的残差分布无法描述山洪灾害导致的流量残差极大值,影响了似然函数的合理性;②参数后验估计为补偿山洪灾害导致的观测径流极大值而产生偏差^[21],降低了后验径流预测分布的整体可靠性;③频繁出现的山洪流量极大值增大了模拟残差的方差,令径流预测分布跨度过大、预测精度差。因此,在山洪频发的山丘区流域,贝叶斯径流模拟方法需要进一步优化和改进。

本文以汶川地震后山洪灾害频发的都江堰白沙河流域为研究区域,探究引发山洪灾害的临界雨量,引入线性放大系数修正山洪灾害对径流量的放大作用,耦合一阶自回归残差模型与不同异方差转换函数构建了6种贝叶斯径流模拟方案。选用适用于山区小流域的概念性水文模型GR4J^[22],模拟不同方案下流域出口日径流过程。通过比较不同模拟方案的表现,分析在山区贝叶斯径流模拟中考虑山洪灾害流量放大作用的必要性。

1 研究方法

1.1 模拟残差模型

令 $t(t=1, 2, \dots, n)$ 时刻山洪灾害流量放大作用

导致的附加洪水流量为 $Q_{F,t}$,则该时刻的观测流量 \tilde{Q}_t 可表示为

$$\tilde{Q}_t = Q_{F,t} + Q_{S,t} + \epsilon_t \quad (1)$$

式中: $Q_{S,t}$ 为 t 时刻模拟流量; ϵ_t 为 t 时刻的原始残差。若假定 $Q_{F,t}$ 是 $Q_{S,t}$ 的线性函数,式(1)可改写为

$$\epsilon_t = \tilde{Q}_t - \alpha Q_{S,t} \quad (2)$$

式中: α 为模拟流量线性放大系数, $\alpha \geq 1$ 。传统的自相关异方差残差模型不考虑山洪对流量的放大作用,可视为 $\alpha=1$ 。进一步假定山洪灾害触发条件为前期累积雨量超过临界雨量,有

$$\epsilon_t = \begin{cases} \tilde{Q}_t - \alpha Q_{S,t}, P_{C,t} > P_{C0} \\ \tilde{Q}_t - Q_{S,t}, P_{C,t} \leq P_{C0} \end{cases} \quad (3)$$

式中: $P_{C,t}$ 为 t 时刻的前期累积雨量; P_F 为相应的山洪临界雨量。为消除 ϵ_t 的异方差性,在计算残差前,选定转化函数 Z 将流量正态化,得到正态化残差 η_t ,即

$$\eta_t = \begin{cases} Z(\tilde{Q}_t) - Z(\alpha Q_{S,t}), P_{C,t} > P_F \\ Z(\tilde{Q}_t) - Z(Q_{S,t}), P_{C,t} \leq P_F \end{cases} \quad (4)$$

利用一阶自回归模型(AR1)描述 η_t 的自相关性,有

$$\eta_t = \phi_1 \eta_{t-1} + y_t \quad (5)$$

式中: ϕ_1 为一阶自回归系数; y_t 为正态化流量残差新息(innovation)。通常假定 y_t 服从均值为零的同方差正态分布^[14, 17, 23],于是

$$y_t \sim N(0, \sigma_y) \quad (6)$$

其中 σ_y 为新息 y_t 的标准差。式(1)至式(6)构建了考虑山洪影响的径流量残差模型,其基本假定为流量残差在去除异方差性、自相关性及山洪放大作用后服从均值为零的同方差正态分布。于是,残差模型参数 $\theta_\epsilon = \{\alpha, \phi_1, \theta_T\}$,其中 θ_T 为转换函数的参数。

1.2 似然函数

令水文模型参数为 θ_H ,根据贝叶斯原理,在给定输入资料序列 \tilde{X} 及观测流量序列 \tilde{Q} 时, θ_H 、 θ_ϵ 的后验分布可表示为

$$p(\theta_H, \theta_\epsilon | \tilde{X}, \tilde{Q}) \propto p(\tilde{Q} | \theta_H, \theta_\epsilon, \tilde{X}) p(\theta_H, \theta_\epsilon) \quad (7)$$

式中: $p(\theta_H, \theta_\epsilon)$ 为水文、残差模型参数的联合先验分布; $p(\tilde{Q} | \theta_H, \theta_\epsilon, \tilde{X})$ 为似然函数,可看作残差的联合分布,即 $p(\tilde{Q} | \theta_H, \theta_\epsilon, \tilde{X}) = p(\epsilon(\theta_H, \tilde{Q}, \tilde{X}) | \theta_\epsilon)$ 。若给定异方差转换函数 Z ,在式(1)至式(6)的假定下,可以得到似然函数

$$p(\tilde{\mathbf{Q}}|\theta_{\text{H}}, \theta_{\varepsilon}, \tilde{X}) = \prod_{i=1}^N Z'(\tilde{Q}_i; \theta_{\varepsilon}) f_N(y_i|0, \sigma_y^2) \quad (8)$$

式中: $Z'(\tilde{Q}_i) = dZ(\tilde{Q}_i)/d\tilde{Q}_i$; f_N 为正态分布的概率密度函数。本文采用的转换函数及其相应的似然函数

表达式见表1, 其中 $A_1, A_2, A_3, \lambda_1, \lambda_2$ 为转换函数参数。若 $\lambda_2 = 0$, 则BC(Box-Cox)转换变为LG(对数)转换。确定转换函数后, 似然函数 $p(\tilde{\mathbf{Q}}|\theta_{\text{H}}, \theta_{\varepsilon}, \tilde{X})$ 通过残差新息构建观测数据与模型、残差参数之间的联系, 并以此更新参数的先验分布。

