

机器学习在 ENSO 预测会商中的应用

李晨彤

(国家海洋环境预报中心,北京 100081)

摘要: 基于多模式集合预报的思想,利用可解释机器学习方法——决策树算法建立了多模式 ENSO 预测结果智能会商系统。分别使用基于 Boosting 的 GBDT、XGBoost、lightGBM 和基于 Bagging 的 RF 4 种决策树模型方法,结合随机搜索交叉验证、网格搜索交叉验证两种超参数调整方法对决策树模型的超参数进行优化调整,根据不同超前预报时效分别建立多模式 ENSO 预测结果智能会商系统,对多模式预测结果进行集合订正,并给出各模式预测结果在智能会商系统中的特征重要性。该智能会商系统模拟了 ENSO 预测会商过程,实现了读取各模式预测结果、训练模型、给出预测结论及预测依据、预测结果可视化等流程的自动化,同时实现了智能调参的功能。

关键词: ENSO;可解释机器学习;多模式;智能会商

中图分类号: P456 **文献标识码:** A **文章编号:** 1003-0239(2022)01-0091-13

1 引言

厄尔尼诺-南方涛动(ElNiño-Southern Oscillation, ENSO)预测模型目前大体可分为动力模式和统计模式两种^[1-4]。动力模式(确定论)利用了物理规律,没有利用或没有充分利用已有的大量实况历史资料;统计模式(概率论)利用了积累的大量实况资料,却没有利用或没有充分利用我们掌握的物理规律。动力模式在短期预测上是成功的,但对长期预测而言,单纯的动力学方法难以奏效。统计模式基于历史资料进行统计分析,无法区分现有资料中哪些联系是本质的,哪些是偶然的。实践表明,动力方法和统计方法都有一定的准确率,两者都能反映大气运动的部分规律^[5-7]。

目前,ENSO 预测模式的预测能力仍然表现出不稳定性^[8-9]。研究显示,采用统计方法对模式预测结果进行集合订正,可以减小模式误差对预报的影响^[10-11]。多模式集合订正是提升 ENSO 预报技巧和预测能力的有效方法^[12-17]。机器学习作为一项新的技术,它的优势之一是对机制尚未清楚的问题进行

分析、联想、记忆、学习和推断^[18]。机器学习方法的引入使得气象预报和气候预测的运算量减小、运算速度加快、运算精度提高,但也存在过拟合和泛化效果差等问题^[19-25]。因此,需要寻找合适的机器学习方法对 ENSO 多模式预测结果进行集合订正,提高 ENSO 预测的准确率,更好地应对 ENSO 事件产生的影响。

近年来,国内外研究人员基于各种机器学习方法开展海表面温度(Sea Surface Temperature, SST)、海表面温度异常(Sea Surface Temperature Anomaly, SSTA)和 Niño3.4 指数的预测研究工作。Ham 等^[26]使用卷积神经网络(Convolutional Neural Networks, CNN)进行了 18 M 超前预报时效的 SSTA 和 Niño3.4 指数预测。1984—2017 年的预报结果表明, CNN 模型的 Niño3.4 指数预报技巧远高于目前最先进的动力模式,同时,该模型也能更好地预测海面温度的详细区域分布,克服了动态预测模型的弱点。Zhang 等^[27]使用长短时记忆模型(Long Short-Term Memory, LSTM)对渤海 SST 进行预测,较传统方法的预报准确度有所提升,均方根误差(Root

收稿日期: 2021-04-21; 修回日期: 2021-06-16。

基金项目: 国家自然科学基金青年科学基金(41907180)。

作者简介: 李晨彤(1995-),男,硕士,主要从事气象预报研究。E-mail:1170964453@qq.com

Mean Squard Error, RMSE)减小。Aguilar-Martinez等^[28-29]分别使用贝叶斯神经网络、支持向量回归和线性回归建立热带太平洋区域 SSTA 预测模型对 ENSO 进行预测,可信度较高,超前预报时效较传统模型可提升至 12 M,且非线性预报模型优于线性预报模型。Nooteboom^[30]等结合自回归整体滑动平均方法(Autoregressive Integrated Moving Average Model, ARIMA)和人工神经网络(Artificial Neural Networks, ANN)建立了一种 Niño3.4 指数预报模型,并根据 Zebiak-Cane(ZC)模式的测试结果进行特征选择,针对不同超前预报时效选取不同输入因子。该模型在 6 M 以内超前时效上的技巧评估优于气候预测系统版本 2(Climate Forecast System version 2, CF Sv2)集成。此外,2010 年 1 月起报的预测结果优于 CFSv2 集合的所有成员。许柏宁等^[31]使用序列到序列模型预测 Niño3.4 指数,可以较好地预测出 SSTA 变化趋势,但在峰值处表现较差。其结果与传统的动力学 ENSO 预报模型相比,在中长期(提前 7 M 以上)预测上的 RMSE 表现更好。何丹丹等^[32]搭建了基于 Attention 机制的序列到序列预报模型进行 Niño3.4 区 SST 预测,在短期预测上其预测结果与其他方法相当,长期预测中 RMSE 比其他方法下降了 0.3~0.4。蒋国荣等^[33]使用后向传播(Back Propagation, BP)算法进行 ENSO 预报,可以较好地预测 ENSO 事件(关键区海温的变化趋势)。但预报技巧评估依赖于预报时效,当预报时效较短时,预报技巧评估较好,反之,则较差。结果还表明,增加隐含层神经元个数对预报技巧评估改进不大。以上研究说明,搭建机器学习模型进行 ENSO 预测,在短期预测和长期预测上预报技巧评估都相当或优于传统的动力模式和统计模式^[34]。跟传统方法一样,也存在随着预报时效增加,预报技巧评估变差的现象。

本研究基于预测会商工作的思想,利用可解释机器学习方法——决策树算法建立了多模式 ENSO 预测结果智能会商系统,模拟会商工作的流程,并给出预测结论及预测依据。在建模过程的可解释方面,该智能会商系统给出了各模式预测结果的特征重要性。不同的决策树模型表现出较为一致的偏好,超前预报时效较短时偏重于动力模式,较长时偏重于统计模式。这与在不同超前预报时效上,动力模式和统计模式预报技巧的高低相吻合。黑

