

基于新安江模型-双向长短期记忆网络与SHAP归因的阜平流域洪水预报残差修正研究

沈欣怡^{1,2}, 史韵琪², 张 轩^{1,2*}, 王高旭¹

(1. 南京水利科学研究院, 210029, 南京; 2. 河海大学, 210024, 南京)

摘 要: 针对华北半干旱半湿润流域洪水预报中传统概念模型洪峰低估、数据驱动模型物理约束不足的问题, 以海河流域大清河水系之阜平流域为研究对象, 构建了新安江模型(XAJ)与双向长短期记忆网络(BiLSTM)耦合的洪水预报残差修正模型, 并结合SHAP方法揭示模型补偿机理。以XAJ输出的预报流量、三层土壤含水量及径流分量等物理变量为输入, 采用递进式特征组合构建BiLSTM残差修正模型, 系统评估不同物理信息对预报效果的影响。结果表明: XAJ在阜平流域存在明显洪峰低估和时序偏差; 将XAJ输出的基准预报流量作为先验特征输入BiLSTM后, 混合模型整体拟合效果显著提升, 且物理边界约束可有效抑制纯数据驱动模型在极端条件下的数值失真; 进一步引入三层土壤含水量后, 模型洪峰捕捉能力最优, 平均洪峰相对误差控制在10%以内。SHAP归因分析表明, 土壤水分状态变量是驱动残差正向补偿的关键因子, 径流分量在本研究样本条件下未表现出稳定增益, 其附加信息可能因误差传播和信息冗余削弱模型泛化能力。研究表明, 基于关键物理状态约束的XAJ-BiLSTM混合框架能够有效提高北方复杂流域洪峰预报精度, 并为物理-数据融合水文模型的可解释构建提供参考。

关键词: 洪水预报; 物理-数据融合; 新安江模型; 双向长短期记忆网络; SHAP归因; 残差修正

Residual correction of flood forecasting in Fuping Catchment based on Xin'anjiang model, bidirectional long short-term memory network, and SHAP attribution

SHEN Xinyi^{1,2}, SHI Yunqi², ZHANG Xuan^{1,2*}, WANG Gaoxu¹

(1. Nanjing Hydraulic Research Institute, Nanjing 210029, China; 2. Hohai University, Nanjing 210024, China)

Abstract: To address the issues of flood forecasting for flood-peak underestimation in traditional conceptual models and insufficient physical constraints in data-driven models for semi-arid and semi-humid catchments in northern China, a residual correction model for flood forecasting was developed by coupling the Xin'anjiang model (XAJ) with a bidirectional long short-term memory network (BiLSTM) for the Fuping Catchment in the Daqing River subsystem of the Haihe River basin. The SHAP method was introduced to interpret the model's compensation mechanism. Physical variables derived from XAJ, including forecasted discharge, three-layer soil moisture, and runoff components, were used as inputs. A progressive feature combination approach was adopted to construct the BiLSTM residual correction model and to evaluate the effects of different physical information on forecasting performance. The results indicate that XAJ exhibits noticeable flood-peak underestimation and temporal deviation in the Fuping catchment. When the baseline forecasted discharge from XAJ is used as a prior input to BiLSTM, the hybrid model achieves significantly improved overall fitting performance, and the physically constrained framework effectively suppresses numerical distortions in purely data-driven models

收稿日期: 2026-03-25 修回日期: 2026-05-06

作者简介: 沈欣怡, 博士研究生, 主要研究方向为水文物理规律模拟及水文预报。

通信作者: 张轩, 工程师, 主要研究方向为水文物理规律模拟及水文预报。E-mail: xzhang@nhri.cn

基金项目: 国家重点研发计划(2023YFC3006501); 中央级公益性科研院所基本科研业务费专项资金重点基金项目(Y524008、Y525015)。

under extreme conditions. After further incorporating the three-layer soil moisture, the model captures flood peaks most accurately, with the mean relative error of flood peaks reduced to within 10%. SHAP attribution analysis reveals that soil moisture state variables are key drivers of positive residual compensation. Runoff component variables do not provide stable additional gains under the current sample conditions, and their inclusion may weaken the model's generalization ability due to error propagation and information redundancy. The research reveals that the proposed XAJ-BiLSTM hybrid framework, constrained by key physical state variables, effectively improves flood-peak forecasting accuracy in complex catchments in northern China and provides a reference for the interpretable construction of physics-data fusion hydrological models.

Keywords: flood forecasting; physics-data fusion; Xin'anjiang model; bidirectional long short-term memory network; SHAP attribution; residual correction

中图分类号: TV124 文献标识码: A 文章编号: 1000-1123(2026)10-0055-10

DOI: 10.3969/j.issn.1000-1123.2026.10.007

一、引言

近年,在全球气候变暖和大气环流调整的共同影响下,我国主雨带呈现明显北移趋势,北方地区极端强降雨事件频发,洪水防御形势日趋复杂。如何提高复杂下垫面条件下流域洪水预报的精度与可靠性,已成为当前水文预报研究的重要课题。

新安江模型(XAJ)作为我国洪水预报业务中广泛应用的概念水文模型之一,在众多流域取得了良好的应用效果。基于XAJ能够提供基准预报流量及土壤含水量、径流分量等内部状态变量,为后续残差修正提供必要的物理约束和可解释信息,本文选取XAJ作为基础物理模型。但是,华北部分流域受长期气候变化及人类活动影响,形成了典型的巨厚包气带结构,其产流过程与标准XAJ的理论假设存在一定偏差。特别是在短历时强降雨条件下,标准XAJ容易高估下渗作用,低估快速汇流量,从而导致洪峰偏低和峰现时间偏差。针对这一问题,已有研究尝试通过引入地下拦蓄水库、多层土壤结构等方式对模型进行物理机制改进,但在下垫面空间异质性较强的条件下,参数率定仍存在明显不确定性,模型结构误差尚难以完全消除。因此,在保留物理模型基本约束的基础上,利用数据驱动方法对其系统残差进行补偿,为提升北方复杂流域洪水预报精度提供了一条可行路径。

