

DOI:10.3969/j.issn.1672-1144.2026.03.028

基于 CNN-BiLSTM-AM 模型的河道洪水水位预测研究

潘景辰, 韩伟刚, 高 涵

(陕西省水利电力勘测设计研究院(集团)有限公司, 陕西 西安 710001)

摘要: 针对传统水动力学模型对糙率率定依赖性强、率定过程耗时长的问题, 提出一种融合卷积神经网络、双向长短期记忆网络和注意力机制的 CNN-BiLSTM-AM 水位预测模型。以某典型河道 73 个实测断面数据为对象, 选取洪水流量、深泓点高程、水面宽度、过水面积和断面流速为输入特征, 按河程顺序前 70% 作训练集、后 30% 作测试集, 并与 SVR、单一 BiLSTM 模型对比。结果表明: 该模型在测试集上的 MAPE、MAE 和 RMSE 分别为 0.36%、3.43m 和 4.02m, 预测精度均优于两种基准模型, 但对下游低水位区间的连续下降趋势仍存在一定高估。该模型可在无需复杂糙率率定条件下, 为防洪调度和河道治理方案比选提供可靠的数据驱动预测手段。

关键词: 河道洪水; 水位预测; CNN-BiLSTM-AM; 深度学习; 注意力机制

中图分类号: TU43

文献标识码: A

文章编号: 1672-1144(2026)03-0203-06

River Flood Water Level Prediction Using CNN-BiLSTM-AM Model

PAN Jingchen, HAN Weigang, GAO Han

(Shaanxi Province Institute of Water Resources and Electric Power Investigation and Design
Institute (Group) Co., Ltd., Xi'an, Shaanxi 710001, China)

Abstract: To address the strong dependence of conventional hydrodynamic models on roughness calibration and the time-consuming nature of the calibration process, a CNN-BiLSTM-AM river water-level prediction model is proposed by integrating convolutional neural networks, bidirectional long short term memory networks and an attention mechanism. Taking 73 measured cross sections of a typical natural river as the study object, flood discharge, thalweg elevation, water surface width, flow area and cross-sectional velocity are selected as input features, the cross sections are divided in along reach order, with the upstream 70% used as the training set and the downstream 30% as the test set, and the proposed model is compared with a support vector regression model and a single BiLSTM model. The results show that the MAPE, MAE and RMSE of the proposed model on the test set are 0.36%, 3.43m and 4.02m, which outperforms both benchmark models. However, the model still systematically overestimates the downward trend in the low-water-level range downstream. The proposed method can provide a reliable data-driven prediction tool for flood-control operation and comparison of river-regulation schemes without complex roughness calibration.

Keywords: river flood; water level prediction; CNN-BiLSTM-AM; deep learning; attention mechanism

随着我国城市化进程不断深入, 叠加全球气候变化引发的极端天气事件频发, 流域防洪安全正面临持续增大的压力, 发展高精度、高时效的河道水位预测模型已成为防洪减灾规划与应急调度决策的核心支撑^[1]。天然河道中, 水位沿程分布受河床地形、断面突变、回水顶托以及人类活动等多种因素共

同作用, 呈现明显的非线性特征^[2], 在断面尺度上的精确模拟一直是防洪水文研究中的难点。一方面, 在缺少可靠糙率参数的情况下, 断面几何突变(如局部束窄、扩宽或坡降变化)往往引起水位的剧烈波动, 如何刻画这一过程仍缺乏成熟手段; 另一方面, 上游来水的顺向演进和下游回水的顶托作用同

收稿日期: 2025-12-24

修稿日期: 2026-02-07

基金项目: 中央高校基本科研业务费专项资金资助(300102255505)

