

杨润祎, 张红武. 基于深度学习的黄河下游水流阻力计算研究[J]. 水利水电技术(中英文), 2026, 57(1): 249-262. DOI: 10.13928/j.cnki.wrahe.2026.01.019

YANG Runyi, ZHANG Hongwu. Flow resistance calculation in lower Yellow River based on deep learning[J]. Water Resources and Hydropower Engineering, 2026, 57(1): 249-262. DOI: 10.13928/j.cnki.wrahe.2026.01.019

基于深度学习的黄河下游水流阻力计算研究

杨润祎, 张红武

(清华大学 水沙科学与水利水电工程国家重点实验室, 北京 100084)

摘要:【目的】冲积河流阻力的准确计算在治河防洪工程中具有重要意义。传统阻力计算公式与现有机器学习方法仍存在诸多不足。为提升阻力模型的性能与泛化能力, 建立一套基于深度学习的阻力计算方法, 【方法】选择弗劳德数、体积含沙量、宽深比、径深比、年径流量和年输沙量等水文特征作为模型输入, 构建一种基于深度森林的水流阻力计算模型。利用黄河下游水文站的实测数据进行模型训练与测试, 并从时空泛化能力、迁移学习表现等方面综合评估模型。【结果】模型在测试集上的纳什效率(NSE)为0.785, 平均绝对误差(MAE)为0.002, 均方根误差($RMSE$)为0.003, 平均绝对百分比误差($MAPE$)为14.618%。加入时空平均特征泛化模型后, 模型的 NSE 从0.6814提升至0.7427, MAE 从0.0023降至0.0021, $RMSE$ 从0.0032降至0.0028, $MAPE$ 从14.978%降至13.689%。在将模型迁移至全新的水沙条件下时, NSE 最大降幅达65.35%, MAE 、 $RMSE$ 和 $MAPE$ 的最大升幅分别为100%、150%和123.98%。【结论】深度森林模型相比于传统的阻力公式和机器学习模型, 在冲积河流一般条件下的阻力计算方面, 能展现出更为优越的预测精度。通过引入大尺度时空平均特征, 模型在不同水文站与水文时期的计算精度有效提升, 泛化能力明显增强。然而, 在应对特殊水沙情势时, 模型仍存在性能波动, 有时计算精度还超不过具有物理背景的糙率公式, 亟需结合物理机制以增强其迁移学习能力, 且在解决复杂环境下的黄河动床阻力计算问题时, 应该注重同良好的传统公式相互印证这一环节。

关键词: 黄河下游; 水流阻力; 深度森林; 迁移学习; 弗劳德数; 径流; 输沙量; 机器学习模型

DOI: 10.13928/j.cnki.wrahe.2026.01.019

开放科学(资源服务)标志码(OSID):

中图分类号: TV213.4

文献标志码: A

文章编号: 1000-0860(2026)01-0249-13



Flow resistance calculation in lower Yellow River based on deep learning

YANG Runyi, ZHANG Hongwu

(State Key Laboratory of Hydrosience and Engineering, Tsinghua University, Beijing 100084, China)

Abstract: [Objective] Accurate calculation of flow resistance in alluvial rivers is of great significance for river regulation and flood control engineering. Conventional resistance formulas and existing machine learning method still have multiple limitations.

收稿日期: 2025-05-14; 修回日期: 2025-07-09; 录用日期: 2025-07-10; 网络出版日期: 2025-08-25

基金项目: 国家重点研发计划项目(2023YFC3208603)

作者简介: 杨润祎(2001—), 男, 博士研究生, 主要从事河流动力学研究。E-mail: yry23@mails.tsinghua.edu.cn

通信作者: 张红武(1958—), 男, 教授, 博士研究生导师, 博士, 主要从事水力学及河流动力学研究。E-mail: zhhw@mail.tsinghua.edu.cn

©Editorial Department of Water Resources and Hydropower Engineering. This is an open access article under the CC BY-NC-ND license.

To improve the performance and generalization ability of resistance models, a flow resistance estimation method based on deep learning is proposed. [Methods] Hydrological features, including Froude number, volumetric sediment concentration, width-to-depth ratio, diameter-to-depth ratio, annual runoff, and annual sediment load, were selected as model inputs, and a flow resistance calculation model based on deep forest was established. The model was trained and tested using measured data from hydrological stations in the lower Yellow River, and comprehensively evaluated in terms of spatiotemporal generalization ability and transfer learning performance. [Results] The model achieved a Nash-Sutcliffe efficiency (NSE) of 0.785, a mean absolute error (MAE) of 0.002, a root mean square error ($RMSE$) of 0.003, and a mean absolute percentage error ($MAPE$) of 14.618% on the test dataset. After the incorporation of spatiotemporal average features, the NSE of the model increased from 0.6814 to 0.7427, the MAE decreased from 0.0023 to 0.0021, the $RMSE$ dropped from 0.0032 to 0.0028, and the $MAPE$ reduced from 14.978% to 13.689%. When the model was transferred to completely new water-sediment conditions, the maximum decline in NSE reached 65.35%, and the maximum increases in MAE , $RMSE$, and $MAPE$ were 100%, 150%, and 123.98%, respectively. [Conclusion] Compared with traditional resistance formulas and machine learning method, the deep forest model demonstrates superior accuracy in predicting flow resistance under general conditions in alluvial rivers. By introducing large-scale spatiotemporal average features, the model's calculation accuracy across different hydrological stations and hydrological periods is effectively improved, and its generalization ability is significantly enhanced. However, under special water-sediment conditions, the model still shows performance fluctuations. In certain cases, its calculation accuracy is even lower than that of physically based roughness formulas. Therefore, it is urgent to incorporate physical mechanisms to enhance its transfer learning capability. When addressing the calculation of movable bed resistance in the Yellow River under complex environments, emphasis should be placed on mutual verification with reliable traditional formulas.

Keywords: lower Yellow River; flow resistance; deep forest; transfer learning; Froude number; runoff; sediment transport volume; machine learning models

0 引言

作为中国人口和经济活动最为集中的区域, 冲积平原承载了全国 54% 的人口和 67% 的国内生产总值, 但与此同时, 该区域也面临着严峻的洪水风险威胁^[1]。合理的防洪规划和洪水风险评估对保护冲积平原地区起至关重要的作用, 其中, 河流阻力作为衡量河道行洪能力的关键水力学参数, 已成为治河防洪工程领域的研究热点^[2-4]。多年以来, 学术界对冲积河流中水流阻力的定量表征进行了广泛而深入的研究。然而, 水流阻力的物理图景高度复杂, 包括水流条件、床面形态、植被覆盖、人类活动在内的多种因素均对其有显著影响, 且不同的影响因子之间也存在着交互作用^[5-7]。当前尚无一套具备严格数学与物理基础的通用阻力计算方法。在此背景下, 阻力系数法作为一种通过单一参数表征河段整体水流阻力的传统方法, 在工程实践与学术研究中广受采用^[8]。

目前, 学术界在冲积河流的阻力系数计算方面已取得丰富的成果^[9-12], 成功通过曼宁系数(n)、谢才系数(C)、达西-魏斯巴赫系数(λ)等多种阻力系数实现了对河段水流阻力的整体描述。其中, 根据 ALAM 等^[13]的研究, 曼宁糙率系数被认为是最为重要的阻力系数。需要指出的是, 现有的阻力系数研

究主要是基于经验性或半经验性方法, 通过水槽试验或河道实测数据拟合出阻力计算的辅助关系式。然而, 此类方法得到的阻力表达式往往具有较强的局限性, 计算性能受限于公式建立时所依据的率定数据, 难以直接用于其他河流, 尤其是多沙河流的水流阻力计算。此外, 传统计算方法常采用试错法来计算有关水文参量, 过程繁琐, 进一步限制了其在实际工程中的应用^[14]。

近年来, 机器学习方法发展迅速^[15-18], 作为一种数据驱动的技术手段, 该方法通过优化模型的内部参数实现数据间隐含关系的深入挖掘, 能够在准确预测的同时实现合理外延。此外, 机器学习的计算需求集中于模型的训练阶段, 优化完毕的模型在实际应用过程中仅需执行单次前向计算, 这实际上将计算负担由实时预测阶段前置到离线训练阶段, 可以显著提升工作效率。目前, 在冲积河流阻力系数计算方面, 机器学习方法也已投入初步应用^[19-22]。

然而, 当前应用于阻力系数计算的人工智能方法主要局限于传统机器学习范畴, 其模型复杂度相对有限, 难以充分捕捉水流阻力与影响因素之间复杂的非线性关系, 在一定程度上制约了模型精度的提升。特别是在黄河等水沙情势复杂多沙河流中, 传统机器学习模型的预测能力与适应性面临挑战。除此以外, 现

