

DOI: 10.20103/j.stxb.202405271227

丁悦凯, 刘睿, 张静. 典型喀斯特河流 CDOM 遥感反演及其驱动因素研究. 生态学报, 2025, 45(16): - .

Ding Y K, Liu R, Zhang J. Colored dissolved organic matter in karst rivers: Remote sensing inversion and driving factors. Acta Ecologica Sinica, 2025, 45(16): - .

## 典型喀斯特河流 CDOM 遥感反演及其驱动因素研究

丁悦凯<sup>1,2</sup>, 刘睿<sup>1,2</sup>, 张静<sup>1,2,\*</sup>

1 重庆师范大学地理与旅游学院, 重庆 401331

2 重庆师范大学地理信息系统应用研究重庆市高校重点实验室, 重庆 401331

**摘要:** 有色溶解有机物 (colored dissolved organic matter, CDOM) 在水生态系统中发挥着重要作用, 影响着河流营养物的循环与污染物质的迁移。本研究针对喀斯特地区芙蓉江流域 CDOM 进行遥感反演, 结合地面实测数据和 Sentinel-2A 遥感影像, 采用 AdaBoost 算法与 LIME 算法构建 CDOM 反演模型。结果表明, 模型反演精度较高, 测试集  $R^2$  为 0.86, RMSE 为 0.495/m。CDOM 在旱季和雨季具有空间分异性, 其中雨季 CDOM 吸光系数普遍高于旱季。此外, 基于可解释性技术揭示了不同的波段比值特征对模型反演的贡献差异, 研究发现波段比值 B4/B3、B6/B4 及 B5/B2 对模型反演的影响较大。旱季与雨季 CDOM 受驱动因素影响作用不同, 较高的气温和较少的降水量易导致 CDOM 吸光系数降低, 土地利用对河流 CDOM 影响显著, 人口密度与 CDOM 的负相关可能与该区域地下水系统的稀释作用及环保措施相关。

**关键词:** 有色溶解有机物 (CDOM); 遥感反演; 机器学习; 可解释性; 芙蓉江流域

## Colored dissolved organic matter in karst rivers: Remote sensing inversion and driving factors

DING Yuekai<sup>1,2</sup>, LIU Rui<sup>1,2</sup>, ZHANG Jing<sup>1,2,\*</sup>

1 School of Geography and Tourism, Chongqing Normal University, Chongqing 401331, China

2 Chongqing Key Laboratory of GIS Application, Chongqing Normal University, Chongqing 401331, China

**Abstract:** Colored dissolved organic matter (CDOM) plays a crucial role in aquatic ecosystems, influencing the cycling of nutrients and the transport of pollutants in water bodies. This study focused on the remote sensing inversion of CDOM in the Furong River Basin, a karst area, utilizing in-situ data and Sentinel-2A remote sensing imagery. The CDOM inversion model was constructed using AdaBoost and Local Interpretable Model-agnostic Explanations (LIME) algorithms. The results exhibited that the model has high inversion accuracy, with an  $R^2$  of 0.86 and an RMSE of 0.495/m for the test-set. There was spatial variability in CDOM between the dry and wet seasons, with a higher absorption coefficient generally observed during the wet season. Additionally, interpretable technology revealed the contribution differences of various band ratios to the model inversion, with band ratios B4/B3, B6/B4, and B5/B2 having a substantial impact. The study also found that different driving factors affect CDOM in different seasons, higher temperatures and lower precipitation tend to reduce CDOM, land use significantly impacted river CDOM, and the negative correlation between population density and CDOM may be related to the dilution effect of the regional groundwater system and environmental protection measures.

**Key Words:** colored dissolved organic matter; remote sensing inversion; machine learning; interpretability; Furong

**基金项目:** 国家自然科学基金项目 (42407094); 重庆市自然科学基金项目 (2024NSCQ-MSX3061); 重庆市教委科学技术研究计划项目 (KJQN202200517); 重庆师范大学基金项目 (20XLB022)

**收稿日期:** 2024-05-27; **网络出版日期:** 2025-00-00

\* 通讯作者 Corresponding author. E-mail: zhangjing@cqnu.edu.cn

## River Basin

河流不仅是陆地与海洋之间的关键连接点,也是陆源物质输送到海洋的主要路径,对全球物质循环和能量交换至关重要<sup>[1]</sup>。近年来,随着我国经济社会的飞速发展,河流作为关键的淡水资源被不断开发与利用,水质生态安全受到严重损害<sup>[2]</sup>,对我国经济社会的可持续发展造成了严重威胁<sup>[3-5]</sup>。有色溶解有机物 (colored dissolved organic matter, CDOM) 是指在水体中的有色有机物质,具有可溶解性<sup>[6-7]</sup>,CDOM 不仅可以吸收紫外线与可见光,为水生生物提供了一定程度的光保护作用<sup>[8]</sup>,同时 CDOM 的分布和变化影响着河流的透明度<sup>[9]</sup>,对水下植物的生长和水生生态系统结构的稳定具有重要意义。为此,了解与分析河流中 CDOM 的分布及变化是保障水资源安全的重要基础,对促进生态平衡及推动经济社会的可持续发展具有现实意义。

传统的 CDOM 监测方法多采用实地采样与实验室分析,该方法费时费力且成本高昂,无法对河流 CDOM 进行空间化。遥感数据因具有高时效性和大范围覆盖等特点,被广泛用于河流 CDOM 的监测研究<sup>[10-12]</sup>,例如,吴铭等<sup>[13]</sup>利用逐步回归模型,结合 Landsat 系列卫星的数据,对北极 Lena 河 CDOM 进行估算;Cai 等<sup>[14]</sup>基于高光谱影像,应用波段比值法反演了京新河、龙泰河等区域的 CDOM。上述研究均表明利用遥感数据可高精度的估算河流 CDOM,证实了遥感数据的可用性,为河流 CDOM 的监测提供了新的思路。

常用的 CDOM 反演算法主要包括以对数、指数、线性回归等为代表的经验算法<sup>[15-16]</sup>,以及考虑物理意义的半分析算法<sup>[17-19]</sup>,随着人工智能技术的不断发展,机器学习技术凭借其卓越的数据发掘能力,被国内外众多学者用于 CDOM 的反演研究,例如,Huang 等<sup>[20]</sup>对比了 6 种机器学习算法于珠江口区域水体 CDOM 反演的差异,结果表明 XGBoost 算法反演精度最高;金则澎等<sup>[21]</sup>采用梯度提升回归树对千岛湖区域的水体 CDOM 进行估算,结果表明基于梯度提升回归树所建立的模型可准确反演水体 CDOM。上述研究的顺利开展极大地促进了机器学习技术于水体 CDOM 反演中的应用。然而,机器学习模型多为黑箱模型,研究人员对模型的内部机理与决策成因无法给出科学的解释,不能充分了解 CDOM 反演的具体计算过程,使得部分训练效果优秀的模型无法取得人们的信任。近年来,利用可解释性算法和机器学习相结合的方法来反演水质参数逐渐备受人们青睐,可解释性算法可对模型的预测结果进行解释,有效推测模型决策原理与成因,提升模型的可信度。例如,Zhang 等<sup>[22]</sup>利用 XGBoost 算法构建叶绿素 a 浓度反演模型,采用 SHAP 算法对模型进行解释,相比 SHAP 算法,LIME 算法具有更高的灵敏性与更低的计算成本,可应用于任何类型的机器学习模型。目前,鲜见前人将可解释性技术结合机器学习用于 CDOM 反演研究。

