

DOI: 10.14042/j.cnki.32.1309.2023.05.002

无资料地区水文模型参数移植不确定性分析

关铁生^{1,2}, 鲍振鑫^{1,2,3}, 贺瑞敏^{1,2}, 杨艳青⁴, 吴厚发^{1,2}

(1. 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; 2. 水利部应对气候变化研究中心, 江苏 南京 210029; 3. 长江保护与绿色发展研究院, 江苏 南京 210098; 4. 中国科学院水利部成都山地灾害与环境研究所, 四川 成都 610299)

摘要: 无资料地区降水径流模拟是水文学研究的国际前沿和热点问题。水文模型参数移植是无资料地区降水径流模拟的重要方法, 对径流模拟精度具有重要的影响。利用核密度估计和蒙特卡罗随机模拟等方法, 构建了一种水文模型参数移植误差驱动的无资料地区径流模拟不确定性定量评估框架。以广西壮族自治区42个有水文监测站点的典型中小河流为研究对象, 率定新安江模型参数并模拟日径流和洪水过程, 将42个典型流域依次假定为无资料流域, 利用基于回归分析、相似流域和机器学习的参数移植方法, 模拟无资料地区的洪水过程并识别最优的参数移植方法, 分析移植法估算的模型参数值和直接率定值相比误差的概率分布特征, 定量评估模型参数移植误差带来的径流模拟不确定性。研究结果表明: ① 基于回归分析的参数移植法模拟无资料地区洪水过程的精度优于相似流域法, 优选的机器学习算法比传统回归分析法和相似流域法的计算精度提高了7%~15%; ② 与模型参数率定值相比, 移植方法计算的模型参数具有一定的误差, 对洪水模拟敏感参数的误差小于不敏感参数; ③ 受模型参数移植误差的影响, 利用蒙特卡罗法随机模拟的洪水过程具有一定的不确定性, 洪量和洪峰相对误差的主要区间分别为10%~30%和10%~40%。相关成果为无资料地区的径流概率模拟及不确定性评估提供了一种新的技术, 对中小河流洪水预报与防洪减灾具有一定的支撑作用。

关键词: 无资料地区; 径流模拟; 不确定性; 新安江模型; 参数移植

中图分类号: TV11 文献标志码: A 文章编号: 1001-6791(2023)05-0660-13

水文模型是对水循环过程的数学描述, 是模拟流域降水径流过程最重要的工具之一, 包括黑箱的统计模型、集总式的概念性模型、分布式的物理模型等多种类型^[1]。模型参数的准确估计对水文模拟精度具有重要的影响, 一般利用流域出口断面的实测流量资料率定求解^[2-3]。然而世界上绝大多数河流特别是中小河流没有监测站点, 缺少水文观测资料。根据第一次全国水利普查成果, 中国流域面积100 km²及以上的河流有22 909条, 其中有水文站和水位站的河流仅1 778条, 即92.2%的河流无水文监测站点^[4]。此外, 近60 a受气候变化和人类活动等因素的影响, 水文资料序列的一致性发生改变, 历史水文观测资料不能很好地反映当前的产汇流关系, 带来了新的资料短缺问题^[5-6]。如何科学估计无资料地区的模型参数, 准确模拟降水径流过程是水文模拟研究领域面临的一项国际前沿和难点问题^[7-9]。

水文学家针对无资料地区的降水径流模拟难题, 开展了大量的研究工作并取得了丰硕的成果, 逐渐认识到模型参数区域化, 即将有资料地区率定的模型参数移植到无资料地区, 是解决这一问题最有效的方法之一^[10-11]。模型参数移植主要包括空间插值、回归分析、相似流域等3种方法^[12-13]。早期由于缺乏流域下垫面特征等相关数据, 基于空间上距离越近的流域具有越相似的水文特征这一基本假设, 根据流域所处的地理位置, 利用最近邻、反距离权重、克里金、最小曲率等空间插值技术将有资料地区率定的模型参数插值计算到无资料地区^[14]。这种方法只利用了流域的空间位置, 未考虑影响水文过程的气候和下垫面等流域特征信

收稿日期: 2022-07-01; 网络出版日期: 2023-09-21

网络出版地址: <https://link.cnki.net/urlid/32.1309.P.20230920.2233.002>

资助项目: 国家重点研发计划资助项目(2022YFC3205200); 国家自然科学基金资助项目(41961124007)

作者简介: 关铁生(1974—), 男, 内蒙古包头人, 正高级工程师, 主要从事防汛抗旱等方面研究。

E-mail: tsguan@nhri.cn

通信作者: 鲍振鑫, E-mail: zxbao@nhri.cn

息,具有一定的误差。随着遥感、航拍等观测技术的不断进步,全球气候、植被、土壤、地形等对流域产汇流过程有重要影响的相关要素数据集逐步完善,为更精细的基于回归分析和相似流域的参数移植方法提供了资料基础^[15-16]。回归分析的核心思想是在有资料的流域利用回归方程等统计方法建立率定的模型参数和流域特征之间的某种定量关系,然后利用构建的统计模型和无资料地区的流域特征来计算模型参数^[17-19]。相似流域法则是先根据影响水文过程的相关物理特征判别流域的相似程度,再将其作为有资料流域的权重来计算无资料地区的模型参数^[20-22]。为了挑选最优的模型参数移植方法,水文学家在全球很多流域作了对比研究,但是未获得一致的结论,一些研究认为相似流域法优于回归分析法,然而另一些研究得到回归分析法优于相似流域法的结果^[23-24]。由于模型参数与流域特征之间的复杂非线性关系、模型参数的异参同效性、模型参数对输入资料的敏感性等多种原因,移植方法计算的模型参数与率定值相比有一定的误差,其径流模拟精度低于基于实测资料率定的结果,具有较大的不确定性^[25-28]。目前在有资料地区的水文模拟中开展了大量的不确定性分析工作,取得了丰硕的成果,但是关于无资料地区水文模拟不确定性的研究较为薄弱^[29-33]。与有资料的流域相比,无资料地区水文模拟最大的不确定性来源于模型参数移植结果的不确定性^[34-36]。如何定量评估模型参数移植过程带来的不确定性是无资料地区水文模拟研究的薄弱环节,也是本论文的重点研究内容。近期,机器学习等人工智能算法快速发展,并成功应用于水文模拟研究,如何利用机器学习算法开展水文模型参数移植,提高无资料地区水文模拟精度,并降低其不确定性是本文拟研究的另一个重要内容。

选取广西壮族自治区42个有水文资料的典型中小河流为研究对象,利用新安江水文模型率定模型参数并模拟降水径流过程,采用留一法将42个典型流域依次假定为无资料的目标流域,采用回归分析、相似流域、机器学习等算法根据其他41个流域率定的模型参数计算目标流域的参数值,分析模型参数移植结果与直接率定值的误差特征,评估水文模型参数移植引起的径流模拟不确定性,为区域洪水预报和水资源评价等工作提供技术支撑。

1 研究区与资料

1.1 研究区概况

广西壮族自治区位于 $20^{\circ}54'N-26^{\circ}24'N$ 、 $104^{\circ}28'E-112^{\circ}04'E$ 之间,地处云贵高原东南边缘,两广丘陵西部,地势西北高、东南低,四周多为山地、高原,地貌以山地丘陵性盆地为主,陆地面积为 23.76 万 km^2 (图1)。广西纬度较低,属亚热带季风气候区,气候温暖,雨水丰沛,光照充足,年平均气温和降水量由东南向西北逐渐减少。区域内河流多随地势由西北流向东南,包括珠江、长江、桂南独流入海、百都河等四大水系,集水面积 1000 km^2 以上的地表河有69条, 100 km^2 以上的地表河有678条,河流总长约 3.4 万 km ,形成了以红水河-西江为主干流的横贯中部以及两侧支流的树枝状分布特征。根据产汇流特征及水文地质条件,可将广西分为10个水文分区(图1(b)),其中,东南部的3、8、9、10为非喀斯特地区,其余地区为喀斯特地区。

