

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

张 珂^{1,2,3,4,5}, 刘 杰^{1,2}, 王宇昊^{1,2}, 申笑萱^{1,2}, 齐千嘉^{1,2}

(1. 河海大学水灾害防御全国重点实验室; 2. 河海大学水文水资源学院; 3. 河海大学长江保护与绿色发展研究院;
4. 中国气象局水文气象重点开放实验室; 5. 水利部水利大数据重点实验室)

摘要:为提高新安江模型(XAJ)在中小流域汇流计算中的精度,构建了结合注意力机制的卷积长短期记忆神经网络(ConvLSTM),用于替代XAJ中的汇流模块,从而建立了结合物理机制与机器学习技术的混合水文模型XAJ-ACL,基于呈村流域实测数据,探究了XAJ-ACL在中小流域有限样本容量条件下的性能,并分别采用ConvLSTM和传统LSTM替代XAJ汇流模块,构建了混合水文模型XAJ-CL和XAJ-LSTM进行对比分析。结果表明:在呈村流域径流模拟中,XAJ-ACL的模拟精度优于XAJ,测试期XAJ-ACL的纳什效率系数为0.85,相关系数为0.93,均高于XAJ;在3组小容量样本训练中,测试期XAJ-ACL的平均纳什效率系数分别为0.847、0.832和0.808,均高于XAJ-CL和XAJ-LSTM,且模拟结果表现出更好的稳定性;与XAJ相比,XAJ-ACL显著提升了有限资料条件下对中小流域汇流过程非线性规律的模拟能力。

关键词:新安江模型;注意力机制;卷积长短期记忆神经网络;混合水文模型;汇流过程;径流模拟;呈村流域

A hybrid hydrological model integrating attention-based convolutional long short-term memory neural network with Xin'anjiang model//Zhang Ke^{1,2,3,4,5}, Liu Jie^{1,2}, Wang Yuhao^{1,2}, Shen Xiaoxuan^{1,2}, Qi Qianjia^{1,2}(1. *The National Key Laboratory of Water Disaster Prevention, Hohai University*; 2. *College of Hydrology and Water Resources, Hohai University*; 3. *Yangtze Institute for Conservation and Development, Hohai University*; 4. *China Meteorological Administration Hydro-Meteorology Key Laboratory*; 5. *Key Laboratory of Water Big Data Technology of Ministry of Water Resources*)

Abstract: To improve the accuracy of the Xin'anjiang model (XAJ) in confluence calculation of small and medium-sized watersheds, a convolutional long short-term memory neural network (ConvLSTM) integrating attention mechanism was constructed to replace the confluence module in XAJ, thereby establishing a hybrid hydrological model XAJ-ACL that combined physical mechanisms and machine learning techniques. The performance of XAJ-ACL under the condition of limited sample capacity in small and medium-sized watersheds was explored based on the measured data of the Chengcun Watershed. To further evaluated the effect of XAJ-ACL, ConvLSTM and the traditional LSTM were respectively adopted to replace the confluence module in XAJ, and two additional hybrid hydrological models, XAJ-CL and XAJ-LSTM, were constructed for comparative analysis with XAJ-ACL. The results show that in runoff simulation of the Chengcun Watershed, the simulation accuracy of XAJ-ACL is superior to that of XAJ. During the testing period, the Nash-Sutcliffe efficiency coefficient (NSE) of XAJ-ACL is 0.85, and the correlation coefficient is 0.93, both of which are higher than those of XAJ. In training of three sets of small-capacity samples, the average NSE values of XAJ-ACL during the testing period were 0.847, 0.832, and 0.808, respectively, all higher than those of XAJ-CL and XAJ-LSTM, and the simulation results showed better stability. Compared with XAJ, XAJ-ACL significantly improves the ability for simulation of nonlinear laws of confluence processes in small and medium-sized watersheds under limited data conditions.

Key words: Xin'anjiang model; attention mechanism; convolutional long short-term memory neural network; hybrid hydrological model; confluence process; runoff simulation; the Chengcun Watershed

洪水是我国最严重的自然灾害之一。在一般年份,全国超过 70% 的洪水灾害损失发生在中小流域^[1],中小流域洪水的突发性和过程不稳定性等特点增加了防汛工作的难度^[2-3]。同时,人为工程因素和自然气候因素的影响也给中小流域内的径流预报带来了挑战^[4]。

新安江模型(XAJ)在我国湿润、半湿润地区洪水径流模拟与预报中有良好的表现^[5-7]。在中小流域应用中,XAJ 的产流模块精度较高^[8],但其汇流计算存在不足。一方面,XAJ 的汇流过程对流域非线性规律的表达能力较弱。研究表明,流域的非线性效应通常随着流域面积的减小而增强^[9],使得中小流域汇流过程有较强的非线性特征,XAJ 在汇流计算中采用线性水库法处理坡面汇流,并使用 Muskingum 法进行河道汇流计算,这些方法本身对非线性规律的刻画存在局限^[10],在中小流域应用中容易引入较大误差。此外,目前采用将各子流域独立演算的洪水过程进行线性叠加的方法未能考虑洪水波之间的非线性相互作用,且 XAJ 汇流模块参数敏感性较高,参数率定难度较大,在工程活动影响显著和观测数据有限的情况下,这些问题尤为突出^[2]。为了对中小流域径流进行更高精度的模拟与预报,有必要对 XAJ 的汇流模块进行改进。

