

王茹冰, 姜翠玲, 孙磊, 等. 基于可解释性深度学习的星云湖表层水温预测研究[J]. 水利水电技术(中英文), 2026, 57(9): 193-204. DOI: 10.13928/j.cnki.wrahe.2026.09.015

WANG Rubing, JIANG Cuiling, SUN Lei, et al. Prediction of surface water temperature in Xingyun Lake based on interpretable deep learning[J]. Water Resources and Hydropower Engineering, 2026, 57(9): 193-204. DOI: 10.13928/j.cnki.wrahe.2026.09.015

基于可解释性深度学习的星云湖 表层水温预测研究

王茹冰¹, 姜翠玲¹, 孙磊¹, 戴昊璋¹, 许睿亭¹, 刘相改¹, 刘洲志², 施椿燕³

(1. 河海大学 水文水资源学院, 江苏 南京 210098; 2. 河海大学 港口海岸与近海工程学院, 江苏 南京 210098; 3. 长江三峡技术经济发展有限公司, 北京 101100)

摘要:【目的】全球变暖背景下, 湖泊表层水温(Lake Surface Water Temperature, LSWT)的持续上升已成为加剧湖泊富营养化进程的关键因素, 因此准确预测其未来变化趋势对湖泊富营养化预警及水环境管理具有重要意义。【方法】基于MODIS遥感影像获取2002—2024年星云湖逐日表层水温数据, 采用随机森林(Random Forest, RF)算法插补其缺失值, 进而构建并行的LSTM-Transformer深度学习模型预测未来不同气候情景(SSPs)下LSWT的变化趋势, 并引入夏普利加性解释法(SHapley Additive exPlanations, SHAP)揭示LSWT的关键驱动因子及其影响方向。【结果】结果表明:(1)2002—2024年期间星云湖表层水温的升温速率为 $0.27\text{ }^{\circ}\text{C}/10\text{ a}$, 秋季升温趋势尤为突出。(2)LSTM-Transformer融合模型具有较高的预测精度, 决定系数(R^2)为0.95, 均方根误差(RMSE)为 $0.89\text{ }^{\circ}\text{C}$; 在未来气候情景(SSP1-2.6、SSP2-4.5和SSP5-8.5)下, 星云湖表层水温预计将以 $0.11\sim 0.42\text{ }^{\circ}\text{C}/10\text{ a}$ 的速率持续上升。(3)近地表气温和地表下行长波辐射是影响星云湖表层水温变化的关键驱动因子。【结论】本研究构建的深度学习模型在高原湖泊表层水温预测中具有良好的适用性, 该框架可为高原湖泊在未来气候情景下的水环境适应性管理与生态风险预警提供依据。

关键词: 湖泊表层水温; MODIS; 星云湖; 深度学习模型; LSTM-Transformer; 气候情景; SHAP; 驱动因子分析

DOI: 10.13928/j.cnki.wrahe.2026.09.015

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)09-0193-12



Prediction of surface water temperature in Xingyun Lake based on interpretable deep learning

WANG Rubing¹, JIANG Cuiling¹, SUN Lei¹, DAI Haozhang¹, XU Ruiting¹, LIU Xianggai¹,
LIU Zhouzhi², SHI Chunyan³

(1. College of Hydrology and Water Resources, Hohai University, Nanjing 210098, Jiangsu, China; 2. College of Harbour, Coastal and Offshore Engineering, Hohai University, Nanjing 210098, Jiangsu, China; 3. Yangtze Three Gorges Technology and Economy Development Co., Ltd., Beijing 101100, China)

收稿日期: 2026-03-23; 修回日期: 2026-05-13; 录用日期: 2026-05-28; 网络出版日期: 2026-07-20

基金项目: 国家重点研发计划(2023YFC3207503)

作者简介: 王茹冰(2001—), 女, 硕士研究生, 研究方向为生态水文与水环境保护。E-mail: ro_icy@163.com

通信作者: 姜翠玲(1966—), 女, 教授, 博士研究生导师, 博士, 研究方向为生态水利。E-mail: cljianghu@163.com

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

Abstract: [Objective] In the context of global warming, the continuous rise in lake surface water temperature (LSWT) has become a key factor in exacerbating lake eutrophication. Therefore, accurately predicting its future trends is of great significance for the early warning of lake eutrophication and water environment management. [Methods] Based on MODIS remote sensing images, daily LSWT data for Xingyun Lake from 2002 to 2024 were acquired. The random forest (RF) algorithm was employed to impute missing values. Subsequently, a parallel LSTM-Transformer deep learning model was developed to predict future LSWT trends under different climate scenarios (SSPs). Furthermore, the SHapley Additive exPlanations (SHAP) method was introduced to identify the key driving factors of LSWT and their directions of influence. [Results] The results indicated that: (1) the warming rate of surface water temperature in Xingyun Lake was $0.27\text{ }^{\circ}\text{C}/10\text{ a}$ from 2002 to 2024, with a particularly pronounced warming trend in autumn. (2) The LSTM-Transformer fusion model exhibited high prediction accuracy, with a coefficient of determination (R^2) of 0.95 and a root mean square error (RMSE) of $0.89\text{ }^{\circ}\text{C}$. Under future climate scenarios (SSP1-2.6, SSP2-4.5, and SSP5-8.5), the LSWT of Xingyun Lake was projected to continue rising at a rate of $0.11\sim 0.42\text{ }^{\circ}\text{C}/10\text{ a}$. (3) Near-surface air temperature and surface downward longwave radiation were identified as the primary driving factors influencing LSWT variations in Xingyun Lake. [Conclusion] The developed deep learning model exhibits good applicability for predicting surface water temperature in plateau lakes. The framework can provide a basis for adaptive water environment management and early warning of ecological risks in plateau lakes under future climate scenarios.

Keywords: lake surface water temperature; MODIS; Xingyun Lake; deep learning model; LSTM-Transformer; climate scenarios; SHAP; driving factor analysis

0 引言

湖泊表层水温 (Lake Surface Water Temperature, LSWT) 是反映湖泊热力状况和区域气候长期演变的重要标志^[1-2], 也是影响湖泊水生态环境的关键因子。已有研究表明, 全球气候变暖导致世界上大多数湖泊的表层水温显著上升^[3], 可能引起蓝藻生长期延长、湖泊富营养化程度加剧等环境问题, 导致水质恶化, 进而影响水生生物生存^[4-5]。因此, 准确预测湖泊表层水温变化, 能够为气候变暖背景下的水质恶化与蓝藻水华防控提供科学支撑。

湖泊表层水温的获取主要包括站点原位观测和红外遥感反演两种途径。传统站点观测虽精度较高, 但多数湖泊缺少长期、连续的历史观测序列。相比之下, 卫星遥感技术具有高时空分辨率和动态连续监测的优势, 已成为研究 LSWT 长期变化规律的重要手段^[6]。其中, 中分辨率成像光谱仪 (MODIS) 的地表温度产品 MOD11A1 能通过质量控制识别、剔除无效像元, 是提取高精度 LSWT 的可靠遥感数据源。然而, 受云覆盖、气溶胶等因素影响, 该数据序列中常存在大量缺失, 因此, 融合遥感观测与机器学习方法已成为构建连续水温序列的有效解决方案^[7-8]。

湖泊表层水温的变化受近地表气温、太阳辐射等多种因素驱动, 各因素间相互作用复杂, 且与 LSWT 呈非线性关系, 使得准确预测 LSWT 具有较高挑战性^[9]。目前, 常见的 LSWT 的预测模型主要包括过程