1.2 资料情况

选取广西壮族自治区内42个有水文监测站点的典型中小河流为研究对象,流域面积为 $23.3 \sim 2762$ km^2 ,平均为 860 km^2 ,典型流域内共有439个雨量站点(图1)。选取的水文站点分布较为均匀,涵盖了山区、丘陵以及盆地等多种地势地形,在广西境内具有较好的代表性。利用的2005—2014年实测径流、降水和蒸发皿蒸发资料来自于广西壮族自治区水文中心整编的水文年鉴,包括日过程和洪水过程2种时段。每个流域根据资料情况选取15场左右的洪水过程。DEM数据来自于日本METI和美国NASA联合发布的ASTER GDEM全球数字高程数据,空间分辨率为 30 $m \times 30$ m ,垂直分辨率为 1 m 。植被叶面积指数数据来源于美国EOS/MODIS数据中心提供的MODIS-Terra/Aqua合成产品。

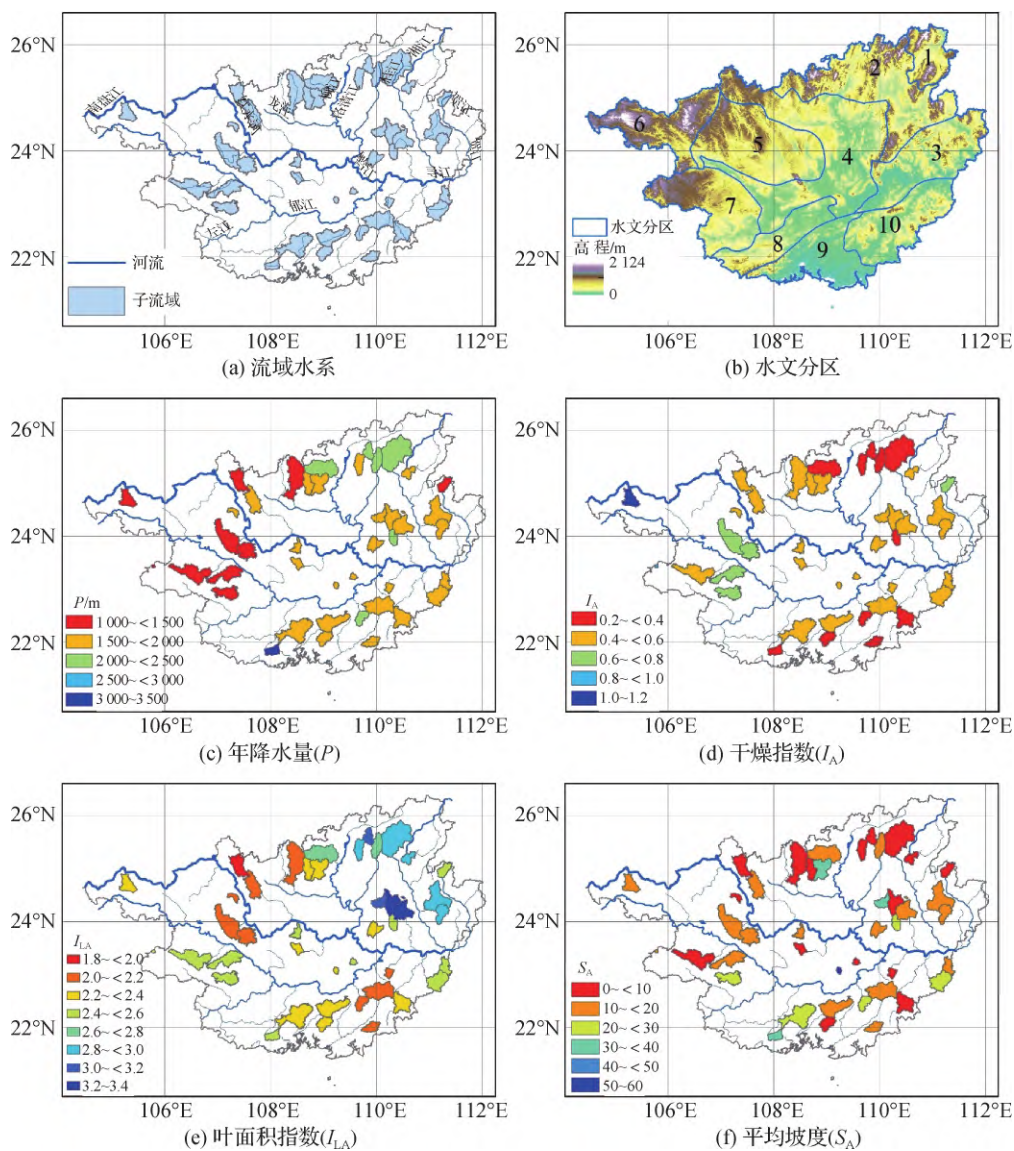


图1 研究区概况

Fig. 1 Basic information of the study area

2 研究方法

2.1 新安江模型简介

新安江模型是一个分单元、分水源、分阶段的概念性水文模型，主要适用于湿润与半湿润地区的降水径流过程模拟^[37]。由于概念清晰、结构合理、调参方便和计算精度较高等优点，新安江模型在中国南方得到了广泛的应用，在广西壮族自治区也有许多成功应用实例，选择新安江模型作为广西中小河流洪水模拟模型。新安江模型主要由蒸散发、产流、分水源和汇流4个部分组成；根据土壤垂向分布的不均匀性将土层分为3层，用3层蒸散发模型计算流域实际蒸散发量；基于蓄满产流原理，利用流域蓄水容量曲线计算总产流量，再利用自由水蓄水库将径流划分为地表径流、壤中流和地下径流3种成分；最后利用单位线、线性水库和马斯京根法等推求流域出口断面的汇流过程。新安江模型共有16个参数需要率定，其物理意义见表1。

表 1 新安江模型参数
Table 1 Parameters of the Xin'anjiang model

参数	物理意义	参数	物理意义
K_C	蒸散发折算系数	E_X	流域自由水容量分布曲线指数
U_M / mm	上层张力水容量	K_g	地下水日出流系数
L_M / mm	下层张力水容量	K_i	壤中流日出流系数
C	深层蒸散发扩散系数	C_s	地表径流日消退系数
W_M / mm	流域张力水容量	C_i	壤中流日消退系数
B	流域蓄水容量分布曲线指数	C_g	地下径流日消退系数
I_M	不透水面积比例	C_r	河网水消退系数
S_M / mm	流域自由水容量	L_r / h	河网滞时