作者简介: 潘景辰(1992—), 女, 硕士, 工程师, 主要从事水利水电工程设计等方面工作。E-mail: tjz13370135093@126.com

时存在,如何在模型中同时反映这两种水力联系,也尚未得到很好解决。

针对河道水位预测,目前主要有两类方法。一是基于圣维南方程组的一维或二维水动力学模型^[3-4]。这类方法物理意义清晰,在常规河段具有较好的可解释性,但其计算精度对糙率和断面几何参数较为敏感。糙率受植被、底质和水位变化影响,本身具有较强的时空变异性,率定过程耗时较长,还常常出现“异参同效”现象^[5],难以满足防洪抢险对实时性的要求。二是数据驱动方法。长短期记忆网络(LSTM)在洪水过程预测中应用较多,擅长捕捉时间序列的长短期依赖^[6]。卷积神经网络(CNN)能够提取局部空间特征,但对长程依赖的刻画不足^[7-8]。单向 LSTM 只能沿正向传递信息,难以表征下游回水的顶托作用^[9],在断面形态突变区域往往精度欠佳。可见,如何在一个数据驱动框架下,把局部空间特征提取、双向水力依赖建模和关键要素自适应加权结合起来,是河道断面水位预测中需要解决的关键问题。

为应对上述挑战,本文基于某典型天然河道的

实际洪水监测数据,提出一种融合卷积神经网络(CNN)、双向长短期记忆网络(BiLSTM)和注意力机制(AM)的混合预测模型(CNN-BiLSTM-AM),并通过与支持向量回归(SVR)、单一 BiLSTM 等基准模型的对比,系统评估其在非均匀断面河段上的预测精度与空间适应性,以期为不依赖糙率率定的河道沿程水位推算提供新的技术途径。

1 研究方法

1.1 研究区域概况

本研究以某典型天然河道防洪治理工程为对象,研究河段全长约 8.3 km。基于实测水文资料,沿程布设 74 个典型断面(序号 0+000—8+305)作为研究对象^[10]。研究河段总体走向自西向东,河谷形态以宽浅型为主,局部夹束窄段与扩宽段。从水力学特征看,研究河段在设计洪水条件下处于亚临界缓流状态,水位的空间分布同时受沿程河床高程、断面收缩与扩散及下游控制断面回水顶托的共同作用,研究河段的空间分布与典型断面位置如图 1 所示。

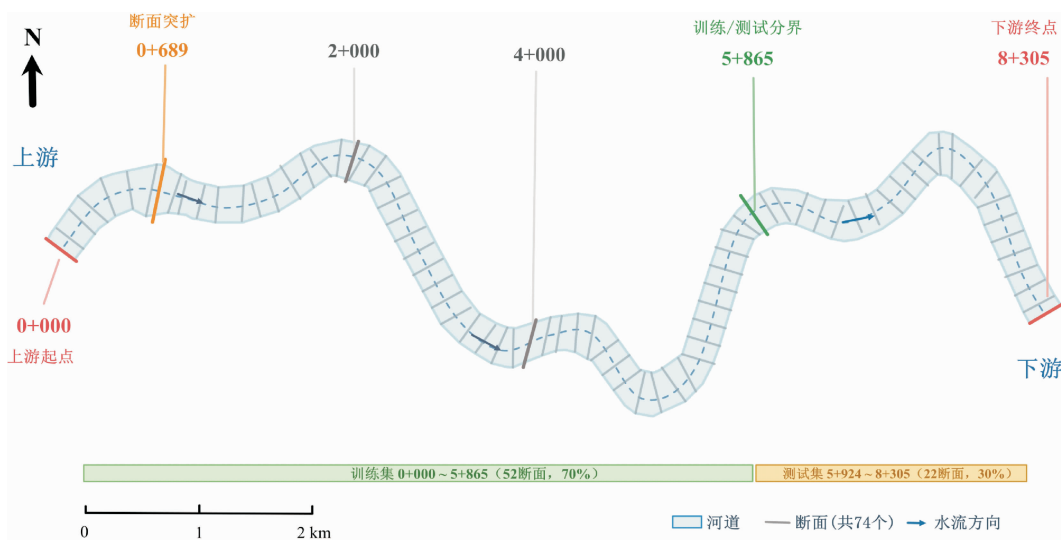


图 1 研究河段空间分布与典型断面布置示意图

1.2 数据集与样本划分

本研究关注的是河道沿程断面水位的空间映射问题,样本以断面里程为序排列,沿程连续分布。为避免随机划分导致相邻断面信息泄漏,采用按河程顺序划分的方式构建训练集与测试集,其中前 70% 样本,即 0+000—5+865 段共 52 个断面作为训练集,后 30% 样本,即 5+924—8+305 段共 22 个断面,作为测试集。模型训练阶段仅使用训练集数据进行参数学习,测试阶段仅对未参与训练的