表1 异方差转换函数与似然函数

Tab. 1 Transformation functions and corresponding likelihood functions

转换方法	转换函数	似然函数
lg-sinh转换	$Z_1(Q) = \frac{\lg(\sinh(\lambda_1 Q + A_1))}{\lambda_1}$	$l_1 = -\frac{N}{2} \lg(2\pi\sigma_y^2) - \frac{1}{2} \sum_{i=1}^N \frac{y_i^2}{\sigma_y^2} + \sum_{i=1}^N \lg(\coth(\lambda_1 y_i + A_1))$
Box-Cox转换	$Z_2(Q) = \frac{(Q + A_2)^{\lambda_2} - 1}{\lambda_2}, \lambda_2 \neq 0$	$l_2 = -\frac{N}{2} \lg(2\pi\sigma_y^2) - \frac{1}{2} \sum_{i=1}^N \frac{y_i^2}{\sigma_y^2} + (\lambda_2 - 1) \sum_{i=1}^N \lg(y_i + A_2)$
LG转换	$Z_3(Q) = \lg(Q + A_3)$	$l_3 = -\frac{N}{2} \lg(2\pi\sigma_y^2) - \frac{1}{2} \sum_{i=1}^N \frac{y_i^2}{\sigma_y^2} - \sum_{i=1}^N \lg(y_i + A_3)$

1.3 径流的预测分布

将式(8)代入式(7), 利用马尔科夫链蒙特卡罗(MCMC)抽样方法, 可以得到水文、残差模型参数的后验分布 $p(\theta_{\text{H}}, \theta_{\varepsilon}|\tilde{X}, \tilde{\mathbf{Q}})$ 。本文选择Vrugt^[24]提出的DREAM(Differential Evolution Adaptive Metropolis algorithm)算法作为抽样算法, 该方法采用自适应随机子空间采样技术(即自适应调整的交叉概率函数), 并能在抽样过程中去除无用的马尔科夫链, 提高了传统MCMC方法的抽样效率和解的精度, 在水文模型的校准方面得到了广泛的应用^[24-26]。基于概念水文模型及残差模型参数的后验分布, 可以生成模拟流量的预测分布 $\hat{\mathbf{Q}} = \{\hat{Q}_{r,i}; r = 1, 2, \dots, N_r, i = 1, 2, \dots, n\}$, 其中 N_r 为采样次数, n 为样本容量, 其具体步骤如下:

(1) 对 $t = 1, 2, \dots, n$, 从参数后验分布中抽取 $\theta_{\text{H},r}, \theta_{\varepsilon,r}$, 对给定的水文模型 f , 计算

$$Q_{\text{S},r,t} = \begin{cases} \alpha_r f(\theta_{\text{H},r}), P_{\text{C},t} > P_{\text{F}} \\ f(\theta_{\text{H},r}), P_{\text{C},t} \leq P_{\text{F}} \end{cases} \quad (9)$$

(2) 根据式(1)-(6), 计算新息标准差 $\sigma_{y,r}$ 。

(3) 生成 t 时刻的随机残差新息 $y_{r,t} \sim N(0, \sigma_{y,r})$ 。

(4) 根据式(5)计算 $\eta_{r,t}$ 。

(5) 根据式(4)对转换流量进行逆转换, 得到预测流量

$$\hat{Q}_{r,t} = Z^{-1}(Z(Q_{\text{S},r,t}) + \eta_{r,t}) \quad (10)$$

完成 N_r 次采样后, 集合 $\hat{\mathbf{Q}}$ 就是径流的预测分布, 其每一个现实 $\hat{\mathbf{Q}}_r = \{\hat{Q}_{r,i}; i = 1, 2, \dots, n\}$ 可以看

作径流过程的一个点估计。于是 t 时刻径流的后验中位数估计 $\hat{Q}_{m,t}$ 即为流量预测分布 $\hat{\mathbf{Q}}$ 的横截面 $\hat{\mathbf{Q}}_t = \{\hat{Q}_{r,t}; r = 1, 2, \dots, N_r\}$ 的中位数, 而 $\hat{\mathbf{Q}}_t$ 的其余分位数可作为相应置信水平下 t 时刻预测流量置信区间上下限。

2 研究区域与模拟方案设置

2.1 研究区域

以汶川地震后山洪灾害频发的都江堰市白沙河流域为研究区域, 其集水面积为 363 km^2 。该流域位于亚热带湿润季风气候区, 降雨充沛, 多年平均降雨量约 1134 mm , 降雨量年内分配不均匀性显著, 5至9月降雨量占全年的80%。同时, 降雨是该流域山洪灾害的主要诱发因素^[27]。白沙河流域内共有3个雨量站(杨柳坪站、大火地站和虹口站), 其中杨柳坪站也是位于流域出口的水文站(见图1)。选用白沙河流域2009—2018年的逐日降雨、蒸发及径流序列进行模拟。

2.2 水文模型

选用Perrin等^[22]建立的GR4J概念水文模型, 该模型适用于山区小流域的日尺度径流模拟^[28]。GR4J模型含有1个产流单元以及1个汇流存储单元, 并使用2个单位线方程来描述降雨与径流产生之间的时间间隔, 能够模拟流域从降水、蒸散发到产流、汇流的整体过程。模型仅含4个参数, 分别为产流单元最大容量 $S_1(\text{mm})$ 、地表水交换系数 $E_x(\text{mm} \cdot \text{d}^{-1})$ 、汇流单元最大容量 $S_2(\text{mm})$ 及单位线系数 $U(\text{d})$ 。