盒模型无法给出该解释,只能给出预测结果。

2 数据与方法

2.1 数据获取及预处理

特征值:从美国哥伦比亚大学气候预测国际研究所(International Research Institute for Climate and Society, IRI)官网(网址:<https://iri.columbia.edu/~forecast/ensofcst/Data/>)下载多个模式的季节性(3 M 滑动平均)Niño3.4 指数预测结果。目前,IRI 官网收录的模式数达十余个,基于特征数据集的连续性和完整性考虑,筛选出预测结果时间序列较长和缺失值较少的模型预测结果作为订正对象,输入智能会商系统。动力模式来自欧洲中期天气预报中心(European Centre for Medium-Range Weather Forecasts, ECMWF)、日本气象厅(Japan Meteorological Agency, JMA)、美国国家航空航天局(National Aeronautics and Space Administration, NASA)、美国国家环境预报中心(National Centers for Environmental Prediction, NCEP)、美国哥伦比亚大学拉蒙地球观测中心(Lamont-Doherty Earth Observatory, LDEO)、韩国气象局和首尔国立大学(Korean Meteorological Administration and Seoul National University, KMA SNU);统计模式有美国气候预测中心马尔可夫(Climate Prediction Center MRKOV),美国气候预测中心模拟构建(Climate Prediction Center Constructed Analog, CPC CA)、美国科罗拉多州立大学气候预测(Colorado State University CLImate PRediction, CSU CLIPR)。

标签值:从美国国家环境预报中心(网址:<https://www.cpc.ncep.noaa.gov/data/indices/>)下载 ERSST v5 季节性(3M 滑动平均)Niño3.4 指数数据,作为监督学习的标签值。

缺失值处理:对各模式预测结果的缺失值进行插值处理。

2.2 数据集构造

依据训练集(2002—2017 年数据)和测试集(2018—2020 年数据)的划分,分别构造多模式超前 1~9 个季节性滑动平均月(下文中分别简称超前 1~9 M)预报时效预测结果的数据集。

2.3 预报技巧评估指标

皮尔逊相关系数(Pearson Correlation Coefficient, PCC) r : 衡量预测值同观测值之间相关程度的量。

$$r = \frac{\sum_{i=1}^n (f_i - \bar{f})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (f_i - \bar{f})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

均方根误差(Root Mean-Square Error, RMSE): 衡量预测值同观测值之间离散程度的量。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (f_i - y_i)^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2} \quad (2)$$

$$e_i = f_i - y_i$$

平均绝对误差(Mean Absolute Error, MAE): 衡量预测值同观测值之间偏差大小的量。

$$MAE = \frac{1}{n} \sum_{i=1}^n |f_i - y_i| = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (3)$$

式中, f_i 为预测值; y_i 为观测值; e_i 为预测值与观测值之间的误差。

2.4 决策树算法

决策树(Decision Trees)是一种用来分类或者回归的监督学习方法^[35-36], 相较其他机器学习方法, 决策树模型具有所需训练数据少、准确性高和可解释性等方面的优势。因此, 本研究选用可解释的决策树模型进行机器学习建模。

单棵决策树容易出现过拟合的问题, 泛化能力有限。通过某种集合策略将多颗决策树组合起来的决策树算法——集成学习(Ensemble Learning), 能够构建并结合多个学习器来完成学习任务, 泛化能力得到提升。根据个体学习器的生成方式, 集成学习方法大致可以分为提升法(Boosting)和套袋法(Bagging)^[37]。

Boosting是一簇可将弱学习器提升为强学习器的算法。基于Boosting的方法主要有: 梯度提升决策树(Gradient Boosting Decision Tree, GBDT)^[38]、极端梯度提升(eXtreme Gradient Boosting, XGBoost)^[39]和轻量梯度提升机(light Gradient Boosting Machine, lightGBM)^[40]。

Bagging基于自助采样法(bootstrap sampling), 也叫有放回重采样法。基于Bagging的方法主要为

随机森林(Random Forest, RF)^[41]。

2.5 超参数调整方法

机器学习算法中的调优参数(Tuning Parameters)需要人为设定, 称为超参数(Hyper Parameter)。在算法中, 它们作为参数传递给估计器类的构造函数。用于搜索超参数最佳配置的策略包括: 网格搜索交叉验证(Grid Search with Cross-Validation, Grid Search CV)和随机搜索交叉验证(Randomized Search with Cross-Validation, Randomized Search CV)等。本文结合两种方法的优势, 扬长避短, 先通过Grid Search CV寻找出超参数最佳配置的大致范围, 然后利用Randomized Search CV在该范围内搜索出超参数的最佳配置。

3 基于决策树模型的ENSO多模式预测结果智能会商系统

在预报预测产品发布前, 通常需要一个必不可少的会商过程来综合汇总各预报结果, 博采众长, 以有效降低或避免因主观性和经验性导致的空报、错报甚至漏报, 提高预报准确率。会商通常包括收集整理、分析处理和可视化等步骤, 并作出最终的预报结论及依据。在此过程中, 存在着大量重复性工作。因此, 本实验借鉴会商工作的思想, 借助机器学习方法在快速完成繁琐工作方面的优势, 模拟ENSO预测会商过程, 建立一个多模式的ENSO预测结果智能会商系统, 实现流程的智能化和智慧决策。该智能会商系统可以实现读取各模式预测结果、智能调参、给出Niño3.4指数预测结论及预测依据和可视化预测结果等一系列流程的自动化。

3.1 基于GBDT算法的ENSO多模式预测结果订正模型

基于GBDT的订正模型中各模式结果特征重要性方面(见表1), 超前1~5 M和超前7 M的动力模式预测结果平均特征重要性高于统计模式, 超前6 M和超前8~9 M的统计模式高于动力模式。

从GBDT模型的订正结果时间序列图可以看出(见图1), 超前1~3 M的相位和强度与标签值基本一致; 超前4~5 M的相位和强度与标签值偏差

表1 GBDT订正模型的模式预测结果特征重要性及订正结果评估

Tab.1 Feature importance of the prediction results of GBDT revised model and evaluation of the revised results of GBDT revised model

模式		超前1 M	超前2 M	超前3 M	超前4 M	超前5 M	超前6 M	超前7 M	超前8 M	超前9 M
动力模式特征重要性	ECMWF	0.302	0.639	0.602	—	—	—	—	—	—
	JMA	0.067	0.228	0.258	0.572	0.509	—	—	—	—
	NASA	0.038	0.004	0.010	0.000	0.072	0.216	0.225	—	—
	GMAO									
	NCEP	0.530	0.011	0.025	0.217	0.115	0.228	0.366	—	—
	CFS									
	LDEO	0.011	0.011	0.000	0.000	0.033	0.033	0.105	—	—
统计模式特征重要性	KMA	0.002	0.017	0.000	0.000	0.046	0.070	0.002	0.193	0.211
	SNU									
	CPC	0.029	0.062	0.064	0.067	0.068	0.148	0.144	0.307	0.309
	MRKOV									
	CPCCA	0.017	0.021	0.001	0.099	0.095	0.289	0.158	0.377	0.408
	CSU	0.002	0.006	0.039	0.045	0.062	0.016	0.000	0.122	0.071
	CLIPR									
评价指标	r	0.960	0.916	0.896	0.805	0.778	0.550	0.608	0.420	0.581
	RMSE	0.205	0.300	0.308	0.454	0.536	0.550	0.539	0.626	0.548
	MAE	0.166	0.236	0.251	0.387	0.461	0.419	0.453	0.466	0.424