5+924—8+305 段断面进行水位预测,并与实测水位进行误差统计,以评价模型在下游测试河段的预测精度与空间泛化能力。

1.3 河道沿程水力要素的空间分布特征

对原始数据进行统计分析后,得到河道沿程水力要素的空间分布如图 2 所示。由图 2 可知,该河段表现出典型的天然河道非线性特征。沿水流方向(0+000—8+305),深泓点高程从上游 956.94 m 逐渐降至下游 939.21 m,形成了缓坡与陡坡交替出

现的形态特征。同时,河宽沿程波动显著,平均河宽约为 140 m。尤其在断面 0+689 附近,河宽自上游相邻 0+561 断面的 143 m 沿水流方向突增至 233 m,扩宽系数约 1.63,紧邻下游 0+707 断面又骤缩至 154 m,形成一个明显的扩宽段。因该断面具有较强局部突变特征,后续建模时将其剔除,以降低单个强突变断面对模型训练和测试评价的影响。

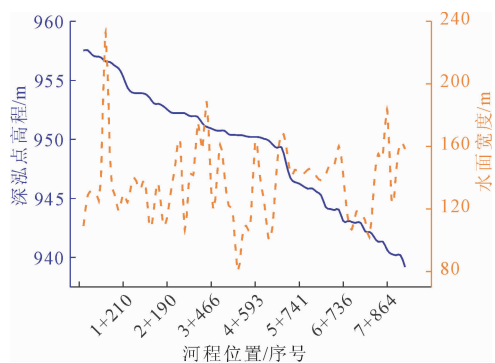


图 2 河道沿程深泓点高程与水面宽度空间变化特征

为验证输入特征选取的合理性,采用皮尔逊相关系数分析了各水力要素与断面水位的线性相关程度,结果如图 3 所示。分析表明,深泓点高程与断面水位的正相关系数高达 0.99,说明水位变化在很大程度上受河床地形控制,这与明渠水流的基本规律一致。断面流速与水位呈负相关,相关系数为 -0.91,反映出过水面积增大时流速减小,符合伯努利方程中流速水头与压强水头的转换关系。洪水流量与水位的相关系数为 -0.78,该负相关主要源于不同河段设计流量的空间差异,不宜简单解释为因果关联。因此,所选五个输入特征在物理机制上与水位具有明确的相关性,能够为模型训练提供有效的特征信息。

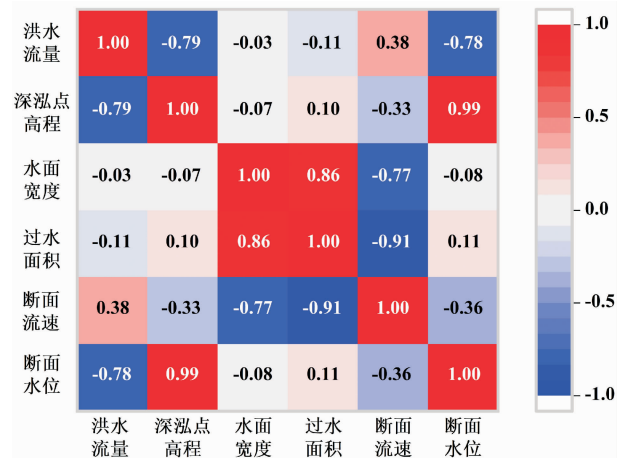


图 3 水力要素与断面水位的皮尔逊相关系数矩阵

1.4 数据归一化处理

由于输入的河道断面特征中,洪水流量、深泓点高程与断面流速等数据的物理量纲差异较大,直接输入神经网络容易导致梯度消失或收敛缓慢。因此,采用最小-最大归一化将所有输入数据映射至 [0, 1] 区间,计算公式为:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

式中: X_{norm} 为归一化后的数据; X 为原始输入数据; X_{min} 和 X_{max} 分别为该特征在全河段样本中的最小值和最大值,测试集采用训练集归一化参数进行同尺度转换。

1.5 模型评价指标

为量化评估模型的预测精度,选取平均绝对误差 (MAE)^[11]、均方根误差 (RMSE)^[12] 和平均绝对百分比误差 (MAPE)^[13] 作为评价指标,计算公式为:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - y_l| \quad (2)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_l)^2} \quad (3)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - y_l}{y_i} \right| \quad (4)$$

