

基于 Sentinel-2 与机器学习的河流流量估算研究 ——以唐乃亥水文站为例

尹文杰^{1,2}, 王雪蕾^{1,2}, 王晨^{1,2}, 王航^{1,2}, 黄才声³, 赵瑞雪^{1,2},
孟凡乐^{1,2}, 刘锦绣³

1. 生态环境部卫星环境应用中心, 北京 100094;
2. 生态环境部卫星遥感重点实验室, 北京 100094;
3. 中国地质大学(北京) 信息工程学院, 北京 100083

摘要: 河流流量是水文循环过程的关键变量, 在洪水预警、水资源调度和生态环境管理中具有重要意义。本文选取黄河源区的唐乃亥水文站为试点, 提出一种融合卫星遥感和机器学习方法的河流流量监测技术。首先利用哨兵2号遥感影像提取河流的水面宽度, 联合全球陆面数据同化系统模拟的气象水文变量(蒸散发、土壤湿度、温度、陆地水储量和径流量)作为6种驱动因子, 分别构建了基于4种统计方法(线性函数、幂函数、指数函数和多项式函数)和4种机器学习方法(XGBoost、Random Forest、LightGBM和CatBoost)的流量监测模型, 评估不同模型监测结果间的差异并利用夏普利加性解释SHAP(Shapley Additive Explanations)方法量化不同驱动因子的重要性。结果表明: 4种统计方法中, 多项式函数模型在测试期的效果优于其他3种模型; 相比于传统统计方法, 机器学习方法在精度和稳定性方面表现出显著优势, 决定系数(R^2)提高了46.15%, 均方根误差(RMSE)和平均绝对误差(MAE)分别减少了54.61%和55.65%。Random Forest模型在测试期的模拟效果优于其他3种模型, 其 R^2 、NSE、RMSE和MAE分别为0.96, 0.89, 172.81 m³/s和147.33 m³/s。SHAP方法表明水面宽度在流量监测模型中具有最显著的贡献(189.02), 其次为土壤湿度(145.11)和温度(97.41)。本研究证实了联合卫星遥感和机器学习方法在复杂地形与资料匮乏区域开展高精度流量估算的可行性和优越性。

关键词: 卫星遥感, 机器学习, 统计模型, SHAP方法, 流量监测, 唐乃亥水文站

中图分类号: TP701/P2

引用格式: 尹文杰, 王雪蕾, 王晨, 王航, 黄才声, 赵瑞雪, 孟凡乐, 刘锦绣. 2026. 基于 Sentinel-2 与机器学习的河流流量估算研究——以唐乃亥水文站为例. 遥感学报, 30(7): 2070-2082

Yin W J, Wang X L, Wang C, Wang H, Huang C S, Zhao R X, Meng F L and Liu J X. 2026. River discharge estimation by integrating Sentinel-2 and machine learning: A case study of the Tangnaihai Hydrometry Station. National Remote Sensing Bulletin, 30(7): 2070-2082 [DOI: 10.11834/jrs.20265294]

1 引言

流量是水文学和水资源管理中的关键要素, 是进行洪水预警、水资源调度以及气候变化研究的基础数据 (Crochemore 等, 2020)。传统地面水文站的监测方法存在站点分布稀疏和数据获取成本高的局限性, 尤其在一些地形复杂或偏远地区, 难以满足精细化水资源管理的需求 (Papa 等, 2023)。因此, 探索新的流量监测技术成为当前研

究的热点。

卫星遥感技术具有监测范围广、时空分辨率高的优势, 为河流流量监测提供了新的数据源和方法 (Filippucci 等, 2022)。例如, 利用光学卫星的水体指数及雷达卫星的后向散射特性进行水体提取, 进一步结合数字高程模型和水力学模型可实现河流流量的遥感监测 (杨胜天 等, 2021; 史宜梦 等, 2022)。遥感技术虽然克服了地面监测的局限性, 但在复杂水文条件下, 结果精度会受到

收稿日期: 2025-08-23; 预印本: 2025-12-03

基金项目: 国家自然科学基金(编号:U23A2015)

第一作者简介: 尹文杰, 研究方向为水生态环境遥感监测。E-mail: wjyin1991@163.com

通信作者简介: 黄才声, 研究方向为水生态环境遥感监测。E-mail: 1250778545@qq.com

遥感影像的分辨率和水体提取误差等因素的影响(孙文超等, 2024)。机器学习方法可以准确地模拟流量与多种驱动因素间的复杂关系, 为处理复杂的水文数据和模型优化提供了新思路(汪静平等, 2023)。Hao和Bai(2023)利用支持向量回归SVR(Support Vector Regression)、梯度提升XGBoost(Xtreme Gradient Boosting)和长短期记忆神经网络LSTM(Long Short-Term Memory)3种机器学习方法, 开展了浙江省鳌江山区流域的日尺度径流监测。结果表明LSTM模型的纳什效率系数(NSE)为0.74, 而SVR和XGBoost的NSE均低于0.4。Madhushani等(2024)以气候数据(降水、气温)、地形指数和土地利用等因素为驱动因子, 联合XGBoost和直方图梯度提升HGB(Histogram-based Gradient Boosting)等机器学习方法开展美国怀俄明州山脉非站点化流域的流量模拟研究。研究表明XGBoost的决定系数为0.856, 在所有模型中表现最好, 并揭示了各变量对流量变化的非线性贡献。

利用机器学习算法对卫星遥感提取的水体宽度和实测流量数据进行拟合, 建立流量反演模型, 可实现高效、准确的流量监测。Sahoo等(2022)将基于MODIS和Landsat影像提取的河道宽度与实测河流流量进行联合分析, 在印度Brahmani河流域内构建了3种机器学习模型用于日尺度流量反演。结果表明SVR模型的纳什效率系数和克林—古普塔效率系数均超过0.85; Wang等(2023)基于哨兵1影像提取的水体面积和Landsat提取的改进型归一化差异水体指数等遥感参数, 利用SVR模型在中国雅砻江流域甘孜站上游河段建立了3种输入组合的流量反演模型。结果表明利用多种遥感水面宽与机器学习融合的方法, 可以实现河流流量的高精度估算。相比于Landsat和哨兵1影像, 哨兵2影像提取河流水面宽度具有更高的空间分辨率和水体敏感性, 通过联合哨兵2影像提取的水面宽度和机器学习算法可有效提高流量监测的准确性, 目前该方面研究较少。

鉴于此, 本文以黄河源区唐乃亥水文站为研究对象, 利用哨兵2号影像提取河流的水面宽度, 结合全球陆面数据同化系统(GLDAS v2.2)提供的陆地水储量变化、降水量、土壤湿度、蒸散发和径流量等驱动因子, 探索联合遥感技术和机器学习的方法开展河流流量监测研究, 以期评估融合多源遥感数据的机器学习模型在流量监测中的

应用潜力, 为复杂地形和数据匮乏地区的流量监测提供科学的技术方案。

2 研究区与数据源

2.1 研究区介绍

唐乃亥水文站位于中国青海省海南藏族自治州共和县与贵德县交界处, 地处青藏高原东北部, 是黄河干流上游的重要控制站之一, 地理坐标 $35^{\circ}30'N$ 、 $100^{\circ}30'E$ (图1)。该水文站主要控制黄河上游流域, 涵盖巴颜喀拉山脉及其周边高寒山地, 流域面积约为 $1.2 \times 10^5 \text{ km}^2$ 。流域内广泛分布有多年冻土、草地、湿地及冰川, 是青藏高原重要的生态屏障区, 也是黄河上游重要的水源涵养区(钱云平等, 2004; 白小晶等, 2024)。该区域属于典型的高原大陆性气候类型, 区域内水系较为发达, 年降水量为250—800 mm, 主要集中在夏季(6—9月), 蒸发强烈且生态环境脆弱(章斯腾等, 2021)。

2.2 数据来源

2.2.1 遥感影像

利用GEE(Google Earth Engine)云平台提供的哨兵2(Sentinel-2) Level-1C多光谱影像, 时间序列为2013年4月4日至2024年4月4日。哨兵2系列是“哥白尼计划”下的重要光学观测卫星, 由双星Sentinel-2A与Sentinel-2B构成, 单星重访周期为10 d, 双星联合观测下可达5 d的重访频率, 具有较高的时空分辨能力, 适合地表水体监测(乐颖等, 2024)。哨兵2搭载的多光谱成像仪(MSI)覆盖13个波段, 本论文主要使用绿色波段(B3)和短波红外波段(B11)计算水体指数用于提取研究区内河流的水面宽度。