将 Nash-Sutcliffe 效率系数 (E_{NS})、年径流量相对误差 (E_R) 和洪水模拟合格率 (R_Q) 作为目标函数来率定新安江模型参数, 评价日径流与洪水过程模拟的精度。洪水模拟合格率包括洪峰许可误差、洪峰出现时间许可误差、径流深许可误差、过程许可误差等评价指标^[38]。Nash-Sutcliffe 效率系数、年径流量相对误差和洪水模拟合格率的计算公式为:

$$E_{NS} = 1 - \frac{\sum_{i=1}^n (Q_{oi} - Q_{si})^2}{\sum_{i=1}^n (Q_{oi} - \bar{Q}_o)^2} \quad (1)$$

$$E_R = \frac{\bar{Q}_s - \bar{Q}_o}{\bar{Q}_o} \times 100\% \quad (2)$$

$$R_Q = \frac{N_R}{N_S} \times 100\% \quad (3)$$

式中: Q_{oi} 和 Q_{si} 分别为观测与模拟的径流序列, m^3/s ; \bar{Q}_o 和 \bar{Q}_s 分别为观测与模拟的径流平均值, m^3/s ; n 为径流序列长度; N_R 和 N_S 分别为洪水模拟的合格次数和总次数。Nash-Sutcliffe 效率系数和合格率越大, 相对误差越小, 径流模拟效果越好。

2.2 模型参数移植方法

利用留一法, 依次将 42 个典型流域假定为无资料的目标流域。对于任意一个目标流域, 基于其他 41 个流域率定的模型参数, 利用回归分析、相似流域和机器学习等参数移植方法, 估算目标流域的模型参数。由于喀斯特地貌对产汇流过程有较大的影响, 在模型参数移植时, 根据水文分区特征分别在喀斯特地区和非喀斯特地区开展研究。基于多元回归方程的回归分析法在模型参数移植中有广泛的应用^[19], 首先在有资料的地区, 构建率定的模型参数和流域下垫面特性之间的多元回归方程; 然后在无资料地区, 由构建的回归方程和下垫面特性来计算模型参数。选取影响流域降水径流过程的 7 个指标来描述流域的物理特征, 包括反映地理位置的纬度 (L_{at}) 和经度 (L_{on})、表征气候的年降水量和干燥指数、表征植被的叶面积指数、表征地形的平均坡度以及反映流域大小的面积 (A)。根据典型流域率定的模型参数和下垫面特征构建的多元回归方程为

$$y = a_0 + \sum_{i=1}^m a_i x_i + \varepsilon \quad (4)$$

式中: y 为 16 个率定的新安江模型参数; x_i 为选取的流域特征; a_0, a_1, \dots, a_m 为回归方程的系数; ε 为随机误差; m 为回归方程的自变量个数。

相似流域法也是一种重要的参数移植方法^[21-22]。基于选取的 7 个流域特征, 利用聚类分析法判别流

域之间的相似程度，将相似流域的模型参数移植到目标流域。由于下垫面特征的单位不同，首先将其标准化。将 42 个流域的 7 个特征看作七维空间中的 42 个点，利用欧氏距离来度量 42 个点之间的接近程度。采用类平均法合并距离最近的 2 类为 1 个新类，重复该步骤直至类的个数为 1。目标流域模型参数的计算公式如下：

$$y_0 = \frac{1}{k} \sum_{i=1}^k y_i r_i + \varepsilon \tag{5}$$

式中： y_0 和 y_i 分别为目标流域和相似流域的模型参数； k 为相似流域的个数； r_i 为每个相似流域的权重，根据量纲一化后相似流域之间的欧氏距离倒数计算。

随着以机器学习为代表的人工智能技术的快速发展，统计模型方法取得了较大的进步。选用机器学习中广泛应用的支持向量机回归(SVR)、随机森林回归(RF)、极端森林回归(ETR)、提升树回归(GBR)和 K 临近回归(KNN)等 5 个机器学习算法来构建模型参数移植技术。将选用的 7 个流域特征作为自变量，率定的模型参数作为因变量，基于均方根误差(E_{RMS})等评价指标来训练 5 个机器学习算法，并将其与回归分析法和相似流域法的计算结果对比分析。针对每一个模型参数挑选最优的参数移植方法，将其作为新安江模型的参数移植方案。

2.3 模型参数移植误差驱动的径流模拟不确定性评估框架

采用基于随机模拟的数值实验方法定量评估参数移植误差驱动的径流模拟不确定性，计算框架见图 2，具体的计算步骤如下：

(1) 计算参数移植法推求的模型参数与直接率定值的误差(ε_p)：

$$\varepsilon_p = p_c - p_r \tag{6}$$

式中： p_r 和 p_c 分别为模型参数的移植和率定结果。

(2) 利用核密度估计方法分析参数移植误差的概率分布，再利用蒙特卡罗随机模拟方法生成误差样本($\hat{\varepsilon}_i$)，将其与模型参数移植结果相加得到模型参数样本(\hat{p}_i)：

$$\hat{p}_i = p_r + \hat{\varepsilon}_i \quad i = 1, 2, \dots, N \tag{7}$$

式中： N 为随机模拟的样本容量，这里取 5 000。

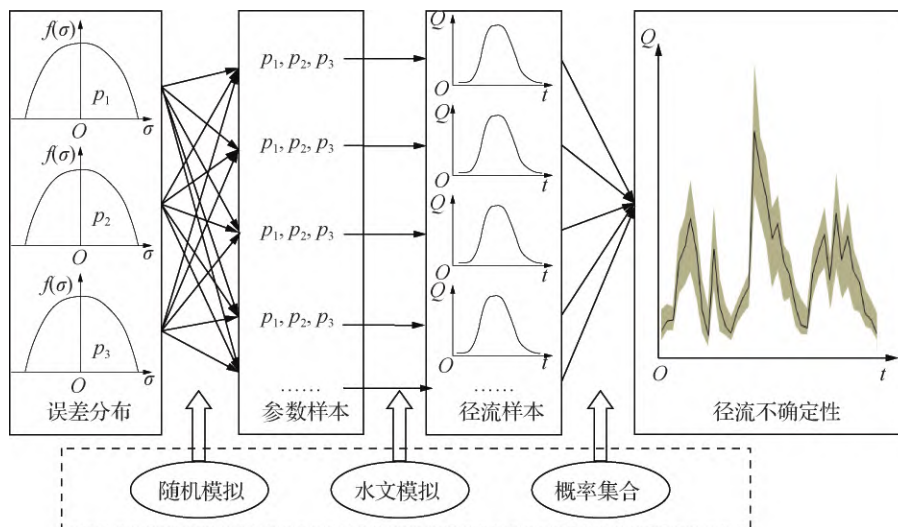


图 2 模型参数移植误差驱动的径流模拟不确定性评估框架

Fig. 2 Framework of streamflow simulation uncertainty introduced by parameters' error estimated by regionalization methodologies

(3) 利用新安江模型模拟每个参数样本下的降水径流过程, 得到径流样本。基于普适似然不确定性估计方法^[39], 根据每个参数样本模拟径流 Nash-Sutcliffe 效率系数的量纲一化结果, 计算每个参数样本的权重系数, 将其作为每组参数样本的概率。对每一场降雨径流过程, 将每个时刻的 5 000 组径流模拟值从小到大排序, 统计 5%、25%、50%、75% 和 95% 累计概率下的径流模拟值, 分析模拟径流的 90% 和 50% 置信区间以及中位数, 得到参数移植误差引起的模拟径流不确定性。

