

基于机器学习模型的东北地区湖冰物候变化特征及驱动因子分析*

彭 灿^{1,4}, 陈立文^{1,2}, 章光新^{1**}, 孙敬轩^{1,3}

(1: 中国科学院东北地理与农业生态研究所, 长春 130102)

(2: 吉林建筑大学测绘与勘查工程学院, 长春 130118)

(3: 中国科学院大学, 北京 10049)

(4: 黑龙江大学水利电力学院, 哈尔滨 150080)

摘要: 东北地区是我国中高纬度湖泊分布区和气候敏感区, 其特有的季节性湖冰物候变化对湖泊生态系统产生重要影响。然而, 东北地区缺少长时间序列湖冰物候数据, 难以识别湖冰物候变化特征。为此, 本研究基于 XGBoost-SHAP 机器学习模型, 构建了 1981-2023 年东北地区重要湖泊(呼伦湖、连环湖、查干湖、兴凯湖和卧龙湖)的湖冰物候数据集, 量化解析了湖冰物候变化特征及其驱动因子。研究结果表明: (1) 基于 XGBoost 的湖冰物候预测模型精度较高, 其中结冰日预测的决定系数(R^2)达到 0.97, 平均绝对百分比误差为 0.5%; 融冰日预测的决定系数(R^2)为 0.97, 平均绝对百分比误差为 1.9%。(2) 湖冰物候呈现“结冰日期推迟、融冰日期提前、冰封期缩短”特征及趋势, 位于高纬度地区的呼伦湖湖冰物候变化明显, 结冰日延后 0.18d/a, 融冰日提前 0.37d/a, 冰封期缩短 0.55d/a; 而纬度相对较低的卧龙湖变化相对缓和, 结冰日延后 0.13d/a, 融冰日提前 0.20d/a, 冰封期缩短 0.33d/a。(3) 湖冰物候变化的关键驱动因子是气温, 其对结冰日和融冰日的贡献度分别达到 40.5% 和 31.2%。研究结果有助于理解全球气候变化背景下寒区湖泊湖冰物候响应机制, 可为湖泊水环境保护和水生态治理提供科学支撑。

关键词: 湖冰物候, XGBoost-SHAP 模型, 驱动因子, 东北地区

A machine-learning model for identifying characteristics and driving factors of lake ice phenology changes in Northeast China*

Peng Can^{1,4}, Chen Liwen^{1,2}, Zhang Guangxin^{1**} & Sun Jingxuan^{1,3}

(1: Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130102, P.R.China)

(2: School of Geomatics and Prospecting Engineering, Jilin Jianzhu University, Changchun 130118, P.R.China)

(3: University of Chinese Academy of Sciences, Beijing 100049, P.R.China)

(4: School of Hydraulic and Electric-Power, Heilongjiang University, Harbin 150080, China)

Abstract: Northeast China, a mid- to high-latitude region with extensive lake distribution and high sensitivity to climate change, experiences unique seasonal lake ice phenology variations that profoundly influence lake ecosystems. However, the scarcity of long-term lake ice phenology records in this region has hindered the identification of changing characteristics. To address this gap, this study applied the XGBoost-SHAP machine learning approach to reconstruct a lake ice phenology dataset for key lakes in Northeast China—Hulun, Lianhuan, Chagan, Xingkai, and Wolong Lakes—from 1981 to 2023, and quantitatively analyzed the trends in lake ice phenology and their driving factors. The results indicate that: (1) the XGBoost-based lake ice phenology prediction model achieves high accuracy, with a coefficient of determination (R^2) of 0.97 for both

* 2025-07-11 收稿;2025-09-30 收修稿。

-中国科学院战略性先导科技专项(A类)项目(XDA28020501)、国家自然科学基金联合基金重点项目(U24A20570)、国家重点研发项目(2021YFC3200203)联合资助。

**通信作者;E-mail: zhgx@iga.ac.cn

freeze-up and break-up dates, accompanied by mean absolute percentage errors of 0.5% and 1.9%, respectively; (2) lake ice phenology exhibits consistent trends toward later freeze-up, earlier break-up, and shorter ice cover duration; specifically, at higher-latitude Hulun Lake, freeze-up is delayed by 0.18 d/a, break-up advances by 0.37 d/a, and ice cover duration shortens by 0.55 d/a, while at the relatively lower-latitude Wolong Lake, the corresponding rates are 0.13 d/a, 0.20 d/a, and 0.33 d/a; (3) air temperature is the dominant driver of these changes, contributing 40.5% to freeze-up date variation and 31.2% to break-up date variation. These findings advance the understanding of cold-region lake ice phenology response mechanisms to global climate change and offer scientific support for lake water environment protection and aquatic ecological management.

Keywords: lake ice phenomenon; XGBoost-SHAP model; driving factors; Northeast China

湖冰物候（包括湖泊的结冰、融冰日期及冰期持续时间）是表征湖冰冻融循环过程的关键参数，与气候变化和湖泊自身属性密切相关^[1]，其年际变化不仅可用于区域气候变化的评估^[2]，还对区域经济发展、水质演变、有害藻华发生及生态系统调节^[3]等过程具有显著影响。因此，长序列湖冰物候变化特征不仅能有效量化气候变化对湖泊的直接影响^[4,5]，也可作为反映湖泊及周围环境对气候变化响应的敏感性指标^[6,7]。近几十年来，全球湖冰物候呈现出一种受气候驱动的趋势，表现为结冰日期推迟、融冰日期提前及冰封期缩短。这种趋势具有显著空间差异性，例如中高纬湖泊受气候变暖产生的冰物候特征变化更为强烈，明显高于低纬度季节性湖泊，并且这种现象在未来有望进一步加剧^[8,9]。因此，系统研究湖冰物候时空变化特征及驱动机制，对深入理解气候变化下湖泊生态系统的生态环境效应具有重要意义。

传统湖冰物候观测主要依赖人工记录，效率低且难以实现系统化监测。虽然光学遥感(如 MODIS)和被动微波遥感(如 SSM/I)的应用显著提升了监测能力^[7,10-12]，但这些方法仍受限于云雾干扰和空间分辨率不足等问题，难以实现高精度的连续监测。近年来，机器学习技术的发展为湖冰物候监测带来了新的突破。基于 ERA5-land 再分析数据集构建基于 XGBoost 的冰覆盖率重建模型^[13] 通过融合多源数据，有效提升监测精度，揭示了气候驱动因子与湖冰物候间的关系，为解析不同规模湖泊对气候变化的响应提供了新方法。湖冰物候作为气候变化的灵敏指标，其形成机制主要由湖泊表面能量平衡过程主导，同时受多种因子共同调控^[14]。已有研究证实，受气候变化影响，近 150 年北半球湖泊的冰封期平均缩短了 28 天，且近几十年来加速趋势明显^[15]。这一过程不仅受气温^[16]、太阳辐射^[17]、风速^[7,18]、降水^[19]等气象因子变化驱动，湖泊地理位置及属性也起了一定作用^[20]。深入解析湖冰物候对气候变化的响应机制，不仅为湖泊生态系统的长期演变趋势预测提供关键参数，也能为湖泊生态系统的气候变化适应性管理提供科学依据。

东北地区是我国湖泊主要分布区和气候变化影响的敏感区，其湖冰物候的动态变化不仅直接反映当地气候特征，也对理解区域气候演变具有重要的指示意义。然而，现有研究受限于监测年限较短、技术方法相对滞后等因素，难以揭示湖冰物候长期变化特征及其对气候变化的响应规律。因此，本研究选取东北地区位于不同纬度的呼伦湖(49°N)、连环湖(47°N)、兴凯湖(45°N)、查干湖(45°N)和卧龙湖(42°N)作为代表性湖泊，创建基于 XGBoost 的湖冰物候预测模型，对其结冰日和融冰日进行预测并构建湖冰物候时序数据集，量化解析湖冰物候演变特征及规律，并利用 SHAP 可视化解释机器学习模型揭示湖冰物候变化的驱动因子。本研究为东北地区湖冰物候变化提供了更全面的数据与方法支撑，对深化湖冰变化规律认识及寒区湖泊水环境保护具有较为重要的指导意义。

1. 研究区概况

东北地区(38°43'-53°34'N, 115°31'-135°05'E)地处中高纬度地带，行政区划包括黑龙江、吉林、辽宁三省及内蒙古东一盟三市，紧邻蒙古国和俄罗斯(图 1)。东北地区属于温带季风气候区，多年平均降水量在 400-800mm 之间，气温在-5-11°C之间^[21]，冬季气温较低，平均气温在-20°C~-10°C之间，冬季降雪时间较长。区内湖泊众多，主要分布在松嫩平原、长白山、呼伦贝尔高原和三江平原等地，季节性湖冰物候特征明显，年均冰封期约 4-5 个月，通常于每年 10 月下旬至 11 月上旬开始结冰，次年 3 月中旬至 5 月中旬开始融冰^[22]。