2.2.2 驱动数据

联合全球陆面数据同化系统GLDAS v2.2(Global Land Data Assimilation System)提供的变量和哨兵2提取的河流宽度数据用作流量监测模型的驱动因子。GLDAS由美国国家航空航天局NASA(National Aeronautics and Space Administration)和美国国家海洋和大气管理局NOAA(National Oceanic and Atmospheric Administration)联合开发(disc.gsfc.nasa.gov[2025-08-23]), 融合了地面实测、遥

感反演与数值模式模拟结果，广泛用于水文模拟与气候影响研究。选取时空分辨率分别为日尺度、0.25°的GLDAS v2.2数据产品，产品名称为GLDAS Catchment Land Surface Model L4 daily 0.25×0.25 degree GRACE-DA1 V2.2（GLDAS_CLSM025_DA1_D 2.2），时间跨度为2013年4月4日至2024年4月4日。这些变量从物理水文过程角度出发，涵盖了

影响河流流量的输入与输出（蒸散发）、存储（陆地水储量和土壤湿度）、响应（径流和河流水面宽）及调节因子（温度）等因素，能够有效表征流域水文状态的时空变化特征（Ji和Yuan，2020）。GLDAS_CLSM025_DA1_D 2.2数据产品的具体参数如表1所示。

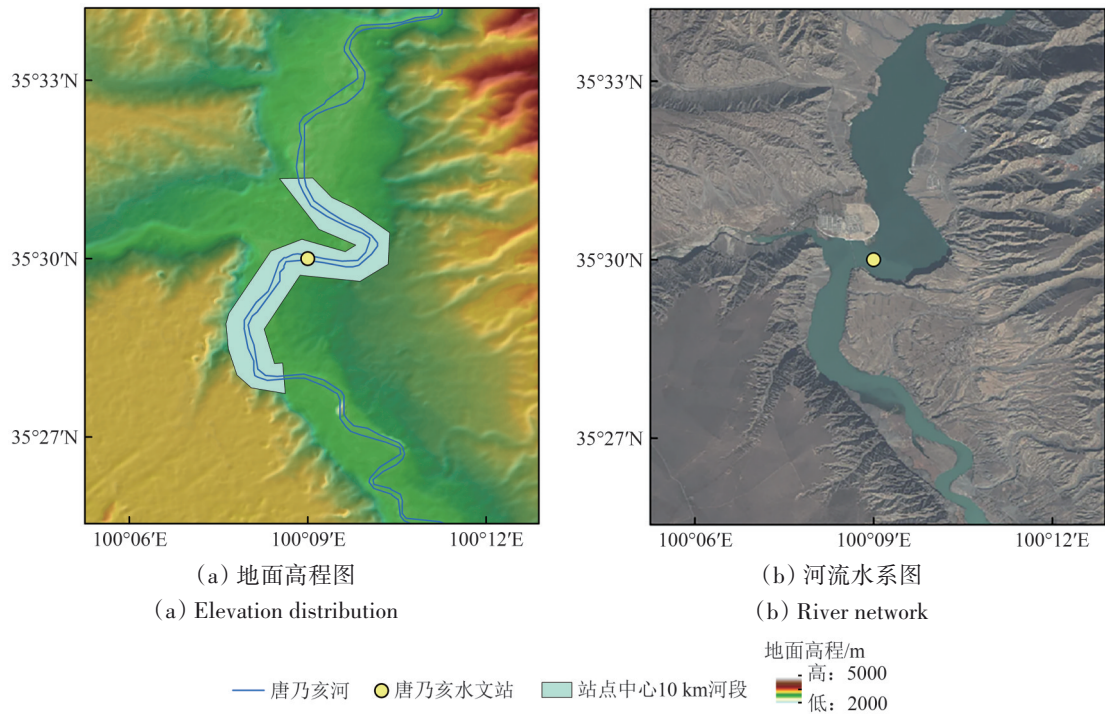


图1 唐乃亥水文站概况
Fig. 1 General situation of the Tangnaihai station

表1 GLDAS_CLSM025_DA1_D 2.2 产品参数

Table1 Parameters of GLDAS_CLSM025_da1_D 2.2 product					
变量	产品中对应名称	单位	含义	时间分辨率/d	空间分辨率
TWSA	TWS_tavg	mm	陆地水储量	1	0.25°×0.25°
Temp	AvgSurfT_tavg	K	平均地表温度	1	0.25°×0.25°
Runoff	Qs_tavg /Qsb_tavg	kg/(m ² s)	地表径流与地下水径流,两者总和即为径流	1	0.25°×0.25°
SM	SoilMoist_RZ_tavg	kg/m ²	土壤湿度	1	0.25°×0.25°
Evap	Evap_tavg	kg/(m ² s)	蒸散发	1	0.25°×0.25°

2.2.3 实测数据

实测数据来自唐乃亥水文站的逐日流量观测数据，数据包括水位（单位：m）和流量（单位：m³/s）两项主要水文要素。数据时间范围为2013年4月4日至2024年4月4日，覆盖了长时间序列的水文信息，具有较强的代表性和连续性。实测数据由中国水利部信息中心提供，属于国家级水文

监测数据，已被广泛应用于流域水文研究、水资源管理和生态环境评估等领域。

3 研究方法

本研究的技术路线详见图2。主要包含4个步骤：（1）获取卫星影像数据、实测流量数据、GLDASv2.2模型数据以及地面高程数据；（2）采用改进的水体指数（MNDWI）以及阈值分割法提

取水体，进而计算水面宽度；(3) 构建基于统计方法的流量模型和基于机器学习的流量模型，并划分相应的训练集、验证集、测试集；(4) 构建评估

指标，并采用SHAP (Shapley Additive Explanations) 方法对不同变量进行重要性评估。

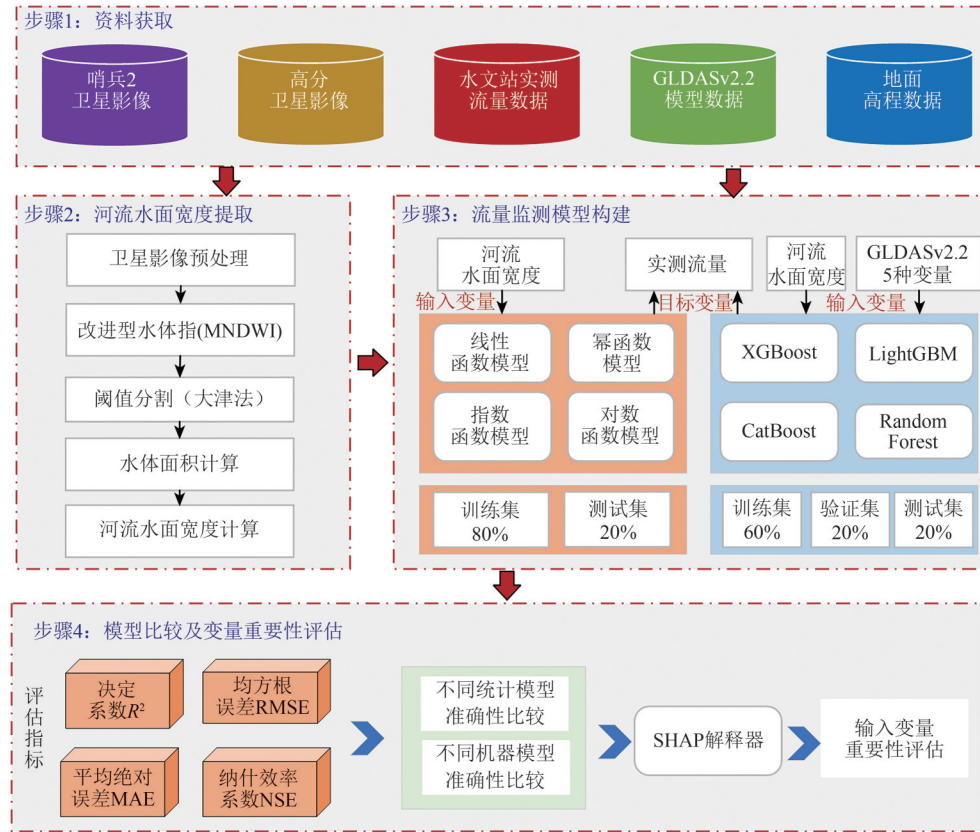


图2 本研究技术路线图

Fig. 2 Technical roadmap of this study

3.1 基于遥感影像的水体提取方法

采用哨兵2多光谱遥感影像，结合改进型归一化差异水体指数 (MNDWI) 与大津法 (Otsu's method) 对站点中心 10 km 河段进行自动水体提取，并以水体面积与河段长度的比值估算河流水面宽度。具体过程包括影像预处理、水体指数计算、阈值分割、水体掩膜生成、水体面积计算以及水面宽度估算等步骤。