近年,深度学习方法在水文时间序列非线性映射建模中表现出较强优势。双向长短期记忆网络(bidirectional long short-term memory network, BiLSTM)能够同时利用前向和后向时序信息,对降雨—径流过程中的长时序依赖关系进行建模。随着

相关研究不断深入,深度学习在水文预报中的应用已由早期的纯数据驱动逐步发展到物理—数据融合阶段。Cui等将XAJ预报流量作为LSTM的附加输入,构建了残差修正框架,显著提升了多步预报精度;Jiang和Zhang通过引入时滞特征建模方法,有效刻画了流域响应过程中的记忆效应;张珂构建了融合注意力机制的卷积长短期记忆网络(ConvLSTM)模型,在有限样本条件下提高了中小流域汇流过程模拟的稳定性;Xiang等进一步将XAJ产汇流微分方程嵌入递归神经网络单元,构建了具有物理一致性的XAJRNN网络结构。

本文以海河流域大清水水系南支沙河上游的阜平流域为研究对象,构建XAJ-BiLSTM洪水预报残差修正模型,设计递进式特征组合方案,分析不同内部物理状态变量对模型修正效果的影响,并引入Shapley Additive Explanations(SHAP)可解释性分析方法揭示模型的补偿机制。研究结果可为复杂下垫面条件下概念水文模型的适应性改进,以及物理—数据融合洪水预报模型的构建提供参考。

二、研究区域与数据

1. 研究区域概况

阜平流域沙河源头位于山西省灵丘县孤山,阜平水文站控制断面以上的集水面积为2191 km²。受太行山脉阻隔影响,流域内降水的空间分布具有显著差异,位于背风坡山西省的区域降水相对匮乏,而位于迎风坡河北省的区域则是暴雨中心的高发地带,全流域多年平均年降水量约为600 mm。从下垫面条件来看,流域内森林面积占总面积的31.3%,总体植被覆盖率为

70%。尽管具有一定的植被覆盖,但流域多为山区,河道平均比降高达5.3‰,地形陡峭。这种特殊的下垫面与气候特征,使得阜平流域大多数洪水呈现“先超后蓄”的特性,地面径流对总径流过程影响较大,而壤中流及地下径流的影响相对微弱,遭遇暴雨时极易形成快速的坡面汇流。

2. 数据来源与处理

本研究的基础水文气象数据来源于水文年鉴,涵盖2016—2023年阜平水文站及其上游控制流域内的降水与流量观测序列。为全面捕捉流域降水的空间异质性,选取了空间布局均匀的阜平、庄旺、下关、冉庄、不老台及龙泉关等12个雨量站点。在数据预处理阶段,所有时序数据均经过严格的一致性检验。针对个别缺失的观测时段,采用相邻站点空间插值法进行延长与补齐,并结合水文规律对异常跳跃值进行“噪声”剔除。流域面雨量的计算采用面平均雨量法进行处理,以确保输入数据的可靠性。

为系统评估物理-数据混合模型在不同降雨情景下的修正效果与泛化能力,本研究从历史水文整编资料中选取了10场典型的暴雨洪水事件(见表1)。样本选择主要依据包括:①资料完整,降雨与流量过程连续可靠;②覆盖不同洪峰规模与总雨量等级,洪峰流量范围为29.8~3100.0 m³/s,累计雨量范围为52.2~300.8 mm;③兼顾年份分布与洪水类型差异,以保证模型训练与验证的代表性。在模型训练与验证的实验设计中,按照时间序列分布与洪水特征的代表性,将这10场洪水划分为两部分:①训练集,包含7场洪水。该数据集用于拟合XAJ和BiLSTM的权重参数与数据归一化特征。②验证集,包含3场洪水。该数据集未参与模型的任何

训练过程,专门用于检验修正模型在未知洪水情景下的预报精度与物理机制的捕捉能力。

三、模型构建与研究方法

1. XAJ的物理驱动机制与结构性偏差分析

在本文构建的物理-数据融合洪水预报框架中,XAJ模型作为基础物理模型,用于提供满足水量平衡约束的过程模拟结果及内部状态信息。在逐时演算过程中,XAJ不仅输出流域出口断面的基准预报流量(Q_{XAJ}),还可同步给出反映流域产汇流演化过程的关键物理状态变量。具体包括:由三层土壤水分结构计算得到的上层土壤含水量(WU)、下层土壤含水量(WL)及深层土壤含水量(WD),以及水源划分阶段得到的地表径流(RS)、壤中流(RI)和地下径流(RG)分量。

标准XAJ以“蓄满产流”为基本假定,在华北巨厚包气带流域应用时,对于短历时强降雨过程,该模型易将部分实际应形成快速径流的水量分配至深层下渗过程,从而造成洪峰偏低和峰现时间偏差等结构性误差。为此,本文在保留XAJ物理约束的基础上,提取其预报流量及内部状态变量作为BiLSTM残差修正模型的输入,以表征流域动态干湿状态、前期蓄水条件和径流组成特征,为后续非线性偏差修正提供物理信息支撑。

2. 基于BiLSTM的残差修正与机理补偿框架

为弥补传统物理模型在跨气候带应用中的局限性,本研究引入BiLSTM构建物理-数据融合的洪水预报残差修正框架。BiLSTM由正向和反向两个LSTM单元组成,能够同时利用序列的前向与后向时序信息,对长时序依赖关系和复杂非线性过程进行建模,因而