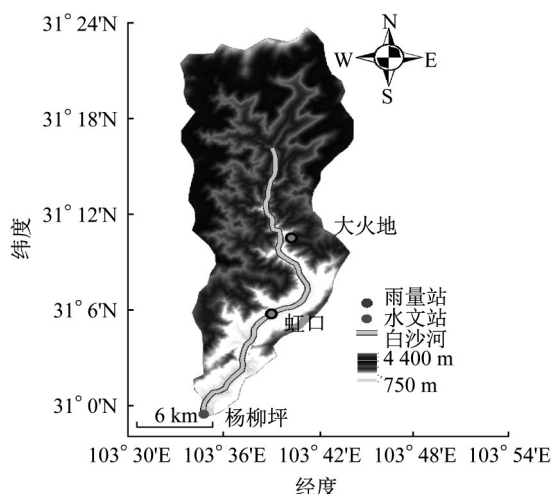


图1 白沙河流域概况

Fig. 1 Water system, gauging sites, and DEM in the Baisha River Basin

2.3 山洪临界雨量

根据白沙河流域内31处山洪沟的监测数据及流域山洪灾害调查报告,白沙河流域内各山洪沟2009—2018年间发生山洪事件天数共347d,包含361次山洪事件。白沙河流域内山洪沟主要分布于上游无人区,少量山洪沟发生洪水时仅会影响河道径流过程,而不会构成严重灾害。

杨柳坪站、大火地站和虹口站3站中,大火地站雨量序列存在缺测情况,为减少缺测资料影响,采用通过等权重的算术平均法估算流域平均逐日面雨量 $\bar{P}_t, t=1, 2, \dots, n$ 。以山洪当日的前 D_c 日(含当日)累积雨量为判别条件,计算日尺度山洪灾害判别准确率

$$A_c = \frac{\hat{F} \cap \tilde{F}}{\hat{F} \cup \tilde{F}} \times 100\% \quad (11)$$

式中: A_c 为山洪灾害判别准确率; \hat{F} 为判别山洪发生的日期集合; \tilde{F} 为实测山洪发生的日期集合。选取 $D_c=1, 2, \dots, 10d$,计算流域前期累积面雨量

$$P_{c,t} = \sum_{i=t-D_c+1}^t \bar{P}_i \quad (12)$$

以累积雨量集合 $P_c = \{P_{c,t}; t=1, 2, \dots, n\}$ 的百分位数作为临界雨量 P_F ,令 P_e 为相应的百分比。根据田述军等^[27]的研究,汶川地震后,白沙河流域泥石流临界雨量较震前有所减小,其中 D_c 取1时,临界雨量为51.2 mm,相应 P_e 值为98.6。针对这一现象,以0.1为步长,取 $P_e \in [90.0, 99.9]$,根据不同前期累积日数 D_c 分组,计算山洪灾害判别准确率 A_c ,绘制每组 A_c 样本的箱型图于图2。

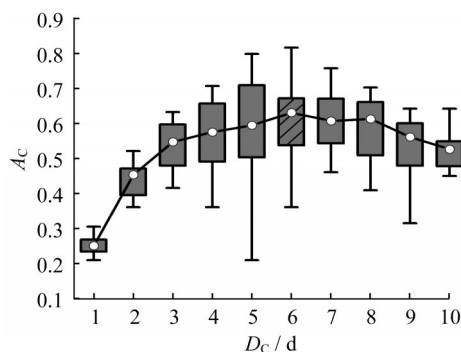


图2 不同累积日数下山洪判别准确率箱型图

Fig. 2 Box plot of accuracy of flash flood identification

由图2可知,山洪判别准确率 A_c 的均值随前期累积日数 D_c 的增加先增后减,其中当 D_c 取6d时, A_c 样本具有最高的中位数0.63。选取 A_c 均值最大的 $D_c=5, 6, 7, 8d$ 共4组,绘制 A_c 随 P_e 的分布于图3。如图所示,当 $D_c=6d, P_e=96.4$ 时,即山洪临界雨量取前6日流域平均雨量18.9mm、累积雨量113.9mm时, A_c 得到最大值0.81。

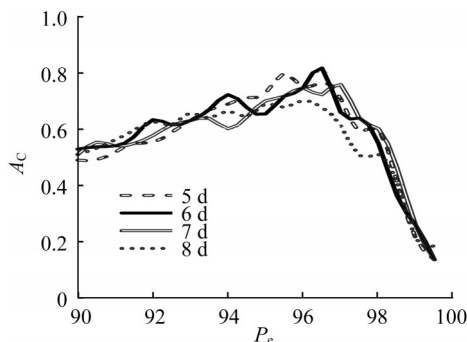


图3 山洪判别准确率随临界雨量百分比分布图

Fig. 3 Distribution of accuracy of flash flood identification with rainfall percentile

为进一步确定山洪灾害对径流的影响,根据图3中的先验信息,本文选择平均判别准确率最高的 $D_c=6d$ 作为计算前期累积雨量及临界雨量的日数,并将 P_e 作为残差模型参数,在校准模型时与其余参数一起估计。

2.4 径流模拟方案

选取山洪灾害发生频繁的2013年1月1日至2018年12月31日为模型校准期,以2009年1月1日至2012年12月31日为模型验证期,模型校准、验证期均以其第1年的资料对模型进行预热,调整模型的状态参数 S_1, S_2 。为探究本文提出的残差模型对贝叶斯径流模拟方法表现的影响,采用6种模拟方

案对白沙河流域杨柳坪站的流量序列进行模拟,模拟方案的设置见表2、表3。

表2 方案设置及残差模型参数范围

Tab. 2 Simulation schemes and corresponding parameter ranges

方案名称	转换函数	残差模型参数范围				
		ϕ_1	λ	A	α	P_c
LSH	lg-sinh	[0,1]	[0,1]	[0,1]		
LSH_F	lg-sinh	[0,1]	[0,1]	[0,1]	[1,30]	[90,99]
BC	Box-Cox	[0,1]	[-2,2]	[0,3]		
BC_F	Box-Cox	[0,1]	[-2,2]	[0,3]	[1,30]	[90,99]
LG	lg	[0,1]		[0,3]		
LG_F	lg	[0,1]		[0,3]	[1,30]	[90,99]

表3 GR4J模型参数

Tab. 3 Parameter ranges of the GR4J model

GR4J模型参数	单位	参数下限	参数上限
S_1	mm	100	1 200
E_X	mm·d ⁻¹	-30	30
S_2	mm	1	500
U	d	0.5	10