较小,存在过拟合问题;超前6~9 M的相位和强度与标签值偏差较大,较标签值幅度逐渐变小。总体而言,随着预报时效的增加,相位滞后程度逐渐加大,强度偏差也越来越大。

GBDT模型的订正结果评估如表1所示。超前1~3 M的 r 在0.9左右,超前4~5 M在0.8左右,超前6~7和超前9 M为0.550~0.581;超前1 M的RMSE为0.205,超前2~3 M在0.3左右,超前4 M为0.454,超前5~9 M为0.5~0.6;超前1~3 M的MAE在0.2左右,超前4~9 M在0.4左右。

3.2 基于XGBoost算法的ENSO多模式预测结果订正模型

基于XGBoost的订正模型中各模式结果特征重要性方面(见表2):超前1~2 M的动力模式预测结果平均特征重要性高于统计模式,超前3~9 M的统计模式高于动力模式。

从XGBoost模型的订正结果时间序列图可以看出(见图2),超前1~3 M的相位和强度与标签值基本一致;超前4~5 M的相位和强度与标签值偏差

较小,存在过拟合问题;超前6~9 M的相位和强度与标签值偏差较大,较标签值幅度逐渐变小。总体而言,随着预报时效的增加,相位滞后程度逐渐加大,强度的偏差也越来越大。

XGBoost模型的订正结果评估如表2所示。超前1~3 M的 r 在0.9左右,超前4~5 M在0.8左右,超前6~7 M达0.6以上,超前8~9 M在0.547以上;超前1 M的RMSE为0.248,超前3~5 M为0.3~0.4,超前5~9 M为0.5~0.6;超前1~4 M的MAE在0.2~0.3,超前4 M为0.308,超前5~9 M为0.408~0.451。

3.3 基于lightGBM算法的ENSO多模式预测结果订正模型

基于lightGBM的订正模型中各模式结果特征重要性方面(见表3):超前1~2 M和超前6 M的动力模式预测结果平均特征重要性高于统计模式,超前3~5 M和超前7~9 M的统计模式高于动力模式。

从lightGBM模型的订正结果时间序列图可以看出(见图3),超前1~3 M的相位和强度与标签值基本一致;超前4~7 M的相位和强度与标签值偏差

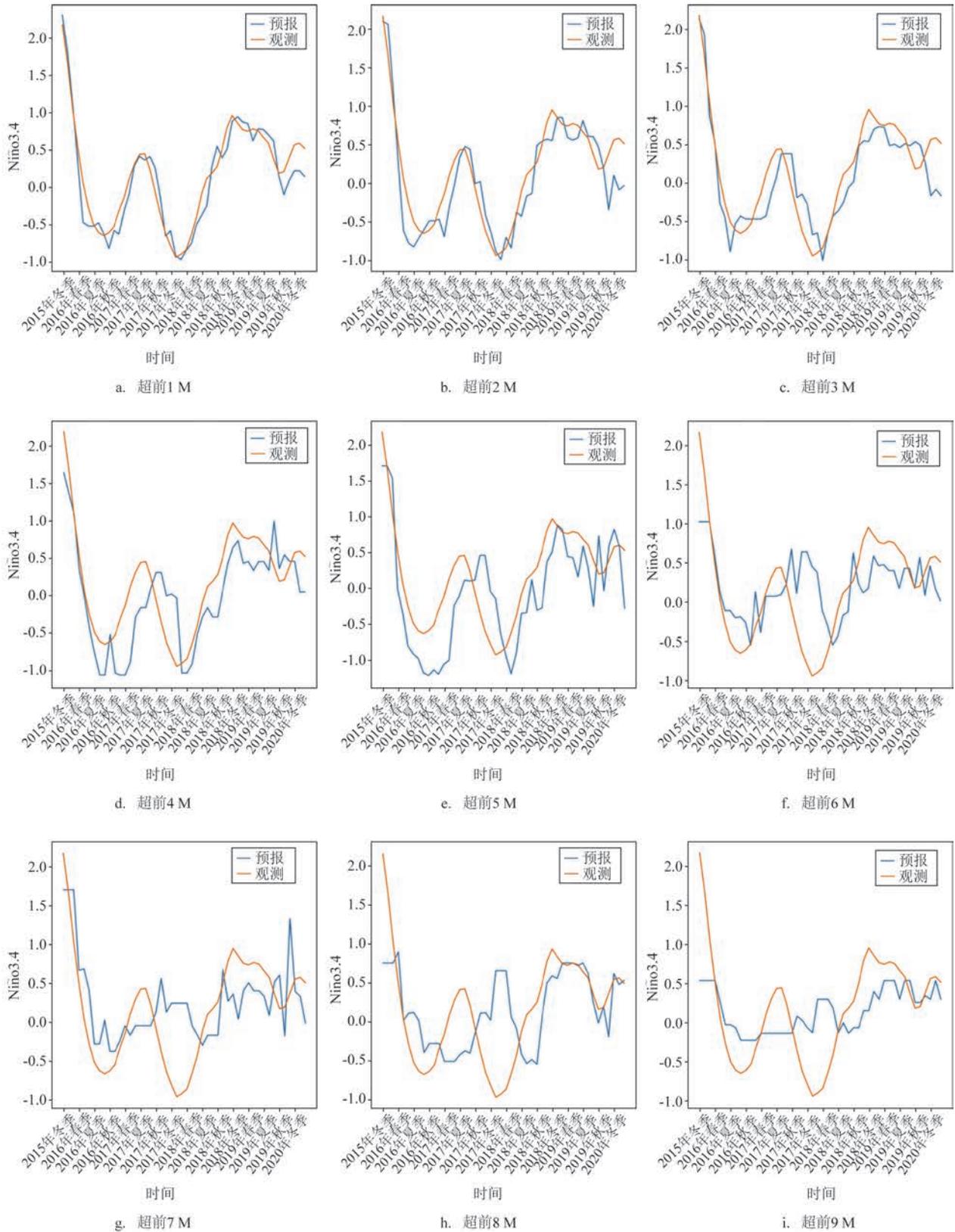


图1 基于GBDT算法的集合订正结果

Fig.1 Correction results based on GBDT algorithm

(注:冬季代表12—2月,春季代表3—5月,夏季代表6—8月,秋季代表9—11月,下同)