有研究中模型输入特征的选取多局限于单一时空尺度, 尚未充分挖掘多源异构信息对模型泛化性能的潜在贡献, 限制了模型在不同流域与不同水文情势下的应用能力。与此同时, 气候变化与人类活动的持续影响使得水文系统呈现出显著的非平稳性特征, 传统模型所依赖的历史数据往往难以正确反映未来水沙关系的演变趋势。因此, 亟需探索已训练人工智能模型在异质性水沙情势下的跨情景适应能力, 以增强其在复杂河流环境中的适用性。

针对上述问题, 本文拟开展以下研究: 首先, 针对传统机器学习方法在预测能力上的局限, 引入深度森林模型作为核心方法, 以其集成式的层级结构和多粒度特征扫描机制提升模型对复杂非线性关系的建模能力。其次, 考虑到现有研究在输入特征集构建方面的不足, 本文在模型训练数据集的构造中将弗劳德数、体积含沙量等日尺度水沙因子与年径流量、年输沙量等水文要素相结合, 利用多尺度信息增强模型在不同流域和不同水文时期的预测能力。最后, 为系统评估所提出模型在非平稳水沙条件下的迁移能力, 本文围绕小浪底水利枢纽运行前后流域水沙格局的显著变化特征, 设计多组具代表性的对比试验工况, 考察模型在不同水沙情势下的表现。

1 研究区域与研究方法

1.1 研究区域与研究数据

黄河全长 5 400 km, 流域面积 75.2 万 km², 是中国第二大河。黄河来水主要由上游供给, 泥沙则主要来源于中、上游地区, 特别是黄土高原^[23]。以内蒙古托克托县双河镇和河南荥阳市桃花峪为分界点, 可将黄河分为上游、中游和下游。其中, 下游河段流经华北平原, 具有典型的冲积河流特征^[24]。

根据平面形态, 可将黄河下游进一步划分: 孟津至高村段属于典型的游荡型河道, 该段河道稳定性差, 河宽变化显著, 堤距范围为 5~20 km, 主槽宽度在 1~3 km 范围内波动, 纵比降为 0.17‰~0.26‰; 高村至艾山段呈现出过渡型特征, 河道纵比降减小至 0.12‰左右; 艾山至利津段则在人工控制下变成限制性弯曲河道, 河宽进一步收窄至 0.03~1.2 km, 河床纵比降维持在 0.1‰的水平^[25]。

游荡型和过渡型河段的存在导致黄河下游的阻力特性复杂多变, 使其成为检验阻力计算方法的优良天然实验室。本文即以黄河下游为研究对象, 从黄河水利委员会整理汇编的《黄河流域水文资料》中获取 1958—2021 年黄河下游花园口、夹河滩、高村、孙

口、艾山、泺口和利津 7 个水文站的 2000 多组实测数据。该数据集包含平均流量、造床流量、比降、含沙量、平均流速、河宽、平均水深、床沙中值粒径、曼宁糙率系数等多维参数, 在此基础上, 本文将系统开展水流阻力的计算研究。

1.2 传统公式

张红武等^[26]针对黄河水沙情势复杂、河床冲淤变形剧烈、水面形态多样、河道周界约束作用差异大等特点^[27-32], 开展黄河动床模型试验, 挖掘了 n 和 Fr 之间的关系。该研究同时参考张罗号公式^[29]和张有龄公式^[30-31], 综合考虑河床自动调整作用、沙粒表面摩擦及含沙量的影响, 建立了如下公式为

$$n = \frac{0.9D_{50}^{1/6}(\kappa/\kappa_0)^{1/5}}{A_{50}(0.1 + 1.85Fr)} \quad (1)$$

式中, D_{50} 为床沙中值粒径(m); A_{50} 为摩阻参数, 对黄河下游取 19; κ_0 为清水卡门常数; κ 为浑水卡门常数, $\kappa = \kappa_0[1 - 4.2\sqrt{S}(0.365 - S)]$ 。

张红武公式没有采用具体的水文实测数据率定, 只是采用黄河下游河道 $Fr = 0.5$ 、 $n = 0.01$ 的特征关系确定其待定系数, 目前在黄河具体工程应用上已经取得成功^[33]。此外, 本文还选用 LIMERINOS 公式^[34]作为西方经验公式的代表, 参与黄河水流阻力的计算。对应计算公式为

$$n = \frac{0.1129R^{1/6}}{1.16 + 2\ln(R/D_{84})} \quad (2)$$

式中, R 为水力半径(m); D_{84} 为粒径累计分布中 84% 的颗粒较之小的粒径(m)。

1.3 深度森林

深度森林是一种基于逐层处理、模型内特征变换等深度学习理念的人工智能算法^[35], 该方法主要由以下两个关键部分构成(见图 1)。

(1)级联森林。为实现“逐层处理”的理念, 深度森林采用了一种由多层树回归器构成的级联结构, 将每一层的回归器设计为随机森林或极端森林。在训练过程中, 每一层通过分析叶节点上的值分布来估计实例的值分布, 并将这些估计值在所有树之间进行平均, 生成一个回归向量, 并将其与原始特征向量拼接, 作为下一层的输入。

(2)多粒度扫描。为实现“模型内特征变换”, 深度森林采用多尺寸的滑动窗口对输入数据进行扫描, 并提取出相应的高维特征向量。这些向量被视为具有相应目标值的新实例, 并生成如“级联森林”部分所述的回归向量。通过向量间的拼接, 可以形成变换后的扩展特征。

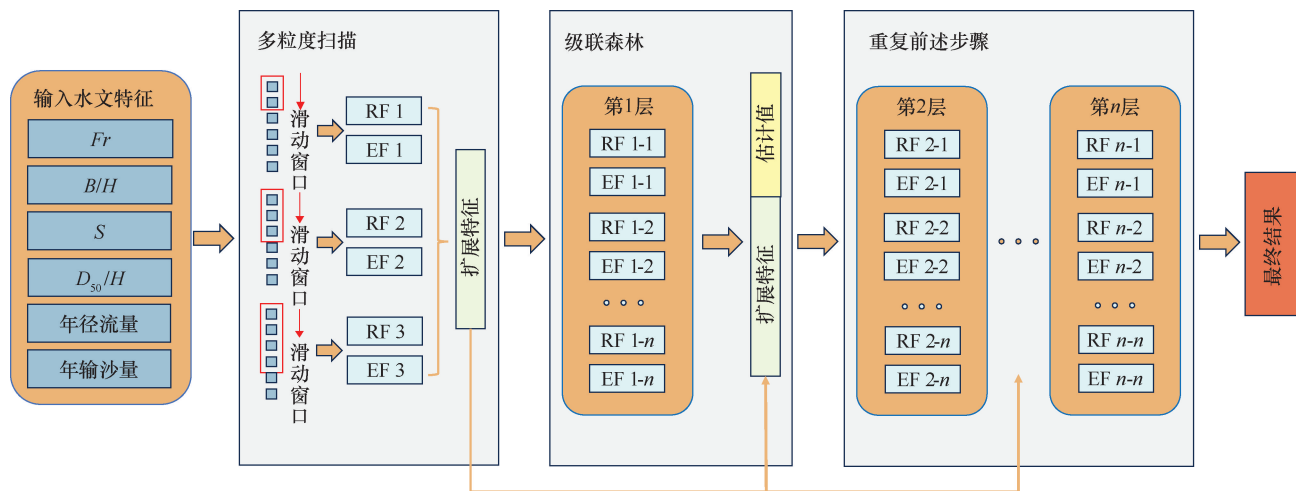


图 1 深度森林模型结构示意图

Fig. 1 Schematic diagram of deep forest model structure

1.4 模型输入特征集构建

为充分挖掘深度森林模型的性能,需审慎构建相应的模型输入特征集。本文综合考虑学界对天然河流曼宁糙率系数 n 的影响因素研究^[36-38],初步选取比降(J)、体积比含沙量(S)、床沙中值粒径(D_{50})、河宽(B)、断面平均水深(h)、雷诺数(Re)、断面平均流速(V)、断面平均流量(Q)等 8 个水文特征作为备选的特征输入。

为考察上述水文特征与目标输出(曼宁糙率系数 n)之间的相关关系,本文采用式(3)分别计算它们之间的互信息数,得到如图 2 所示的结果。可以发现, V 和 n 之间的相关性最高,互信息数达到 0.677。其余水文特征与 n 之间的互信息数均较低,不超过 0.45。进一步利用 Mann-Whitney U 检验开展的显著性测试表明,在 95% 的置信度下,上述 8 个水文特征均与 n 具有相关性。互信息数公式为

$$I(X, Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \left[\frac{p(x, y)}{p(x)p(y)} \right] \quad (3)$$

式中, $I(X, Y)$ 为随机变量 X 和 Y 的互信息数; $p(x)$ 为 X 的边际概率分布函数; $p(y)$ 为 Y 的边际概率分布函数; $p(x, y)$ 为 X 和 Y 的联合概率分布函数。

因此,从相关性的角度来说,这 8 个水文特征都应作为模型的输入。但相关研究^[39-41]显示,模型输入之间较强的相关性将导致模型参数估计的不稳定,从而影响模型的解释力和预测性能(即“多重共线性”问题)。观察图 2 可知,现有的 8 个水文特征之间,确有多对特征间的相关性较高(如 Q 和 V 等),所以,有必要对现有的输入特征集进行精简。根据各

