

# 基于多种机器学习方法订正大城市温度精细化预报试验\*

刘嘉慧敏<sup>1,2</sup> 赵声蓉<sup>3</sup> 林建<sup>3</sup> 唐健<sup>3</sup> 王青霞<sup>4</sup> 尚可<sup>5</sup>

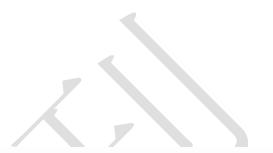
1 陕西省气象台, 西安 710014

2 秦岭和黄土高原生态环境气象重点实验室, 西安 710016

3 国家气象中心, 北京 100081

4 湖南省气象台, 长沙 410118

5 西安航空学院, 西安 710077



**提 要:** 利用 2021—2024 年欧洲中期天气预报中心(ECMWF)模式预报产品及西安城区站点 2m 温度实况, 针对西安关键点温度预报构建了多模型集成预报订正流程及方案。2021 年 9 月 1 日至 2023 年 12 月 31 日数据作为训练集, 用于因子筛选、参数调优与模型集成, 2024 年 1 月 1 日至 4 月 30 日数据作为测试集, 用于评估数值模式及不同训练方案下模型的预报性能。通过主观经验筛选与时滞相关分析, 优选了与温度变化密切相关的 7 个模式预报物理量以及不同时效高空关键区变量等特征因子, 采用 XGBoost、LightGBM、CatBoost 三种梯度提升树模型进行单模型偏差订正, 最终通过 Stacking 集成实现模型融合优化。结果表明: ECMWF 模式对西安温度预报存在系统性冷偏差, 夜间误差显著大于白天, 降温降水过程中冷偏差加剧。三种机器学习模型经贝叶斯优化调参后(均方根误差分别降低 0.039°C、0.030°C、0.027°C)均能有效订正模式偏差。优选特征因子后, 单模型均方根误差平均降低约 0.25°C。Stacking 集成较传统加权集成表现更优, 集成后均方根误差降低 0.034°C, 2°C 内预报准确率提升 3.1%, 在明显降温降水天气过程中均方根误差较单模型最大减小 0.491°C。

**关键词:** 气象保障, 温度预报, 机器学习, 集成订正

**中图分类号:** P456, P457      **文献标志码:** A

**DOI:** 10.7519/j.issn.1000-0526.2025.072902

## Experimental Study on the Temperature Refined Forecasting in Large

### City Based on Multiple Machine Learning Methods

LIU Jiahuimin<sup>1,2</sup> ZHAO Shengrong<sup>3</sup> LIN Jian<sup>3</sup> TANG Jian<sup>3</sup> WANG Qingxia<sup>4</sup> SHANG Ke<sup>5</sup>

1 Shaanxi Meteorological Observatory, Xi'an 710014

2 Key Laboratory of Eco-Environment and Meteorology for the Qinling Mountains and Loess Plateau, Xi'an 710016

3 National Meteorological Centre , Beijing 100081

\* 中国气象局复盘总结专项(FPZJ2024-131)、中国气象局创新发展专项(CXFZ2025Q020)、秦岭和黄土高原生态环境气象重点实验室重点课题(2023K-2)、陕西省科技计划项目(2025JC-YBMS-331)、国家自然科学基金(42205037)共同资助

第一作者: 刘嘉慧敏, 主要从事数值预报模式客观订正预报研究. E-mail:857392276@qq.com

通讯作者: 赵声蓉, 主要从事数值预报产品释用方法研究. E-mail:zhaosr@cma.gov.cn

35

36 **Key words:** meteorological support, temperature forecast, machine learning, ensemble correction  
37 A multi-model integrated forecast correction process and scheme are constructed for 2 m  
38 temperature forecasts at Xi'an Station based on the 2021-2024 European Centre for  
39 Medium-Range Weather Forecasts (ECMWF) model forecasts and the 2 m temperature  
40 observations from Xi'an Station. The data from 1 September 2021 to 31 December 2023 is used  
41 as the training set for factor screening, parameter tuning, and model ensemble, while the data from  
42 1 January to 30 April 2024 are taken as the test set to assess the forecast performance of numerical  
43 models and models trained under different schemes. Through subjective experience screening and  
44 time-lag correlation analysis, seven model prediction physical variables closely related to  
45 temperature changes, as well as different lead-time high-altitude key zone variables and other  
46 characteristic factors get optimized. XGBoost, LightGBM and CatBoost are used for single model  
47 bias correction, and finally model fusion optimization is achieved through Stacking ensemble. The  
48 results show that ECMWF model exhibits a systematic cold bias in temperature forecasts at Xi'an  
49 Station, with the error being significantly greater at night than during the day and the cold bias  
50 intensifying during cooling and precipitation processes. After Bayesian optimization and  
51 parameter tuning, all the three machine learning models are able to effectively correct mode bias  
52 with root mean square errors (RMSE) reduced by 0.039°C, 0.030°C, and 0.027°C, respectively.  
53 Subsequent feature factor optimization further improves the single-model forecast accuracy by  
54 approximately 0.25pi. The Stacking ensemble surpasses the traditional weighted ensemble. After  
55 ensemble, the RMSE of temperature forecasts is reduced by 0.034°C, and the forecast accuracy  
56 within 2°C is improved by 3.1%. During the significant cooling and precipitation process, the  
57 forecast RMSE has a maximum reduction of 0.491°C compared to that by the single-factor  
58 model.

59 **Key words:** meteorological support, temperature forecast, machine learning, ensemble correction

## 60 引 言

61 随着社会经济的发展和城市化进程的加快,大城市举办的外事活动、国际会议、体育赛  
62 事等重大活动呈现出大型化、室外化的特点(唐钧等, 2021; 轩春怡等, 2022; 漆梁波, 2025),  
63 这些重大活动社会关注度较高,受天气因素制约大,因此气象保障成为大城市重大活动组织  
64 实施和运行体系中必不可少的重要部分(甘璐等, 2021)。西安近年来承办了如 2021 年第十  
65 四届全国运动会, 2023 年 5 月中国-中亚峰会等高规格的重大活动(刘华等, 2023)。这些重  
66 大活动室外保障点多,气象保障要求高精度、高响应,给大城市气象保障带来巨大挑战。因  
67 此,开展大城市气象要素客观预报技术研究与应用,为重大活动气象保障提供精细化的客观  
68 预报产品支撑具有迫切需求。

69 温度与降水的精细化预报对重大活动气象保障同样至关重要。一方面对于滑雪、马拉松、  
70 马术和山地自行车等运动室外活动,温度的急剧变化会直接影响活动的顺利举行(李嘉睿,  
71 等,2022; 王在文等, 2023; 张芳等, 2024);另一方面极端温度会影响设备性能,同样会对  
72 活动造成重大影响。然而,由于初值场的不确定性、物理过程参数化方案的局限性及大气混  
73 沌性等诸多原因,即便是最先进的气象预报数值模式,其预报精度仍然有限,温度预报与实  
74 际观测值间会产生明显偏差(李妮娜等, 2024; 沈学顺等, 2025),特别是在复杂地形区,随  
75 着地形高度增加,模式温度预报偏差一般呈增大趋势(智协飞等, 2019; 秦庆昌等, 2022)。

76 例如, 北京冬奥会期间, ECMWF、中国气象局区域数值天气预报(CMA-MESO)温度预报在  
77 京津冀地区北部山区和平原地区均存在明显暖偏差, 预报较实况均偏高 4℃以上(佟华等,  
78 2022)。因此, 模式直接的温度预报必须经过偏差订正才能满足重大活动气象保障需求, 亟  
79 需研发更精准的订正方法以提升预报性能。

80 MOS 方法(model output statistics, Glahn and Lowry, 1972)在模式要素预报偏差订正中  
81 得到广泛应用, 其最常用的统计方法是线性回归方法, 包括一元线性回归、多元线性回归、  
82 逐步回归。同时, 在经典的 MOS 方法基础上还进一步发展了最优子集回归、递减回归、准  
83 对称滑动回归等一系列线性统计后处理方法(钱莉等, 2010; Cui et al, 2012; 吴启树等, 2016;  
84 曾晓青等, 2019)。上述方法均是通过建立预报对象(如温度)和多个预报因子(模式输出的温、  
85 压、湿、风等要素场)之间的定量线性关系, 对原始数值预报结果进行统计订正, 适用于在  
86 一定时间和空间范围内气温连续变化与预报因子呈线性相关的场景, 能够有效订正模式预报  
87 的系统性偏差。近年来随着计算技术快速发展, 神经网络、随机森林等机器学习方法在气象  
88 要素预报偏差订正中发挥了较大作用, 订正效果较 MOS 等传统统计方法有一定优势(朱育雷  
89 等, 2024)。刘杰等(2024)利用随机森林算法订正 ECMWF 模式 72h 内温度预报, 并将其与  
90 传统 MOS 订正方法和主观预报产品进行对比分析发现, 随机森林对最高温度预报改进效果  
91 明显。XGBoost(extreme gradient boosting)、LightGBM(light gradient boosting machine)和  
92 CatBoost(categorical boosting)作为高效的梯度提升树机器学习算法, 可以将多个弱学习器组  
93 合成一个强学习器, 通过不断地训练和调整多个弱学习器, 然后将其组合起来, 以提高整体  
94 的预测精度(Chen and Guestrin, 2016; Cui et al, 2021; 杨璐等, 2021)。在要素偏差订正中,  
95 XGBoost 可以通过自定义损失函数和正则化项, 提高预报的准确性和稳定性; LightGBM 通  
96 过直方图优化策略, 可以显著提高训练速度和预报精度; CatBoost 可以自动处理类别特征,  
97 减少数据预处理的复杂性。徐景峰等(2023)针对冬奥会复杂山地百米尺度 10m 风速预报采用  
98 XGBoost 设计订正试验, 发现根据风速等级表归类, 针对每个分类单独构建 XGBoost 模型,  
99 每个区间模型合并后形成的 L-XGBoost 较原始风速预报误差最大可减少 73.28%。将  
100 LightGBM 框架应用于温度精细化预报订正算法中, 对雨雪天气中温度预报质量有明显改进  
101 (谭江红等, 2018)。王珊珊等(2022)发现基于 CatBoost 的长江中游降水相态预报模型对雨、  
102 雪、冻雨有较好的分类和预报效果。

