白薇扬,柳 睿,江豪浩.水-气界面汞通量动态预测模型构建——以长寿湖为例 [J]. 中国环境科学, 2026,46(1):20-28.
Bai W Y, Liu R, Jiang H H. Dynamic prediction model for mercury flux at the water-air interface: A case study of Changshou Lake [J]. China Environmental Science, 2026,46(1):20-28.

水-气界面汞通量动态预测模型构建——以长寿湖为例

白薇扬[1,2*],柳 睿[1],江豪浩[1] (1.重庆理工大学化学化工学院,重庆 400054；2.西南大学资源环境学院,重庆 400715)

摘要：针对湖泊系统汞通量估算受多环境变化因素协同影响,不确定性明显,以喀斯特岩溶的人工水库长寿湖为研究对象,综合运用野外采样,通径分析与机器学习方法来揭示汞通量交换机制.研究发现:汞通量受多因子直接/间接的协同调控.其中,水体汞浓度(0.897)、紫外线强度(0.463)和光照强度(0.446)的直接效应主导着水体溶解气态汞 Hg^0(DGM)生成,而风速通过光照强度(0.226)的间接影响强于其直接扩散作用(0.197).基于机器学习构建的汞通量预测混合模型(0.9GBR(梯度提升机)+0.05RF(随机森林)+0.05SVR(支持向量机回归)),在保持预测精度的同时提升鲁棒性和稳定性,实现湖泊多参数耦合下汞通量的动态预测,为区域汞环境风险评估与管控提供决策支持.

关键词：汞通量；水-气界面；机器学习；长寿湖
中图分类号：X142 **文献标识码**：A 文章编号：1000-6923(2026)01-0020-09

Dynamic prediction model for mercury flux at the water-air interface: A case study of Changshou Lake. BAI Wei-yang[1,2*], LIU Rui[1], JIANG Hao-hao[1] (1.College of Chemistry and Chemical Engineering, Chongqing University of Technology, Chongqing 400054, China；2.College of Resources and Environment, Southwest University, Chongqing 400715, China). *China Environmental Science*, 2026,46(1)：20~28

Abstract：Quantifying Hg flux in lake systems remains challenging due to the synergistic influences of multiple environmental drivers. To address this, we investigated Changshou Lake, an artificial reservoir situated in a karst region, through an integrated approach that included field sampling, path analysis, and machine learning, to elucidate the mechanisms of Hg flux exchange. The results indicated that Hg flux was jointly regulated by multiple factors via both direct and indirect pathways. Water Hg concentration (0.897), UV radiation (0.463), and solar radiation (0.446) were identified as exerting dominant direct effects on the production of dissolved gaseous mercury (DGM, Hg^0). In contrast, the influence of wind speed on Hg flux was observed to be stronger through its indirect enhancement of solar radiation (0.252) than through direct diffusion (0.197). A hybrid machine-learning prediction mode [0.9 Gradient Boosting Regression (GBR) + 0.05 Random Forest (RF) + 0.05Support Vector Regression (SVR)] was developed. This model was shown to improve generalization ability and stability while maintaining predictive accuracy, enabling dynamic prediction of Hg flux under the coupling of multiple lake parameters. The approach provides decision support for regional Hg risk assessment and management.

Key words：mercury flux；water-air interface；machine learning；Changshou Lake

汞作为具有神经毒性、生物累积性与跨境迁移能力的全球性污染物,其环境释放通量直接关系着环境生态的暴露风险.研究显示,全球约 60%的人为源汞排放通过大气-水体界面交换进入水生生态系统,并在沉积物-水界面经甲基化作用生成剧毒甲基汞(MeHg),通过食物链放大对人类健康构成严重威胁[1-7].汞的环境归趋不仅取决于大气汞含量,更受控于汞在水体赋存形态.因此加强水-气界面的汞通量研究对于提供汞源数据支持和区域汞环境评价具有重要意义.

汞的生物地球化学循环研究逐步揭示其形态转化与迁移机制的关键过程.早期研究发现汞的络合形态对其环境行为具有决定性影响[1].随后研究证实,Hg^{2+}在日光辐射下可发生光化学还原生成气态 Hg^0,显著增强水-气界面汞交换通量[2].全球观测数据进一步印证,光照强度是导致汞迁移纬度分异的关键因素,强光照区域水体中 Hg^0 比例可达 45%,远高于高纬度地区的 12%[3].

近年来,对光化学反应参数与环境汞通量的耦合建模,推动了机制研究的深入[8],仍受限于模型的静态架构,对温度、pH 值、风速等多种变量与汞通

收稿日期：2025-05-14
基金项目：国家自然科学基金面上项目(41373113)
* 责任作者, 副教授, baiwy@cqut.edu.cn

量变化的交互作用缺乏系统解析,难以全面揭示实际环境因子间的动态非线性相互作用.因此,发展动态耦合模型对于准确模拟湖泊汞通量至关重要.

针对上述局限,对喀斯特岩溶的长寿湖水体−大气界面的汞通量,开展“驱动识别−模型优化”的多参数耦合动态预测研究.通过 pH 值、溶解氧、有机质等水化学参数,结合风速、光照、温湿度等环境要素实时监测数据,通径分析定量解析各因子的交互作用权重,实现汞通量随环境驱动因子的响应演变.基于此,根据四季汞通量时序数据集,利用机器学习,进一步对比并优化随机森林(RF)、支持向量回归(SVR)、梯度提升机(GBR)、决策树(DT)及 LightGBM 等机器学习算法的时空适应性,通过特征工程优化与超参数调节提升模型泛化能力.研究结果可提供不同水气环境条件下的汞释放通量动态预测,为区域汞环境风险评估与管控提供科学依据和技术支撑.

1　研究方法

1.1　研究区概况

长寿湖是三峡库区东北部重要的支流型水库,位于川东褶皱带与川中丘陵过渡区,基底和部分区域夹杂三叠系石灰岩,是研究喀斯特岩溶梯级水库汞生物地球化学行为的典型场景[7].该水域属亚热带季风气候区,年均水温 18.3℃,其分层型水文特征与周期性蓄排水过程共同塑造了独特的汞迁移环境.