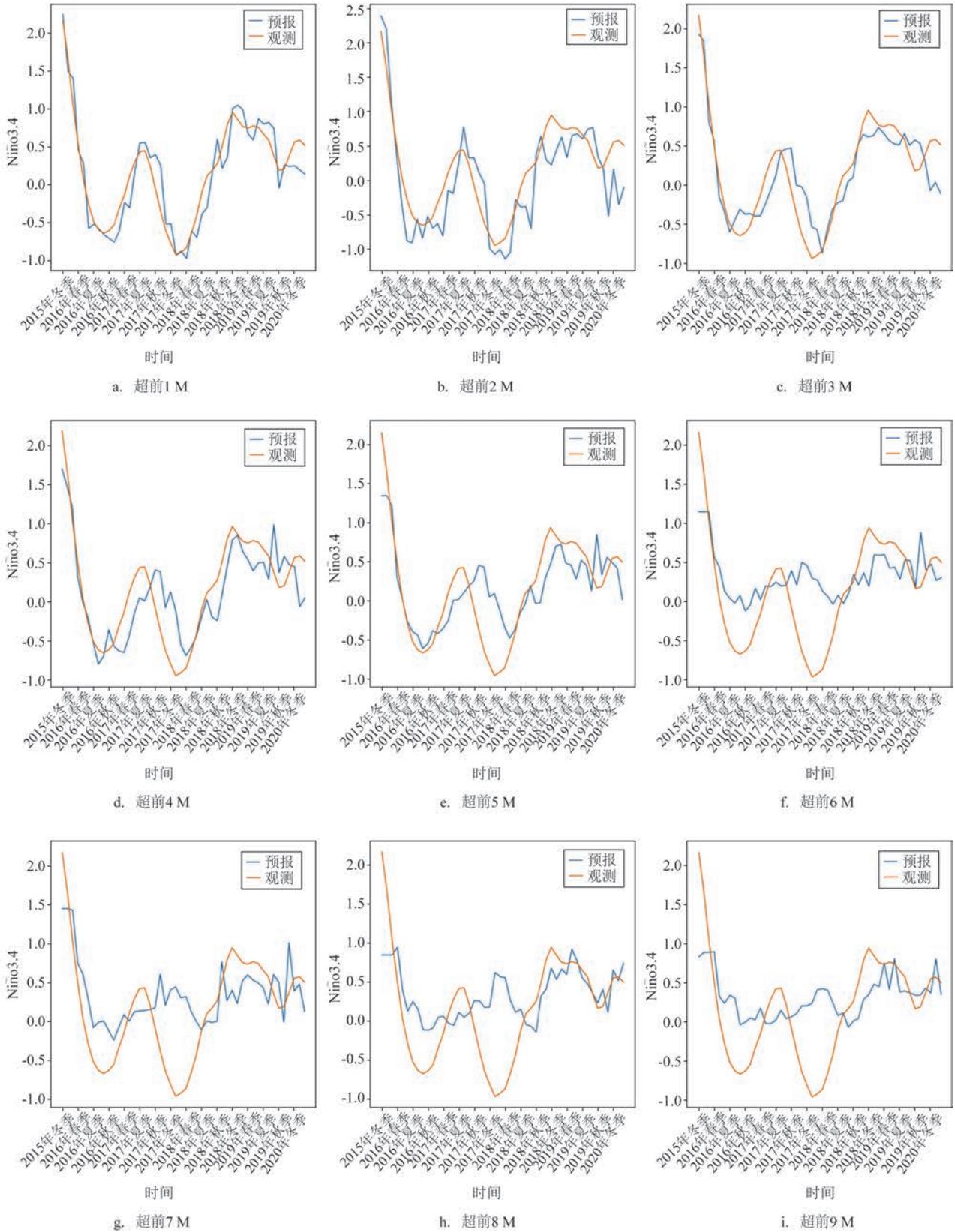


图2 基于XGBoost的集合订正结果

Fig.2 Correction results based on XGBoost algorithm

表 2 XGBoost 订正模型的模式预测结果特征重要性及订正结果评估

Tab.2 Feature importance of the prediction results of XGBoost revised model and evaluation of the revised results of XGBoost revised model

模式		超前 1 M	超前 2 M	超前 3 M	超前 4M	超前 5 M	超前 6 M	超前 7 M	超前 8 M	超前 9 M
动力模式特征重要性	ECMWF	0.593	0.168	0.165	—	—	—	—	—	—
	JMA	0.008	0.342	0.154	0.235	0.302	—	—	—	—
	NASA	0.034	0.118	0.099	0.137	0.154	0.244	0.229	—	—
	GMAO									
	NCEP	0.260	0.015	0.086	0.217	0.194	0.187	0.234	—	—
	CFS									
	LDEO	0.037	0.022	0.071	0.039	0.049	0.074	0.131	—	—
	KMA	0.006	0.074	0.010	0.036	0.036	0.043	0.065	0.185	0.215
统计模式特征重要性	SNU									
	CPC	0.042	0.069	0.124	0.137	0.099	0.146	0.129	0.302	0.329
	MRKOV									
	CPCCA	0.014	0.174	0.222	0.124	0.109	0.241	0.151	0.329	0.255
	CSU	0.006	0.018	0.070	0.075	0.057	0.065	0.061	0.183	0.200
评价指标	CLIPR									
	r	0.937	0.876	0.893	0.829	0.782	0.679	0.627	0.547	0.548
	RMSE	0.248	0.404	0.297	0.373	0.413	0.535	0.541	0.581	0.577
	MAE	0.202	0.330	0.243	0.299	0.308	0.408	0.432	0.432	0.451

表 3 lightGBM 订正模型的模式预测结果特征重要性及订正结果评估

Tab. 3 Feature importance of the prediction model results of lightGBM revised model and evaluation of the revised results of lightGBM revised model