103 相比于单模型, 多模型集成不仅可以发挥各模型的预报优势, 且不会因某个模型性能调  
104 整而导致最终结果出现较大变动。常用多模型集成技术主要有集合平均法、权重分配法、  
105 BP 神经网络等(赵声蓉, 2006; 智协飞等, 2015)。相较传统集成方法, 机器学习 Stacking  
106 模型将多个单模型堆叠在一起, 利用预测结果作为新特征来训练一个新的模型, 其通过多层  
107 学习结构和模型多样性的结合, 动态调整基础模型和元模型的参数, 提高模型的适应性  
108 (Wolpert, 1992; 苏刚等, 2021)。韩念霏等(2022)基于 Stacking 模型构建了适合于误差分析  
109 的集成学习订正模型, 发现该方法对京津冀地区 4 个季节的风温湿要素预报效果改善最明显。

110 为提高重大活动大城市温度精细化预报服务能力, 采用 XGBoost、LightGBM、CatBoost  
111 等 3 种梯度提升树机器学习算法, 结合主观先验知识, 利用时滞相关分析法挖掘不同预报时  
112 效的高空区域特征因子构建西安大城市关键点 2m 温度预报订正模型, 再基于机器学习  
113 Stacking 集成方法融合单模型的预报优势, 探索机器学习方法在大城市温度预报误差订正方  
114 面的作用和效果, 为业务应用提供参考。

115

## 116 1 数据集构建

117 实况数据为西安城区方新气象观测站(以下简称西安站)地面 2m 温度观测资料, 该站  
118 是西安重大活动气象保障的关键站点。模式数据使用 ECMWF 模式地面和高空预报产品,

119 空间分辨率分别为  $0.125^{\circ} \times 0.125^{\circ}$ 、 $0.25^{\circ} \times 0.25^{\circ}$ ，时间分辨率为 3h。本文预报起报时间为 00:00  
120 (世界时，下同)。由于 ECMWF 数据获取时间滞后于模式起报时间，故将前一日 ECMWF12:00  
121 起报的 12~36h 预报产品截断 12h 作为当日 00:00 起报的 0~24h 预报。文中温度若无特殊  
122 说明均指 2m 温度。

123 对于 ECMWF 预报数据，当同一起报时次同一预报要素预报时效缺失达到或超过 4 个、  
124 第一时效或最后一个时效缺测时，则认为该起报时次缺测，否则缺测时效以相邻预报时效数  
125 据插值代替。若一条样本的预报或实况数据之一缺测则剔除该样本。

126 由于建模所需的特征因子来自于模式不同类别的预报变量，存在量级和量纲差异，因此  
127 在训练模型前，需要对原始预报变量进行标准化处理，便于不同量级或量纲的特征值进行比  
128 较，以提升模型收敛速度和精度。本文使用数据标准化对原始变量进行变换，将其变换到均  
129 值为 0，标准差为 1 的范围内：

130 
$$z_i = \frac{x_i - \mu}{\sigma} \quad (1)$$

131 式(1)中， $z_i$  为预处理后的数据， $x_i$  为原始数据， $\mu$  为平均值， $\sigma$  为标准差。

132 本文使用 2021 年 9 月 1 日至 2023 年 12 月 31 日数据作为训练集，用于模型训练，剔  
133 除无效样本后得到训练样本 66172 组。由于 1—4 月为冬季向春季的过渡阶段，冷空气活动  
134 频繁且强度多变。此阶段气温波动剧烈，数值模式对气温的转折点及升降幅度的预报偏差较  
135 大。因此，选取 2024 年 1 月 1 日至 4 月 30 日数据作为测试集，用于模型预报性能测试及评  
136 估，测试样本 7533 组。为提高模型的泛化能力，通过交叉验证方法选择 10 次训练中误差最  
137 小的参数作为模型最优参数(姜红等，2021)。

## 138 2 ECMWF 模式预报评估

139 在评估不同预报订正模型性能之前，首先检验 ECMWF 模式对西安站温度的预报性能，  
140 分析模式温度预报的整体效果及在明显天气过程中的温度预报表现，作为模型对比评估的基  
141 础。

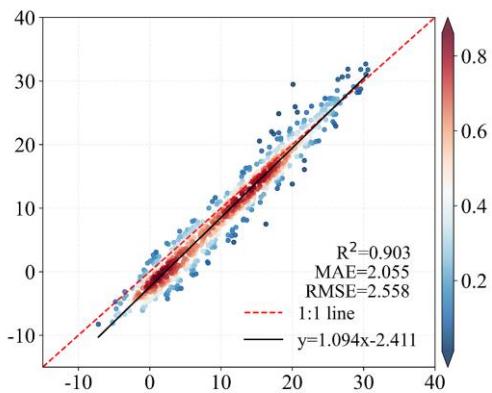
### 142 2.1 评估方法

143 评估方法适用于数值模式评估和订正模型评估。采用平均绝对误差(MAE)、均方根误差  
144 (RMSE)、 $2^{\circ}\text{C}$  以内预报准确率(AUC)和决定系数( $R^2$ )等指标(贾俊平，2019；薛谌彬等，2019；  
145 闫文杰等，2022)对模型具体预报效果进行评估。平均绝对误差(MAE)是预报与实况值偏差  
146 的绝对值的平均，均方根误差(RMSE)是检验模式与实况的离散程度。 $2^{\circ}\text{C}$  以内预报准确率  
147 (AUC)是计算两组数据的误差绝对值小于  $2^{\circ}\text{C}$  的比例，检验模式与实况资料的准确度。决定  
148 系数( $R^2$ )表示回归模型可以解释因变量的方差占比，数值越接近 1，拟合越好。

149

### 150 2.2 整体评估

151 从西安站逐 3h 温度实况与 ECMWF 温度预报的散点回归分布及时序演变可见(图 1)，  
152 模式温度预报趋势整体上与实况变化较为一致， $R^2$  为 0.903，RMSE 和 MAE 分别为  $2.558^{\circ}\text{C}$ 、  
153  $2.055^{\circ}\text{C}$ 。散点密度大值区有 2 个区间，即  $10\sim20^{\circ}\text{C}$ 、 $-2\sim8^{\circ}\text{C}$ ，在这 2 个区间内预报较实况  
154 偏低。进一步对比 ECMWF 模式各时效温度预报与实况的差异(图 2)，发现模式温度预报误  
155 差存在白天小夜间大的日变化特征。3h 时效模式温度预报表现最好，RMSE 和 AUC 分别为  
156  $1.370^{\circ}\text{C}$ 、0.860(图略)。18h 时效模式预报偏差最大，RMSE 达  $3.206^{\circ}\text{C}$ ，AUC 仅为 0.355(图  
157 2d)。



注：散点颜色代表其核密度，红色表示在对应区间数据分布密度大，蓝色则相反，下同；黑色线为实况与预报拟合线。

图 1 2024 年 1 月 1 日至 4 月 30 日西安站温度实况和 ECMWF 预报散点回归

Fig.1 Scatter regression of temperature observations at Xi'an Station and forecasts by ECMWF from 1 January to 30 April 2024

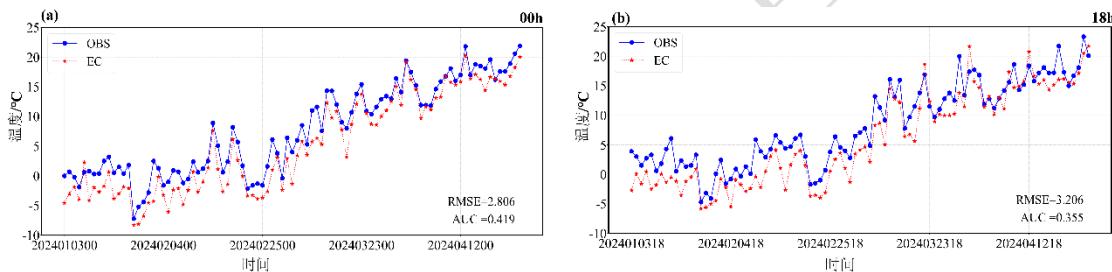


图 2 2024 年 1 月 1 日至 4 月 30 日 (a) 00:00 和 (b) 18:00 西安站温度实况与 ECMWF 预报的时间序列

Fig.2 Time series of temperature observations at Xi'an Station and forecasts by ECMWF from 1 January to 30 April 2024