本研究选择 5 个采样点,采样点 S1(29°55′3″N, 107°15′29″E)位于水库大坝前,为长寿湖出水口.S2 (29°59′3″N, 107°16′23″E)位于水库西部,为龙溪河入水口,龙溪河流经垫江县和梁平县,河两岸分布密集村落和大片果园.S3 (29°58′21″N,107°18′58″E)位于水库北部回水区,水力交换时间长.S4 (29°57′5″N, 107°20′19″E)位于水库东部,为另一主要涞水水口,两岸果园分布广泛,村落分布较少.S5(29°56′30″N, 107°17′45″E)位于水库中部,北部岛屿分布成片果园,西部和南部岛屿旅游业发达(图 1).根据长寿湖地区的气候特征,分四季进行监测采样,春季采样时间为 5 月 5~6 日、夏季采样时间为 7 月 1~2 日、秋季采样时间为 10 月 24~25 日、冬季采样时间为 1 月 8~9 日.在各采样点利用水样采集器(Nisiki)每隔 4m 分层采集水样,测定溶解气态汞、各项水质参数.其中,在采样点 S1 利用动力通量箱(DFC)法连续 48h 监测水−气界面汞交换通量.

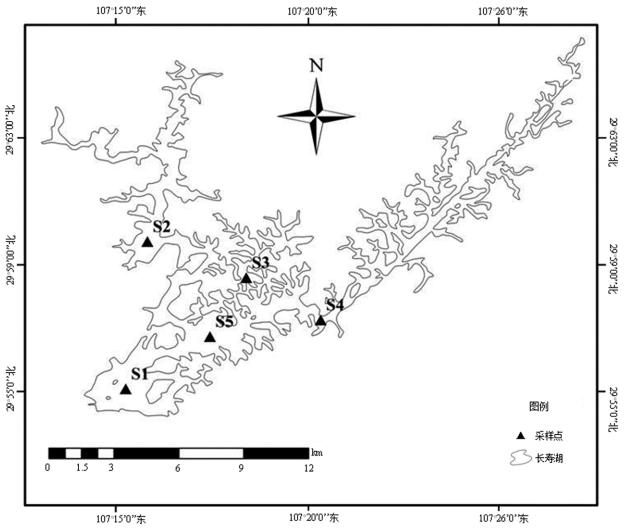


图 1　长寿湖采样地点

Fig.1　Sampling locations in Changshou Lake

审图号:GS(2019)3333,底图无修改

1.2　数据来源及采样方法

水−气界面汞交换通量用 DFC 结合 Tekran 2537X 在线汞蒸气分析仪[9]测量.半圆柱型的 DFC 采用厚度为 5mm 的石英玻璃制作,底面积 0.08m^2 (200mm×600mm),体积为 0.024m^3.

汞通量交换(F)用如下公式进行计算:

$$F=\frac{(C_0-C_i)Q}{A} \tag{1}$$

其中: F 为汞交换通量,ng/(m^2·h);C_o 和 C_i 分别为流入和流出通量箱气态汞含量,ng/m^3;Q 为通量箱内空气流量,m^3/h;A 为通量箱的底面积,m^2.

太阳辐射用 TES−II Digital Lumino−meter[10]进行测定.水温用 Flash Check Pocket Probe Digital Thermometer 11000[11]进行测定.其他环境参数(包括大气温度,大气湿度以及大气压)用 Kestrel4000pocket Weather Tracker[12]测定.pH 值用 PT−10[13]测定.DO 用 YSI 550A[14]测定;DOC 使用 Sievers InnovOx Laboratory TOC Analyzer[15]进行分析.

溶解气态汞采样取分层水样 800mL,迅速转移至超净处理的硼硅玻璃气泡瓶中,然后通入 300mL/min 流速纯净无汞氩气将水体中溶解气态

汞富集至空白金富集管中,通气 30min.取下金管,两头用 PTFE 塞密封,立刻运回实验室用冷原子荧光光谱法测定[16].

水样分析参照 Method 1631[9]进行.分别取未过滤水样和过滤水样经氯化溴氧化,氯化亚锡还原,金管富集-冷原子荧光光谱法(CVAFS)测定,得出水样 THg、DHg 浓度,将 THg-DHg=PHg.取未过滤水样直接用氯化亚锡还原,金管富集- CVAFS 测定,得出水样 RHg 浓度.采用不大于样品最小值 20%的方法空白、实验器皿空白、野外空白确保样品各形态汞的质量分析准确性[16].分析过程中的质量控制采用平行样控制及加标回收,加标回收率为 80%~112%,并对 5%的样品进行重复测定,相对标准偏差<9%.

1.3 模型原理

GBR 通过分阶段迭代构建决策树序列,逐步优化预测残差.在回归问题中,梯度提升算法的目标是最小化预测值与真实值之间的均方误差(MSE).每一轮训练,模型计算残差(真实值与当前模型预测值之间的差异),然后拟合出新的弱学习器来预测这些残差,从而逐步减小整体模型的误差.

RF 基于 Bootstrap 重采样生成多棵独立决策树,通过平均化预测结果以降低模型方差,如图 2 所示.每棵树在节点分裂时仅随机选择部分特征(如总特征数的平方根),以此增强基学习器多样性.其集成策略通过“多数投票”或“均值聚合”提升泛化能力,但对序列依赖性数据的适应性较弱.

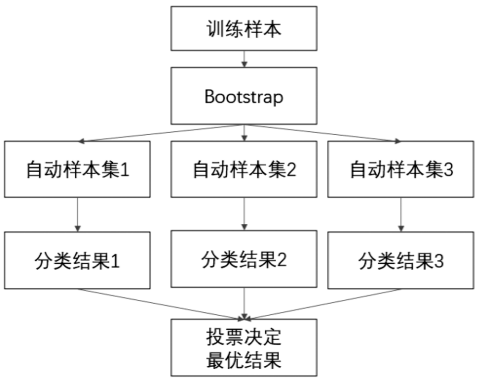


图 2 RF 原理示意

Fig.2 Schematic diagram of RF principle

SVR 利用核函数将输入数据映射至高维特征空间,构建 ε-insensitive 损失函数下的最优超平面.该函数对预测值在真实值±ε 范围内的误差不予惩罚,以此平衡模型复杂度与噪声抑制.其全局优化特性可避免局部极小陷阱,但对异质误差分布的敏感性较高.