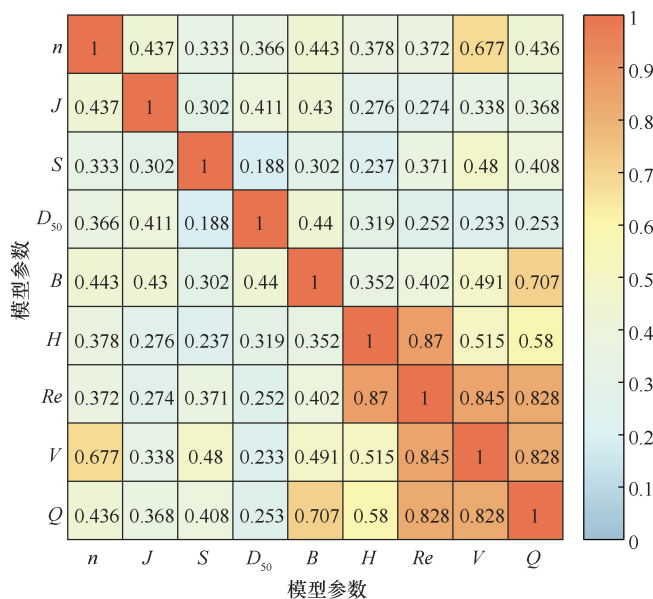


图 2 模型输入特征与曼宁糙率系数 n 之间互信息数

Fig. 2 Mutual information values between model input features and Manning's roughness coefficient n

特征的物理含义,本研究最终选取流速、水深、河宽、含沙量和代表粒径等 5 个基本特征作为模型输入。

为保证计算过程中物理意义的准确,本文将上述物理量进行无量纲化处理,分别得到弗劳德数(Fr)、宽深比(B/H)、体积含沙量(S)和径深比(D_{50}/H)等 4 个无量纲数。除此之外,年径流量和年输沙量也将作为模型输入,以提高模型泛化能力(具体分析见 2.2 节)。在开展模型训练之前,本研究将数据集的 80% 划分为训练集,其余 20% 划分为测试集,并对输入数据进行归一化处理。相关模型输入的统计指标如表 1 所列。

表 1 各输入水文特征的统计指标

Table 1 Statistical indicators of different input hydrological features

水文特征	计数	最大值	最小值	极差	均值	标准差	峰度	偏度
n	2 581	0.058	0.004	0.054	0.015	0.007	5.772	1.424
S	2 581	0.117	3.019×10^{-5}	0.117	0.008	0.014	27.646	4.277
D_{50}/H	2 581	1.898	0.007	1.891	0.106	0.139	41.430	4.811
B/H	2 581	4 197.531	17.197	4 180.334	352.814	427.353	22.970	3.912
Fr	2 581	0.811	0.056	0.755	0.355	0.128	2.392	0.157
年径流量/亿 m^3	2 581	973.100	20.000	953.100	312.524	152.477	4.945	1.152
年输沙量/亿 t	2 581	702.500	0.100	702.400	5.365	14.808	1 904.811	40.523

1.5 模型评价指标

本文采用如表 2 所列的评价指标评估各方法的性能。

表 2 评价指标

Table 2 Evaluation metrics

统计参数	表达式
均方根误差	$RMSE = \sqrt{\frac{\sum_{i=1}^N (P_i - O_i)^2}{N - 1}}$
纳什-萨特克里夫效率率	$NSE = 1 - \frac{\sum_{i=1}^N (P_i - O_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2}$
平均绝对误差	$MAE = \frac{1}{N} \sum_{i=1}^N P_i - O_i $
平均绝对百分比误差	$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{ P_i - O_i }{O_i} \times 100\%$

注: P 为计算值; O 为实测值。

2 模型验证结果

2.1 深度森林模型与其他方法的性能对比

根据系统性的比较研究^[42-44], 本研究选择随机森林、轻量级梯度提升机 (LightGBM) 和人工神经网络 (ANN) 等 3 种机器学习算法和张红武公式、LIMERINOS 公式等 2 种传统公式作为深度森林模型的对照, 基于黄河下游 7 个水文站的实测资料, 开展动床阻力计算对比研究, 结果如图 3 所示。

从图 3 看出, 随机森林、LightGBM 以及 ANN 等 3 种机器学习算法的拟合精度均较高, NSE 都接近或超过了 0.7。对于传统公式来说, 张红武公式的 NSE 达到 0.668, 其精度与机器学习算法较为接近, LIMERINOS 公式计算精度不理想, NSE 为 -0.685。深度森林作为一种深度学习算法, 其 NSE 为 0.785, 分别比随机森林、LightGBM 和 ANN 高出 0.77%、

0.38% 和 13.28%; $MAPE$ 为 14.618%, 分别比随机森林、LightGBM 和 ANN 低 2.48%、3.63% 和 16.03%, 说明该方法刻画非线性关系的能力相较于传统机器学习算法有较大优势。

前述分析聚焦于各方法在一般性的中水流量下的计算表现。然而, 对于黄河等多沙河流, 更需重点关注直接关乎防洪减灾问题的极端情势下的阻力规律。为此, 本研究设计以下 4 种典型极端工况进行深入分析: (a) 质量含沙量 $S_m > 100 \text{ kg/m}^3$ 的高含沙工况; (b) 糙率系数 $n > 0.03$ 的大糙率工况; (c) 流量 $Q > 5000 \text{ m}^3/\text{s}$ 的大流量工况; (d) 流量 $Q < 200 \text{ m}^3/\text{s}$ 的小流量工况。具体结果如图 4 所示。

由散点图看出, 在高含沙和大流量工况下, 河道的阻力系数较小, 均保持在 0.02 以下。在此情况下, 除 LIMERINOS 公式外, 本文所研究的各类阻力计算方法均展现出良好的计算精度, 相对误差控制在 33% 以内。特别是深度森林, 其结果点据基本分布在 45° 线附近, 表明预测结果与真实情况的高精度拟合。而在大糙率工况下, 各方法的计算精度均出现不同程度的下降。绝大部分结果点据分布在 45° 线下方, 表现出对极值系统性低估的现象, 且包括深度森林在内的 4 种人工智能算法部分结果的相对误差也超过了 33%。对于小流量工况, 各方法的计算精度介于前述两种情况之间, 可以得到合理的糙率计算结果。对比其他工况, LIMERINOS 公式在该工况下表现最优, 有效地将相对误差控制在 33% 以内。

值得一提的是, 在图 4 所展示的各类工况下, 张红武公式和人工智能算法之间的精度差异均较小, 在大糙率和小流量工况下, 张红武公式的计算精度甚至超越了部分机器学习模型。这一现象可以从训练数据的角度解释: 极端工况的出现频率低, 在训练集中占比少, 对模型的训练作用有限, 而人工智能模型难以完全消化极端现象背后的物理图景, 不能适宜特殊或稀遇洪水来临时河床摩阻规律的变化, 导致此类情形下的模拟结果不理想。