式中: y_i 和 y_l 分别为实测与预测水位; n 为样本数。

考虑到绝对高程作为分母可能导致 MAPE 偏小,主要以 RMSE 指标进行评价。

1.6 CNN-BiLSTM-AM 模型架构

本文构建的 CNN-BiLSTM-AM 模型是一种端到端的混合深度神经网络,旨在捕捉河道水力要素中复杂的非线性与空间异质性特征。整体技术路线与模型架构分别如图 4 和图 5 所示。模型主要包括以下三个依次连接的功能模块:

(1) 基于 CNN 的局部特征提取模块。输入层接收归一化后的多维水力特征向量,通过一维卷积提取洪水流量、深泓点高程、河宽、过水面积和流速等输入变量之间的局部组合关系,从而增强对关键水力特征的表达能力。

(2) 基于 BiLSTM 的非线性关系表征模块。该模块通过双向 LSTM 单元对输入特征序列进行双向编码,进一步刻画不同水力要素之间的非线性依赖关系。

(3) 基于 AM 的关键特征自适应加权模块。该模块计算不同特征对水位预测的相对贡献,并进行加权融合,使模型更加关注对断面水位变化贡献较大的关键变量。

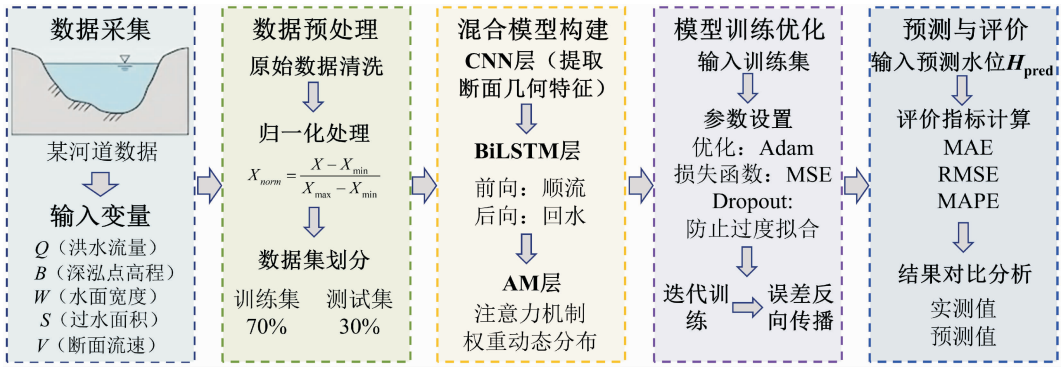


图 4 基于 CNN-BiLSTM-AM 的河道水位预测技术路线图

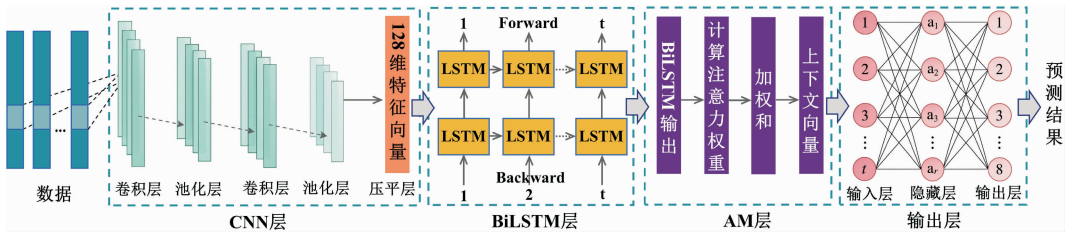


图 5 CNN-BiLSTM-AM 模型网络架构图

1.7 模型参数设置

本模型基于 TensorFlow 框架搭建,优化器选用 Adam,损失函数采用均方误差(MSE)。经多组超参数试验确定:CNN 卷积核数为 32,BiLSTM 隐藏单元为 64,批次大小为 8。训练轮数设为上限 200,并引入学习率自适应衰减以抑制过拟合并提高训练效率,所有试验固定随机种子以保证结果可重复。基准对比模型的参数设置同步说明,SVR 采用径向基核函数,通过网格搜索确定惩罚因子与核宽度,单一 BiLSTM 模型保持隐藏单元数量与本文模型一致以确保对比的可比性。训练过程中从训练集内部划分部分样本作为验证集,用于早停判断和学习率调整,测试集不参与模型训练与超参数调整。