DT 通过递归二分法划分特征空间,以信息增益(ID3)、增益率(C4.5)或基尼指数(CART)为分裂准则,构建树形结构,如图 3 所示.其预测结果为叶节点区域的样本均值或众数.单棵树因结构简单易解释,但易因深度限制导致欠拟合或噪声敏感性.

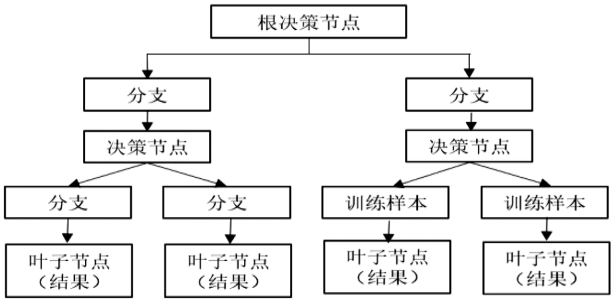


图 3 DT 原理示意

Fig.3 Schematic diagram of DT principle

LGBM 采用 leaf-wise 垂直生长策略,优先分裂损失下降最大的叶子节点,以提升计算效率.通过直方图算法离散化连续特征,减少内存占用与分裂点计算量.其基于梯度单边采样(GOSS)与专属特征绑定(EFB)优化数据分布处理,但对噪声数据的过拟合风险需通过正则化约束.

1.4 评价指标

模型性能评价采用决定系数(R^2)、均方根误差(RMSE)、平均绝对误差(MAE)和性能偏差比(RPD)进行评价.

决定系数(R^2)来评估模型对数据的拟合程度.R^2 值越高,表示模型对数据的拟合越好.R^2 的判断标准为: R^2>0.90 表示模型回归结果出色;R^2 在 0.81~0.90 之间表示模型回归结果很好; R^2 在 0.66~0.80 之间为模型回归结果一般; R^2<0.66 表示模型回归很差.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(f_i - y_i)^2}{\sum_{i=1}^{n}(f_i - \overline{y}_i)^2} \tag{2}$$

式中:f_i 为真实值,y_i 为预测值,$\overline{y}_i$ 为真实值的平均值.

均方根误差(RMSE):

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(f_i-y_i\right)^2} \tag{3}$$

平均绝对误差(MAE):

$$\mathrm{MAE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left|f_i-y_i\right|} \tag{4}$$

RMSE 和 MAE 越小,模型的性能越好.

性能偏差比(RPD):

RPD 是评估回归模型预测能力的重要指标,反映模型预测误差与数据自身变异程度的相对关系.

$$\mathrm{RPD}=\frac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(f_i-\overline{f}\right)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(f_i-y_i\right)^2}} \tag{5}$$

RPD>3:模型预测能力极佳;2<RPD≤3:模 RPD≤1.5:模型预测能力较差.

2　结果与分析

本研究采用 DFC 结合 Tekran 2537X 在线汞蒸气分析仪测量了研究区域四季汞通量 96 组,协同同步收集的风速、气温湿度、气压、光照、紫外线和水温等环境数据,用以探究环境因素对长寿湖水体-大气界面汞交换的静态和动态影响效果.

2.1　通径分析

通径分析是一种基于回归统计学的因果关系解析方法,旨在解析环境因子对汞交换过程的差异化调控机制.通过计算标准化通径系数,将总效应分解为直接作用与间接作用,并依据标准化回归系数量化路径权重.从而有效揭示多因子系统中各变量的调控强度及交互机制.

从表 1 中可以看出,长寿湖水/气界面汞交换影响总效应最大的是光照强度(1.697),其次为紫外线强度(1.641).影响因素的效应分为两种:直接效应和间接效应.直接效应前三位影响因子分别为水体汞浓度、紫外线强度、光照强度,影响值为 0.897、0.463 和 0.446,即水体-大气汞交换通量由水体汞浓度、光照强度、紫外线强度因素直接引发.根据量子理论,太阳辐射通过提供紫外光将低能态 HgCl 激发至高能态,从而将表层水中 Hg^{2+}向 Hg^0 转化,同时增加水体热量提高水体温度,进而驱动过饱和的 DGM 从水体表面向大气的释放[17].长寿湖水库夏季光照强度(29.7±1.837)klux 和紫外线强度 (1485.4±1174)μW/cm^2 远高于冬季的光照强度(1.7±1.8) klux 和紫外线强度(50.5±16.6)μW/cm^2,且光照强度和紫外线强度与水气界面汞交换通量呈显著正线性相关关系,相关系数分别为 0.606 和 0.543(P<0.01; ANOVA).说明光照强度和紫外线强度对 DGM 生成有着重要促进影响.Vette 等[18]发现 UV-A (320~400nm)辐射增加,DGM 的形成速率明显增加.阴永光等[2]研究结果指出光照条件能够促进 Hg(II) 溶液中溶解性气态汞的产生.何天容等[19]在红枫湖研究显示光照促进汞光致还原,从而表层水中产生更多 DGM.因此,长寿湖水库夏季的高光照强度以及高紫外线强度直接促使水体汞光致还原反应发生,从而增加水气界面汞交换通量.