表1 阜平流域典型场次洪水特征统计

场次编号	样本用途	起始时间	终止时间	历时(h)	累计雨量(mm)	洪峰流量(m ³ /s)	洪水级别
20230729	训练	2023-07-29 12:00	2023-08-05 23:00	180	300.8	3100.0	特大洪水
20220806	训练	2022-08-06 20:00	2022-08-12 16:00	140	168.0	102.0	中等洪水
20210720	训练	2021-07-20 00:00	2021-08-05 19:00	403	114.9	1130.0	大洪水
20200716	训练	2020-07-16 14:00	2020-07-25 00:00	202	60.8	57.0	小洪水
20190804	验证	2019-08-04 08:00	2019-08-14 19:00	251	52.2	110.0	中等洪水
20180806	验证	2018-08-06 00:00	2018-08-18 23:00	311	104.4	66.0	小洪水
20180716	验证	2018-07-16 08:00	2018-07-31 08:00	360	119.7	130.0	中等洪水
20170720	训练	2017-07-20 01:00	2017-07-29 20:00	235	70.5	29.8	小洪水
20160724	训练	2016-07-24 00:00	2016-07-30 02:00	146	77.4	174.0	中等洪水
20160719	训练	2016-07-19 00:00	2016-07-23 23:00	119	191.9	560.4	大洪水

适用于降雨—径流演化过程的动态表征。结合山区流域汇流历时较短、洪水响应迅速的特点,本文将输入特征滑动窗口设定为 12 h,即利用当前时刻及其前 11 h ($t-11$ 至 t 时刻)的气象驱动信息与 XAJ 内部状态变量构成输入矩阵,从而对短期流量残差进行预测。

利用 BiLSTM 学习和修正传统物理模型的系统性残差。设 R_t 为时刻 t 的流量残差,其定义为实测流量 $Q_{obs,t}$ 与 XAJ 预报流量 $Q_{XAJ,t}$ 之间的差值。

$$R_t = Q_{obs,t} - Q_{XAJ,t} \quad (1)$$

在此基础上,利用 BiLSTM 建立气象驱动因子、物理模型内部状态变量与残差序列之间的映射关系,得到残差预测值 \hat{R}_t 。最终洪水预报结果 $Q_{final,t}$ 由 XAJ 输出的基准预报流量与 BiLSTM 预测残差叠加得到(为符合水文物理现实,最终流量被约束为非负值)。

$$Q_{final,t} = \max(Q_{XAJ,t} + \hat{R}_t, 0) \quad (2)$$

通过上述混合建模方式, XAJ 为洪水过程模拟提供满足水量平衡约束的物理基准, BiLSTM 则利用多源输入特征对传统概念模型难以准确描述的非线性偏差进行补偿。二者结合后,既保留了物理模型在水文过程模拟中的约束作用,又增强了模型对巨厚包气带调蓄作用、干旱下垫面入渗效应及极端降雨条件下异常响应过程的适应能力。

3. 物理—数据融合的递进式特征组合方案设计

为系统分析不同水文物理状态变量对深度学习残差修正效果的影响,并评估 XAJ 内部状态变量在北方半干旱半湿润流域洪水预报中的补偿作用,设计了 1 组物理基准模型和 6 组递进式 BiLSTM 输入方案(见表 2)。通过逐步引入不同类型的气象驱动信息和物理状态变量,对物理约束与数据驱动融合条件下的模型响应特征进行对比分析。

其中, M1 为物理基准模型,直接采用 XAJ 预报结果,用于量化传统概念模型在阜平流域的结构性误差。M2 为纯数据驱动方案, M2-a 仅输入降雨特征, M2-b 在此基础上增加历史实测流量,以检验缺乏物理约束条件下 BiLSTM 的洪水过程拟合能力。考虑到阜平流域汇流历时较短,本文选取 3 h、6 h 和 12 h 前期累计降雨 P_{3h} 、 P_{6h} 、 P_{12h} 作为输入特征,以表征不同时间尺度上的前期降雨影响。

M3-a 和 M3-b 为基础物理—数据融合方案,通过引入 XAJ 预报流量,考察物理模型输出对 BiLSTM 残差修正的引导作用。M3-c 在 M3-b 基础上进一步引入 WU 、 WL 和 WD ,以增强模型对流域干湿状态及下垫

面调蓄作用的表征能力。该设计有助于网络识别传统 XAJ 在强降雨条件下的系统偏差,并在本研究流域中主要表现为对洪峰低估的正向补偿。M3-d 则在 M3-c 基础上进一步加入地表径流、壤中流和地下径流量,用于分析冗余物理变量是否会引入信息“噪声”并影响模型泛化能力。

4. 基于 SHAP 的混合模型归因分析原理

尽管深度学习模型在洪水预报中具有较强的非线性拟合能力,但其复杂的网络结构也使模型预测过程缺乏直观的物理可解释性,这在一定程度上限制了其在实际防洪预报中的应用与推广。为提高物理—数据融合模型的可解释性,本文引入 SHAP 方法,对各输入特征在模型预测过程中的贡献进行定量分析。

SHAP 方法由 Lundberg 和 Lee 基于合作博弈论中的 Shapley 值提出,是一种统一的加性特征归因方法。其基本思想是将模型对单一样本的预测结果表示为各输入特征贡献值的线性组合。对于任意给定样本,其解释模型可表示为:

$$g(z') = \phi_0 + \sum_{j=1}^M \phi_j z'_j \quad (3)$$

式中, $g(z')$ 为解释模型; M 为输入特征的最大维度(即参与博弈的特征总数); z'_j 表示第 j 个特征是否被观察到(1 为存在, 0 为缺失); ϕ_0 为模型在没有任何特征输入时的基准预测值(通常为所有样本的预测均值); ϕ_j 为第 j 个特征的 SHAP 归因值,表征该特征对最终预测结果的贡献大小与方向。

为保证特征归因结果满足一致性和局部准确性, SHAP 值按照 Shapley 值的边际贡献分配原则进行计算,其表达式为:

$$\phi_j = \sum_{S \subseteq N \setminus \{j\}} \frac{|S|! (|N| - |S| - 1)!}{|N|!} [v(S \cup \{j\}) - v(S)] \quad (4)$$

