

DOI:10.7522/j.issn.1000-0240.2023.0078

YUE Handong, DOU Tingfeng, LI Runkui, et al. Monthly-scale Arctic sea ice extent prediction based on multiple machine learning and stacking ensemble methods[J]. Journal of Glaciology and Geocryology, 2023, 45(3):893-901. [岳瀚栋, 窦挺峰, 李润奎, 等. 基于多种机器学习及其堆叠式集成方法的月尺度北极海冰预测研究[J]. 冰川冻土, 2023, 45(3):893-901.]

基于多种机器学习及其堆叠式集成方法的 月尺度北极海冰预测研究

岳瀚栋¹, 窦挺峰¹, 李润奎¹, 丁明虎², 效存德³

(1. 中国科学院大学 资源与环境学院, 北京 100049; 2. 中国气象科学研究院 灾害天气国家重点实验室/全球变化与极地研究所, 北京 100081; 3. 北京师范大学 地表过程与资源生态国家重点实验室, 北京 100875)

摘 要: 北极海冰范围退缩已对区域和全球气候变化、北极通航性乃至地缘政治格局产生了深远影响, 开展北极海冰预测具有重要意义。本文使用美国冰雪数据中心(NSIDC)发布的海冰密集度、海冰冰龄遥感资料以及ERA5再分析资料, 结合前人研究和海气耦合模式结果选择预测参量方案, 开展了月尺度的海冰冰情预测。比较了支持向量机(SVR)、深度森林(DF)、LightGBM(LGB)、XGBoost(XGB)和CatBoost(CAT)等5种机器学习算法和以树模型LGB、XGB和CAT作为基模型, 以贝叶斯回归、岭回归、套索回归和深度森林作为元模型的4种堆叠式集成学习模型, 以及深度神经网络(DNN)、卷积神经网络(CNN)、时空卷积网络(ConvLSTM)3种深度学习模型在2000年测试集上对海冰范围和密集度空间分布的预测效果。结果表明: 在海冰密集度预测中, ConvLSTM表现最优, 套索堆叠集成学习模型预测效果次之。集成学习模型相较于三种单一树模型在预测效果上有约1%~4%的提升。在海冰范围预测中, 堆叠式集成学习模型的预测效果最好。本研究为开展机器学习海冰预测奠定了重要基础。

关键词: 海冰预测; 机器学习; 北极; 海冰范围; 海冰密集度

中图分类号: P731.15 **文献标志码:** A **文章编号:** 1000-0240(2023)03-0893-09

0 引言

近几十年北极升温速率达到全球平均的2倍以上^[1-2]。北极海冰范围对气候变化极为敏感, 不仅表现为海冰范围退缩, 海冰的厚度和冰龄也呈明显减小趋势。气候模式预测结果显示, 北冰洋夏季无冰期将有可能在2030年左右提前到来^[3]。北极海冰快速消融给气候变化和生态环境带来诸多负面影响。比如, 增加了北极熊等生物灭绝风险, 加剧了北半球中高纬极端天气气候的发生。与此同时, 海冰退缩也使得北极航道的通航成为可能。东北航道的开通可使亚欧间的航行距离缩短约40%^[4]。北极海冰预测研究, 既有助于更好地开展北极地区和

北半球中高纬短期气候预测^[5-6], 也有助于对北极资源进行合理开发和利用, 保障科考、渔业和航运安全。

受海洋因素和大气因素的影响, 海冰变化机制异常复杂, 预测极富挑战性。目前海冰预测主要有三种方法: 统计预测方法、模式预测方法和近年来兴起的机器学习方法。统计方法主要是基于历史数据建立海冰变量随时间或特征量变化的统计模型, 对未来的海冰范围进行预测。如 Schröder 等^[7]选取了海冰表面融池作为预测特征, 通过回归分析发现融池可作为9月海冰最小范围的良好预测指标。统计方法的优点在于运算简单, 缺点是无法把握特征与预测目标之间的非线性关系。模式预测

收稿日期: 2022-09-07; 修订日期: 2022-12-16

基金项目: 国家自然科学基金项目(42222608); 国家重点研发计划项目(2018YFC1406100)资助

作者简介: 岳瀚栋, 硕士研究生, 主要从事机器学习与气候变化研究. E-mail: yuehandong19@mails.ucas.ac.cn

通信作者: 窦挺峰, 教授, 主要从事极地海冰与气候变化、大气冰冻圈研究. E-mail: doutf@ucas.ac.cn

方法是在模式动力框架、初始和边界条件的约束下,充分考虑影响海冰变化的热动力过程和机制的基础上开展预测。Wayand等^[8]发现不同海冰模式之间的预测结果差异较大,但是通过多模式集合平均方法可提升预测技能。模式预测的优点在于其预测结果的可解释性较强,缺点在于针对海冰变化过程和机理的参数化方案尚不完善,为预测结果带来较大不确定性,且对计算机算力的要求较高^[9]。相较而言,机器学习可较好地把握特征与变量之间的非线性关系,快速获得预测结果,且消耗算力较小,应用前景广泛^[10],但缺点是对预测结果的可解释性较差。