表 1　长寿湖水气汞通量影响因素通径分析

Table 1　Path analysis of influencing factors of water-air mercury flux in Changshou Lake

变量	路径效应								通径系数		总效应	效应路径比重 (%)							
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	直接效应	间接效应		X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
X_1	0.197	0.005	0.069	0.033	0.252	0.226	0.075	0.215	0.197	0.809	1.006	18.37	0.50	6.46	3.08	23.54	21.04	6.97	20.05
X_2	0.010	0.027	0.002	0.008	0.008	0.013	0.012	0.020	0.027	0.030	0.057	10.03	27.08	2.46	8.00	7.85	12.83	12.03	19.72
X_3	0.064	0.000	0.212	0.012	0.232	0.222	0.052	0.312	−0.212	−0.896	−1.107	5.79	0.00	19.17	1.08	20.98	20.07	4.70	28.21
X_4	0.051	0.002	0.020	0.129	0.083	0.070	0.013	0.088	0.129	−0.108	0.021	11.19	0.37	4.31	28.36	18.21	15.36	2.85	19.34
X_5	0.111	0.000	0.110	0.024	0.446	0.409	0.134	0.510	0.446	1.251	1.697	6.38	0.03	6.32	1.38	25.56	23.42	7.69	29.22
X_6	0.096	0.001	0.102	0.020	0.394	0.463	0.149	0.456	0.463	1.178	1.641	5.71	0.05	6.06	1.19	23.44	27.55	8.87	27.14
X_7	0.046	0.001	0.034	0.005	0.186	0.215	0.321	0.252	0.321	0.727	1.048	4.32	0.09	3.25	0.47	17.58	20.26	30.27	23.76
X_8	0.047	0.001	0.074	0.013	0.253	0.235	0.090	0.897	0.897	0.713	1.610	2.93	0.04	4.58	0.79	15.73	14.60	5.60	55.73

注:X_1为风速;X_2为气温;X_3为相对湿度;X_4为大气压;X_5为光照强度;X_6为紫外线强度;X_7为水体温度;X_8为汞通量浓度

间接效应前三位影响因子分别为光照强度、紫外线强度、风速,影响值分别为 1.251,1.178,0.809.其

中风速通过光照强度间接作用(0.252)对水−气界面汞交换通量的影响高于风速本身(0.197)直接影响,说明长寿湖水库水气界面汞交换通量受到水体中DGM生成以及Hg^0浓度梯度扩散两方面制约:一方面光照强度、紫外线强度、水体汞浓度直接或间接促进着水体DGM的生成;另一方面风速越大,降低了水面上方大气汞浓度,有助于Hg^0从水体向大气释放扩散,从而增加水−气界面交换通量.

2.2 机器学习模型构建与分析

汞的生物地球化学循环受多环境因素协同影响,在通径分析的基础上,考虑到汞通量的动态变化性和为了充分捕捉复杂变量间的交互效应,本研究采用随机森林(RF)、支持向量机回归(SVR)、梯度提升机(GBR)、决策树(DT)和LightBGM这5种机器学习模型整合多源环境参数与汞形态数据,构建汞通量回归模型.

2.2.1 数据预处理 本研究首先进行系统化的数据预处理,如图4所示.首先对数据进行垂直合并与特征筛选,保留前12列核心环境参数;然后剔除含缺失值(NaN)及无效标识符,同步实施数值标准化处理.为消除数据采集时序性偏差,采用随机排列算法(random_state=42)对样本顺序进行重排,进一步建立基于K−最近邻算法(KNNImputer, n_neighbors=5)的非线性插补模型,实现缺失值的稳健估计,最终生成规范化的数据集.

这一系列处理步骤有效解决了原始数据中的格式混乱、单位不统一和缺失值问题,为后续建立汞元素迁移模型提供了可靠的数据基础.

2.2.2 实验结果讨论 在预处理后,本研究通过不同的算法对数据进行回归,各评价指标如表2所示.基于表2的各模型评价指标,发现GBR在测试集上表现最优,模型的R^2=0.93,RMSE=3.50、MAE=1.26,显著优于其他模型,表明其对汞通量时空变异具有较好的解释力与预测精度.RF模型测试集R^2=0.83,RMSE=8.08,体现较强的泛化能力,但残差(RMSE)较GBR高130%,反映其对复杂非线性关系的捕捉能力有限.SVR模型的测试集R^2=0.78,RMSE=10.18,显示预测性能较不理想,DT与LGBM的测试集R^2分别为0.70与0.48,LGBM出现过拟合(训练集RMSE=16.66,测试集RMSE=24.15)现象,表明单一树模型对多因子交互效应建模存在局限性.