### 2.3 个例表现

选取 2024 年 2 月 15—16 日降温过程(简称 0215 过程)和 3 月 9—11 日降水过程(简称 0309 过程)分析 ECMWF 模式预报在典型天气过程中的表现。前一次过程受 500hPa 冷涡东移南下影响，西安站日平均气温最大降幅达 5℃ 以上，后一次过程受短波槽东移影响，西安城区出现了明显降水过程，降水持续近 6 h，西安站日平均气温最大降幅达 5.2℃。

从 2 次过程西安站逐 3h 气温与 ECMWF 预报的对比演变来看(图 3)，模式对温度降幅存在较大的预报偏差。0215 过程中，降温阶段模式温度预报表现为明显冷偏差。特别是在夜间偏差更明显，2 月 15 日 18:00 温度预报较实况偏低 8.5℃(图 3a)。该过程 ECMWF 预报 RMSE 达 3.693℃，AUC 仅为 0.333。在 0309 过程中，ECMWF 预报偏差较 0215 过程有所减小，但仍表现为冷偏差，偏差最大达 6.5℃(图 3b)，整个过程 ECMWF 预报 RMSE 和 AUC 分别为 3.102℃、0.407。

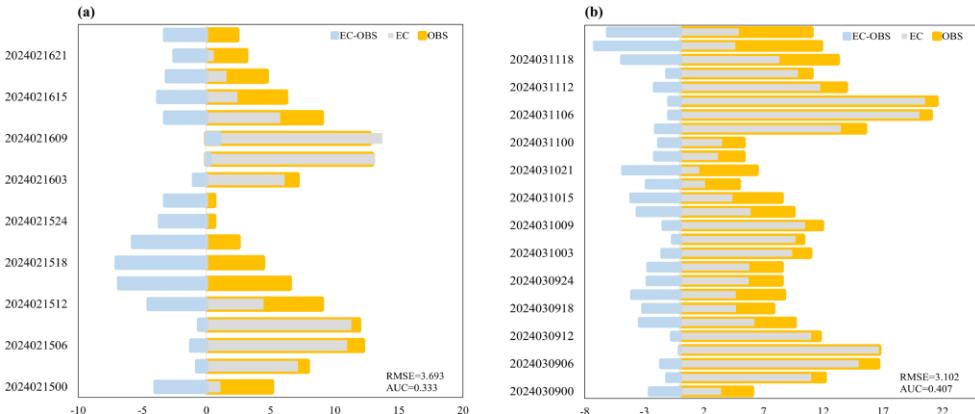


图3 2024年(a)2月15—16日、(b)2024年3月9—11日西安站温度实况与ECMWF预报的时间序列

Fig.3 Time series of temperature observations at Xi'an Station and forecasts by ECMWF during (a) 15–16 February and (b) 9–11 March 2024

由上述分析可见，尽管ECMWF温度预报整体上与西安站实况演变趋势一致，但在夜间预报误差较白天偏大。特别是在明显天气过程中，模式温度预报不能较好地反映关键点温度的降幅，存在显著冷偏差。

### 3 模型训练

在评估ECMWF模式的基础上，尝试利用XGBoost(简称XGB)、LightGBM(简称LGB)和CatBoost(简称CAT)三种机器学习算法对模式预报进行订正，以提高西安大城市关键点温度预报效果。XGB是一种集成学习决策树模型，是基于梯度提升树算法的树结构增强模型(杨璐等，2021)，其将多个弱回归树模型集成形成一个强分类器，回归树每次迭代可减少上次迭代的残差，并在残差减少的梯度方向上训练新的模型。LGB是基于梯度提升决策树框架提出的改进模型。相较XGB算法，LGB拥有训练效率高、内存使用低、准确率高、支持并行化学习等优点(Ke et al, 2017)。该算法使用基于直方图的分割算法取代了传统的预排序遍历算法，能有效防止过拟合。CAT也是基于梯度提升决策树的机器学习框架，该算法的改进之处就在于在学习的时候处理这些特征，而不是在数据预处理阶段，不需要任何显式的预处理来将类别转换为数字(宋慧娟等，2022)。本文主要应用这3种机器学习算法从超参数调优、优选特征因子两方面进行模型训练。

#### 3.1 超参数调优

超参数调优是机器学习模型训练的重要部分。通过优化超参数，可以增强模型泛化能力，显著提高模型性能。常见的超参数调优方法包括网格搜索、随机搜索、贝叶斯优化和进化算法等(Nguyen, 2019; 刘佳星, 2022)。贝叶斯优化通过构建概率模型来估计超参数与模型性能之间的关系，选择最有可能提高性能的超参数进行试验，逐步逼近最优解，适用于各种类型的超参数调优，且能动态调整策略，高效利用计算资源加速优化流程(Li et al, 2017)。基于ECMWF模式温度预报单一特征因子，分别构建3种单模型输入方案，定量评估参数优化前后各模型预报性能的改进程度。

图4给出了2024年1月1日至4月30日3种方法超参数优化前后整体的预报效果对比。直接建模对模式预报性能有一定改进，RMSE降至2℃以内。通过贝叶斯优化超参数，3种方法预报误差可进一步降低，RMSE分别由1.956℃、1.949℃、1.937℃降至1.917℃、1.919℃、1.910℃，较ECMWF模式最大降低0.648℃。

从不同预报时效的改进效果可见(图5)，3种方法超参数优化后对夜间温度预报误差较

214 ECMWF 模式有明显降低。以 XGB 模型为例(图 5a), 优化参数后在 0、12、15、18、21、  
 215 24h 时效误差降幅最大, RMSE 减小率达 22.8%~46.6%, 这与 LGB 和 CAT 模型调参后得到  
 216 的定性结果一致。另外, 3 种方法调参前后在 3、6、9h 时效的预报误差均较 ECMWF 模式  
 217 偏高。这一现象可能揭示了模型在特定时段面临的物理复杂性挑战。这些时效对应着白天边  
 218 界层的发展阶段, 湍流混合增强, 局地环流(如城市热岛环流、山谷风)具有高度的非线性特征  
 219 (朱丽等, 2020; 王倩倩等, 2022)。单一的模式温度预报因子虽然包含了模式对这些过程  
 220 的模拟结果, 但其本身是一个高度集成的输出量, 并未体现驱动这些变化的关键物理信息,  
 221 机器学习模型难以充分捕捉这些复杂非线性物理过程的内在关联和变化规律, 更容易陷入欠  
 222 拟合或过拟合。因此, 超参数只是影响模型质量的重要因素之一, 合理优选特征因子能否进  
 223 一步改进模型性能? 这将在下文进行分析。

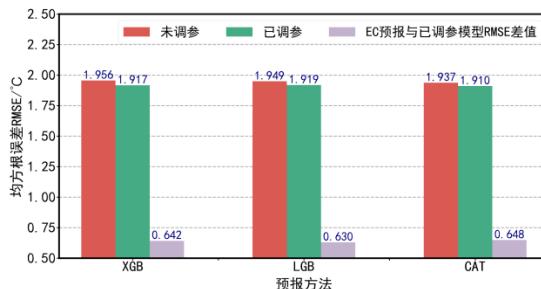


图 4 2024 年 1 月 1 日至 4 月 30 日 3 种模型调参前后温度预报的平均 RMSE

Fig.4 The average RMSE of the temperature forecasts by three models before and after hyperparameter tuning  
from January 1 to April 30 2024

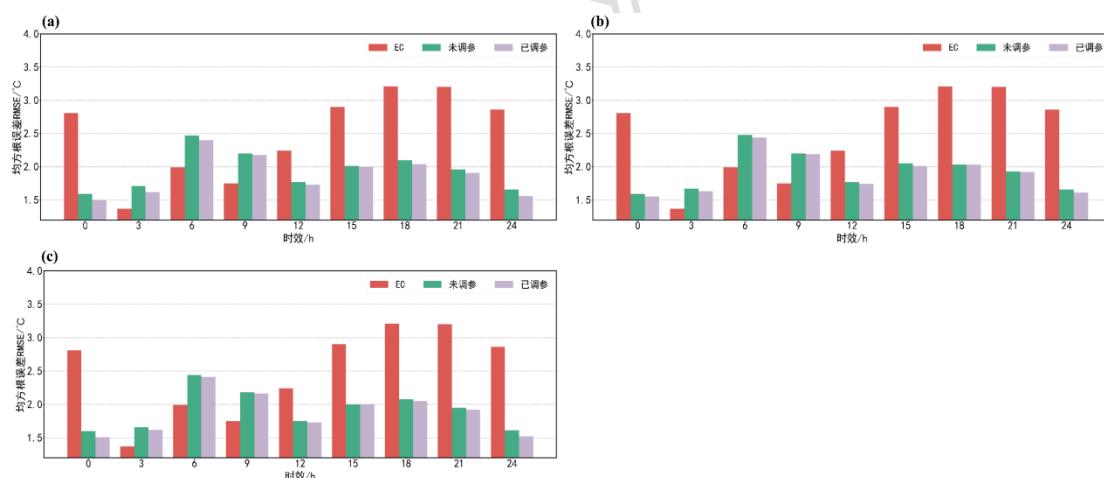


图 5 2024 年 1 月 1 日—4 月 30 日 (a)XGB, (b)LGB, (c)CAT 模型调参前后及 ECMWF 不同预报时效温度  
预报的平均 RMSE