表2给出了模拟方案的名称,使用的异方差转换函数及残差模型参数范围,以后缀F表示方案采用考虑山洪影响的残差模型。表3给出了GR4J的模型参数取值范围。本文假定所有参数在其取值范围内服从均匀先验分布。同时,设定DREAM算法马尔可夫链数为各方案参数总个数的2倍,参数后验分布样本均取MCMC收敛后的2 000次抽样。

2.5 径流模拟结果评价指标

根据选定的径流模拟方案,为系统性比较不同异方差转换方法及残差模型的模拟表现,选取了不同的统计指标来描述各径流模拟方案的准确性、预测分布可靠性、预测分布精度及残差模型合理性。

2.5.1 模拟准确性

通过径流的后验中位数估计的纳什效率系数(Nash-Sutcliffe model efficiency coefficient)描述径流模拟的准确性,其表达式

$$N_m = 1 - \frac{\sum_{t=1}^n (\tilde{Q}_{m,t} - \hat{Q}_t)^2}{\sum_{t=1}^n (\tilde{Q}_{m,t} - \bar{Q})^2} \quad (13)$$

式中: \bar{Q} 为观测径流序列的平均值。 N_m 的取值范围为 $(-\infty, 1]$,越接近1则模拟准确度越高,反之模拟准确度越差。

2.5.2 径流预测分布可靠性

若径流预测分布能捕捉观测资料的分布特性,即径流观测值能看作从径流预测分布中抽取的样

本,则称该预测分布是可靠的。采用预测QQ(predictive quantile-quantile plot)图^[18, 29-30]可视化表征径流预测分布的可靠性,并采用Evin等^[17]提出的可靠性指标 π 对其定量计算,即

$$\pi = \frac{2}{n} \sum_{t=1}^n \left| F_U[F_{\hat{Q}(t)}(\tilde{Q}_t)] - F_n[F_{\hat{Q}(t)}(\tilde{Q}_t)] \right| \quad (14)$$

式中: F_U 为均匀分布 $U(0, 1)$ 的累积分布函数; $F_{\hat{Q}(t)}$ 为径流预测分布的经验累积分布函数; F_n 为集合 $\Omega = \{F_{\hat{Q}(t)}(\tilde{Q}_t), t = 1, 2, \dots, n\}$ 的经验累积分布函数。 π 的取值范围从0(完全可靠)到1(完全不可靠)。

2.5.3 径流预测分布精确性

径流预测分布精确性指1.3节中由径流预测分布分位数得到的置信区间宽度,可用精确性指标 P_c 计算。

$$P_c = \frac{\sum_{t=1}^n \text{std}(\hat{Q}_t)}{\sum_{t=1}^n \tilde{Q}_t} \quad (15)$$

式中: $\text{std}(\hat{Q}_t)$ 为径流预测分布在 t 时刻的横截面标准差。 $P_c \geq 0$, P_c 值越小,流量预测分布的精度越高、预测置信区间越窄。

3 径流模拟结果分析

3.1 不同方案的径流预测分布

根据表2、表3,利用贝叶斯方法对白沙河流域杨柳坪站的日径流资料进行模拟。不同方案去除模型预热期后的验证期(2010年1月1日至2012年12月31日)径流模拟结果如图4所示。

图4分别展示了考虑山洪灾害影响(图4a、4c、4e)与未考虑山洪影响(图4b、4d、4f)情况下模拟径流分布的95%置信区间与其后验中位数估计,同时,图4a与4b、图4c与4d、图4e与4f分别代表同一种异方差转换函数。

比较图4b、4d、4f可知,在未考虑山洪影响时,3种异方差转换函数中,BC方案低流量模拟的95%置信区间最窄,然而在模拟高流量时出现了较大的不确定性。图5给出了BC、BC_F方案下残差模型参数 λ_2 的后验核密度分布,如图所示,BC方案 λ_2 的后验众数为-1.24,且其后验样本均小于零。图6给出了残差模型后验众数下($\lambda_2 = -1.24$, $A_2 = 0.009$, $A_1 = 0.017$)BC与LG方案的转换流量分布图。如图所示,相比于LG转换,BC转换大幅减