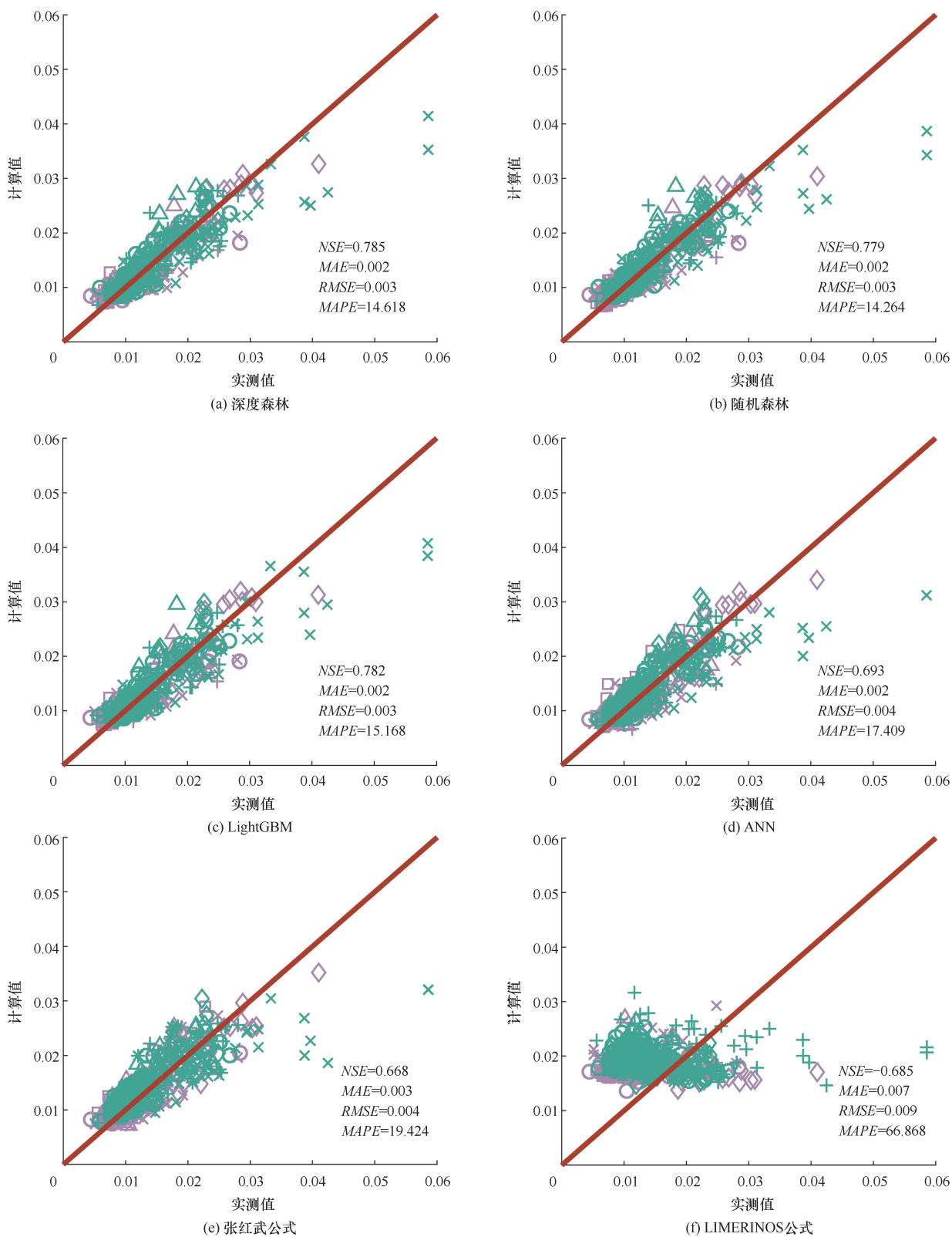


图3 各方法的计算结果散点图

Fig. 3 Scatter plots of calculation results from different methods

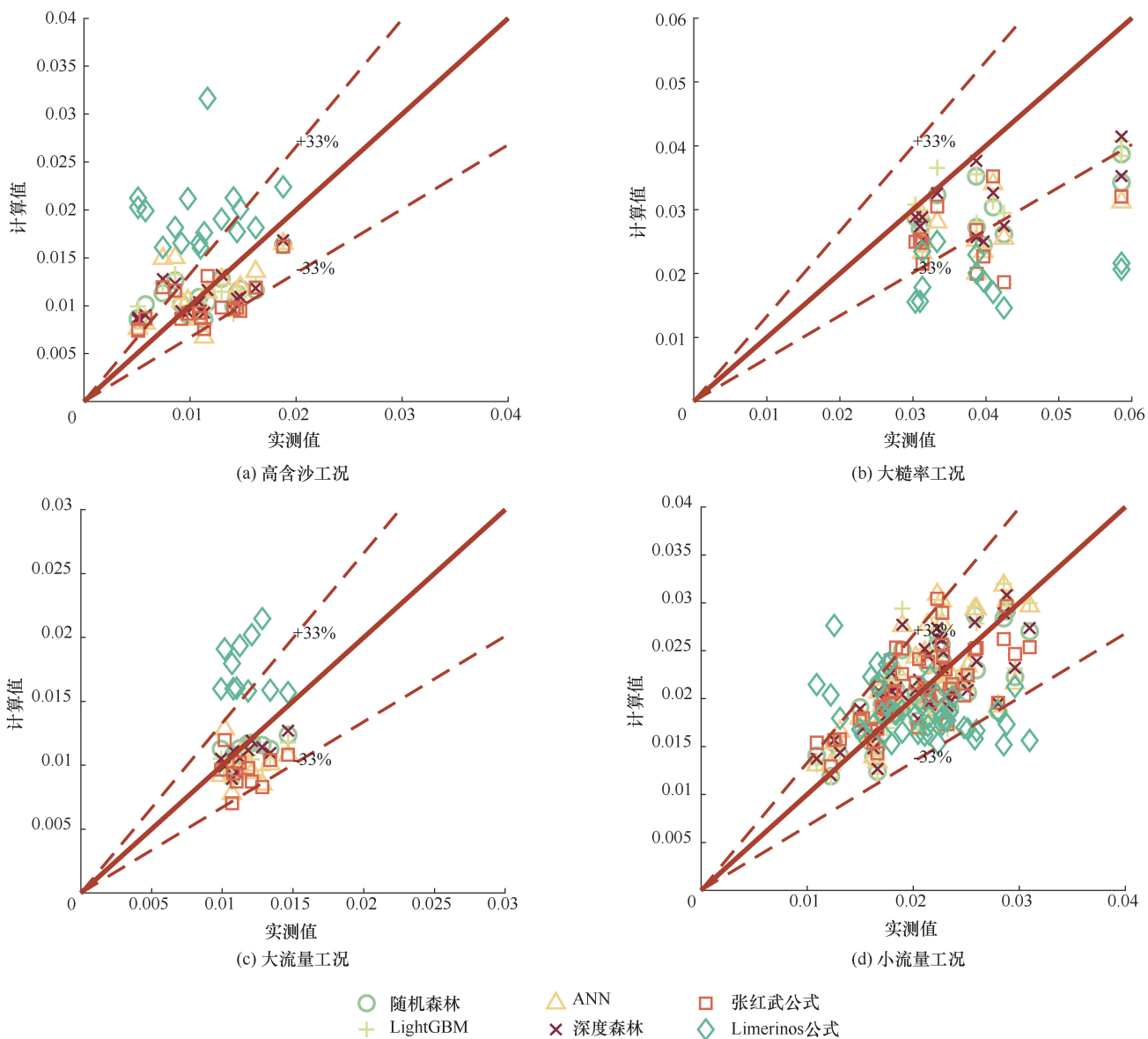


图4 各方法在极端情况下的计算结果散点图

Fig. 4 Scatter plots of calculation results from different methods under extreme conditions

附带指出,大糙率工况主要集中在黄河下游窄河段。该河段内的水文站主要位于两岸坝垛密布的区域,其综合糙率表现为河岸阻力而非动床阻力^[42]。这实际上是大糙率工况中各方法精度下降的主要原因。

2.2 泛化与非泛化模型的效果对比

根据水文模型区域化理论^[45],耦合不同空间尺度的水文特征有助于提升模型在不同流域内的预测精度和泛化能力。本文从时间维度上推广了该理论,在原有模型日尺度输入特征集的基础上,引入年尺度的水沙因子(年径流量和年输沙量)作为模型输入,以期提升现有深度森林模型在不同测站和不同水文时期内的模拟精度。为区别于已有的模型区域化理论,下文讨论中称这种引入了年径流量和年输沙量的模型为

泛化模型。

为考察泛化模型的空间泛化能力,本文对比了不同模型在黄河下游7个测站上的精度评价指标。从图5看出,相比于原有模型,泛化模型在各测站上的预测精度均有所提升。其中,提升幅度最大的是孙口站, NSE 从0.4586升至0.6230, $RMSE$ 从0.0032降至0.0027, MAE 从0.0022降至0.0020, $MAPE$ 从13.6426%降至12.7535%;相比之下,夹河滩站的优化幅度最小,其 NSE 从0.6259升至0.6308, $RMSE$ 从0.0029降至0.0028, MAE 从0.0023降至0.0022, $MAPE$ 从16.8753%降至15.5032%。