此外,已有研究表明 CDOM 受气温、降水、土地利用等因素影响,Zhuang 等<sup>[23]</sup>通过冗余分析阐明了黄土高原 CDOM 的驱动因素,结果表明降水量、水蚀面积及人类活动强度对 CDOM 影响较大;Shao 等<sup>[24]</sup>利用多元通用线性模型探究了人类活动强度、总磷及总碱度对 CDOM 的影响作用,结果表明总磷和总碱度影响最为强烈。然而,目前多数研究仅针对点源 CDOM 进行分析,并未与遥感反演结果进行有效结合,且旱季与雨季间驱动因素的变化机理尚不明晰。通过分析旱季与雨季 CDOM 的驱动因素,可更好的理解河流生态的发展规律,为水质管理和保护提供科学依据。

综上,本研究选择典型喀斯特地区—芙蓉江流域作为研究对象,结合地面实测数据与 Sentinel-2A 卫星影像作为主要数据,利用机器学习技术构建 CDOM 反演模型,采用 LIME 算法对模型进行解释,对比分析旱季与雨季河流 CDOM 的时空异质性及驱动因素变化,旨在为河流 CDOM 监测研究提供新的思路与参考,并为喀斯特地区河流生态管理提供科学支撑。

## 1 材料与方法

### 1.1 研究区概况

本研究选取芙蓉江流域(28°0'57"—29°14'23"N,107°10'3"—107°52'42"E)作为研究区域(图 1),芙蓉江地处长江流域中游,起源于贵州省遵义市,流经重庆市彭水县和武隆区后注入乌江。河流全长约 231 km,流

域面积约为 7,406 km<sup>2</sup>。芙蓉江流域拥有复杂的水系网络,涵盖了三江、梅江和清溪河等主要河流。该地区地质结构多样,广泛分布着碳酸盐岩,形成了典型的喀斯特地貌<sup>[25]</sup>。气候方面,芙蓉江流域属于亚热带季风气候,年均降雨量介于 900—1200 mm 之间,其中大约 86% 的年度降雨集中于 5 月至 9 月<sup>[26]</sup>,年平均气温在 16.14 °C 左右。芙蓉江流域土地利用类型主要为耕地与林地,其中耕地面积占比 31.9%,林地面积占比 60.4%,河流附近多分布经济林与果树,芙蓉江流域水质表现出显著的时空异质性<sup>[27]</sup>。

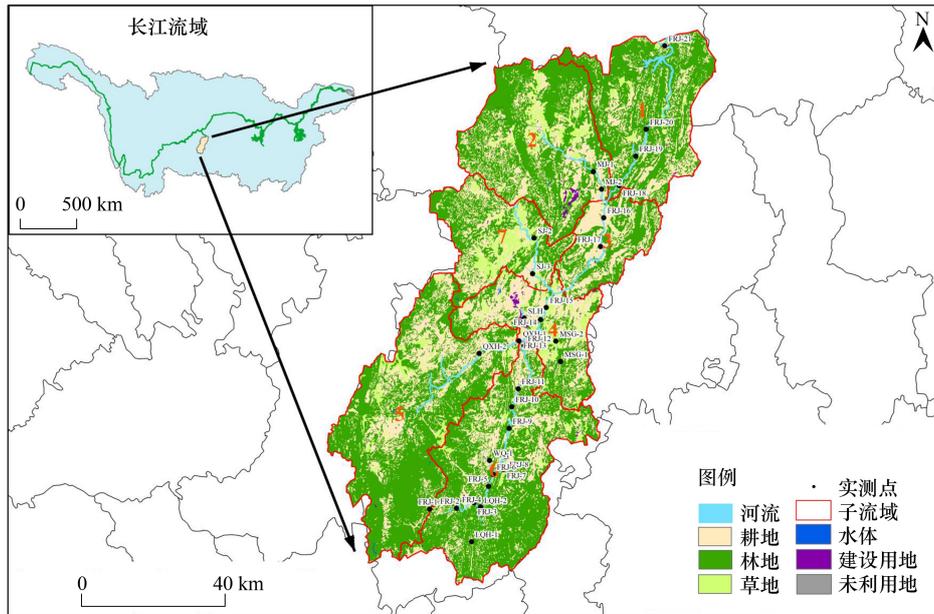


图 1 研究区概况

Fig.1 The Location of study area

## 1.2 数据来源与处理

### 1.2.1 CDOM 实测数据

本研究分别于旱季(2019 年 12 月)和雨季(2020 年 7 月)在芙蓉江流域采集水样共 65 个,选取河道表层(10—20 cm)的水体进行采样,其中包括上游采样点 15 个(FRJ-1、FRJ-2、FRJ-3、FRJ-4、FRJ-5、FRJ-6、FRJ-7、FRJ-8、FRJ-9、FRJ-10、FRJ-11、LQH-1、LQH-2、WQ-1 和 QXH-2),中游采样点 11 个(FRJ-12、FRJ-13、FRJ-14、FRJ-15、FRJ-17、MSG-1、MSG-2、QXH-1、SJ-2、SJ-3 和 SLH)及下游采样点 7 个(FRJ-16、FRJ-18、FRJ-19、FRJ-20、FRJ-21、MJ-1 和 MJ-2)。采样时天气晴朗,通过玻璃纤维膜(GF/F 47 mm, 0.45 μm, Whatman)对水样进行过滤处理后进行冷藏,并在一周内运送至实验室进行测定。针对所采集的水样数据,通过 UV-1500PC 分光光度计(Macy, 上海)完成紫外—可见光谱分析,以 1 nm 作为间隔,200-700 nm 为范围进行扫描,使用 Milli-Q 超纯水作为空白对照。

此外,采用光谱响应函数将各实测点的遥感反射率计算为 Sentinel-2A 各波段的光谱值,计算公式如下:

$$R_{rs}(\lambda_i) = \frac{\sum_{j=1}^n F_i(\lambda_j) R_{rs}(\lambda_j)}{\sum_{j=1}^n F_i(\lambda_j)} \quad (1)$$

式中: $R_{rs}(\lambda)$ 表示实测遥感反射率( $\text{sr}^{-1}$ ), $F_i$ 表示数据的第  $i$  波段的光谱响应函数, $\lambda_i$ 表示实测光谱数据的第  $i$  个波段, $\lambda_j$ 表示影像光谱数据的第  $j$  个波段。