驱动模型^[10]、基于物理统计的混合模型^[11]以及统计模型^[12]。其中, 过程驱动模型依赖大量的湖泊特征信息, 在资料缺乏的湖泊中应用受限。为克服该问题, 以 Air2water 为代表的物理统计混合模型^[13-15]被广泛应用, 但其仅以气温作为输入, 难以刻画实际水温受多因素共同影响的复杂过程。统计模型往往通过建立 LSWT 与气象变量间的回归关系进行预测, 包括传统多元回归和机器学习模型等^[16-17]。近年来, 以长短期记忆网络 (Long Short-Term Memory, LSTM) 为代表的深度学习模型凭借其对序列依赖关系的有效捕捉, 在 LSWT 预测中得到了广泛应用^[18-20]。与此同时, 基于自注意力机制的深度学习模型 (如 Transformer) 在时间序列任务中也表现出显著优势, 尤其在特征全局交互方面表现突出^[21]。然而, 深度学习模型的“黑箱”性质限制了对模型内部驱动因子响应机制的深入解析。夏普利加性解释法 (SHapley Additive exPlanations, SHAP) 作为一种模型后解释方法, 能够有效量化各输入特征对预测结果的贡献, 并揭示其作用方向与强度, 为提升模型可解释性提供了有力工具。AMNUAYLOJAROEN 等^[22]借助 SHAP 法量化了气象因子对 LSWT 的贡献, 证实了气温的主导作用。

伴随着全球气候模式 (GCMs) 的不断发展, 第六次国际耦合模式比较计划 (CMIP6) 提供了一系列基于共享社会经济路径 (SSPs) 与典型浓度路径 (RCPs) 的未来气候情景数据集, 为湖泊表层水温的预测提供了

重要的驱动基础^[2,15]。但大多数气候模式的空间分辨率较低,难以满足湖泊等局地尺度的气候响应研究需求^[23]。鉴于此,美国国家航空航天局(NASA)通过降尺度和偏差校正/空间离散化方法生成了一套高分辨率的日尺度数据集(NEX-GDDP-CMIP6),显著提高了区域及更小尺度上的气候变化相关研究的可靠性^[24]。

星云湖作为典型的高原湖泊,在区域水资源保障、气候调节与生态平衡维持等方面具有重要作用。然而,该湖泊长期面临着严峻的富营养化问题,蓝藻水华频发,严重威胁其生态系统健康^[25]。已有研究表明,湖泊表层水温是调控藻类生长过程的关键环境因子之一,水温升高通过影响藻类生长速率与群落结构,会进一步加剧水华暴发的风险^[26]。因此,准确预测湖泊表层水温,并评估不同气候情景下的水温演变趋势,是开展星云湖富营养化风险预警、制定精准防控策略的重要基础。然而,目前针对该湖泊在不同气候情景下的表层水温预测及驱动机制研究仍较为有限。鉴于此,本研究基于星云湖 2002—2024 年 MODIS 遥感水温数据与同期气象资料,构建了 LSTM-Transformer 融合模型,系统评估了其相较于单一 LSTM 与 Transformer 模型的预测性能,进而利用该融合模型开展不同气候情景下的 LSWT 预测,并结合 SHAP 方法解析关键驱动因子,以期为气候变暖背景下高原湖泊的生态风险评估与适应性管理提供科学依据与技术支持。

1 研究区域与数据来源

1.1 研究区概况

星云湖(102°45'E—102°48'E, 24°17'N—24°23'N)位于云南省玉溪市江川区境内,属于滇中高原陷落型浅水湖泊(见图1)。湖泊南北长约 9.09 km,东西平均宽 3.80 km,最大水深 10.81 m,平均水深 6.01 m,整体呈茄形,湖面面积为 34.33 km²。星云湖水体主要靠地表径流和降水补给,共有东西大河、螺蛳铺河等 12 条主要入湖河流,河道总长 132.30 km。星云湖属于亚热带半湿润高原季风气候,具有四季不分明、干湿季明显的气候特征,年均气温 15.90℃,降水主要集中在 5—10 月。

1.2 数据来源

1.2.1 星云湖水温数据

2002—2024 年期间的星云湖表层水温数据基于 Google Earth Engine(GEE)平台提供的 MODIS MOD11A1 V6.1 温度产品获取。该产品空间分辨率为 1 km×1 km,时间分辨率为 1 d,主要科学数据集包括日间(LST_Day_1 km)与夜间(LST_Night_1 km)地表温度及其对应的质量控制波段(QC_Day、QC_Night)。首先提取星云湖区域的日间和夜间 LST 数据,并根据质量控制波段剔除受云覆盖或气溶胶等因素影响的低质量像元;将保留的像元值(DN)转换为摄氏温度($LSWT=DN \times 0.02 - 273.15$),对各像元点

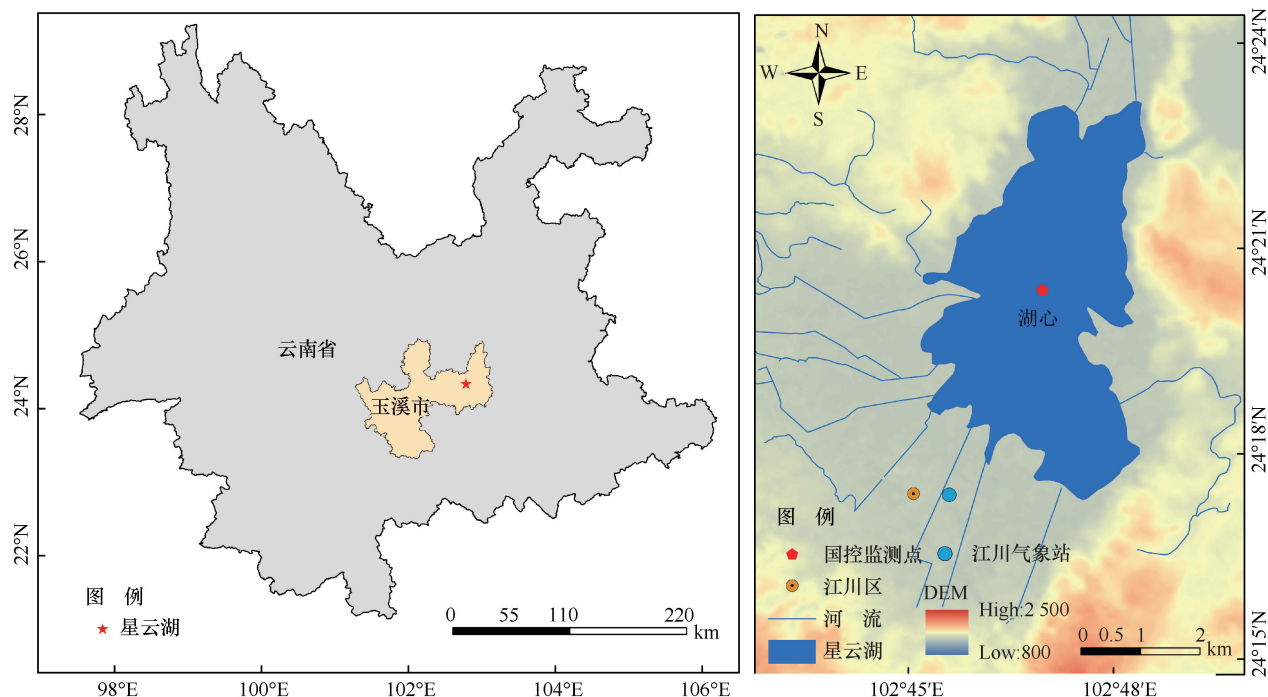


图1 星云湖区域概况

Fig. 1 Overview map of Xingyun Lake

分月份采用四分位距法(IQR)识别并剔除异常水温值;分别对日间和夜间各像元的表层水温进行空间平均,再计算出日均值(若日间或夜间任一时段数据存在缺失,则该日表层水温标记为缺失)。由此得到2002—2024年的星云湖表层水温序列,缺失比例约为63%。

星云湖实测水温数据源自国家地表水水质自动监测实时数据发布系统(<https://www.cnemc.cn>),考虑到数据的连续性,选取2021—2024年星云湖心监测站点逐4h记录的水温数据计算出日均水温,用以验证插补后MODIS水温序列的精度。需要指出的是,本文采用的MODIS湖泊表层水温为全湖有效水体像元的空间平均值,而实测水温为湖心站点单点观测值。受湖泊表层水温空间异质性影响,全湖平均水温与湖心单点水温之间存在一定空间尺度差异。