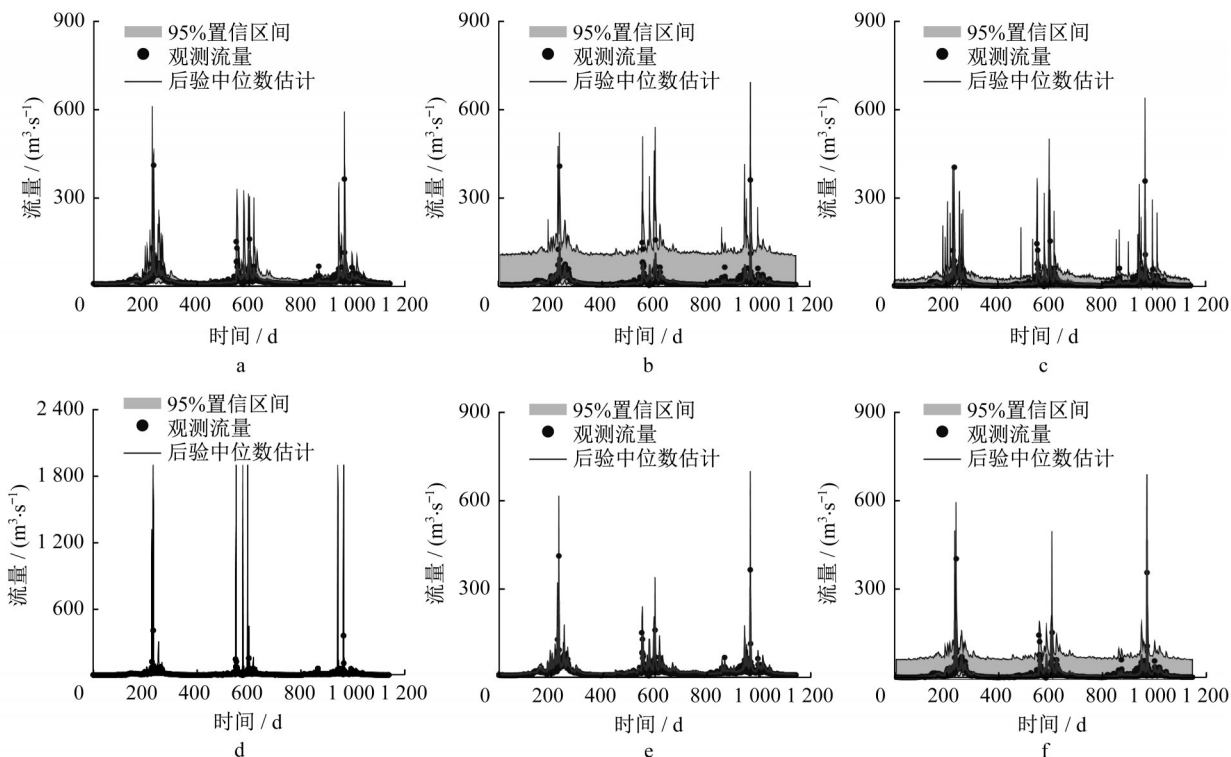


图4 杨柳坪站验证期不同方案径流模拟结果

Fig. 4 Simulation performance of different schemes in validation period

小了高低流量差,这弥补了山洪灾害的流量放大作用,减小了残差的整体方差,使BC方案的整体精度在未考虑山洪影响的径流模拟方案中最优。然而,BC方案按式(10)利用逆转换生成流量预测分布时,有预测流量

$$\hat{Q}_t = \left[\lambda_2 \eta_t + (Q_{s,t} + A_2)^{\lambda_2} \right]^{1/\lambda_2} - A_2 \quad (16)$$

由于 A_2 的后验估计趋于零,当 $\lambda_2 < 0$ 且 $Q_{s,t}$ 较大时, $\lambda_2 \eta_t + (Q_{s,t} + A_2)^{\lambda_2}$ 可能接近零,此时 $\hat{Q}_t \rightarrow 0^{1/\lambda_2}$,导致流量预测分布频繁采样到不合理的大值,因此出现图4d中的高流量区的高尖峰,降低了高流量区径流预测分布的可靠性和精确性。LSH及LG方案的逆转换函数较为稳定^[31],不易产生不合理的高流量预测,但也无法弥补山洪灾害导致的残差方差增大,导致其低流量区域的置信区间较宽。

对比图4a、4c、4e与图4b、4d、4f,易见考虑山洪灾害对流量的放大作用后,模拟径流的95%置信区间宽度皆明显减小。其中,BC_F方案参数 λ_2 的后验众数为0.12(图5),仅19%的 λ_2 后验样本小于零,显著减小了流量预测分布采样到异常值的概率,增加了BC逆转换下高流量径流预测分布的可靠性和精确性。

除转换函数本身的性质外,对应于残差标准差

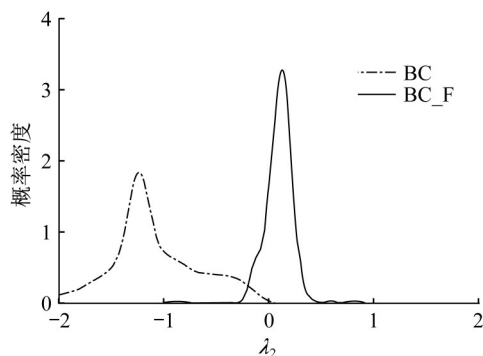
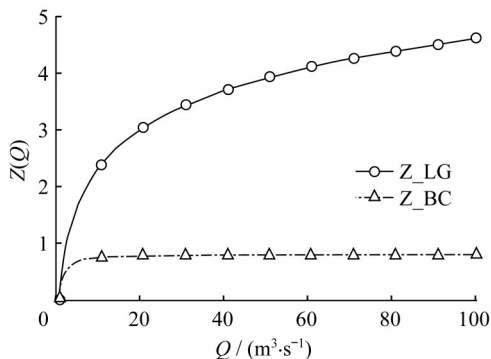
图5 λ_2 后验核密度分布Fig. 5 Posterior kernel density of λ_2 

图6 LG、BC方案转换流量对比

Fig. 6 Transformed runoff of LG and BC scheme

的残差新息标准差 σ_y 对流量预测分布的精确性也有着直接影响,较大的 σ_y 值对应较宽的径流预测分布及较大的预测不确定性。表4给出了6种方案下 σ_y 的后验样本的5%分位数、中位数及95%分位数。根据表4,采用考虑山洪影响的残差模型后,LSH、BC、LG这3种方案的 σ_y 中位数分别减小84%、25%、23%, σ_y 的90%置信区间宽度分别减小15%、90%、17%。

表4 不同方案后验 σ_y 分位数表

Tab. 4 Quantiles of posterior distribution of σ_y

方案名称	5%分位数	中位数	95%分位数
LSH	8.22×10^2	8.35×10^2	8.55×10^2
LSH_F	1.18×10^2	1.30×10^2	1.46×10^2
BC	9.15×10^{-3}	9.88×10^{-3}	1.07×10^{-2}
BC_F	7.30×10^{-3}	7.38×10^{-3}	7.46×10^{-3}
LG	1.42×10^{-1}	1.43×10^{-1}	1.45×10^{-1}
LG_F	1.16×10^{-1}	1.17×10^{-1}	1.18×10^{-1}

表5列举了6种模拟方案的准确性指标 N_m 、可靠性指标 π 及精确性指标 P_c 。由表5可知,与前述一致,未考虑山洪影响时,BC方案在径流预测分布精确性指标 P_c 及准确性指标 N_m 上表现最优。综合考虑全部6种方案后,模拟准确性最优方案为LSH_F,而预测分布可靠性及精确性的最优方案皆为LG_F。