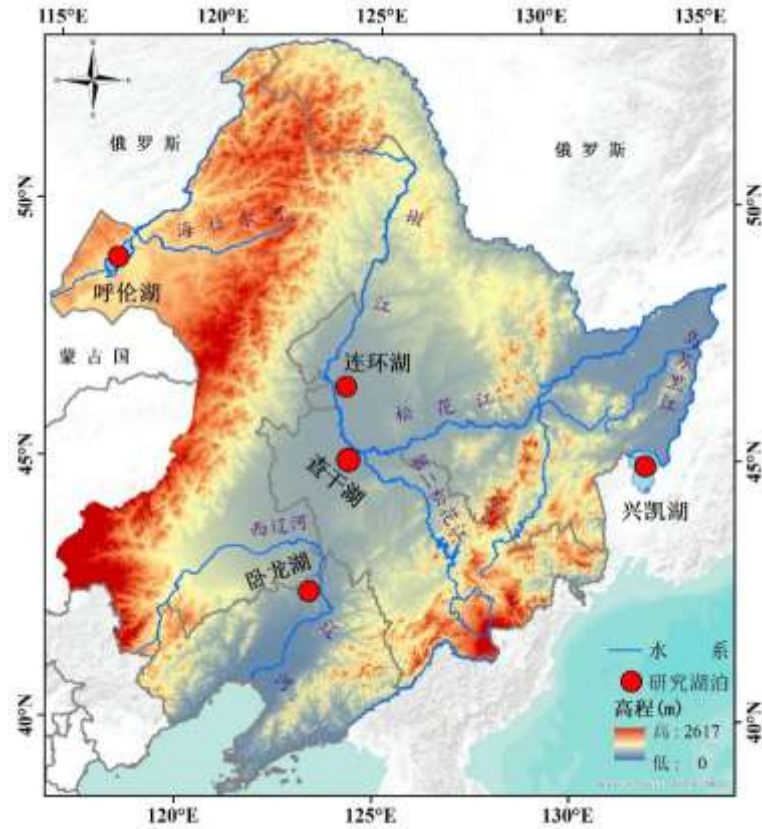


图 1 研究区湖泊位置

Fig.1 Locations of lakes in the study area

东北地区整体处于中高纬度地带，但研究对象集中在 42°N-49°N 的中纬度地带。为了揭示不同纬度湖泊的湖冰物候对气候变化的响应机制，选取呼伦湖(49°N)、连环湖(47°N)、兴凯湖(45°N)、查干湖(45°N)和卧龙湖(42°N)五个代表性湖泊作为研究对象（表 1），这些湖泊纬度间隔约 2°，在各纬度带内面积相对较大且具有连续观测数据，其中呼伦湖、兴凯湖和查干湖均位于国家级自然保护区内。

表 1 湖泊地理位置及其属性

Tab.1 The geographical location and attributes of lakes

序号	湖泊名称	中心经度 (°E)	中心纬度 (°N)	湖泊面积 (km ²)	海拔 (m)	容积 (10 ⁸ ×m ³)	平均水深 (m)
1	呼伦湖	117.40	48.94	2202.40	540.49	47.31	5.75
2	连环湖	124.11	46.60	330.57	136.00	1.36	5.20
3	兴凯湖	132.57	45.34	4190.00	68.00	18.30	4.50
4	查干湖	124.26	45.27	345.00	125.50	1.26	3.15
5	卧龙湖	123.28	42.75	56.67	88.00	0.96	1.70

2. 数据与方法

2.1 数据来源

ERA5-Land 作为欧洲中期天气预报中心（European Centre For Medium-Range Weather Forecasts, ECMWF）研发的高分辨率再分析数据集(<https://cds.climate.copernicus.eu/>)，通过先进的数据同化系统融合多源观测数据，提供了 1950 年以来全球范围内 0.1°空间分辨率和每小时时间分辨率的气候要素。虽然站点观测与网格估计值之间存在不可避免的系统差异，但大量研究^[23,24]证实该数据能准确捕捉气候要素的时空格局和长期趋势。基于此，提取 1981-2023 年春季(3-5 月)和秋冬季(10-12 月)气温、地表温度、太

阳辐射、风速和降水数据进行分析，其中，地表温度用于表征冰面的表层温度。

为分析湖泊属性对湖冰物候变化的影响，通过卫星影像中获取各湖泊中心点的经度和纬度，并基于 HydroLAKES 数据集(<https://www.hydrosheds.org/>)^[25] 获取东北地区 5 大湖泊的地理位置、面积、海拔、容积和平均深度。

此外，利用王春绪^[22]等发布 2000-2022 年中国东北地区湖冰物候数据集(<http://www.ncdc.ac.cn>)，选取呼伦湖、连环湖、查干湖和兴凯湖的湖冰物候数据，用于验证基于 XGBoost 的湖冰物候预测模型的预测结果。

2.2 方法

本研究技术流程见图 2，主要包括数据获取与处理、模型训练与输出和变化驱动因子分析三个核心部分。针对湖泊结冰日和融冰日对气象因子的响应机制的差异性，本研究创新性地构建了独立的结冰日预测模型和融冰日预测模型。关键技术环节包括：(1) 在构建结冰日和融冰日预测模型时，采用地表温度阈值法以首次连续 10 日地表温度低于 0℃ 和高于 0℃ 的日期作为起始日期，通过滑动时间窗口累计计算后续 10 日的高分辨率气象因子和湖泊属性数据作为模型输入参数，以便准确表征气象因子的滞后效应。(2) 为保证结冰日和融冰日预测模型可靠性，整合研究湖泊 1981-2023 年结冰日和融冰日通过滑动时间窗口的方式计算的累计气象数据和湖泊属性数据，共构建约 4000 组样本数据，其中结冰日和融冰日样本各 2000 组。(3) 为优化模型性能，将样本数据按 7:3 比例随机划分为训练集与验证集，通过系统评估模型的预测性能，最终保留预测性能最优的模型组合。(4) 在此基础上，采用 SHAP 模型量化解析影响湖冰物候的关键驱动因子重要性排序与贡献度。该方法较传统遥感监测平均提前 5-8 天识别冻融信号，且显著降低了人工观测的主观偏差，为湖冰物候研究提供了可靠的技术方法。

2.2.1 Delta 空间降尺度方法。Delta 降尺度方法是常用的统计降尺度技术，在全球气候模式(Global Climate Models, GCMs)降尺度过程中表现出良好的适用性。为提高研究结果的精度，采用 Delta 降尺度方法将所有气象数据统一处理至 0.01° 分辨率。首先将低分辨率的模拟数据空间插值至高分辨率网格，然后基于历史观测数据计算 1-12 月各月多年平均值间的栅格尺度校正关系，最终实现对低分辨率模拟数据的系统偏差校正和降尺度处理。具体计算公式如下：

$$\Delta(P) = P_{obs} / P_{gcms_his} \quad (1)$$

$$\Delta(T) = T_{obs} - T_{gcms_his} \quad (2)$$

$$P'_{gcms_rcp} = \Delta(P) \times P_{gcms_rcp} \quad (3)$$

$$T'_{gcms_rcp} = \Delta(T) + T_{gcms_rcp} \quad (4)$$

其中， P_{obs} ， T_{obs} 表示由观测数据得到的降水、温度等数据， P_{gcms_his} ， T_{gcms_his} 表示 GCMs 下的历史降水、温度等数据， P_{gcms_rcp} ， T_{gcms_rcp} 表示未校正的降水、气温等数据， P'_{gcms_rcp} ， T'_{gcms_rcp} 表示偏差校正后的降水、温度等数据， $\Delta(P)$ 表示降水差异系数， $\Delta(T)$ 表示温度差。

已有研究证实，Delta 降尺度方法在东北地区具有较好的适用性^[26]。因此，选择降尺度后的高分辨率气象数据作为本文分析数据。

2.2.2 XGBoost-SHAP 模型 采用 XGBoost(eXtreme Gradient Boosting)算法构建湖冰物候预测模型，该算法基于梯度提升决策树(GBDT)框架，通过集成多个弱学习器并优化目标函数，在处理非线性关系和特征交互方面有优势^[27]。在模型构建过程中，分别建立了结冰日预测模型和融冰日预测模型(表 2)，两个模型均整合了 1981-2023 年累积气象因子(气温、地表温度、太阳辐射、风速和降水数据)与湖泊属性特征作为模型输入变量，将结冰日和融冰日分别作为两个模型的预测目标变量。在数据预处理阶段，采用系统排列组合方法优化气象因子组合，并运用 Z-score 标准化方法对所有输入变量进行归一化处理。随后，通过随机分层抽样将数据集按 7:3 比例随机划分为训练集与验证集，输入基于 XGBoost 的湖冰物候预测模型进行训练和验证。模型性能通过决定系数(R^2)、均方根误差(RMSE)、平均绝对误差(MAE)、平均绝对百分比误差(MPAE)进行综合评估，最终确定基于 XGBoost 的湖冰物候预测模型的最优组合^[13]。在趋势分析方面，运用线性回归方法定量分析 1981-2023 年湖冰物候的长期变化趋势。