1.2.2 气象数据

本研究所用的逐日气象数据包括气温(TEM,℃)、相对湿度(RHU,%)、降水量(PRE,mm)、风速(WS,m/s)、地表下行短波辐射(DSR,W/m²)及地表下行长波辐射(DLR,W/m²),时间范围为2002—2024年。其中,TEM、RHU、PRE及WS数据来源于国家气象科学数据中心(<http://data.cma.cn>),站点为云南玉溪江川站;DSR及DLR数据来源于欧洲中期天气预报中心(ECMWF,<https://cds.climate.copernicus.eu>)提供的ERA5再分析数据集,其空间分辨率为0.25°×0.25°,时间分辨率为1h,通过Climate Data Store API编程批量下载,辐射数据取当日累积值。

1.2.3 未来气候情景数据

星云湖区域未来气候情景数据来源于NASA地球交换全球每日降尺度预测数据集(NEX-GDDP-CMIP6),该数据集基于CMIP6多模型输出,采用每日版本的月度偏差校正和空间离散化(BCSD)方法,生成空间分辨率为0.25°×0.25°的逐日格网数据。考虑到ACCESS-CM2模式在东亚地区具有良好的气候模拟能力,适用于区域尺度的气候影响评估,本研究选取该模式在三种未来共享社会经济路径情景(SSP1-2.6、SSP2-4.5、SSP5-8.5)下的数据资料。变量包括近地表气温(TEM)、相对湿度(RHU)、地表下行短波辐射(DSR)、地表下行长波辐射(DLR)、降水量(PRE)及风速(WS),时间跨度为2025—2100年。

2 研究方法

2.1 LSWT 缺失值插补方法

由于云覆盖、气溶胶等原因导致MODIS遥感提

取的湖泊逐日表层水温序列不连续,本研究采用随机森林(Random Forest, RF)算法对缺失值进行插补。RF是一种基于决策树的集成学习算法,通过自助采样(Bootstrap)和随机特征选择构建多棵决策树,并集成其输出结果以提升模型预测精度和泛化能力^[27]。

为充分表征水温序列中的时序规律,本研究构建了多维时间特征作为模型输入,包括年、月、日序数、季节分类等。在此基础上,对经线性插值预处理的日尺度表层水温数据进行STL(Seasonal-Trend decomposition using Loess)分解,以进一步捕捉序列的长期趋势、周期性波动及随机扰动。该方法基于鲁棒局部加权回归,能够将序列稳健地分解为趋势项、季节项和残差项,广泛适用于时序数据的季节性分析^[28-29]。最终,将上述多维时间特征与STL分解特征共同作为RF的输入变量。数据集按时间顺序划分为训练集(2002—2020年)和测试集(2021—2024年),采用超参数调优和时间序列交叉验证方法,得到模型的最优参数组合:树的数量为413,最大深度为30,最小分裂样本数为13。基于该配置重新训练模型,估算出缺失日期的水温数据,获得完整的湖泊逐日表层水温序列,为后续湖泊表层水温预测提供可靠的数据基础。

2.2 湖泊表层水温预测模型

本研究基于气象数据与MODIS水温数据,构建深度学习模型开展湖泊表层水温预测。为优化模型输入,选择RF算法进行关键特征筛选,该算法依据信息增益等准则进行节点分裂,并以各特征在所有决策树中的平均分裂贡献度作为重要性评估指标^[30]。在特征筛选前,首先通过方差膨胀因子(VIF)对各特征的多重共线性进行诊断,结果表明各特征的VIF均小于10。后按时间顺序将数据集划分为训练集(2002—2020年)和测试集(2021—2024年),利用网格搜索和时间序列交叉验证确定最佳超参数(树的数量为400,最大深度为10,最小分裂样本数为2),并基于最优模型在训练集上计算特征重要性。最终筛选出重要性得分大于5%的特征,包括近地表气温(TEM, 45.05%)、地表下行长波辐射(DLR, 30.54%)、降水量(PRE, 8.49%)、地表下行短波辐射(DSR, 7.23%)和相对湿度(RHU, 5.07%),作为后续预测模型的输入变量。

在模型构建方面,综合考虑水温数据的时序性与非线性特征,选取LSTM和Transformer作为基础预测模型。LSTM模型通过门控机制(细胞状态、遗忘门、输入门、输出门)有效降低了梯度消失风险^[31-32],是

处理时序数据的常用模型。Transformer 模型完全基于自注意力机制 (Self-Attention Mechanism) 处理序列数据, 其利用位置编码为输入序列注入顺序信息, 通过多头自注意力捕捉长程依赖, 具备强大的并行计算能力^[33]。

鉴于单一类型模型难以同时有效捕捉时序依赖与长程交互特征, 本研究构建了一种并行架构的 LSTM-Transformer 融合模型。该模型采用双分支并行处理机制: LSTM 分支提取序列中的时序依赖特征, Transformer 分支捕捉特征变量间的长程依赖与交互关系。模型将两分支的输出在特征维度上进行拼接后, 经全连接融合层进行非线性整合, 最终通过输出投影

层映射为水温预测值, 模型结构如图 2 所示。该并行架构实现了两种深度学习模型在时序局部性与全局关联性建模上的优势互补, 有助于提升模型在复杂气象条件下的鲁棒性与预测精度。

将筛选出的气象变量数据及 MODIS 水温数据按时间顺序划分为训练集 (2002—2020 年) 和测试集 (2021—2024 年)。各模型均采用单步预测, 时间步长设定为 30 d, 基于 PyTorch2. 8. 0 框架构建。超参数优化采用网格搜索结合五折时间序列交叉验证, 优化器为 Adam, 损失函数为 MSELoss。各模型的超参数搜索范围及调参结果如表 1 所列, 为平衡搜索充分性与计算效率, 三个模型分别测试 12、24、24 组参