### 1.2.2 CDOM 吸光系数测定

针对实验所测的 CDOM 吸光度,采用公式(2)计算 CDOM 吸光系数,参考相关研究<sup>[28—30]</sup>,选取 355 nm 处的吸光系数用于表征河流 CDOM。

$$a(\lambda) = 2.303A(\lambda)b \quad (2)$$

式中: $A(\lambda)$ 为 CDOM 吸收波长为  $\lambda$  时的吸光度, $b$ 为吸光路径长度。

### 1.2.3 遥感影像数据

本研究主要使用与采样时间相近的 Sentinel-2A 遥感影像进行 CDOM 反演,具有 10 m 的空间分辨率,10 天的时间分辨率,主要运用该影像的海岸波段(B1)、蓝光波段(B2)、绿光波段(B3)、红光波段(B4)、红边波段(B5)、红边波段(B6)。本研究通过 Google Earth Engine 平台去除影像堆栈中的无效像素值,数据级别为 L2A 级,已通过 Sen2Cor 软件进行大气校正等预处理。此外,结合归一化差分水体指数与目视解译法提取研究区域的河流边界。

### 1.2.4 驱动因素数据

依据河流的空间分布及区域地形特征,结合 Shuttle Radar Topography Mission 数字高程模型和 Soil and Water Assessment Tool 模型提取流域河网,确定流域的主河道<sup>[31]</sup>,再通过对比芙蓉江的主干及支流分布,结合实际地理空间特征,将芙蓉江流域划分为 7 个子流域(图 1)。参考相关研究<sup>[32-34]</sup>,选取对 CDOM 影响较大的月均气温、月均降水、人口密度、耕地覆盖率、林地覆盖率、建设用地覆盖率作为驱动因素进行分析。本研究采用土地利用数据源于中国科学院资源环境科学与数据中心(<https://www.resdc.cn/>),空间分辨率 30 m,该数据将研究区域的土地利用类型分为耕地、林地、建设用地等 6 类,具有较高的精确性;同时,本研究采用高分辨率月网格气温数据集<sup>[35]</sup>(<https://zenodo.org/>)与高分辨率月网格降水数据集<sup>[36]</sup>(<https://www.geodata.cn/>)表征芙蓉江流域的气象因素,该类数据基于气象站点实测值进行计算,空间分辨率均为 1 km;研究区人口密度数据则选用 WorldPop 人口数据集<sup>[37]</sup>(<https://hub.worldpop.org/>),空间分辨率为 100 m,该数据通过整合土地利用类型、地形地貌、夜间灯光指数等多源地理数据,利用机器学习算法进行计算,精度较高,被广泛用于各类研究中。本研究采用 ArcGIS 10.4 平台的分区统计工具统计各子流域下的驱动因素。

## 1.3 研究方法

### 1.3.1 AdaBoost 算法

AdaBoost 是由 Freund 等<sup>[38]</sup>于 1995 年提出的集成学习方法,该算法的核心思想是结合多个弱学习器,通过加权投票或加权平均的方式构建出一个性能较强的集成模型。AdaBoost 算法通过对每个训练样本分配权重,迭代选择错误计算成本最高的样本从而确定权重,优化模型性能。AdaBoost 算法适用于机器学习中的回归问题,能有效提高弱学习器的准确率,同时对噪声数据和异常值具有较高的鲁棒性。当前尚未有研究将该算法应用于 CDOM 反演,算法计算公式如下:

$$\hat{y}_i = \text{sign}\left(\sum_{k=1}^K \alpha_k f_k(x_i)\right) \quad (3)$$

式中: $K$ 表示决策树的数量, $x_i$ 表示第  $i$  个样本, $f_k(x_i)$ 表示对样本的第  $i$  棵决策树的运算结果, $\alpha_k$ 是第  $k$  个模型的权重。

此外,本研究通过决定系数( $R^2$ )与均方根误差(RMSE)对模型精度进行评价,其中  $R^2$ 越接近 1,表明模型拟合效果越好, RMSE 越小,表明模型反演精度越高,计算公式如下:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (5)$$

式中: $n$ 表示样本点的数量, $i$ 代表单个数据点, $y_i$ 和  $\hat{y}_i$ 分别表示实测值和模型反演值。

### 1.3.2 LIME 算法

LIME 算法是一种用于解释复杂机器学习模型的技术,由 Ribeiro 等<sup>[39]</sup>于 2016 年提出。LIME 算法的原理是在模型的预测区域附近创建一个局部数据集,对数据集进行扰动并观察预测结果的变化,使用相关模型

对数据集进行局部拟合,从而揭示特征对模型预测的重要作用。LIME 算法计算公式如下:

$$f_{x_s}(x_s) = E_{x_c} [\hat{f}[(x_s, x_c)]] dP(x_c) = \int \hat{f}(x_s, x_c) dP(x_c) \quad (6)$$

式中: $x_s$ 为特征变量, $f_{x_s}(x_s)$ 是基于特征 $x_s$ 所计算的值, $x_c$ 是除 $x_s$ 以外的其它特征变量。

### 1.3.3 Pearson 相关性分析

相关性分析可用于表征两个因素间的相关程度,本研究通过 Pearson 相关性分析探究芙蓉江流域 7 个子流域不同驱动因素对 CDOM 的影响作用,相关系数  $r$  的值范围从 -1 到 1,当  $r > 0$  时,表明该因素对 CDOM 起促进作用,当  $r < 0$  时,表明该因素对 CDOM 起抑制作用。参考相关研究<sup>[40-41]</sup>,采用插值技术对样本数据进行扩充。

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (7)$$

式中: $x_i$ 和 $y_i$ 分别是两个变量的观测值, $\bar{x}$ 和 $\bar{y}$ 分别是这两个变量的平均值。

## 2 结果

### 2.1 CDOM 吸光系数分析

芙蓉江流域所有采样点的实测吸光度都随波长的增加而逐渐降低,显现出明显的下降趋势,表明在波长较短的区域即紫外波段附近吸收能力较强,在波长较长的区域吸收能力较弱(图 2)。此外,不同时段内 CDOM 吸光系数的异质性较为显著,旱季 CDOM 吸光系数均值为  $(4.18 \pm 0.72)/m$ ,CDOM 吸光曲线较为平坦,吸光度变化幅度较小,表明其中的河流 CDOM 相对稳定且含量较低,雨季 CDOM 吸光系数均值为  $(5.28 \pm 1.74)/m$ ,较早季吸光度整体较高,曲线也显现出更明显的下降趋势(图 2)。