2 CNN-BiLSTM-AM 模型预测结果与分析

2.1 模型收敛性分析

CNN-BiLSTM-AM 模型在训练过程中呈现出“快速下降-局部波动-稳定平台”的收敛特征,如图 6 所示。训练初期,训练损失迅速降低,说明模型能够较快捕捉断面水位与深泓点高程、河宽、过水面积及流速等水力要素之间的主要映射关系。验证损失在初始阶段出现短暂波动,随后逐渐趋于稳定,表明模型在完成主要特征学习后,对未参与训练样本的预测误差基本保持在相对稳定范围内。需要指出

的是,验证损失整体高于训练损失,说明模型在训练集和验证集之间存在一定泛化差距。这与研究样本数量有限、河道断面形态沿程变化复杂以及水力要素空间分布非均匀有关。尽管如此,训练后期验证损失未出现持续上升或明显发散,说明模型未发生严重失稳。CNN 模块有助于提取多维水力输入中的局部组合特征,BiLSTM 模块用于刻画特征间的非线性依赖关系,注意力机制进一步突出对水位预测贡献较大的关键变量^[14-15]。综合来看,CNN-BiLSTM-AM 模型具有较好的训练收敛性,但其预测精度仍受到小样本和空间异质性条件的制约。

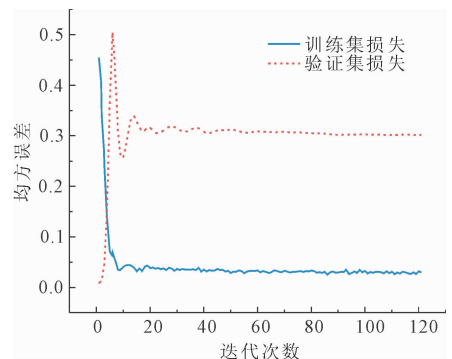


图 6 CNN-BiLSTM-AM 模型训练收敛曲线

2.2 预测精度与沿程变化分析

为直观评估 CNN-BiLSTM-AM 模型在未参与训练的下游河段测试集上的表现,将测试集水位预测

值与实测值沿程进行对比,结果如图7所示。由图7可知,测试河段实测水位整体呈沿程下降趋势,由5+924断面的954.58 m逐渐降低至8+305断面的948.73 m,反映出下游河床高程降低及断面水力条件变化对水位分布的控制作用。模型预测水位整体稳定在955~956 m附近,在测试集前段与实测水位较为接近,能够反映该河段水位处于较高水位区间的总体特征。随着河程向下游推进,实测水位下降幅度逐渐增大,而预测水位变化幅度相对较小,导致模型在测试集后段出现一定程度的系统性高估。这说明 CNN-BiLSTM-AM 模型能够学习训练河段中水位与主要水力要素之间的总体映射关系,但对下游测试河段中持续降低的水位趋势响应不足。该现象可能与训练样本数量有限、训练集与测试集之间存在空间分布差异,以及下游河段水力要素组合特征发生变化有关。结合评价指标, CNN-BiLSTM-AM 模型在测试集上的 MAE 、 $RMSE$ 和 $MAPE$ 分别为 3.43 m、4.02 m 和 0.36%,表明模型在总体预测精度上具有一定可用性,但在低水位下游断面仍存在进一步改进空间。