本文将基于不同机器学习方法开展北极海冰密集度和范围预测。受计算能力的限制,本文开展了月尺度预测。Kim等^[11]选取喀拉海和巴伦支海作为研究区域,对比了深度神经网络和多元回归算法对海冰密集度的预测效果,指出前者具有明显优势。Chi等^[12]对比了多层感知机(MLP),长短时记忆神经网络(LSTM),和自回归模型(autoregressive model)在海冰密集度上的预测效果,发现LSTM精度最高,均方根误差仅为8.89%。总体而言,现阶段运用机器学习方法开展的海冰预测研究较为初步,特别缺乏对不同机器学习算法在海冰预测效果方面的全面评价。本文对比分析了支持向量机(SVR)、深度森林(DF)、LGB、XGB和CAT等5种机器学习算法和贝叶斯回归、岭回归、套索回归和深度森林作为元模型、LGB、XGB和CAT作为基模型的等4种堆叠式集成学习模型,以及深度神经网络(DNN)、卷积神经网络(CNN)、时空卷积网络(ConvLSTM)3种深度学习模型对北极海冰密集度和范围的预测效果,给出了未来开展月尺度北极海冰预测的优势模型建议。

1 数据与方法

1.1 数据

本文使用的海冰密集度数据为美国冰雪数据中心(NSIDC)提供的被动微波海冰密集度气候数据记录(Climate Data Record of Passive Microwave Sea Ice Concentration),简称NSIDC海冰数据集(<https://nsidc.org/data/G02202/versions/4>)。该数据集主要基于被动微波传感器反演得到,所用算法集成了NASA戈达德中心(GSFC)开发的两种传统算法:NASA Team (NT)算法^[13]和Bootstrap (BT)算

法^[14]。该数据集空间覆盖范围为北极和南大洋海区。本文所用数据的空间分辨率为25 km(北极地区空间网格:304×448),文中使用的其他数据均通过格网匹配与上述海冰密集度网格保持一致。

海冰冰龄数据采用NSIDC发布的EASE-Grid Sea Ice Age (Version 4)(<https://nsidc.org/data/NSIDC-0611/versions/4>),其空间分辨率为12.5 km,时间分辨率为7天。该套数据融合了多套卫星遥感数据、浮标数据以及再分析数据,以拉格朗日法获得海冰运动,结合海冰的冻融生消变化计算冰龄。该数据通过取平均和格网转换与遥感数据时空分辨率相匹配。

气温、降水、气压、风速、辐射、热通量采用ERA5再分析数据(<https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview>)。ERA5为欧洲中尺度天气预报中心第五代大气再分析产品,与其前身ERA-Interim数据相比,同化了更多的遥感和台站实测数据,在时空分辨率和同化数据集的数量上有了大幅提升^[15-16],输出变量从上一代ERA-Interim的100种增加至240种,包括了更多与海冰相关的变化。相较其他再分析资料,ERA5对北极地区海-冰-气相关要素的刻画具有一定的优势,数据精度更高^[17-18]。

影响海冰变化的特征主要包括:海冰自身物理特性,以及影响海冰冻融过程的热力因素和动力因素。物理特性主要包括冰龄、皮温和反照率等^[19]。以反照率为例,主要受海冰表面形态和上覆积雪特征的影响,可用于表征海冰当前的冻融状态和吸热效率^[7]。热力影响因素方面,近年来有研究表明,降水相态变化和降雨首日是指示消融期海冰退缩速率的重要前兆信号^[20]。除海表温度和气温外,感热潜热通量和净辐射通量也是影响海冰能量平衡的重要因素。此外,大气环流的变化会对大尺度海冰动力输送产生影响。本文选取10 m风速和海表气压作为决定海冰冻融过程的主要动力特征参量。除上述因素外,大气和海洋向极热输送,陆面入海径流,冰-海耦合过程,上升流等也是促进海冰消融的重要因素^[21-22],但由于缺乏上述变量相关的大尺度观测资料,目前尚无法将这些因素全部考虑在内。本文所使用的海冰预测特征库详见表1。

1.2 方法

机器学习方法可分为传统的机器学习和深度学习两大类,集成学习是机器学习的一种组合式优

表1 海冰预测参量特征库

Table 1 Sea ice prediction feature database				
变量名	数据源	时间分辨率	空间分辨率	单位
海冰密集度	NSIDC	日均/月均	25 km	%
上月海冰密集度	NSIDC	日均/月均	25 km	%
上年海冰密集度	NSIDC	日均/月均	25 km	%
海冰冰龄	NSIDC	周均/月均	12.5 km	a
海表温度	ERA5	月均	0.25°	K
海表气压	ERA5	月均	0.25°	Pa
10 m 经纬风速	ERA5	月均	0.25°	m·s ⁻¹
总云覆盖	ERA5	月均	0.25°	%
长/短波辐射通量	ERA5	月均	0.25°	W·m ⁻²
感/潜热通量	ERA5	月均	0.25°	W·m ⁻²
2 m 气温	ERA5	月均	0.25°	K
预报反照率	ERA5	月均	0.25°	%
海冰温度	ERA5	月均	0.25°	K
总降水	ERA5	月均	0.25°	mm·d ⁻¹
降雪	ERA5	月均	0.25°	mm·d ⁻¹