Fig.5 The average RMSE of the temperature forecasts by (a)XGB, (b)LGB, and (c)CAT models before and after  
hyperparameter tuning and ECMWF with different forecastlead times from January 1 to April 30 2024

### 3.2 优选特征因子

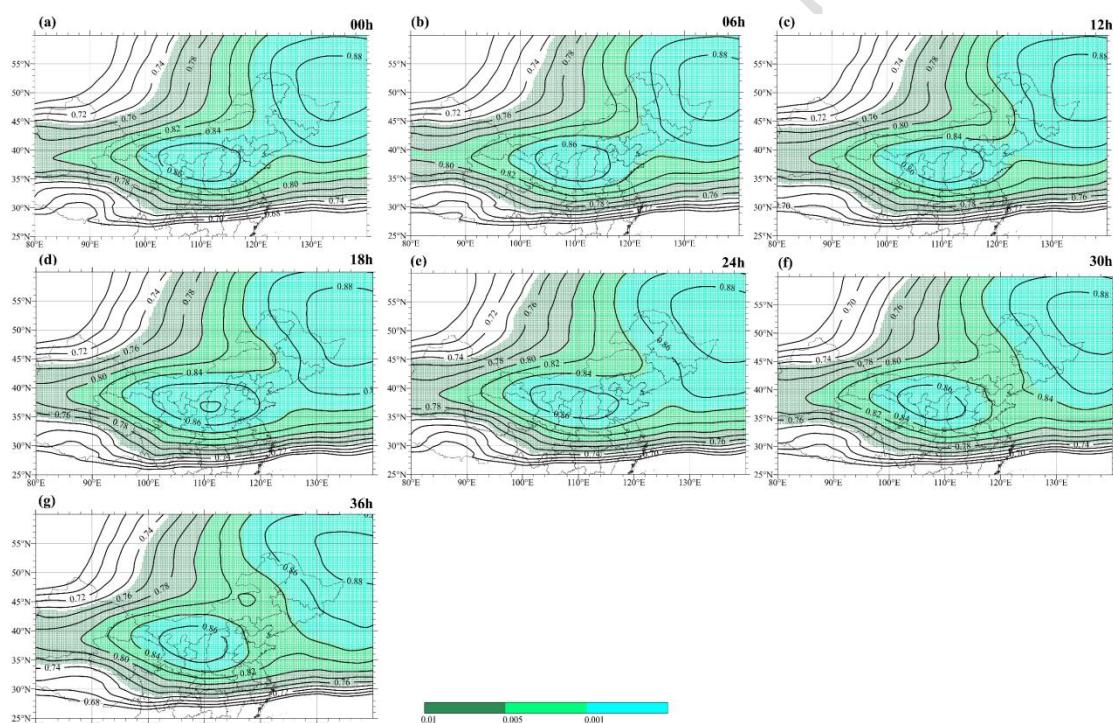
在机器学习算法确定的情况下, 如何选取特征因子对提高模型质量至关重要。尽管输入海量的特征因子能够使模型更好地捕捉数据的内在规律, 但过量的特征因子可能带来特征冗余、数据噪声等问题(Chandrashekhar and Sahin, 2014)。就温度预报订正模型而言, 在构建地面和高空区域特征因子时, 如何挖掘出各个预报时效上对应的重要区域因子, 是影响模型质

240 量的关键。

### 241 3.2.1 方案设计

242 在温度精细化预报建模中，除数值模式输出的温度预报产品外，还需综合考虑大尺度环  
243 流异常、水汽分布等相关物理量场特征。 $500\text{hPa}$  位势高度场的槽脊异常通过冷暖平流机制  
244 直接影响地面温度，槽区冷平流导致地面降温，脊区暖平流导致地面增温。 $850\text{hPa}$  比湿则  
245 直接影响低层水汽含量，进而通过影响云和辐射过程来改变地面温度。因此，本文设计了 2  
246 组对比试验。试验 1 将模式  $2\text{m}$  温度预报临近点插值到西安站作为唯一特征因子，即单因子  
247 方法。试验 2 为多因子方法，根据主观先验知识在试验 1 的基础上增加了两类优选特征因子：  
248 第一类是由模式预报临近点插值到西安站点的物理量场，包括总云量、低云量、 $10\text{m}$  纬向风、  
249  $10\text{m}$  经向风、 $850\text{hPa}$  温度、 $700\text{hPa}$  温度；第二类为通过时滞相关分析法筛选得到的模式不同预报时效  
250 不同关键区平均值，涵盖  $500\text{hPa}$  位势高度和  $850\text{hPa}$  比湿。

251 时滞相关分析是通过计算站点实况与不同时效模式预报场的相关，选取通过显著性检验  
252 ( $\alpha \leqslant 0.001$ ) 的区域作为关键区域，提取该区域的平均值作为预报特征因子。以西安站 00:00  
253 起报的 0h 预报为例，图 6 给出了 2021 年 9 月 1 日至 2023 年 12 月 31 日该站 00:00 温度实  
254 况与 ECMWF 模式不同时效的  $500\text{hPa}$  位势高度场的相关分布，发现在内蒙古中部—陕西北  
255 部、东北亚存在 2 个高相关区。对不同时效内 2 个高相关区进行区域平均后再与 00:00 温度  
256 实况序列进行相关分析，发现与 ECMWF 模式零场的相关系数最大，分别达到 0.915、0.892，  
257 故将模式零场作为该站 0h 预报  $500\text{hPa}$  位势高度因子的关键时效。以此类推，西安站 00:00  
258 起报的其他时效同样按照此方法，对每一个时效单独进行时滞相关分析，最终获得西安站未  
259 来 24h 内不同预报时效的高空区域因子(图略)。



260 注：填色区域分别表示通过  $\alpha=0.01$ 、 $\alpha=0.005$  和  $\alpha=0.001$  显著性水平检验的区域。

261  
262 图 6 2021 年 9 月 1 日至 2023 年 12 月 31 日西安站 00:00 温度实况与 ECMWF 模式前一日 12:00 起报的不  
263 同预报时效  $500\text{hPa}$  位势高度场的相关系数

264 Fig.6 Correlation between temperature observation at 00:00 UTC of Xi'an Station and the  $500\text{hPa}$   
265 geopotential height field forecasted by ECMWF initiated at 12:00 UTC of the previous day from 1 September  
266 2021 to 31 December 2023

267

268

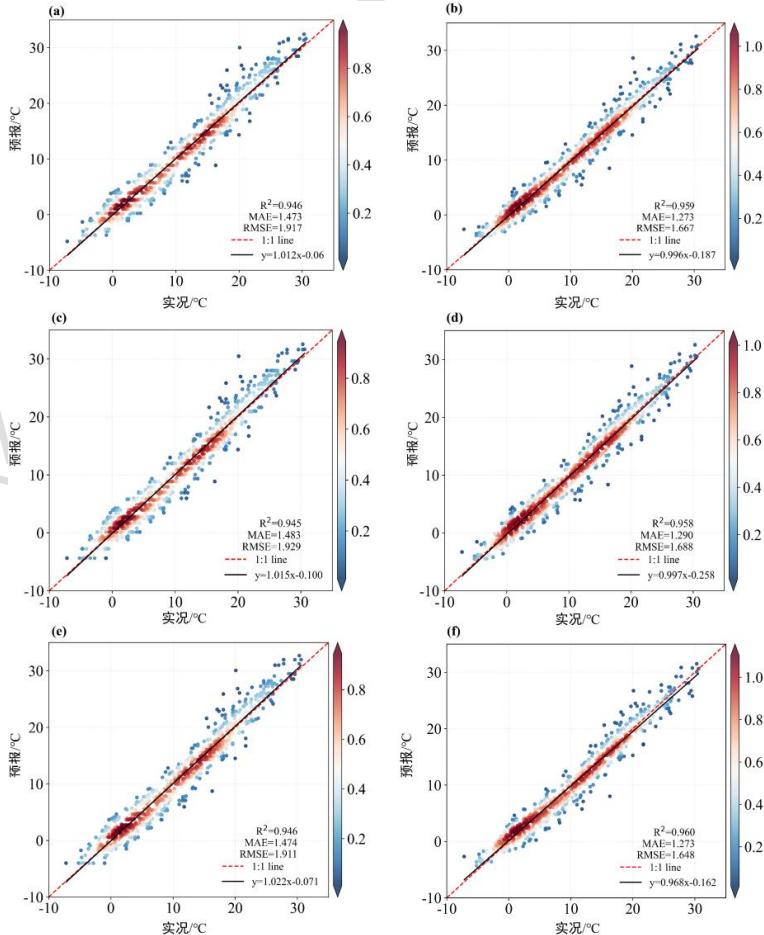
图 7 基于时滞相关分析法和主观先验知识确立的不同预报时效的特征因子  
 Fig.7 Characteristic factors for each forecast lead time established based on the time-lag correlation analysis  
 method and subjective prior knowledge