3 结果及讨论

3.1 新安江模型率定与验证

利用广西 42 个水文站 2005—2014 年实测日径流和洪水过程资料率定新安江模型参数, 验证模型在区域的适用性。其中, 2005—2011 年的资料用于模型参数率定, 2012—2014 年的资料用于模型验证。42 个典型流域在模拟期与检验期的日过程和洪水过程精度评价统计结果见表 2。总体而言, 新安江模型对 42 个典型流域的日径流过程和洪水过程具有较好的模拟效果。日径流模拟中, 在率定期和检验期, 平均 E_{NS} 分别为 0.84 和 0.75, 其中, 最大值为 0.947, 最小值为 0.601; 平均相对误差分别为 3.25% 和 11%; $E_{NS} \geq 0.8$ 的流域个数分别占 74% 和 57%; $E_{NS} \leq 0.7$ 的流域个数分别仅占 7% 和 21%。在洪水模拟中, 在率定期和检验期, 平均 E_{NS} 分别为 0.58 和 0.63; 平均合格率分别为 61.5% 和 58.3%, $E_{NS} \geq 0.6$ 的流域个数分别占 55% 和 71%。与一般的大江大河相比, 研究区洪水模拟精度相对不高, 这主要是由于本研究选用的 42 个典型流域是中小河流, 流域面积都小于 3 000 km², 中小河流的调蓄作用较小, 洪水模拟精度普遍小于大江大河。同时选取的大部分典型流域位于喀斯特地区, 对流域汇流有较大的影响, 本文采用的新安江模型结构中未考虑喀斯特的作用, 此外很难精确获取详细的喀斯特地区地下河汇流路径, 这也是典型流域洪水模拟精度相对不高的一个重要原因。

表 2 新安江模型模拟日过程和洪水径流过程精度评价统计

Table 2 Performance of the Xin'anjiang model in daily streamflow and flood simulation

时间	日过程模拟流域个数				洪水过程模拟流域个数			
	$E_{NS} \geq 0.9$	$0.8 \leq E_{NS} < 0.9$	$0.7 \leq E_{NS} < 0.8$	$E_{NS} < 0.7$	$E_{NS} \geq 0.8$	$0.7 \leq E_{NS} < 0.8$	$0.6 \leq E_{NS} < 0.7$	$E_{NS} < 0.6$
率定期	13	18	8	3	8	10	5	19
检验期	7	17	9	9	14	9	7	12

3.2 模型参数移植

率定的日过程和洪水过程的 16 个模型参数与 7 个流域特征的相关系数见图 3。大部分模型参数与流域特征的相关性较弱。相对而言, 洪水过程中控制汇流的参数, 例如 C_s 、 C_r 、 L_r 等与流域大小的相关性强于日过程参数。针对每个模型参数, 利用 AIC 准则遴选回归方程的自变量个数。利用回归分析法、相似流域法和 5 种机器学习算法移植计算的新安江模型参数与直接率定值的均方根误差见图 4。从图 4 中可以看出回归分析法计算的 16 个模型参数的均方根误差总体小于相似流域法。16 个模型参数相比, 对径流模拟敏感的参数移植结果与率定值的误差较小。这是因为模型参数之间不独立, 存在异参同效现象, 不敏感参数尤其显著。同时敏感参数的物理意义较强, 与流域特征的物理关系较密切; 而不敏感参数的概化程度较高, 与流域特征的物理关系较弱^[19, 23]。因此, 基于模型参数与流域特征回归方程计算的敏感参数相关性高于不敏感参数。基于机器学习算法移植的模型参数误差总体上小于回归分析法和相似流域法的计算结果。7 种参数移植方法对比结果表明, 支持向量机回归和提升树回归计算的模型参数误差最小。根据均方根误差对比结果, 针对每个