近年来,深度学习方法已被广泛应用于复杂非线性规律的学习与表达^[11-14],使用该方法构建的数据驱动模型输入参数较少且容易率定。在洪水预报领域,卷积长短期记忆神经网络(ConvLSTM)在捕捉时空依赖特征方面有出色表现^[15-16]。与传统长短期记忆神经网络(LSTM)相比,ConvLSTM 在学习时间依赖关系的同时,还可以在空间维度上捕捉不同区域局部依赖关系^[17-18],在极端洪水预报中也有出色表现。Ha 等^[19]使用 ConvLSTM 对长江流域大同站极端洪水预报结果表明,模型在测试期的平均决定系数(R^2)达到 0.86,远优于传统方法。因此,使用 ConvLSTM 替代 XAJ 中的汇流模块进行径流模拟与预报有较大的潜力。

然而,ConvLSTM 结构复杂,需要训练大量样本来学习规律和调整模型内部节点参数^[20]。而中小流域往往观测数据有限甚至缺少数据,增加了 ConvLSTM 训练的难度。注意力机制是一种模仿人类在处理信息中注意力分配过程的技术,通过关注关键信息提高模型处理复杂序列的能力^[21]。Zhang 等^[22-23]在使用注意力机制改进 LSTM 模型进行洪水预报的研究中指出,在相同样本容量情况下,注意力机制可以更好地关注时间序列中的关键信息,从而提高模型整体预报精度。但在径流模拟中,注意力

机制模块采用对时间、空间分别计算注意力的方法,会导致时空关联信息的丢失,不利于 ConvLSTM 的训练。因此,有必要构建一种针对 ConvLSTM 计算特点的注意力机制算法。

考虑到中小流域的汇流特征,本文构建了一种适用于 ConvLSTM 的注意力机制,建立了结合注意力机制的卷积长短期记忆神经网络(attention-ConvLSTM,ACL),以替代新安江模型汇流模块,从而构建了结合物理机制与机器学习技术的混合水文模型 XAJ-ACL,并基于呈村流域实测数据对模型模拟效果与整体性能进行评估。

1 XAJ-ACL

1.1 模型构建

XAJ-ACL 与 XAJ 均使用 3 层蒸散发模型进行蒸散发和产流计算,模型计算参数一致。但在汇流计算中,XAJ-ACL 使用数据驱动的 ACL 替代 XAJ 的汇流模块,其中,数据驱动模型的产流数据来源于单独率定的 XAJ,汇流神经网络的输入为 XAJ 的产流结果,输出为出口站流量。XAJ-ACL 计算过程主要包括注意力计算、特征映射和 ConvLSTM 计算 3 部分。XAJ 与 XAJ-ACL 结构如图 1 所示。

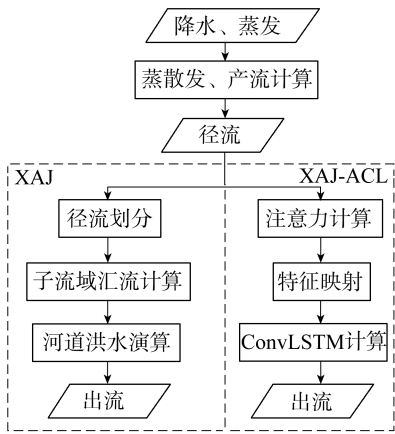


图 1 XAJ 与 XAJ-ACL 结构

Fig. 1 Structures of XAJ and XAJ-ACL

1.1.1 注意力机制

本文构建的注意力机制实现了对多个子流域产流输入过程的注意力捕捉,可以帮助 ConvLSTM 在小样本情况下完成学习。每个样本中 m 个子流域序列长度为 n 的产流过程经卷积计算、Sigmoid 函数激活^[20-21]并经 Softmax 函数^[20-21]概率化后与原输入产流过程进行 Hadamard 求积^[24],输出的特征张量被用于下一阶段的计算。该结构通过 Sigmoid 函数识别洪水事件,使用 Softmax 函数对权重矩阵进行概率化计算,最后通过 Hadamard 积的形式将权重赋予最初的输入特征,实现洪水注意力的计算。

1.1.2 特征映射

ConvLSTM 可以实现对不同子流域产流依赖关系的学习,但由于卷积计算的特点,这种关系仅是局部的。而产流从子流域汇入主河道后,每个子流域洪水波间都存在相互作用。为此,本文提出了一种特征映射方式,通过调整输入 ConvLSTM 的特征张量结构,实现任意两个子流域间相关关系的学习。特征映射的具体公式为:

p_{ijk} = \begin{cases} a_{ik} & j \text{ 为奇数} \\ a_{i\lambda} & j \text{ 为偶数} \end{cases} \tag{1}

其中 $\lambda = (j/2 + k) \bmod (m)$
式中: p_{ijk} 为特征映射结果; a 为输入特征映射层的特征,即注意力计算结果; i 为样本索引; j 为时间索引; k, λ 为空间索引; $\bmod(\cdot)$ 为取模计算符号。特征映射过程经式(1)对子流域顺序进行重新排序与组合,并通过窗口滑动形式将单一子流域的输入映射为子流域输入的组合,以提取子流域汇流的交互作用。