化方法,主要通过构建组合多个学习器的方式完成数据学习。集成学习用给定的数据样本训练一组基学习器,用不同的模型融合策略组合基学习器得到增强模型,其往往比单个学习器具有更好的预测

精度和泛化性能^[23]。集成学习算法主要有三大类: Boosting(提升式)类, Bagging(装袋式)类和 stacking(堆叠式)类。Boosting 类集成学习算法是指通过将相似度较高的模型进行串联的算法(如 GB-DT, Adaboost, XGBoost, LightGBM, CatBoost);而个体学习器之间不存在强依赖关系进行并联的被称为 Bagging 类算法(如随机森林)。最后一类是 stacking 模型,其基本思想是将前一层模型的预测结果作为后层模型的训练集,后层模型拟合预测结果和真实值之间的关系,相当于把不同类型学习器进行串并联组合,得到更好的预测结果。

本文将全面对比不同机器学习方法在北极海冰预测中的效果,选择线性回归模型作为参照模型。机器学习算法选择 SVR(Support Vector Regression)、决策树算法、LGB 算法、XGB 算法和 CAT 算法。LGB、CAT 和 XGB 三类 Boosting 类集成学习算法的特点和差异如表 2 所示。

深度学习算法特指网络结构大于两层的神经网络算法,例如 CNN、RNN、LSTM 和 ConvLSTM 等模型,各模型特点对比如表 3 所示。其中,适用于时空序列预测的模型有 DNN、CNN 和 ConvLSTM 三种。

表2 树模型类机器学习方法对比
Table 2 Comparison of tree model machine learning methods

算法	数据处理	节点分割策略	树生长策略
LGB	GOES(单边梯度抽样算法)+列采样	直方图算法 EFB(互斥特征捆绑)	Leaf-wise(先生成增益最大的子节点)
CAT	类别特征自动处理	无偏梯度分裂	对称树(镜像二叉树)
XGB	列采样	精确贪心+近似算法	Level-wise(逐层生成子节点)

表3 深度学习模型对比

Table 3 Comparison of deep learning models		
模型名称	基本结构	特点
CNN	卷积层+池化层+全连接层	参数共享、捕捉数据细节
RNN	NN+时间层	处理时间序列,存在梯度爆炸
LSTM	门控单元(遗忘,输出,更新)	通过门控单元解决梯度爆炸问题
ConvLSTM	卷积层+LSTM	更好处理数据的时空相关性

在基于 stacking 集成学习思想的创新模型中,效果较好的是深度森林。深度森林是基于集成学习的思想,以树模型作为基础模型,参考神经网络的结

构创新而得的一种新的机器学习算法^[24-25]。其结构特点为鲜明的层联结构:将神经网络中的每一个神经元节点替换成了树模型,通过每层的训练处理自动提取特征,层层堆叠集成增加模型复杂度减少误差;其数据处理特点为多粒度扫描,即类比 CNN 卷积的方法,使用不同大小的滑窗对原始特征进行池化平均(pooling)操作,然后再输出决策树层中训练。深度森林的优势首先在于无需调参,深度森林每新增一层就会自动验证精度,当精度不再提升会自动停止增加层数,结束训练;其次在于该模型通过数据的差异性采样和模型的集成处理能够更好地防止过拟合。本文拟采用第一层为 LightGBM、CatBoost 和 XGBoost 的基模型,第二层为使用增加 L1 正则化损失函数的贝叶斯回归、岭回归、Lasso 回归和深度森

林作为元模型的4种 stacking 集成学习作为对比,其模型名分别简称为 Bayeslr_stacking、Ridge_stacking、Lasso_stacking 和 DF_stacking。

本文拟选取 1989—1999 年共计 10 年的海冰预测参量特征库数据作为训练集,以北极海冰开始显著退缩的过渡年份(2000 年)的数据作为验证集。数据集的特征 X 为 $T-1$ 月的海冰预测参量特征(表 1 中 15 个变量),预测目标 y 为 T 月的海冰密集度,通过对训练集的学习,模型拟合前一个月特征变量和本月海冰密集度之间的关系,进行提前一个月的海冰密集度预测。在控制训练数据相同的条件下,对比线性回归、SVR、LightGBM、CatBoost 和 XGBoost 以及深度森林和 stacking 集成学习模型,DNN、CNN 和 ConvLSTM 三种深度学习模型在海冰预测上的效果。海冰预测的目标量是海冰密集度,是回归问题,故采用 MAE[平均绝对误差,预测值与真实值绝对偏差的均值,其计算方法如式(1)所示]、RMSE[均方根误差,真实值与预测值误差平方的均值,该参量对异常值较敏感,能很好地反映预测数据与真值之间的偏差程度,计算方法如式(2)所示]和 R^2 [决定系数,反映预测值的全部变异性能通过回归关系被真值解释的比例,计算方法如式(3)所示]作为模型精度和效果的评估标准。

