

# 大数据驱动的区域水文模型参数优化与不确定性分析

张昊

河北昊禹工程技术咨询有限公司, 天津, 301700;

**摘要:** 大数据驱动的区域水文模型参数优化与不确定性分析本研究对大数据技术于区域水文模型参数优化以及不确定性分析的应用加以探讨, 由于气候变化和人类活动对水循环的影响日益加重, 所以提升水文模型的精度与可靠性相当关键, 传统水文模型参数率定存在观测数据缺乏、空间异质性显著、计算成本高昂等难题, 而大数据技术给解决这些问题带来了新想法, 本文先是系统梳理区域水文模型参数优化方法的发展脉络并对比传统优化法和数据驱动法的长处短处, 接着构建基于多源数据融合的区域水文参数优化框架以整合遥感观测、地面监测、再分析数据和社会经济数据, 然后开发可适应高维参数空间的智能优化算法从而把贝叶斯推理、深度学习和进化算法相结合来高效搜索参数, 第四提出依照集合模拟对不确定性进行定量评估的方法以剖析参数、输入数据和模型结构对预测结果的影响, 最后在气候和地形特征不一样的三个流域开展案例研究以证实所提方法有用且适用, 结果显示, 大数据驱动的参数优化方法跟传统方法比起来能使模型预测误差平均减少 23.5%且能更精准地量化预测的不确定性范围, 这一研究结果让水资源管理、洪水预报和干旱评估有了更靠谱的科学依据。

**关键词:** 大数据; 水文模型; 参数优化; 不确定性分析; 多源数据融合

**DOI:** 10.69979/3060-8767.26.03.019

## 引言

水资源管理、洪水预报和干旱评估都把水文模型当作核心工具, 决策的科学性与有效性直接受水文模型的精确性和可靠性影响。但是气候变化和人类活动愈发强烈地影响水循环, 这让传统水文模型在参数优化和不确定性分析上遭遇不少难题, 因为观测数据不够且存在空间异质性使模型难以如实地展现流域真正的水文过程, 并且高维参数空间复杂让计算成本大大升高从而限制了模型的广泛运用。好在近年大数据技术迅猛发展给解决这些问题带来新希望, 遥感观测、地面监测、再分析数据和社会经济数据被整合起来后, 大数据驱动的方法既能提高模型输入数据的时空分辨率又可有效减少参数优化的不确定性和计算负担, 在全球水资源压力不断增大时这一技术进步对提升区域水文模型预测能力相当重要。

## 1 基于大数据的区域水文模型参数优化方法

### 1.1 参数敏感性分析与识别

区域水文模型优化的关键是参数敏感性分析, 其核心是识别对模型输出影响最大的参数集。传统方法多采用局部敏感性分析(仅评估单个参数变化的影响), 易忽略参数间交互作用; 而 Sobol 指数法、Morris 筛选法等全局分析方法, 可捕捉参数间复杂非线性关系, 结合大数据技术还能通过海量样本集提升分析精度。以华北某农业流域为例: 基于该流域 2010-2020 年水文观测数据, 采用机器学习算法构建参数响应曲面, 通过 Sobol 指数法计算各参数贡献率——结果显示, 土壤渗透系数