1.1.3 ConvLSTM

ConvLSTM 将卷积过程嵌入每个 LSTM 计算单元中,提高了模型对具有时空结构序列数据的处理能力,并实现了对时间和空间数据特征的同步捕捉^[17,25]。与 LSTM 计算单元不同,ConvLSTM 计算单元对来自上一单元的隐藏状态与输入数据进行卷积计算,这也是 ConvLSTM 对时空数据有更强学习能力的关键。计算时,不同时间步的数据输入到不同计算单元中,每个计算单元对单元内计算结果进行选择“记忆”与“遗忘”,并按时间顺序传递给下一个单元,完成对时空序列特征的学习。

1.2 模型参数设置

为了得到更合理、准确的模型评估结果,对获取的 38 场洪水按照 3 : 1 的比例划分数据集,前 28 场洪水数据用于模型率定/训练,后 10 场洪水数据用于模型测试。XAJ 的参数见表 1,模型预热期为 48 h。对于神经网络模型数据集,以 25 h 为样本长度,根据时间过程滑动窗口生成样本,以训练数据集为基准,对数据进行最大-最小归一化处理。训练过程中,以平均绝对误差 (MAE) 作为损失函数,学习率设置为 0.001,使用早停法保存模型,早停迭代周期数为 200。

为了更好地评估 XAJ-ACL 性能,分别采用 ConvLSTM 和传统 LSTM 替代 XAJ 汇流模块,构建混合水文模型 XAJ-CL 和 XAJ-LSTM。3 个混合水文模型均采用 32 个神经元全连接层作为输入层和

表 1 XAJ 模型参数
Table 1 Parameters of XAJ

参数	物理意义	参数值
KC	流域蒸散发折算系数	0.93
UM	上层张力水容量	20 mm
LM	下层张力水容量	60 mm
C	深层蒸散发折算系数	0.18
WM	流域平均张力水容量	120 mm
B	张力水蓄水容量曲线方次	0.3
IM	不透水面积占比	0.01
SM	表层自由水蓄水容量	42 mm
EX	表层自由水蓄水容量曲线方次	1.5
KG	表层自由水库对地下水的日出流系数	0.3
KI	表层自由水库对壤中流的日出流系数	0.57
CI	壤中流消退系数	0.76
CG	地下水消退系数	0.98
CS	河网蓄水消退系数	0.5
L	滞时	0
KE	马斯京根法演算洪水传播滞时参数	3 h
XE	马斯京根法演算入流-出流权重参数	0.4

输出层。在隐藏层,三者均使用 Sigmoid 作为激活函数,区别在于 XAJ-ACL 使用注意力模块与 32 个神经元的 ConvLSTM 单元;XAJ-CL 仅使用 32 个神经元的 ConvLSTM 单元;XAJ-LSTM 则使用 32 个神经元的 LSTM 单元。

1.3 评价指标

采用纳什效率系数 (NSE)、相关系数 (r)、相对误差 (RE)、MAE 和均方误差 (MSE) 等 5 种指标对模型结果进行评估,其中 NSE 和 r 主要用来评价模型的整体回归程度。

2 实例验证

2.1 研究区概况与数据来源

呈村流域属钱塘江水系,位于安徽省休宁县境内,地处 29°32′40″N ~ 29°44′13″N、117°38′35″E ~ 117°53′21″E,总面积约 298 km²,是典型的中小流域。流域地形以丘陵为主,地势南高北低;降水主要集中在 4—7 月,降水、径流年内分配不均匀,年际变化大;流域洪水流速快,破坏力大,洪水过程陡涨陡落,历时较短。流域境内有棣甸、左龙等 7 个雨量站和 1 个位于流域出口处的水文站,各站位置如图 2 所示。

数字高程数据来自地理空间数据云 (<https://www.gscloud.cn/>),空间分辨率为 30 m,用于提取水系并结合观测站点进行子流域划分。洪水观测数据来自水文年鉴,选取呈村流域 1989—1996 年 38 场洪水数据进行模型率定/训练与测试。经统计,呈村流域年均降水量约 1 600 mm,年均径流量为 4.5 亿 m³,径流模数为 48.2 L/(s · km²)。