$$MAE = \frac{1}{n} \sum_{i=1}^n |(y_{i_{pred}} - y_{i_{real}})| \quad (1)$$

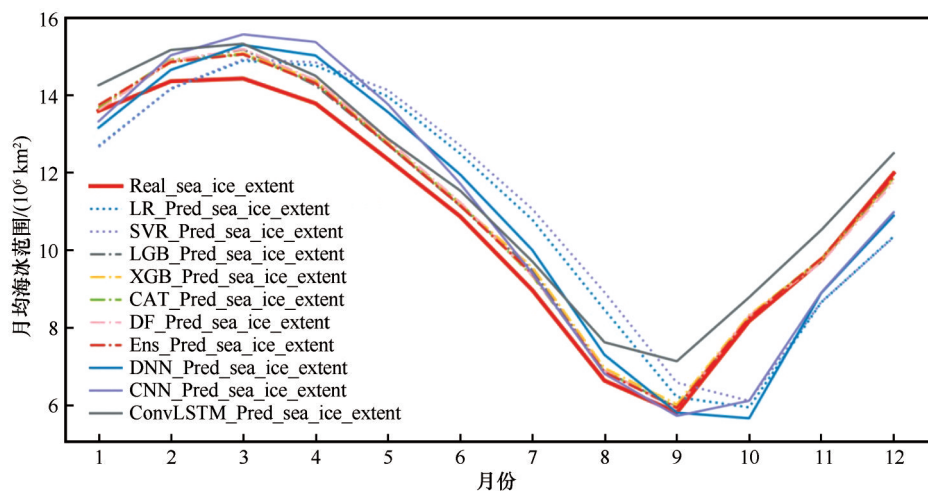


图1 各模型给出的2000年逐月海冰范围预测结果

Fig. 1 Predicted monthly sea ice extents in 2000 by various models

图2展示了各模型预测的2000年逐月海冰密集度 RMSE 值的变化,发现各模型对海冰消融期

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{i_{pred}} - y_{i_{real}})^2} \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_{i_{pred}} - y_{i_{real}})^2}{\sum_{i=1}^n (y_{i_{real}} - \bar{y}_{i_{real}})^2} \quad (3)$$

式中: $y_{i_{pred}}$ 为某格点中预测的海冰密集度值; $y_{i_{real}}$ 为对应格点中遥感观测的海冰密集度值。

海冰范围根据密集度空间数据进行计算,首先筛选出密集度大于15%的格点,然后将这些格点对应面积相加即得到海冰范围值的大小。

2 结果与分析

在控制训练使用数据量相同的前提下,经过3次重复训练,最终在测试集上的结果取平均值,从精度评价指标和可视化结果两方面对基于不同算法的海冰预测效果进行对比和评价。

2.1 预测精度对比

以2000年12个月各模型所预测的密集度结果为基础,计算了逐月海冰范围。由图1可知,ConvLSTM模型在一整年都存在高估,其他模型高估了2月到9月海冰范围,低估了1月、10月至12月海冰范围。从表4中可以看出:在海冰范围预测精度方面,Lasso_stacking 式集成学习模型预测表现最好(Ens),LGB、XGB 和 CAT 稍次之。深度学习模型海冰范围预测精度同树模型相比较差,但优于 LR 和 SVR。

(4—8月)和冻结期(10—12月、1月和2月)密集度的预测精度较低。在机器学习模型中,集成学习模

表4 各模型给出的2000年海冰范围预测精度
(单位: $\times 10^6 \text{ km}^2$)

Table 4 Prediction accuracies of sea ice extent in 2000 of various models (unit: $\times 10^6 \text{ km}^2$)

模型名	RMSE	MAE
LR	1.50	1.35
SVR	1.38	1.23
LGB	0.37	0.31
XGB	0.39	0.33
CAT	0.36	0.30
DF	0.52	0.46
Ens	0.35	0.29
DNN	1.22	1.03
CNN	1.05	0.88
ConvLSTM	0.80	0.77

型的预测精度最高,且在各月 RMSE 值变化不大。LR 和 SVR 的预测精度最差,且这两类给出的逐月 RMSE 值存在突变情况。深度学习模型中,CNN 在

3月和7月的表现优于 ConvLSTM,但在10月突增,稳定性较差。结合表5各模型的平均表现可知,ConvLSTM 在 SIC 预测中精度最高,集成学习(Ens)次之,三种树模型和 CNN 精度相近,预测效果略差于集成学习。

2.2 海冰预测可视化结果

2.2.1 机器学习可视化结果

同遥感观测相比,各种算法都可以较好地刻画海冰的空间分布。但在海冰边缘区域存在较为明显的预测偏差。结合差值图可以看出,在2000年9月的海冰密集度预测结果中,不同算法预测偏差存在较大的空间差异。三种树模型偏差在空间分布上较为类似,均在北冰洋中心区域存在低估,而在东西伯利亚海偏北地区存在高估(图3)。总体而言,深度森林(DF)、LightGBM (LGB)、XGBoost (XGB)和 CatBoost (CAT)四种模型的预测结果与观测更为接近,其在9月的 RMSE 值分别为 5.82%、5.61%、5.53% 和 4.99%,而 LR 和 SVR 的 RMSE 值分别是 6.13% 和 6.07%。