270

271

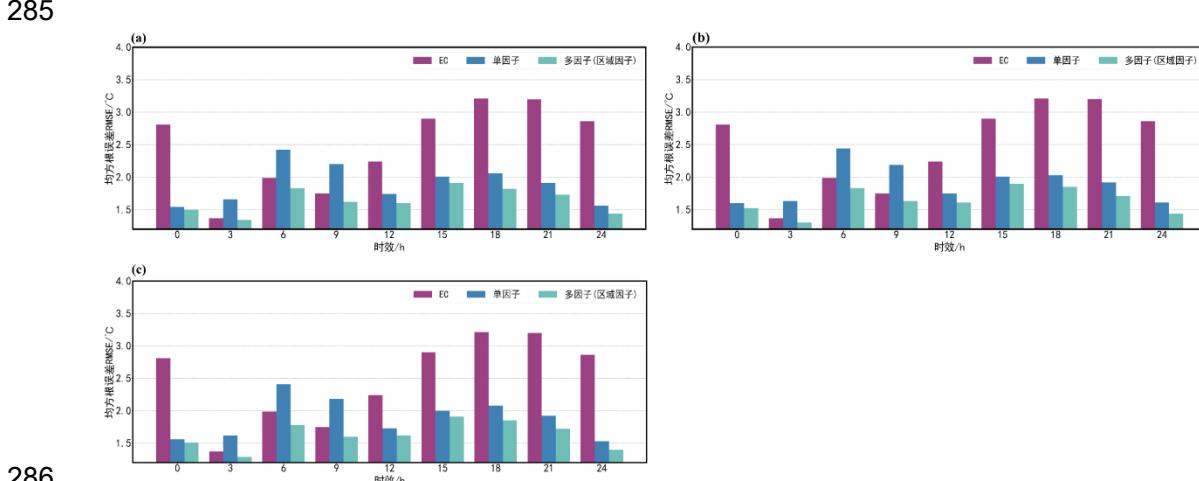
### 3.2.2 预报效果对比

由图 8 可见, 3 种方法优选特征因子后在 2024 年 1 月 1 日至 4 月 30 日整体时段的预报效果均有所提升, XGB、LGB、CAT 模型  $R^2$  分别由 0.946、0.945、0.946 提升至 0.959、0.958、0.960, RMSE 和 MAE 分别下降了 0.250°C、0.241°C、0.263°C 和 0.200°C、0.193°C、0.201°C。从不同时效预报表现可见(图 9), 3 种方法优选特征因子后在夜间较单因子方法预报误差进一步降低。以 XGB 模型为例(图 9a), 12~24h 时效 RMSE 平均减小 0.156°C, 特别是在 3、6、9h 时效均有明显改进, RMSE 低于 ECMWF 和单因子模型预报。以 6h 时效为例, XGB、LGB、CAT 方法 RMSE 分别较 ECMWF 预报降低了 0.16°C、0.16°C、0.21°C。



280

281                  图 8 2024 年 1 月 1 日至 4 月 30 日实况和 (a,c,e) 单因子与(b,d,f)多因子建模预报散点回归  
 282                  (a,b)XGB、(c,d)LGB、(e,f)CAT  
 283          Fig.8 Scatter regression between observations and (a,c,e) single-factor , and (b,d,f) multi-factor modeling for  
 284          (a,b) XGB, (c,d) LGB, and (e,f) CAT from 1 January to 30 April 2024



286          图 9 2024 年 1 月 1 日至 4 月 30 日 (a)XGB、(b)LGB、(c)CAT 模型单因子与多因子建模及 ECMWF 不同  
 287          预报时效温度预报的平均 RMSE  
 288  
 289          Fig.9 The average RMSE of temperature forecast with different forecast lead times by ECMWF and single-factor  
 290          and multi-factor modeling for (a) XGB, (b) LGB, and (c) CAT models from 1 January to 30 April 2024

## 291      4 模型评估

292          基于先验知识和时滞相关分析优选特征因子后,通过贝叶斯超参数优化对单模型预报效  
 293          果有明显改进。但由于不同单模型存在一定的局限性,因而在模型训练的基础上采用  
 294          Stacking 集成方法,通过融合三种单模型输出有效降低对单一模型的依赖,最终实现预报精  
 295          度的进一步提升。

### 296      4.1 集成方案设计

297          为了评估 Stacking 集成方法对预报效果的影响,设计 2 组对比试验。试验 1: 加权平均,  
 298          计算前 20 天优选特征因子后的 XGB、LGB 和 CAT 预报的均方根误差,利用均方根误差倒  
 299          数对 3 种方法预报进行加权平均得到集成预报结果,具体见式(2)。试验 2: 采用 Stacking 方  
 300          法对优选特征后的 XGB、LGB 和 CAT 单模型进行集成,对比评估两组集成方案模型预报性  
 301          能。

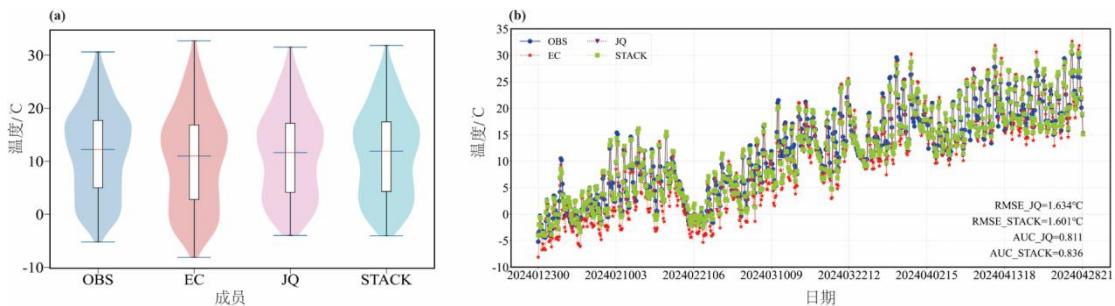
$$302 \quad JQ = \frac{M_X \times \frac{1}{RMSE_X} + M_L \times \frac{1}{RMSE_L} + M_C \times \frac{1}{RMSE_C}}{\frac{1}{RMSE_X} + \frac{1}{RMSE_L} + \frac{1}{RMSE_C}} \quad (2)$$

303          式中:  $JQ$  为加权集成模型预报,  $M_X$ 、 $M_L$ 、 $M_C$  分别为 XGB、LGB 和 CAT 模型预报,  
 304           $RMSE_X$ 、 $RMSE_L$ 、 $RMSE_C$  分别为 XGB、LGB 和 CAT 模型前 20d 预报均方根误差。由于加权  
 305          平均法是取前 20d 单模式均方根误差倒数进行集成,故文中模型集成部分测试集取 2024 年  
 306          1 月 23 日至 4 月 30 日数据,有效测试样本 6318 组。

### 307      4.2 结果评估

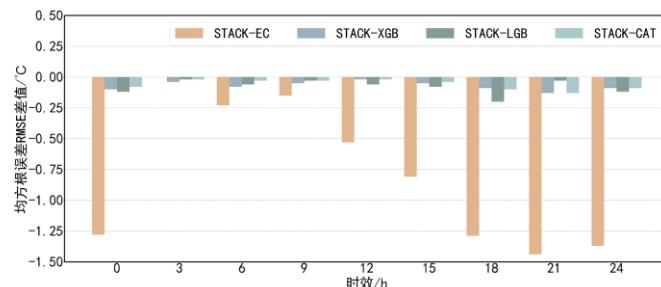
308          图 10a 给出了 2024 年 1 月 23 日至 4 月 30 日加权集成和 Stacking 集成方案与西安西安  
 309          站温度实况的整体对比,可以看出在测试集整体时段内西安站温度在 8~20°C 数据密度高,  
 310          即此处样本更为密集,温度最低、最高值分别为 -8.5°C、30.2°C。ECMWF 温度预报样本密  
 311          集区间值整体较实况略偏低,但极端温度预报强度较实况偏强,即极端高(低)温较实况偏高  
 312

313 (低), 2 种集成预报方案温度整体分布与实况较为一致。2 种集成方案温度预报与实况变化  
 314 趋势基本一致, 但 Stacking 集成方案预报效果优于加权集成方案, 其 RMSE 和 AUC 分别为  
 315 1.601°C、0.836, 低(高)于后者对应指标(图 10b)。

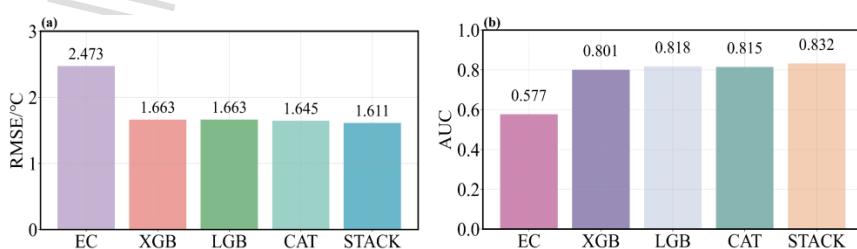


316  
 317 图 10 2024 年 1 月 23 日至 4 月 30 日西安站温度实况和预报的(a)小提琴图 及 (b)时间序列  
 318 Fig.10 (a) Violin plots and (b) time series of temperature observations at Xi'an Station and forecasts from  
 319 23 January to 30 April 2024