表5 不同方案模拟结果评价指标

Tab. 5 Performance metrics of different schemes

方案名称	N_m	π	P_c
LSH	0.61	0.37	0.89
LSH_F	0.80	0.23	0.34
BC	0.68	0.44	0.71
BC_F	0.74	0.37	0.35
LG	0.63	0.36	0.80
LG_F	0.76	0.19	0.22

值得注意的是,除最优方案外,亦可发现在采用考虑山洪影响的残差模型后,径流模拟结果评价指标都有所改进,其中, N_m 平均提升19%, π 平均降低32%, P_c 平均降低62%,而对于不同的异方差转换函数,LG方案3项评价指标平均提升45%,LSH及BC方案分别平均提升43%和26%。

3.2 残差新息的后验分布

以LG方案为例,分别绘制LG、LG_F方案标准化残差新息 v 的直方图于图7、图8。如图7所示,与标准正态分布相比,LG方案的标准化残差新息分布呈较强的正偏态,右侧有长尾并有较大的峰度。相较之下,图8中LG_F方案的标准化残差新息与标准正态分布更为一致。为进一步比较,将LG和LG_F

方案下 v 的均值、偏态系数及峰度系数列于表6。

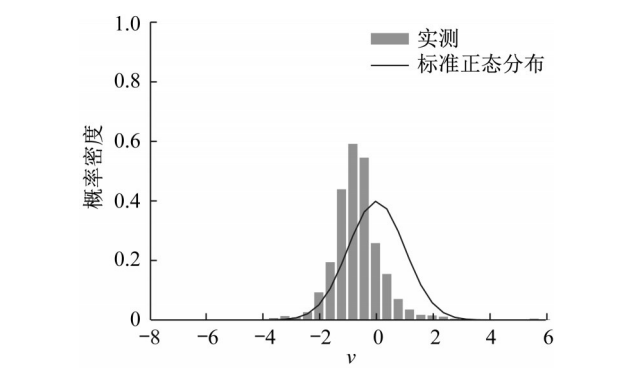


图7 LG方案标准化残差新息直方图

Fig. 7 Histogram of standardized innovation of LG scheme

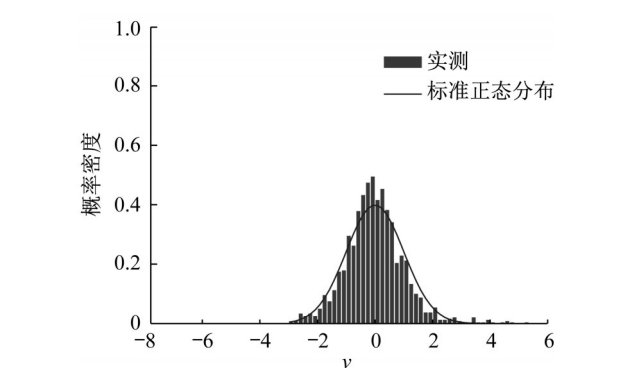


图8 LG_F方案标准化残差新息直方图

Fig. 8 Histogram of standardized innovation of LG_F scheme

表6 考虑山洪灾害影响前后标准化残差新息 v 统计参数

Tab. 6 Statistics of standardized innovation of LG and LG_F schemes

方案名称	均值	偏态系数	峰度系数
LG	-0.342	0.460	15.533
LG_F	-0.017	0.116	5.772
标准正态分布	0	0	3

由表6可知,相较于LG方案,LG_F方案下 v 的统计参数更接近残差模型所假定的标准正态分布的统计参数。由于2种方案中 v 的标准差均为1,因此具有更大峰度系数的LG方案中 v 包含更多的特大值,根据式(3)-(6)可知,某些时间点的模拟径流远小于观测流量。分别记LG方案 v 中较大的前5%、3%为 v_5 及 v_3 。 v_5 与山洪灾害发生时间重合的天数占 v_5 总数的65%, v_3 相应重合比例为83%,表明 v 出现特大值与山洪灾害发生具有较强的相关性。而这些特大值也是LG方案下 v 的分布出现较强正偏性的主要原因。同时,MCMC参数估计方法过多关注这些特大值^[32],导致参数的后验估计使模型模拟径

流 Q_s 整体偏大,造成低流量区域 Q_s 频繁大于观测流量 \tilde{Q} ,使LG方案下 v 的均值小于零。

另一方面,LG_F方案通过流量线性放大系数 α 放大山洪灾害发生时的模拟径流 Q_s ,减小了这些时间点上 Q_s 与观测流量 \tilde{Q} 的偏差;LG_F方案下 v_5 、 v_3 的均值较LG方案分别降低39%、57%。 v 的特大值的下降进一步提升了MCMC参数估计的准确性,降低的低流量区域负残差的出现概率使 v 的均值由-0.342增至-0.017。综上,考虑山洪影响的残差模型通过识别并修正山洪灾害导致的转换残差特大值,使其更接近假定的同方差正态分布,LSH与BC方案下得到的结论与LG方案相似。

3.3 山洪修正参数的后验分析

相较于未考虑山洪灾害影响的自相关异方差残差模型,本文通过引入参数 P_e 、 α 识别并修正山洪灾害对流量的放大作用。图9、图10给出了3种考虑山洪影响的残差模型参数 P_e 、 α 的后验核密度分布。