模型参数分别挑选误差最小的方法将其作为优选的参数移植方案。

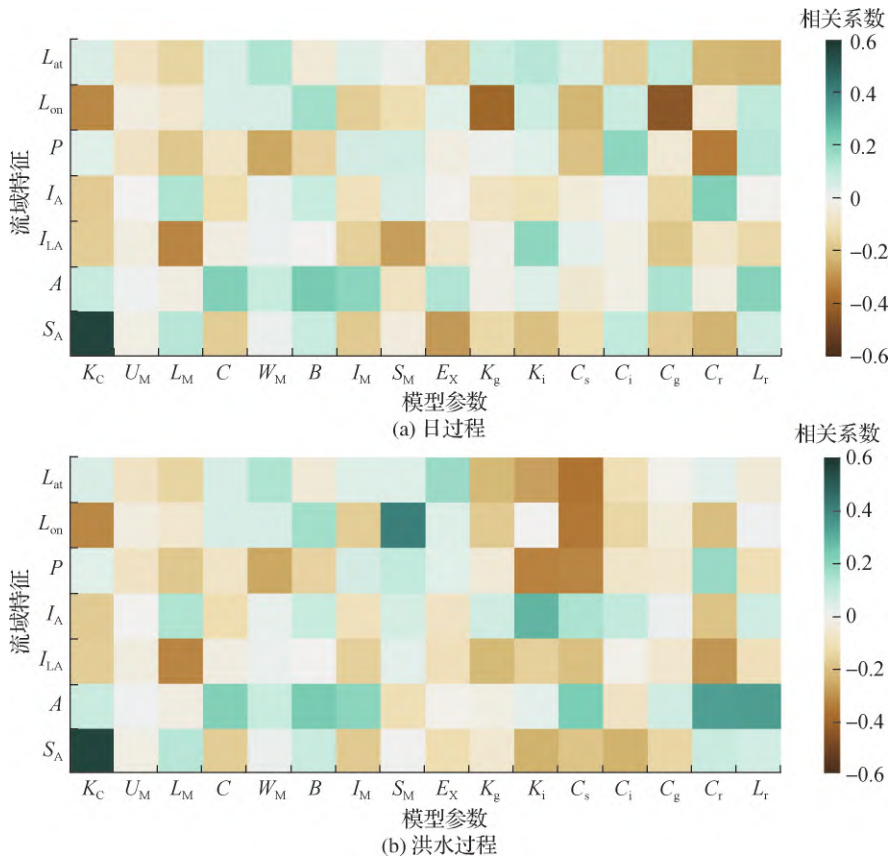


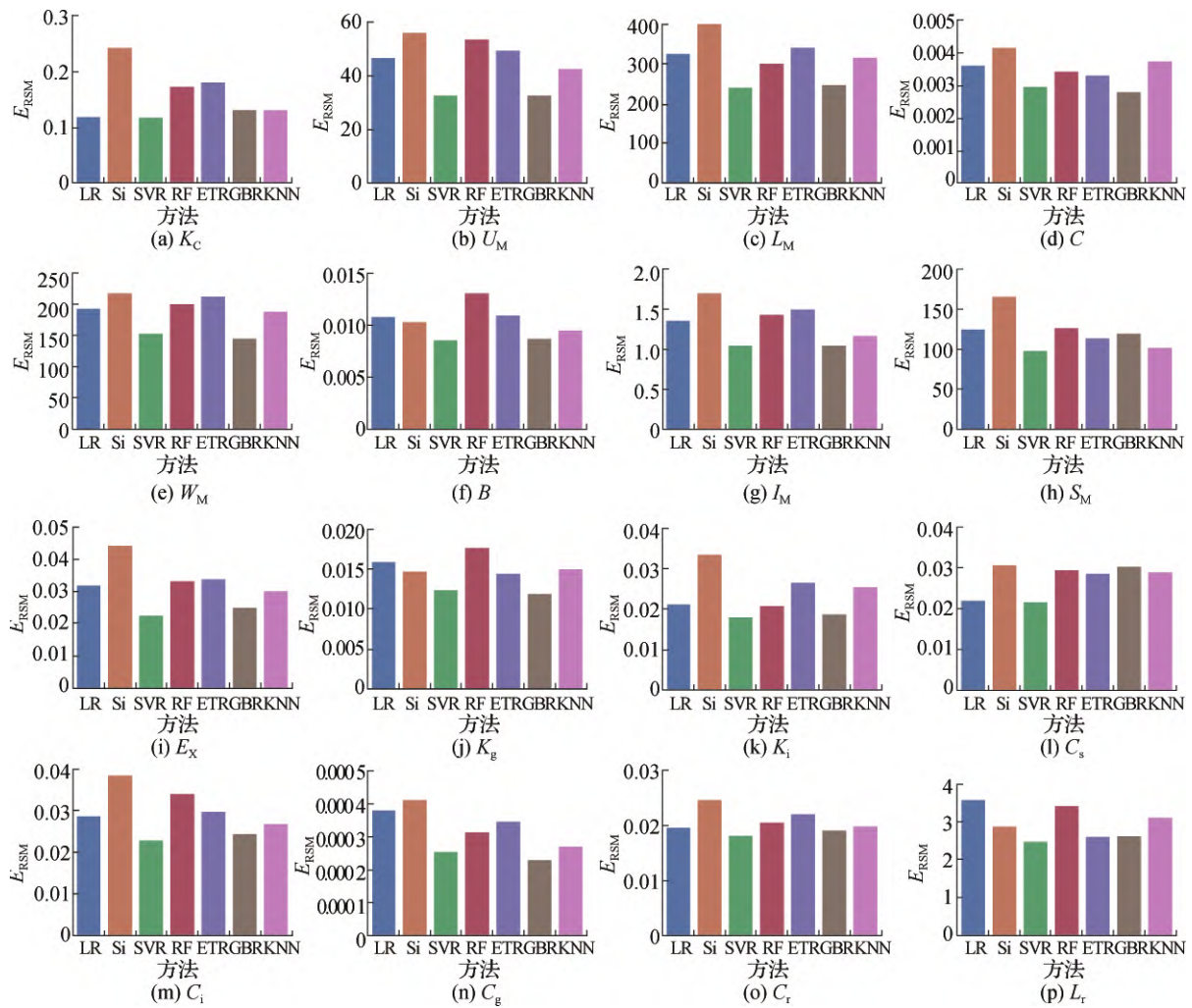
图 3 率定的模型参数与流域特征相关系数

Fig. 3 Correlation coefficients of calibrated model parameters versus catchment characteristics

基于回归分析、相似流域和优选的参数移植方法模拟洪水过程的 Nash-Sutcliffe 效率系数与合格率对比结果见图 5。从图 5 中可以看出，总体而言基于回归分析法模拟洪水过程的精度优于相似流域法，优选的机器学习算法的模拟精度最高。在 42 个典型流域中，对于 Nash-Sutcliffe 效率系数，回归分析法、相似流域法和机器学习优选法计算的 50% 分位数分别为 0.467、0.438 和 0.503。回归分析法和相似流域法相比，31 个典型流域的回归分析法模拟效果较好，11 个典型流域的相似流域法模拟效果较好。对于合格率，回归分析法模拟效果较好的有 19 个典型流域，相似流域法模拟效果较好的有 10 个，另有 13 个流域 2 种方法的模拟效果相当。优选的机器学习算法模拟洪水过程的合格率比回归分析法和相似流域法的合格率分别提高了 7% 和 15%。无论是 Nash-Sutcliffe 效率系数还是合格率，参数移植方法的模拟结果与参数率定的结果都有一定的差距，即基于模型参数移植的无资料地区径流模拟具有较大的误差，需要进一步评估其不确定性。

3.3 模型参数移植误差分析

利用核密度估计方法计算基于参数移植法估计的模型参数误差概率分布见图 6。从图 6 中可以看出，相对而言，参数 K_C 、 W_M 、 I_M 、 S_M 、 E_X 、 K_g 、 C_i 、 C_g 和 C_r 等的相对误差较小，25% ~ 75% 区间在 -20% ~ 20% 范围内；参数 U_M 、 L_M 、 C 、 B 、 K_i 、 C_s 等的误差较大，25% ~ 75% 区间超过 20%。图 6 中参数误差的概率分布越集中，表示计算的参数误差越小。多种参数移植方法相比，优选的机器学习算法计算模型参数的误差最小，其次是回归分析法，相似流域法的误差最大。 K_C 等大部分参数的误差呈对称分布， W_M 等部分参数的误差呈偏态分布。此外由于模型参数的不独立性会影响径流模拟不确定性的评估结果^[40]，对 16 个模型参数移



LR: 回归分析法; Si: 相似流域法

图 4 基于 7 种参数移植法计算的模型参数与直接率定值相比的均方根误差

Fig. 4 Root mean square errors of regionalized and calibrated model parameters

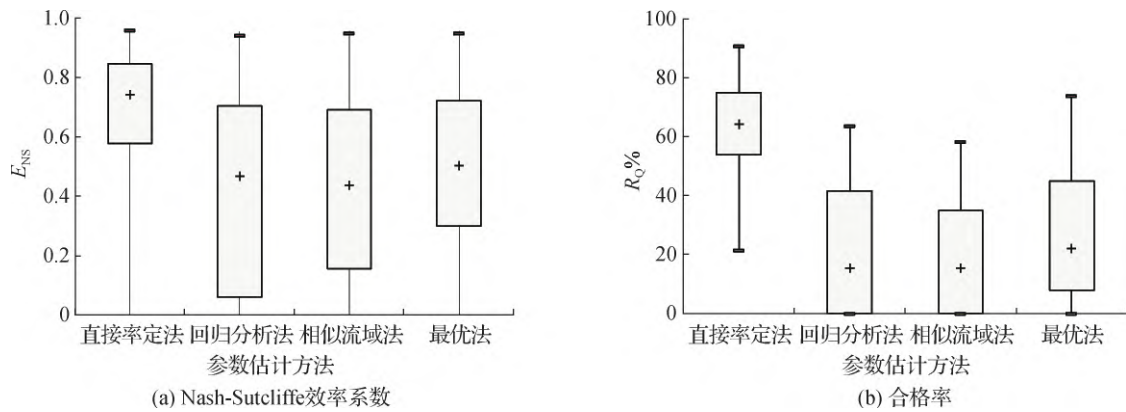


图 5 基于参数移植法模拟的洪水过程精度评价箱线图

Fig. 5 Boxplot of Nash-Sutcliffe efficiency coefficient and accuracy rate of flood simulation by parameters regionalization methodologies

植误差的相关性作统计检验，结果表明模型参数移植误差之间的相关性不显著，具有较好的独立性。在此基础上，基于核密度估计方法计算的概率分布特征，对每个模型参数分别独立利用蒙特卡罗方法随机模拟误差样本。