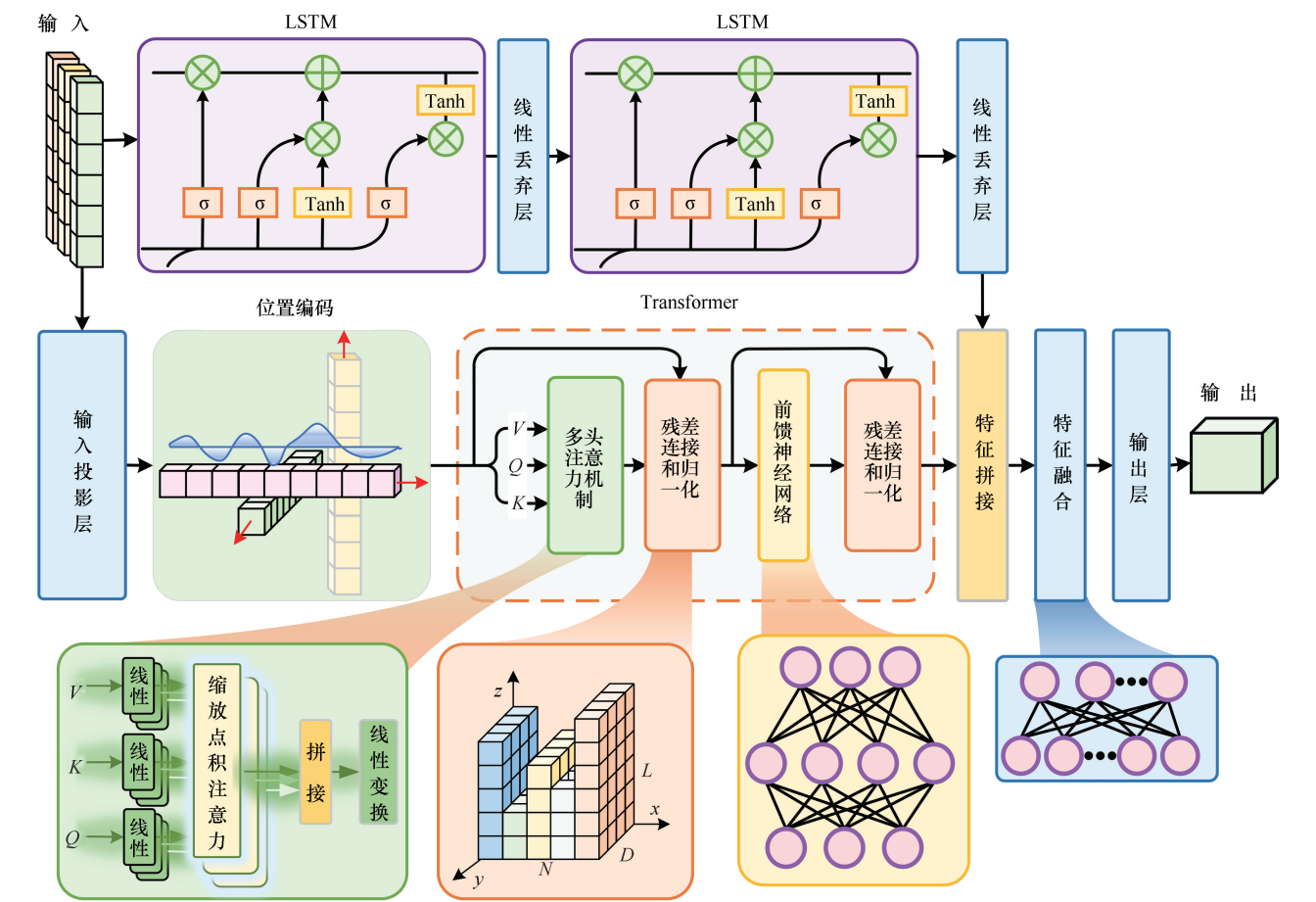


图 2 LSTM-Transformer 融合模型结构

Fig. 2 Architecture of LSTM-Transformer parallel fusion model

表 1 模型参数搜索范围及最佳参数

Table 1 Parameter search ranges and optimal parameters of models

参 数	参数搜索范围	最佳参数		
		LSTM	Transformer	LSTM-Transformer
学习率 (learning_rate)	[0.000 1, 0.000 5, 0.001]	0.001	0.000 1	0.000 1
隐藏层 (num_layers)	[2, 4]	2	4	LSTM (2) + Transformer (4)
注意力头数 (nhead)	[4, 8]	—	8	8
丢弃率 (dropout)	[0.1, 0.2]	0.2	0.1	0.1

数组组合。对于每一组超参数, 计算每折上验证集的均方误差(MSE), 以 5 折 MSE 的均值最小作为最优超参数的选取准则。

2.3 模型可解释性分析方法

SHAP 法由 LUNDBERG 等于 2017 年提出^[34], 是一种基于博弈论的模型解释方法, 其拥有独特的理论基础和强大的解释能力, 能为每个特征在模型预测中做出的贡献提供一个公平且合理的量化评估。特征对模型预测结果的贡献程度越大, 对应的 Shapley 值越大。其中, Shapley 值的计算公式为

$$\phi_i = \sum_{S \subseteq N} \frac{|S|! (|N| - |S| - 1)!}{|N|!} [f(S \cup \{i\}) - f(S)] \tag{1}$$

式中, ϕ_i 为第 i 个特征的 Shapley 值; N 为所有特征的集合; S 为 N 的一个不包含特征 i 的子集; $|N|$ 为所有特征的数量; $|S|$ 为集合 S 中特征的数量; $f(S)$ 为模型在特征集合 S 上的预测值; $f(S \cup \{i\})$ 为在集合 S 中加入特征 i 后模型的预测值。

本研究使用 Python 库 shap 0.47.2 进行模型 SHAP 评估。

2.4 模型性能评估指标

本研究采用决定系数 (R^2)、均方根误差 (RMSE)、平均绝对误差 (MAE)、平均绝对百分比误差 (MAPE) 来全面评估模型的模拟效果。 R^2 越接近 1, RMSE、MAE 和 MAPE 越小, 说明模型性能越好。各指标计算公式为

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

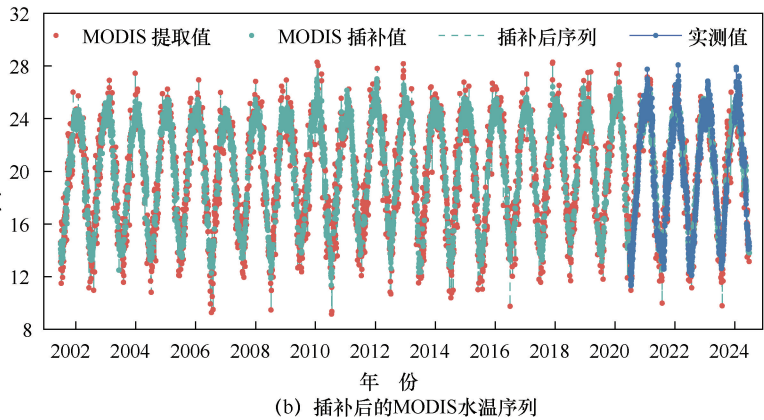


图 3 MODIS 水温精度验证散点图及完整序列

Fig. 3 Scatter plots of MODIS water temperature accuracy verification and complete sequence

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{3}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{4}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{\bar{y}} \right| \times 100\% \tag{5}$$

式中, n 为样本总数; y_i 、 $\hat{y}_i$ 、 $\bar{y}$ 分别为实测值、模拟值、实测值的均值。

3 结果与分析

3.1 LSWT 插补及序列变化特征分析

采用 RF 算法对 2002—2024 年 MODIS 湖泊表层水温序列中的缺失值进行插补, 为验证遥感水温数据的可靠性, 将 2021—2024 年期间的 MODIS 湖泊表层水温序列与星云湖心站点实测水温数据进行对比[见图 3(a)]。可见遥感水温与实测值吻合较好, 二者之间的 R^2 为 0.90, RMSE 为 1.37 °C, MAE 为 1.09 °C, MAPE 为 5.70%, 且数据点均匀分布在 1:1 线两侧。插补后的序列清晰地呈现出 LSWT 的季节性特征和年际周期性变化[见图 3(b)], 说明基于该插补方法的 MODIS 湖泊表层水温序列可靠。

星云湖表层水温每年从 9~12 °C 开始上升, 在 6—8 月达到最大值 (25~29 °C), 随后逐渐降低。2002—2024 年期间, 星云湖多年平均表层水温为 20.11 °C。其中, 2012 年年均水温最高, 达 20.58 °C; 2004 年年均水温最低, 为 19.62 °C。线性趋势分析结果显示, 研究期间星云湖年均表层水温呈显著上升趋势, 升温速率为 0.27 °C/10 a ($p < 0.05$), 受年际波动影响, 线性拟合优度较低 ($R^2 = 0.40$)。在季节尺度上, 表层水温的变化趋势存在显著差异。夏季和秋季升温趋势较为明显, 速率分别为 0.27 °C/10 a 和

0.46 °C/10 a, 多年季均水温分别为 24.39 °C 和 20.76 °C; 春季和冬季的升温速率相对较低, 分别为 0.25 °C/10 a 和 0.08 °C/10 a, 多年季均水温分别为 20.47 °C 和 14.77 °C。这表明星云湖表层水温的升高主要由秋季贡献, 湖泊热状况呈现出明显的季节性非对称特征。

3.2 LSWT 预测模型精度评估与比较

本节对比评估了单一模型 (LSTM、Transformer)