2000年一方面小浪底水库开始蓄水拦沙,另一方面受到流域泥沙治理和气候变化的影响,黄河下

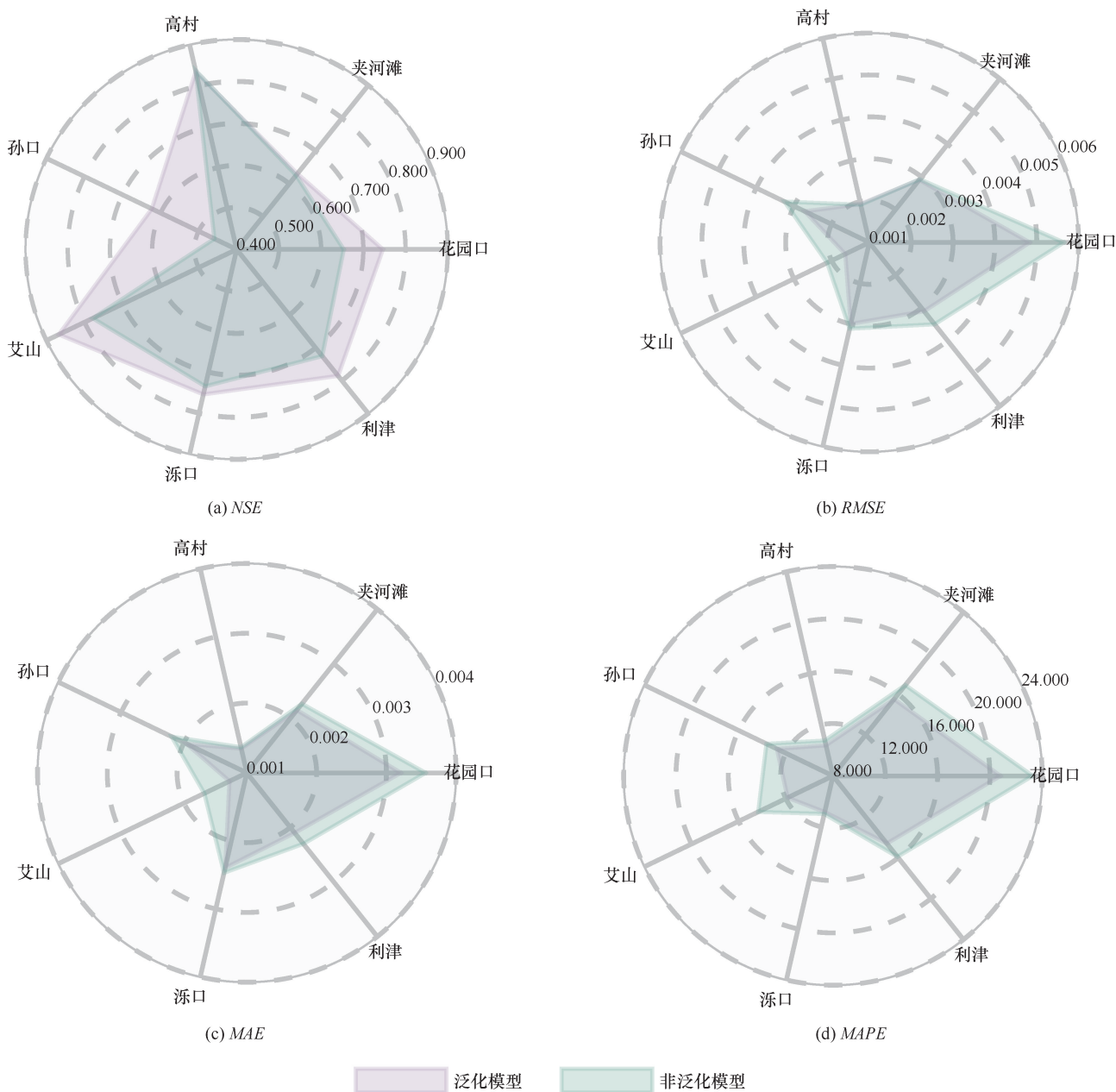


图 5 泛化与非泛化模型在各水文站的模拟精度对比

Fig. 5 Comparison of simulation accuracy between generalized and non-generalized models at different hydrological stations

游的水沙情势发生了显著的变化。本文以该年为时间节点, 将测试集划分为水库拦沙期和自然状况期等 2 个水文时期, 以进一步考察泛化模型的时间泛化能力。从图 6 看出, 在不同的水文时期内, 泛化模型的计算精度均有所提升。直方图分析表明, 泛化模型的计算结果分布更加集中于低误差区间内, 相比之下, 非泛化模型的计算结果在高误差区间分布更多。除此之外, 误差累积曲线显示, 泛化模型的累计误差分布普遍居左, 其对应的平均相对误差为 0.146 2, 低于非泛化模型的 0.159 1。

不同量级的年径流量和年输沙量过程, 其对应的

水流阻力关系势必不一致。因此, 本文绘制了如图 7 所示的光谱图以考察泛化模型在不同来水来沙条件下的预测精度。其中, RE_{diff} 定义为泛化模型相对误差和非泛化模型相对误差的差值, 其值为正表示相应条件下泛化模型表现劣于非泛化模型; 反之则表示相应条件下泛化模型表现优于非泛化模型。

由图 7(a) 看出, 在年径流量为 100 亿~500 亿 m^3 的区间内, RE_{diff} 的负值分布略多于正值分布, 即在平水年和枯水年的条件下, 泛化模型性能相比于非泛化模型有部分提升。而在年径流量大于 500 亿 m^3 的区间内, RE_{diff} 的负值分布显著多于正值分布。这表

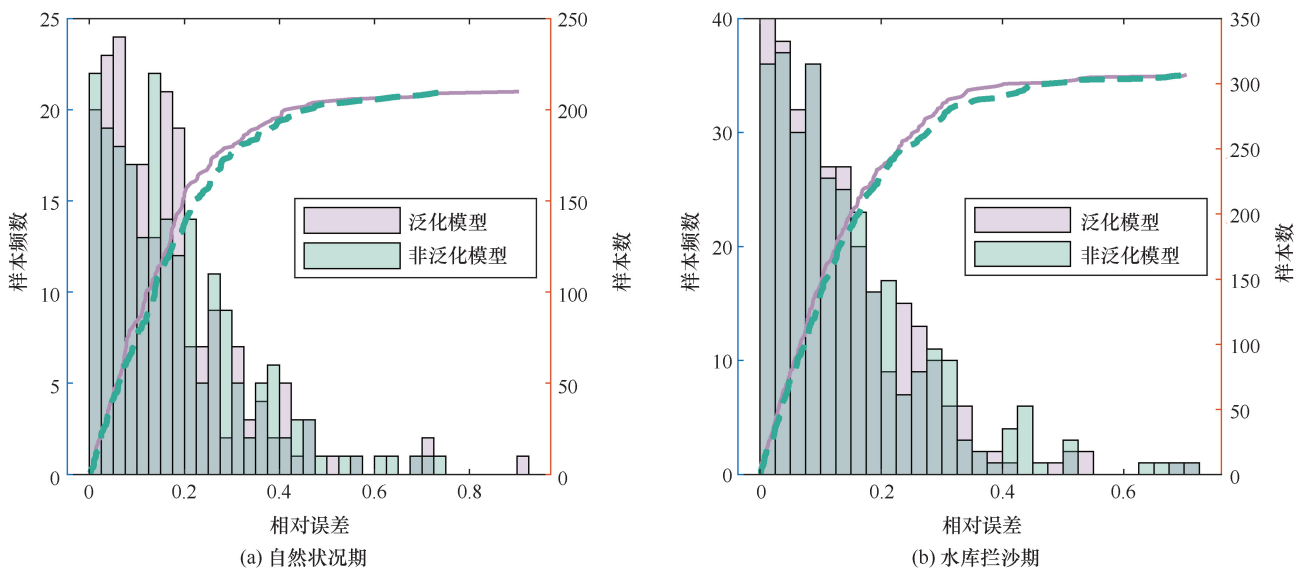


图6 泛化与非泛化模型在不同时期的模拟精度对比

Fig. 6 Comparison of simulation accuracy between generalized and non-generalized models during different periods

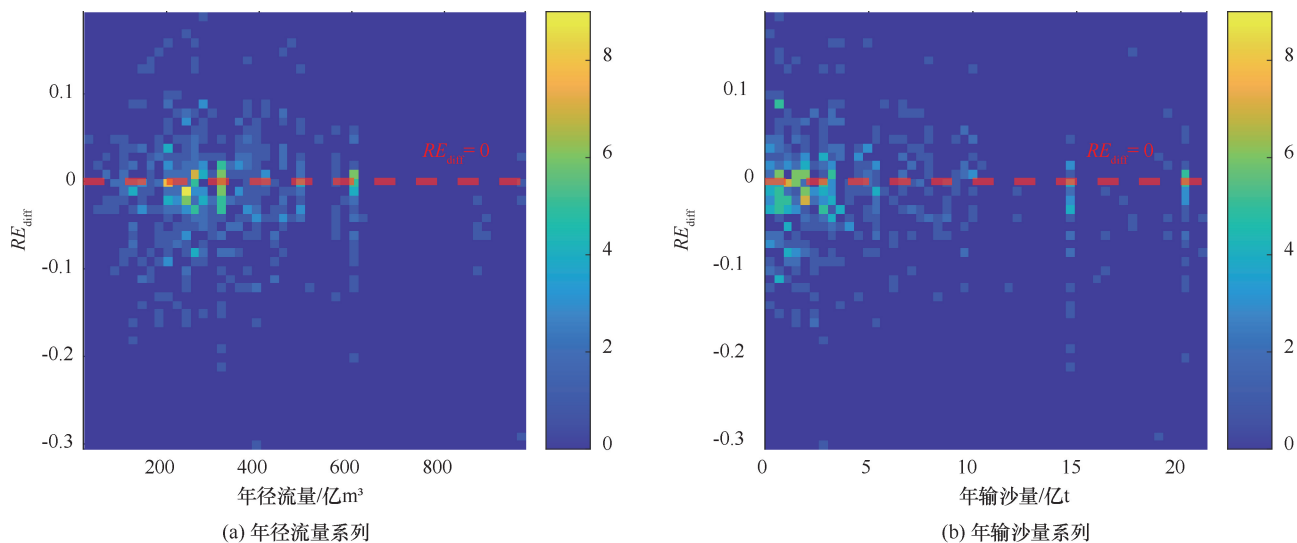


图7 不同年径流量和年输沙量量级对泛化与非泛化模型的影响

Fig. 7 Influence of different magnitudes of annual runoff and sediment load on generalized and non-generalized models