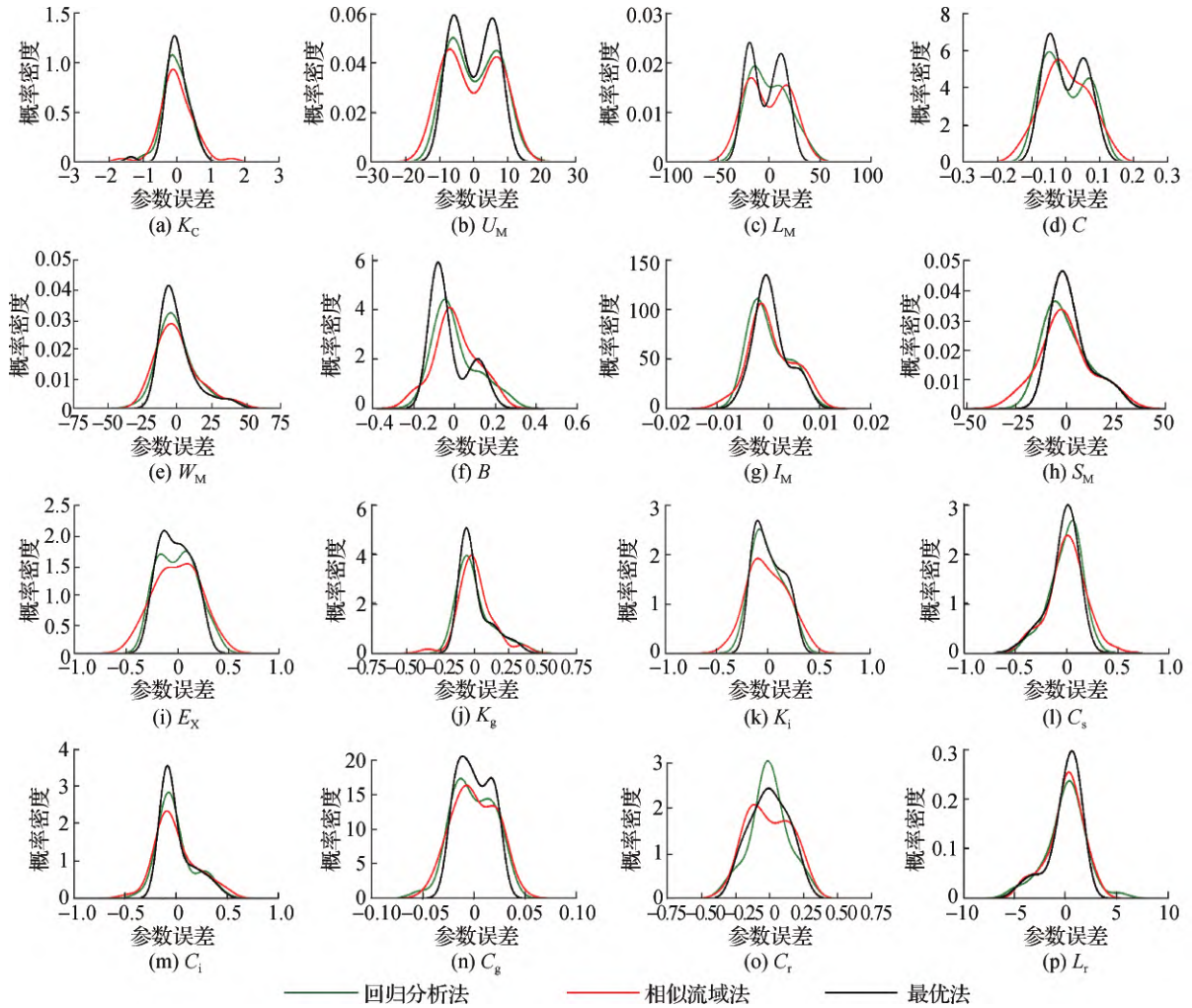


图6 基于参数移植方法计算的模型参数误差概率分布

Fig. 6 Probability distribution of parameters' error estimated by regionalized-based methodology

3.4 径流模拟不确定性评估

根据图2的不确定性评估框架，利用蒙特卡罗随机模拟方法对参数移植误差进行重采样，形成5 000组参数样，在此基础上模拟每组参数样本下的洪水过程，得到 Nash-Sutcliffe 效率系数、洪量相对误差和洪峰相对误差等洪水过程模拟精度评价指标的统计特征及典型场次洪水过程，见图7。总体上，由模型参数移植误差驱动模拟的洪水过程具有较大的不确定性。从图7可以看出，Nash-Sutcliffe 效率系数主要位于0.25~0.75之间，其50%分位数为0.445；洪量相对误差主要位于10%~30%之间，其50%分位数为21.3%；洪峰相对误差主要位于10%~40%之间，其50%分位数为24.1%。该方法可以展示模型参数移植误差带来的洪水模拟不确定性分布区间。因此，在利用参数移植方法模拟无资料地区径流过程时，应充分考虑模型参数移植误差带来的径流模拟不确定性^[34-36]。

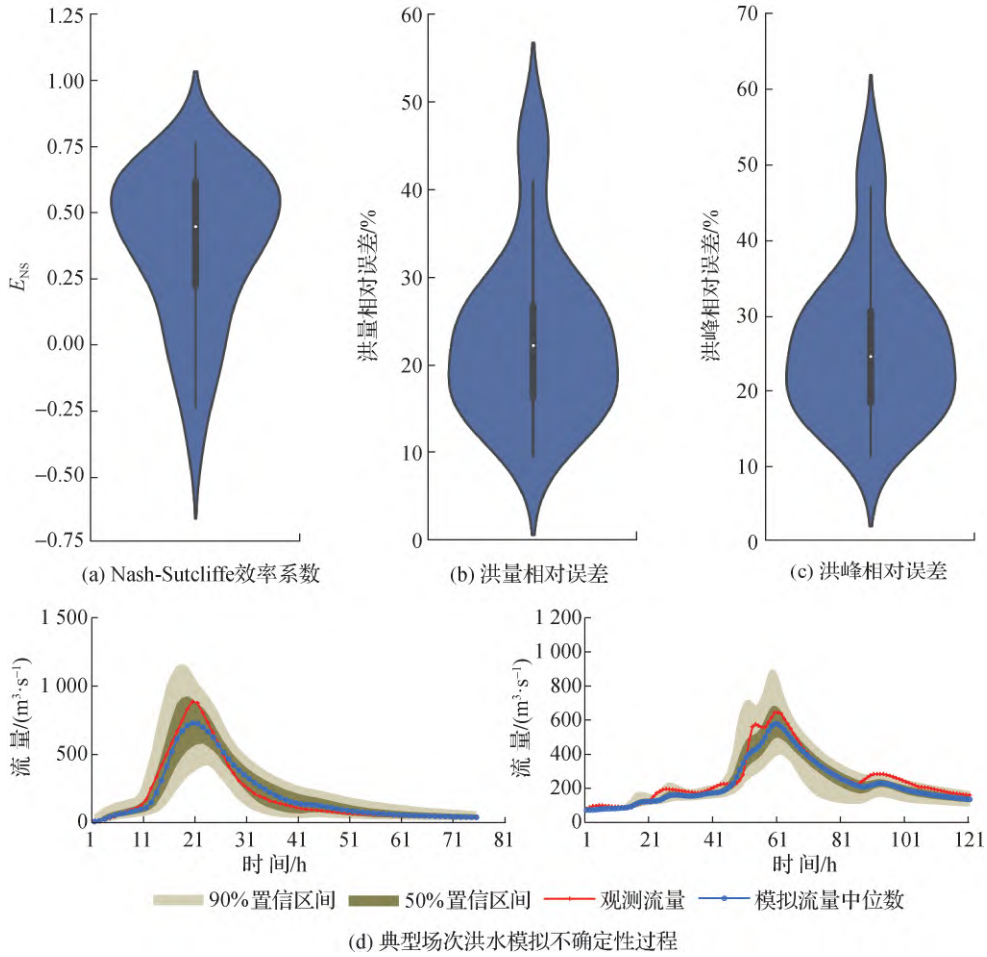


Fig. 7 Violin plot of accuracy index of stochastically modeled flood and uncertainties of typical floods