与融合模型 (LSTM-Transformer) 在星云湖表层水温预测任务中的性能表现。如图 4 所示, 各模型点群基本沿 1:1 线分布, 均能较好地捕捉水温的整体变化趋势。从各评价指标来看, LSTM 模型[图 4(a)]和 Transformer 模型[图 4(b)]的预测精度较为接近, 而 LSTM-Transformer 融合模型[图 4(c)]取得了最优的预测结果, R^2 由 0.93 提升至 0.95, $RMSE$ 由 1.00 °C 降至 0.89 °C, 降幅达 11.00%; 同时, MAE 及 $MAPE$

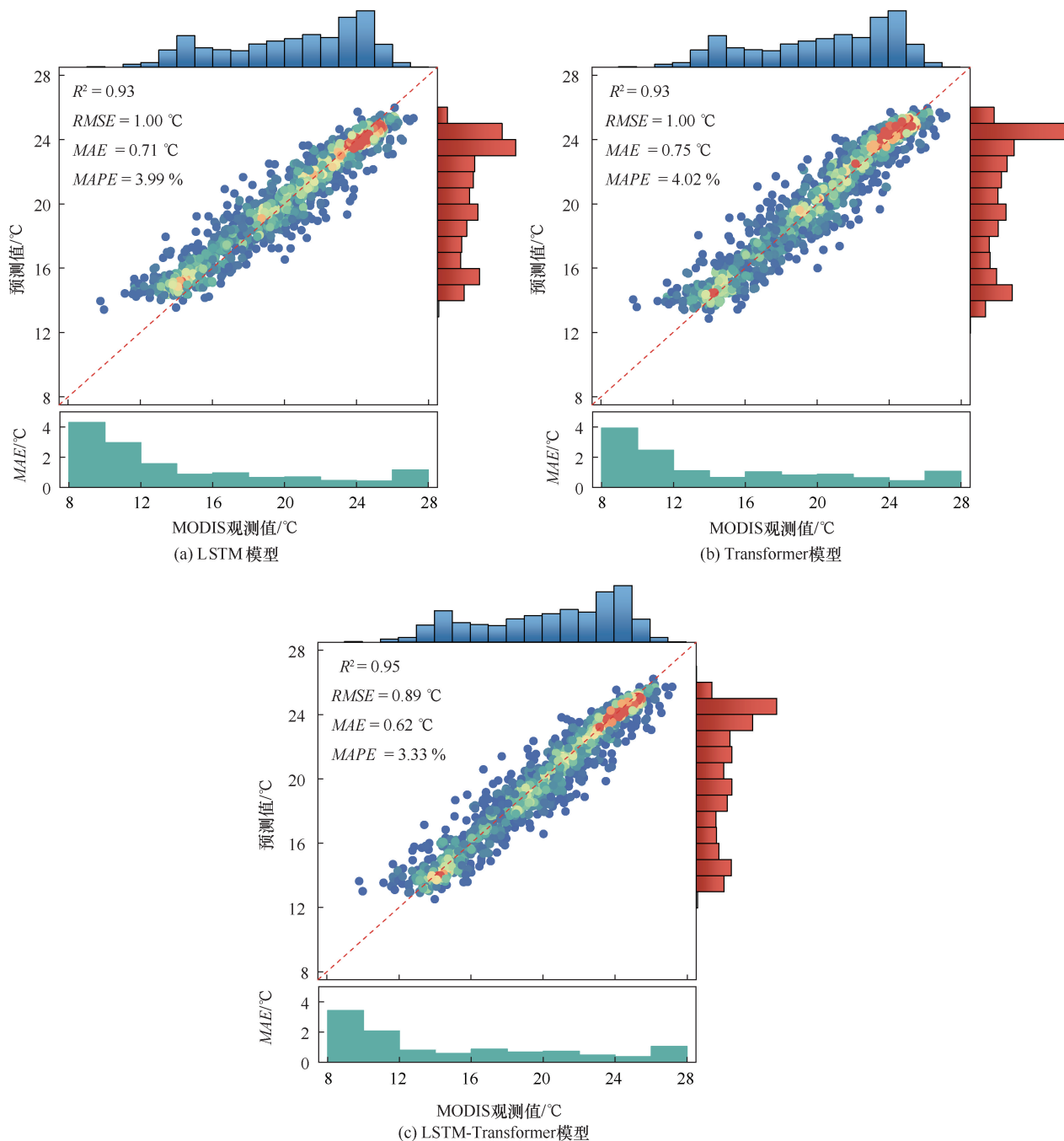


图 4 模型测试集预测结果散点图及 MAE 分布

Fig. 4 Scatter plots of model prediction results and MAE distribution on test set

也均呈现不同程度的下降。

进一步分析各模型 *MAE* 的区间分布特征, 结果表明, 各模型在模拟 8~12 ℃ 区间内的表层水温时均存在一定程度的系统性偏高误差, 可能与该区间内训练样本较少、模型对低值水温的拟合能力有限有关。其中, 单一深度学习模型在该低温区间的 *MAE* 相对较高, 8~10 ℃ 区间甚至超过 4 ℃, 反映出其对极端水温事件的模拟能力仍存在不足。而 LSTM-Transformer 融合模型在低温区间的误差控制能力明显提升, 其在 8~12 ℃ 区间内的 *MAE* 相比 LSTM 和 Transformer 模型分别降低了 24.39% 和 13.85%。这表明, 尽管当前单一深度学习模型对于极端水温事件的拟合能力仍存在一定局限, 但本研究构建的并行融合模型在一定程度上缓解了单一模型在极端区间表现出的系统性偏差, 展现出更强的稳健性与泛化能力, 为湖泊水温的精准预测提供了更为可靠的建模思路。

3.3 不同气候情景下的 LSWT 变化趋势

基于最优模型 (LSTM-Transformer) 对星云湖表层水温在不同气候情景下的变化进行预测, 结果如图 5 所示, 图 5 中阴影区域表示年均值的 90% 置信区间。预测结果表明, 在不同共享社会经济路径下, 湖泊表层水温对排放情景的响应存在显著差异, 未来 *LSWT* 的增温速率将随温室气体排放水平的升高而持续加剧。在低温室气体排放情景 (可持续发展路径, SSP1-2.6) 下, 2025—2100 年星云湖表层水温将以 0.11 ℃/10 a 的速率缓慢上升; 若延续当前中等排放路径 (SSP2-4.5), 同期增温速率将达到 0.24 ℃/10 a;

而在高温温室气体排放情景 (以化石燃料为主的高速发展路径, SSP5-8.5) 下, 增温速率将进一步加剧至 0.42 ℃/10 a, 约为 SSP1-2.6、SSP2-4.5 情景下的 3.82、1.75 倍。至 21 世纪末 (2100 年), SSP1-2.6、SSP2-4.5、SSP5-8.5 情景下星云湖的年均表层水温预计将达到 21.36 ℃、22.36 ℃ 和 23.83 ℃, 较 2024 年分别升高 1.28 ℃、2.28 ℃、3.75 ℃。可见, 不同气候情景对未来湖泊热状况的演变轨迹具有决定性影响, 尤其在 high 排放情景下, 湖泊增温幅度显著加大。

在未来气候变暖背景下, 星云湖表层水温的持续上升将对湖泊水生态系统产生深远影响。较高的水温会为蓝藻等适温藻类生长创造竞争优势, 改变原有藻类群落结构, 进而加剧蓝藻水华暴发的频率与强度^[26]。星云湖作为典型的高原浅水湖泊, 水体交换周期较长、营养盐本底浓度较高^[35], 对水温变化尤为敏感。在此背景下, 即使外源营养盐输入得到有效控制, 由气候变暖驱动的表层水温持续升高仍可能抵消外源削减的治理成效。因此, 未来星云湖的生态修复与富营养化防控策略需关注水温上升对湖泊生态过程的影响, 以提升治理措施的适应性与针对性。