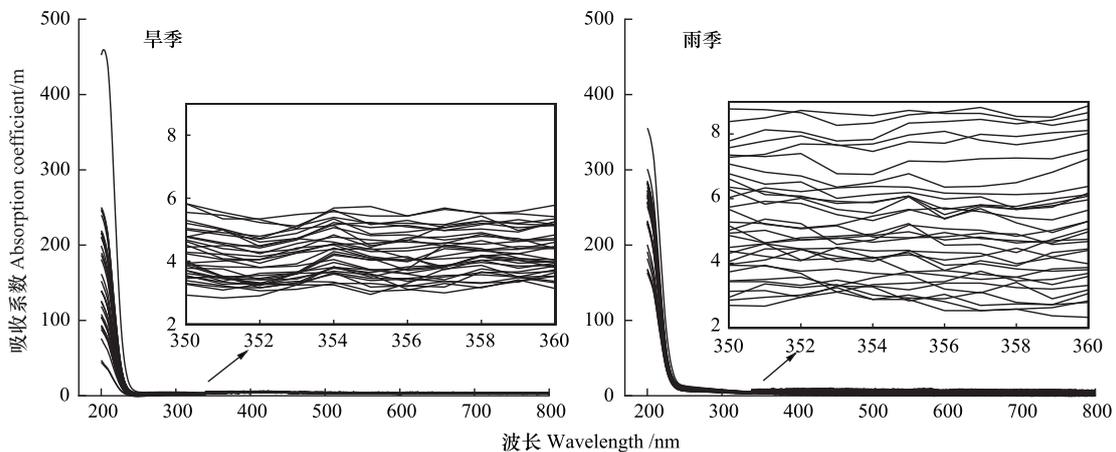


图 2 CDOM 吸光系数曲线

Fig.2 CDOM absorption coefficient curve

### 2.2 CDOM 反演

选取合适的特征对构建机器学习模型至关重要,本研究共选择了 11 个特征,包括 B1、B2、B3、B4、B5 共 5 个单波段,以及它们的比值 B4/B3、B5/B4、B6/B3、B6/B4、B4/B2 以及 B5/B2,按照 7:3 划分训练集与测试集,采用 AdaBoost 算法构建 CDOM 反演模型。

通过  $R^2$  与 RMSE 对模型进行评价,训练集 RMSE 为  $0.395/m$ ,测试集 RMSE 为  $0.495/m$ ,训练集  $R^2$  为 0.89,测试集  $R^2$  为 0.86,显现出较高的拟合度,模型散点图如图 3 所示,表明模型精度较好,具有良好的预测能力与稳定性,可用于反演芙蓉江流域河流的 CDOM。

基于所建立的反演模型,对 Sentinel-2A 影像进行计算,反演得到芙蓉江流域河流旱季与雨季两个时间段

的 CDOM 空间分布图,旱季 CDOM 吸光系数均值为  $(4.26 \pm 0.40)/m$ , 整个流域内 CDOM 吸光系数相对较低且比较均匀,雨季 CDOM 吸光系数均值较高,为  $(4.62 \pm 0.53)/m$  (图 4)。此外,通过对比可以发现,芙蓉江流域下游雨季 CDOM 较早季有所降低,而中游 CDOM 有所升高,上游变化较小(图 4)。可能是由于下游的大量水量与雨季降水稀释 CDOM<sup>[42]</sup>, 同时泥沙吸附导致 CDOM 降低<sup>[43]</sup>, 而中游区域河流距离建设用地较近(图 1), 人类活动影响强度较大,导致 CDOM 吸光系数有所升高。

### 2.3 模型可解释性分析

通过 LIME 算法可对反演模型进行局部解释, 分别从研究区域内的上、中、下游各选取同一实测点 (FRJ-7、FRJ-13、FRJ-21) 进行分析, 且在不同时间段内相同的波段特征对反演模型的预测具有不同作用(图 5)。例如, 对于旱季的实测点 FRJ-21 与 FRJ-13, 当  $B_6/B_4 > 0.83$  以及  $0.49 < B_6/B_3 \leq 0.72$  时对反演结果产生负向作用, 而对于雨季的实测点, 当  $B_4 \geq 0.17$  时对模型反演产生负面影响, 当  $0.06 < B_4 \leq 0.11$  时, 对模型反演产生正面影响, 表明在这个特征范围内  $B_4$  波段对模型预测的正确性有积极的促进作用。同时对于雨季的实测点 FRJ-21, 当  $B_4/B_3 \leq 0.59$  时对模型正确预测的贡献度最大, 对于雨季的实测点 FRJ-7, 当  $B_6/B_3 \leq 0.37$  时影响模型的正确反演, 即较低的波段比值的重要性有所增加。由此可得, 河流水体的物理与化学性质有着重要的季节性变化, 不同的光谱特征在不同季节下对模型影响作用不同。

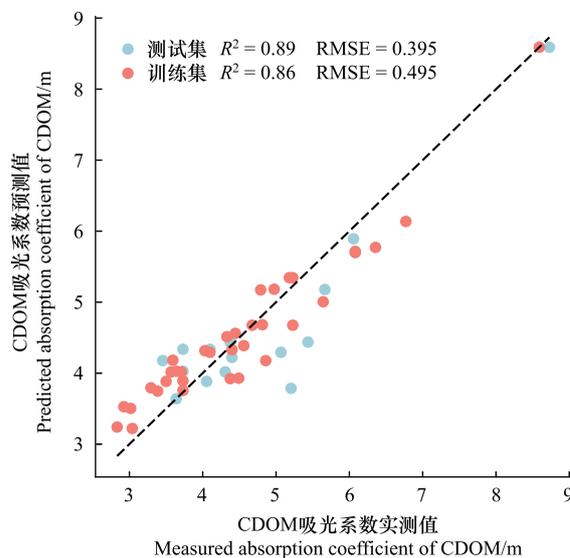


图 3 CDOM 反演模型散点图

Fig.3 Scatter plot of CDOM inversion model

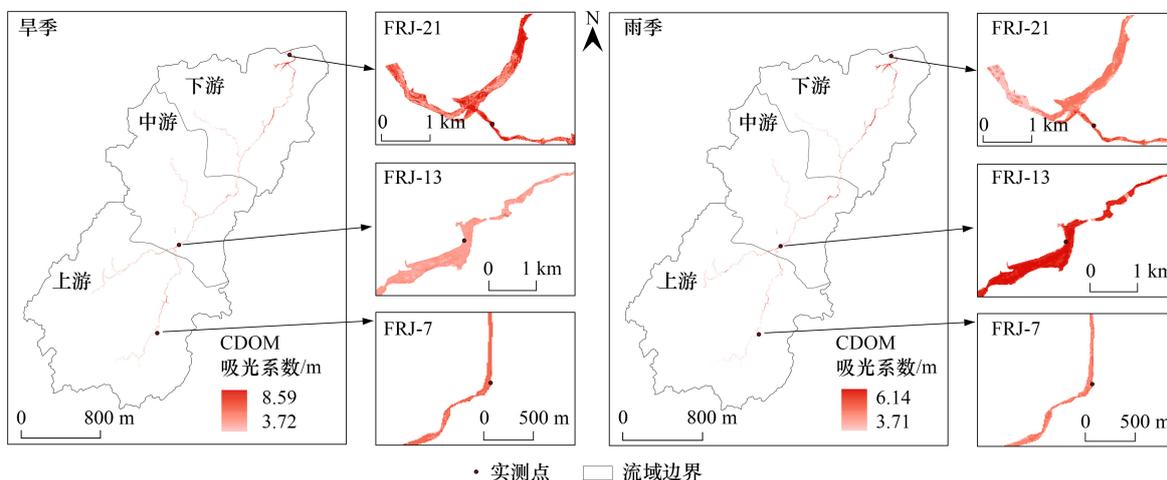


图 4 CDOM 空间分布图

Fig.4 Spatial distribution of CDOM

FRJ-7、FRJ-13、FRJ-21 为芙蓉江上、中、下游处实测点

此外,不同波段及其比值不同程度地影响模型的预测效果。例如,  $B_4$  波段影响模型反演的结果, 而  $B_3$ 、 $B_5$ 、 $B_6$  波段的重要性相对较小。相比之下, 当  $B_4/B_3 \leq 0.59$  时下对模型的预测贡献度较大, 当  $B_6/B_4 \geq 0.83$  时, 在旱季中对模型产生了负向影响, 且重要性较大,  $B_5/B_2$  对模型的贡献度较大, 即波段比值影响着模型反演的性能, 验证了波段比值在遥感反演中的重要性。