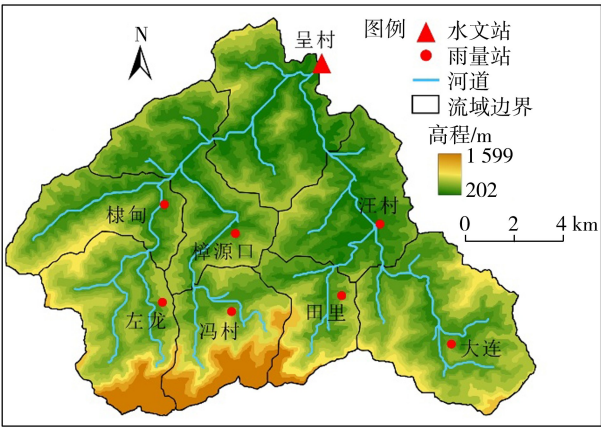


图 2 研究区概况

Fig. 2 Overview of study area

2.2 结果与分析

2.2.1 XAJ 与混合水文模型结果对比

XAJ 汇流模块使用深度学习模型改进后,传统 XAJ 汇流模块的参数率定过程变为神经网络的学习过程。XAJ-ACL 汇流模块学习曲线如图 3 所示,损失函数在模型训练前期快速下降,500 次迭代后减小缓慢,损失函数变化过程中存在小范围波动。从训练样本和测试样本的对比看,训练样本 MAE 值在约 100 次迭代开始逐渐低于测试样本,存在一定程度的过拟合现象。这一问题也反映在表 2 不同指标的结果中,混合水文模型测试期的 NSE, r , MAE 指标均比训练过程略差。这些现象说明,模型在样本容量有限的情况下,参数学习与调整较为困难,同时增加了模型训练的耗时。从模拟结果来看, XAJ-ACL 的性能整体上优于 XAJ。混合水文模型测试期 NSE 相较于传统 XAJ 提高了 4%~7%, MAE 提高超过 15%, MSE 提高超过 50%。但混合水文模型测试期 RE 结果较差,在一定程度上反映出混合水文模型在泛化能力和对汇流过程物理规律理解中的问题。混合水文模型在训练期整体表现相近,即训练期三者都能较好地学习汇流过程的复杂规律并合理表达。测试期尽管混合水文模型性能有所下降,但大多数指标优于 XAJ。混合水文模型模拟结果与观测值的趋势具有一致性,相较于传统 XAJ 在模拟精

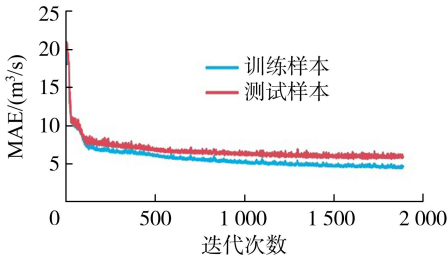


图 3 XAJ-ACL 汇流模块学习曲线

Fig. 3 Learning curve of confluence module in XAJ-ACL

度上有较大优势,且 XAJ-ACL 模拟效果最好。

表 2 XAJ 与混合水文模型评价指标比较

Table 2 Comparison of evaluation indexes of XAJ and hybrid hydrological models

模型	时期	NSE	r	RE/%	MAE/ (m^3/s)	MSE/ (m^3/s) ²
XAJ	率定期	0.83	0.90	13	5.82	224.77
	测试期	0.80	0.88	5	5.23	132.74
XAJ-ACL	训练期	0.87	0.94	15	3.53	98.44
	测试期	0.85	0.93	8	3.99	59.75
XAJ-CL	训练期	0.86	0.89	7	3.60	100.95
	测试期	0.82	0.88	12	4.17	76.52
XAJ-LSTM	训练期	0.86	0.92	1	3.59	93.15
	测试期	0.83	0.89	8	4.43	83.52

图 4 为 XAJ-ACL 的拟合效果,训练期和测试期流量实测值与模拟值呈现明显的线性相关性,大部分点在 1:1 线附近,离群值分布较为均匀,且测试期的误差小于训练期。

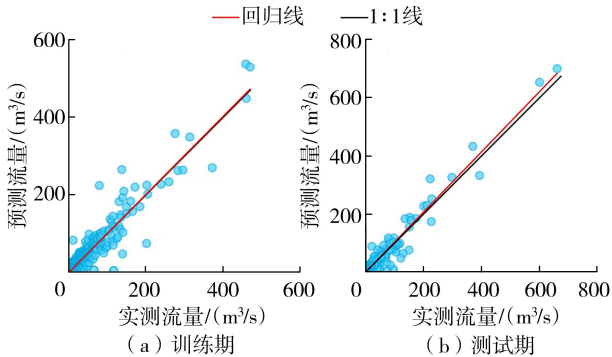


图 4 XAJ-ACL 拟合效果

Fig. 4 Fitting results of XAJ-ACL

上述指标更侧重于从整体角度评价模型,而径流模拟结果的优劣与每个样本时序过程密切相关。本文选取测试期 6 个主要洪水过程,针对传统 XAJ 与混合水文模型进行径流过程拟合结果的对比与分析,结果如图 5 所示。

从整体上看,混合水文模型对小洪量和大洪量的模拟效果均优于传统 XAJ,特别是对洪峰流量和峰现时间的模拟效果更好。混合水文模型的模拟结果存在一定差异。XAJ-LSTM 在径流量较小时表现更好,而对较大洪水的模拟结果在洪峰流量和峰现时间上存在误差。如图 5(e) 和 (f) 所示, XAJ-LSTM 模拟的峰现时间与观测结果相比有较大误差;图 5(e) 中, XAJ-LSTM 洪峰流量模拟结果相较于实测值的相对误差超过 40%,远未达到水文情报预报规范要求。XAJ-CL 对较大洪水的模拟结果远优于 XAJ-LSTM,但 XAJ-CL 的模拟结果并不稳定,特别是对于连续发生的洪水。如图 5(b) 所示, XAJ-CL 在连续两场洪水的模拟中模拟精度有所下降;同样的情况出现在图 5(f) 中,在经历过一场较大洪水