(1) MNDWI 计算：采用 MNDWI 以增强水体与其他地物（特别是建筑物、植被等）的光谱区分度。其计算公式如下 (徐涵秋, 2005)：

$$\text{MNDWI} = \frac{B_{\text{Green}} - B_{\text{SWIR}}}{B_{\text{Green}} + B_{\text{SWIR}}} \quad (1)$$

式中， B_{Green} 为绿光波段的地表反射率， B_{SWIR} 为短波红外波段的地表反射率。MNDWI 值越大，表示该像元为水体的可能性越高。

(2) 阈值分割 (大津法)：为实现自适应地提

取水体区域，引入 Otsu 法对 MNDWI 图像进行自动阈值分割。该方法通过最大化类间方差自动确定最优阈值 T ，将像元划分为水体与非水体两类 (Ren 等, 2024)，最终选取使类间方差最大的 T 作为分割阈值，生成水体掩膜，计算公式如下：

$$\text{WaterMask}(x, y) = \begin{cases} 1, \text{MNDWI}(x, y) \geq T \\ 0, \text{MNDWI}(x, y) < T \end{cases} \quad (2)$$

式中， $\text{WaterMask}(x, y)$ 表示坐标 (x, y) 处的水体二值掩膜，1 表示水体，0 表示非水体。

(3) 水体面积计算：对分割后水体掩膜中的水体像元进行面积统计，水体总面积 A (单位： m^2) 计算如下：

$$A = \sum_{(x, y) \in \text{WaterMask} = 1} \text{PixelArea}(x, y) \quad (3)$$

式中， $\text{PixelArea}(x, y)$ 为影像中每个像元对应的实际地面面积。

(4) 河流水面宽度估算：假设研究区内河段

长度为 L (单位: m), 则该河段的平均水面宽度 W 计算如下:

$$W = \frac{A}{L} \quad (4)$$

式中, A 为水体面积, m^2 ; L 为河道中心线长度, m , 由 Google Earth 中采用距离标尺测量得出; W 为估算的平均河面宽度, m 。该方法适用于河道形态较为规则、无多段分流的中小尺度河流区域。

3.2 基于统计方法的流量监测

将河流水面宽度和实测流量作为自变量和因变量, 分别利用线性函数、幂函数、指数函数、多项式函数4种方法建立河流水面宽—河流流量之间的映射关系, 通过训练集数据拟合、测试集数据验证的方法构建流量监测模型 (方升等, 2021), 具体计算公式如下:

$$y = ax + b \quad (5)$$

$$y = a \cdot x^b \quad (6)$$

$$y = a \cdot e^{bx} \quad (7)$$

$$y = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \quad (8)$$

模型采用的数据集为2013年4月4日到2024年4月4日共450个数据, 4种统计模型将数据按4:1比例划分为训练集和测试集, 具体地, 2013年4月4日至2023年5月7日为训练集 (80%), 2023年5月15日至2024年4月4日为测试集 (20%)。

3.3 基于机器学习方法的流量监测

采用4种集成学习模型对河流流量进行监测, 分别包括RF (XGBoost、Random Forest)、LightGBM和CatBoost。各模型以一组具有物理意义的驱动因子 (河流水面宽度、陆地水储量、温度、土壤湿度、径流量和蒸散发) 为输入变量, 按3:1:1的比例将数据划分训练集、验证集和测试集, 通过训练—验证—测试流程构建流量响应的机器学习框架。

XGBoost 是一种基于梯度提升决策树 GBDT (Gradient Boosting Decision Tree) 框架的增强型机器学习方法, 它通过迭代训练多个回归树来最小化误差。与传统 GBDT 相比, XGBoost 引入了正则化项以防止过拟合, 尤其适用于构建高精度的流量或径流监测模型 (苏林鹏等, 2024)。Random Forest 是一种基于 Bagging 策略的集成学习算法, 它通过构建大量相互独立的决策树, 并对每棵树的输出进行回归或分类以提高模型的整体精度和

稳定性 (兰小机等, 2024), 尤其适用于在变量数量多且具有强噪声背景的流域水文数据中具有较高的适应性。LightGBM 是一种基于梯度提升框架的高效机器学习算法, 通过采用基于直方图的决策树构建方式显著降低了内存消耗和训练时间 (Dai等, 2023)。在水文建模中, LightGBM 能够快速处理包含大量气候驱动因子的复杂数据集, 适合用于高频率、高分辨率的流量监测。CatBoost 是一种专为类别特征优化设计的梯度提升算法, 通过使用“目标编码”与有序 Boosting 技术来高效处理类别型变量。CatBoost 尤其适合处理带有多个类别型或离散型气象、水文特征的时序数据 (Samat等, 2020)。

3.4 SHAP值解释算法

由于多数机器学习模型属于黑箱模型, 虽然具备良好的性能, 但其内部决策机制难以理解, 限制了模型在科学研究和决策支持中的应用 (莫锦英等, 2025; 刘东岳等, 2026)。为解决这一问题, 本文引入了SHAP值解释算法对模型进行可解释性分析。SHAP由Lundberg和Lee (2017) 提出, 基于博弈论中的Shapley值原理, 能够量化每个特征在给定样本上的边际贡献, 从而揭示模型监测的依据。SHAP假设模型模拟值可表示为特征贡献的线性组合:

$$f(\mathbf{x}) = \phi_0 + \sum_{i=1}^M \phi_i \quad (9)$$

式中, $f(\mathbf{x})$ 为模型对样本 $\mathbf{x}$ 的输出; ϕ_0 为模型在所有特征缺失时的基准输出 (通常为训练集的均值); ϕ_i 表示特征 i 对该监测结果的贡献, 即其Shapley值。Shapley值的计算公式如下 (Li, 2022):

$$\phi_i = \sum_{S \subseteq M \setminus \{i\}} \frac{|S|!(M - |S| - 1)!}{M!} [f_{S \cup \{i\}}(x_{S \cup \{i\}}) - f_S(x_S)] \quad (10)$$

式中, S 为不包含特征 i 的特征子集; $f_S(x_S)$ 表示仅使用特征子集 S 时的模型模拟结果。Shapley值反映了特征在各种组合中加入时带来的边际变化, 并对所有组合情况进行加权平均。其正负号表示该特征对模型结果的推动方向 (正值表示推动增大, 负值表示推动减小), 数值大小则代表其影响的强弱 (Van den Broeck等, 2022)。通过SHAP, 可以实现对单一样本或整个模型的全局解释, 为复杂模型的可信性、可视化和变量重要性分析提

供理论支持和实践依据。

3.5 精度评价

选取决定系数 R^2 (coefficient of determination)、均方根误差 RMSE (Root Mean Square Error) 和平均绝对误差 MAE (Mean Absolute Error) 作为主要评价指标。评估不同模型效果间的差异。 R^2 用于衡量模型对观测数据的拟合优度，越接近 1 表示模型越能解释目标变量的变化。RMSE 衡量的是模拟值与真实值之间偏差的标准差，其值越小表示模型结果与真实值越接近。与 RMSE 相比，MAE 对异常值的敏感性较低，更能反映整体误差水平的稳定性 (李昕潼等, 2023; 石锦等, 2019)，计算方法如下：

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (11)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (12)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (13)$$

式中， y_i 为第 i 个样本的真实值； $\hat{y}_i$ 为第 i 个样本的模拟值； $\bar{y}$ 为所有真实值的平均值， n 为样本总数。

4 结果与分析

4.1 4种不同统计方法的评估

图 3 显示了 2013 年 4 月—2024 年 4 月不同统计方法与实测流量值间的时间序列曲线。可见各统计方法在整体趋势上基本能反映实测流量的变化特征，2016 年—2020 年的丰水期能够有效捕捉流量峰值。其中多项式函数在高流量阶段表现最优，拟合曲线与实测值吻合度较高，幂函数与指数函数在中低流量阶段表现稳定；然而，线性函数偏离较大，甚至出现了负流量的情况，拟合能力相对较弱。总之，在模型测试期，多项式函数、幂函数与指数函数的结果相近，趋势上与实测流量保持一致，但是数值上仍有一定的差别。