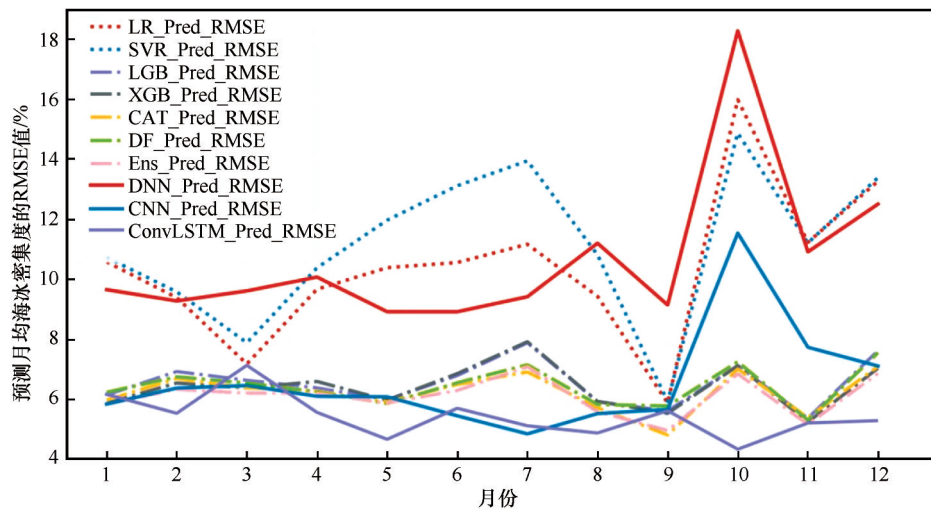


图2 各模型对2000年逐月海冰密集度的预测偏差(RMSE)

Fig. 2 RMSE of monthly sea ice concentrations in 2000 predicted by various models

2.2.2 集成学习可视化结果

不同堆叠式(stacking)集成学习的预测结果非常类似。以2000年3月为例,4种堆叠模型的偏差结果高度一致(图4)。从3月 RMSE 值量化来看, Bayeslr_stacking 为 6.21%, Ridge_stacking 为 6.20%, Lasso_stacking 为 6.17%, DF_stacking 为 6.22%,故 Lasso 堆叠集成模型的预测效果最优,而深度森林堆叠模型的预测效果稍差,故选取 Lasso 堆叠集成模型作为 Ens 集成学习模型在表4和表5

中进行统计。从表5一整年逐月 RMSE 均值来看,堆叠模型值为 6.12%, LGB 为 6.53%, XGB 为 6.41%, CAT 为 6.22%,同三种表现较优的机器学习的结果相比,堆叠集成学习模型(Ens)的预测精度更高,在海冰密集度预测精度上有约 1%~4% 的提升。

2.2.3 深度学习可视化结果

由图5可知,各类深度学习模型都能很好地刻画2000年3月和9月海冰冰情的空间分布情况。预

表5 不同模型预测的2000年逐月海冰密集度偏差
(以RMSE和 R^2 表示)

Table 5 Monthly RMSE and R^2 of sea ice concentrations in 2000 predicted by various models

模型名	逐月 RMSE 平均值/%	逐月 R^2 平均值
LR	10.39	0.91
SVR	11.14	0.90
LGB	6.53	0.96
XGB	6.41	0.97
CAT	6.22	0.97
DF	7.07	0.95
Ens	6.12	0.97
DNN	10.65	0.90
CNN	6.56	0.94
ConvLSTM	5.79	0.97

测偏差方面,深度神经网络在3月和9月有着相似的误差分布,均是在北冰洋及周边海域出现了海冰密集度的低估情况,而在较低纬度海域有所高估。而CNN和ConvLSTM预测误差的分布区域类似,但是高低估情况相反。在2000年3月的预测结果

中,CNN在大西洋一侧存在低估,在鄂霍次克海存在高估,而ConvLSTM在上述区域有着相反的高低估情况。而在9月,ConvLSTM低估了边缘区域海冰密集度,而CNN却明显高估了这些区域的海冰密集度,但ConvLSTM在低估区域面积和偏差上都小于CNN的结果。从RMSE的结果量化来看,9月CNN为5.67%,ConvLSTM为5.51%,故ConvLSTM预测结果较优(图5)。

3 结论与展望

本文基于遥感和再分析数据,开展了超前一个月的北极海冰冰情预测,对12种不同机器学习算法在海冰密集度和海冰范围上的预测效果进行了对比分析。结果显示在对海冰密集度的预测中,ConvLSTM表现最优,套索堆叠集成学习模型预测效果次之;在海冰范围预测中,堆叠式集成学习模型的预测表现最好。此外,集成学习是提升机器学习预测效果的有效方法。基于stacking方法的集成学习模型相较于三种单一树模型(LGB、XGB和CAT)在海冰密集度预测效果上有约1%~4%的提升。