表 2 基于 XGBoost 的湖冰物候预测模型参数配置

参数	描述	范围
N_estimators	训练树的数量	100-200
Learning_rate	防止拟合的步长缩减值	0.01-0.30
Max_depth	树的最大深度	3-10
Subsample	训练数据的子采样比例	0.10-0.90
Colsample_by_tree	每棵树特征的子采样比例	0.1-0.90
Gamma	进一步分裂节点所需的最小损失减少量	0-5
Min_child_weight	叶节点中实例权重的最小和	1-10
Reg_alpha	L1 正则化项, 用于限制权重	0-1
Reg_lambda	L2 正则化项, 用于限制权重	0-1

SHAP(Shapley Additive Explanations) 方法是一种基于博弈论 Shapley 值的机器学习模型解释技术, 其通过量化各驱动因子在不同特征组合条件下对模型输出的边际贡献, 实现驱动因子重要性评估的公平性分配^[28,29]。本研究通过计算 SHAP 值定量解析各驱动因子对湖冰物候变化的相对贡献度。SHAP 算法的公式为:

$$\varphi_j = \frac{|S|!(p-|S|-1)!}{p!} \times (f_x(S \cup \{x_j\}) - f_x(S)) \quad (5)$$

其中, φ_j 为驱动因子 j 的 SHAP 值, S 为模型中使用的子集, p 为所有输入驱动因子的数目, x_j 为驱动因子 j 的样本数据, $\{x_j\}$ 为 j 的样本数据集合, $f_x(S)$ 为子集 S 的预测。

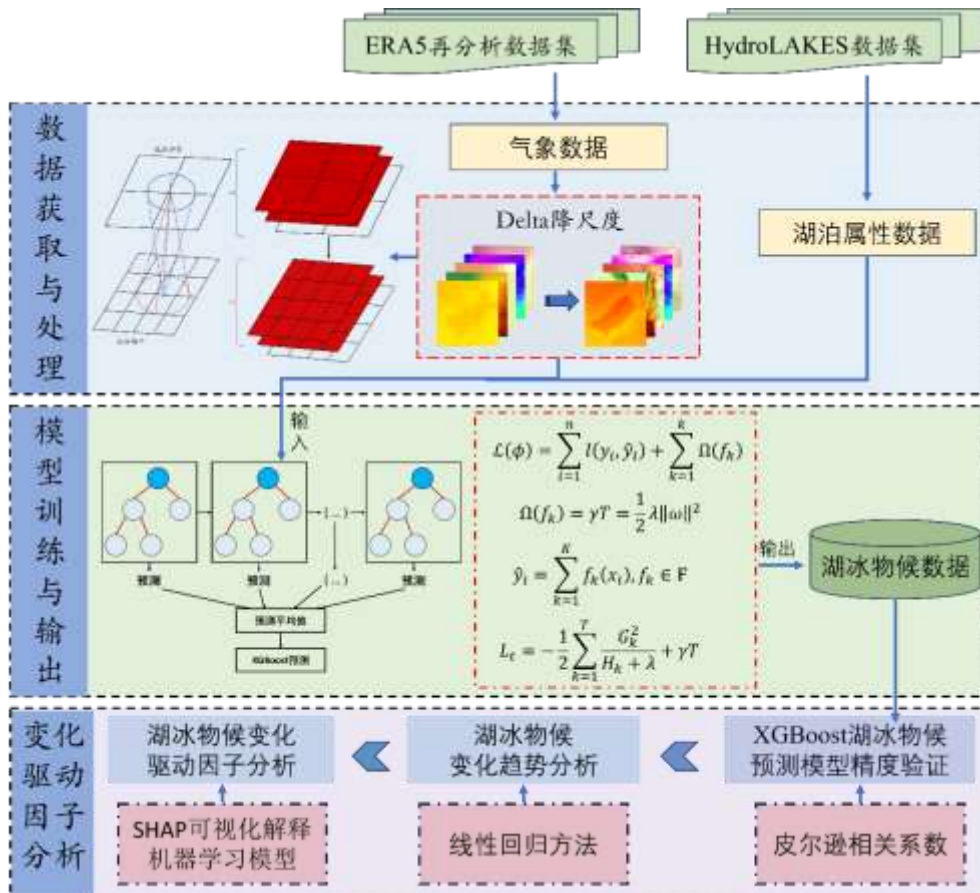


图 2 研究工作流程图

Fig.2 Research workflow chart

3. 结果

3.1 基于 XGBoost 的湖冰物候预测模型精度评估

基于 1981-2023 年高分辨率气象数据与湖泊属性数据,采用均方根误差(RMSE)和决定系数(R^2)作为评价指标,系统评估了不同驱动因子组合对基于 XGBoost 的湖冰物候预测模型性能的影响。最终确定结冰日预测的最佳特征组合为:3 日累积气温、5 日累积地表温度、10 日累积太阳辐射、3 日累积风速和 6 日累积降水;融冰日预测的最佳特征组合则为:9 日累积气温、7 日累积地表温度、9 日累积太阳辐射、4 日累积风速和 7 日累积降水。

利用基于 XGboost 的湖冰物候预测模型获取了 1981-2023 年东北地区五个代表性湖泊的湖冰物候,并与 2000-2022 年东北地区湖冰物候数据集^[22]和实测数据进行交叉验证。图 3 显示了基于 XGBoost 模型的预测结果和东北地区湖冰物候数据集^[22]之间的强相关性,验证结果显示:对于结冰日期,模型训练集预测精度达到 $R^2=0.998$ (MAPE=0.0009, RMSE=0.37d, MAE=0.29d),验证集同样保持较高精度($R^2=0.85$, MAPE=0.0009, RMSE=3.81d, MAE=2.88d);尽管融冰日期的预测模型在训练集上表现出较高的拟合优度($R^2=0.94$),在独立验证集上也保持了可接受的精度($R^2=0.74$),其预测结果的均方根误差(RMSE)却显著高于结冰日期,这表明融冰过程的预测更具复杂性,可能与融冰期冰层内部不稳定的能量变化密切相关。

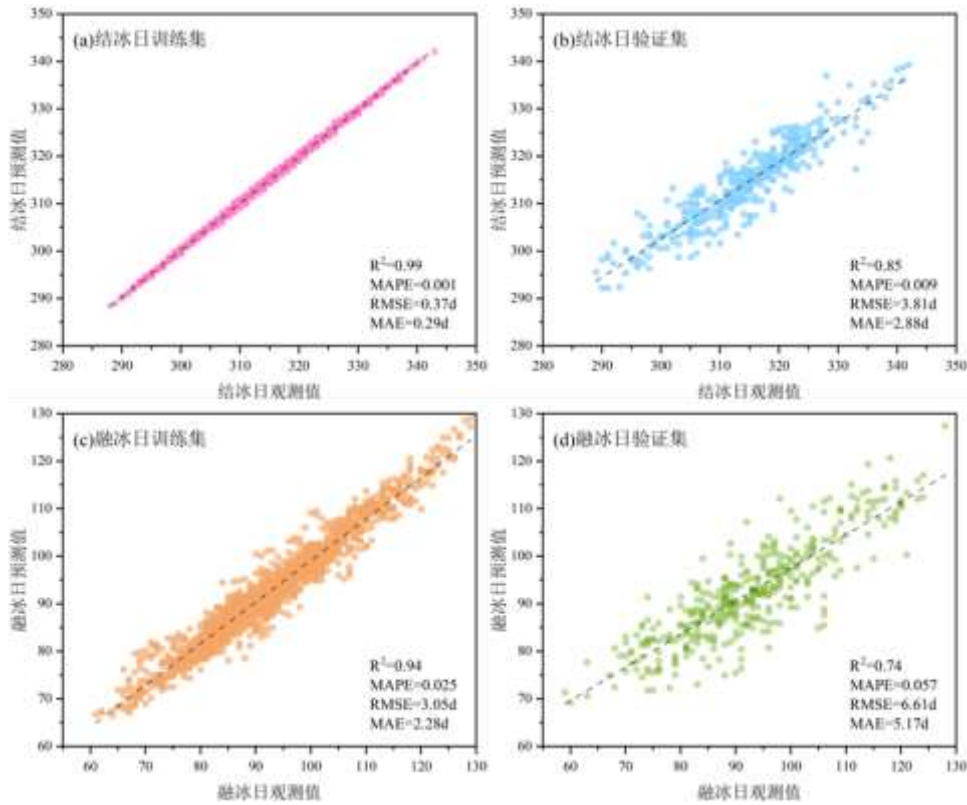


图 3 模型预测值与观测值的拟合度

Fig.3 Model prediction fit to observations

图 4 显示了基于 XGBoost 模型的预测结果和实测数据之间的强相关性,验证结果显示:对于结冰日期,模型预测精度达到 $R^2=0.97$ (MAPE=0.0045, MAE=1.36d, RMSE=1.72d);融冰日期的预测精度同样优异($R^2=0.97$, MAPE=0.019, MAE=1.71d, RMSE=2.27d)。交叉验证结果充分证明了基于 XGBoost 的湖冰物候预测模型的可靠性和准确性,为长时间序列湖冰动态研究提供了有力的方法支撑。