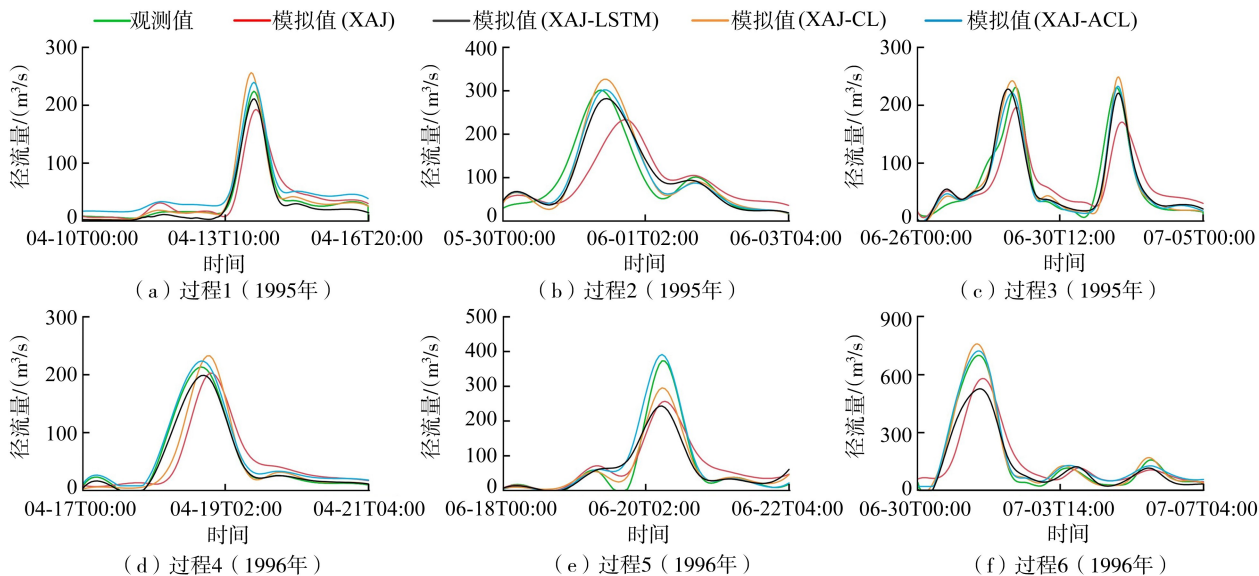


图 5 测试期主要洪水过程拟合结果

Fig. 5 Fitting results of main flood processes during testing period

后,XAJ-CL 在后续两场小洪水模拟中不仅精度有所下降,对于峰现时间的模拟也出现了偏差。结合中小流域实际情况看,导致这种现象最可能的原因是训练样本质量不高或容量不足。相应的情况在 XAJ-ACL 中得到了改善。在上述两场复杂洪水过程中(图 5(b)和(f)),XAJ-ACL 对径流量和峰现时间的模拟更加准确,模型稳定性有所提高。此外,XAJ-ACL 基本能准确模拟峰现时间,尽管在模拟较大流量洪水时存在高估洪量的情况,但整体平均相对误差小于 10%,符合水文情报预报规范要求。XAJ-ACL 模拟的精准度和稳定性说明,注意力机制提高了混合水文模型对有限样本特征的学习能力。而注意力机制对洪水事件过度关注导致模型在模拟较大洪水时洪量偏高,虽然从流域风险的角度考虑留出了安全冗余度,但过分高估洪量会导致不必要的防洪投入。

混合水文模型的优势在于数据驱动可以捕捉洪峰特征,而 XAJ-LSTM、XAJ-CL 缺陷源于小样本下难以通过学习获得大洪水的复杂和非线性规律,导致两个模型在测试期 RE 值增大。XAJ-ACL 引入注意力机制虽提升了学习能力,但过度关注洪水事件导致洪量高估,暴露出数据驱动与物理约束(如洪量平衡)的矛盾。未来研究需结合降水径流物理机制优化模型,权衡模拟精度、稳定性与水量平衡的合理性,降低防洪资源的过度投入与模拟结果的不确定性。

对率定/训练期和测试期样本模拟结果统计发现,混合水文模型对样本的拟合精度远优于 XAJ,且训练期混合水文模型对各样本的拟合精度相差不

大,差别主要出现在测试期(图 6)。测试期 XAJ-CL 存在拟合结果 NSE 远小于 0 的样本,且存在较多离群异常值,说明 XAJ-CL 在径流模拟稳定性方面存在较大问题。对于 XAJ-ACL 和 XAJ-LSTM,二者异常值和离群值的数量相差不大,但前者总体精度和平均 NSE 值均优于后者,说明 XAJ-ACL 能够更好地利用有限的学习产汇流过程中的相关规律,并具有更好的泛化能力。混合水文模型优势源于物理机制与数据驱动的融合,训练期差异较小体现出三者对训练数据的适配性相近,测试期差异体现出模型核心能力的分化。XAJ-CL 稳定性差,可能是由于其耦合逻辑未能在物理约束与数据拟合间达到平衡,导致其对陌生数据拟合精度降低;而 XAJ-ACL 的优势在于其自适应学习机制(如注意力、动态调整),使其能精准捕捉产汇流关键规律,这对缺资料流域的径流模拟具有重要价值,克服了传统模型泛化能力弱的不足。

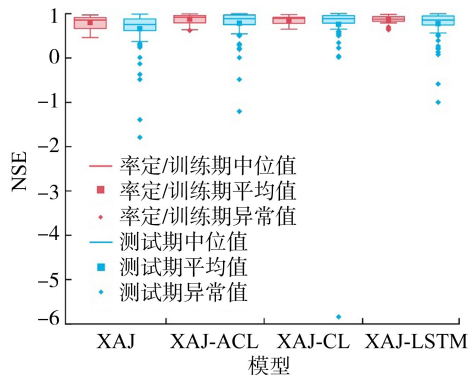


图 6 XAJ 与混合水文模型样本模拟结果箱线图

Fig. 6 Boxplots of simulation results of samples obtained by XAJ and hybrid hydrological models