模式		超前 1 M	超前 2 M	超前 3 M	超前 4 M	超前 5 M	超前 6 M	超前 7 M	超前 8 M	超前 9 M
动力模式特征重要性	ECMWF	0.112	0.324	0.238	—	—	—	—	—	—
	JMA	0.098	0.296	0.138	0.286	0.193	—	—	—	—
	NASA	0.119	0.056	0.063	0.163	0.298	0.161	0.183	—	—
	GMAO									
	NCEP	0.150	0.042	0.075	0.102	0.000	0.219	0.238	—	—
	CFS									
	LDEO	0.116	0.014	0.050	0.020	0.035	0.110	0.143	—	—
	KMA	0.087	0.028	0.050	0.041	0.070	0.142	0.063	0.167	0.200
统计模式特征重要性	SNU									
	CPC	0.099	0.155	0.213	0.143	0.175	0.077	0.127	0.333	0.300
	MRKOV									
	CPCCA	0.121	0.042	0.088	0.143	0.123	0.187	0.135	0.500	0.500
	CSU	0.099	0.042	0.088	0.102	0.105	0.103	0.111	0.000	0.000
评价指标	CLIPR									
	r	0.955	0.901	0.884	0.803	0.608	0.699	0.632	0.513	0.449
	RMSE	0.219	0.301	0.317	0.441	0.608	0.522	0.585	0.584	0.614
	MAE	0.190	0.233	0.241	0.369	0.513	0.408	0.477	0.456	0.500

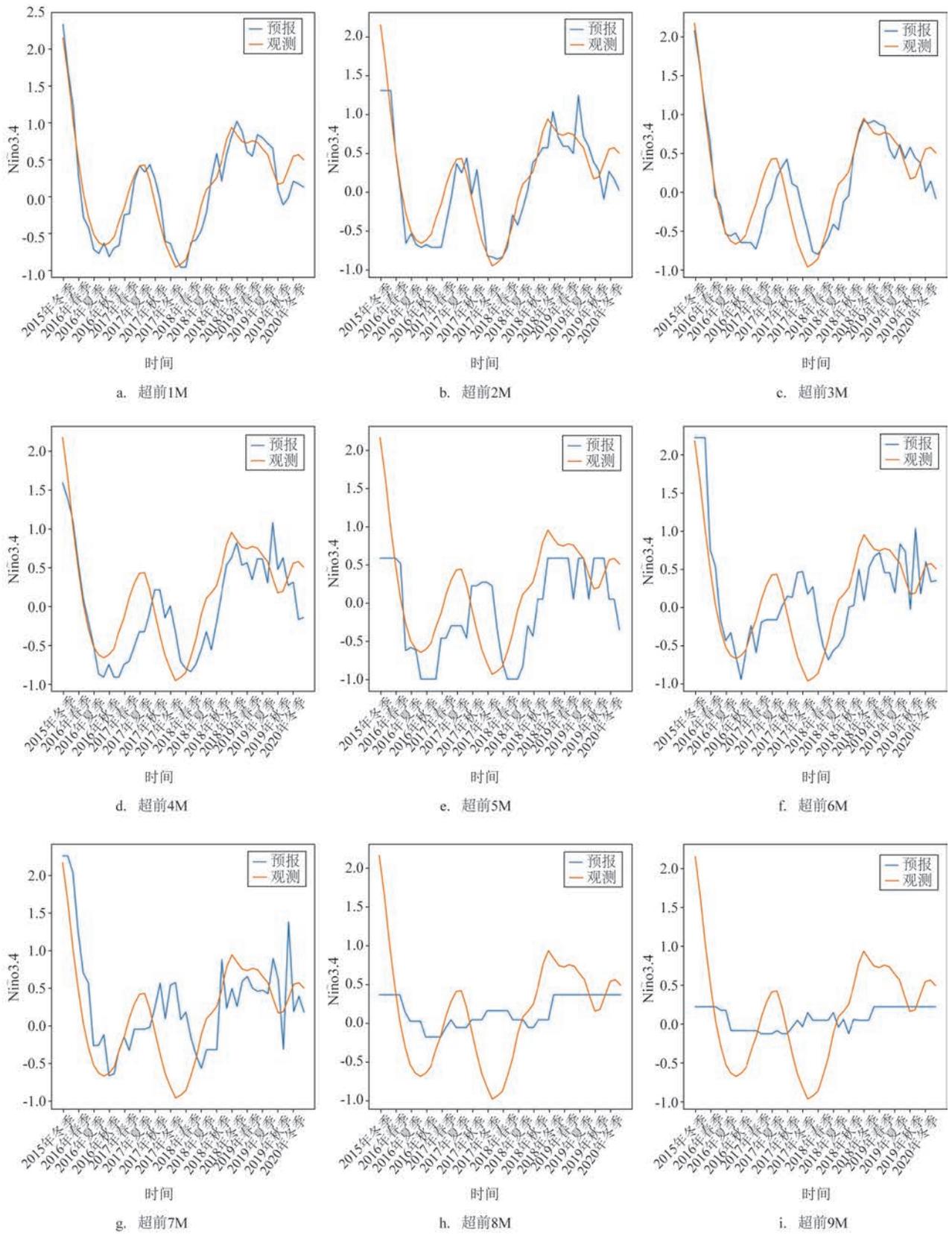


图3 基于lightGBM的集合订正结果
Fig.3 Correction results based on lightGBM algorithm

较小,存在过拟合问题;超前8~9 M的相位和强度与标签值偏差较大,较标签值幅度逐渐变小且趋平。总体而言,随着预报时效的增加,相位滞后程度逐渐加大,强度偏差也越来越大。

lightGBM模型的订正结果评估如表3所示。超前1~3 M的 r 在0.9左右,超前4 M达到0.8,超前5~7 M为0.6~0.7,超前8~9 M为0.5左右;超前1~3 M的RMSE的 r 在0.2~0.3,超前4 M为0.441,超前5~9 M为0.5~0.6;超前1~3 M的MAE在0.2左右,超前4 M为0.369,超前5~9 M为0.4~0.5。

3.4 基于RF算法的ENSO多模式预测结果订正模型

基于RF的订正模型中各模式结果特征重要性方面:超前1~5 M和超前7 M的动力模式预测结果平均特征重要性高于统计模式,超前6 M和超前8~9 M的统计模式高于动力模式。具体而言,超前1~3 M动力模式明显高于统计模式;超前4~7 M动力模式和统计模式基本相当;超前8~9 M动力模式明显低于统计模式。

从RF模型的订正结果时间序列图可以看出

(见图4),超前1~3 M的相位和强度与标签值基本一致;超前4~7 M的相位和强度与标签值偏差较小,存在过拟合问题;超前8~9 M的相位和强度与标签值偏差较大,相较标签值幅度逐渐变小。总体而言,随着预报时效的增加,相位滞后程度逐渐加大,强度的偏差也越来越大。