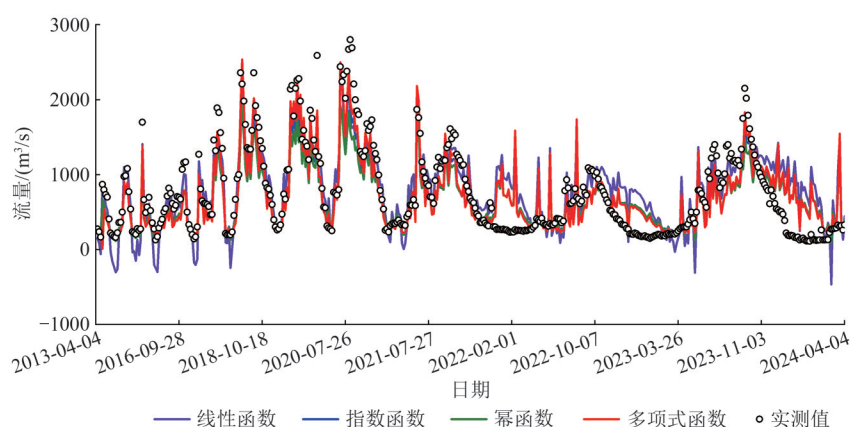


图 3 2013 年—2024 年不同统计方法河流流量模拟模拟与实测流量的时间序列

Fig. 3 Time series of simulated river flow from different statistical methods and in-situ measurements during 2013—2024

图 4 为固定测试期不同统计方法模拟河流流量与实测流量间的结果对比。可见，4 种模型均呈现“低流量高估、高流量低估”的系统偏差。其中当实测流量低于 $1000 \text{ m}^3/\text{s}$ 时，模拟值普遍位于 1:1 线之上；而当实测流量超过 $1000 \text{ m}^3/\text{s}$ 时，模拟值又明显落于 1:1 线之下。相比较而言，线性函数的 R^2 为 0.61，NSE 值仅为 0.17，表明其对流量变异的解释能力较弱，MAE 和 RMSE 均高于其他模型；指数函数与幂函数拟合的 R^2 均达到 0.65 以上，NSE 为 0.42，表明模型性能明显改善，且 MAE 和 RMSE 均显著降低；多项式函数的综合性能表现最好， R^2

达到了 0.67，误差指标 (MAE: $319.01 \text{ m}^3/\text{s}$, RMSE: $393.14 \text{ m}^3/\text{s}$) 均低于其他模型。

图 5 展示了 4 种统计模型在训练期与测试期上的性能比较，NMAE 和 NRMSE 分别表示指标 MAE 和 RMSE 归一化后的结果。考虑到 R^2 越大表示模型拟合效果越好，而 NRMSE 和 NMAE 越小表示误差越小，为便于在同一雷达图中直观对比，采用 $1-\text{NRMSE}$ 和 $1-\text{NMAE}$ 形式进行指标标准化处理，使得所有指标均为“值越大表示模型性能越优”。

模型训练期 (2013 年 4 月—2023 年 3 月)：多项式函数的 R^2 达到 0.79 以上，优于指数函数 (0.77)、

幂函数 (0.71) 和线性函数 (0.70), 其他系数相差不大 (0.91—0.93); 在模型测试期 (2023 年 3 月—2024 年 4 月), 幂函数拟合性能有所改善, 与多项式拟合、指数拟合的系数接近, 只有线性

拟合的 R^2 较低。整体来看, 模型训练期的性能明显优于测试期, 且多项式模型在拟合精度和稳健性方面均优于其他统计模型。

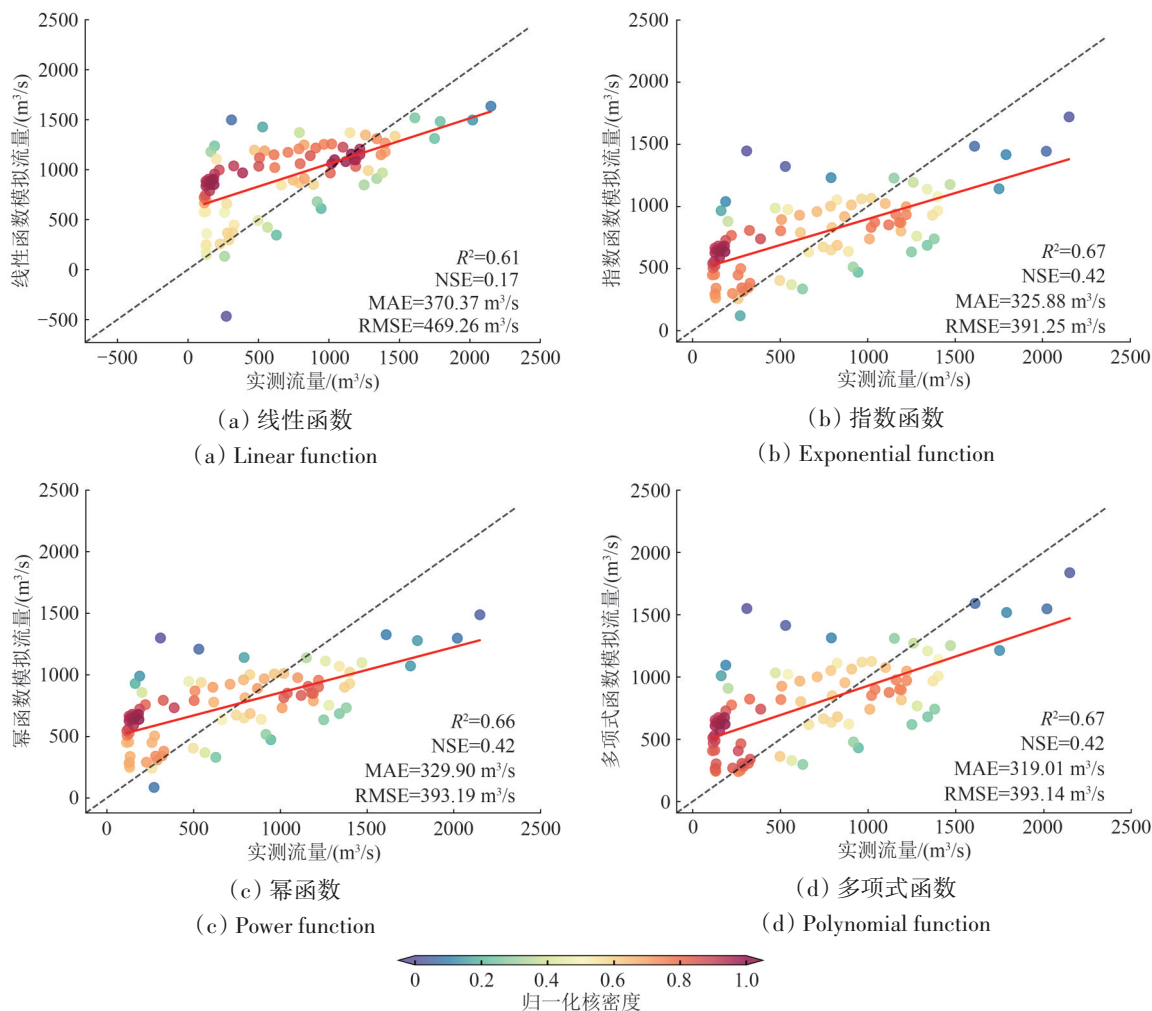


图4 2023年5月—2024年4月模型测试期4种不同统计方法河流量模拟结果与实测流量的拟合图
Fig. 4 Fitting plots of simulated and observed river flow based on four statistical methods during the testing period May 2023 to April 2024