3.2 湖冰物候变化特征及趋势

利用线性趋势法分析了 1981-2023 年东北地区五个代表性湖泊结冰日、融冰日和冰封期的变化趋势(图 5 至图 7)。整体来看,东北地区湖泊开始结冰时间一般在 10 月中旬至 11 月中旬之间,开始融化

时间一般在3月下旬至5月中旬，直至冰层完全融化。

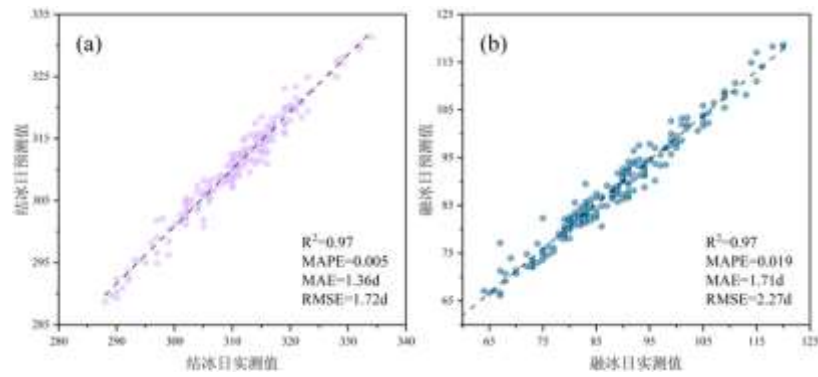


图4 模型预测值与实测值的拟合度

Fig.4 Model prediction fit to actual measurements

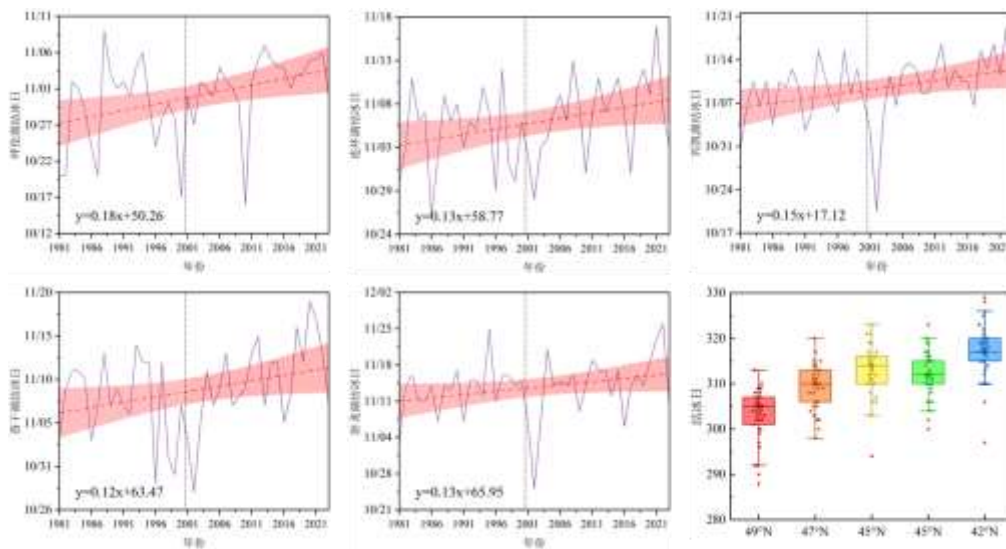


图5 1981-2023年东北地区五大湖泊结冰日变化趋势图

Fig.5 Daily variation trends of freeze-up dates in the five major lakes in Northeast China from 1981 to 2023

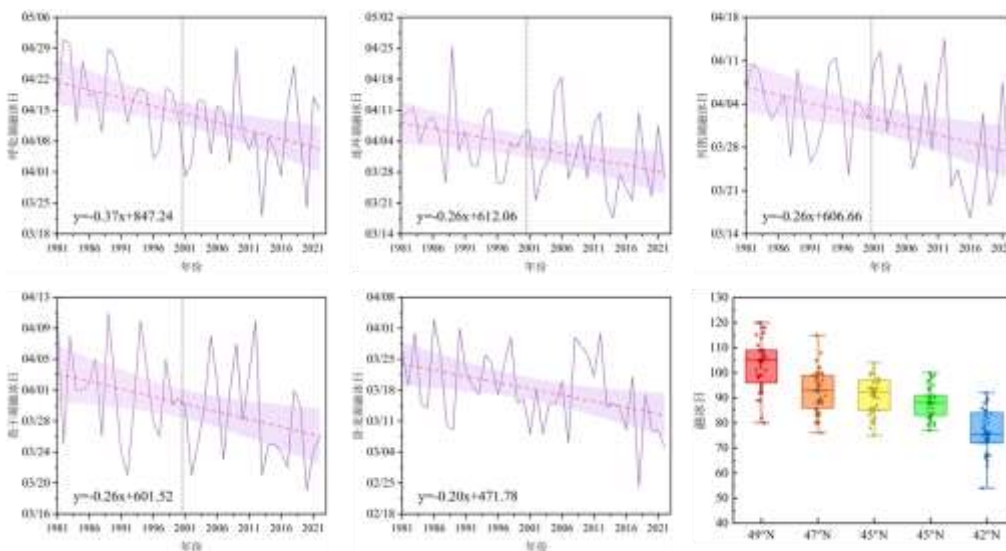


图6 1981-2023年东北地区五大湖泊融冰日变化趋势图

Fig.6 Daily variation trends of break-up dates in the five major lakes in Northeast China from 1981 to 2023

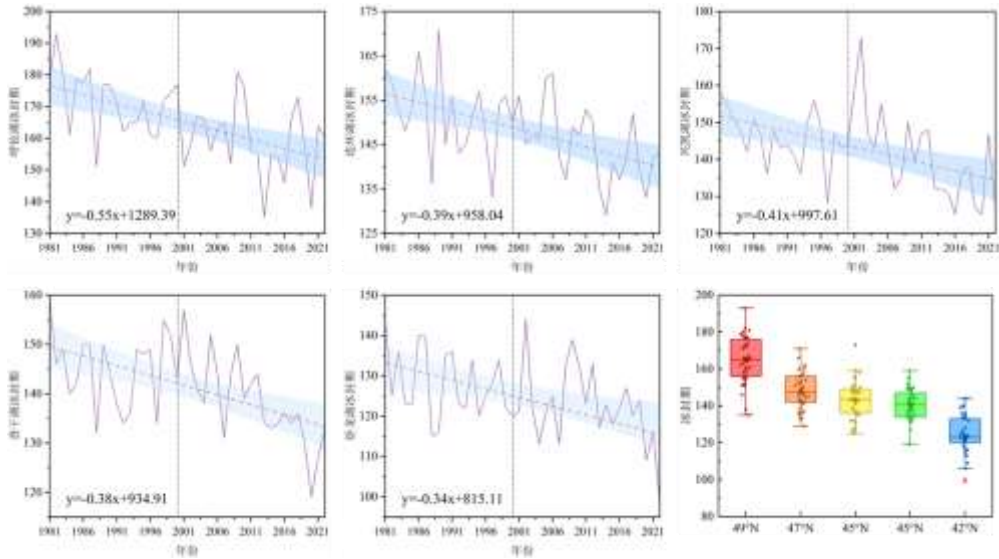


图 7 1981-2023 年东北地区五大湖泊冰封期变化趋势图

Fig.7 Daily variation trends of lake ice-cover duration in the five major lakes in Northeast China from 1981 to 2023

1981-2023 年东北地区五个代表性湖泊结冰日变化趋势图（图 5）显示不同湖泊开始结冰日期存在显著空间差异，呼伦湖开始结冰日期最早(10 月下旬至 11 月上旬)，最晚的是卧龙湖(11 月中旬)。从长期变化趋势来看，过去 43 年湖泊结冰日均呈现推迟的趋势，其中呼伦湖推迟速率最大(0.18 d/a)，其次为兴凯湖(0.15 d/a)，连环湖和卧龙湖均为 0.13d/a，查干湖推迟速率相对最小(0.12d/a)。