320 进一步对比 Stacking 集成模型在不同时效的预报指标及其与单模型的差异可见(图 11),  
 321 Stacking 集成模型不同预报时效的预报误差均小于单模型预报, 特别是夜间预报偏差较单模  
 322 型及 ECMWF 进一步降低, 对于模式预报偏差最大的 21h, Stacking 集成模型预报 RMSE 为  
 323 1.70°C, 较 ECMWF、XGB、LGB、CAT 模型预报分别减小 1.44°C、0.125°C、0.014°C、0.119°C。  
 324 从整体时段检验来看(图 12), Stacking 集成模型预报效果优于单模型, RMSE 和 AUC 为  
 325 1.611°C、0.832, RMSE(AUC)较单模型最大降低(提高)0.034°C(0.031)。



326  
 327 图 11 2024 年 1 月 23 日至 4 月 30 日 Stacking 集成与 ECMWF 和单模型不同预报时效温度预报平均 RMSE  
 328 的差值  
 329 Fig.11 Difference of the average RMSE for temperature forecasts with different forecast lead times between the  
 330 Stacking ensemble and ECMWF and single model from 23 January to 30 April 2024  
 331



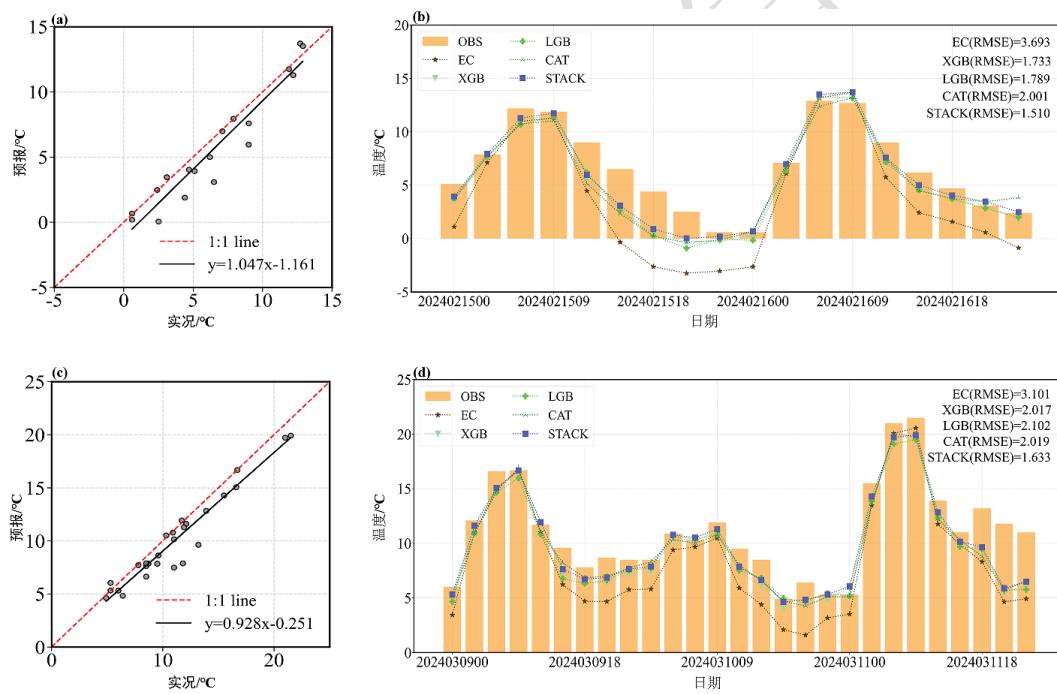
332  
 333 图 12 2024 年 1 月 23 日至 4 月 30 日温度预报平均(a )RMSE,、(b)AUC 对比  
 334 Fig.12 Comparison of the average (a) RMSE and (b) AUC for temperature forecasts from 23 January to 30 April  
 335 2024  
 336

## 337 5 模型应用

338 上述分析表明，Stacking 集成模型能够有效提升单模型在整体时段的温度预报效果，特  
 339 别是西安夜间温度预报偏差较单模型及 ECMWF 大幅降低。本节基于前文选取的 0215 降温  
 340 过程和 0309 降水过程，进一步评估 Stacking 集成模型在明显天气过程中温度预报的改进效  
 341 果。

342 由图 13 可见，尽管 Stacking 集成对 2 次过程西安关键点温度预报整体较实况略偏低(图  
 343 13a,c)，但相较 ECMWF 模式和单模型，Stacking 集成预报效果有明显提升。在 0215 降温过  
 344 程中，15 日夜间模式较实况温度预报显著偏低，Stacking 集成预报值较模式预报升高 2~3℃，  
 345 更接近实况(图 13b)。从 2 次过程整体预报表现来看(图 14)，ECMWF 预报误差最大，RMSE  
 346 超过 3℃，3 种单模型与 Stacking 集成分别较 ECMWF 预报有明显改善，Stacking 集成误差  
 347 减小幅度最大。2 次过程中 Stacking 集成模型 RMSE 较 ECMWF 分别降低了 2.18℃、1.57℃，  
 348 AUC 分别提高了 0.47、0.48(图 14a,14c)。0309 降水过程 Stacking 集成 AUC 最高，达 0.8(图  
 349 14b,d)。

350 综上所述，在明显天气过程中，Stacking 集成能够一定程度上改进模式温度预报偏低的  
 351 问题。因此，Stacking 集成模型应用于西安大城市关键点温度精细化预报中，可为重大活动  
 352 温度预报提供客观支撑。



353  
 354 图 13 2024 年(a,b)2 月 15—16 日，(c,d)3 月 9—11 日 (a,c)西安站温度实况和 Stacking 集成预报散点回归，  
 355 (b,d)西安站温度实况和预报的时间序列

356 Fig.13 (a,c) Scatter plots with regression line for temperature observations at Xi'an Station and forecasts by  
 357 Stacking ensemble, and (b,d) time series of observations at Xi'an Station and forecasts during (a,b) 15—16  
 358 February and (b,d) 9—11 March 2024  
 359

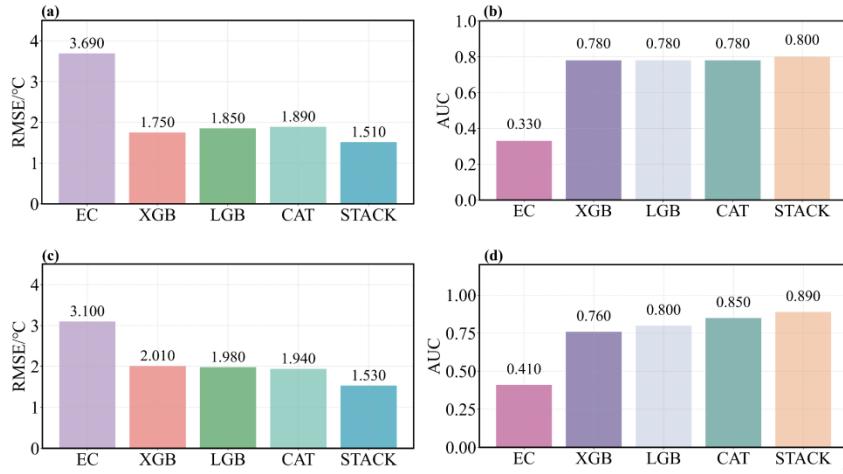


图 14 2024 年(a,b)2 月 15—16 日, (c,d)3 月 9—11 日温度预报 平均(a,c) RMSE、(c,f) AUC 对比

Fig.14 Comparison of the average (a,c) RMSE and (b,d) AUC of temperature forecasts during (a,b) 15—16 February and (b,d) 9—11 March 2024

## 6 结论与讨论

利用西安城区站点温度实况和 ECMWF 模式高空和地面预报数据, 基于 XGB、LGB、CAT 模型和 Stacking 集成方法, 建立了西安大城市关键点未来 24h 温度预报集成模型。通过对模型预报效果进行评估, 得到以下结论:

(1)ECMWF 模式对西安关键点温度预报整体上与实况变化趋势较为一致, 但在 10~20°C、-2~8°C 这 2 个区间预报较实况明显偏低。模式温度预报误差存在日变化特征, 即白天小夜间大。在明显天气过程中, 模式温度预报不能较好地反映西安关键点温度的变化, 存在显著冷偏差。

(2)基于 ECMWF 模式温度预报单一特征构建的 XGB、LGB 和 CAT 模型能够有效降低模式对西安大城市关键点温度预报偏差, 且夜间均方根误差减小明显。采用贝叶斯方案优化超参数后 3 种模型夜间温度预报误差可进一步降低。以 XGB 模型为例, 优化超参数后夜间温度预报 RMSE 减小率达 22.8%~46.6%。