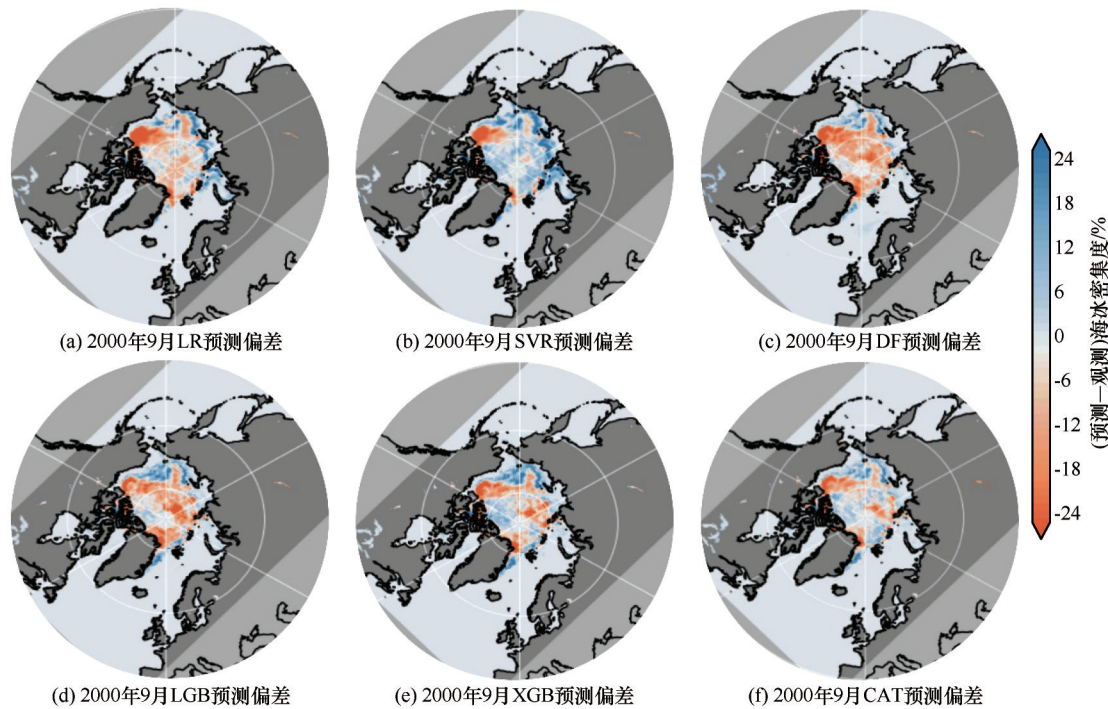


图3 不同机器学习模型预测的2000年9月海冰密集度偏差

Fig. 3 The bias of sea ice concentration between different machine learning models predicted and remote sensing observed in September, 2000

在实际应用中,除了超前一个月的预测需求外,对更长预测时效(如跨季节尺度预测)也有较大

需求,未来需针对不同预测时效,对模型预测效果进行深入对比分析,以便为特定预测时效匹配最优

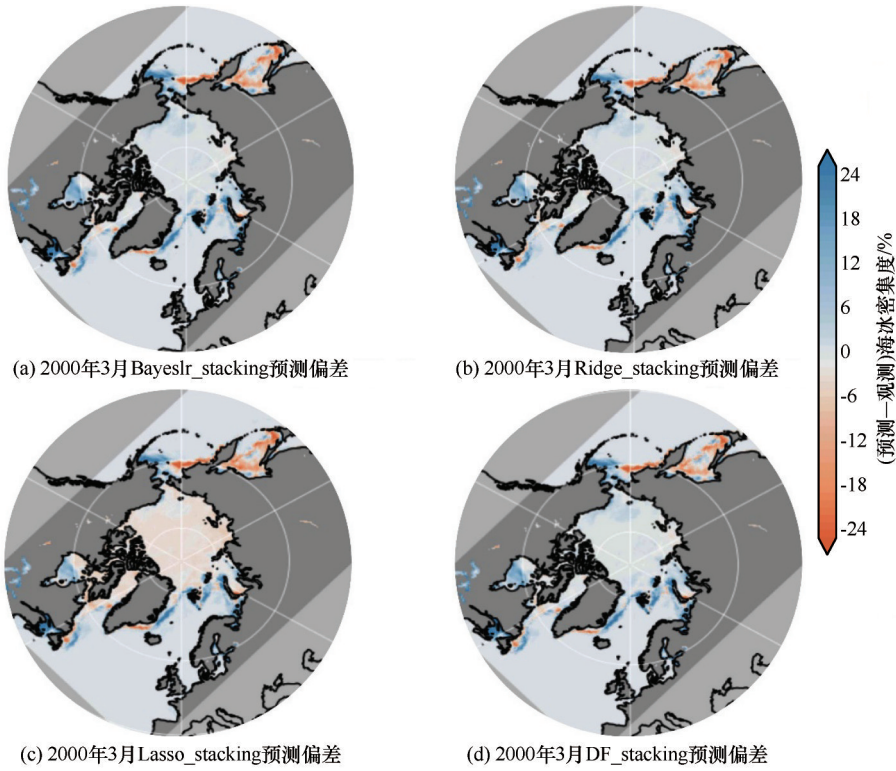


图4 不同堆叠式集成学习模型预测的2000年3月海冰密集度偏差

Fig. 4 The bias of sea ice concentration between different stacking ensemble learning models predicted and remote sensing observed in March, 2000