1981-2023 年东北地区五个代表性湖泊融冰日变化趋势图（图 6）显示不同湖泊开始融冰日期同样存在显著空间差异，卧龙湖开始融冰时间最早(3 月下旬至 4 月初)，最晚的是呼伦湖(5 月上旬)。趋势分析表明，过去 43 年湖泊融冰日均呈现提前的趋势，其中呼伦湖提前速率最大(0.37d/a)，连环湖、兴凯湖和查干湖平均每年提前 0.26d，卧龙湖提前速率相对最小(0.20d/a)。

1983-2023 年东北地区五个代表性湖泊冰封期变化趋势图（图 7）显示呼伦湖(49°N)冰封期最长(平均 168d)，卧龙湖(42°N)冰封期最短(平均 123d)，而连环湖(47°N)、兴凯湖(45°N)、查干湖(45°N)、冰封期介于其间(平均 140d)，可见冰封期持续时间与纬度息息相关。结合结冰日推迟和融冰日提前的双重变化特征，东北地区湖泊冰封期呈现明显的缩短趋势。整体而言，研究湖泊冰封期以平均 0.41d/a 的速率缩短，其中呼伦湖最大(0.55d/a)，卧龙湖最小(0.34d/a)。这种差异主要受纬度和湖泊物理特性的共同调控，位于较高纬度的呼伦湖作为大型浅水湖泊，其湖冰物候对气候的响应更加敏感；而面积较小的卧龙湖所处纬度较低，热力学过程相对稳定，因此其响应速率相对较低。

4. 讨论

4.1 基于 XGBoost 的湖冰物候预测模型精度及局限性

基于 XGBoost 的湖冰物候预测模型在东北地区五个代表性湖泊的预测值与历史观测值拟合度较高。以呼伦湖为例，模型预测的结冰日和融冰日与历史观测值偏差均小于一天^[30]，准确预测了湖泊的晚结冰和早融冰这一由气候变暖导致的典型物候变化特征。对查干湖 2021-2022 年的预测结果（结冰日 11 月 12 日，融冰日 3 月 23 日）与基于正交距离回归算法(结冰日 11 月 12 日，融冰日 3 月 22 日)高度吻合^[22]，验证了模型在东北地区湖冰物候研究中的适用性。为进一步验证模型可靠性，采用两种独立方法进行交叉验证，图 8 a 展示了查干湖 2021 年 9 月至 2022 年 7 月期间每日地表温度变化过程及基于 XGBoost 湖冰物候预测模型确定的湖冰物候，其中确定的结冰日和融冰日分别为 11 月 12 日和 3 月 23 日；图 8b-c 为使用 10 米分辨率的哨兵多光谱仪影像(RGB 真彩色)观察查干湖的冻融过程。哨兵影像显示结冰日期和融冰日期分别在 11 月 10 日和 3 月 23 日左右，与 XGBoost 预测模型所确定的结果吻合。基于哨兵影像的验证结果表明，XGBoost 湖冰物候预测模型预测值与实际观测值具有良好的一致性，结冰日预测误差

为 ± 2 天，融冰日预测结果与观测值完全吻合。这些验证结果证实了模型具有较为可靠的预测精度，为解析不同气象因子对湖冰物候的影响机制提供了重要数据支撑和理论依据。

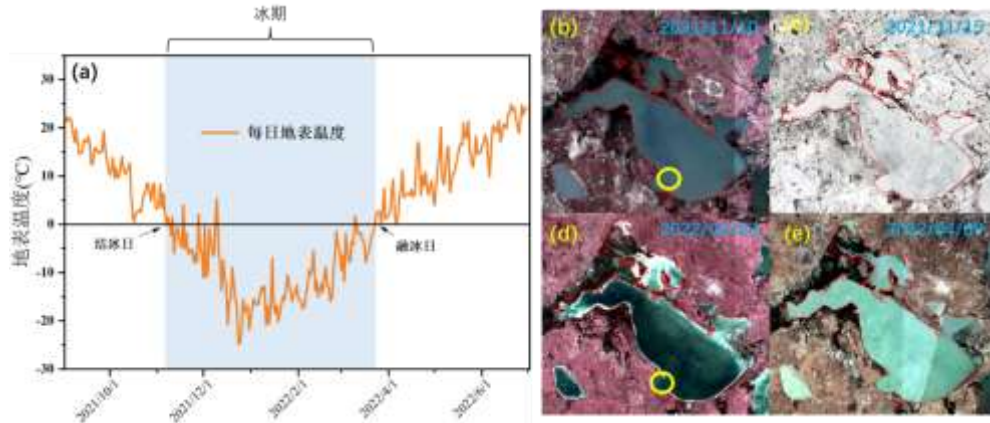


图 8 (a) 基于 XGBoost 的湖冰物候预测结果; (b-e) 10 米分辨率的哨兵多光谱影像

Fig.8 (a) XGBoost-based lake ice phenology prediction results; (b-e) Sentinel multispectral imagery with 10-meter resolution

XGBoost 湖冰物候预测模型虽然展现出较好的预测性能，但仍存在局限性。首先，模型对突发性天气波动的响应能力有限^[13]，例如强风或极端温度波动等短时气象波动可能会中断湖冰形成和融化的自然进程，这在呼伦湖观测到的阶段性冻融振荡现象中得到印证^[30]。其次，由于模型采用的日均气象数据时间分辨率不足，难以捕捉到日尺度气象要素对湖冰物候动态的瞬时响应，特别是在天气快速更替的过渡季节，可能导致 1-2 天的预测偏差，但这在湖冰物候研究中属于可接受范围，主要源于湖冰形成或消融过程本身的物理特性。相较于日尺度气象要素的瞬时波动，湖泊固有属性（如面积、水深、地形等）的空间异质性对预测结果的影响更为显著。因此，未来研究应着重改进湖泊形态特征的参数化方案，并加强区域尺度验证，以进一步提升模型的适用性和预测精度。

4.2 湖冰物候变化的驱动因子贡献度分析

湖冰物候的变化受到气象因子和湖泊自身属性的共同作用^[4,31]。基于 SHAP 模型对湖冰物候变化的驱动因子的重要性排序与贡献度分析结果表明，气温是影响湖冰物候变化最主要的气象驱动因子，在结冰日和融冰日的驱动因子重要性中均居首位。进一步通过 SHAP 模型量化各因子对湖冰物候变化的贡献度发现(图 9)，气温作为最关键的影响因子，对结冰和融冰过程分别表现出显著的抑制和促进作用，即气温升高(SHAP 值 >0)，会延迟湖泊结冰，但会加速融冰(SHAP 值 <0)。纬度在结冰和融冰过程中均是次要因子，体现了其通过调节区域热力学环境及太阳辐射季节性变化而对湖冰物候产生间接影响。纬度较高的湖泊（如呼伦湖）受太阳辐射的季节性振幅和冬季气温波动更大的共同作用，使得湖冰物候表现出更强的年际波动^[32,33]。此外，风速和太阳辐射在不同物候阶段的重要性有所差异：风速在结冰过程中作用更显著，而太阳辐射在融冰过程中占据更关键地位，反映出其在湖泊热量收支调节中的差异化作用。值得注意的是，降水在融冰阶段的贡献度相对提高，这可能与积雪层的参与有关^[34]。结冰前期湖泊主要受气-水热交换控制，而在融冰期，积雪通过改变气-雪-冰-水的能量传输过程参与表面能量收支，从而调节气温和辐射对冰层消融的作用。

除气象因子外，湖泊形态特征(如湖泊面积、海拔和平均水深)也是影响湖冰物候的重要因子^[35]。湖泊形态主要通过影响风力、水循环及热量储存等物理过程进而影响湖冰物候^[36]。然而，与气象因子和纬度的贡献度相比，湖泊形态对湖冰物候变化的贡献度相对较弱^[37]。值得注意的是，尽管研究区五大湖泊在形态上存在差异，但仍表现出相似的湖冰物候变化趋势。SHAP 分析结果进一步揭示了湖泊形态(湖泊面积与平均水深的比值)在不同阶段的重要性差异，在结冰期，湖泊形态重要性相对较高，而在融冰期，其作用明显减弱。这一现象可能源于研究区湖泊数量有限，限制了对湖泊形态敏感性的准确评估；同时在区域尺度分析中，气候因子的主导作用可能会部分掩盖湖泊形态的局地影响^[38]，从而使形态各异的湖泊表现出相似的长期变化趋势。