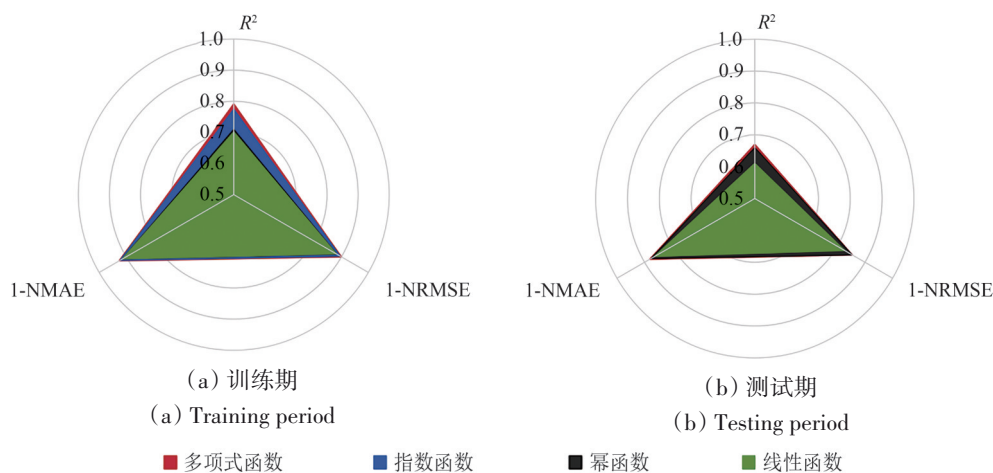


图5 4种统计模型效果参数雷达图
Fig. 5 Radar charts of performance metrics based on four statistical models

4.2 4种不同机器学习方法的效果评估

图 6 为 2013 年 4 月—2024 年 4 月期间，基于 4 种机器学习模型河流流量模拟结果与实测流量间的时间序列曲线。可见 4 种模型均较好地捕捉到了流量的变化规律，尤其是在汛期，模拟值能够较为准确地跟随实测值的峰值波动。其中在河流流量较大的阶段（2017 年—2021 年），4 种模型对实测流量的拟合效果较为一致，体现出较强的模拟

能力；然而，在部分年份（如 2018 年和 2020 年），各模型在流量峰值的捕捉上存在一定偏差，可能与极端事件或观测误差有关；在模型测试期，所有模型均表现出较高的一致性和精度，表明其对未来流量变化具有良好的泛化能力。整体而言，机器学习方法在河流流量模拟中展现出良好的适应性，明显优于统计模型结果的精度。

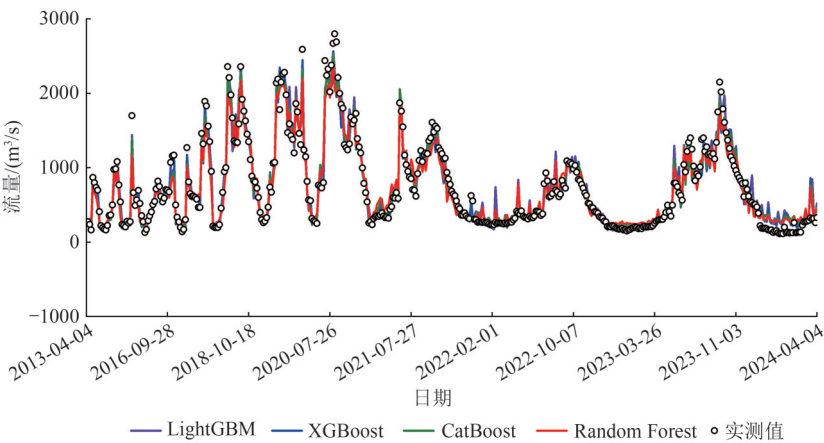
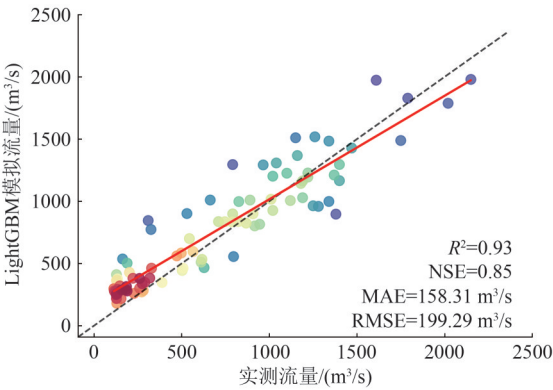


图 6 2013 年—2024 年不同机器学习方法模拟结果与实测流量的时间序列图

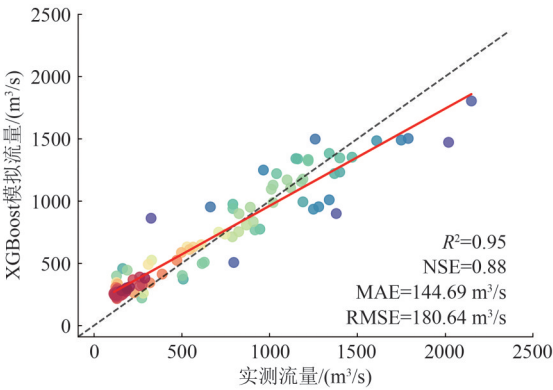
Fig. 6 Time series graph of simulated river flow from different machine learning methods and in-situ measurements during 2013—2024

图 7 表示模型测试期，4 种机器学习模型模拟结果与实测流量之间的拟合图。与实测流量相比，LightGBM 模型的 R^2 和 NSE 最低，MAE 和 RMSE 均高于其他模型，表明该模型在误差控制方面相对较弱。相比之下，XGBoost、CatBoost 和 Random Forest 模型的评估指标表现相对接近，拟合效果更加优越。其中，Random Forest 的 R^2 和 NSE 最高，

分别为 0.96 和 0.89，RMSE 最低，仅为 172.81 m^3/s ，显示出其在整体拟合精度和误差控制方面的优势；XGBoost 的 MAE 最低，表明其在减少小误差方面具有更强的表现能力。尽管各模型均能有效模拟河流流量过程，但综合评估结果显示，Random Forest 模型在监测流量的结果精度与稳定性方面更具优势。



(a) 轻量梯度提升机
(a) LightGBM



(b) 极端梯度提升
(b) XGBoost

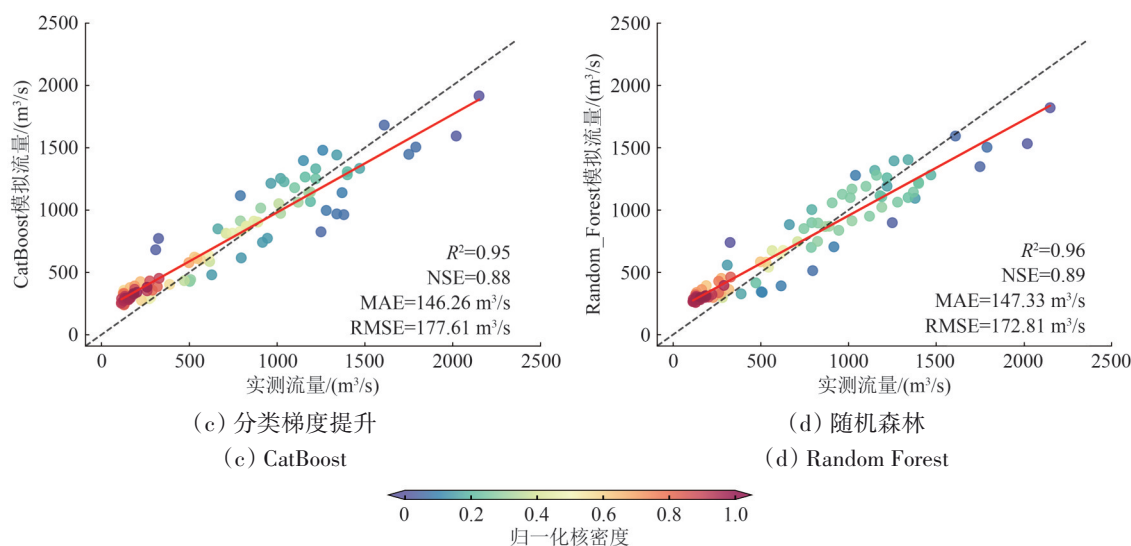


图7 模型测试期(2023年5月—2024年4月)4种机器学习模型模拟结果与实测流量间的拟合图
Fig. 7 Fitting plots of simulated and observed discharge based on four machine learning methods during the testing period (May 2023–April 2024)