RF模型的订正结果评估如表4所示。超前1~3 M的 r 在0.9左右,超前4~5 M在0.8左右,超前6 M和超前8~9 M基本达到或超过0.6;超前1~3 M的RMSE为0.2~0.3,超前4~6 M和超前8~9 M在0.55以下;超前1~3 M的MAE在0.2左右,超前4~6 M和超前8~9 M在0.46以下。

3.5 智能会商系统结果分析

4种决策树算法订正结果的时序趋势随着预报时效的增加而趋平,这与输入订正模型的模式预测结果有关。特别地,随着超前预报时效的增加,起报时间逐渐提前,输入特征的极值范围不断缩小。对于2015年冬季发生的超强厄尔尼诺事件,起始时段的订正结果时间序列呈直线状,且强度明显偏

表4 RF订正模型的模式预测结果特征重要性及订正结果评估

Tab. 4 Feature importance of the prediction model results of RF revised model and evaluation of the revised results of

RF revised model

模式		超前1 M	超前2 M	超前3 M	超前4 M	超前5 M	超前6 M	超前7 M	超前8 M	超前9 M
动力模式 特征重要性	ECMWF	0.371	0.674	0.659	—	—	—	—	—	—
	JMA	0.085	0.184	0.217	0.549	0.489	—	—	—	—
	NASA	0.012	0.028	0.015	0.070	0.109	0.245	0.219	—	—
	GMAO									
	NCEP	0.443	0.011	0.025	0.137	0.108	0.263	0.305	—	—
	CFS									
	LDEO	0.011	0.011	0.010	0.020	0.028	0.047	0.113	—	—
	KMA	0.007	0.010	0.011	0.020	0.029	0.030	0.038	0.050	0.050
	SNU									
	统计模式 特征重要性	CPC	0.054	0.048	0.022	0.048	0.088	0.179	0.135	0.310
MRKOV										
CPCCA		0.008	0.017	0.017	0.100	0.097	0.192	0.148	0.640	0.400
CSU		0.010	0.016	0.024	0.056	0.052	0.045	0.043	0.000	0.000
CLIPR										
评价指标	r	0.954	0.914	0.885	0.818	0.795	0.635	0.517	0.598	0.636
	RMSE	0.215	0.293	0.328	0.533	0.549	0.523	0.646	0.533	0.513
	MAE	0.179	0.237	0.268	0.451	0.456	0.399	0.524	0.394	0.428

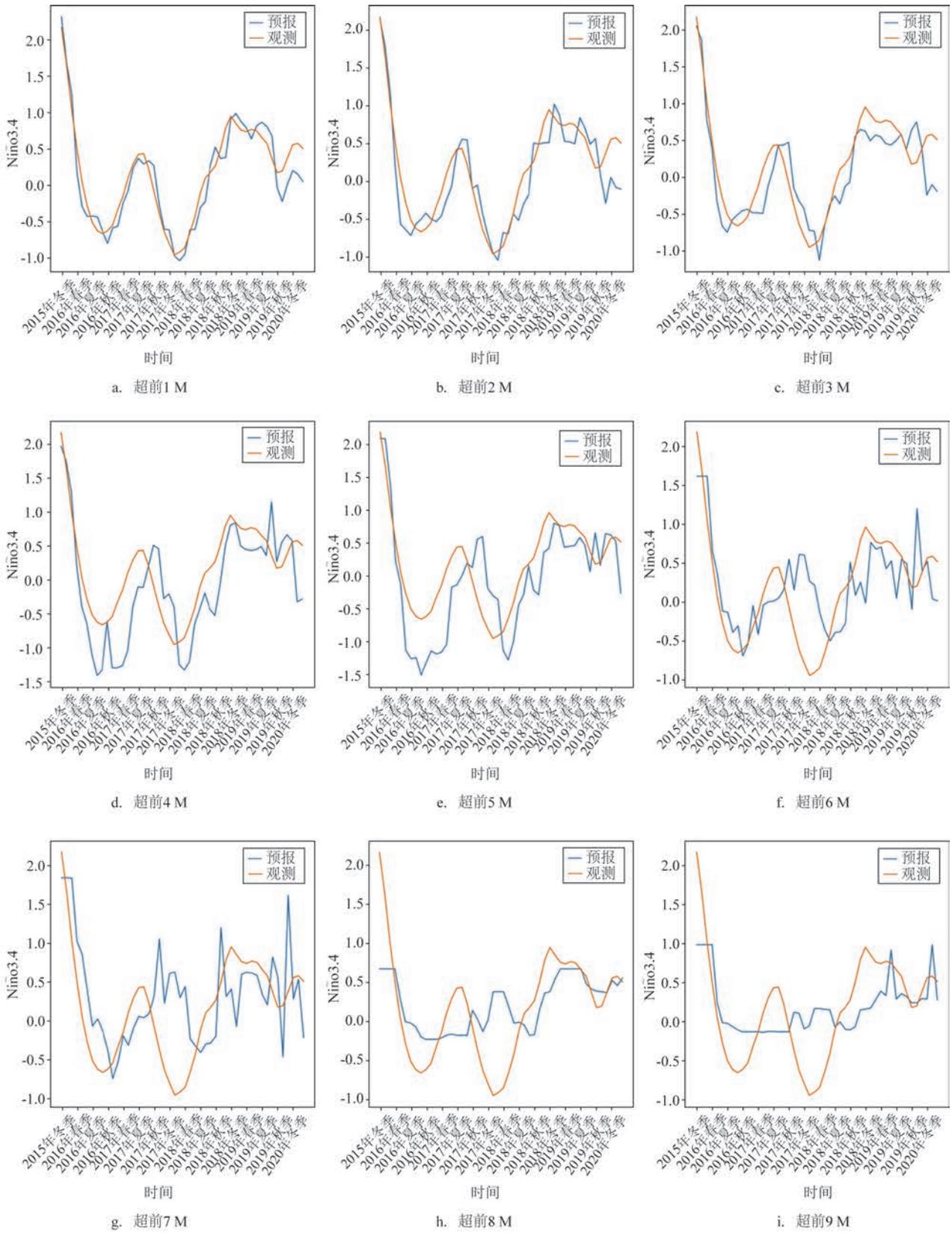


图4 基于RF的集合订正结果

Fig.4 Correction results based on RF algorithm

弱。这跟机器学习模型的预测效果依赖于训练集的丰富程度密切相关。

4种决策树算法在不同超前预报时效上的预报技巧表现不一,这可能与4种算法在数学方法上选用的目标函数和损失函数不同,以及在工程实现方面采用的技术手段不同有关。XGBoost对异常点敏感;GBDT精度更高且灵活性更强,但运算过程消耗内存过大;lightGBM所需内存更小且运算速度更快;RF综合表现良好,相对于其他算法有较大的优势。