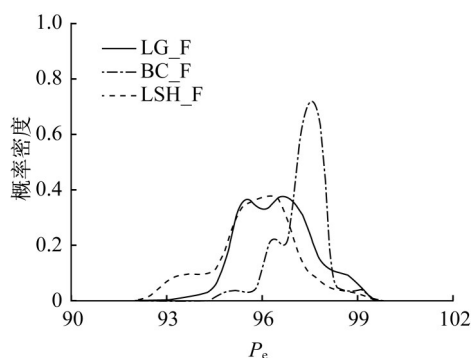


图9 各方案 P_e 后验核密度分布

Fig. 9 Posterior kernel density of P_e

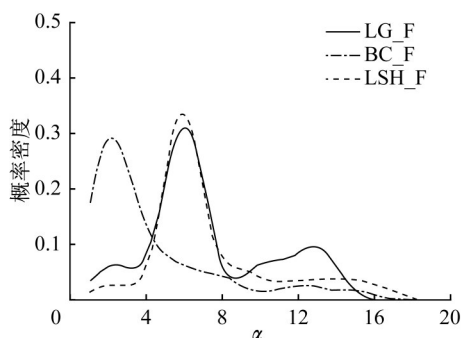


图10 LG_F方案 α 后验核密度分布

Fig. 10 Posterior kernel density of α

根据图9,LG_F方案 P_e 的后验分布集中于95至98之间,其最大后验估计为96.5,对应的前6日临界累积雨量为115.1mm。值得注意的是,该雨量并

非传统意义上的山洪临界雨量,其含义是,当前期雨量超过该值时,流域响应径流发生明显增大,需要进行修正。因此,该雨量应称为“流量修正临界雨量”。同时,根据图10,LG_F方案 α 的后验众数为6.0,在式(3)的假定下,即判断山洪发生时,径流量将被放大6倍,这一结果与胡凯衡等^[33]得到的非线性流量修正系数结果相近。

LSH_F方案 P_e 、 α 的后验分布与LG_F方案相似,而BC_F方案的 P_e 后验众数为97.7,判别验证山洪灾害发生的日数较LG_F、LSH_F方案减少约14d,山洪灾害识别准确率 A_c 为0.59。同时,BC_F方案 α 的后验众数为2.8明显小于LG_F、LSH_F方案。造成这一差异的一个可能原因是BC转换函数 Z_2 (表1)在 λ_2 较小的情况下,减小了高流量观测值与模拟值,因此平缓了山洪灾害导致的较大模拟残差与普通残差的差别,使残差模型无法完全识别、修正某些场次的模拟残差。同时,这种残差平缓效应也使得参数 α 的后验估计偏小。

4 结论

山洪灾害壅高河道水位带来的流量放大作用影响了贝叶斯径流模拟方法在山丘区流域的表现。以汶川地震后山洪灾害频发的都江堰白沙河流域为研究区域,基于引发流域山洪灾害的临界雨量,引入线性放大系数修正山洪灾害对流量的放大作用,耦合一阶自回归残差模型与不同异方差转换函数构建了6种贝叶斯径流模拟方案,利用GR4J模型模拟不同方案下流域出口日径流过程,得到如下主要结论:识别并修正山洪灾害的流量放大作用后,①山丘区流域径流残差新息更接近残差模型假定的同方差正态分布;②贝叶斯径流模拟表现明显提升,其中准确性、可靠性及精确性指标平均提升19%、32%、62%;③不同异方差转换函数中,LG转换得到了最优的径流预测分布可靠性及精确性,log-sinh转换得到了径流过程的最优后验中位数估计,而BC转换受限于其转换函数特性,对山洪灾害的识别及修正效果较差,径流模拟表现的提升有限。

提出的考虑山洪影响的贝叶斯径流模拟方法理论基础扎实,修正了山洪灾害对山丘区流域径流量的放大作用,优化了径流模拟结果,可在汶川地震后山洪频发的山丘区流域推广应用,为径流预报提供更为准确的科学依据。同时,为进一步提升山丘区贝叶斯径流模拟的表现,除临界雨量、线性放大系数

外,更加高效、准确的山洪灾害识别及流量修正方法值得继续深入研究。

作者贡献声明:

梁冀雨:数据收集整理、模型脚本编写、模拟结果分析及论文撰写。

刘曙光:提供研究思路及设计模拟实验方案。

周正正:数据校核及论文修订

钟桂辉:数据校核及论文修订

方琦:数据校核及论文修订

甄亿位:数据校核及论文修订。

参考文献:

- [1] 刘志雨. 山洪预警预报技术研究与应用[J]. 中国防汛抗旱, 2012, 22(2):41.
LIU Zhiyu. Study and application of flash flood forecast and warning[J]. China Flood and Drought Management, 2012, 22(2): 41.
- [2] 赵刚, 庞博, 徐宗学, 等. 中国山洪灾害危险性评价[J]. 水利学报, 2016, 47(9):1133.
ZHAO Gang, PANG Bo, XU Zongxue, *et al.* Assessment on the hazard of flash flood disasters in China [J]. Journal of Hydraulic Engineering, 2016, 47(9):1133.
- [3] 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.
LIANG Zhongmin, DAI Rong, LI Binqun. Research progress of hydrological uncertainty analysis based on Bayesian theory [J]. Advances of Water Science, 2010, 21(2):274.
- [4] BEVEN K, FREER J. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology [J]. Journal of Hydrology, 2001, 249(1):11.
- [5] YANG J, REICHERT P, ABBASPOUR K C. Bayesian uncertainty analysis in distributed hydrologic modeling: A case study in the Thur River basin (Switzerland) [J]. Water Resources Research, 2007, 43(10):W10401.
- [6] SCHOUPS G, VRUGT J A. A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors[J]. Water Resources Research, 2010, 46(10):W10531.
- [7] HAN S, COULIBALY P. Bayesian flood forecasting methods: A review[J]. Journal of Hydrology, 2017, 551:340.
- [8] 刘章君, 郭生练, 李天元, 等. 贝叶斯概率洪水预报模型及其比较应用研究[J]. 水利学报, 2014, 45(9):1019.
LIU Zhangjun, GUO Shenglian, LI Tianyuan, *et al.* Comparative study of Bayesian probabilistic flood forecasting models[J]. Journal of Hydraulic Engineering, 2014, 45(9): 1019.
- [9] SCHOUPS G, VAN DE GIESEN N C, SAVENIJE H H G. Model complexity control for hydrologic prediction[J]. Water Resources Research, 2008, 44(12): W00B03.
- [10] ZHANG Z, WAGENER T, REED P, *et al.* Reducing uncertainty in predictions in ungauged basins by combining hydrologic indices regionalization and multiobjective optimization [J]. Water Resources Research, 2008, 44(12): W00B04.
- [11] SOROOSHIAN S, DRACUP J A. Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: Correlated and heteroscedastic error cases [J]. Water Resources Research, 1980, 16(2):430.
- [12] HUNTER J, THYER M, MCINERNEY D, *et al.* Achieving high-quality probabilistic predictions from hydrological models calibrated with a wide range of objective functions [J]. Journal of Hydrology, 2021, 603:126578.
- [13] EVIN G, KAVETSKI D, THYER M, *et al.* Pitfalls and improvements in the joint inference of heteroscedasticity and autocorrelation in hydrological model calibration [J]. Water Resources Research, 2013, 49:4518.
- [14] MCINERNEY D, THYER M, KAVETSKI D, *et al.* Improving probabilistic prediction of daily streamflow by identifying Pareto optimal approaches for modeling heteroscedastic residual errors [J]. Water Resources Research, 2017, 53(3):2199.
- [15] LIN K, CHEN X, ZHANG Q, *et al.* A modified generalized likelihood uncertainty estimation method by using Copula function [C]// Proceedings of the 2008 Symposium on Hydrological Modelling and Integrated Water Resources Management in Ungauged Mountainous Watersheds. Chengdu: IAHS Publication, 2009: 51-56.
- [16] WANI O, SCHEIDEGGER A, Cecinati F, *et al.* Exploring a copula-based alternative to additive error models—for non-negative and autocorrelated time series in hydrology [J]. Journal of Hydrology, 2019, 575:1031.
- [17] EVIN G, THYER M, KAVETSKI D, *et al.* Comparison of joint versus postprocessor approaches for hydrological uncertainty estimation accounting for error autocorrelation and heteroscedasticity [J]. Water Resources Research, 2014, 50(3):2350.
- [18] MCINERNEY D, THYER M, KAVETSKI D, *et al.* Improving probabilistic prediction of daily streamflow by identifying Pareto optimal approaches for modeling heteroscedastic residual errors [J]. Water Resources Research, 2017, 53(3):2199.
- [19] SUN R, YUAN H, LIU X. Effect of heteroscedasticity treatment in residual error models on model calibration and prediction uncertainty estimation [J]. Journal of Hydrology, 2017, 554:680.
- [20] 崔鹏. 中国山地灾害研究进展与未来应关注的科学问题[J]. 地理科学进展, 2014, 33(2):145.
CUI Peng. Progress and prospects in research on mountain hazards in China [J]. Progress in Geography, 2014, 33(2):145.
- [21] KRAUSE P, BOYLE D P, BÄSE F. Comparison of different

- efficiency criteria for hydrological model assessment [J]. *Advances in Geosciences*, 2005, 5:89.
- [22] PERRIN C, MICHEL C, ANDRÉASSIAN V. Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments[J]. *Journal of Hydrology*, 2001, 242:275.
- [23] YANG J, REICHERT P, ABBASPOUR K C, *et al.* Hydrological modelling of the Chaohe Basin in China: Statistical model formulation and Bayesian inference [J]. *Journal of Hydrology*, 2007, 340(3):28:167.
- [24] VRUGT J A. Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation [J]. *Environmental Modelling & Software*, 2016, 75:273.
- [25] SADEGH M, MAJD M S, HERNANDEZ J, *et al.* The quest for hydrological signatures: Effects of data transformation on Bayesian inference of watershed models [J]. *Water Resources Management*, 2018, 32(5):1867.
- [26] SADEGH M, VRUGT J A, XU C, *et al.* The stationarity paradigm revisited: Hypothesis testing using diagnostics, summary metrics, and DREAM(ABC) [J]. *Water Resources Research*, 2016, 51(11):9207.
- [27] 田述军, 樊晓一. 基于临界雨量的震后泥石流预测预报研究——以白沙河流域为例[J]. *自然灾害学报*, 2015, 24(4):176.
- TIAN Shujun, FAN Xiaoyi. Forecast of post-earthquake debris flow based on critical rainfall: Taking the Basin as an example[J]. *Journal of Natural Disasters*, 2015, 24(4):176.
- [28] PERRIN C, MICHEL C, ANDRÉASSIAN V. Improvement of a parsimonious model for streamflow simulation [J]. *Journal of Hydrology*, 2003, 279(1):275.
- [29] EVIN G, THYER M, KAVETSKI D, *et al.* Comparison of joint versus postprocessor approaches for hydrological uncertainty estimation accounting for error autocorrelation and heteroscedasticity [J]. *Water Resources Research*, 2014, 50(3):2350.
- [30] THYER M, RENARD B, KAVETSKI D, *et al.* Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis.[J]. *Water Resources Research*, 2009, 45(12): 1211.
- [31] WANG Q J, SHRESTHA D L, ROBERTSON D E, *et al.* A log-sinh transformation for data normalization and variance stabilization [J]. *Water Resources Research*, 2012, 48(5): W05514.
- [32] KRAUSE P, BOYLE D, BÄSE F. Comparison of different efficiency criteria for hydrologic models [J]. *Advances in Geosciences*, 2005, 5:89.
- [33] 胡凯衡, 崔鹏, 游勇, 等. 汶川灾区泥石流峰值流量的非线性雨洪修正法 [J]. *四川大学学报(工程科学版)*, 2010, 42(5):52.
- HU Kaiheng, CUI Peng, YOU Yong, *et al.* Nonlinear modified flood method for calculating the debris flow peak discharge in Wenchuan earthquake region [J]. *Journal of Sichuan University (Engineering Science Edition)*, 2010, 42(5):52.