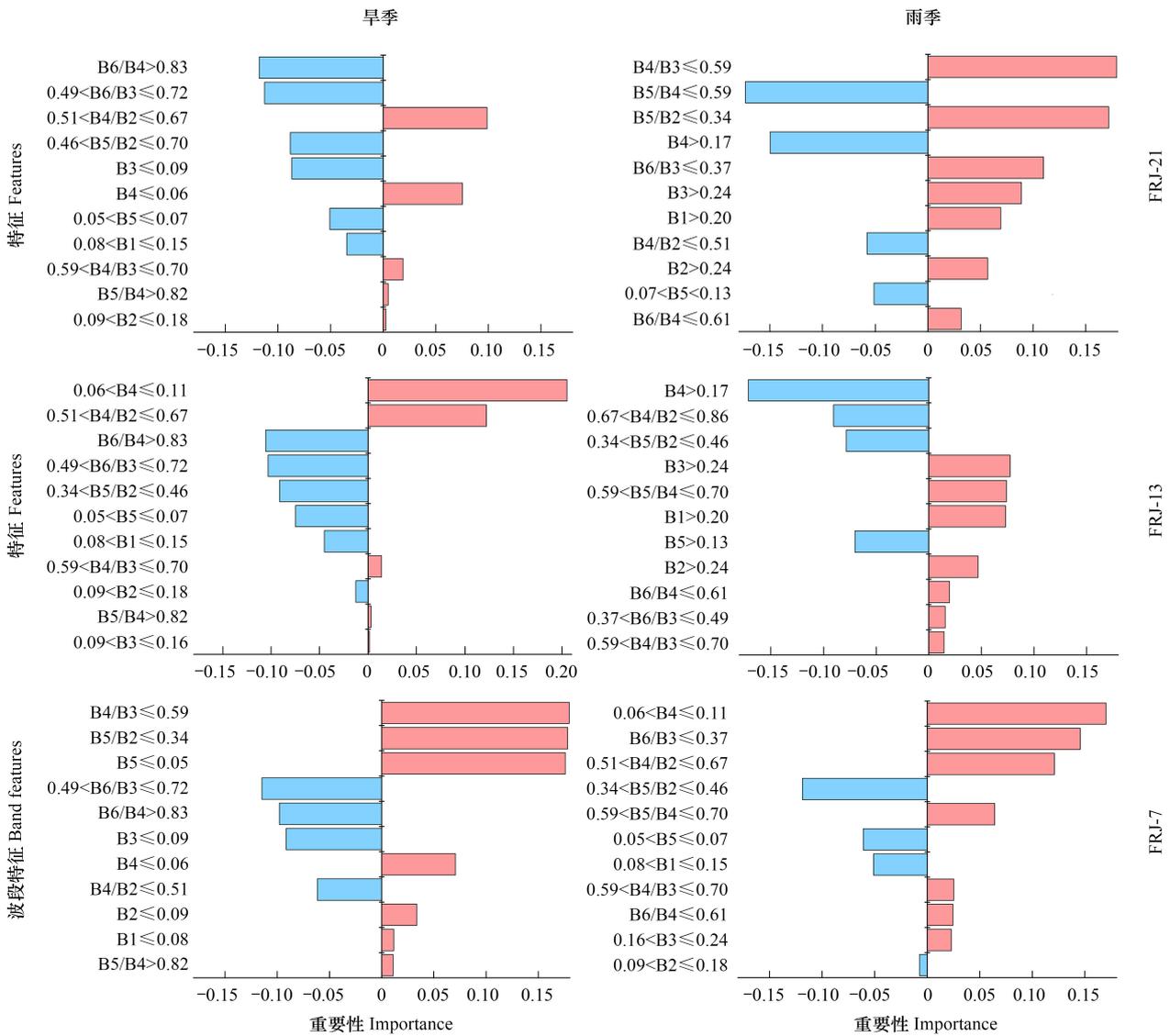


图 5 LIME 对不同季节样本的解释  
 Fig.5 LIME explanation of different seasonal samples

### 2.4 相关性分析

统计各子流域下的驱动因素(表 S1),计算它们与 CDOM 间的 Pearson 相关性系数,以探究不同因素在旱季与雨季对 CDOM 的驱动机制。相关性分析结果显示除旱季人口密度与 CDOM 间相关性分析的  $P$  值 $>0.05$  以外,其它驱动因素的统计意义均具有显著性(图 6)。

月均降水与 CDOM 在旱季和雨季均呈显著正相关,旱季相关系数为 0.206,雨季降低至 0.186(图 6)。月均气温则表现出显著负相关,旱季相关系数为-0.374,雨季降低至-0.215(图 6)。人口密度在雨季与 CDOM 呈负相关,相关系数为-0.073,但在旱季影响不显著(图 6)。耕地覆盖率和林地覆盖率在旱季中表现出显著的负相关和正相关关系,旱季耕地覆盖率为-0.652,雨季则为-0.074,旱季林地覆盖率为 CDOM 的相关系数为 0.573,雨季降低至-0.076(图 6)。建设用地覆盖率在旱季对 CDOM 影响为负,但在雨季呈现显著负相关,相关系数为 0.200(图 6)。上述结果表明不同驱动因素对 CDOM 的影响及其季节变化存在显著差异。

### 3 讨论

#### 3.1 基于可解释性模型的 CDOM 反演

本研究通过采用 AdaBoost 与 LIME 算法对芙蓉江流域内河流的 CDOM 进行反演,结果表明,该模型不仅可以较高精度的反演河流 CDOM,同时也可增强研究人员对模型决策过程的理解,能够更好地识别和解释 CDOM 反演模型的关键因素。目前,越来越多的研究采用机器学习算法反演水体 CDOM。然而,机器学习模型通常被视为黑箱模型,降低了反演模型及其结果的可信性与透明度<sup>[44-45]</sup>。可解释性技术可以定性地反映遥感数据的各个特征与 CDOM 之间的关系,为解决这一问题提供了思路。

本研究采用 LIME 算法对模型进行解释,揭示了 Sentinel-2A 影像的不同波段及组合特征对 CDOM 预测的具体贡献。结果表明,波段比值 B4/B3、B6/B4 及 B5/B2 对模型反演影响较大,CDOM 在较短的波长有较强的吸收,而在较长波长的吸收相对较弱,通过波段比值运算则可更为敏感的捕捉 CDOM 的空间分布(图 2)。例如,CDOM 强烈吸收蓝光而对红边波段的吸收较弱,使得 B5/B2 在 CDOM 变化时更为显著,Shang 等<sup>[15]</sup>研究表明利用 Sentinel-2 影像的 B5/B2 可准确反演 CDOM,并减少叶绿素 a 浓度所造成的误差,同时,较低的波段比值在雨季的影响较为显著,旱季中较高的波段比值特征对模型的预测产生负向作用,反映了不同水文条件下光谱特征的敏感性变化。