3.4 气象因子对 LSWT 的贡献分析

基于 SHAP 法对 LSTM-Transformer 融合模型的预测结果进行气象因子贡献分析, 通过计算各因子 *Shapley* 值的绝对平均值得到全局重要性排序, 并进一步计算特征权重以量化各因子对湖泊表层水温预测的贡献比例 [见图 6(a)]。此外, 借助 SHAP 摘要图刻画各因子对 *LSWT* 预测值的相对贡献和影响方向

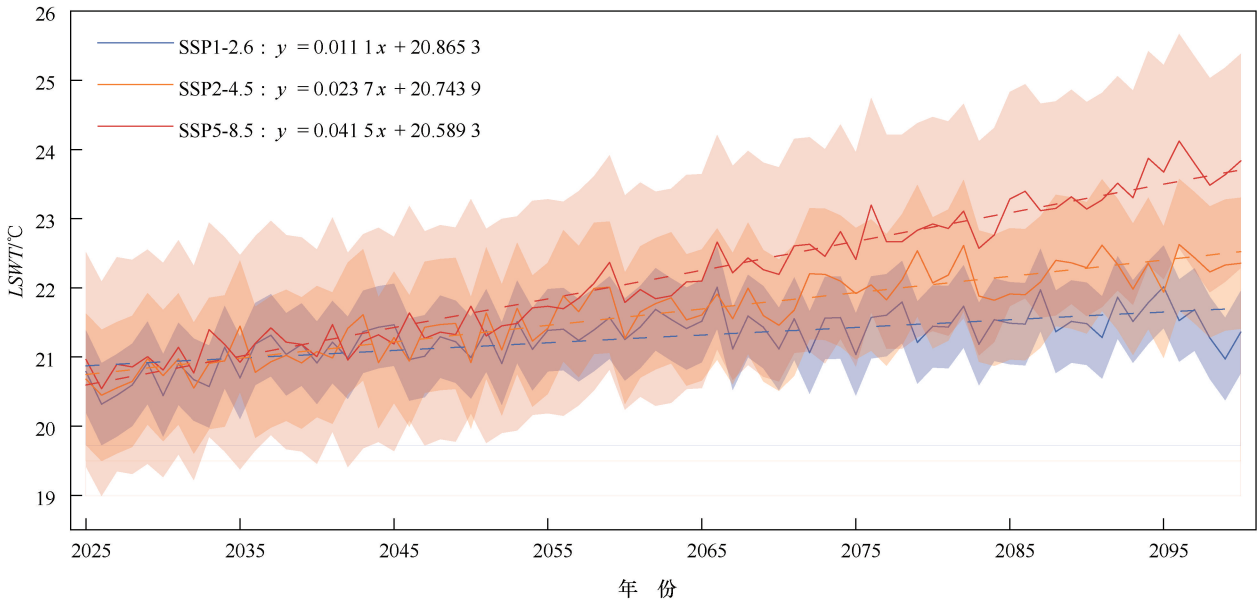


图 5 不同气候情景下星云湖年均表层水温的变化趋势

Fig. 5 Trends of annual average lake surface water temperature in Xingyun Lake under different climate scenarios

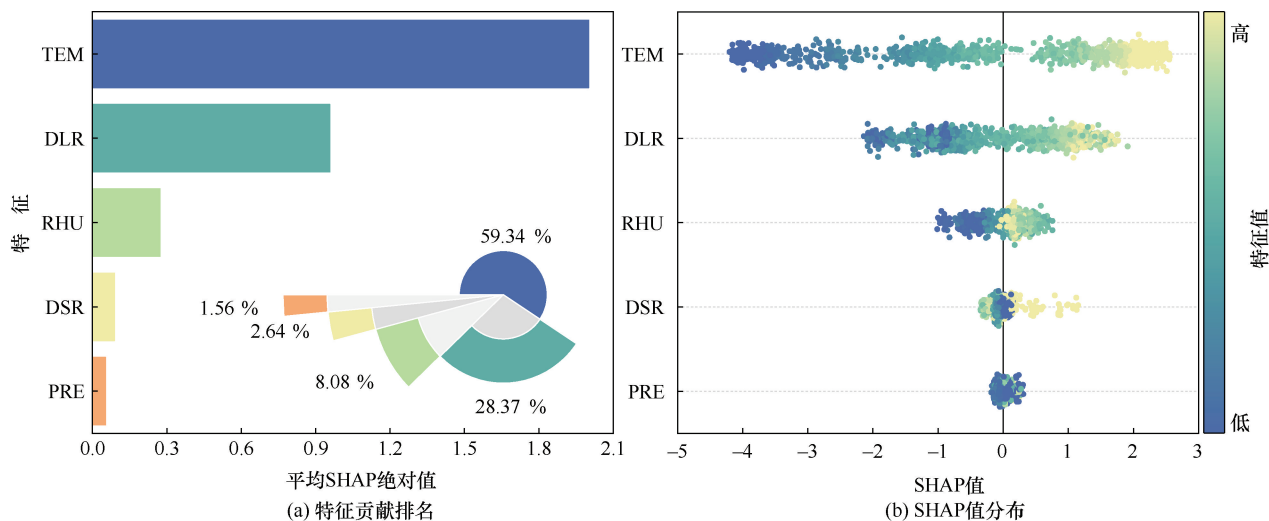


图6 特征贡献排名及 SHAP 值分布

Fig. 6 Feature contribution ranking and SHAP value distribution

[见图 6(b)], 散点的颜色越浅代表该样本的特征值越高。当较高的特征值对应的散点多分布在 SHAP 正值区域, 表明相应的样本在 *LSWT* 预测中起正向作用; 相反, 当较高的特征值对应的散点多分布在 SHAP 负值区域, 则表明样本在 *LSWT* 预测中起负向作用。

结果表明, TEM 和 DLR 是影响 *LSWT* 最主要的因素, 贡献率分别高达 59.34% 和 28.37%, 且均对 *LSWT* 具有正向影响。气温升高与地表下行长波辐射增加会共同加剧大气与表层水体间的热量交换, 促使热量持续由大气向水体输送^[36]。RHU 对 *LSWT* 具有正向作用(贡献率为 8.08%), 较高的湿度会抑制湖面蒸发进而减少潜热散失^[37]。SHAP 摘要图表明中等偏高的湿度水平对 *LSWT* 的提升作用最为显著, 而极高湿度条件下的正向影响则减弱, 这意味着 RHU 对 *LSWT* 的影响是非线性的且具有阈值效应。此外, 较高的 DSR 对 *LSWT* 起到增温作用(贡献率为 2.64%), 其能量可直接穿透水面并被水体吸收^[36]。但较低的 DSR 并不会显著降低 *LSWT*, 原因是其伴随的云层覆盖会增加长波辐射, 从而抑制低 DSR 对湖面的冷却效应。PRE 对 *LSWT* 的贡献最低(1.56%), 大多 PRE 为 0 的样本 SHAP 值分散在零值两侧主要是受其他气象条件调制。云贵高原地区的高海拔、强辐射特征, 使湖泊表层水温对气候变暖的响应更加敏感^[18], 因此气温与辐射共同构成了星云湖 *LSWT* 变化的核心热力驱动。

4 结果讨论

4.1 研究方法对比

为客观评价本研究提出的 LSTM-Transformer 融合

模型的性能, 将其与湖泊表层水温预测中的两类代表性方法——物理统计模型(Air2water)和过程驱动模型(FLake)进行系统比较。

在预测精度方面, 潘惠敏等^[11]基于 Air2water 模型重构了洞庭湖城陵矶和南咀站的日均水温序列, 验证期的 R^2 分别为 0.98 和 0.95, $RMSE$ 分别为 1.13 °C 和 1.46 °C; 孙静静等^[10]采用 FLake 模型模拟太湖的表层水温, R 为 0.97(对应 $R^2=0.94$)。本文 LSTM-Transformer 融合模型的精度(测试集 $R^2=0.95$, $RMSE=0.89$ °C)整体优于单一的深度学习模型及 Air2water、FLake 模型。在驱动因子分析方面, Air2water 模型仅以气温作为输入, 无法解释太阳辐射等其他因子对水温的影响; FLake 模型能够通过情景模拟方法定量分离各气候因素的贡献, 但需要预设湖泊平均水深、底部沉积物热活性层深度及消光系数等参数, 在数据资料匮乏的湖泊中应用受限。本文模型虽为数据驱动架构, 但引入 SHAP 方法量化了各气象要素对表层水温的复杂非线性贡献, 有效弥补了数据驱动模型可解释性不足的缺陷。且无需预设湖泊物理参数, 更适用于数据资料有限的湖泊。