明年径流量和年输沙量等额外信息的辅助可有效提高泛化模型在丰水及极端来水条件下的预测能力。由图7(b)可知, RE_{diff} 的负值分布相较于正值分布在年输沙量为0~12亿t的区间内略有优势, 在年输沙量大于12亿t的区间内, 优势则更为明显。这种分布特征与年径流量情况较为相似。因此可以推论, 泛化模型在不同量级的来水来沙条件下, 均实现了对非泛化模型的性能提升, 并且性能优化的幅度随年径流量和年输沙量的增加而增加。

2.3 迁移学习能力分析

小浪底枢纽工程的建设、中上游水保水利工程的拦沙作用以及气候变化的影响显著改变了黄河下游的

水沙情势。以2000年为界, 该节点前后的水文系列所对应的统计指标差别显著, 具有明显的非平稳特征。通过20世纪资料训练的模型可能难以正确反映新时期的黄河阻力关系。为定量研究深度森林模型应对非平稳水沙条件的迁移学习能力, 本研究设计并开展了四组对比试验: (1) 自然状况期训练并模拟; (2) 水库拦沙期训练并模拟; (3) 全时期训练并模拟; (4) 自然状况期训练、水库拦沙期模拟。

图8的试验结果显示, 试验工况(1)~(3)的误差分布具有相似性, 均表现为左偏态的单峰分布, 概率密度峰值位于0.1附近, 且绝大多数测试样本的相对误差集中在-20%~20%的低误差区间。相比之下,

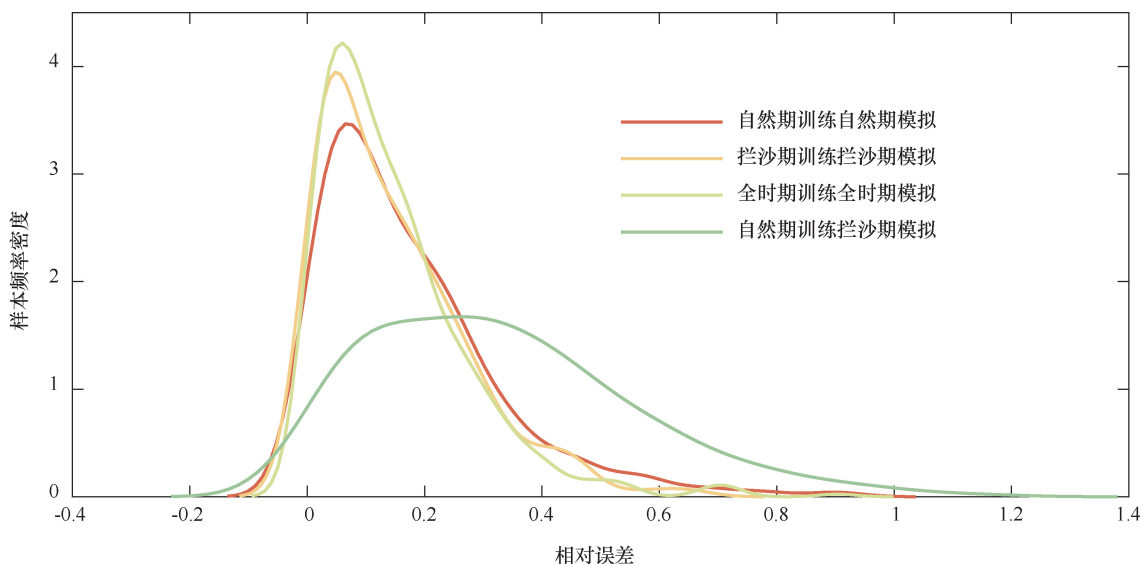


图 8 不同试验工况的相对误差分布

Fig. 8 Distribution of relative errors under different experiment conditions

表 3 迁移学习能力分析结果

Table 3 Analysis results of transfer learning capabilities

方法	评价指标	自然状况期训练 自然状况期模拟	水库拦沙期训练 水库拦沙期模拟	全时期训练 全时期模拟	自然状况期训练 水库拦沙期模拟
深度森林	<i>NSE</i>	0.762	0.743	0.785	0.272
	<i>RMSE</i>	0.003	0.004	0.003	0.006
	<i>MAE</i>	0.002	0.002	0.002	0.005
	<i>MAPE/%</i>	17.254	14.643	14.618	32.742
张红武公式	<i>NSE</i>	0.712	0.599	0.668	0.599
	<i>RMSE</i>	0.003	0.004	0.003	0.004
	<i>MAE</i>	0.002	0.003	0.004	0.003
	<i>MAPE/%</i>	21.387	19.355	19.424	19.355

注：表中的“训练”和“模拟”分别指的是利用给定水文时期的 $(Fr, B/H, S, D_{50}/H, n)$ 数据集进行训练和利用给定水文时期的 $(Fr, B/H, S, D_{50}/H)$ 输入组合进行 n 的计算。

试验工况(4)的误差分布表现出更低的峰度，且存在较多样本分布于相对误差超过40%的高误差区间，表明模型预测精度出现显著下降。与此同时，由表3看出，经过与测试集水沙条件一致的数据训练后，深度森林模型可以展现出优异的计算精度，试验工况(1)–(3)的*NSE*均超过了0.7。横向对比这三种工况可以发现，试验工况3的*NSE*较工况(1)和工况(2)分别提升了3.02%和5.65%。这一结果表明，采用包含不同水文时期数据的训练集所建立的模型，在各水文时期的计算精度均优于仅采用单一对应时期数据训练的模型。然而，在第四种迁移学习工况中，深度森林模型的*NSE*指标相比工况(3)降低了65.35%，出现了显著的性能下降。该现象可归因于模型训练集与测试集在水沙关系及其相应水流阻力表现等方面的一致性。即模型仅学习到自然状况下水沙条件对动床阻力的影响规律，而不能把握小浪底水库调水调沙

后黄河下游水沙情势下糙率的变化趋势，导致其对水库拦沙期的水沙条件下河床阻力预测能力不足。相比之下，基于基本物理图形建立的张红武公式因其不依赖数据率定的特性，在各项评价指标上均优于深度森林模型。

上述分析揭示了当前人工智能技术在水利科学应用中的关键短板。为此，在水利数据具有灾害实时数据极少、数据力学背景差异大等特殊性的客观条件下^[46]，如何提升模型在非平稳水沙条件下的迁移学习能力，将是未来防洪减灾最需攻克的难题，也是智慧水利建设长期的努力方向。同时，研究结果也强调了传统水沙理论的重要价值，即在数据受限或环境剧变条件下，具有正确物理背景的公式往往能提供更可靠的预测结果。因此，在推进深度学习技术的水利应用的同时，应持续深化水沙运动的机理研究。

3 结果讨论

目前, 针对黄河下游水流阻力的研究依然以传统经验性或半经验性的方法为主。举例而言, 李琳琪等^[27]基于相关性分析方法, 选取流速、水深、能坡、浑水涡团参数及床沙中值粒径作为公式主参数, 并通过大样本回归确定了各参数的指数, 最终建立的阻力公式的 *MAPE* 为 17.3%。侯琳等^[42]则从河流过程原理出发, 系统分析了多因素耦合作用机制, 最终构建了以水深、比降、含沙量及河道纵向稳定系数为核心变量的计算公式, 其 *MAPE* 降至 16.5%。此外, 麻妍妍等^[46]从工程实用性角度考虑, 建立了曼宁系数与弗劳德数之间的单变量经验关系, 在黄河下游测试数据集上也能表现出较好的预测性能, *RMSE* 为 0.004 9。可喜的是, 已有学者借助流速分布公式的理论突破^[48], 成功开展了黄河阻力计算。本文的深度森林模型对于黄河下游的阻力计算, 其计算精度相对于现有公式有较多提升。

而在机器学习水流阻力计算领域, 现有研究主要分为基于天然河流实测资料和基于水槽试验资料两大派别。前者大都采用日尺度水沙因子作为模型输入特征。例如, MILUKOW 等^[20]选取河道比降和相对粗糙度(即水力半径与特征粒径之比)作为输入变量; NIAZKAR 等^[19]则选择包含水力半径、弗劳德数以及颗粒剪应力与起动剪应力之比等在内的多维水沙特征进行输入。不难看出, 尽管不同研究采用的输入特征在形式上存在差异, 但其本质上均可溯源至流速、水深、河宽、含沙量及特征粒径等基础水沙因子的非线性组合与无量纲变换受制于水沙测验方法及其整编精度的影响^[49]。本文在实施无量纲化处理前所构建的输入特征空间, 可以完整包含上述各类参数。从模型表现来讲, MILUKOW 基于极限学习机构建的最优模型 *NSE* 为 0.78; 而 NIAZKAR 采用人工神经网络获得的 *NSE* 仅为 0.52, 二者均低于本文所采用的深度森林模型。