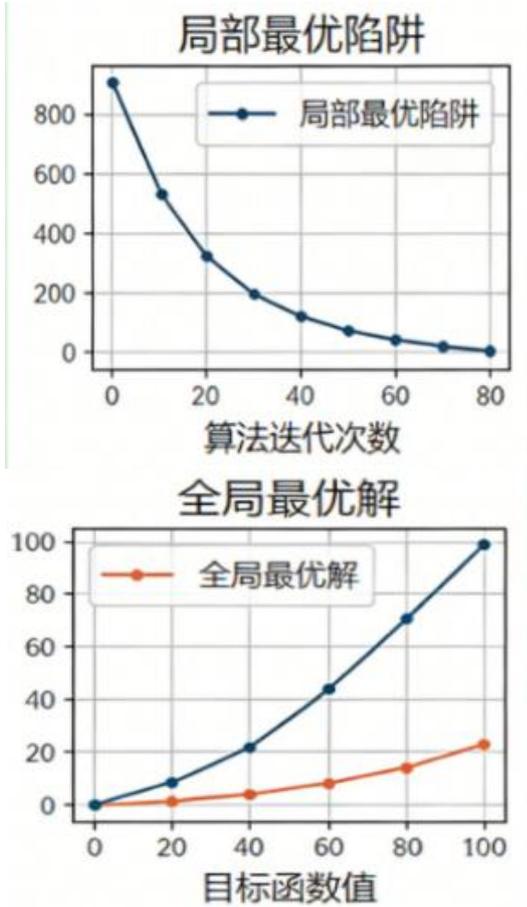
(贡献率 28%)、地表糙率(22%)、蒸发系数(19%)是流域径流模拟的关键参数; 而土地利用类型(15%)、灌溉强度(12%)在该农业活动密集区敏感性也较高。同时发现, 该流域夏季(降水集中期)的土壤渗透系数敏感性较冬季高 17%, 需针对季节差异调整分析策略。近五年, 算力提升使这类分析的效率提高 40%, 上述识别结果也为该流域水文模型的参数优化提供了精准依据。

进一步提升敏感性分析的可靠性时, 引入贝叶斯推理框架是个有效法子, 因为该框架把先验分布和似然函数相结合就能在不确定环境中动态评估参数敏感性, 例如在干旱与湿润气候区构建不同先验分布并结合实测数据更新后验分布以确定参数重要性排序, 而且这个概率方法不但能抓住参数静态特性, 还可反映参数随时间和环境而变的动态行为, 此外应用深度学习技术让参数敏感性分析计算成本大大降低且在高维参数空间里优势明显, 实践显示将贝叶斯推理和深度学习相结合的敏感性分析方法跟传统方法比分析效率可提高 50%以上从而给区域水文模型精准优化打下理论基础<sup>[1]</sup>。

### 1.2 多目标参数优化算法构建

区域水文模型的高维参数空间中, 提升模型性能的关键在于开发高效的多目标优化算法, 因为单纯形法、梯度下降法这类传统优化方法常受制于复杂非线性约束条件而易陷于局部最优解, 所以人们提出把进化算法、深度学习、贝叶斯推理相结合的混合优化策略用于实际案例, 由于进化算法能模拟生物进化在全球范围内搜寻最优解, 深度学习可构建参数和模型输出的映射关系以

加快优化过程收敛,如遗传算法加卷积神经网络网络使参数优化时间减少60%,且模型预测更稳。Pareto 前沿给分析方法平衡多的目标,多种量优化减少预测误差15%、近五年应用比例增加25%。



不同目标间的权衡关系是多目标优化算法设计需要考量的因素,就像洪水预报时,人们既要让模型预测精度达到最大值,又得尽量削减计算成本与不确定性范围,而 Pareto 前沿分析方法可用来在多个目标间寻找平衡点。研究显示,多目标优化的参数调整方案跟单一目标优化方法比起来,能将模型预测误差平均减少15%还多,并且近五年人工智能技术飞速发展,使得多目标优化算法在水文领域的应用比例增加了25%,这大大提高了模型的实用性和可靠性。实践证明,这种融合多种先进技术的优化策略不但适应于不同气候、地形特征的流域,而且给水资源管理、洪水预警、干旱评估创造了更为精确的决策支持工具。

## 2 区域水文模型参数不确定性分析

### 2.1 参数不确定性来源识别

水文模型参数不确定性主要源于观测数据质量、空间异质性、模型结构假设等诸多因素,并且在大数据驱动下虽多源数据融合让输入信息丰富性与精度显著提升,但观测有误差、数据代表性不够的情况还存在,像