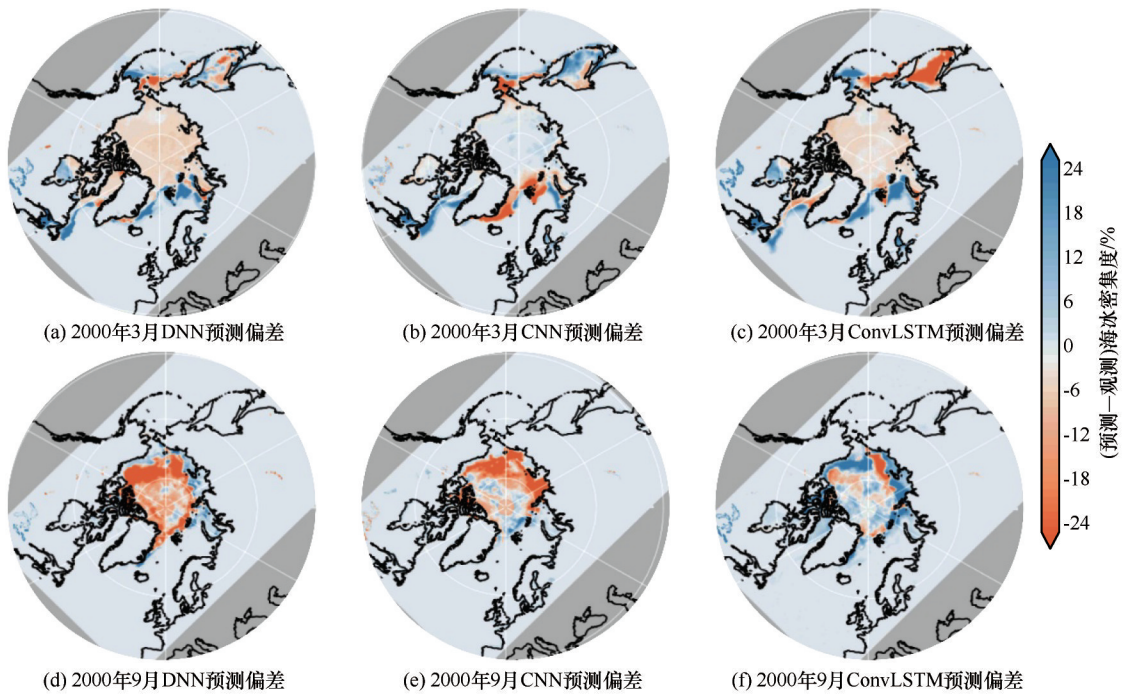


图5 不同深度学习模型预测的2000年3月、9月海冰密集度偏差

Fig. 5 The prediction bias of sea ice concentration from different deep learning models relative to remote sensing observations in March and September, 2000

模型。同时在现有工作的基础上,结合海冰预测实际及其变化机制研究进展,对现有的机器学习模型

进行结构创新,如增加物理机制约束,尝试更多集成学习模型结构,得到实际应用中预测效果更好的

机器学习模型。此外,有必要对机器学习模型的解释方法进行研究,探索通用的模型解释方法对不同模型进行分析,将机器学习海冰预测中的“黑箱”变为“灰箱”。

参考文献(References):