本研究通过引入 LIME 算法对 CDOM 反演模型进行解释,提高了方法的实用性,增强了模型的解释能力和透明度,降低了模型的复杂度,为水质监测和管理提供了新的方法与参考。

#### 3.2 CDOM 驱动因素分析

研究表明,旱季与雨季两个时段内芙蓉江流域河流的 CDOM 存在显著差异,体现了不同的环境条件对河流有机物质浓度的影响作用,深入分析 CDOM 的季节变化及驱动因素可为区域水资源管理提供主要依据。

本研究通过 Pearson 相关性分析探究了不同驱动因素在旱季与雨季对 CDOM 变化的影响作用(图 6),结果表明,气象因素对 CDOM 影响较大,在旱季,研究区域 CDOM 吸光系数为 4.26/m,较高的气温通常会加速有机物的分解<sup>[46]</sup>,从而降低河流中的 CDOM,因而气温与 CDOM 为负相关。此外,旱季的降水量较少,减少了地表径流,较少的陆源有机物被搬运到河流中,CDOM 吸光系数降低,而在雨季,CDOM 吸光系数升高至 4.62/m,大量降水增加了地表径流,从而增加了从陆地到河流的有机物搬运<sup>[47]</sup>,导致 CDOM 吸光系数增高,即降水与 CDOM 间呈正相关。

人类活动是影响区域 CDOM 的重要因素,在旱季,耕地覆盖率为 CDOM 呈显著负相关(图 6),其原因可能是由于喀斯特区域土层浅薄且贫瘠,河流周边的有机质含量普遍较低<sup>[48-50]</sup>,同时耕地植被通过根系作用影响土壤微生物群落的结构和活性<sup>[51-52]</sup>,通过提高微生物多样性和活性,导致更多 CDOM 在土壤中被消耗<sup>[53-54]</sup>,且喀斯特地貌促使地表水与地下水之间物质交换频繁<sup>[55]</sup>,耕地中施肥、灌溉、使用农药等方式产生的有机质易通过土壤的快速下渗作用进入地下水系统,减少了对河流 CDOM 的直接影响。林地覆盖率为 CDOM 呈正相关(图 6),则可能与林地植物残体及土壤中有有机物质的输入密切相关<sup>[56-58]</sup>,河流附近由于存在经济林与果树的种植活动,施肥和施药等行为导致有机物质进入水体,影响河流中 CDOM 的含量。而雨季中,耕地和林地对 CDOM 的影响较弱,大量的降水将导致地表径流增强,水流速度加快,部分区域可能发生土

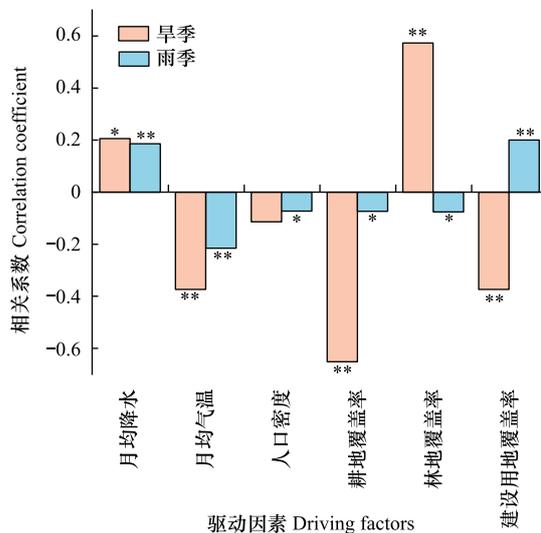


图 6 CDOM 与驱动因素相关性

Fig.6 Correlation between CDOM and driving factors

\*表示  $P \leq 0.05$ , \*\*表示  $P \leq 0.01$

壤侵蚀,影响着有机物质的淋溶和迁移模式<sup>[59]</sup>。建设用地覆盖率为人口密度在旱季与 CDOM 均呈负相关,一方面,可能是由于旱季降水少,地表径流不足,难以将有机物质冲刷进入河流,另一方面,芙蓉江流域建设用地面积仅占总面积的 0.39%(表 S2),人类活动强度较弱,其影响作用可能受降水、气温等环境因素的限制。雨季由于降水量的增加导致地表径流加强,促使更多的有机质流入河流,导致人口密度与 CDOM 间的负相关性减弱,而建设用地受城市化影响,排放的大量污水与工业废料等流入河流,对 CDOM 的影响更为直接,因而建设用地覆盖率为 CDOM 为正相关。此外,由于芙蓉江流域为典型喀斯特地貌,地表径流可通过溶隙快速下渗至地下水系统<sup>[55,60]</sup>,减少了生活污水在地表水中的滞留时间,且地方政府近年来推行的农村污水集中处理项目有效降低人为源有机物的直接输入<sup>[61-62]</sup>,因此人口密度与 CDOM 的负相关可能与该区域地下水系统的稀释作用及环保措施相关。

总体而言,旱季与雨季的环境条件和人类活动的变化对芙蓉江流域河流 CDOM 的影响存在显著差异。理解驱动因素的影响作用对于制定有效的水资源管理策略至关重要。旱季应注重保护水源和控制污染物,而雨季则需加强径流管理和污染物的拦截,以维持河流生态系统的可持续发展。

#### 4 结论

(1) 芙蓉江流域河流 CDOM 时空分异较强,雨季 CDOM 吸光系数高于旱季,且雨季 CDOM 吸光系数整体高于旱季,时空异质性显著。

(2) 本研究利用 Sentinel-2A 卫星影像与地面实测数据,通过 AdaBoost 算法构建 CDOM 反演模型,结果表明模型具有良好的预测能力与稳定性,能够准确地反演芙蓉江流域河流在不同时段的 CDOM。

(3) 通过 LIME 算法对 CDOM 反演模型进行深入解析,研究表明不同的波段特征对模型具有不同的影响作用,提高了模型的透明度和可解释性,为 CDOM 反演研究提供了科学参考。

(4) 芙蓉江流域旱季与雨季 CDOM 受驱动因素影响作用不同,较高的气温和较少的降水量易导致 CDOM 吸光系数降低,土地利用在旱季对 CDOM 吸光系数有显著影响,而在雨季影响较弱,人口密度对 CDOM 吸光系数的负面影响可能与该区域地下水系统的稀释作用及环保措施相关。