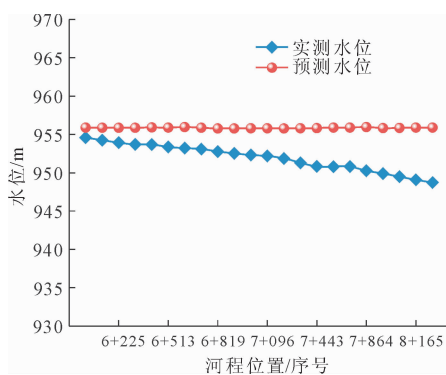


图7 测试集水位预测值与实测值沿程对比

2.3 与基准模型的性能对比分析

为进一步评价 CNN-BiLSTM-AM 模型的预测性能,本文选取 SVR 模型和单一 BiLSTM 模型作为基准模型进行对比分析^[16-17],并采用 MAE 、 $RMSE$ 和 $MAPE$ 作为评价指标,结果如图8所示。由图8可知 CNN-BiLSTM-AM 模型在三项评价指标上均取得最低误差,其 MAE 、 $RMSE$ 和 $MAPE$ 分别为 3.43 m、4.02 m 和 0.36%,SVR 模型对应指标分别为 3.93 m、4.28 m 和 0.41%。单一 BiLSTM 模型误差相对较大, MAE 、 $RMSE$ 和 $MAPE$ 分别为 4.60 m、4.89 m 和 0.48%。整体来看, CNN-BiLSTM-AM 模型在测试河段上的预测精度优于两个基准模型。从模型结构角度看,SVR 模型能够通过核函数建立输入水力

要素与断面水位之间的非线性映射关系,在小样本条件下具有一定稳定性,因此其预测误差低于单一 BiLSTM 模型。然而,SVR 本质上仍属于点对点静态回归模型,难以进一步提取多维水力特征之间的局部组合关系。单一 BiLSTM 模型虽然具备非线性序列建模能力,但在样本数量有限、断面水力要素变化较复杂的条件下,容易受到输入特征尺度差异和空间分布非均匀性的影响,导致预测误差相对较大。

相比之下, CNN-BiLSTM-AM 模型通过 CNN、BiLSTM 和注意力机制的组合,提高了对复杂水力要素的表征能力。其中, CNN 模块可对洪水流量、深泓点高程、河宽、过水面积和流速等多维输入特征进行局部组合特征提取,降低原始特征噪声对后续预测的影响。BiLSTM 模块进一步刻画水力要素之间的非线性关联。注意力机制则对不同输入特征进行自适应加权,使模型能够更关注对水位变化贡献较大的关键变量。与 BiLSTM 模型相比, CNN-BiLSTM-AM 模型的 MAE 、 $RMSE$ 和 $MAPE$ 分别降低约 25.4%、17.8% 和 25.0%。与 SVR 模型相比,三项指标分别降低约 12.7%、6.1% 和 12.2%。结果表明,所提出的 CNN-BiLSTM-AM 混合模型在当前测试集上具有更优的综合预测性能。考虑到深度学习模型存在随机初始化的固有波动,本文试验结果取自固定随机种子下的单次训练,后续若结合多次重复试验可进一步评估模型间差异的统计稳健性。

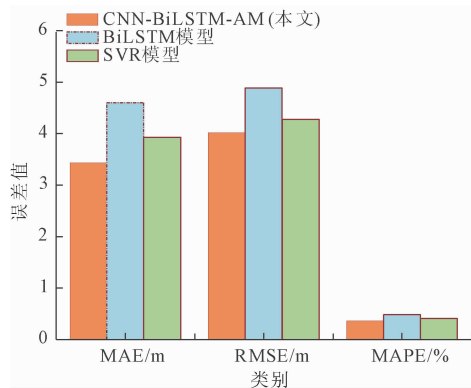


图8 三种预测模型的性能评价指标对比图

2.4 测试集预测残差的沿程空间分布

河道水位预测的难点不仅在于整体误差大小,更在于模型能否准确刻画水位沿程变化趋势。图9展示了 CNN-BiLSTM-AM 模型在测试河段的预测残差空间分布,其中残差定义为预测水位与实测水位之差。由图9可知,测试集残差整体为正,说明模型在5+924—8+305测试河段存在一定程度的系统

性高估。残差由测试河段前段的约 1.3 m 逐渐增大至末端的约 7.2 m,表明随着河程向下游推进,模型对实测水位下降趋势的跟踪能力逐渐减弱。

结合图 7 可知,测试河段实测水位沿程持续降低,而模型预测水位变化幅度相对较小,导致预测误差在下游低水位区间逐渐增大。这一现象说明,CNN-BiLSTM-AM 模型能够学习训练河段中水力要素与水位之间的总体映射关系,但在顺序划分条件下,训练集与测试集之间存在一定空间分布差异,模型对测试河段后段水位持续降低特征的外推能力仍然有限。其原因可能与样本数量较少、下游河床高程降低幅度较大以及水面宽度、过水面积和流速等水力要素组合关系发生变化有关。