为进一步评估4种机器学习模型模拟效果的差异,图8分别展示了模型训练期、验证期和测试期4种机器学习模型不同指标的对比。在训练期,各个模型的指标均在0.95以上,其中XGBoost模型3个指标均优于其他模型,说明其在拟合阶段具有较强的学习能力和对复杂非线性关系的刻画能力;验证期间,4种模型的性能差距趋于收敛,整体表现相对稳定,均保持较高的拟合精度;测试期不

同模型的NMAE和NRMSE指标较为接近,但Random Forest的 R^2 最高(0.96),其次是CatBoost(0.95)、XGBoost(0.95)和LightGBM(0.93)。整体来看,4种机器学习模型均表现出较强的拟合能力,所有模型的决定系数 R^2 均超过0.93,其他指标均在0.85以上,明显优于统计模型,证实了联合机器学习方法和卫星遥感技术在流量监测中具有良好的精度和稳定性。

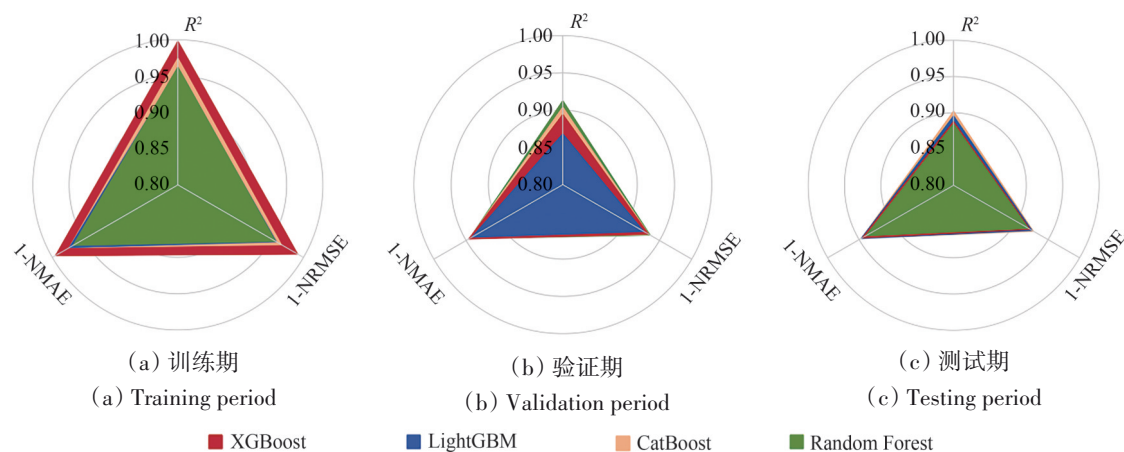


图8 4种机器学习模型效果参数雷达图
Fig. 8 Radar charts of performance metrics for the four machine-learning models

4.3 驱动变量重要性评估

图9表示不同驱动变量的平均SHAP值分布,反映了各变量对模型输出的影响强度和方向。特征值的颜色反映了原始变量数值的高低,右侧呈

红色表示变量值较高,左侧蓝色表示变量值较低。从图9中可以看出,水面宽度是流量变化最主要的控制因子,其SHAP值分布范围广且高值对应正向SHAP值,说明较大的河宽会正向推动流量模拟结

果。此外，水面宽度的平均 SHAP 绝对值最大 (189.02)，说明其在流量变化模拟中具有最显著的贡献，其次为土壤湿度 (145.11) 和温度 (97.41)，这些气象水文因子在流量形成过程中起到重要调节作用。相较之下，陆地水储量与地表径流的贡献相对较低，对流量模拟的影响较小。值得注意的是，作为流量直接物理组成部分的径流量，其贡献度最低这一现象初看有悖于直觉。我们分析认为，这主要源于 GLDAS 径流量数据与站点观测之间的尺度差异。GLDAS 数据相对粗粝的空间分辨率 (0.25°) 导致其与实测流量间的相关性较弱，难以精确捕捉唐乃亥站点所在位置的实际径流过

程。同时，径流量与贡献度最高的水面宽度之间相关性也较弱，表明二者信息重叠有限，水面宽度并未简单地“替代”径流量的信息。这反而说明，在本研究的数据框架下，遥感直接提取的水面宽度作为对河道水力状态响应更直接、观测不确定性更低的信号，被机器学习模型优先作为核心特征；而 GLDAS 径流量则因其代表性局限，所提供的增量信息相对有限，从而导致其统计贡献度较低。这一发现凸显了在融合多源数据时，数据本身的代表性与质量对模型解释结果具有决定性影响。

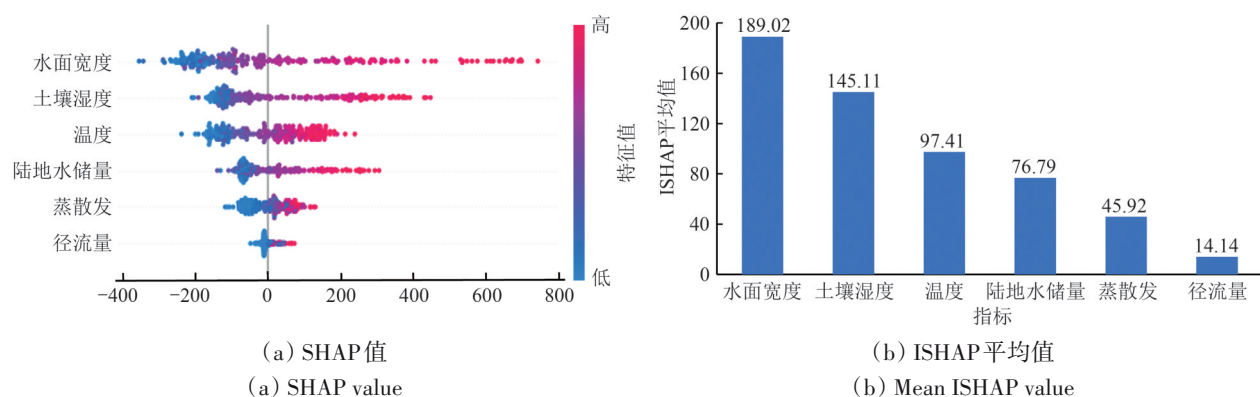


图9 不同驱动因素的重要性评估
Fig. 9 Importance of different driving factors

4.4 局限性及展望

本论文利用卫星遥感技术提取的河流水面宽度，联合统计方法和机器学习模型构建了河流水面宽度与河流流量间的线性和非线性联系，在唐乃亥水文站实现了河流流量的长时序监测，并取得了较好的反演效果，但仍存在以下局限性：

(1) 哨兵2影像10 m的空间分辨率在低水位时期对河道边界的细微变化捕捉能力有限，可能影响低流量阶段的估算精度；同时，光学影像易受云层覆盖影响，导致部分时段数据缺失，对水文序列的连续性构成挑战。

(2) 本文所采用的机器学习模型虽在精度上优于传统统计方法，但缺乏对水文物理过程（如质量守恒、动量传输等）的显式表达，因此在机制解释性和外推适用性方面仍存在不足。

针对上述问题，后续研究可从以下方面加以改进：一方面通过融合哨兵1号雷达遥感数据，弥

补光学影像在云雨条件下的观测盲区，提升数据获取的连续性与稳定性；另一方面，可探索将水文物理约束嵌入机器学习模型之中，发展耦合物理机制与数据驱动的混合建模框架，以增强模型的泛化能力与机理透明度。

5 结 论

本文以黄河源区唐乃亥水文站为研究区，系统对比了4种传统统计回归模型与4种机器学习模型在融合哨兵2遥感水面宽度及GLDAS气象水文变量进行流量估算中的表现。基于结果分析与模型解释，得出以下主要结论：

(1) 传统统计模型不适用于高精度流量估算。线性、幂、指数及多项式函数均存在“低流量高估、高流量低估”的系统性偏差，且对极端流量响应迟钝。其中多项式函数虽在统计模型中表现最优（测试期 $R^2=0.67$ ），但整体精度仍无法满足实际监测需求。因此，在复杂地形与数据稀缺流域，

不建议单独使用此类纯经验回归模型进行流量估算。

(2) 机器学习方法显著优于统计模型,且随机森林(Random Forest)为该研究区的最适模型。与最优统计模型相比,机器学习使 R^2 平均提升46.15%,RMSE和MAE分别降低54.61%和55.65%。4种集成学习中,随机森林在测试期综合性能最佳($R^2=0.96$, $NSE=0.89$, $RMSE=172.81\text{ m}^3/\text{s}$),兼具高精度与强泛化能力;XGBoost在极小误差控制方面略优,但整体稳定性稍逊。因此,推荐将随机森林作为基于多源遥感数据的日尺度流量监测首选模型。