#### 参考文献(References):

- [1] 杨平,傅伯杰. 土地利用变化与陆地—海洋的相互作用. 地球科学进展, 2000, 15(6): 723-728.
- [2] Huang J C, Zhang Y J, Bing H J, Peng J, Dong F F, Gao J F, Arhonditis G B. Characterizing the river water quality in China: recent progress and on-going challenges. Water Research, 2021, 201: 117309.
- [3] 左其亭,崔国韬. 人类活动对河湖水系连通的影响评估. 地理学报, 2020, 75(7): 1483-1493.
- [4] 吴志强,刘晓畅,刘琦,乔壬路,刘治宇,杨婷,严娟,姬凌云. 基于水资源约束的我国城市发展策略研究. 中国工程科学, 2022, 24(5): 75-88.
- [5] Zhang F T, Xie A Y, Jiang C X, Chen J, An Y Z, Yang P R, Ma D L. Coupling coordination analysis and spatiotemporal heterogeneity between urban land green use efficiency and ecosystem services in Yangtze River Economic Belt, China. Humanities and Social Sciences Communications, 2024, 11(1): 1328.
- [6] 张运林,秦伯强. 梅梁湾、大太湖夏季和冬季 CDOM 特征及可能来源分析. 水科学进展, 2007, 18(3): 415-423.
- [7] 高洁,江韬,李璐璐,陈雪霜,魏世强,王定勇,闫金龙,赵铮. 三峡库区消落带土壤中溶解性有机质(DOM)吸收及荧光光谱特征. 环境科学, 2015, 36(1): 151-162.
- [8] 张运林. 湖泊光学研究进展及其展望. 湖泊科学, 2011, 23(4): 483-497.
- [9] 张运林,秦伯强,杨龙元. 太湖梅梁湾水体悬浮颗粒物和 CDOM 的吸收特性. 生态学报, 2006, 26(12): 3969-3979.
- [10] 杨一鹏,王桥,王文杰,高士平. 水质遥感监测技术研究进展. 地理与地理信息科学, 2004, 20(6): 6-12.
- [11] 张运林. 水体中有色可溶性有机物的研究进展. 海洋湖沼通报, 2006, 28(3): 119-127.
- [12] 王波,黄津辉,郭宏伟,许旺,曾清怀,麦有全,祝晓瞳,田上. 基于遥感的内陆水体水质监测研究进展. 水资源保护, 2022, 38(3): 117-124.
- [13] 吴铭,黄珏,宫丽娇,江涛. 长时序 Landsat 的北极 Lena 河 DOC 浓度变化及驱动力分析. 遥感学报, 2021, 25(3): 830-845.

- [14] Cai X L, Wu L Y, Li Y M, Lei S H, Xu J, Lyu H, Li J D, Wang H J, Dong X Z, Zhu Y X, Wang G L. Remote sensing identification of urban water pollution source types using hyperspectral data. *Journal of Hazardous Materials*, 2023, 459: 132080.
- [15] Shang Y X, Liu G, Wen Z D, Jacinthe P A, Song K S, Zhang B, Lyu L L, Li S J, Wang X, Yu X F. Remote estimates of CDOM using Sentinel-2 remote sensing data in reservoirs with different trophic states across China. *Journal of Environmental Management*, 2021, 286: 112275.
- [16] Ling Z B, Sun D Y, Wang S Q, Qiu Z F, Huan Y, Mao Z H, He Y J. Remote sensing estimation of colored dissolved organic matter (CDOM) from GOCI measurements in the Bohai Sea and Yellow Sea. *Environmental Science and Pollution Research*, 2020, 27(7): 6872-6885.
- [17] Wang Y C, Shen F, Sokoletsky L, Sun X R. Validation and calibration of QAA algorithm for CDOM absorption retrieval in the Changjiang (Yangtze) estuarine and coastal waters. *Remote Sensing*, 2017, 9(11): 1192.
- [18] Zhu W N, Tian Y Q, Yu Q, Becker B L. Using Hyperion imagery to monitor the spatial and temporal distribution of colored dissolved organic matter in estuarine and coastal regions. *Remote Sensing of Environment*, 2013, 134: 342-354.
- [19] Li J W, Yu Q, Tian Y Q, Becker B L, Siqueira P, Torbick N. Spatio-temporal variations of CDOM in shallow inland waters from a semi-analytical inversion of Landsat-8. *Remote Sensing of Environment*, 2018, 218: 189-200.
- [20] Huang Y H, Pan J Y, Devlin A T. Enhanced estimate of chromophoric dissolved organic matter using machine learning algorithms from landsat-8 OLI data in the Pearl River estuary. *Remote Sensing*, 2023, 15(8): 1963.
- [21] 金则澎, 毛峰, 程乾, 李军, 张轩豪. 梯度提升回归树在千岛湖水体 CDOM 反演中的应用. *遥感信息*, 2022, 37(1): 110-118.
- [22] Zhang J W, Meng F, Fu P J, Jing T T, Xu J, Yang X Y. Tracking changes in chlorophyll-a concentration and turbidity in Nansi Lake using Sentinel-2 imagery: a novel machine learning approach. *Ecological Informatics*, 2024, 81: 102597.
- [23] Zhuang D Y, Shao T T, Zheng K. Characteristics, sources and driving factors of riverine CDOM in a severe erosion basin on the Loess Plateau, China. *Ecological Indicators*, 2023, 148: 110080.
- [24] Shao T T, Liang X W, Zhuang D Y, Zheng K, Wang T. Seasonal variations in CDOM characteristics and effects of environmental factors in coastal rivers, NorthEast China. *Environmental Science and Pollution Research International*, 2023, 30(11): 29052-29064.
- [25] 蒋筌, 陈文奇, 宁凡, 孙玉慧, 敖进, 罗维均, 况光显. 贵州高原北部河流阶地发育与喀斯特地貌演化. *地理研究*, 2021, 40(1): 81-92.
- [26] 刘睿, 张静, 陈祖胜, 倪茂飞, 刘文胜. 典型喀斯特河流水-气界面二氧化碳交换特性及其营养调控因素. *环境科学*, 2021, 42(2): 740-748.
- [27] 张柳柳, 刘睿, 郑达燕, 郑财贵, 张静. 土地利用对典型喀斯特河流溶解性有机质的直接和间接影响. *环境科学*, 2023, 44(10): 5511-5523.
- [28] Shao T T, Song K S, Du J, Zhao Y, Liu Z M, Zhang B. Retrieval of CDOM and DOC using in situ hyperspectral data: a case study for potable waters in NorthEast China. *Journal of the Indian Society of Remote Sensing*, 2016, 44(1): 77-89.
- [29] Xu J, Fang C Y, Gao D, Zhang H S, Gao C, Xu Z C, Wang Y Q. Optical models for remote sensing of chromophoric dissolved organic matter (CDOM) absorption in Poyang Lake. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, 142: 124-136.
- [30] Kim J, Jang W, Kim J H, Lee J W, Cho K H, Lee Y G, Chon K, Park S, Pyo J, Park Y, Kim S. Application of airborne hyperspectral imagery to retrieve spatiotemporal CDOM distribution using machine learning in a reservoir. *International Journal of Applied Earth Observation and Geoinformation*, 2022, 114: 103053.
- [31] Arnold J G, Moriasi D N, Gassman P W, Abbaspour K C, White M J, Srinivasan R, Santhi C, Harmel R D, van Griensven A, Van Liew M W, Kannan N, Jha M K. SWAT: model use, calibration, and validation. *Transactions of the ASABE*, 2012, 55(4): 1491-1508.
- [32] Shao T T, Wang T. Effects of land use on the characteristics and composition of fluvial chromophoric dissolved organic matter (CDOM) in the Yiluo River watershed, China. *Ecological Indicators*, 2020, 114: 106332.
- [33] Olmanson L G, Page B P, Finlay J C, Brezonik P L, Bauer M E, Griffin C G, Hozalski R M. Regional measurements and spatial/temporal analysis of CDOM in 10, 000+ optically variable Minnesota lakes using Landsat 8 imagery. *Science of the Total Environment*, 2020, 724: 138141.
- [34] Lyu L L, Wen Z D, Jacinthe P A, Shang Y X, Zhang N, Liu G, Fang C, Hou J B, Song K S. Absorption characteristics of CDOM in treated and non-treated urban lakes in Changchun, China. *Environmental Research*, 2020, 182: 109084.
- [35] He Q, Wang M, Liu K, Li K W, Jiang Z Y. GPRChinaTemp1km: a high-resolution monthly air temperature data set for China (1951—2020) based on machine learning. *Earth System Science Data*, 2022, 14(7): 3273-3292.
- [36] Peng S Z, Ding Y X, Liu W Z, Li Z. 1 km monthly temperature and precipitation dataset for China from 1901 to 2017. *Earth System Science Data*, 2019, 11(4): 1931-1946.
- [37] Tatem A J. WorldPop, open data for spatial demography. *Scientific Data*, 2017, 4: 170004.
- [38] Freund Y, Schapire R E. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 1997, 55(1): 119-139.
- [39] Ribeiro M T, Singh S, Guestrin C. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. // *Proceedings of the 22nd ACM*

- SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco California USA. ACM, 2016; 1135-1144.
- [40] 杨延梅, 向维, 苏靖, 陈文婷, 傅雪梅, 虞敏达, 孙源媛, 郑明霞. 基于 BP 神经网络的白洋淀水环境承载力研究. 中国农村水利水电, 2021(7): 61-66, 71.
- [41] 陈岩, 冯亚中. 基于 RS-SVR 模型的流域水资源脆弱性评价与预测研究——以黄河流域为例. 长江流域资源与环境, 2020, 29(1): 137-149.
- [42] Clark J M, Lane S N, Chapman P J, Adamson J K. Export of dissolved organic carbon from an upland peatland during storm events; Implications for flux estimates. *Journal of Hydrology*, 2007, 347(3/4): 438-447.
- [43] 孙启元, 江娟, 吴春山, 王菲凤, 吴延培. 山仔水库间隙水不同分子质量 CDOM 来源及分布规律. 福建师范大学学报: 自然科学版, 2017, 33(4): 92-99.
- [44] Barredo Arrieta A, Diaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, Garcia S, Gil-Lopez S, Molina D, Benjamins R, Chatila R, Herrera F. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 2020, 58: 82-115.
- [45] McGovern A, Lagerquist R, John Gagne D, Jergensen G E, Elmore K L, Homeyer C R, Smith T. Making the black box more transparent: understanding the physical implications of machine learning. *Bulletin of the American Meteorological Society*, 100(11): 2175-2199.
- [46] 袁博, 郭梦京, 周孝德, 郑兴. 气候变化及人类活动对河流溶解性有机质(DOM)影响的研究进展. 长江流域资源与环境, 2018, 27(7): 1604-1614.
- [47] Xu J, Jian Z J, Wang Y Q, Fang C Y, Hu Q W. Spatial-seasonal characteristics and influencing factors of dissolved organic carbon and chromophoric dissolved organic matter in Poyang Lake. *Environmental Earth Sciences*, 2023, 82(1): 44.
- [48] 张珍明, 周运超, 黄先飞, 田潇. 喀斯特小流域土壤有机碳密度空间异质性及影响因素[J]. 自然资源学报, 2018, 33(02): 313-324.
- [49] 吴敏, 刘淑娟, 叶莹莹, 张伟, 王克林, 陈洪松. 喀斯特地区坡耕地与退耕地土壤有机碳空间异质性及其影响因素. 生态学报, 2016, 36(6): 1619-1627.
- [50] Xiao S S, Ye Y Y, Xiao D, Chen W R, Zhang W, Wang K L. Effects of tillage on soil N availability, aggregate size, and microbial biomass in a subtropical Karst region. *Soil and Tillage Research*, 2019, 192: 187-195.
- [51] 习丹, 翁浩东, 胡亚林, 吴建平. 林冠氮添加和林下植被去除对杉木林土壤有机碳组分的影响. 生态学报, 2021, 41(21): 8525-8534.
- [52] Ushio M, Balsler T C, Kitayama K. Effects of condensed tannins in conifer leaves on the composition and activity of the soil microbial community in a tropical montane forest. *Plant and Soil*, 2013, 365(1): 157-170.
- [53] Zhou Q, Zhang P, Wang Z Q, Wang L X, Wang S B, Yang W T, Yang B J, Huang G Q. Winter crop rotation intensification to increase rice yield, soil carbon, and microbial diversity. *Heliyon*, 2023, 9(1): e12903.
- [54] Zhou L, Zhou Y Q, Tang X M, Zhang Y L, Jang K S, Székely A J, Jeppesen E. Resource aromaticity affects bacterial community successions in response to different sources of dissolved organic matter. *Water Research*, 2021, 190: 116776.
- [55] 肖旭芳, 张双龙, 郭雯, 王敬富, 杨海全, 廖鹏. 西南喀斯特地区地表水和地下水环境污染特征与研究展望. 地球与环境, 2023, 51(5): 564-573.
- [56] Wu R, Wang H Y, Li W, Chen X W, Zhou Q C, Chang J J. Temporal variations in dissolved organic matter properties in a eutrophic lake under water dilution and aquatic plant litter decomposition: a simulated microcosm study revealing the environmental and microbial driving factors. *Journal of Water Process Engineering*, 2024, 65: 105837.
- [57] 邓继峰, 丁国栋, 魏亚伟, 赵国平, 张若菡, 邓舸, 周永斌, 殷有. 毛乌素沙地南缘樟子松人工林土壤理化性质差异及其与林分生长的关系. 干旱区资源与环境, 2017, 31(8): 160-166.
- [58] 王立英, 张润宇, 陈敬安. 贵州高原红枫湖水理化特征与碳氮硅的时空分布研究. 地球与环境, 2017, 45(4): 383-389.
- [59] 张羿, 王晗, 马永梅, 杜佳杰, 龙靖潇, 郎家锐, 倪茂飞. 溶解性有机碳耦合营养状态解析河流 CO<sub>2</sub> 动态及驱动因素. 湖泊科学, 2024, 36(3): 794-806.
- [60] 苑浩达, 刘睿, 倪茂飞, 王志康, 吴昌妮. 典型喀斯特河流溶解性有机质生物降解特性及其温度敏感性研究. 环境科学学报, 2022, 42(3): 218-226.
- [61] 吴用, 李畅游, 张成福, 史小红, 赵胜男, 林涛涛. 基于 ArcGIS 和 SWAT 的山区、平原区复合地貌流域划分方法研究. 干旱区地理, 2016, 39(2): 413-419.
- [62] 邓光耀, 张忠杰. 基于网络 SBM-DEA 模型和 GML 指数的中国各省工业用水效率研究[J]. 自然资源学报, 2019, 34(07): 1457-1470.