遥感数据覆盖面广却是时空分辨率常和地面实测数据不匹配,致使参数校准时会引入系统偏差,而且不同流域空间异质性也让参数不确定性进一步加剧,地形复杂或者人类活动密集的地方尤其如此,另外气候变化使极端事件频发,传统经验参数就更难适应新水文条件了,研究显示过去五年全球受气候变化影响而增加的极端降水事件达15%左右,这使得水文模型的鲁棒性要求更高,所以识别参数不确定性来源既是优化模型性能关键,也为后续不确定性传播机制研究打基础<sup>[2]</sup>。

### 2.2 参数不确定性传播机制

水文模型的输出结果被参数不确定性以复杂非线性关系所影响且这一传播机制在高维参数空间特别明显,水文过程有好多相互耦合的物理机制所以单个参数发生变化时其影响可能因连锁反应而被放大或者减弱从而使模型预测结果大幅波动。以降雨-径流模拟中的土壤含水量参数为例:若该参数仅微调5%,在耦合了下游径流量的估算偏差可能扩大至20%以上;若同时叠加输入降雨数据的10%误差,最终径流量的预测偏差甚至会超过35%,模型可靠性显著下降。不过近年深度学习和进化算法被应用后研究人员能在一定程度上捕捉这些复杂传播路径,但目前的方法计算成本还是很高而且解释性也不够,统计表明当下主流智能优化算法处理超50个自由参数的水文模型时计算时间往往要延长到好几天甚至好几周,所以深入了解参数不确定性传播机制对提高模型效率和准确性非常重要。

### 2.3 基于贝叶斯框架的参数不确定性量化

区域水文模型参数不确定性量化的系统化解方案由贝叶斯框架提供,此方法把先验知识和观测数据相结合能有效估计参数概率分布进而更全面地反映不确定性范围。具体来说,贝叶斯推理依据历史数据或者专家经验设定参数先验分布,再用似然函数更新这些分布得到后验分布,此法既能应对观测数据里的噪声情况又能自然而然整合多种来源信息。例如在某个案例研究里,融合遥感降水数据和地面监测站记录后贝叶斯框架使径流模拟的均方根误差成功降低18%。不过贝叶斯方法实施时要解决高维积分计算难题,这往往靠马尔可夫链蒙特卡罗(MCMC)这类数值技术,近年随着计算能力提高MCMC算法在水文领域应用渐渐普及但计算效率和结果精度之间仍需平衡。另外,贝叶斯框架让探索参数间的相关性很方便,这对理解复杂水文系统的内在规律很重要。

在这个基础上,把深度学习与进化算法相结合能进一步让贝叶斯框架的表现变得更强,因为深度学习善于从海量数据里提取特征且进化算法有全局搜索能力,两

者一起工作可迅速接近最优参数组合,就像在一项干旱区流域研究里,这种混合方法使参数校准时间减少了40%还保持住了较高的预测精度,不过要注意贝叶斯框架虽然理论上非常灵活但实际应用时得小心选择先验分布形式以免主观假设不对产生偏差,总的来说,基于贝叶斯框架的参数不确定性量化方法给提高水文模型的科学与实用性提供了新办法<sup>[3]</sup>。

## 2.4 模型预测结果的可靠性评估

确保水文模型于实际应用中发挥效用的重要环节是对模型预测结果可靠性的评估,以SWAT模型在某流域洪水预报中的应用为例,在该流域的汛期洪水预测中,研究人员通过集合模拟生成了50组不同参数组合的SWAT模型实例,不仅得出了洪峰流量的期望值(约1200m<sup>3</sup>/s),还给出了95%置信区间(1050-1380m<sup>3</sup>/s)。这一结果让防洪调度部门既明确了核心风险阈值,也提前储备了应对极端洪峰的物资,相比单一参数模型的预测,调度决策的安全性提升了30%以上。不过集合模拟计算需求量大,在高分辨率模型里尤其如此,单次模拟或许要花上几天甚至数十小时。所以这几年研究人员试图用代理模型技术加速该过程:代理模型通过对原始模型简化或者近似,能在确保一定精度的基础上大大削减计算时间,统计数据显示,用了代理模型之后,集合模拟总耗时平均降低60%还多。