(3)基于主观经验确立模式 2m 温度、总云量、低云量、10m 纬向风、10m 经向风、850hPa 温度、700hPa 温度等 7 个直接预报特征因子, 利用时滞相关分析方法确立不同预报时效的高空区域特征因子(500hPa 位势高度、850hPa 比湿)。优选特征因子后 3 种方法在整体时段上预报效果均有所提升, RMSE 分别下降了 0.250°C、0.241°C、0.263°C。在不同预报时效 3 种方法温度预报偏差均小于原始模式。

(4)基于 Stacking 方法对 3 种方法预报进行集成, 测试集整体时段内 RMSE(AUC)低(高)于加权集成方案相应指标值, 预报效果优于后者。特别是夜间 Stacking 集成温度预报模型偏差较单模型大幅降低。相较单模型, Stacking 集成可以进一步降低西安城区明显天气过程温度预报误差。

本研究通过筛选优化特征因子、调整目标区域的贝叶斯超参数, 采用 Stacking 集成融合优化单模型预报结果, 为西安大城市关键点温度预报提供了一种客观参考方法。预报订正流程和方案具有一定的普适性, 可以移植推广到其他站点, 具备潜在的业务应用价值。本文构建的基于机器学习的站点尺度温度精细化预报模型在冬春过渡期展现出一定优势, 其性能可能得益于该季节冷空气活动频繁且大尺度环流相对稳定的特点。然而, 不同季节主导温度变化的大尺度环流系统及物理过程存在显著差异: 夏季高温热浪与西太平洋副热带高压位置强

392 度、大陆高压与伊朗高压的阶段性合并有关；秋季晴空辐射降温主要受夜间边界层逆温发展  
393 影响。这些差异可能导致模型在不同季节的误差来源和预报订正难度不同，其优势能否稳定  
394 推广至其他季节仍需通过纳入全年代表性样本进行交叉验证。另外，在样本数较少的极端转  
395 折性天气过程中，该模型的订正能力有限。一方面是因为机器学习模型性能与训练样本数量  
396 密切相关，过少的样本数使得模型无法学习到有效特征；另一方面与机器学习集成模型自身  
397 算法框架有关。因此，后期需要在数据增强、深度学习建模等方面（代刊等，2025；金荣花  
398 等，2025）深入研究以提高不同大气环流背景下模型温度预报精度。

399 致谢：感谢国家气象中心天气预报技术研发室对本文的数据支持和技术指导。

400

## 401 参考文献

- 402 代刊,杨绚,周康辉,等,2025.深度学习在数字智能天气预报中的应用[J].气象,51(11):1477-1494.Dai K,Yang X,Zhou K H ,et  
403 al,2025.Application of deep learning in digital intelligent weather forecasting[J].Meteor Mon,51(11):1477-1494 (in Chinese).
- 404 甘璐,郭金兰,雷蕾,等,2021.北京世园会开幕式期间弱降水天气成因[J].气象与环境学报,37(3): 12-18. Gan L, Guo J L, Lei L,  
405 et al, 2021. Causes of a weak rainfall event during the opening ceremony of the Beijing International Horticultural Exhibition[J]. J Meteor  
406 Environ, 37(3): 12-18 (in Chinese).
- 407 韩念霏,杨璐,陈明轩,等,2022.京津冀站点风温湿要素的机器学习订正方法[J].应用气象学报,33(4): 489-500. Han N F, Yang L,  
408 Chen M X, et al, 2022. Machine learning correction of wind, temperature and humidity elements in Beijing-Tianjin-Hebei region[J].  
409 J Appl Meteor Sci, 33(4): 489-500 (in Chinese).
- 410 贾俊平,2019.统计学基础:第4版[M].北京:中国人民大学出版社. Jia J P, 2019. Fundamental Statistics[M]. 4th ed. Beijing: China  
411 Renmin University Press (in Chinese).
- 412 姜红,何清,曾晓青,等,2021.基于随机森林和卷积神经网络的FY-4A号卫星沙尘监测研究[J].高原气象,40(3): 680-689. Jiang  
413 H, He Q, Zeng X Q, et al, 2021. Sand and dust monitoring using FY-4A satellite data based on the random forests and convolutional  
414 neural networks[J]. Plateau Meteor, 40(3): 680-689 (in Chinese).
- 415 金荣花,曹勇,赵瑞霞,等,2025.中央气象台智能数字天气预报业务的技术进展[J].气象,51(11):1321-1334.Jin R H, Cao Y ,Zhao R  
416 X,et al,2025.Technological advances in the intelligent digital weather forecasting operational system of National Meteorological  
417 Centre[J].Meteor Mon,51(11):1321-1334 (in Chinese).
- 418 李嘉睿,符娇兰,陶亦为,等,2022.冬奥会张家口赛区气温与风的特征分析[J].气象,48(2):149-161.LI J R,FU J L,TAO Y W, et  
419 al,2022.Temperature and wind characteristic analysis in Zhangjiakou Olympic Area for the Winter Olympic Games[J].Meteor  
420 Mon,48(2):149-161(in Chinese).
- 421 李妮娜,刘湊华,林建,等,2024.2019—2021年北京春季温度预报的精细化检验评估[J].气象, 50(11): 1397-1408. Li N N, Liu C H,  
422 Lin J, et al, 2024. Refined evaluation of spring temperature forecast in Beijing during 2019-2021[J]. Meteor Mon, 50(11): 1397-1408  
423 (in Chinese).
- 424 刘华,杨依军,朱超,等,2023.续写千年友谊开辟崭新未来习近平主席在西安主持首届中国-中亚峰会纪实[J].中亚信息, (5): 4-9.  
425 Liu H, Yang Y J, Zhu C, et al, 2023. Continuation of millennium friendship opens up a new future[J]. Cent Asian Inf, (5): 4-9 (in  
426 Chinese).
- 427 刘佳星,2022.基于网格搜索超参数优化的支持向量回归[J].科学技术创新, (13): 71-74. Liu J X, 2022. Support vector regression  
428 based on grid search hyperparameter optimization[J]. Sci Technol Innovation, (13): 71-74 (in Chinese).
- 429 刘杰,刘高平,安晶晶,等,2024.基于机器学习的模式温度预报订正方法[J].沙漠与绿洲气象,18(3): 96-104. Liu J, Liu G P, An J J,  
430 et al, 2024. Correction method of model temperature forecast based on machine learning[J]. Desert Oasis Meteor, 18(3): 96-104 (in  
431 Chinese).
- 432 漆梁波,2025.重大活动气象保障中的预报难点回顾及展望[J].气象,51(11):1417-1432.Qi L B,2025.Review and outlook on challenges  
433 in meteorological service for major events[J].Meteor Mon,51(11):1417-1432(in Chinese).
- 434 钱莉,兰晓波,杨永龙,2010.最优子集神经网络在武威气温客观预报中的应用[J].气象, 36(5): 102-107. Qian L, Lan X B, Yang Y  
435 L, 2010. The application of optimal subset neural network to temperature objective forecast in Wuwei[J]. Meteor Mon, 36(5):