从残差沿程分布看,模型误差并非随机波动,而是表现出随河程增加而逐步增大的趋势。这表明模型在当前数据划分方式下仍存在空间累积偏差。尽管 CNN 模块和注意力机制有助于提取多维水力特征中的关键变量信息,BiLSTM 模块也能够增强模型对非线性关系的表达能力,但在小样本和明显空间异质性条件下,仅依赖现有输入特征仍难以完全消除下游河段的系统性高估。

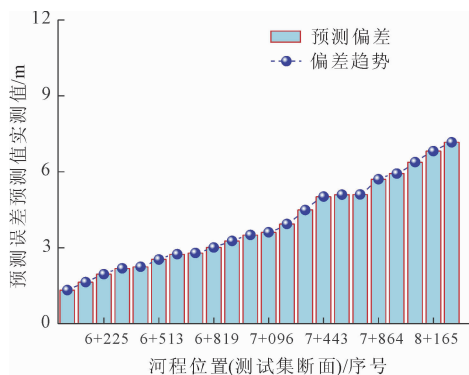


图 9 预测误差沿程空间分布特征

3 结 论

(1) 本文构建的 CNN-BiLSTM-AM 组合模型能够在不依赖糙率率定的前提下,建立河道水力要素与断面水位之间的非线性映射关系,在测试河段取得较高预测精度,MAPE、MAE 和 RMSE 分别为 0.36%、3.43 m 和 4.02 m,可在小样本与空间异质性条件下提供数据驱动的沿程水位估算。

(2) 与传统 SVR 模型和单一 BiLSTM 模型的对比结果表明,CNN-BiLSTM-AM 模型在三项指标上均有明显改善。与 SVR 相比,MAE、RMSE 和 MAPE 分别降低 12.7%、6.1% 和 12.2%,与单一 BiLSTM 相

比,三项指标分别降低 25.4%、17.8% 和 25.0%。表明 CNN 的局部特征提取、BiLSTM 的非线性关系表征和注意力机制的关键特征加权能够提高模型对复杂水力要素组合关系的表达能力,从而改善河道断面水位预测精度。

(3) 残差沿程分布结果表明 CNN-BiLSTM-AM 模型在测试河段整体存在一定系统性高估,预测残差由测试河段前段的约 1.3 m 逐渐增大至末端的约 7.2 m,说明模型对下游低水位区间的连续下降趋势刻画仍不充分。该现象与训练集和测试集之间的空间分布差异、样本数量有限以及下游河段水力要素组合变化有关。

参考文献:

- [1] 梅超,刘家宏,王佳,等.统筹城市防洪体系和内涝治理的难点及对策探讨[J].中国水利,2025(17):22-26.
- [2] 董立.南川河流域延安城区段防洪治理工程主体施工方案探讨[J].地下水,2023,45(03):313-314.
- [3] 薛文宇.城市暴雨积水及街道洪水模拟模型研究[D].天津:天津大学,2016.
- [4] 朱思鹏,孙怀卫,郑启明,等.基于一维二维耦合水动力学模型的库区水位响应规律研究[J].水电能源科学,2025,43(11):207-210.
- [5] 瞿聪,卢沈煜,顾秋鹏.一维河网模型构建及其在排涝计算中的应用[J].江苏水利,2025(04):44-47,52.
- [6] 徐晓芹,赵文聚,仇东山,等.长短时记忆神经网络模型在兴济河洪水预报中的应用[J].济南大学学报(自然科学版),2025,39(06):843-851.
- [7] 宋桃云,盖优普,郭易,等.深度学习在洪水预测与管理中的应用进展与挑战[J].吉林水利,2025(12):77-82.
- [8] 曹锦阳,刘梦,李嘉铮,等.基于随机森林和卷积神经网络的风-光伏-抽水蓄能电站联合优化运行[J].水利与建筑工程学报,2023,21(04):30-37.
- [9] 欧阳特.基于 BiLSTM 和迁移学习的水质预测方法研究[D].武汉:武汉大学,2021.
- [10] 许晓,索宁珠,付曼蓉,等.延河延安城区核心段河道治理工程可行性研究报告[R].西安:陕西省水利电力勘测设计研究院(集团)有限公司,2025:427-428,479.
- [11] 王文辉.基于 CNN-Transformer 神经网络反演的深基坑变形预测模型[J].水利与建筑工程学报,2025,23(05):17-24.