式中, N 为所有输入特征的集合; S 为剔除第 j 个特征后剩余特征的所有可能子集; $|N|$ 与 $|S|$ 分别表示集合 N

表 2 多源状态变量输入与混合预报方案配置

方案名称	模型属性	基础输入	增量信息
M1	物理模型	XAJ 输出流量	—
M2-a	纯数据驱动	$P, P_{3h}, P_{6h}, P_{12h}$	—
M2-b	纯数据驱动	$P, P_{3h}, P_{6h}, P_{12h}$	Q_{obs}
M3-a	物理—数据融合	$P, P_{3h}, P_{6h}, P_{12h}$	Q_{XAJ}
M3-b	物理—数据融合	$P, P_{3h}, P_{6h}, P_{12h}$	Q_{obs}, Q_{XAJ}
M3-c	物理—数据融合	M3-b	WU, WL, WD
M3-d	物理—数据融合	M3-c	RS, RI, RG

和 S 中包含的特征数量; $v(s)$ 为仅在子集 S 的特征参与下模型输出的预测期望值; $[v(s \cup \{j\}) - v(s)]$ 表示在特征子集 S 的基础上,加入第 j 个特征后所带来的边际预测增量。

在本文中,SHAP方法主要用于分析不同物理状态变量和气象驱动因子对预测结果残差的影响方向及相对重要性,从而揭示BiLSTM在残差修正过程中对关键水文物理信息的响应特征,并为后续混合模型的机理解释与特征筛选提供依据。

5. 模型评价指标体系

为评价基准物理模型及各混合预报方案的模拟性能,本文从整体过程拟合、洪量平衡和极值捕捉三个方面构建模型评价指标体系。具体选取纳什效率系数(NSE)、均方根误差($RMSE$)、洪峰流量相对误差、洪量相对误差及峰现时间绝对误差,对不同方案的预报效果进行综合比较(见图1)。

在整体过程拟合度方面,选取 NSE 与 $RMSE$ 作为核心评价指标。其中, NSE 用于衡量模拟流量过程与实测流量过程之间的整体吻合程度, $RMSE$ 用于反映整个洪水过程中模拟值与实测值之间的平均偏差。其计算公式分别为:

$$NSE = 1 - \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^n (Q_{obs,i} - Q_{obs})^2} \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2} \quad (6)$$

式中, n 为洪水序列的时段总数; $Q_{obs,i}$ 与 $Q_{sim,i}$ 分别为第 i 时刻的实测流量与模拟流量, m^3/s 。

在洪量平衡和极值刻画方面,进一步选取洪峰流量相对误差(RE_p)、洪量相对误差(RE_v)及峰现时间绝对误差(ΔT)作为补充评价指标。其中,洪峰流量相对误差和峰现时间绝对误差反映模型对洪峰量级及其出现时刻的预测能力,洪量相对误差用于检验模型对洪水总体水量的模拟精度。考虑到山区流域洪水预警更强调洪峰预报的准确性和时效性,洪峰相关指标在模型方案优选中具有更高权重。通过上述指标的综合分析,可对不同方案在过程拟合能力、极值捕捉能力及水量平衡表现方面进行系统评价。

四、结果与分析

1. 传统概念水文模型在半干旱半湿润流域的适应性分析

为分析传统概念水文模型在华北半干旱半湿润流

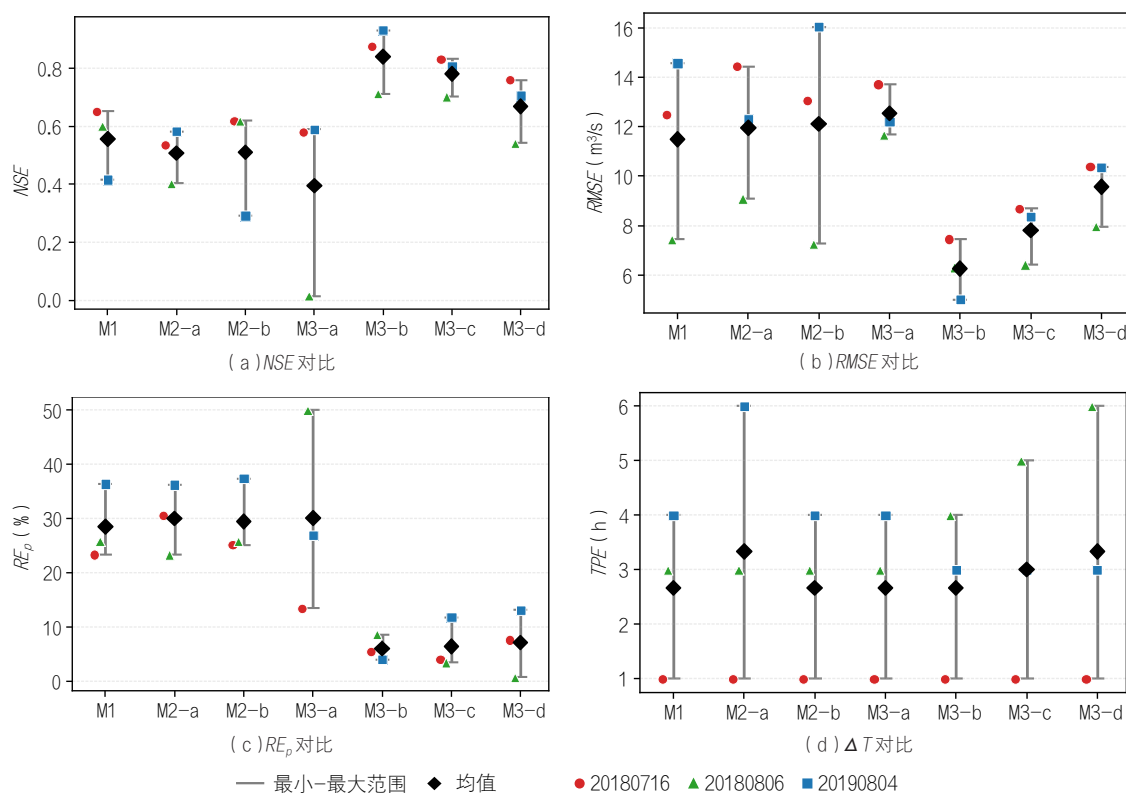


图1 各混合预报方案的精度指标综合对比

域中的适用性,本文选取基准模型M1对阜平流域验证期洪水过程进行了模拟评价。结合表3的定量指标与图2的流量过程线可知,M1在3场验证洪水中的平均NSE仅为0.56,平均RMSE为11.49 m³/s,且不同场