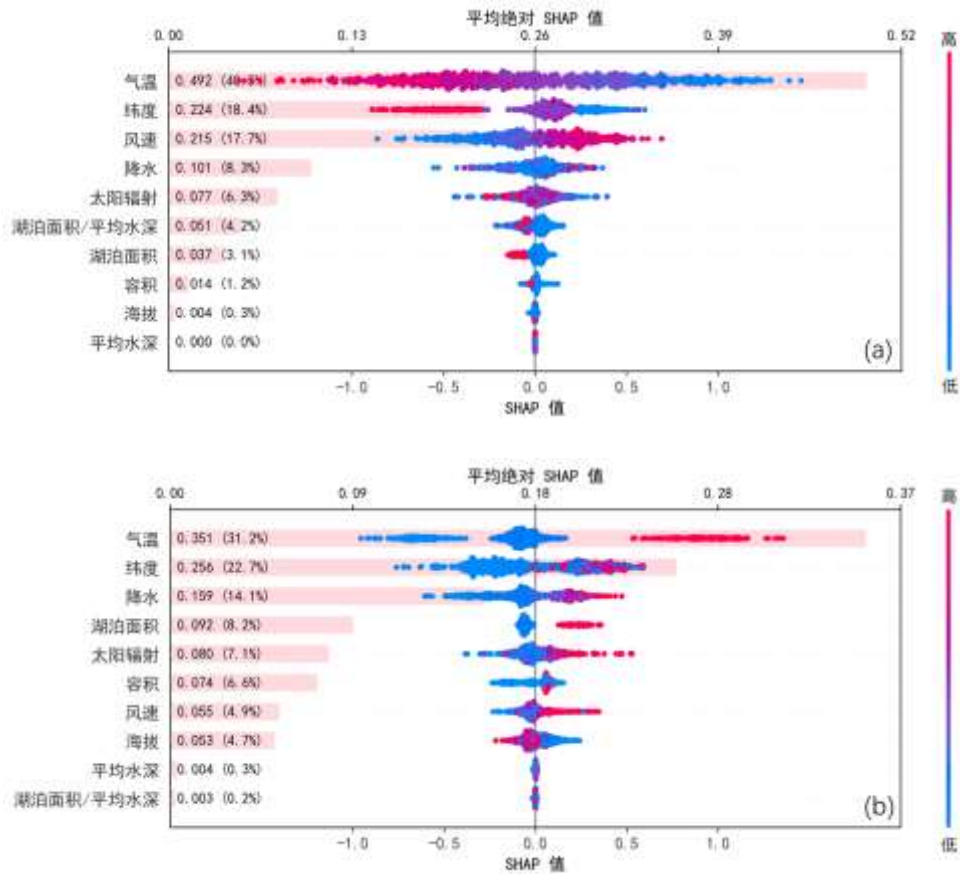


图 9 (a) 东北地区结冰日变化的驱动因子贡献度; (b) 东北地区融冰日变化的驱动因子贡献度

Fig.9 (a) Contributions of driving factors to changes in freeze-up dates in the Northeast China; (b) Contributions of driving factors to changes in break-up dates in the Northeast China.

湖冰物候与气象因子间存在动态的相互作用关系，湖冰物候可以反映各气象因子的变化，而气象因子变化可以提前或延后湖冰物候^[39]。基于 SHAP 模型解析结果，本研究以查干湖为例，结合 1981-2023 年高分辨率气象数据，通过分析结冰日与融冰日前十天关键窗口期的气象因子动态特征，发现查干湖气温、太阳辐射及风速与湖冰物候存在显著关联(图 10)，结合物理过程讨论，这些气象因子对湖冰物候的调控得到了进一步阐释。研究发现，气温区间平均值与结冰日呈显著正相关(图 10a)，与融冰日呈显著负相关关系(图 10b)，表明气温升高会显著推迟结冰并促进融冰，是影响湖冰物候变化最关键气候因子之一。太阳辐射同样表现出显著影响，其与结冰日呈正相关(图 10c)、与融冰日呈负相关(图 10d)，其作用机制在于温度升高促使冰层含水量增加，进而提升太阳辐射的吸收效率，穿透冰层的太阳辐射导致冰下水体增温，通过促进冰层底部融化进一步加速湖冰的融化^[40,41]。风速是调控湖冰物候的关键气象因子，对结冰与融冰过程具有双重作用(图 10e、f)^[42]，在结冰期，较高的风速会增强湖-气热交换并阻碍浮冰聚集，从而延迟稳定冰盖的形成；而在融冰期，风速增强会促进湖水混合，增加冰盖的融化速度^[43,44]。这些发现可为深入理解寒区湖泊对气候变化的响应提供一定的理论支撑。

4.3 湖冰物候对气候变化的差异分析

湖冰物候对气候变化的响应不仅表现为整体上的结冰日推迟和融冰日提前趋势，还存在显著的湖泊间差异^[15]。在本研究的五个代表性湖泊中，结冰日差异主要受气温、纬度、湖泊形态的综合调控。兴凯湖与查干湖虽处于近似纬度(约 45°N)，海拔差异亦有限，但结冰日推迟速率分别为 0.15 d/a 与 0.12 d/a，差异很可能源于湖泊面积和体积的差异：大型湖泊(如兴凯湖)因面积较大而降温缓慢，从而推迟结冰，而中小型湖泊(如查干湖、卧龙湖)则更易在低温条件下快速冻结^[7,20]。相比之下，呼伦湖作为纬度最高的湖泊，冬季气温更低，结冰期相对更早，但年际波动更显著，反映出高纬度湖泊对气候变化的

敏感性。

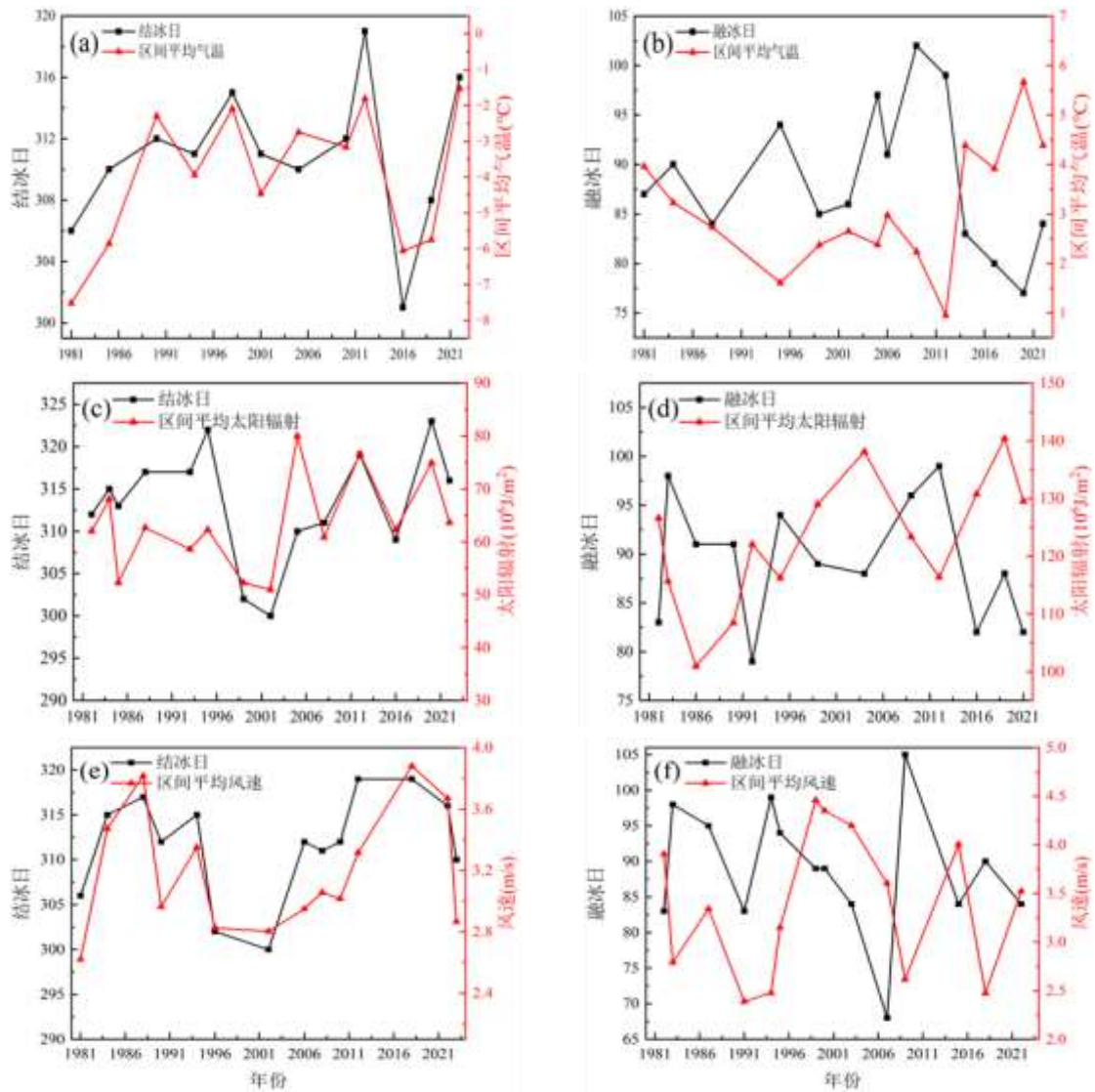


图 10 1981-2023 年查干湖冰物候变化与关键驱动因子的相关关系

Fig.10 Analysis of the relationship between lake ice phenology changes and key driving factors in Chagan Lake from 1981 to 2023