相比之下, 后者通常进一步引入了床面形态参数作为输入变量。举例而言, KARBASI 等^[21]将水力半径与沙波几何特征(波长、波宽等)的比值纳入模型输入特征集; 而 YARAHMADI 等^[43]则系统考虑了沙波高与水深之比、沙波高与特征粒径之比以及沙波高与沙波长之比等多元形态参数。从模型评价指标来看, KARBASI 采用扩展卡尔曼滤波优化的人工神经网络模型, 其最优 *RMSE* 为 0.004 1, 略高于本研究结果; 而 YARAHMADI 基于支持向量机构建的模型

则表现出优异的拟合能力, 其 *NSE* 达到 0.99。

这类研究最大缺陷是受制于河工模型相似律的影响^[31], 毕竟水槽试验量测的床面形态^[50]受水流挟沙能力^[51]、水流阻力特性^[52]直接影响, 与真实河流差距过大。另外, 值得注意的是, YARAHMADI 等采用的输入特征集将流速、水深、比降 3 个变量全部包含在内, 这是有待商榷的。因为所谓“实测曼宁糙率系数”实际上是利用曼宁公式通过实测的流速、水深和比降反算得到的, 当可以获得上述全部 3 个变量时, 研究本身将失去意义。从另一方面讲, 此时模型所需学习和拟合的也不再是复杂的水沙关系, 而仅仅是曼宁公式而已, 这对于机器学习来说并不是难事。

此外, 从上述分析可以推论, 在输入特征集中额外加入床面形态参数有助于模型性能的提升。如果资料符合实际, 这在基于水槽试验的研究中是可行的, 在水槽内床面形态相关的几何参数的非同步测量是可以实现的, 特别是部分研究中的沙波还是人为规则布设的。但是, 在天然河流, 特别是大型冲积河流中, 这种方法的应用则会受限: 一是床面形态的测绘十分困难, 黄河的观测精度更无法保证^[51], 二是难以用简单的参数合理描述复杂的天然河床形态。综上所述, 本文通过互信息数筛选并构造的无量纲化输入特征集是相对合理的, 通过该数据集训练的模型展示出较强的性能优势。

尽管深度森林模型在黄河下游不同测站以及不同水沙条件下均表现出较强的适应能力, 但由于未能专门考虑研究对象的物理背景, 在无资料区的推广应用仍面临不小的挑战。具体来说, 该模型需要包括流速、水深、河宽、河床颗粒粒径及体积含沙率在内的众多水文特征, 而在无测站的区域根本无法获取这些数据。事实上, 上述文献中的其他模型在应用于无资料区时亦存在类似困难。因此, 研发适用于无资料区域的水流阻力计算模型, 是本领域长期难以攻克的难题。

4 结论

冲积河流的综合阻力系数是衡量河道行洪能力的重要指标, 对于防洪规划工作有重要意义。然而, 目前的水流阻力计算方法依然局限于传统公式和机器学习模型, 其预测精度有限且泛化能力不足。本研究引入深度森林模型对黄河下游的曼宁系数进行计算, 并采用年径流量和年输沙量等水文特征对模型进行泛化。在此基础上, 本文设置多组对比试验, 考察所构建模型的迁移学习能力。得到的主要结论如下。

(1)模型性能比较研究表明,深度森林模型受黄河下游7个水文站丰富的洪水测验资料训练后,在阻力计算中展现出良好的精度,且相较于机器学习方法(随机森林、LightGBM、ANN)和传统公式(张红武公式、LIMERINOS公式),在NSE、RMSE、MAE和MAPE四项评价指标上均取得不同的优势。

(2)模型泛化机制研究发现,通过引入年径流量、年输沙量和宽深比(B/H)等时空特征参数,可在一定程度上体现水沙条件对洪水阻力的影响,有效提升模型的时空适应能力。对于不同量级的来水来沙条件,泛化模型均实现了性能的提升。

(3)模型迁移学习能力评估显示,当前深度森林模型在应对非平稳水文条件尤其是极端特殊洪水时,由于受训作用小而依然存在局限性,其精度低于不受训练资料制约、有明确物理背景的传统公式,表明在推进人工智能技术应用的大趋势下,必须持续深化传统水沙理论的研究,二者相辅相成才能全面解决复杂环境下的河流阻力计算问题。

参考文献 (References):

[1] JIANG R J, LU H, YANG K, et al. Substantial increase in future fluvial flood risk projected in China's major urban agglomerations [J]. *Communications Earth & Environment*, 2023, 4(1): 389.

[2] MOTAMEDI A, AFZALIMEHR H, SINGH V P. Estimation of friction factor in open channels [J]. *Journal of Hydrologic Engineering*, 2010, 15(3): 249-254.

[3] BELLOS V, NALBANTIS I, TSAKIRIS G. Friction modeling of flood flow simulations [J]. *Journal of Hydraulic Engineering*, 2018, 144(12): 04018073.

[4] 翟明杰, 焦彩虹, 庞慧霞, 等. 糙率及地面沉降对河道行洪能力影响的研究 [J]. *水利发展研究*, 2024, 24(11): 29-34.
ZHAI Mingjie, JIAO Caihong, PANG Huixia, et al. Discussion on the influence of roughness and land subsidence on flood discharge capacity of watercourse [J]. *Water Resources Development Research*, 2024, 24(11): 29-34.

[5] KUMAR S, KHUNTIA J R, KHATUA K K. Prediction of flow resistance in an open channel over movable beds using artificial neural network [J]. *Journal of Hydrologic Engineering*, 2021, 26(5): 04021015.

[6] STAUDT F, MULLARNEY J C, PILDITCH C A, et al. Effects of grain-size distribution and shape on sediment bed stability, near-bed flow and bed microstructure [J]. *Earth Surface Processes and Landforms*, 2019, 44(5): 1100-1116.

[7] 穆祥鹏, 张哲琦, 李军强, 等. 纳米胶在混凝土输水工程中的减阻降糙特性 [J]. *南水北调与水利科技(中英文)*, 2025, 23(2): 473-480.

MU X P, ZHANG Z Q, LI J Q, et al. Preliminary study on the drag reduction and roughness reduction properties of Nano-Rubber applications in concrete water conveyance engineering [J]. *South-to-North Water Transfers and Water Science & Technology*, 2025, 23(2): 473-480.

[8] LIU J Z, ZHANG J L, HUANG Z, et al. Prediction of movable bed resistance in open channels with dune forms incorporating the effects of flow separation and water temperature [J]. *Journal of Hydrology*, 2024, 634: 131073.

[9] RAUDKIVI A J. Analysis of resistance in fluvial channel [J]. *Journal of the Hydraulics Division*, 1967, 93(5): 73-84.

[10] YALIN S. On the average velocity of flow over a movable bed [J]. *La Houille Blanche*, 1964 (1): 45-51.

[11] VAN RIJN L C. Sediment transport, part III: bed forms and alluvial roughness [J]. *Journal of Hydraulic Engineering*, 1984, 110(12): 1733-1754.

[12] WHITE W R, PARIS E, BETTESS R, et al. The frictional characteristics of alluvial streams: A new approach [J]. *Proceedings of the Institution of Civil Engineers*, 1980, 69(3): 737-750.

[13] ALAM A M Z, KENNEDY J F. Friction factors for flow in sand-bed channels [J]. *Journal of the Hydraulics Division*, 1969, 95(6): 1973-1992.

[14] YEN B C. Open channel flow resistance [J]. *Journal of Hydraulic Engineering*, 2002, 128(1): 20-39.

[15] GUO S L, WEN Y H, ZHANG X Q, et al. Research on precipitation prediction based on a complete ensemble empirical mode decomposition with adaptive noise-long short-term memory coupled model [J]. *Water Supply*, 2022, 22(12): 9061-9072.

[16] FENG Z K, NIU W J, WAN X Y, et al. Hydrological time series forecasting via signal decomposition and twin support vector machine using cooperation search algorithm for parameter identification [J]. *Journal of Hydrology*, 2022, 612: 128213.

[17] ZHANG X Q, YANG Y. Suspended sediment concentration forecast based on CEEMDAN-GRU model [J]. *Water Supply*, 2020, 20(5): 1787-1798.

[18] ZHI W, APPLING A P, GOLDEN H E, et al. Deep learning for water quality [J]. *Nature Water*, 2024, 2(3): 228-241.

[19] NIAZKAR M, TALEBBEYDOKHTI N, AFZALI S H. Novel grain and form roughness estimator scheme incorporating artificial intelligence models [J]. *Water Resources Management*, 2019, 33: 757-773.

[20] MILUKOW H A, BINNS A D, ADAMOWSKI J, et al. Estimation of the Darcy-Weisbach friction factor for ungauged streams using Gene Expression Programming and Extreme Learning Machines [J]. *Journal of Hydrology*, 2019, 568: 311-321.