次间模拟精度波动较大。总体来看,基准模型在整个过程拟合和极值刻画方面均存在一定局限,尤其在洪峰模拟上表现出较为明显的系统性低估特征。

以20180716和20190804两场洪水为例,实测洪

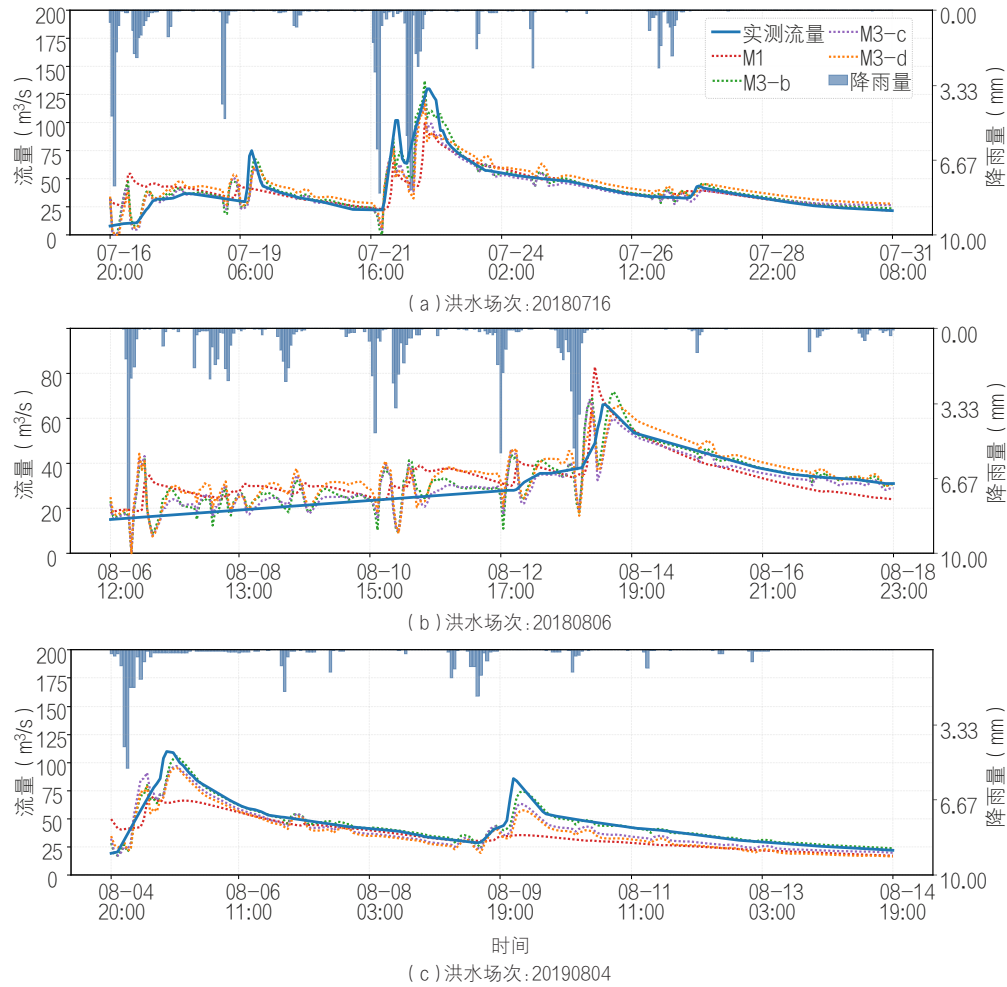


图2 基准模型与核心混合方案典型场次洪水模拟过程线对比

表3 阜平流域新安江模型洪水模拟结果

时期	洪水编号	评价指标				
		NSE	RMSE (m ³ /s)	RE _p (%)	ΔT (h)	RE _v (%)
训练期	20230729	0.938	152.672	4.72	0	21.48
	20220806	0.640	14.706	11.66	7	0.65
	20210720	0.706	91.539	42.11	0	14.84
	20200716	0.219	9.220	12.90	7	11.66
	20170720	-0.769	5.427	5.45	10	39.41
	20160724	0.188	175.034	77.04	1	30.24
	20160719	0.574	78.744	4.95	7	10.43
验证期	20190804	0.481	14.375	36.40	4	21.24
	20180806	0.618	7.393	25.81	3	8.70
	20180716	0.678	12.300	23.35	1	0.92

峰分别为 $130.0 \text{ m}^3/\text{s}$ 和 $110.0 \text{ m}^3/\text{s}$, M1 的洪峰流量相对误差分别达到 23.35% 和 36.40%。从流量过程线对比结果看, 实测洪水呈现出典型的北方山区洪水“陡涨陡落”特征, 而模拟过程则相对平缓, 峰值偏低, 且起涨和峰现时间普遍提前, 难以满足该区域防洪预警对洪峰量级和峰现时刻的预报要求。特别是在 20170720 场次洪水中, M1 的 NSE 为 -0.769 , 说明模型对该场洪水过程的模拟效果较差。其主要原因在于明显的相位偏差: 模型对初期降雨响应过快, 导致模拟洪峰提前出现, 而流域受巨厚包气带调蓄作用影响, 实测洪峰具有一定滞后性, 从而使时间错位被 NSE 指标进一步放大。

上述“洪峰偏低、过程平缓、峰现提前”的误差特征, 反映出传统 XAJ 结构假定与研究区实际产汇流机制之间存在一定偏差。模型在数值上表现为总径流量偏小、洪峰偏低及过程线平滑化。该结果表明, 传统概念模型在该类北方流域中存在结构性适应不足问题, 也说明有必要引入能够表征复杂非线性响应关系的数据驱动方法, 对其系统性残差进行进一步修正。