与结冰期不同，融冰日期的差异相对较小，且多个湖泊(如兴凯湖和查干湖)呈现一致的提前速率。这说明在融冰期，湖泊属性的作用被其他因子削弱。结合 SHAP 分析结果，湖泊形态(湖泊面积与平均水深的比值)在融冰期的重要性相对降低，降水的重要性相对提升。表明在春季融冰阶段，降水及其相关的积雪效应可能发挥关键作用。积雪通过改变气-雪-冰-水的能量传输过程，影响气温和辐射对冰层消融的作用，从而导致不同湖泊在融冰期表现出更为一致的提前趋势。

总体来看，东北地区湖泊的结冰日期平均以 0.14d/a 的速率推迟，融冰日期以 0.27d/a 的速率提前，整体呈现冰封期缩短趋势，与北半球湖泊的普遍规律一致^[15]。与中亚地区湖泊的接近趋势不同，青藏高原湖泊表现出更为敏感响应^[45,46]。这种空间异质性揭示了不同地理背景下主要驱动因子的区别，高原湖泊的快速变化可能与高海拔环境下的强辐射有关^[2]，而东北地区湖泊的变化则更多由气温主导。总体而言，冰封期缩短受区域气候背景、地貌特征及大尺度环流等多重因素调控^[47]。

5. 结论与展望

本文利用降尺度的高分辨率气候因子和湖泊属性数据建立了基于 XGBoost 的东北地区重要湖泊湖冰物候预测模型，揭示了湖冰物候变化特征及趋势，并利用 SHAP 可视化解释机器学习模型量化解析湖

冰物候变化的驱动因子贡献率。主要结论如下:

- (1) 基于 XGBoost 的东北地区湖冰物候预测模型具有较高的精度和适应性,证实了湖冰物候变化是气象因子持续作用的结果而非瞬时响应。该模型可为寒区湖泊冻融过程的气候响应研究提供可靠方法。
- (2) 东北地区湖冰物候呈现“结冰日期推迟、融冰日期提前、冰封期缩短”特征及趋势,与全球湖冰物候变化一致,但湖冰物候变化速率随纬度升高而增大。
- (3) 东北地区湖冰物候变化主要受气候因子和湖泊属性的驱动,其中气温是最关键的驱动因子,因冰层融化期的物理过程更为复杂,所以气温在融冰日的贡献度(31.2%)要略低于结冰日(40.5%)。
- (4) 建议开展未来气候变化情景下东北地区湖冰物候变化趋势研究,为湖泊生态系统的气候变化适应性管理提供科学支持。同时在类似的气候区推广应用本文建立的基于 XGBoost 的湖冰物候预测模型,进一步提高模型的精度和适用性。

6 参考文献

- [1] Qin QY, Li XM, Zhang B *et al.* Change of ice phenology in the Sayram Lake from 2000 to 2019. *Arid Land Geography*, 2022, **45**(1): 37-45. DOI: 10.12118/j.issn.1000-6060.2021.029.[秦启勇,李雪梅,张博等.2000—2019 年赛里木湖湖冰物候特征变化.干旱区地理,2022, **45**(1): 37-45.]
- [2] Kropáček J, Maussion F, Chen F *et al.* Analysis of ice phenology of lakes on the Tibetan Plateau from MODIS data. *The Cryosphere*, 2013, **7**(1): 287-301. DOI: 10.5194/tc-7-287-2013.
- [3] Wang W, Lee XH, Xiao W *et al.* Global lake evaporation accelerated by changes in surface energy allocation in a warmer climate. *Nature Geoscience*, 2018, **11**: 410. DOI: 10.1038/s41561-018-0114-8.
- [4] Brown, L.C.; Duguay, C.R. The response and role of ice cover in lake-climate interactions. *Progress In Physical Geography-Earth and Environment*, 2010, **34**: 671-704. DOI: 10.1177/0309133310375653.
- [5] Weber H, Riffler M, Noges H *et al.* Lake ice phenology from AVHRR data for European lakes: An automated two-step extraction method. *Remote Sensing of Environment*, 2016, **174**: 329-340. DOI: 10.1016/j.rse.2015.12.014.
- [6] Ke CQ, Tao AQ, Jin X *et al.* Variability in the ice phenology of Nam Co Lake in central Tibet from scanning multi channel microwave radiometer and special sensor microwave/imager: 1978 to 2013. *Journal of Applied Remote Sensing*, 2013, **7**: 073477. DOI: 10.1117/1.JRS.7.073477.
- [7] Cai Y, Ke CQ, Li XG *et al.* Variations of Lake Ice Phenology on the Tibetan Plateau From 2001 to 2017 Based on MODIS Data. *Journal of Geophysical Research: Atmospheres*, 2019, **124**: 825-843. DOI: 10.1029/2018jd028993.
- [8] Duguay CR, Prowse TD, Bonsal BR *et al.* Recent trends in Canadian lake ice cover. *Hydrological Processes*, 2006, **20**: 781-801. DOI: 10.1002/hyp.6131.
- [9] Latifovic R, Pouliot DR. Analysis of climate change impacts on lake ice phenology in Canada using the historical satellite data record. *Remote Sensing of Environment*, 2007, **106**: 492-507. DOI: <https://doi.org/10.1016/j.rse.2006.09.015>.
- [10] Chen XZ, Wang GY, Li WJ *et al.* Lake ice and its remote sensing monitoring in the Tibetan Plateau. *Glaciol Geocryology*, 1995, **17**: 241 - 246.[陈贤章,王光宇,李文君等.青藏高原湖冰及其遥感监测.冰川冻土,1995, **17**: 241-246.]
- [11] Jeffries MO, Morris K, Weeks WF *et al.* Structural and stratigraphic features and ERS 1 synthetic aperture radar backscatter characteristics of ice growing on shallow lakes in NW Alaska, winter 1991-1992. *Journal of Geophysical Research: Oceans*, 2012, **99**: 22459-22471. DOI: 10.1029/94jc01479.
- [12] Duguay CR, Pultz TJ, Lafleur PM *et al.* Radarsat backscatter characteristics of ice growing on shallow sub - Arctic lakes, Churchill, Manitoba, Canada. *Hydrological Processes*, 2002, **16**: 1631-1644. DOI: 10.1002/hyp.1026.
- [13] Tong J, Lin YL, Fan CY *et al.* Fine-scale monitoring of lake ice phenology by synthesizing remote sensed and climatic features based on high-resolution satellite constellation and modeling. *Science of The Total Environment*, 2024, **912**, DOI: 10.1016/j.scitotenv.2023.169002.
- [14] Reed B, Budde M, Spencer P *et al.* Integration of MODIS-derived metrics to assess interannual variability in snowpac