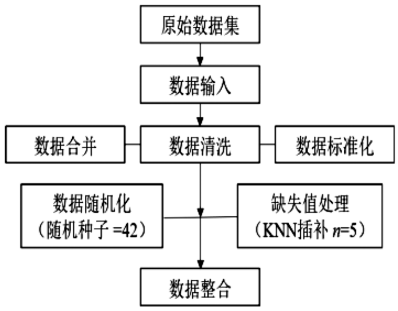


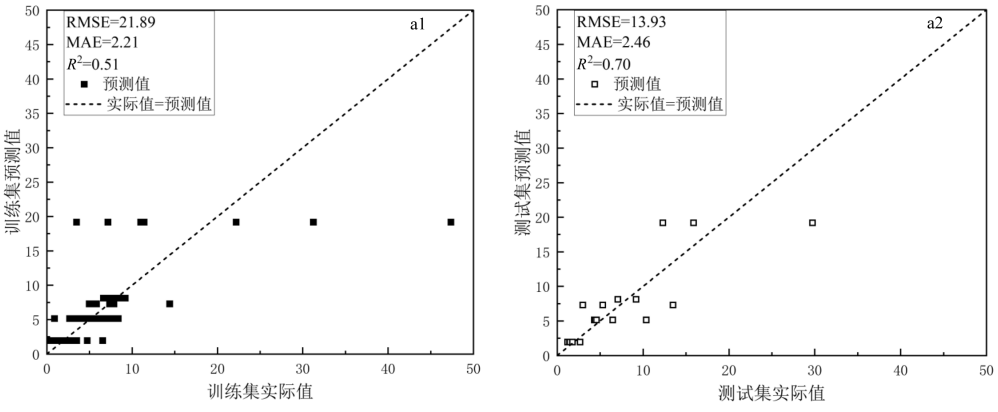
图4 数据预处理流程示意

Fig.4 Schematic diagram of data preprocessing process

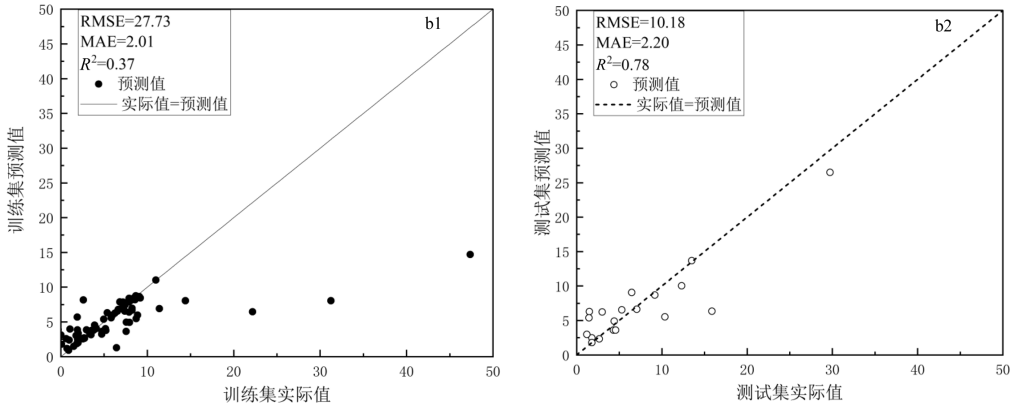
表2 各模型的评价指标

Table 2 Evaluation indicators of each model

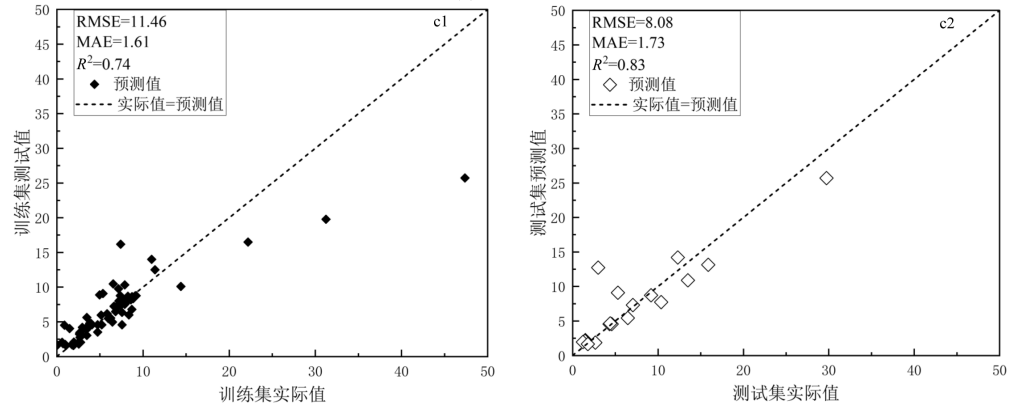
模型	训练集评价指标			测试集评价指标			RPD
	RMSE	MAE	R^2	RMSE	MAE	R^2	
DT	21.89	2.21	0.51	13.93	2.46	0.70	1.47
SVR	27.73	2.01	0.37	10.18	2.20	0.78	1.35
RF	11.46	1.61	0.74	8.08	1.73	0.83	2.02
GBR	8.37	1.50	0.81	3.50	1.26	0.93	2.45
LGBM	16.66	1.96	0.62	24.15	2.69	0.48	1.56