2. 物理-数据融合混合模型修正效果与方案对比

为开展残差修正, 构建基于 BiLSTM 的物理-数据融合网络。模型由输入层、2 层 BiLSTM 隐藏层和全连接输出层组成, 其中每层隐藏层包含 256 个神经元, 用于输出流量残差预测值。在数据预处理阶段, 采用 MinMaxScaler 对多源输入特征进行归一化处理, 以减小不同量纲变量对模型训练的影响。结合前文设定, 输入时间步长取 12 h, 即利用当前时刻及前 11 h 的输入信息预测未来 1 h 的流量残差。

在模型训练过程中, 采用均方差作为损失函数, 并选用 Adam 优化器进行参数更新, 初始学习率设定为 0.0005。考虑到训练样本数量有限, 为降低过拟合风险, 训练最大迭代轮次设为 100, 并引入早停机制: 当验证集损失在连续 10 轮训练中未进一步下降时, 停止模型训练。

结合图 1、图 2 和表 4 可知, 纯数据驱动方案 (M2-a、M2-b) 相较于基准物理模型 M1 在部分指标上有所改善, 但整体提升幅度有限, 且不同验证场次间精度波动仍较明显。其中, M2-a 的平均 NSE 为 0.48, 平均 $RMSE$ 为 $12.54 \text{ m}^3/\text{s}$; 在引入历史实测流量后, M2-b 的平均 NSE 提高至 0.53, 平均 $RMSE$ 降至 $11.64 \text{ m}^3/\text{s}$, 说明历史流量信息能够在一定程度上增强模型对洪水过程演进的刻画能力。然而, 在缺乏物理约束条件下, 纯数据驱动模型对极端洪水过程的适应性仍然不足, 其平均 RE_p

表 4 不同方案的洪水模拟精度评价指标对比

方案编号	NSE	$RMSE$ (m^3/s)	RE_p (%)	RE_v (%)
M1	0.56	11.50	28.52	10.66
M2-a	0.48	12.54	30.52	11.42
M2-b	0.53	11.64	29.46	10.12
M3-a	0.49	11.96	30.16	13.45
M3-b	0.74	8.32	15.65	7.27
M3-c	0.70	8.48	8.60	11.83
M3-d	0.57	10.61	8.13	20.74

分别为 30.52% 和 29.46%, 并未显著优于基准模型。

当进一步引入 XAJ 预报流量形成基础物理-数据融合方案后, 模型性能得到明显改善。特别是 M3-b 方案, 在整体过程拟合方面表现最佳, 平均 NSE 提高至 0.74, 平均 $RMSE$ 降至 $8.32 \text{ m}^3/\text{s}$, 平均 RE_p 也由 M1 的 28.52% 降低至 15.65%。这表明, XAJ 提供的基准预报流量能够为 BiLSTM 残差修正提供有效的物理边界信息, 使网络在保持整体过程合理性的基础上进一步修正传统概念模型的系统偏差。与纯数据驱动方案相比, M3-b 在三场验证洪水中表现更为稳定, 说明物理约束的引入有助于提升模型在未知洪水过程中的泛化能力。

在递进式特征组合中, 土壤水分状态变量的引入进一步改善了模型对极值过程的识别能力。M3-c 在 M3-b 的基础上增加 3 层土壤含水量后, 平均 NSE 由 0.74 略降至 0.70, 平均 $RMSE$ 由 $8.32 \text{ m}^3/\text{s}$ 小幅增至 $8.48 \text{ m}^3/\text{s}$, 但其平均 RE_p 进一步由 15.65% 显著降低至 8.60%。这说明, 尽管 M3-c 在整体过程拟合上略逊于 M3-b, 但在洪峰量级捕捉方面表现更优。考虑到山区流域防洪预警更强调洪峰预报的准确性和时效性, 极值指标在模型优选中具有更高权重, 因此本文将 M3-c 作为核心优选方案。从物理意义上看, 3 层土壤含水量变量为网络提供了流域动态干湿状态信息, 使模型能够更准确地识别强降雨条件下传统 XAJ 可能出现的洪峰低估情形, 并通过输出正向残差对其进行补偿。

相比之下, M3-d 在 M3-c 的基础上进一步加入地表径流、壤中流和地下径流分量后, 模型性能并未继续提升。由表 4 可见, M3-d 的平均 RE_p 为 8.13%, 与 M3-c 接近, 但其平均 RE_v 增至 20.74%, 明显高于其他方案, 平均 NSE 也回落至 0.57。该结果表明, 物理变量的引入并非越多越好。对于由传统模型进一步分解得到且本身存在较大不确定性的径流分量, 盲目纳入输入特征可能会引入冗余信息, 干扰网络对降雨过程与

土壤干湿状态等关键控制因子的学习,进而削弱模型的整体泛化能力。

3. 基于SHAP归因的物理-数据融合模型机制分析

为进一步揭示深度学习网络对传统概念水文模型结构性偏差的补偿机制,本文引入SHAP方法,对各混合预报方案的输入特征进行了归因分析,以量化不同物理状态变量和气象驱动因子对残差预测结果的贡献。

由不同递进式方案的SHAP特征重要性摘要图可知,在基础混合模型M3-b(见图3a)中,XAJ预报流量与输入滑动窗口内的历史实测流量序列对残差预测的贡献占主导地位。说明该方案主要依赖流量序列的自回归信息以及物理模型提供的基准预报值实现残差修正。由于缺乏刻画流域干湿状态的内部物理变量,M3-b虽然能够改善整体过程拟合效果,但对引起预报偏差的关键水文过程识别相对有限,因此在极值捕捉方面仍存在一定不足。