总体上,混合水文模型整体模拟精度和预测稳定性相较于 XAJ 均有所提高,特别是提高了对洪峰模拟的准确度。然而,受到训练数据集大小的限制,混合水文模型存在一定的过拟合现象,依旧存在提升空间。

2.2.2 不同样本容量下混合水文模型的稳定性

为进一步探究 XAJ-ACL 面对不同样本容量数据集时的性能,分别设置了 100%、85% 和 70% 训练样本容量大小的 3 组梯度容量数据集进行分析。图 7 为混合水文模型测试期对 3 组梯度容量数据集运行 5 次得到的模拟结果箱线图。

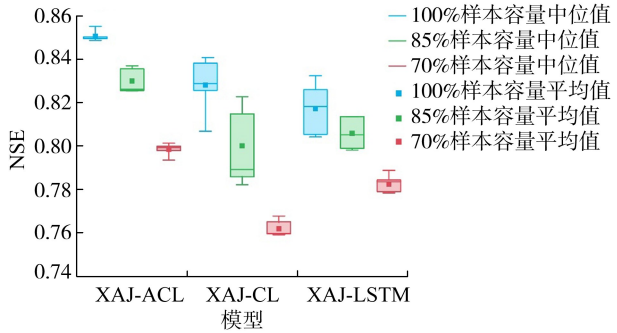


图 7 测试期混合水文模型不同样本容量模拟结果箱线图
Fig. 7 Boxplots of simulation results of hybrid hydrological models for different sample sizes in testing period

从样本容量角度看,较大样本容量训练得到的模型在验证期表现较好,说明中小流域中有限样本容量给数据驱动的深度学习模型带来了一定的难度。从模型角度来看,XAJ-ACL 在预测稳定性和准确度方面均为最佳,3 组样本容量下的平均 NSE 分别为 0.847、0.832 和 0.808,优于 XAJ-LSTM。XAJ-CL 表现最差,甚至在 100% 样本容量时,NSE 表现不如 70% 训练样本容量时的 XAJ-ACL,且稳定性较差。因此,注意力机制很大程度上提高了 ConvLSTM 在面对小样本容量时的学习能力,有助于构建性能更强的物理-神经网络混合水文模型。

小样本给中小流域深度学习模型带来了数据瓶颈,使模型难以充分学习产汇流关键特征。XAJ-ACL 的优势在于注意力机制能聚焦核心水文过程,弥补小样本下特征提取不足;而 XAJ-CL 因缺乏该机制,无法筛选有效信息,导致其性能较差。这印证了注意力机制对物理-神经网络混合水文模型的关键作用,即平衡数据稀缺性与特征学习效率,为缺资料流域构建可靠模型提供有效路径。

数据驱动的深度学习模型一般为黑箱模型,模型内部不包含实际的物理信息,给模拟结果的可靠性带来了挑战^[26]。本文计算了 XAJ 和混合水文模型模拟结果与实测值的 RE,从径流过程总流量的角度评估混合水文模型对实际物理规律的学习与表达

能力。测试期不同样本容量下各模型 RE 与 NSE 的相关关系如图 8 所示,图中浅色阴影部分表示各模型在 3 种样本容量下的稳定程度。

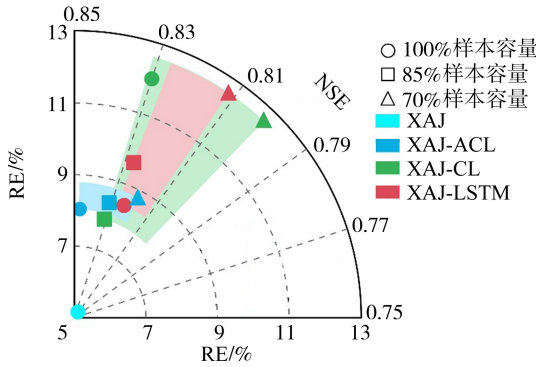


图 8 测试期模型 RE 与 NSE 相关关系
Fig. 8 Correlation between RE and NSE for models in testing period

相较于 XAJ,3 个混合水文模型的 NSE 值普遍较高,但径流过程总流量与实际观测值存在较大偏差。XAJ-ACL 对于不同样本容量都有出色表现,其平均 RE 仅为 8%,低于 XAJ-LSTM 和 XAJ-CL(平均 RE 值均为 10%)。此外,在 3 组样本容量的对比中,XAJ-ACL 的 RE 和 NSE 两个指标的变化幅度较小,模型稳定性较高,在稳定性和预报精度方面相较于其他混合水文模型有较大的优势。值得注意的是,XAJ 虽然在 NSE 值上仅为 0.84,但其 RE 值却在所有模型中最小,这在一定程度上体现了 XAJ 在物理规律模拟中的优势,也是混合水文模型需要完善的地方。

深度学习模型“黑箱”特性使其难于精准匹配水量平衡等物理规律,虽然混合水文模型的 NSE 值高却存在水量偏差。XAJ-ACL 因注意力机制聚焦关键水文过程,在精度与物理规律间取得较好平衡。但 XAJ 在 RE 值上的优势表明,混合水文模型需进一步融合物理机制(如水量平衡约束),以解决“黑箱”特性与物理机制一致性的矛盾,这是提升混合水文模型可靠性的核心方向。