4 结 论

无资料地区径流模拟不确定性评估是当前研究的薄弱环节,以广西壮族自治区 42 个有水文监测站点的典型中小河流为研究对象,在新安江模型模拟日径流过程和洪水过程的基础上,对比了回归分析法、相似流域法、机器学习算法等模型参数移植方法模拟洪水过程的精度,分析了模型参数移植误差的统计特征,评估了无资料地区模型参数移植误差驱动的径流模拟不确定性,主要结论如下:

(1) 新安江模型在广西壮族自治区具有一定的适用性,日过程和洪水过程径流模拟的 Nash-Sutcliffe 效率系数分别为 0.75 ~ 0.84 和 0.58 ~ 0.63。

(2) 相比而言,基于回归分析的参数移植法计算的模型参数和模拟洪水过程的精度优于相似流域法,优选的机器学习算法模拟洪水过程的合格率比传统回归分析法和相似流域法的精度提高了 7% ~ 15%。

(3) 与模型参数率定值相比,参数移植法计算的模型参数具有一定的误差,总体上敏感参数的相对误差小于不敏感参数,应重点关注敏感参数的移植计算。基于模型参数移植误差概率分布,利用蒙特卡罗随机模拟的洪水过程具有一定的不确定性,洪量和洪峰相对误差的主要区间分别为 10% ~ 30% 和 10% ~ 40%。

参数移植是估算无资料地区水文模型参数的重要方法之一,但是计算结果具有一定的误差,引起了径流

模拟的不确定性。如何利用更多的流域特征信息和人工智能等技术研究新的无资料地区径流模拟方法、降低其不确定性是未来研究的发展方向。

参考文献:

- [1] 李致家, 梁珂, 阚光远, 等. 新安江模型中河网汇流参数 C_s 的一种计算方法[J]. 水科学进展, 2016, 27(5): 652-661. (LI Z J, LIANG K, KAN G Y, et al. A method for deriving the river network flow concentration parameter C_s of the Xin'anjiang model[J]. Advances in Water Science, 2016, 27(5): 652-661. (in Chinese))
- [2] XU X, SUN C, HUANG G H, et al. Global sensitivity analysis and calibration of parameters for a physically-based agro-hydrological model[J]. Environmental Modelling & Software, 2016, 83: 88-102.
- [3] 张静文, 郭家力, 王敬斌, 等. 水文模型参数选取对模拟径流的年际年内分布影响评估[J]. 气候变化研究进展, 2020, 16(3): 325-335. (ZHANG J W, GUO J L, WANG J B, et al. Appraising the effect of choosing hydrological model parameters on the inter-annual and intra-annual distribution of simulated runoff[J]. Climate Change Research, 2020, 16(3): 325-335. (in Chinese))
- [4] 《第一次全国水利普查成果丛书》编委会. 河湖基本情况普查报告[M]. 北京: 中国水利水电出版社, 2017. (Editorial board of "A Series of the First National Census for Water". Bulletin of census for river and lake[M]. Beijing: China Water & Power Press, 2017. (in Chinese))
- [5] MONTANARI A, YOUNG G, SAVENIJE H H G, et al. "Panta rhei—everything flows": change in hydrology and society: the IAHS scientific decade 2013—2022[J]. Hydrological Sciences Journal, 2013, 58(6): 1256-1275.
- [6] 熊立华, 刘烁楠, 熊斌, 等. 考虑植被和人类活动影响的水文模型参数时变特征分析[J]. 水科学进展, 2018, 29(5): 625-635. (XIONG L H, LIU S N, XIONG B, et al. Impacts of vegetation and human activities on temporal variation of the parameters in a monthly water balance model[J]. Advances in Water Science, 2018, 29(5): 625-635. (in Chinese))
- [7] SIVAPALAN M, TAKEUCHI K, FRANKS S W, et al. IAHS decade on Predictions in Ungauged Basins (PUB), 2003—2012: shaping an exciting future for the hydrological sciences[J]. Hydrological Sciences Journal, 2003, 48(6): 857-880.
- [8] GRIMALDI S, NARDI F, PISCOPIA R, et al. Continuous hydrologic modelling for design simulation in small and ungauged basins: a step forward and some tests for its practical use[J]. Journal of Hydrology, 2021, 595: 125664.
- [9] 李巧玲, 马亚楠, 李致家, 等. 无资料小水库影响下的中型水库入库洪水模拟[J]. 河海大学学报(自然科学版), 2021, 49(3): 213-219. (LI Q L, MA Y N, LI Z J, et al. Inflow flood simulation of medium reservoirs under impact of ungauged small reservoirs[J]. Journal of Hohai University (Natural Sciences), 2021, 49(3): 213-219. (in Chinese))
- [10] HRACHOWITZ M, SAVENIJE H H G, BLÖSCHL G, et al. A decade of Predictions in Ungauged Basins (PUB): a review[J]. Hydrological Sciences Journal, 2013, 58(6): 1198-1255.
- [11] 王璟京, 高红凯, 刘敏, 等. 参数区域化在全球水文模型 FLEX-Global 中的应用[J]. 中国科学: 地球科学, 2021, 51(5): 805-823. (WANG J J, GAO H K, LIU M, et al. Parameter regionalization of the FLEX-Global hydrological model[J]. Scientia Sinica Terrae, 2021, 51(5): 805-823. (in Chinese))
- [12] WASEEM M, AJMAL M, KIM T W. Ensemble hydrological prediction of streamflow percentile at ungauged basins in Pakistan[J]. Journal of Hydrology, 2015, 525: 130-137.
- [13] 李红霞, 王瑞敏, 黄琦, 等. 中小河流洪水预报研究进展[J]. 水文, 2020, 40(3): 16-23, 50. (LI H X, WANG R M, HUANG Q, et al. Advances on flood forecasting of small-medium rivers[J]. Journal of China Hydrology, 2020, 40(3): 16-23, 50. (in Chinese))
- [14] WOLFF W, DUARTE S N. Toward geostatistical unbiased predictions of flow duration curves at ungauged basins[J]. Advances in Water Resources, 2021, 152: 103915.
- [15] ATIEH M, TAYLOR G, SATTAR A M A, et al. Prediction of flow duration curves for ungauged basins[J]. Journal of Hydrology, 2017, 545: 383-394.
- [16] DU T L T, LEE H, BUI D D, et al. Streamflow prediction in "geopolitically ungauged" basins using satellite observations and regionalization at subcontinental scale[J]. Journal of Hydrology, 2020, 588: 125016.
- [17] KRATZERT F, KLOTZ D, HERRNEGGER M, et al. Toward improved predictions in ungauged basins: exploiting the power of machine learning[J]. Water Resources Research, 2019, 55(12): 11344-11354.