- [1] Serreze M C, Barrett A P, Stroeve J C, et al. The emergence of surface-based Arctic amplification [J]. *The Cryosphere*, 2009, 3(1): 11-19.
- [2] Serreze M C, Francis J A. The Arctic amplification debate[J]. *Climatic Change*, 2006, 76(3/4): 241-264.
- [3] Kim Y H, Min S K, Gillett N P, et al. Observationally-constrained projections of an ice-free Arctic even under a low emission scenario[J]. 2023, 14: 3139.
- [4] Wang C, Ding M, Yang Y, et al. Risk assessment of ship navigation in the northwest passage: historical and projection[J]. *Sustainability*, 2022, 14(9): 5591.
- [5] Xiao Cunde, Yang Jiao, Zhang Tong, et al. The predictability, irreversibility and deep uncertainty of cryospheric change [J]. *Climate Change Research*, 2022, 18(1): 1-11. [效存德, 杨佼, 张通, 等. 冰冻圈变化的可预测性、不可逆性和深度不确定性[J]. *气候变化研究进展*, 2022, 18(1): 1-11.]
- [6] Xiao Cunde, Su Bo, Dou Tingfeng, et al. Interpretation of IPCC SROCC on polar system changes and their impacts and adaptations [J]. *Climate Change Research*, 2020, 16(2): 153-162. [效存德, 苏勃, 窦挺峰, 等. 极地系统变化及其影响与适应新认识[J]. *气候变化研究进展*, 2020, 16(2): 153-162.]
- [7] Schröder D, Feltham D L, Flocco D, et al. September Arctic sea-ice minimum predicted by spring melt-pond fraction[J]. *Nature Climate Change*, 2014, 4(5): 353-357.
- [8] Wayand N E, Bitz C M, Blanchard-Wrigglesworth E. A year-round subseasonal-to-seasonal sea ice prediction portal [J]. *Geophysical Research Letters*, 2019, 46(6): 3298-3307.
- [9] Liu Quanhong, Zhang Ren, Wang Yangjun, et al. Application of deep learning method in Arctic sea ice prediction[J]. *Transactions of Atmospheric Sciences*, 2022, 45(1): 14-21. [刘泉宏, 张韧, 汪杨俊, 等. 深度学习方法在北极海冰预报中的应用[J]. *大气科学学报*, 2022, 45(1): 14-21.]
- [10] Reichstein M, Camps-Valls G, Stevens B, et al. Deep learning and process understanding for data-driven Earth system science [J]. *Nature*, 2019, 566(7743): 195-204.
- [11] Kim J, Kim K, Cho J, et al. Satellite-based prediction of Arctic sea ice concentration using a deep neural network with multi-model ensemble[J]. *Remote Sensing*, 2018, 11(1): 19.
- [12] Chi J, Kim H. Prediction of arctic sea ice concentration using a fully data driven deep neural network [J]. *Remote Sensing*, 2017, 9(12): 1305.
- [13] Cavalieri D J, Gloersen P, Campbell W J. Determination of sea ice parameters with the NIMBUS 7 SMMR[J]. *Journal of Geophysical Research: Atmospheres*, 1984, 89 (D4): 5355-5369.
- [14] Comiso J C. Characteristics of Arctic winter sea ice from satellite multispectral microwave observations [J]. *Journal of Geophysical Research*, 1986, 91(C1): 975.
- [15] Hersbach H, Bell B, Berrisford P, et al. The ERA5 global reanalysis[J]. *Quarterly Journal of the Royal Meteorological Society*, 2020, 146(730): 1999-2049.
- [16] Dee D P, Uppala S M, Simmons A J, et al. The ERA-Interim reanalysis: configuration and performance of the data assimilation system[J]. *Quarterly Journal of the Royal Meteorological Society*, 2011, 137(656): 553-597.
- [17] Graham R M, Hudson S R, Maturilli M. Improved performance of ERA5 in Arctic gateway relative to four global atmospheric reanalyses[J]. *Geophysical Research Letters*, 2019, 46 (11): 6138-6147.
- [18] Lindsay R, Wensnahan M, Schweiger A, et al. Evaluation of seven different atmospheric reanalysis products in the Arctic [J]. *Journal of Climate*, 2014, 27(7): 2588-2606.
- [19] Kwok R. Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958—2018)[J]. *Environmental Research Letters*, 2018, 13(10): 105005.
- [20] Dou T, Xiao C, Liu J, et al. A key factor initiating surface ablation of Arctic sea ice: earlier and increasing liquid precipitation[J]. *The Cryosphere*, 2019, 13(4): 1233-1246.
- [21] Li S, Lin P, Dou T, et al. Upwelling of Atlantic Water in Barrow Canyon, Chukchi Sea [J]. *Journal of Geophysical Research: Oceans*, 2022, 127(3): e2021JC01783.
- [22] Li Shutong, Dou Tingfeng, Xiao Cunde. Research progress for the changes of Arctic Ocean surface wave with diminishing sea ice[J]. *Transactions of Atmospheric Sciences*, 2021, 44(1): 118-127. [李姝彤, 窦挺峰, 效存德. 海冰覆盖度变化下北冰洋海浪研究进展[J]. *大气科学学报*, 2021, 44(1): 118-127.]
- [23] Gagne D J, McGovern A, Xue M. Machine learning enhancement of storm-scale ensemble probabilistic quantitative precipitation forecasts [J]. *Weather and Forecasting*, 2014, 29(4): 1024-1043.
- [24] Zhou Z H, Feng J. Deep forest[J]. *National Science Review*, 2019, 6(1): 74-86.
- [25] Feng J, Yu Y, Zhou Z H. Multi-layered gradient boosting decision trees[J]. *Advances in Neural Information Processing Systems*, 2018, 31: 1-11.

Monthly-scale Arctic sea ice extent prediction based on multiple machine learning and stacking ensemble methods

YUE Handong¹, DOU Tingfeng¹, LI Runkui¹, DING Minghu², XIAO Cunde³

(1. College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China; 2. State Key

Laboratory of Severe Weather / Institute of Global Change and Polar Meteorology, Chinese Academy of Meteorological

Sciences, Beijing 100081, China; 3. State Key Laboratory of Earth Surface Processes and

Resource Ecology, Beijing Normal University, Beijing 100875, China)

Abstract: The rapid retreat of Arctic sea ice has attracted extensive international attention, with far-reaching impacts on navigation periods, regional and global climate change, and geopolitical Pattern. It is of great significance to predict Arctic sea ice. In this study, the sea ice concentration, ice age data from NSIDC and ERA5 re-analysis were used to select prediction parameter schemes based on previous studies and coupled Ocean-Atmosphere model results. This study focused on the monthly-scale sea ice prediction. Five machine learning algorithms, including support vector machine (SVR), Deep forest (DF), LightGBM (LGB), XGBoost (XGB) and CatBoost (CAT), and four stacked ensemble learning models using Bayesian regression, ridge regression, Lasso regression and deep forest as meta models, three tree models LGB, XGB and CAT as base models, as well as deep neural network (DNN), convolutional neural network (CNN) and spatio-temporal Convolution network (ConvLSTM) were compared and evaluated for the predictions of sea ice concentration and extent in 2000 test set. The results show that ConvLSTM has the best performance in the prediction of sea ice concentration, and lasso stacking ensemble learning model takes the second place. The ensemble learning model improved the prediction performance by about 1% to 4% compared with the three single tree models. In sea ice extent prediction, stacking ensemble learning model performs best. This study has laid an important foundation for the monthly-scale sea ice prediction based on the machine learning.

Key words: sea ice prediction; machine learning; Arctic; sea ice extent; sea ice concentration

(责任编辑: 杨清华)