整体来看,在采用较小样本容量训练时,XAJ-ACL 在性能上虽有所下降,但依然保持了较高预测精度,且在模型稳定性方面优于 XAJ-LSTM。此外,在几乎所有评价指标中,XAJ-ACL 表现均优于 XAJ-CL,证明了注意力机制在样本容量有限的中小流域 XAJ-ACL 构建中具有重要作用。

3 结 论

a. XAJ-ACL 能显著提升中小流域径流模拟精度,尤其在对洪峰关键特征(洪峰流量、峰现时间)预测中表现出精准的捕捉能力,可有效降低中小流

域因洪峰误判导致的应急响应滞后或过度响应风险,为短历时、高风险中小流域洪水预报提供更精准的技术支撑。

b. 在面对不同样本容量数据集时,相较于 XAJ-LSTM 和 XAJ-CL, XAJ-ACL 在模拟精度和稳定性方面表现出较大优势。XAJ-ACL 可优先聚焦洪水演进的关键特征,减少无关信息干扰,在有限样本中高效提取非线性汇流规律。

c. XAJ-ACL 作为一种完全基于数据驱动的汇流计算改进方案,在物理机制方面相较于 XAJ 仍存在不足。后续研究中,可在产流计算模块中增加对产流时空动态变化的考虑,并将水量平衡、能量守恒等物理关系作为神经网络汇流模块的约束条件,以提高模型的可解释性。

参考文献:

[1] 刘志雨,刘玉环,孔祥意. 中小河流洪水预报预警问题与对策及关键技术应用[J]. 河海大学学报(自然科学版),2021,49(1):1-6. (Liu Zhiyu, Liu Yuhuan, Kong Xiangyi. Problems, strategies and key technology research of flood forecasting and early warning for small and medium-sized rivers [J]. Journal of Hohai University (Natural Sciences), 2021, 49 (1) : 1-6. (in Chinese))

[2] Zang Shuaihong, Li Zhijia, Zhang Ke, et al. Improving the flood prediction capability of the Xin' anjiang model by formulating a new physics-based routing framework and a key routing parameter estimation method [J]. Journal of Hydrology, 2021, 603: 126867.

[3] 晁丽君,张珂,陈新宇,等. 基于多源降水融合驱动的 WRF-Hydro 模型在中小河流洪水预报中的适用性 [J]. 河海大学学报(自然科学版),2022,50(3):55-64. (Chao Lijun, Zhang Ke, Chen Xinyu, et al. Applicability of WRF-Hydro model based by multi-source precipitation merging in flood forecasting for small and medium-sized watersheds [J]. Journal of Hohai University (Natural Sciences), 2022, 50 (3) : 55-64. (in Chinese))

[4] 张珂,牛杰帆,李曦,等. 洪水预报智能模型在中国半干旱半湿润区的应用对比[J]. 水资源保护,2021,37(1):28-35. (Zhang Ke, Niu Jiefan, Li Xi, et al. Comparison of artificial intelligence flood forecasting models in China's semi-arid and semi-humid regions [J]. Water Resources Protection, 2021, 37 (1) : 28-35. (in Chinese))

[5] 李致家,臧帅宏,刘志雨,等. 新安江模型中河道汇流方法的改进[J]. 河海大学学报(自然科学版),2020,48(3):189-194. (Li Zhijia, Zang Shuaihong, Liu Zhiyu, et al. Improvement of channel routing method for Xin' anjiang model [J]. Journal of Hohai University (Natural Sciences), 2020, 48 (3) : 189-194. (in Chinese))

[6] Luo Yuning, Zhang Ke, Wang Yuhao, et al. iRainSnowHydro v1. 0: a distributed integrated rainfall-runoff and snowmelt-runoff simulation model for alpine watersheds [J]. Journal of Hydrology, 2024, 645: 132220.

[7] Wang Yuhao, Zhang Ke, Luo Yuning, et al. GHydroMod-SIM-DAI v1. 0: a new distributed hydrological model capable of dynamically recognizing runoff generation mechanisms and accounting for aboveground and underground anthropogenic impacts [J]. Journal of Hydrology, 2025, 663: 134184.

[8] 阚光远,刘志雨,李致家,等. 新安江产流模型与改进的 BP 汇流模型耦合应用 [J]. 水科学进展, 2012, 23 (1) : 21-28. (Kan Guangyuan, Liu Zhiyu, Lizhijia, et al. Coupling Xinanjiang runoff generation model with improved BP flow concentration model [J]. Advances in Water Science, 2012, 23 (1) : 21-28. (in Chinese))

[9] 李巧玲. 基于 DEM 的水文模型及流域汇流非线性问题研究 [D]. 南京:河海大学,2007.

[10] 杨雪. 非线性马斯京根模型研究与应用 [D]. 西安:西安理工大学,2015.

[11] Sabzipour B, Arsenault R, Troin M, et al. Comparing a long short-term memory (LSTM) neural network with a physically-based hydrological model for streamflow forecasting over a Canadian catchment [J]. Journal of Hydrology, 2023, 627: 130380.

[12] Moishin M, Deo R C, Prasad R, et al. Designing deep-based learning flood forecast model with ConvLSTM hybrid algorithm [J]. IEEE Access, 2021, 9: 50982-50993.