- [18] PRIETO C, le VINE N, KAVETSKI D, et al. Flow prediction in ungauged catchments using probabilistic random forests regionalization and new statistical adequacy tests[J]. *Water Resources Research*, 2019, 55(5): 4364–4392.
- [19] FOUAD G, LOÁICIGA H A. Independent variable selection for regression modeling of the flow duration curve for ungauged basins in the United States[J]. *Journal of Hydrology*, 2020, 587: 124975.
- [20] TEGEGNE G, KIM Y O. Modelling ungauged catchments using the catchment runoff response similarity[J]. *Journal of Hydrology*, 2018, 564: 452–466.
- [21] CHOUBIN B, SOLAIMANI K, REZANEZHAD F, et al. Streamflow regionalization using a similarity approach in ungauged basins: application of the geo-environmental signatures in the Karkheh River basin, Iran[J]. *CATENA*, 2019, 182: 104128.
- [22] 彭安帮, 刘九夫, 马涛, 等. 辽宁省资料短缺地区中小河流洪水预报方法[J]. *水力发电学报*, 2020, 39(8): 79–89. (PENG A B, LIU J F, MA T, et al. Flood forecasting method for medium-and small-size rivers short of hydrological data in Liaoning Province[J]. *Journal of Hydroelectric Engineering*, 2020, 39(8): 79–89. (in Chinese))
- [23] BAO Z X, ZHANG J Y, LIU J F, et al. Comparison of regionalization approaches based on regression and similarity for predictions in ungauged catchments under multiple hydro-climatic conditions[J]. *Journal of Hydrology*, 2012, 466/467: 37–46.
- [24] 龚珺夫, 陈红兵, 朱芳, 等. 新安江模型在资料匮乏的长江中下游山区中小流域洪水预报应用[J]. *湖泊科学*, 2021, 33(2): 581–594, 650. (GONG J F, CHEN H B, ZHU F, et al. Application of Xin'anjiang model in the flow prediction of ungauged small-and mediumsized catchments in the middle and lower reaches of the Yangtze River basin[J]. *Journal of Lake Sciences*, 2021, 33(2): 581–594, 650. (in Chinese))
- [25] HAQUE M M, RAHMAN A, HAGARE D, et al. Parameter uncertainty of the AWBM model when applied to an ungauged catchment[J]. *Hydrological Processes*, 2015, 29(6): 1493–1504.
- [26] WANG H, GONG W, DUAN Q Y, et al. Evaluation of parameter interaction effect of hydrological models using the sparse polynomial chaos (SPC) method[J]. *Environmental Modelling & Software*, 2020, 125: 104612.
- [27] LIU X P, LU M Z, CHAI Y Z, et al. A comprehensive framework for HSPF hydrological parameter sensitivity, optimization and uncertainty evaluation based on SVM surrogate model: a case study in Qinglong River watershed, China[J]. *Environmental Modelling & Software*, 2021, 143: 105126.
- [28] 宋晓猛, 孔凡哲, 占车生, 等. 基于统计理论方法的水文模型参数敏感性分析[J]. *水科学进展*, 2012, 23(5): 642–649. (SONG X M, KONG F Z, ZHAN C S, et al. Sensitivity analysis of hydrological model parameters using a statistical theory approach[J]. *Advances in Water Science*, 2012, 23(5): 642–649. (in Chinese))
- [29] BLÖSCHL G, SIVAPALAN M, WAGENER T, et al. Runoff prediction in ungauged basins: synthesis across processes, places and scales[M]. Cambridge: Cambridge University Press, 2013.
- [30] ZHANG J L, LI Y P, HUANG G H, et al. Assessment of parameter uncertainty in hydrological model using a Markov-Chain-Monte-Carlo-based multilevel-factorial-analysis method[J]. *Journal of Hydrology*, 2016, 538: 471–486.
- [31] LIU Y R, LI Y P, HUANG G H, et al. A Bayesian-based multilevel factorial analysis method for analyzing parameter uncertainty of hydrological model[J]. *Journal of Hydrology*, 2017, 553: 750–762.
- [32] KIANG J E, GAZOORIAN C, MCMILLAN H, et al. A comparison of methods for streamflow uncertainty estimation[J]. *Water Resources Research*, 2018, 54(10): 7149–7176.
- [33] 刘章君, 郭生练, 许新发, 等. 贝叶斯概率水文预报研究进展与展望[J]. *水利学报*, 2019, 50(12): 1467–1478. (LIU Z J, GUO S L, XU X F, et al. Bayesian probabilistic hydrological forecasting: progress and prospects[J]. *Journal of Hydraulic Engineering*, 2019, 50(12): 1467–1478. (in Chinese))
- [34] WAGENER T, MONTANARI A. Convergence of approaches toward reducing uncertainty in predictions in ungauged basins[J]. *Water Resources Research*, 2011, 47(6): W06301.
- [35] WESTERBERG I K, WAGENER T, COXON G, et al. Uncertainty in hydrological signatures for gauged and ungauged catchments[J]. *Water Resources Research*, 2016, 52(3): 1847–1865.
- [36] BRUNNER M I, SIKORSKA A E, FURRER R, et al. Uncertainty assessment of synthetic design hydrographs for gauged and ungauged catchments[J]. *Water Resources Research*, 2018, 54(3): 1493–1512.
- [37] 陆曼皎. 新安江模型研究的回顾和展望[J]. *水利学报*, 2021, 52(4): 432–441. (LU M J. Recent and future studies of the Xin'anjiang model[J]. *Journal of Hydraulic Engineering*, 2021, 52(4): 432–441. (in Chinese))

- [38] 水文情报预报规范: GB/T 22482—2008 [S]. 北京: 中国标准出版社, 2009. (Standard for hydrological information and hydrological forecasting: GB/T 22482—2008 [S]. Beijing: Standards Press of China, 2009. (in Chinese))
- [39] ZHOU R R, LI Y, LU D, et al. An optimization based sampling approach for multiple metrics uncertainty analysis using generalized likelihood uncertainty estimation [J]. *Journal of Hydrology*, 2016, 540: 274–286.
- [40] GHAITH M, LI Z, BAETZ B W. Uncertainty analysis for hydrological models with interdependent parameters: an improved polynomial chaos expansion approach [J]. *Water Resources Research*, 2021, 57(8): e2020WR029149.

Uncertainties of model parameters regionalization in ungauged basins*

GUAN Tiesheng^{1 2}, BAO Zhenxin^{1 2 3}, HE Ruimin^{1 2}, YANG Yanqing⁴, WU Houfa^{1 2}

- (1. *The National Key Laboratory of Water Disaster Prevention, Nanjing Hydraulic Research Institute, Nanjing 210029, China;*
2. *Research Center for Climate Change of MWR, Nanjing 210029, China;* 3. *Yangtze Institute for Conservation and Development, Nanjing 210098, China;* 4. *Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610299, China*)

Abstract: Prediction in ungauged basins is a challenge and hot issue. Parameters regionalization is a useful methodology estimating hydrological model parameters in ungauged basins and has a critical effect on streamflow simulation. With kernel density estimation and Monte Carlo stochastic simulation methods, a framework was constructed to assess the uncertainty of simulated streamflow caused by parameters' error estimated by regionalization methodology. The Xin'anjiang model was applied for streamflow simulation in 42 small and medium-sized catchments with observed hydrologic stations located in the Guangxi Province. As each catchment being supposed an ungauged basin, the parameters of the Xin'anjiang model were calculated by regionalization methodologies including regression-based, similarity-based, and machine learning-based methodology. The performance of flood simulation using regression-based methodology was better than that of the similarity-based methodology. Using optimized machine learning-based regionalization methodology, the flood simulation accuracy was improved by 7%—15%. Compared with calibrated values, there were pronounced errors of model parameters estimated by parameters regionalization methodologies. The errors of sensitive parameters were lower than non-sensitive ones. The results indicated that there were significant uncertainties of randomly modeled floods by Monte Carlo methodology. The relative errors of simulated flood volumes and peak discharges were 10%—30% and 10%—40%, respectively. The results could provide a new technique for streamflow probability modeling and uncertainty assessment in ungauged basins. And this would be useful for flood forecasting and disaster prevention in small and medium-sized rivers.

Key words: ungauged basin; streamflow simulation; uncertainty; Xin'anjiang model; parameters regionalization

* The study is financially supported by the National Key R&D Program of China (No. 2022YFC3205200) and the National Natural Science Foundation of China (No. 41961124007) .