随着超前预报时效的增加,预报技巧整体呈下降趋势。但是,4种决策树模型在超前6~9 M预报时效上的预报技巧存在着不同程度的起伏,并不严格随预报时效的增加而降低,这可能与决策树算法的不稳定性有关。

4 总结与展望

本研究主要得到以下结论:

(1)在建模特征重要性方面,不同的决策树模型表现出不同的偏好。整体而言,超前预报时效较短时,偏重于动力模式,较长时,偏重于统计模式。这与在不同超前预报时效上,动力模式和统计模式预报技巧的高低相吻合。

(2)决策树模型同传统的动力模式和统计模式一样,随着预报时效的增加,相位滞后程度逐渐加大,强度的偏差也越来越大,甚至有时间序列趋于平缓的趋势,同时还存在过拟合的问题。不同决策树模型在不同超前预报时效上的预报技巧表现不一,这与机器学习算法的特性有关。

(3)基于4种可解释机器学习算法建立的多模式ENSO预测结果智能会商系统,对多模式结果进行集合订正,预报技巧达到或优于传统的动力模式和统计模式水平。这说明使用机器学习算法进行多模式集合订正具有优势,同Ren等^[42]的调研结果相一致。

此外,由于机器学习算法本身存在一定的缺陷,机器学习回归预测模型的预报技巧依赖于训练数据的丰富程度,对于未经训练的极值,无法给出有效预测结果。例如,对于2015—2016年发生的超强厄尔尼诺事件,该智能会商系统的预测能力有限。

参考文献:

- [1] 李崇银,穆穆,周广庆,等. ENSO机理及其预测研究[J]. 大气科学, 2008, 32(4): 761-781.
Li C Y, Mu M, Zhou G Q, et al. Mechanism and prediction studies of the ENSO[J]. Chinese Journal of Atmospheric Sciences, 2008, 32(4): 761-781.
- [2] 丑纪范. 天气和气候的可预报性[J]. 气象科技进展, 2011, 1(2): 11-14.
Chou J F. Predictability of weather and climate[J]. Advances in Meteorological Science and Technology, 2011, 1(2): 11-14.
- [3] 陶祖钰,赵翠光,陈敏. 谈谈统计预报的必要性[J]. 气象科技进展, 2016, 6(1): 6-13.
Tao Z Y, Zhao C G, Chen M. The necessity of statistical forecasts [J]. Advances in Meteorological Science and Technology, 2016, 6 (1): 6-13.
- [4] 许小峰. 从物理模型到智能分析——降低天气预报不确定性的新探索[J]. 气象, 2018, 44(3): 341-350.
Xu X F. From physical model to intelligent analysis: a new exploration to reduce the uncertainty of weather forecast[J]. Meteorological Monthly, 2018, 44(3): 341-350.
- [5] Jin E K, Kinter III J L, Wang B, et al. Current status of ENSO prediction skill in coupled ocean - atmosphere models[J]. Climate Dynamics, 2008, 31(6): 647-664.
- [6] 任福民,袁媛,孙丞虎,等. 近30年ENSO研究进展回顾[J]. 气象科技进展, 2012, 2(3): 17-24.
Ren F M, Yuan Y, Sun C H, et al. Review of progress of ENSO studies in the past three decades[J]. Advances in Meteorological Science and Technology, 2012, 2(3): 17-24.
- [7] Clarke A J. El Niño physics and El Niño predictability[J]. Annual Review of Marine Science, 2014, 6: 79-99.
- [8] 陈大可,连涛. 厄尔尼诺-南方涛动研究新进展[J]. 科学通报, 2020, 65(35): 4001-4003.
Chen D K, Lian T. Frontier of El Niño-southern oscillation research [J]. Chinese Science Bulletin, 2020, 65(35): 4001-4003.
- [9] Barnston A G, Tippett M K, L'Heureux M L, et al. Skill of real-time seasonal ENSO model predictions during 2002-11 — is our capability increasing?[J]. Bulletin of the American Meteorological Society, 2012, 93(5): 631-651.
- [10] 陈静,陈德辉,颜宏. 集合数值预报发展与研究进展[J]. 应用气象学报, 2002, 13(4): 497-507.
Chen J, Chen D H, Yan H. A brief review on the development of ensemble prediction system[J]. Journal of Applied Meteorological Science, 2002, 13(4): 497-507.
- [11] Mylne K R, Evans R E, Clark R T. Multi-model multi-analysis ensembles in quasi-operational medium-range forecasting[J]. Quarterly Journal of the Royal Meteorological Society, 2002, 128 (579): 361-384.
- [12] 郑飞. ENSO集合预报研究[D]. 北京: 中国科学院研究生院(大