[13] 申笑萱,张珂,张兆安,等. 水文多模型拆分组配与协同优化研究进展 [J]. 水资源保护, 2025, 41 (4) : 189-196. (Shen Xiaoxuan, Zhang Ke, Zhang Zhaoan, et al. Research progress on hydrological multi-model decomposition allocation and collaborative optimization [J]. Water Resources Protection, 2025, 41 (4) : 189-196. (in Chinese))

[14] Wang Chao, Jiang Shijie, Zheng Yi, et al. Distributed hydrological modeling with physics-encoded deep learning: a general framework and its application in the Amazon [J]. Water Resources Research, 2024, 60 (4) : e2023WR036170.

[15] Chen Chen, Jiang Jiange, Liao Zhan, et al. A short-term flood prediction based on spatial deep learning network: a case study for Xi County, China [J]. Journal of Hydrology, 2022, 607: 127535.

[16] Fang Wei, Qin Hui, Liu Guanjun, et al. A method for spatiotemporally merging multi-source precipitation based on deep learning [J]. Remote Sensing, 2023, 15 (17) : 4160.

(区)为例[J]. 水资源与水工程学报, 2024, 35(6): 33-41. (Jia Mengyuan, Zhang Ziqi, Zhang Xinran, et al. Coupling mechanism and measurement method of urban development and water resources environment: a case study of 51 counties (districts) in the Daqing River Basin [J]. Journal of Water Resources and Water Engineering, 2024, 35(6): 33-41. (in Chinese))

[19] 董怡, 邹磊, 夏军, 等. 武汉市人水系统耦合协调发展及障碍因素识别[J]. 水资源保护, 2023, 39(6): 87-94. (Dong Yi, Zou Lei, Xia Jun, et al. Coupling and coordinated development of human-water system in Wuhan City and identification of obstacle factors [J]. Water Resources Protection, 2023, 39(6): 87-94. (in Chinese))

[20] 刘沛衡. 中原城市群水资源-经济社会-生态环境互馈机制及耦合协调模拟[D]. 郑州: 华北水利水电大学, 2024.

[21] 左其亨, 王娇阳, 马军霞, 等. 匹配度、协调度与和谐度的概念区别与计算方法对比[J]. 华北水利水电大学学报(自然科学版), 2023, 44(1): 1-9. (Zuo Qiting, Wang Jiaoyang, Ma Junxia, et al. Conceptual differences and calculation methods comparison of matching degree, coordination degree and harmony degree [J]. Journal of North China University of Water Resources and Electric Power (Natural Science Edition), 2023, 44(1): 1-9. (in Chinese))

[22] 张恬姿, 王小军, 齐广平, 等. 甘肃省水资源-经济社会-生态环境耦合协调驱动因素研究[J]. 水利水运工程学报, 2023(3): 38-46. (Zhang Tianzi, Wang Xiaojun, Qi Guangping, et al. Driving factors of water resources-economic society-ecological environment coupling in Gansu Province [J]. Hydro-Science and Engineering, 2023(3): 38-46. (in Chinese))

[23] 胡浩, 康玲芬, 尚雪英, 等. 基于 ESRE 系统评价的甘肃区域协调发展时空分析[J]. 甘肃行政学院学报, 2021(6): 111-121. (Hu Hao, Kang Lingfen, Shang Xueying, et al. Spatial-temporal analysis of the coordinated development of Gansu based on the ESRE system evaluation [J]. Journal of Gansu Administration Institute, 2021(6): 111-121. (in Chinese))

(收稿日期: 2025-04-11 编辑: 施业)

(上接第 143 页)

[17] Dehghani A, Moazam H M Z H, Mortazavizadeh F, et al. Comparative evaluation of LSTM, CNN, and ConvLSTM for hourly short-term streamflow forecasting using deep learning approaches [J]. Ecological Informatics, 2023, 75: 102119.

[18] Liao Yaoxing, Wang Zhaoli, Yu Haijun, et al. Accelerating urban flood inundation simulation under spatio-temporally varying rainstorms using ConvLSTM deep learning model [J]. Water Resources Research, 2025, 61(8): e2025WR040433.

[19] Ha Si, Liu Darong, Mu Lin. Prediction of Yangtze River streamflow based on deep learning neural network with El Niño-Southern oscillation [J]. Scientific Reports, 2021, 11(1): 11738.

[20] Niu Zhaoyang, Zhong Guoqiang, Yu Hui. A review on the attention mechanism of deep learning [J]. Neurocomputing, 2021, 452: 48-62.

[21] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017: 6000-6010.

[22] Zhang Jinyu, Yan Hua. A long short-term components neural network model with data augmentation for daily runoff forecasting [J]. Journal of Hydrology, 2023, 617: 128853.

[23] Cui Zhen, Guo Shenglian, Zhou Yanlai, et al. Exploration of dual-attention mechanism-based deep learning for multi-step-ahead flood probabilistic forecasting [J]. Journal of Hydrology, 2023, 622: 129688.

[24] Chrysos G G, Wu Yongtao, Pascanu R, et al. Hadamard product in deep learning: introduction, advances and challenges [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025, 47(8): 6531-6549.

[25] Shi Xingjian, Chen Zhouong, Wang Hao, et al. Convolutional LSTM network: a machine learning approach for precipitation nowcasting [C]//Proceedings of the 29th International Conference on Neural Information Processing Systems. Montreal: MIT Press, 2015: 802-810.

[26] Ismail A A, Gunady M, Corrada Bravo H, et al. Benchmarking deep learning interpretability in time series predictions [C]//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver: NeurIPS, 2020: 6441-6452.

(收稿日期: 2024-12-16 编辑: 施业)