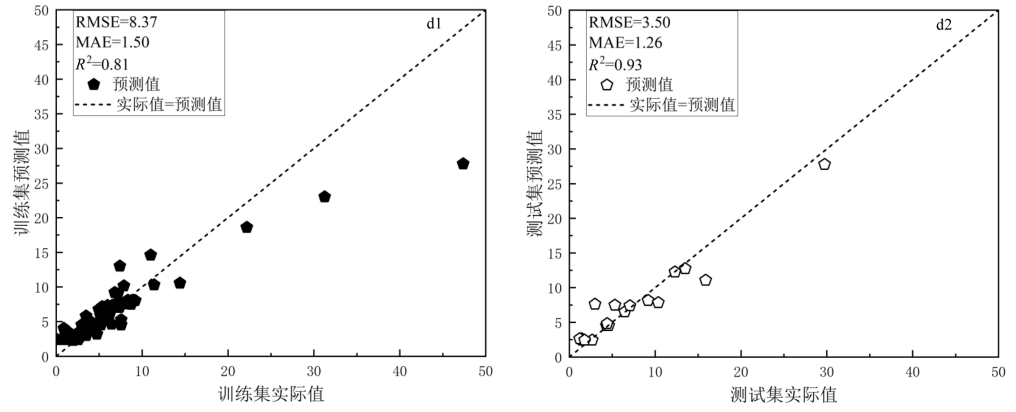
(a) DT模型



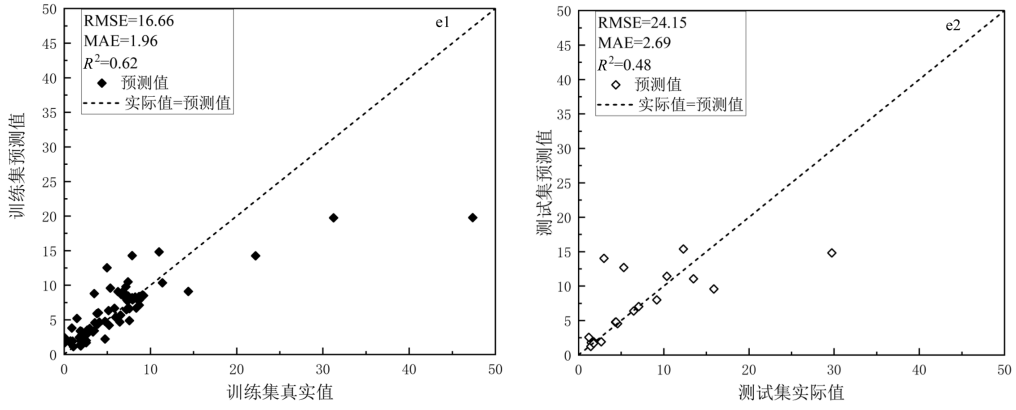
(b) SVR 模型



(c) RF 模型



(d) GBR 模型



(e) LGBM 模型

图 5　各模型的训练集和测试集数据对比

Fig.5　Comparison of training and testing Set Data of the Models

与传统的水环境汞及重金属研究相比,如表 3 所示, Janssen 等[20]通过单一的静态数学统计模型主成分分析法(PCA)和同位素分馏模型解析汞来源及其在鱼类中的累积机制.Atangana 等[21]通过主成分分析和层次聚类分析(HCA)评估喀麦隆金矿流域地表水质受汞污染主导并且揭示了污染源.Zhou 等[8]耦合 Princeton Ocean Model 的三维数值模型模拟了水体汞的分布和迁移.Crocker 等[22]通过为期一年的野外观测,记录了 Cane Creek 湖汞气水交换通量的季节变化,用双薄膜模型计算汞排放通量的季节变化.Sullivan 等[23]运用线性混合效应模型和逻辑回归模型,系统分析康涅狄格州 56 个水体中大口黑鲈汞浓度的长期变化(1995~2021 年),并且研究鱼类生长、食物网结构及遗留汞释放对污染趋势的影响.近年来,随着机器学习方法在水体重金属研究日渐广泛,呈现从单一模型向多模型耦合的发展趋势.Huynh 等[24]以台湾 10 个地下水盆地为研究区,基于低成本理化参数,用机器学习模型预测浅层含水层中 As、Fe、Mn 的浓度,评估模型不确定性和可解释性.其中 RF 模型最优,在 As 预测中 R^2=0.80;在 Fe 预测中 R^2=0.65;在 Mn 预测中 R^2=0.75.陈亚松等[25]基于水质指标、气候指标、湿地运行参数,收集以往研究文献数据,通过 3 种机器学习模型预测人工湿地出水氨氮(NH_4^+-N)、COD、磺胺甲噁唑(SMX)以及部分重金属的浓度.其中,RF 模型在研究重金属浓度中表现最好 R^2 在 0.94~0.99 之间.本研究对汞通量进行多环境因素的动态分析预测,如图 5 所示,GBR 模型的 R^2 为 0.93,说明有较好的预测分析能力.

针对出现的单一树模型对多因子交互效应建模存在局限性的问题,本研究进一步采用了融合模型对汞通量进行回归.

表 3 不同研究模型在水环境汞及重金属中的应用

Table 3 Application of different research models in mercury and heavy metals in aquatic environments

年份	主要模型	研究内容	参考文献
2019	主成分分析	通过 PCA 和同位素分馏模型解析汞来源及其在鱼类中的累积机制	[20]
2019	Princeton Ocean Model 的三维数值模型	构建三维数值模型模拟水体汞分布与迁移	[8]
2020	DFC、双薄膜模型	用双薄膜模型计算汞排放通量	[22]
2022	RF、SVR、KNN、MLP、GBR、LR	基于低成本理化参数,用机器学习模型预测浅层含水层中 As、Fe、Mn 的浓度	[24]
2023	PCA、HCA	评估水体汞污染和影响因素	[21]
2023	线性混合效应模型 逻辑回归模型	分析水体中大口黑鲈汞浓度的长期变化	[23]
2025	RF、XGB、LGB	基于水质指标、气候指标、湿地运行参数,通过 3 种机器学习模型预测人工湿地出水氨氮、COD、磺胺甲噁唑以及部分重金属的浓度	[25]

注:PCA:主成分分析;DFC:动态通量室; RF:随机森林; SVR:支持向量回归; KNN:K近邻算法; MLP:多层感知机; GBR:梯度提升回归; LR:逻辑回归; HCA:层次聚类分析; XGB:极端梯度提升; LGB:轻量级梯度提升.

表 4 新模型的评价指标对比

Table 4 Comparison of evaluation indicators for the new model

模型	训练集评价指标			测试集评价指标			RPD
	RMSE	MAE	R^2	RMSE	MAE	R^2	
DT	21.89	2.21	0.51	13.93	2.46	0.70	1.47
SVR	27.73	2.01	0.37	10.18	2.20	0.78	1.35
RF	11.46	1.61	0.74	8.08	1.73	0.83	2.02
GBR	8.37	1.50	0.81	3.50	1.26	0.93	2.45
LGBM	16.66	1.96	0.62	24.15	2.69	0.48	1.56
0.9GBR+0.05RF+0.05SVR	9.13	1.53	0.79	3.79	1.32	0.92	1.77

2.2.3 机器学习模型的优化及结果评价 在探究多环境因素影响的建模中,算法选择需同时考虑数据统计特性和过程物理本质的匹配度,单一模型难以全面捕捉其规律.考虑到影响汞通量数据的环境因素较为复杂,以及各模型在研究汞通量的优点的差异性,本研究采用加权融合策略(0.9GBR+0.05RF+0.05SVR),其评价指标如表 4 所示.融合模型在测试集上 R^2=0.92,图 6 所示,和单一模型相比降低了过拟合风险(测试集 R^2 从 0.81 降至 0.79).

本研究采用 GBR 作为主导模型,主导了环境因素对汞通量变化回归分析的解释力; SVR 的 ε-insensitive 损失函数[26] (ε=0.05,匹配传感器测量误差阈值±5%),通过容忍预测值在真实值±ε 区间内的浮动,有效吸收光照强度波动引发的随机脉冲噪声,避免 GBR 对低信噪比数据的过拟合倾向.RF 的

多数投票机制[27-30]可部分抵消 GBR 因梯度迭代产生的路径依赖性偏差,该架构在保留 GBR 核心预测能力的同时,利用 SVR 吸收测量噪声[31]与 RF 抑制异常扰动,测试集 R^2 从 0.93 降至 0.92,验证了多算法协同对汞通量复杂动力学的解析优势.