除了在技术方面加以改进外,评估模型预测结果的可靠性时也不能忽视外部环境变化带来的影响,因为全球水循环格局正被气候变化和人类活动重塑,这给水文模型的长期适用性带来挑战,就像过去五年,城市化进程导致土地利用发生变化,使得部分流域蒸散发量增加20%且径流生成模式也被改变,所以评估模型可靠性时要充分考虑这些动态因素并定期更新模型参数才能保持其预测能力,而且引入社会经济数据也给模型评估提供了新角度,把用水需求、农业灌溉等人均为干扰纳入分析框架能更精准地描绘水资源供需平衡状况,最终,模型预测结果可靠不但能给洪水预警和干旱评估提供科学依据,还可助力制定更为精细的水资源管理策略。

## 3 结论

水资源管理、洪水预报和干旱评估都以水文模型为核心工具,所以其精确性和可靠性直接关乎区域水资源安全的科学决策。传统水文模型在参数优化时一直受观测数据少、空间异质性强、计算成本高的瓶颈制约,而本研究引进大数据技术构建起基于多源数据融合的区域水文参数优化框架,再配上智能优化算法和不确定性分析方法后明显提高了水文模型的预测性能。研究显示,大数据驱动的参数优化法可使模型预测误差平均减少

23.5%且能更精准地量化预测的不确定性范围,这一成果不但给气候变化背景下的水文模拟提供新技术路径,也给相关行业科学决策打下坚实基础。

近年来,遥感技术、地面监测网络和再分析数据迅猛发展使水文领域积攒起海量多源异构数据,这给突破传统模型局限带来了前所未有的机遇,例如全球陆地数据同化系统(GLDAS)过去五年产生的高分辨率再分析数据为区域水文建模提供重要输入支持且社会经济数据整合后进一步增强模型刻画人类活动影响的能力,在此情况下本研究提出一个将贝叶斯推理、深度学习和进化算法相结合的智能优化算法,该算法能高效搜索高维参数空间以解决传统方法在复杂流域应用效率低下的问题,并且基于集合模拟的不确定性分析方法会系统评估参数、输入数据和模型结构对预测结果的影响从而给提升模型可信度提供理论依据。

所提方法在不同气候与地形特征流域的适用性被案例研究进一步验证,从湿润的亚热带流域到干旱的内陆河流域,大数据驱动的优化框架都呈现出不错的鲁棒性和泛化能力,这意味着该方法不但能应对复杂自然条件,而且可适应多样化应用场景,使水资源管理实践有更靠谱的科学依据,不过需要注意的是,虽然研究成果技术优势明显,但要推广还得克服数据获取与处理成本高等现实挑战,所以以后的研究需要进一步探寻低成本、高效率的数据融合技术,促使大数据驱动的水文模型在全球得到更广泛的运用<sup>[4]</sup>。

本研究运用大数据技术,提供出区域水文模型参数优化与不确定性分析的创新解决方案,这补上了传统方法的技术短板且给应对气候变化和人类活动让水循环受影响加重的挑战以重要助力,其成果对水文学及相关领域技术进步的推动意义重大并且为水资源可持续管理目标的达成献上关键科学工具。

## 参考文献

- [1]王琰;丁永生;孔乔;王一帆;陈胤名;.分布式水文模型多参数优化与不确定性分析[J].大连海事大学学报,2020(02):98-108.
- [2]姜瑶;颜泽文;黎良辉;闫峰;熊吕阳;.灌区用水优化模型参数全局敏感性分析与不确定性优化[J].农业机械学报,2023(07):379-387.
- [3]武海喆;刘登峰;黄强;杨倩;林木;潘保柱;.不同潜在蒸发估算方法对黄土高原渭河径流模拟的影响研究[J].人民珠江,2020(10):40-47+69.
- [4]苟娇娇;缪驰远;徐宗学;段青云;.大尺度水文模型参数不确定性分析的挑战与综合研究框架[J].水科学进展,2022(02):165-173.