(3) 遥感提取的水面宽度是流量估算的最关键驱动因子,GLDAS再分析径流量的贡献可忽略。SHAP分析表明,水面宽度的平均贡献值(189.02)远高于土壤湿度(145.11)和温度(97.41),而GLDAS径流量的贡献值仅14.14。这说明:1)卫星遥感直接观测的河道形态信息对流量具有主导性指示作用;2)粗分辨率(0.25°)全球再分析径流数据在站点尺度上代表性不足,无法有效表征实际径流过程。未来融合多源数据时,应优先选用高分辨率遥感水力变量,并谨慎对待粗网格模拟的径流量。

(4) 融合光学遥感水面宽度与机器学习的方法,为无资料/缺资料地区提供了一种高效、可靠的流量估算技术路径。该方法无需复杂水力学参数与地面标定,自动化程度高,且在唐乃亥站这类高寒、地形复杂的源区表现出优异性能。建议后续研究进一步融合雷达遥感(如哨兵1号)以克服云雨影响,并探索将物理约束(如水量平衡)嵌入机器学习模型,以提升外推能力与可解释性。

志 谢 感谢中国水利部信息中心提供的唐乃亥水文站点实测流量数据;感谢美国宇航局NASA、戈达德太空飞行中心GSFC、美国国家海洋和大气管理局NOAA和欧洲航天局ESA的卫星数据支持。

参考文献(References)

Bai X J, Wang Z Y, Liu T X, Li X D and Wan T Z. 2024. Response to climate changes and prediction model of runoff in the upper Yelow River Basin. *Science Technology and Engineering*, 24(18): 7502-7509 (白小晶, 王中玉, 刘泰兴, 李晓丹, 万铁庄. 2024. 黄

河上游径流量对气候变化的响应和预测模型. *科学技术与工程*, 24(18): 7502-7509 [DOI: 10.12404/j.issn.1671-1815.2305557]

Crochemore L, Isberg K, Pimentel R, Pineda L, Hasan A and Arheimer B. 2020. Lessons learnt from checking the quality of openly accessible river flow data worldwide. *Hydrological Sciences Journal*, 65(5): 699-711 [DOI: 10.1080/02626667.2019.1659509]

Dai J J, Liu T Y, Zhao Y Y, Tian S F, Ye C Y and Nie Z. 2023. Remote sensing inversion of the Zabuye Salt Lake in Tibet, China using LightGBM algorithm. *Frontiers in Earth Science*, 10: 1022280 [DOI: 10.3389/feart.2022.1022280]

Fang S, Liang F B and Liu Y J. 2021. A survey on statistical regression models and optimization algorithms. *Journal of Fuzhou University (Natural Science Edition)*, 49(5): 638-654 (方升, 梁飞豹, 刘勇进. 2021. 统计回归模型及其优化算法综述. *福州大学学报(自然科学版)*, 49(5): 638-654) [DOI: 10.7631/issn.1000-2243.21283]

Filippucci P, Brocca L, Bonafoni S, Saltalippi C, Wagner W and Tapanelli A. 2022. Sentinel-2 high-resolution data for river discharge monitoring. *Remote Sensing of Environment*, 281: 113255 [DOI: 10.1016/j.rse.2022.113255]

Hao R N and Bai Z X. 2023. Comparative study for daily streamflow simulation with different machine learning methods. *Water*, 15(6): 1179 [DOI: 10.3390/W15061179]

Ji P and Yuan X. 2020. Underestimation of the warming trend over the Tibetan Plateau during 1998-2013 by global land data assimilation systems and atmospheric reanalyses. *Journal of Meteorological Research*, 34(1): 88-100 [DOI: 10.1007/s13351-020-9100-3]

Lan X J, He Y L and Wu S W. 2024. River water quality prediction based on RF-BiLSTM model. *Journal of Changjiang River Scientific Research Institute*, 41(7): 57-63, 71 (兰小机, 贺永兰, 武帅文. 2024. 基于RF-BiLSTM模型的河流水质预测. *长江科学院院报*, 41(7): 57-63, 71) [DOI: 10.11988/ckyyb.20230244]

Le Y, Liu J T and Wen H. 2024. Extraction and spatiotemporal variation of Poyang Lake water body based on multi-source and multi-phase images. *Journal of Yangtze River Scientific Research Institute*, 41(8): 164-171, 206 (乐颖, 刘聚涛, 文慧. 2024. 基于多源多时相影像的鄱阳湖水体提取及时空变化分析. *长江科学院院报*, 41(8): 164-171, 206) [DOI: 10.11988/ckyyb.20230268]

Li X T, Li Z L and Han R C. 2023. Evaluations of different bias correction methods on the GCM precipitation data. *Journal of China Hydrology*, 43(3): 93-100, 117 (李昕潼, 李占玲, 韩孺村. 2023. 不同偏差校正法对GCM降水数据的应用效果分析. *水文*, 43(3): 93-100, 117) [DOI: 10.19797/j.cnki.1000-0852.20220060]

Li Z Q. 2022. Extracting spatial effects from machine learning model using local interpretation method: an example of SHAP and XGBoost. *Computers, Environment and Urban Systems*, 96: 101845 [DOI: 10.1016/j.compenvurbsys.2022.101845]

Liu D Y, Dong W Z, Gou R and Su W C. 2026. Spatial and temporal evolution of ecosystem service supply and demand and analysis of driving factors in Chongqing based on InVEST model and XGBoost-SHAP. *Environmental Science*, 47(5): 3312-3323 (刘东岳, 董文卓, 勾容, 苏维词. 2026. 基于InVEST模型和XGBoost-SHAP的重庆市生态系统服务供需时空演变及驱动因素分析.