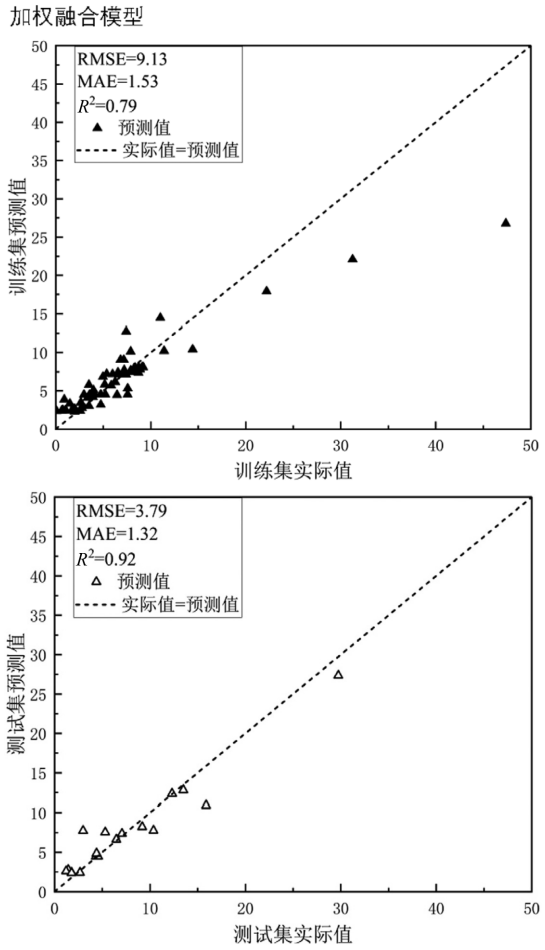


图 6　加权融合模型的训练集和测试集数据对比

Fig.6　Comparison of training and testing data of weighted fusion model

3　结论

3.1　多因子耦合作用路径:通径分析表明,汞通量受直接与间接效应双重调控.直接效应主要来自水体汞浓度(0.897)、紫外线强度 (0.463)和光照强度(0.446),通过光化学还原路径主导 DGM 生成.间接效应方面,紫外线(1.178)和光照强度(1.251)通过增强光反应活性产生倍增效应,而风速对通量的影响主要体现为通过增强光照强度(0.252)的间接作用,其贡献度显著高于直接扩散效应(0.197).

3.2　动态预测模型构建:通过整合机器学习算法,构建汞通量多尺度预测模型.梯度提升机(GBR)模型展现最优解释力(R^2=0.93),但存在过拟合风险.开发混合模型(0.9GBR+0.05RF+0.05SVR),在保持预测精度(测试集 R^2=0.92)的同时提升模型鲁棒性和稳定性,降低对特定数据模式的过度依赖.该模型提供不同水气环境条件下的汞释放通量预测,为区域汞环境风险评估与管控提供决策支持.

参考文献:

[1] Stumm W, Morgan J J. Aquatic chemistry: Chemical equilibria and rates in natural waters [M]. 3rd ed. New York: Wiley, 1996.

[2] 阴永光,李雁宾,蔡　勇,等.汞的环境光化学 [J]. 环境化学, 2011, 30(1):84−91.
Yin Y G, Li Y B, Cai Y, Jiang G B. Environmental photochemistry of mercury [J]. Environmental Chemistry, 2011,30(1):84−91.

[3] Schroeder W H, Anholt B R. Global mercury observation system: latitudinal gradient of atmospheric mercury speciation [J]. Environmental Science & Technology, 2017,51(12):6813−6822.

[4] 何天容,冯新斌,郭艳娜,等.红枫湖沉积物中汞的环境地球化学循环 [J]. 环境科学, 2008,29(7):1768−1774.
He T R, Feng X B, Guo Y N, et al. Geochemical cycling of mercury in the sediment of Hongfeng Reservior [J]. Environmental Science, 2008, 29(7):1768−1774.

[5] 姚　珩,冯新斌,仇广乐,等.贵州洪家渡水库水体不同形态汞的分布特征 [J]. 生态学杂志, 2011,30(5):961−968.
Yao H, Feng X B, Qiu G L, et al. Distribution of different mercury forms of Hongjiadu Reservoir from Guizhou province [J]. Chinese Journal of Ecology, 2011,30(5):961−968.

[6] 方志青,陈秋禹,尹德良,等.三峡库区支流河口沉积物重金属分布特征及风险评价 [J]. 环境科学, 2018,39(6):2607−2614.
Fang Z Q, Chen Q Y, Yin D L, et al. Distribution characteristics and risk assessment of heavy metals in the sediments of the estuary of the tributaries in the Three Gorges Reservoir, SW China [J]. Environmental Science, 2018,39(6):2607−2614.

[7] 张　成,宋　丽,王定勇,等.干湿交替条件下三峡水库消落带土壤汞形态变化 [J]. 应用生态学报, 2013,24(12):3531−3536.
Zhang C, Song L, Wang D Y, et al. Mercury speciation transformation in soil of the water−level−fluctuating zone in the Three Gorges Area under alternative dry−wet condition [J]. Chinese Journal of Applied Ecology, 2013,24 (12):3531−3536.

[8] Zhou C, Tada A, Yano S, et al. Mercury dynamic simulation of Minamata Bay by using a three-dimensional numerical model [J]. Environmental Science and Pollution Research, 2019,26(36):36732−36747.

[9] USEPA. Method 1631: Mercury in water by oxidation, Purge and trap, and cold vapor atomic fluorescence spectrometry [S]. Washington DC: United States Environmental Protection Agency, 2002.

[10] TES Electrical Electronic Corp. TES-II digital illuminometer: product specifications [M]. Taiwan, China: TES Electrical Electronic Corp., 2017.

[11] DeltaTRAK Inc. FlashCheck pocket probe digital thermometer 11000: Specifications and calibration guide [M]. California, USA: DeltaTRAK, 2008.

[12] Nielsen-Kellerman. Kestrel 4000Pocket weather tracker: User manual [M]. Commonwealth of Pennsylvania, USA: Nielsen- Kellerman Co. 2012.