- 气物理研究所), 2007.
- Zheng F. Research on ENSO ensemble prediction[D]. Beijing: Institute of Atmospheric Physics, Chinese Academy of Sciences, 2007.
- [13] Yan X Q, Tang Y M. An analysis of multi-model ensembles for seasonal climate predictions[J]. Quarterly Journal of the Royal Meteorological Society, 2013, 139(674): 1179-1198.
- [14] 王琳. ENSO的多模式集合预报研究[D]. 成都: 成都信息工程大学, 2017.
- Wang L. Study on ENSO multi-model ensemble predictions with the statistical correction method[D]. Chengdu: Chengdu University of Information Technology, 2017.
- [15] 郭炜豪, 温文, 王晓春, 等. NINO3.4指数的多模式集合预报方法[J]. 热带气象学报, 2019, 35(2): 262-267.
- Guo W H, Wen W, Wang X C, et al. A multimodel ensemble method for NINO3.4 index forecast[J]. Journal of Tropical Meteorology, 2019, 35(2): 262-267.
- [16] 王琳. 基于多模式集合的季节-年际气候预测方法研究[D]. 武汉: 中国地质大学, 2020.
- Wang L. Study on seasonal to interannual climate prediction with multi-model ensemble method[D]. Wuhan: China University of Geosciences, 2020.
- [17] 陈溢豪, 张蕴斐, 周倩, 等. 集合预报及其在季节尺度气候预测中的应用[J]. 海洋预报, 2020, 37(6): 102-111.
- Chen Y H, Zhang Y F, Zhou Q, et al. Ensemble forecasting and its application in seasonal climate forecast[J]. Marine Forecasts, 2020, 37(6): 102-111.
- [18] 焦李成, 杨淑媛, 刘芳, 等. 神经网络七十年: 回顾与展望[J]. 计算机学报, 2016, 39(8): 1697-1717.
- Jiao L C, Yang S Y, Liu F, et al. Seventy years beyond neural networks: retrospect and prospect[J]. Chinese Journal of Computers, 2016, 39(8): 1697-1717.
- [19] 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.
- [20] Sebastian S, Gabriele M. Weather and climate forecasting with neural networks: using general circulation models (GCMs) with different complexity as a study ground[J]. Geoscientific Model Development, 2019, 12(7): 2797-2809.
- [21] Ruti P M, Tarasova O, Keller J H, et al. Advancing research for seamless Earth system prediction[J]. Bulletin of the American Meteorological Society, 2020, 101(1): E23-E35.
- [22] Wheeling K. Machine learning improves weather and climate models[J]. Eos, 2020, 101, doi: 10.1029/2020EO142422.
- [23] De Vos M G, Hazeleger W, Bari D, et al. Open weather and climate science in the digital era[J]. Geoscience Communication, 2020, 3(2): 191-201.
- [24] Bauer P, Dueben P D, Hoefler T, et al. The digital revolution of Earth-system science[J]. Nature Computational Science, 2021, 1(2): 104-113.
- [25] 贺圣平, 王会军, 李华, 等. 机器学习的原理及其在气候预测中的潜在应用[J]. 大气科学学报, 2021, 44(1): 26-38.
- He S P, Wang H J, Li H, et al. Machine learning and its potential application to climate prediction[J]. Transactions of Atmospheric Sciences, 2021, 44(1): 26-38.
- [26] Ham Y G, Kim J H, Luo J J. Deep learning for multi-year ENSO forecasts[J]. Nature, 2019, 573(7775): 568-572.
- [27] Zhang Q, Wang H, Dong J Y, et al. Prediction of sea surface temperature using long short-term memory[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(10): 1745-1749.
- [28] Aguilar-Martinez S, Hsieh W W. Forecasts of Tropical Pacific Sea Surface temperatures by neural networks and support vector regression[J]. International Journal of Oceanography, 2009, 2009: 167239.
- [29] Feng Q Y, Vasile R, Segond M, et al. ClimateLearn: a machine-learning approach for climate prediction using network measures [J]. Geoscientific Model Development Discussions, 2016, doi: 10.5194/gmd-2015-273.
- [30] Nooteboom P D, Feng Q Y, López C, et al. Using network theory and machine learning to predict El Niño[J]. Earth System Dynamics Discussions, 2018, doi: 10.5194/esd-2018-13.
- [31] 许柏宁, 姜金荣, 郝卉群, 等. 一种基于区域海表面温度异常预测的 ENSO 预报深度学习模型[J]. 科研信息化技术与应用, 2017, 8(6): 65-76.
- Xu B N, Jiang J R, Hao H Q, et al. A deep learning model of ENSO prediction based on regional sea surface temperature anomaly prediction[J]. e-Science Technology & Application, 2017, 8(6): 65-76.
- [32] 何丹丹, 姜金荣, 郝卉群, 等. 基于深度学习的 ENSO 预报方法研究[J]. 科研信息化技术与应用, 2019, 10(1): 38-47.
- He D D, Jiang J R, Hao H Q, et al. Research on ENSO forecasting method based on deep learning[J]. e-science Technology & Application, 2019, 10(1): 38-47.
- [33] 蒋国荣, 张韧, 沙文钰. 用 EOF 展开和人工神经网络方法预测 ENSO 的研究[J]. 海洋预报, 2001, 18(3): 1-11.
- Jiang G R, Zhang R, Sha W Y, et al. The study of forecasting ENSO by using EOF approach and neural network method[J]. Marine Forecasts, 2001, 18(3): 1-11.
- [34] Dijkstra H A, Petersik P, Hernández-García E, et al. The application of machine learning techniques to improve El Niño prediction skill[J]. Frontiers in Physics, 2019, 7: 153.
- [35] Breiman L. Random forests[J]. Machine Learning, 2001, 45(1): 5-32.
- [36] Quinlan R J. Induction of decision trees[J]. Machine Learning, 1986, 1(1): 81-106.
- [37] Dietterich T G. An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization[J]. Machine Learning, 2000, 40(2): 139-157.

- [38] Friedman J H. Greedy function approximation: a gradient boosting machine [J]. *The Annals of Statistics*, 2001, 29(5): 1189-1232.
- [39] Chen T Q, Guestrin C. XGBoost: a scalable tree boosting system [C]//*Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco, California, USA: ACM, 2016.
- [40] Ke G L, Meng Q, Finley T, et al. LightGBM: a highly efficient gradient boosting decision tree[C]//*Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach, CA, USA: ACM, 2017: 3149-3157.
- [41] Liaw A, Wiener M. Classification and Regression by random Forest [J]. *R News*, 2002, 2(3): 18-22.
- [42] Ren H L, Zheng F, Luo J J, et al. A review of research on tropical Air-Sea interaction, ENSO dynamics, and ENSO prediction in China[J]. *Journal of Meteorological Research*, 2020, 34(1): 43 - 62.

The application of machine learning in ENSO prediction consultation

LI Chentong

(National Marine Environmental Forecasting Center, Beijing 100081 China)

Abstract: Banned on the concept of multi-model ensemble forecasting, this study establishes an intelligent consultation system for multi-model intelligent consultation system of ENSO prediction using the interpretable machine learning method named decision tree algorithm. The hyper parameters of four decision tree models of GBDT based on Boosting, XGBoost, lightGBM and Random Forest (RF) based on Bagging are optimized and adjusted by using two hyper parameter adjustment methods of random search cross-validation and grid search cross-validation. The intelligent consultation system of multi-model ENSO prediction results is established according to different prediction leading time, which makes integrated correction on the multi-model prediction results and provides the feature importance of the prediction result of each model in the intelligent consultation system. The intelligent consultation system simulates the consultation process of ENSO prediction, which realizes the automation of the processes of reading the prediction results of each model, training the model, giving the prediction conclusion and prediction basis and the visualization of the prediction results, and realizes the function of intelligent parameter tuning. The intelligent consultation system collectively revises the multi-modal ENSO prediction results. The results show that machine learning also has some advantages in the multi-modal result consultation, which provide a reference for consultations of ENSO prediction in the future.

Key words: ENSO; interpretable machine learning; multi-model; intelligent consultation