- 环境科学, 47(5): 3312-3323 [DOI: 10.13227/j.hj.kx.202504257]
- Lundberg S M and Lee S I. 2017. A unified approach to interpreting model predictions//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc.: 4768-4777
- Madhushani C, Dananjaya K, Ekanayake I U, Meddage D P P, Kantamaneni K and Rathnayake U. 2024. Modeling streamflow in non-gauged watersheds with sparse data considering physiographic, dynamic climate, and anthropogenic factors using explainable soft computing techniques. *Journal of Hydrology*, 631: 130846 [DOI: 10.1016/J.JHYDROL.2024.130846]
- Mo J Y, Tian Y C, Wang J L, Du J Z, Zhang Q, Zhang Y L, Tao J and Lin J L. 2025. Remote sensing inversion of COD in Maowei Sea and nearshore aquaculture ponds based on machine learning. *Haiyang Xuebao*, 47(5): 128-140 (莫锦英, 田义超, 王家乐, 杜金泽, 张强, 张亚丽, 陶进, 林俊良. 2025. 基于机器学习的茅尾海及近岸养殖池水体 COD 遥感反演. *海洋学报*, 47(5): 128-140) [DOI: 10.12284/hyxb2025046]
- Papa F, Crétaux J F, Grippa M, Robert E, Trigg M, Tshimanga R M, Kitambo B, Paris A, Carr A, Fleischmann A S, de Fleury M, Gbetkom P G, Calmettes B and Calmant S. 2023. Water resources in Africa under global change: monitoring surface waters from space. *Surveys in Geophysics*, 44(1): 43-93 [DOI: 10.1007/s10712-022-09700-9]
- Qian Y P, Lin Y P, Jin S Y, Song R P and Jiang X H. 2004. Analysis of water resources changes in the source region of the Yellow River. *Water Resources and Hydropower Engineering*, 35(5): 8-10 (钱云平, 林银平, 金双彦, 宋瑞鹏, 蒋秀华. 2004. 黄河河源区水资源变化分析. *水利水电技术*, 35(5): 8-10) [DOI: 10.13928/j.cnki.wrahe.2004.05.003]
- Ren J, Wang J K, Chen R, Li H, Xu D L, Yan L H and Song J Y. 2024. Remote sensing identification of shallow landslide based on improved Otsu algorithm and multi feature threshold. *Frontiers in Earth Science*, 12: 1473904 [DOI: 10.3389/FEART.2024.1473904]
- Sahoo D P, Sahoo B, Tiwari M K and Behera G K. 2022. Integrated remote sensing and machine learning tools for estimating ecological flow regimes in tropical river reaches. *Journal of Environmental Management*, 322: 116121 [DOI: 10.1016/J.JENVMAN. 2022. 116121]
- Samat A, Li E Z, Du P J, Liu S C, Miao Z L and Zhang W. 2020. CatBoost for RS image classification with pseudo label support from neighbor patches-based clustering. *IEEE Geoscience and Remote Sensing Letters*, 19: 8004105 [DOI: 10.1109/LGRS.2020.3038771]
- Shi J, Zhou J G, Wang H, Gan L, Shen J L, Li X, Li Y Y and Wu J S. 2019. Analyzing the uncertainty induced by methods used to calculate the missing data in time series: a case study based on meteorological and hydrological data in small watershed. *Journal of Irrigation and Drainage*, 38(2): 84-92 (石锦, 周脚根, 王辉, 甘蕾, 沈健林, 李希, 李裕元, 吴金水. 2019. 点源时间序列数据缺失值的估值不确定性分析——以小流域气象和水文数据为例. *灌溉排水学报*, 38(2): 84-92) [DOI: 10.13522/j.cnki.ggps.2017.0421]
- Shi Y M, Liu X S, Zhu W B and Song H L. 2022. Research on inversion of river discharge in high mountain region based on GEE platform. *Remote Sensing Technology and Application*, 37(1): 186-195 (史宜梦, 刘希胜, 朱文彬, 宋宏利. 2022. 基于 GEE 云平台的黄河源区河流径流量遥感反演研究. *遥感技术与应用*, 37(1): 186-195) [DOI: 10.11873/j.issn.1004-0323.2022.1.0186]
- Su L P, Zhou X and Zhang S P. 2024. Research on small and medium-sized watershed flow prediction based on GRU and XGBoost algorithm. *Jilin Water Resources*, (10): 72-78 (苏林鹏, 周祥, 张守平. 2024. 基于 GRU 和 XGBoost 算法的中小流域流量预测研究. *吉林水利*, (10): 72-78) [DOI: 10.15920/j.cnki.22-1179/tv. 2024.10.014]
- Sun W C, Wang X C and Xu Z X. 2024. Estimating streamflow using remote sensing: progress and prospects. *Acta Geographica Sinica*, 79(3): 565-583 (孙文超, 王星灿, 徐宗学. 2024. 河道流量遥感估算研究进展与展望. *地理学报*, 79(3): 565-583) [DOI: 10.11821/dlxb202403002]
- Van den Broeck G, Lykov A, Schleich M and Suciu D. 2022. On the tractability of SHAP explanations. *Journal of Artificial Intelligence Research*, 74: 851-886 [DOI: 10.1613/jair.1.13283]
- Wang J P, Wu X D, Ma D J, Wen J G and Xiao Q. 2023. Remote sensing retrieval based on machine learning algorithm: uncertainty analysis. *National Remote Sensing Bulletin*, 27(3): 790-801 (汪静平, 吴小丹, 马杜娟, 闻建光, 肖青. 2023. 基于机器学习的遥感反演: 不确定性因素分析. *遥感学报*, 27(3): 790-801) [DOI: 10.11834/jrs.20221172]
- Wang X C, Sun W C, Lu F and Zuo R. 2023. Combining satellite optical and radar image data for streamflow estimation using a machine learning method. *Remote Sensing*, 15(21): 5184 [DOI: 10.3390/RS15215184]
- Xu H Q. 2005. A study on information extraction of water body with the modified normalized difference water index (MNDWI). *National Remote Sensing Bulletin*, 9(5): 589-595 (徐涵秋. 2005. 利用改进的归一化差异水体指数(MNDWI)提取水体信息的研究. *遥感学报*, 9(5): 589-595) [DOI: 10.11834/jrs.20050586]
- Yang S T, Wang P F, Wang J, Lou H Z and Gong T L. 2021. River flow estimation method based on UAV aerial photogrammetry. *National Remote Sensing Bulletin*, 25(6): 1284-1293 (杨胜天, 王鹏飞, 王娟, 娄和震, 巩同梁. 2021. 结合无人机航空摄影测量的河道流量估算. *遥感学报*, 25(6): 1284-1293) [DOI: 10.11834/jrs. 20209082]
- Zhang S T, Lu X, Lu Y, Cheng L, Li M C and Yang K. 2021. Tracking dynamic river networks in the Tibetan Plateau with high-resolution CubeSat imagery. *National Remote Sensing Bulletin*, 25(10): 2142-2152 (章斯腾, 陆欣, 陆瑶, 程亮, 李满春, 杨康. 2021. 青藏高原河流网络高分 CubeSat 遥感监测. *遥感学报*, 25(10): 2142-2152) [DOI: 10.11834/jrs.20219268]

River discharge estimation by integrating Sentinel-2 and machine learning: A case study of the Tangnaihai Hydrometry Station

YIN Wenjie^{1,2}, WANG Xuelei^{1,2}, WANG Chen^{1,2}, WANG Hang^{1,2}, HUANG Caisheng³,
ZHAO Ruixue^{1,2}, MENG Fanle^{1,2}, LIU Jinxiu³

1. Satellite Application Center for Ecology and Environment, Ministry of Ecology and Environment, Beijing 100094, China;

2. Key Laboratory of Satellite Remote Sensing, Ministry of Ecology and Environment, Beijing 100094, China;

3. School of Information Engineering, China University of Geosciences, Beijing 100083, China

Abstract: River discharge is a pivotal variable within the hydrological cycle, holding significant importance for flood warning, water resource allocation, and eco-environmental management. Traditional ground-based methods are limited by the sparse distribution of stations and the high cost of data acquisition, particularly in areas with complex terrains or remote regions, making it difficult to meet the demands of precise water resource management. Satellite remote sensing technology offers extensive coverage and high spatiotemporal resolution, providing new data sources and methodologies for river discharge monitoring. Machine Learning (ML) approaches can accurately simulate complex relationships between river discharge and multiple driving factors, offering novel avenues for processing intricate hydrological data and optimizing models. Integrating ML algorithms with remote sensing and *in-situ* river discharge can provide an innovative measure for efficient and reliable river discharge monitoring.

This study selected the Tangnaihai Hydrometry Station as the research area and proposed a river discharge monitoring method by integrating satellite remote sensing and ML methods. Sentinel-2 imagery was utilized to extract river water surface width on the basis of the Google Earth Engine cloud platform. Five GLDAS v2.2 model-simulated variables, namely, evapotranspiration, soil moisture, temperature, terrestrial water storage, and runoff, served as predictor variables. Discharge monitoring models were subsequently developed on the basis of four statistical methods (linear, power, exponential, and polynomial functions) and four ML algorithms (XGBoost, random forest, LightGBM, and CatBoost). The discrepancies among different models were assessed, and the shapley additive explanation (SHAP) method was employed to quantify the importance of different input variables.

Results indicated that the polynomial function model demonstrated superior performance over the other three statistical models during the testing period, with a coefficient of determination (R^2) of 0.67, and its error metrics (mean absolute error [MAE]: 319.01 m³/s, root-mean-square error [RMSE]: 393.14 m³/s) were lower than those of the other three models. Compared with the traditional statistical approaches, the ML models exhibited significant improvements overall in simulation accuracy and stability; the R^2 increased by 46.15%, while the RMSE and MAE decreased by 54.61% and 55.65, respectively. Notably, the random forest model achieved the optimal performance in the testing phase, with an R^2 of 0.96, a Nash-Sutcliffe efficiency coefficient of 0.89, an RMSE of 172.81 m³/s, and a MAE of 147.33 m³/s, reflecting robust generalization capability and stability. SHAP analysis revealed that water surface width contributed most significantly to the discharge monitoring model (189.02), followed by soil moisture (145.11) and temperature (97.41). The runoff variable exhibited the minimum degree of influence on the river discharge monitoring model with a value of 14.14%.

This study confirms the feasibility and superiority of integrating satellite remote sensing and ML approaches for high-accuracy discharge estimation in regions characterized by complex topography and data scarcity. Future work could be optimized by integrating higher-resolution satellite imagery and mechanistic models with physical processes.

Key words: satellite remote sensing, machine learning, statistical models, SHAP method, discharge monitoring, Tangnaihai Hydrometry Station

Supported by National Natural Science Foundation of China (No. U23A2015)