[13] Sartorius AG. PT−10pH meter: Technical documentation [M].

Göttingen, Germany: Sartorius Group, 2015.

[14] YSI Incorporated. YSI 550A dissolved oxygen instrument: User manual [M]. Washington, USA: YSI Environmental, 2005.

[15] GE Analytical Instruments. Sievers innovOx laboratory TOC analyzer: Operation manual [M]. Colorado, USA: General Electric Company, 2010.

[16] 闫海鱼,冯新斌,商立海,等.天然水体中痕量汞的形态分析方法研究 [J]. 分析测试学报, 2003,22(5):10−13.
Yan H Y, Feng X B, Shang L H, et al. Research on the morphological analysis method of trace mercury in natural water bodies [J]. Journal of Instrumental Analysis, 2003,(5):10−13.

[17] Lopez S A, Travnikov O, Sonke J E, et al. Photochemistry of oxidized Hg(I) and Hg(II) species suggests missing mercury oxidation in the troposphere [J]. Proceedings of the National Academy of Sciences, 2020,117(49):30949−30956.

[18] Vette A F, Landis M S, Keeler G J. Mercury photochemistry in the marine boundary layer: Modeling and experimental evidence of rapid destruction [J]. Environmental Science & Technology, 2002,36(1):5−9.

[19] 何天容,冯新斌,郭艳娜,等.红枫湖水体中活性汞和溶解气态汞的分布特征及其控制因素 [J]. 环境科学研究, 2008,21(2):14−17.
He T R, Feng X B, Guo Y N, et al. Distributions of reactive and dissolved gaseous mercury and controlling factors in Hongfeng Reservoir [J]. Research of Environmental Sciences, 2008,21(2):14−17.

[20] Janssen S E, Murray R K, Dewild J F, et al. Chemical and physical controls on mercury source signatures in stream fish from the Northeastern United States [J]. Environmental Science & Technology, 2019,53(17):10110−10119.

[21] Atangana B M S, Ngoupayou N J R, Deliege J F. Hydrogeochemistry and mercury contamination of surface water in the Lom Gold Basin (East Cameroon): Water quality index, multivariate statistical analysis and spatial interpolation [J]. Water. 2023,15(13):2502.

[22] Crocker W C, Zhang H. Seasonal and diurnal variation of air/water exchange of gaseous mercury in a southern reservoir lake (Cane Creek Lake, Tennessee, USA) [J]. Water, 2020,12(8):2102.

[23] Sullivan C J, Vokoun J C, Perkins C R. Spatiotemporal changes in largemouth bass mercury concentrations from Connecticut waterbodies, 1995~2021 [J]. Environmental Monitoring and Assessment, 2023,195(6):780.

[24] Huynh T M T, Ni C F, et al. Predicting heavy metal concentrations in shallow aquifer systems based on low−cost physiochemical parameters using machine learning techniques [J]. International Journal of Environmental Research and Public Health, 2022,19(19).12180.

[25] 陈亚松,刘家雯,赵云鹏,等.基于机器学习的人工湿地出水水质预测与影响因素 [J]. 中国环境科学, 2025,45(6):3161−3170.
Chen Y S, Liu J W, Zhao Y P, et al. Prediction of effluent water quality and analysis of influencing factors in constructed wetlands based on machine learning [J]. China Environmental Science, 2025,45(6):3161−3170

[26] 成 鹏,汪西莉.SVR 参数对非线性函数拟合的影响 [J]. 计算机工程, 2011,37(3):189−191.
Cheng P, Wang X L. Influence of SVR parameter on non−linear function approximation [J]. Computer Engineering, 2011,37(3):189−191.

[27] 钟 原,张 泰,李 平,等.随机森林的融合模型在压井方法分类中的研究 [J]. 西南石油大学学报(自然科学版), 2022,44(1):165−173.
Zhong Y, Zhang T, Li P, et al. Research on classification of well killing method based on Random Forest fusion model [J]. Journal of Southwest Petroleum University(Science & Technology Edition), 2022, 44(1):165−173.

[28] 刘丽贞,黄 琪,迟殿委.基于随机森林回归模型和高频数据的鄱阳湖子湖电导率预测 [J]. 水电能源科学, 2023,41(10):50−53.
Liu L Z, Huang Q, Chi D W, Prediction of electrical conductivity in sub−lake of Poyang Lake based on random forest regression model and high−frequency Data [J]. Water Resources and Power, 2023,41 (10):50−53.

[29] 张瑞杰,周春艳,陈 辉,等.基于 GF−5B 卫星的随机森林模型反演京津冀地区 $PM_{2.5}$ [J]. 中国环境科学, 2024,44(11):5961−5970.
Zhang R J, Zhou C Y, Chen H, et al. Inversion of $PM_{2.5}$ in the Beijing−Tianjin−Hebei region using a random forest model based on GF−5B satellite [J]. China Environmental Science, 2024,44(11):5961−5970.

[30] 吴志明,李建超,王 睿.基于随机森林的内陆湖泊水体有色可溶性有机物(CDOM)浓度遥感估算 [J]. 湖泊科学, 2018,30(4):979−991.
Wu Z M, Li J C, Wang R. Estimation of CDOM concentration in inland lake based on random forest using Sentinel3A OLCI [J]. Journal of Lake Sciences, 2018,30(4):979−991.

[31] 孔宪喻,苏荣国.基于支持向量机的黄东海富营养化快速评价技术 [J]. 中国环境科学, 2016,36(1):143−148.
Kong X Y, Su R G. A support vector machine−based technology for rapidly assessing trophic status of the Yellow Sea and the East China Sea [J]. China Environmental Science, 2016,36(1):143−148.

致谢：本实验的现场采样工作由重庆长寿湖区环保局协助完成,在此表示感谢.

作者简介：白薇扬(1980−),女,重庆人,副教授,博士,主要研究方向为环境重金属污染控制.发表论文 20 余篇.baiwy@cqut.edu.cn.