- 436 102-107(in Chinese).
- 437 秦庆昌,于波,徐路扬,等,2022.造成北京春季气温预报失误的持续性低云过程分析[J].气象,48(6): 705-718. Qin Q C, Yu B, Xu  
438 L Y, et al, 2022. Analysis of a continuous low-cloud process causing spring temperature forecast errors in Beijing[J]. Meteor Mon,  
439 48(6): 705-718(in Chinese).
- 440 沈学顺,苏勇,李兴良,等,2025.中国业务数值天气预报自主发展中的理论和方法创新[J].气象,51(11):1293-1320.
- 441 Shen X S,Su Y,Li X L,et al ,2025.Innovation of theory and methodology in the independent development of operational numerical  
442 weather prediction in China[J].Meteor Mon,51(11):1293-1320(in Chinese).
- 443 宋慧娟,陈耀登,欧阳霖,等,2022.基于Catboost和Stacking融合模型的长江中下游短时临近降水预报研究[J].气象科学,42(5):  
444 569-580. Song H J, Chen Y D, Ouyang L, et al, 2022. Short-range forecast of precipitation over the middle-lower reaches of the  
445 Yangtze River based on the Catboost and Stacking model[J]. J Meteor Sci, 42(5): 569-580 (in Chinese).
- 446 苏刚,秦胜伍,乔双双,等,2021.基于Stacking集成学习的泥石流易发性评价:以四川省雅江县为例[J].世界地质,40(1): 175-184.
- 447 Su G, Qin S W, Qiao S S, et al, 2021. Debris flow susceptibility evaluation based on Stacking ensemble learning: a case study in  
448 Yajiang, Sichuan Province[J]. Global Geol, 40(1): 175-184 (in Chinese).
- 449 谭江红,陈伟亮,王珊珊,2018.一种机器学习方法在湖北定时气温预报中的应用试验[J].气象科技进展,8(5): 46-50. Tan J H,  
450 Chen W L, Wang S S, 2018. Using a machine learning method for temperature forecast in Hubei province[J]. Adv Meteor Sci  
451 Technol, 8(5): 46-50 (in Chinese).
- 452 唐钧,安东元会,王勇,2021.大型体育赛事风险控制:促精细化、增集成度、提承受力[J].中国减灾,(15): 40-45. Tang J, An D Y H,  
453 Wang Y, 2021. Risk management in major sports events: promoting refinement, enhancing integration, and improving resilience[J].  
454 Disaster Reduct China, (15): 40-45 (in Chinese).
- 455 佟华,张玉涛,齐倩倩,等,2022.基于CMA模式体系的京津冀地区复杂地形下冬季的精细化地面要素多模式集成预报研究[J].  
456 气象,48(12): 1539-1549. Tong H, Zhang Y T, Qi Q Q, et al, 2022. The multi-model blending forecasts of near-surface parameters  
457 based on CMA model system[J]. Meteor Mon, 48(12): 1539-1549 (in Chinese).
- 458 王倩倩,权建农,程志刚,等,2022.2019年冬季北京海陀山局地环流特征及机理分析[J].气象学报,80(1): 93-107. Wang Q Q,  
459 Quan J N, Cheng Z G, et al, 2022. Local circulation characteristics and mechanism analysis of Haituo mountain in Beijing during  
460 winter 2019[J]. Acta Meteor Sin, 80(1): 93-107 (in Chinese).
- 461 王珊珊,雷彦森,方鸿斌,等,2022.基于CatBoost算法的长江中游冬季降水相态预报方法研究[J].气象,48(9): 1153-1161. Wang S  
462 S, Lei Y S, Fang H B, et al, 2022. Method of winter precipitation type prediction in the middle reaches of the Yangtze River based on  
463 CatBoost algorithm[J]. Meteor Mon, 48(9): 1153-1161 (in Chinese).
- 464 王在文,全继萍,张鑫宇,2023.基于CMA-BJ数值预报模式产品的复杂地形下冬奥站点地面气温和风速预报方法研究[J].气象  
465 学报,81(6): 926-942. Wang Z W, Quan J P, Zhang X Y, 2023. Forecasting surface temperature and wind speed at Winter Olympics  
466 stations over complex terrain based on the CMA-BJ model products[J]. Acta Meteor Sin, 81(6): 926-942 (in Chinese).
- 467 吴启树,韩美,郭弘,等,2016.MOS温度预报中最优训练期方案[J].应用气象学报,27(4): 426-434. Wu Q S, Han M, Guo H, et al,  
468 2016. The optimal training period scheme of MOS temperature forecast[J]. J Appl Meteor Sci, 27(4): 426-434 (in Chinese).
- 469 徐景峰,宋林烨,陈明轩,等,2023.冬奥会复杂山地百米尺度10 m风速预报的机器学习订正对比试验[J].大气科学,47(3):  
470 805-824. Xu J F, Song L Y, Chen M X, et al, 2023. Comparative machine learning-based correction experiment for a 10 m wind  
471 speed forecast at a 100 m resolution in complex mountainous areas of the Winter Olympic Games[J]. Chin J Atmos Sci, 47(3):  
472 805-824 (in Chinese).
- 473 轩春怡,吴春艳,刘勇洪,等,2022.基于风险矩阵的重大活动气象风险评估[J].大气科学学报,45(5): 791-800. Xuan C Y, Wu C Y,  
474 Liu Y H, et al, 2022. Meteorological risk assessment of major activities based on risk matrix[J]. Trans Atmos Sci, 45(5): 791-800 (in  
475 Chinese).
- 476 薛湛彬,陈娴,张瑛,等,2019. ECMWF高分辨率模式2 m温度预报误差订正方法研究[J].气象,45(6): 831-842. Xue C B, Chen X,  
477 Zhang Y, et al, 2019. Bias correction method for the 2 m temperature forecast of ECMWF high resolution model[J]. Meteor Mon,  
478 45(6): 831-842 (in Chinese).
- 479 闫文杰,刘圣军,刘新儒,等,2022.基于高相关区域上最小角回归的华南初夏暴雨日数预测[J].数学理论与应用,42(2): 35-46.

- 480 Yan W J, Liu S J, Liu X R, et al, 2022. Prediction of early summer rainstorm days in South China based on least angle regression on  
481 high correlation regions[J]. *Math Theory Appl.*, 42(2): 35-46 (in Chinese).
- 482 杨璐, 南刚强, 陈明轩, 等, 2021. 基于三种机器学习方法的降水相态高分辨率格点预报模型的构建及对比分析[J]. *气象学报*,  
483 79(6): 1022-1034. Yang L, Nan G Q, Chen M X, et al, 2021. The construction and comparison of high resolution precipitation type  
484 prediction models based on three machine learning methods[J]. *Acta Meteor Sin*, 79(6): 1022-1034 (in Chinese).
- 485 曾晓青, 薛峰, 赵瑞霞, 等, 2019. 几种格点化温度滚动订正预报方案对比研究[J]. *气象*, 45(7): 1009-1018. Zeng X Q, Xue F, Zhao  
486 R X, et al, 2019. Comparison study on several grid temperature rolling correction forecasting schemes[J]. *Meteor Mon*, 45(7):  
487 1009-1018 (in Chinese).
- 488 张芳, 王刚, 张朝飞, 等, 2024. GA-BP 神经网络在发射场的气温预报应用研究[J]. *载人航天*, 30(2): 219-223. Zhang F, Wang G,  
489 Zhang C F, et al, 2024. Research on application of GA-BP artificial neural network model for launch Centre temperature forecast[J].  
490 Manned Spaceflight, 30(2): 219-223 (in Chinese).
- 491 赵声蓉, 2006. 多模式温度集成预报[J]. *应用气象学报*, 17(1): 52-58. Zhao S R, 2006. Multi-model consensus forecast for  
492 temperature[J]. *J Appl Meteor Sci*, 17(1): 52-58 (in Chinese).
- 493 智协飞, 孙晶, 周文友, 2015. 2009 年夏季西太平洋台风的集合预报和多模式集成预报试验[J]. *大气科学学报*, 38(5): 633-640. Zhi  
494 X F, Sun J, Zhou W Y, 2015. Ensemble and multimodel ensemble forecasts of western Pacific typhoons during summer 2009[J].  
495 *Trans Atmos Sci*, 38(5): 633-640 (in Chinese).
- 496 智协飞, 吴佩, 俞剑蔚, 等, 2019. GFS 模式地形高度偏差对地面 2 m 气温预报的影响[J]. *大气科学学报*, 42(5): 652-659. Zhi X F,  
497 Wu P, Yu J W, et al, 2019. Impact of topographic altitude bias of the GFS model on the 2 m air temperature forecast[J]. *Trans Atmos  
498 Sci*, 42(5): 652-659 (in Chinese).
- 499 朱丽, 苗峻峰, 高阳华, 2020. 重庆城市热岛环流结构和湍流特征的数值模拟[J]. *大气科学*, 44(3): 657-678. Zhu L, Miao J F, Gao Y  
500 H, 2020. A numerical simulation of urban breeze circulation structure and turbulence in Chongqing[J]. *Trans Atmos Sci*, 44(3):  
501 657-678 (in Chinese).
- 502 朱育雷, 杨静, 钟水新, 等, 2024. 基于多神经网络的动态权重集成温度预报订正研究[J]. *热带气象学报*, 40(1): 156-168. Zhu Y L,  
503 Yang J, Zhong S X, et al, 2024. Research on temperature forecast correction by dynamic weight integration based on multi-neural  
504 networks[J]. *J Trop Meteor*, 40(1): 156-168 (in Chinese).
- 505 Chandrashekhar G, Sahin F, 2014. A survey on feature selection methods[J]. *Comput Electr Eng*, 40(1): 16-28.
- 506 Chen T Q, Guestrin C, 2016. XgBoost: a scalable tree boosting system[C]//Proceedings of the 22nd ACM SIGKDD International  
507 Conference on Knowledge Discovery and Data Mining. San Francisco: Association for Computing Machinery: 785-794.
- 508 Cui B, Toth Z, Zhu Y J, et al, 2012. Bias correction for global ensemble forecast[J]. *Weather Forecast*, 27(2): 396-410.
- 509 Cui Z J, Qing X X, Chai H X, et al, 2021. Real-time rainfall-runoff prediction using light gradient boosting machine coupled with  
510 singular spectrum analysis[J]. *J Hydrol*, 603: 127124.
- 511 Glahn H R, Lowry D A, 1972. The use of Model Output Statistics (MOS) in objective weather forecasting[J]. *J Appl Meteorol Climatol*,  
512 11(8): 1203-1211.
- 513 Ke G L, Meng Q, Finley T, et al, 2017. LightGBM: a highly efficient gradient boosting decision tree[C]//Proceedings of the 31st  
514 International Conference on Neural Information Processing Systems. Long Beach: ACM: 3149-3157.
- 515 Li L S, Jamieson K, DeSalvo G, et al, 2017. Hyperband: a novel bandit-based approach to hyperparameter optimization[J]. *J Mach Learn*,  
516 18(1): 6765-6816.
- 517 Nguyen V, 2019. Bayesian optimization for accelerating hyper-parameter tuning[C]//Proceedings of 2019 IEEE Second International  
518 Conference on Artificial Intelligence and Knowledge Engineering. Sardinia: IEEE: 302-305.
- 519 Wolpert D H, 1992. Stacked generalization[J]. *Neural Netw*, 5(2): 241-259.
- 520