[21] KARBASI M, GHASEMIAN M, JAMEI M, et al. Developing extended and unscented Kalman Filter-based neural networks to

- predict cluster-induced roughness in gravel bed rivers [J]. *Water Resources Management*, 2024, 38(8): 3023-3048.
- [22] SAGHEBIAN S M, ROUSHANGAR K, OZGUR KIRCA V S, et al. Modeling total resistance and form resistance of movable bed channels via experimental data and a kernel-based approach [J]. *Journal of Hydroinformatics*, 2020, 22(3): 528-540.
- [23] XIA X H, DONG J W, WANG M H, et al. Effect of water-sediment regulation of the Xiaolangdi reservoir on the concentrations, characteristics, and fluxes of suspended sediment and organic carbon in the Yellow River [J]. *Science of the Total Environment*, 2016, 571: 487-497.
- [24] YANG R Y, PENG Y, ZHANG H W, et al. Estimation of flow resistance in natural rivers based on deep forest [J]. *Earth Science Informatics*, 2025, 18(1): 1-16.
- [25] 侯琳, 张红武, 李琳琪, 等. 黄河水流阻力及河床比降计算方法研究 [J]. *水利学报*, 2024, 55(9): 1020-1032.
- HOU L, ZHANG H W, LI L Q, et al. Study on calculation of flow resistance and riverbed slope of the Yellow River [J]. *Journal of Hydraulic Engineering*, 2024, 55(9): 1020-1032.
- [26] 张红武, 张罗号, 彭昊, 等. 冲积河流糙率由来与计算方法研究 [J]. *水利学报*, 2020, 51(7): 774-787.
- ZHANG H W, ZHANG L H, PENG H, et al. Research on cognition and calculation method of alluvial river roughness [J]. *Journal of Hydraulic Engineering*, 2020, 51(7): 774-787.
- [27] 李琳琪, 张红武, 侯琳, 等. 河流动床阻力影响参数分析及其计算研究 [J]. *水利学报*, 2023, 54(12): 1496-1506.
- LI L Q, ZHANG H W, HOU L, et al. Parameter analysis and calculation method of moving bed resistance [J]. *Journal of Hydraulic Engineering*, 2023, 54(12): 1496-1506.
- [28] 张罗号. 黄河河槽糙率异常原因及其解决途径 [J]. *水利学报*, 2012, 43(11): 1261-1270.
- ZHANG L H. Reasons for the abnormal channel roughness of the Yellow River and the solution to its problems [J]. *Journal of Hydraulic Engineering*, 2012, 43(11): 1261-1270.
- [29] 张罗号, 卜海磊, 田依林. 细沙河流床沙中径计算公式研究 [J]. *人民黄河*, 2011, 33(12): 38-41.
- ZHANG L H, BU H L, TIAN Y L. Study on the formula of medium size of fine sand of bed material [J]. *Yellow River*, 2011, 33(12): 38-41.
- [30] CHANG Y L. Laboratory investigation of flume traction and transportation [J]. *Transactions of the American Society of Civil Engineers*, 1939, 104(1): 1246-1284.
- [31] 张红武, 江恩惠, 白咏梅, 等. 黄河高含沙洪水模型的相似律 [M]. 郑州: 河南科学技术出版社, 1994.
- ZHANG H W, JIANG E H, BAI Y M, et al. The Scale Law of Model Similarity of Hyperconcentrated Flow [M]. Zhengzhou: Henan Science and Technology Press, 1994.
- [32] 彭昊, 黄河清, 张红武, 等. 体现水流与沙粒摩擦阻力相对作用的糙率公式 [J]. *泥沙研究*, 2022, 47(2): 1-8.
- PENG H, HUANG H Q, ZHANG H W, et al. A roughness formula based on the relativity of flow intensity and grain roughness [J]. *Journal of Sediment Research*, 2022, 47(2): 1-8.
- [33] HOU L, ZHANG H W, LI L Q, et al. A method for calculating water demand for sediment transport based on the principles of river dynamics [J]. *Water*, 2023, 15(19): 3514.
- [34] LIMERINOS J T. Determination of the Manning Coefficient from Measured Bed Roughness in Natural Channels [M]. Washington, D. C.: US Government Printing Office, 1970.
- [35] ZHOU Z H, FENG J. Deep forest [J]. *National Science Review*, 2019, 6(1): 74-86.
- [36] DAVID G C L, WOHL E, YOCHUM S E, et al. Controls on spatial variations in flow resistance along steep mountain streams [J]. *Water Resources Research*, 2010, 46(3): 1-21.
- [37] WANG Z Y, QI P, MELCHING C S. Fluvial hydraulics of hyperconcentrated floods in Chinese rivers [J]. *Earth Surface Processes and Landforms*, 2009, 34(7): 981-993.
- [38] DAVIES T R, SUTHERLAND A J. Resistance to flow past deformable boundaries [J]. *Earth Surface Processes*, 1980, 5(2): 175-179.
- [39] GARG A, TAI K. Comparison of statistical and machine learning methods in modelling of data with multicollinearity [J]. *International Journal of Modelling, Identification and Control*, 2013, 18(4): 295-312.
- [40] CHAN J Y L, LEOW S M H, BEA K T, et al. Mitigating the multicollinearity problem and its machine learning approach: a review [J]. *Mathematics*, 2022, 10(8): 1283.
- [41] FARRAR D E, GLAUBER R R. Multicollinearity in regression analysis: The problem revisited [J]. *The Review of Economic and Statistics*, 1967, 49(1): 92-107.
- [42] 侯琳, 张红武. 黄河下游河道摩阻计算方法研究 [J]. *水利学报*, 2025, 56(1): 63-72.
- HOU L, ZHANG H W. Research on the calculation method of frictional resistance in the lower reaches of the Yellow River [J]. *Journal of Hydraulic Engineering*, 2025, 56(1): 63-72.
- [43] YARAHMADI M B, PARSAIE A, SHAFAI-BEJESTAN M, et al. Estimation of Manning roughness coefficient in alluvial rivers with bed forms using soft computing models [J]. *Water Resources Management*, 2023, 37(9): 3563-3584.
- [44] MIR A A, PATEL M. Machine learning approaches for adequate prediction of flow resistance in alluvial channels with bedforms [J]. *Water Science & Technology*, 2024, 89(2): 290-318.
- [45] 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.

- [46] 张红武, 赵晨苏, 张罗号, 等. 内蒙古自治区水利行业数字孪生总体构架方案[R]. 北京: 清华大学, 2024.
ZHANG H W, ZHAO C S, ZHANG L H, et al. Overall Framework for Digital Twins in the Water Resources Sector of the Inner Mongolia Autonomous Region[R]. Beijing: Tsinghua University, 2024.
- [47] 麻妍妍, 夏军强, 张晓雷. 现有动床阻力计算公式验证与比较[J]. 武汉大学学报(工学版), 2017, 50(4): 481-486.
MA Y Y, XIA J Q, ZHANG X L. Verification and comparison of flow resistance formulae over movable beds [J]. Engineering Journal of Wuhan University, 2017, 50(4): 481-486.
- [48] 杨润祎, 张红武. 基于涡团模式的紊流流速与阻力系数的统一规律[J]. 中国科学: 物理学 力学 天文学, 2026, 56(3): 234711.
YANG R Y, ZHANG H W. Unified law of turbulent flow velocity and drag coefficient based on turbulent eddy model[J]. Scientia Sinica (Physica, Mechanica & Astronomica), 2026, 56(3): 234711.
- [49] 杨润祎, 张红武, 朱金峰. 河流泥沙颗粒分析规程修编对悬移质级配分布测算的影响研究[J]. 水利学报, 2025, 56(3): 317-327.
YANG R Y, ZHANG H W, ZHU J F. Effect of the revision of technical standard for determination of sediment particle size on the estimation of particle size distribution [J]. Journal of Hydraulic Engineering, 2025, 56(3): 317-327.
- [50] 张红武, 李琳琪, 彭昊, 等. 基于流域高质量发展目标的黄河相关问题研究[J]. 水利水电技术(中英文), 2021, 52(12): 60-68.
ZHANG Hongwu, LI Linqi, PENG Hao, et al. Study on related problems of the Yellow River based on the goal of high-quality development of the basin [J]. Water Resources and Hydropower Engineering, 2021, 52(12): 60-68.
- [51] 张红武, 张清. 黄河水流挟沙力的计算公式[J]. 人民黄河, 1992, 14(11): 7-9.
ZHANG H W, ZHANG Q. Formula of sediment carrying capacity of the Yellow River[J]. Yellow River, 1992, 14(11): 7-9.
- [52] 张罗号, 卜海磊. 黄河下游糙率奇小的原因分析[J]. 人民黄河, 2010, 32(12): 84-86.
ZHANG L H, BU H L. Analysis of the extremely low roughness coefficient in the lower reaches of the Yellow River [J]. Yellow River, 2010, 32(12): 84-86.

(责任编辑 王海锋)