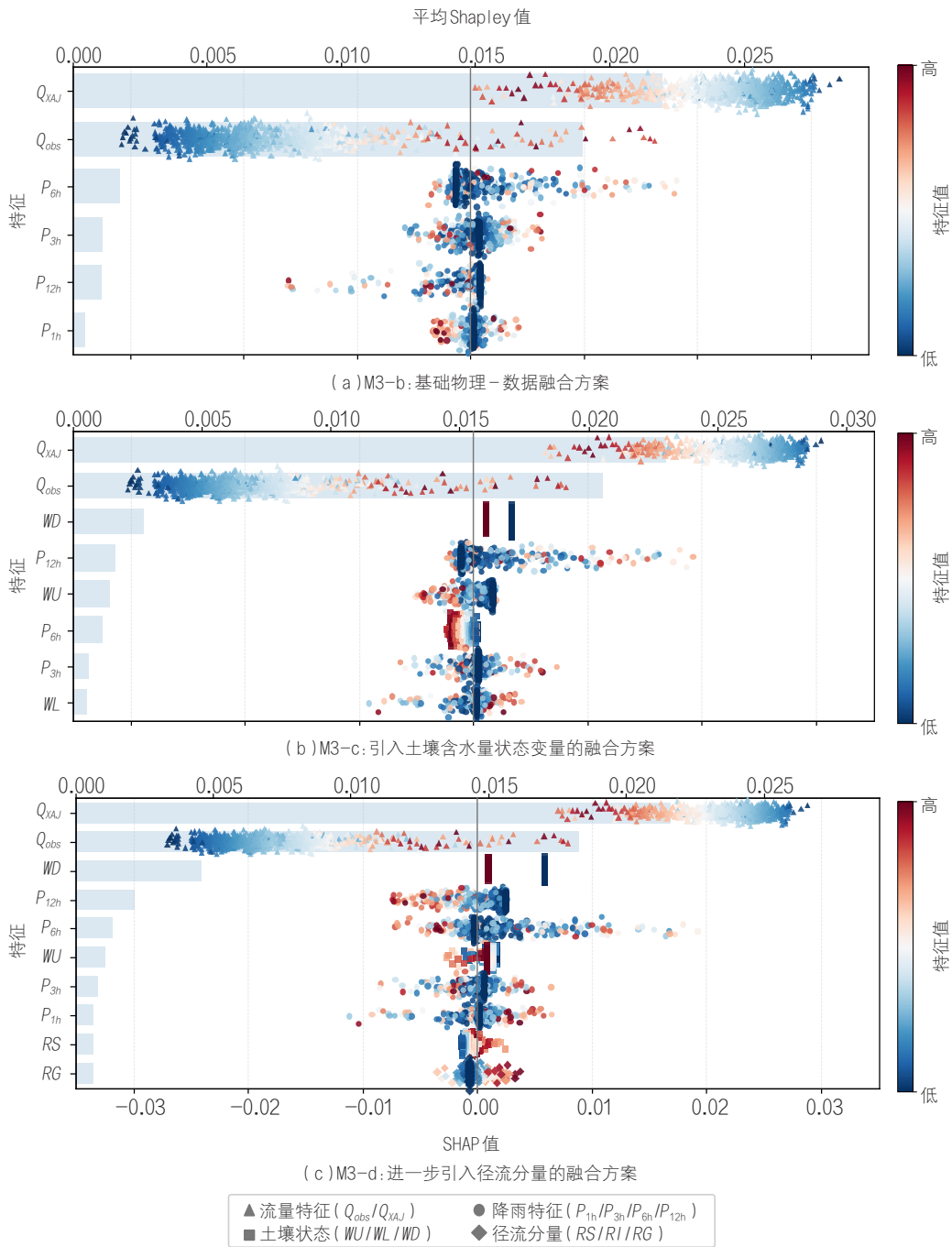


图3 基于SHAP归因的混合模型内部状态变量贡献度分布

在优选方案M3-c(见图3b)中,引入3层土壤含水量后,特征重要性分布发生明显变化。其中, WD 和 WL 的重要性显著提高,并超过部分短历时降雨特征。结合SHAP值分布可知,在高强度降雨条件下,当深层土壤 WD 含水量处于较低水平时,模型更倾向于输出正向残差。该现象表明,在具有巨厚包气带特征的北方流域中,标准XAJ受“蓄满产流”假定约束,可能将部分本应形成快速径流的水量分配至深层下渗和蓄存过程,从而造成洪峰偏低。引入关键土壤水分状态变量后,BiLSTM能够更有效地识别这种“强降雨-低深层含水量”组合所对应的偏差情景,并通过正向残差补偿修正传统模型的洪峰低估问题。由此说明,土壤水分状态变量不仅提高了模型对极值过程的识别能力,也增强了混合模型对北方流域实际产流响应特征的表征能力。

相比之下,在包含附加径流分量的M3-d方案(见图3c)中,新增的 RS 、 RI 和 RG 的SHAP值大多分布在0附近,特征重要性排序相对靠后,表明其对残差预测的增量贡献有限。考虑到北方复杂流域径流成分划分本身具有较高不确定性,这些由传统模型进一步分解得到的径流变量可能挟带系统误差信息。将其直接作为输入未表现出稳定的补充约束作用,反而可能形成信息冗余,进而放大洪量误差并削弱整体泛化能力。该认识目前仅针对本研究流域与样本条件成立,结论的广泛性仍需更多北方流域进一步验证。

五、结 论

针对传统概念水文模型在华北半干旱半湿润流域中的适用性不足问题,本文提出了一种耦合XAJ、BiLSTM与SHAP归因分析的洪水预报残差修正方法,并在阜平流域进行验证。结果表明:

①标准XAJ在该流域存在洪峰偏低、过程线平缓及峰现时间提前等结构性偏差;将XAJ输出的基准预报流量作为BiLSTM的输入特征后,物理-数据融合模型整体性能明显提升,说明物理模型输出的基准过程能够为深度学习残差修正提供有效约束。

②在递进式特征组合中,融合3层土壤含水量的M3-c方案表现出最优的洪峰捕捉能力,平均洪峰相对误差控制在10%以内,表明关键土壤水分状态变量能够有效改善传统XAJ对洪峰的低估问题。