4.2 研究结果比较

本文经 RF 算法填补 MODIS 水温缺失值, 得到 2002—2024 年期间星云湖表层水温的升温速率为 0.27 °C/10 a。YANG 等^[38]讨论了 2001—2018 年云贵高原不同类型湖泊的表层水温变化, 其自然湖泊(包括星云湖)的综合增温速率与本研究结果较为接近。在驱动因子识别方面, YANG 等^[38]基于多元回归贡献率分析法指出, 近地表气温和太阳辐射是影响湖泊表层水温的最主要因子, 本文基于 SHAP 方法的

分析得出一致结论。

对比未来气候情景下的预测结果, DUAN 等^[15]采用 Air2water 模型和 CMIP6 气候情景, 预测了滇池 2015—2100 年的表层水温变化趋势。由于预测模型、湖泊物理属性及外部强迫的差异, 本文星云湖在 SSP 情景下的升温速率普遍高于滇池的预测结果(对应 RCP1-2.6、RCP2-4.5、RCP5-8.5 情景的升温速率分别为 0.05℃/10 a、0.10℃/10 a、0.28℃/10 a)。但两项研究揭示的趋势具有一致性: 高温温室气体排放情景将导致湖泊表层水温以数倍于历史时期的速率持续上升, 这将加剧湖泊富营养化及蓝藻水华暴发的风险。

4.3 研究不足与展望

尽管本研究在星云湖表层水温预测及驱动因子分析方面取得了一定进展, 但仍存在一些不足, 有待后续研究进一步改进。其一, 本文的深度学习模型为纯数据驱动模型, 缺乏热力学约束, 对训练样本稀疏的极端水温事件的预测能力有限, 容易产生系统性预测偏差。其二, 本文采用的 ERA5 辐射数据及 NEX-GDPP-CMIP6 未来气候数据的空间分辨率显著大于星云湖的区域尺度, 较粗的分辨率网格难以精确表征湖区局地的气象条件, 可能引入一定的系统性误差。其三, 本研究仅考虑了气象因子对湖泊表层水温的影响, 未纳入人类活动等潜在驱动因子, 可能遗漏重要驱动机制, 影响模型的解释能力和预测精度。

针对上述不足, 未来研究将从以下方向重点突破: 一是探索结合物理机制和数据驱动的混合模型, 增强模型在极端条件下的泛化能力; 二是利用降尺度方法或更高分辨率的区域气候数据集, 提高气候强迫数据的空间代表性; 三是开展多源数据融合, 量化人类活动指标(土地利用变化、人口密度、取排水工程等)并将其纳入预测框架, 以更全面地揭示自然与人为因素共同驱动下的湖泊表层水温演变机制。

5 结 论

(1)RF 算法能有效填补 MODIS 湖泊表层水温产品中的缺失值, 与星云湖心实测值的验证精度达到 $R^2 = 0.90$, $RMSE = 1.37\text{ }^{\circ}\text{C}$; 基于完整时序数据, 2002—2024 年期间星云湖表层水温呈显著升高趋势, 升温速率为 $0.27\text{ }^{\circ}\text{C}/10\text{ a}$, 秋季升温趋势尤为突出 ($0.46\text{ }^{\circ}\text{C}/10\text{ a}$)。

(2)LSTM-Transformer 融合模型在星云湖表层水温预测中表现出最优性能 ($R^2 = 0.95$, $RMSE = 0.89\text{ }^{\circ}\text{C}$, $MAE = 0.62\text{ }^{\circ}\text{C}$, $MAPE = 3.33\%$), 有效缓

解了单一模型在低温区间的系统性偏差。未来情景预测显示, 2025—2100 年期间星云湖表层水温在 SSP1-2.6、SSP2-4.5 和 SSP5-8.5 情景下的升温速率分别为 $0.11\text{ }^{\circ}\text{C}/10\text{ a}$ 、 $0.24\text{ }^{\circ}\text{C}/10\text{ a}$ 和 $0.42\text{ }^{\circ}\text{C}/10\text{ a}$ 。表层水温的持续升高将加剧湖泊富营养化风险, 对星云湖的治理与长效管理构成新的挑战。

(3)TEM (贡献率 59.34%) 和 DLR (贡献率 28.37%) 是驱动星云湖表层水温变化的主导因子, 二者均表现为稳定的增温效应; RHU、DSR 及 PRE 的贡献率较低, 但呈现出更为复杂的非线性响应特征与阈值效应。

参考文献 (References):

[1] SONG K S, WANG M, DU J, et al. Spatiotemporal variations of lake surface temperature across the Tibetan Plateau using MODIS LST product[J]. Remote Sensing, 2016, 8(10): 854.

[2] LI Z H, ZHANG Z J, XIONG S Q, et al. Lake surface temperature predictions under different climate scenarios with machine learning methods: A case study of Qinghai Lake and Hulun Lake, China[J]. Remote Sensing, 2024, 16(17): 3220.

[3] O'REILLY C M, SHARMA S, GRAY D K, et al. Rapid and highly variable warming of lake surface waters around the globe [J]. Geophysical Research Letters, 2015, 42(24): 10773-10781.

[4] MEI X Y, GAO S S, LIU Y, et al. Effects of elevated temperature on resources competition of nutrient and light between benthic and planktonic algae[J]. Frontiers in Environmental Science, 2022, 10: 908088.

[5] SHUVO A, O'REILLY C M, BLAGRAVE K, et al. Total phosphorus and climate are equally important predictors of water quality in lakes[J]. Aquatic Sciences, 2021, 83(1): 16.

[6] PICCOLROAZ S, ZHU S, LADWIG R, et al. Lake water temperature modeling in an era of climate change: Data sources, models, and future prospects[J]. Reviews of Geophysics, 2024, 62: e2023RG000816.

[7] 杨逸帆, 岳林坦, 邓建明, 等. 基于 BP 神经网络与遥感反演的太湖热浪规律研究[J]. 湖泊科学, 2025, 37(3): 989-999.

YANG Y F, YUE L T, DENG J M, et al. Patterns of heatwaves in Lake Taihu based on BP neural network and remote sensing inversion [J]. Journal of Lake Sciences, 2025, 37(3): 989-999.

[8] 霍俐君, 张淑萍, 王贺崑元, 等. 基于 MODIS 湖泊表面温度和 FLake 模型的 2000—2023 年青海湖热储量变化研究[J]. 冰川冻土, 2025, 47(3): 668-679.

HUO L J, ZHANG S P, WANG H, et al. Heat storage variations in Qinghai Lake from 2000 to 2023 based on MODIS lake surface temperature and FLake model [J]. Journal of Glaciology and Geocryology, 2025, 47(3): 668-679.