- k, lake ice, and NDVI in southwest Alaska. *Remote Sensing of Environment*, 2009, **113**(7): 1443-1452, DOI:10.1016/j.rse.2008.07.020.
- [15] Sharma S, Blagrove K, Magnuson JJ *et al.* Widespread loss of lake ice around the Northern Hemisphere in a warming world. *Nature Climate Change*, 2019, **9**(3): 227-+, DOI: 10.1038/s41558-018-0393-5.
- [16] Woolway RI, Kraemer BM, Lenters JD *et al.* Global lake responses to climate change. *Nature Reviews Earth & Environment*, 2020, **1**(8): 388-403, DOI: 10.1038/s43017.020.0067.5.
- [17] Cai Y, Ke CQ, Xiao Y *et al.* What caused the spatial heterogeneity of lake ice phenology changes on the Tibetan Plateau? *Science of The Total Environment*, 2022, **836**: DOI: 10.1016/j.scitotenv.2022.155517.
- [18] Guo LN, Wu YH, Zheng HX *et al.* Uncertainty and Variation of Remotely Sensed Lake Ice Phenology across the Tibetan Plateau. *Remote Sensing*, 2018, **10**(10): 1534. DOI: 10.3390/rs10101534.
- [19] Liu Y, Chen HP, Wang HJ *et al.* The Impact of the NAO on the Delayed Break-Up Date of Lake Ice over the Southern Tibetan Plateau. *Journal of Climate*, 2018, **31**(22): 9073-9086, DOI: <https://doi.org/10.1175/JCLI-D-18-0197.1>.
- [20] Sharma S, Richardson DC, Woolway RI *et al.* Loss of Ice Cover, Shifting Phenology, and More Extreme Events in Northern Hemisphere Lakes. *Journal of Geophysical Research-Biogeosciences*, 2021, **126**(10), DOI: 10.1029/2021JG006348.
- [21] Li L, Li XF, Zheng XM *et al.* The effects of declining soil moisture levels on suitable maize cultivation areas in Northeast China. *Journal of Hydrology*, 2022, **608**, 127636, DOI: 10.1016/j.jhydrol.2022.127636.
- [22] Wang Y, Zhao XS, Mamtimin AL *et al.* Evaluation of Reanalysis Datasets for Solar Radiation with In Situ Observations at a Location over the Gobi Region of Xinjiang, China. *Remote Sensing*, 2021, **13**(21), DOI: 10.3390/rs13214191.
- [23] Liu L, Gu HT, Xie JK *et al.* How well do the ERA-Interim, ERA-5, GLDAS-2.1 and NCEP-R2 reanalysis datasets represent daily air temperature over the Tibetan Plateau? *International Journal of Climatology*, 2021, **41**(2): 1484-1505, DOI: 10.1002/joc.6867.
- [24] Messenger ML, Lehner B, Grill G *et al.* Estimating the volume and age of water stored in global lakes using a geostatistical approach. *Nature Communications*, 2016, **7**, DOI: 10.1038/ncomms13603.
- [25] Wang CX, Yang Q, Shi XG *et al.* A dataset of lake ice phenology in Northeast China from 2000 to 2022. *China Scientific Data*, 2024, **9**(3): 202-216. DOI: 10.57760/sciencedb.j00001.00991.[王春绪, 杨倩, 石晓光等. 2000 - 2022 年中国东北地区湖冰物候数据集. 中国科学数据(中英文网络版), 2024, **9**(3): 202-216.]
- [26] Sun JX, Zhang GX, Wu YF *et al.* Risk assessment of agricultural green water security in Northeast China under climate change. *Science China Earth Sciences*, 2024, **67**(7): 2178-2194, DOI: 10.1007/s11430-023-1278-2. [孙敬轩, 章光新, 吴燕锋等. 气候变化下东北地区农业绿水安全风险评估. 中国科学: 地球科学, 2024, **54**(7): 2210-2226, DOI: 10.1360/SSTe-2023-0240.]
- [27] Chen TQ, Guestrin C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, 785-794. DOI: 10.1145/2939672.2939785.
- [28] Abdollahi A, Pradhan B. Explainable artificial intelligence (XAI) for interpreting the contributing factors feed into the wildfire susceptibility prediction model. *Science of The Total Environment*, 2023, **879**, 163004, DOI: 10.1016/j.scitotenv.2023.163004.
- [29] Karathanasopoulos N, Singh A, Hadjidoukas P. Machine learning-based modelling, feature importance and Shapley additive explanations analysis of variable-stiffness composite beam structures. *Structures*, 2024, **62**, DOI:10.1016/j.istruc.2024.106206.
- [30] Wu QH, Li CY, Sun B *et al.* Change of ice phenology in the Hulun Lake from 1986 to 2017. *Progress in Geography*, 2019, **38**(12): 1933-1943. DOI: 10.18306/dlkxjz.2019.12.009. [吴其慧, 李畅游, 孙标等. 1986—2017 年呼伦湖湖冰物候特征变化. 地理科学进展, 2019, **38**(12), 1933-1943.]
- [31] Cai Y, Ke CQ, Duan Z. Monitoring ice variations in Qinghai Lake from 1979 to 2016 using passive microwave rem

- ote sensing data. *Science of the Total Environment*, 2017, **607**, 120-131, DOI: 10.1016/j.scitotenv.2017.07.027.
- [32] Williams GP. Correlating freeze-up and break-up with weather conditions. *Canadian Geotechnical Journal*. 1965, **2**(4): 313-326, DOI: 10.1139/t65-047.
- [33] Palecki MA, Barry RGM. Freeze-up and Break-up of Lakes as an Index of Temperature Changes during the Transition Seasons: A Case Study for Finland. *Journal of Applied Meteorology and Climatology*, 1986, **25**(7): 893-902, DOI: 10.1175/1520-0450(1986)025.
- [34] Lei RB, Leppäranta M, Cheng B *et al.* Changes in ice-season characteristics of a European Arctic lake from 1964 to 2008. *Climatic Change*, 2012, **115**: 725-739. DOI: 10.1007/s10584-012-0489-2.
- [35] Zhang KX, Zhao L, Wang TG *et al.* Lake Ice Phenology Variations and Influencing Factors in the Qinghai-Xizang Plateau During 1978-2016. *Resources and Environment in the Yangtze Basin*, 2024, **33**(5): 1089-1098. DOI: 10.11870/cjlyzyyhj202405016. [张克新, 赵莉, 汪田归等. 1978—2016 年青藏高原湖冰物候时空变化特征及其影响因素分析. 长江流域资源与环境, 2024, **33**(5): 1089-1098.
- [36] Jeffries MO, Morris KM. Some aspects of ice phenology on ponds in Central Alaska, USA. *In Annals of Glaciology*, 2007, **46**, 397-+.
- [37] Williams SG, Stefan HG. Modeling of lake ice characteristics in North America using climate, geography, and lake bathymetry. *Journal of Cold Regions Engineering*, 2006, **20**(4): 140-167, DOI: 10.1061/(ASCE)0887-381X(2006)20:4(140).
- [38] Huang WF, Li R, Li ZJ *et al.* Analysis on the variation of ice-on date of Lakes Ngoring and Gyaring from 1979-2021 and its influencing factors. *Advances in Water Science*, 2023, **34**, 102-114. DOI: 10.14042/j.cnki.32.1309.2023.01.010. [黄文峰, 李瑞, 李志军等. 1979—2021 年鄂陵湖和扎陵湖结冰日变化及主控因素分析. 水科学进展, 2023, **34**, 102-114.
- [39] Wang ZY, Wu YH, Chang J *et al.* Temporal and Spatial Variation of Lake Ice Phenology and Its Influencing Factors in the Tibetan Plateau. *Journal of Beijing University of Technology*, 2017,**43**(5): 701-709. DOI: 10.11936/bjtxb2016110002. [王智颖, 吴艳红, 常军等. 青藏高原湖冰物候的时空变化及其影响因素. 北京工业大学学报, 2017, **43**(5): 701-709.]
- [40] Jeffries MO, Morris K, Kozlenko NM. Ice characteristics and processes, and remote sensing of frozen rivers and lakes. *Remote Sensing in Northern Hydrology: Measuring Environmental Change*, 2005, **163**, 63-90.
- [41] B JF. Climate and Time in their Geological Relations: a theory of Secular Changes of the Earth's Climate. *Nature*, 1875, **12**, 141-144, DOI: 10.1038/012141a0.
- [42] Liston GE, Hall DK. Sensitivity of lake freeze-up and break-up to climate change: A physically based modeling study. *In Annals of Glaciology*, 1995, **21**, 387-393.
- [43] Yang Q, Song KS, Wen ZD *et al.* Recent trends of ice phenology for eight large lakes using MODIS products in Northeast China. *International Journal of Remote Sensing*, 2019, **40**(14): 5388-5410, DOI: 10.1080/01431161.2019.1579939.
- [44] Leppäranta M. Modelling the Formation and Decay of Lake Ice. *The Impact of Climate Change on European Lakes*, 2010,63-83.
- [45] Xu YC, Long D, Li XD *et al.* Unveiling lake ice phenology in Central Asia under climate change with MODIS data and a two-step classification approach. *Remote Sensing of Environment*, 2024, 301. DOI: 10.1016/j.rse.2023.113955.
- [46] Tang H, Zhao YX, Niu RJ *et al.* A Study on the Evolution Characteristics of Qinghai Lake Ice in Recent 40 Years Based on an Analytical Model. *Plateau Meteorology*, 2024, **43**(5): 1152-1162. DOI: 10.7522/j.issn.1000-0534.2024.00015. [唐鸿, 赵仪欣, 牛瑞佳等. 基于分析模型的青海湖近 40 年湖冰演变特征研究. 高原气象, **43**(5): 1152-1162.
- [47] Weyhenmeyer GA, Livingstone DM, Meili M *et al.* Large geographical differences in the sensitivity of ice-covered lakes and rivers in the Northern Hemisphere to temperature changes. *Global Change Biology*, 2011, **17**, 268-275, DOI:10.1111/j.1365-2486.2010.02249.x.