③SHAP归因分析进一步表明,历史观测流量、XAJ基准流量和土壤水分状态是驱动残差补偿的重要因素,而额外引入的径流分量在本研究流域、样本和参

数条件下未表现出稳定增益,可能因误差传播和信息冗余影响模型性能,其广泛适用性仍有待多流域验证。

总体而言,基于关键物理状态约束的XAJ-BiLSTM混合框架能够有效提高北方复杂流域的洪水预报精度,并为物理-数据融合水文模型的可解释构建提供参考。

参考文献:

- [1] 丁一汇,任国玉,石广玉,等.气候变化国家评估报告(I):中国气候变化的历史和未来趋势[J].气候变化研究进展,2006(1):3-8+50.
- [2] 马强,涂泽辉,李邦森,等.雨带北移影响下大河北支防洪情势[J].南水北调与水利科技(中英文),2025,23(5):1127-1133.
- [3] ZHAO L, WANG J, XIAO Z, et al. Solar 11-year cycle-modulated north-south contrasting patterns of summer precipitation in China[J]. Journal of Climate, 2025, 38(14):3277-3294.
- [4] 刘志雨.雨水情监测预报“三道防线”耦合贯通建设应用与探讨[J].中国水利,2025(10):1-7.
- [5] 国家防灾减灾救灾委员会办公室.应急管理部发布2025年上半年全国自然灾害情况[J].安全与健康,2025(8):58-59.
- [6] REN-JUN Z. The Xin'anjiang model applied in China[J]. Journal of Hydrology, 1992, 135(1):371-381.
- [7] 仲志余,张建云,王焰新,等.华北地区深层地下水回补修复试点研究与思考[J].中国水利,2025(15):1-5.
- [8] 姚成,李致家,张珂,等.基于栅格型新安江模型的中小河流精细化洪水预报[J].河海大学学报(自然科学版),2021,49(1):19-25.
- [9] 邵景力,白国营,刘翠珠,等.我国地下水管理面临的问题与对策——兼谈地下水“双控”管理[J].水文地质工程地质,2023,50(5):1-9.
- [10] 晋华.双超式产流模型的理论及应用研究[D].北京:中国地质大学(北京),2006.
- [11] 魏玉涛.基于新安江-海河模型的漳河流域洪水预报和调度研究[D].济南:济南大学,2022.
- [12] 钟粟,姚成,李致家,等.应用新安江-海河模型研究下垫面变化对设计洪水的影响[J].湖泊科学,2015,27(5):975-982.
- [13] 李致家,张心愿,白云鹏,等.海河“23·7”流域性特大洪水复盘模拟[J].河海大学学报(自然科学版),2024,52(5):13-19+92.

- [14] 李巧玲, 李旻喆, 李致家, 等. 新安江-海河模型参数物理意义分析及应用[J]. 河海大学学报(自然科学版), 2025, 53(1):1-9.
- [15] CHEN X, ZHANG K, LUO Y, et al. A distributed hydrological model for semi-humid watersheds with a thick unsaturated zone under strong anthropogenic impacts: A case study in Haihe River Basin[J]. Journal of Hydrology, 2023, 623: 129765.
- [16] KRATZERT F, KLOTZ D, BRENNER C, et al. Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks[J]. Hydrology and Earth System Sciences, 2018, 22(11): 6005-6022.
- [17] KRATZERT F, KLOTZ D, SHALEV G, et al. Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets[J]. Hydrology and Earth System Sciences, 2019, 23: 5089-5110.
- [18] 瞿思敏, 余裕, 方正, 等. 机器学习模型与物理机制模型在长诏水库流域实时洪水预报中的比较研究[J]. 水资源保护, 2025, 41(5): 73-78+88.
- [19] 王俊, 程海云, 郭生练, 等. 智慧流域水文预报技术研究进展与开发前景[J]. 人民长江, 2023, 54(8): 1-8+59.
- [20] 张建云, 谢康, 刘艳丽, 等. 融合物理机制的机器学习水文模型研究进展[J]. 人民长江, 2025, 56(10): 37-46.
- [21] LI W, LIU C, XU Y, et al. An interpretable hybrid deep learning model for flood forecasting based on Transformer and LSTM[J]. Journal of Hydrology: Regional Studies, 2024, 54: 101873.
- [22] CUI Z, ZHOU Y, GUO S, et al. A novel hybrid XAJ-LSTM model for multi-step-ahead flood forecasting[J]. Hydrology Research, 2021, 52(6): 1436-1454.
- [23] JIANG H, ZHANG C. A Hybrid XAJ-LSTM-TFM Model for improved runoff simulation in the Poyang Lake basin: integrating physical processes with temporal and lag feature learning[J]. Water, 2025, 17(14): 2146.
- [24] 张珂, 刘杰, 王宇昊, 等. 结合注意力机制的ConvLSTM与新安江模型相融合的混合水文模型[J]. 水资源保护, 2026, 42(1): 137-143+151.
- [25] XIANG X, GUO S, LI C, et al. An explainable deep learning model based on hydrological principles for flood simulation and forecasting[J]. Hydrol. Earth Syst. Sci., 2025, 29: 7217-7239.
- [26] 陈暘, 雷晓辉, 蒋云钟, 等. 分布式水文模型Easy-DHM在海河阜平流域的应用[J]. 南水北调与水利科技, 2010, 8(4): 111-114+132.
- [27] SCHUSTER M, PALIWAL K K. Bidirectional recurrent neural networks[J]. IEEE transactions on Signal Processing, 1997, 45(11): 2673-2681.
- [28] LUNDBERG S M, LEE S I. A unified approach to interpreting model predictions[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2017: 4768-4777.
- [29] XU Y, LIN K, HU C, et al. Uncovering the dynamic drivers of floods through interpretable deep learning[J]. Earth's Future, 2024, 12(10): e2024EF004751.

责任编辑 吕彩霞