[9] YANG K, YU Z Y, LUO Y, et al. Spatial-temporal variation of lake

- surface water temperature and its driving factors in Yunnan-Guizhou Plateau[J]. *Water Resources Research*, 2019, 55(6): 4688-4703.
- [10] 孙静静, 王梦竹, 黄朝, 等. 典型浅水湖泊水温长期变化特征及其与气候因素相关性分析[J]. *环境污染与防治*, 2022, 44(8): 979-985.
- SUN J J, WANG M Z, HUANG C, et al. Characteristic of long-term water temperature variation in typical shallow lake and correlation analysis with climate factors [J]. *Environmental Pollution and Control*, 2022, 44(8): 979-985.
- [11] 潘惠敏, 蔡华阳, 王博芝, 等. 1973—2020 年洞庭湖水温演变特征[J]. *湖泊科学*, 2023, 35(1): 326-337.
- PAN H M, CAI H Y, WANG B Z, et al. Evolution characteristics of water temperature in Lake Dongting from 1973 to 2020[J]. *Journal of Lake Sciences*, 2023, 35(1): 326-337.
- [12] DI NUNNO F, ZHU S L, PTAK M, et al. A stacked machine learning model for multi-step ahead prediction of lake surface water temperature[J]. *Science of the Total Environment*, 2023, 890: 164323.
- [13] PICCOLROAZ S, HEALEY N C, LENTERS J D, et al. On the predictability of lake surface temperature using air temperature in a changing climate: A case study for Lake Tahoe (U. S. A.) [J]. *Limnology and Oceanography*, 2018, 63(1): 243-261.
- [14] ZHU S L, PTAK M, YASEEN Z M, et al. Forecasting surface water temperature in lakes: A comparison of approaches [J]. *Journal of Hydrology*, 2020, 585: 124809.
- [15] DUAN H M, SHANG C X, YANG K, et al. Dynamic response of surface water temperature in urban lakes under different climate scenarios: A case study in Dianchi Lake, China [J]. *International Journal of Environmental Research and Public Health*, 2022, 19(19): 12142.
- [16] BACHMANN R W, SHARMA S, CANFIELD D E, et al. The distribution and prediction of summer near-surface water temperatures in lakes of the coterminous United States and southern Canada [J]. *Geosciences*, 2019, 9(7): 296.
- [17] ZHU S L, PIOTROWSKI A P. River/stream water temperature forecasting using artificial intelligence models: A systematic review [J]. *Acta Geophysica*, 2020, 68(5): 1433-1442.
- [18] XU S S, YANG K, ZHOU X L, et al. Prediction of surface water temperature and its spatial-temporal variation characteristics of 11 main lakes in Yunnan-Guizhou Plateau [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024, 17: 2189-2200.
- [19] 陶家欣, 王远坤, 尤扬, 等. 近 60 年来洞庭湖流域水温情势演变 [J]. *生态学报*, 2025, 45(15): 7440-7454.
- TAO J X, WANG Y K, YOU Y, et al. The evolution of surface water temperature in the Dongting Lake Basin over the past 60 years [J]. *Acta Ecologica Sinica*, 2025, 45(15): 7440-7454.
- [20] HAO Z H, LI W D, WU J R, et al. A novel deep learning model for mining nonlinear dynamics in lake surface water temperature prediction [J]. *Remote Sensing*, 2023, 15(4): 900.
- [21] LIU J T, SHEN C P, O' DONNCHA F, et al. From RNNs to Transformers: Benchmarking deep learning architectures for hydrologic prediction [J]. *Hydrology and Earth System Sciences*, 2025, 29(23): 6811-6828.
- [22] AMNUAYLOJAROEN T, PTAK M, SOJKA M. Assessment of the impact of meteorological variables on lake water temperature using the SHapley additive exPlanations method [J]. *Water*, 2024, 16(22): 3296.
- [23] 温姗姗, 汪辰钰, 翟建青, 等. 基于 NEX-GDDP-CMIP6 的不同升温情景下中国气候要素变化 [J]. *水土保持学报*, 2025, 39(5): 219-233.
- WEN S S, WANG C Y, ZHAI J Q, et al. Changes in climate elements in China under different warming scenarios based on NEX-GDDP-CMIP6 [J]. *Journal of Soil and Water Conservation*, 2025, 39(5): 219-233.
- [24] THRASHER B, WANG W L, MICHAELIS A, et al. NASA global daily downscaled projections, CMIP6 [J]. *Scientific Data*, 2022, 9: 262.
- [25] ZENG Y B, CHANG F Q, WEN X Y, et al. Seasonal variation in the water quality and eutrophication of Xingyun Lake in southwestern China [J]. *Water*, 2022, 14(22): 3677.
- [26] DUAN Z Z, GAO W, CHENG G W, et al. Warming surface and Lake heatwaves as key drivers to harmful algal blooms: A case study of Lake Dianchi, China [J]. *Journal of Hydrology*, 2024, 632: 130971.
- [27] ISLES P D F. A random forest approach to improve estimates of tributary nutrient loading [J]. *Water Research*, 2024, 248: 120876.
- [28] XU L, OU Y X, CAI J J, et al. Offshore wind speed assessment with statistical and attention-based neural network methods based on STL decomposition [J]. *Renewable Energy*, 2023, 216: 119097.
- [29] 甘容, 马超鑫, 高勇, 等. 基于二次分解和支持向量机的月径流预测方法 [J]. *郑州大学学报(工学版)*, 2024, 45(6): 32-39.
- GAN R, MA C X, GAO Y, et al. Monthly runoff prediction method based on secondary decomposition and support vector machine [J]. *Journal of Zhengzhou University (Engineering Science)*, 2024, 45(6): 32-39.
- [30] 林天宙, 彭杨, 罗诗琦, 等. 基于 RF-LSTM 模型的长江上游汛期日含沙量预测 [J]. *中国农村水利水电*, 2024(10): 32-39.
- LIN T Z, PENG Y, LUO S Q, et al. Prediction of daily sediment concentrations in the upper reaches of Yangtze River during flood season based on RF-LSTM model [J]. *China Rural Water and Hydropower*, 2024(10): 32-39.
- [31] 龙鸿元, 王丽霞, 张珈玮, 等. 基于 CMIP6 气候情景的泾河流域水文模拟及预测 [J]. *水利水电技术 (中英文)*, 2025, 56(2): 89-103.
- LONG H Y, WANG L X, ZHANG J W, et al. Hydrological

simulation and prediction of the Jinghe River Basin based on CMIP6 climate scenario[J]. Water Resources and Hydropower Engineering, 2025, 56(2): 89-103.

[32] 孙祥瑜, 王超, 杨一旸, 等. CEEMDAN-Pyraformer-LSS 模型在永定河径流预测中的应用[J]. 南水北调与水利科技(中英文), 2025, 23(2): 363-374.

SUN X Y, WANG C, YANG Y Y, et al. Application of CEEMDAN-Pyraformer-LSS model in runoff forecasting of Yongding River[J]. South-to-North Water Transfers and Water Science & Technology, 2025, 23(2): 363-374.

[33] OROZCO LÓPEZ E, KAPLAN D, LINHOSS A. Interpretable transformer neural network prediction of diverse environmental time series using weather forecasts[J]. Water Resources Research, 2024, 60(10): e2023WR036337.

[34] YANG S W, CHEN J G, LIANG R F, et al. Interpretable machine learning model for reservoir outflow water temperature prediction[J]. Renewable Energy, 2026, 256(PC): 124080.

[35] WU C H, JIANG C L, JU M S, et al. Influence of meteorological elements and inflow pollution loads on cyanobacterial blooms trends; An integrated environmental fluid dynamics code model and comprehensive trend detection method [J]. Journal of Hydrology, 2025, 654: 132789.

[36] JIA T F, YANG K, PENG Z Q, et al. Review on the change trend, attribution analysis, retrieval, simulation, and prediction of lake surface water temperature[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 6324-6355.

[37] 潘梅娥, 杨昆, 邹天乐, 等. 区域气候变化下洞里萨湖表面水温时空变化的归因[J]. 地理科学, 2022, 42(4): 739-750.

PAN M E, YANG K, ZOU T L, et al. Determining factors in spatio-temporal variation of the Tonle Sap Lake's surface water temperature under regional climate change [J]. Scientia Geographica Sinica, 2022, 42(4): 739-750.

[38] YANG K, YU Z Y, LUO Y. Analysis on driving factors of lake surface water temperature for major lakes in Yunnan-Guizhou Plateau [J]. Water Research, 2020, 184: 116018.

(责任编辑 王璐)