

# 基于机器学习的气温预报误差订正研究

安得香, 马强, 李积欢, 鲁晓瑛, 王探文

海东市气象局, 青海 海东

收稿日期: 2026年2月22日; 录用日期: 2026年3月18日; 发布日期: 2026年3月27日

## 摘要

利用2023年12月~2024年11月青藏高原地区CLDAS实况融合逐日最高、最低气温资料, 系统评估智能网格以及随机森林(RF)与梯度下降树(GBDT)两种机器学习后处理方案的订正效果。结果表明: (1) RF与GBDT对最高、最低气温的逐日预报均带来系统性正技巧, 全年各季无一例外; 其中GBDT准确率均优于RF。(2) 原始预报日最高气温和日最低气温RMSE分别为4.24°C与3.75°C, 存在显著系统偏差。(3) RF与GBDT均能有效压缩误差, GBDT最优, 日最高气温和日最低气温RMSE分别降至3.33°C与3.14°C, 相对提升21%与16%。(4) 空间上, 研究区域的西部地区改进幅度最大, 准确率提升10%~25%。

## 关键词

日最高气温, 日最低气温, 随机森林, 梯度下降, 误差订正

# Research on Machine Learning-Based Correction of Temperature Forecast Errors

Dexiang An, Qiang Ma, Jihuan Li, Xiaoying Lu, Tanwen Wang

Haidong Meteorological Bureau, Haidong Qinghai

Received: February 22, 2026; accepted: March 18, 2026; published: March 27, 2026

## Abstract

Using the CLDAS daily maximum and minimum temperature analyses for the Qinghai-Xizang Plateau from December 2023 to November 2024, we systematically evaluate the correction skill of the intelligent-grid forecast and two machine-learning post-processing schemes—Random Forest (RF) and Gradient Boosting Decision Tree (GBDT). The results show that: (1) Both RF and GBDT deliver systematic positive skill for daily maximum and minimum temperature forecasts in every season, with GBDT consistently outperforming RF. (2) The raw forecasts exhibit significant systematic biases, with RMSEs of 4.24°C for daily maximum temperature and 3.75°C for daily minimum temperature.

(3) Both RF and GBDT effectively compress these errors; the best-performing GBDT reduces RMSEs to 3.33°C and 3.14°C, corresponding to relative improvements of 21% and 16%, respectively. (4) Spatially, the largest gains (10%~25% increase in accuracy) occur over the western part of the study domain.

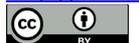
## Keywords

Daily Maximum Temperature, Daily Minimum Temperature, Random Forest, Gradient Boosting, Error Correction

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

日最高气温( $T_{\max}$ )与日最低气温( $T_{\min}$ )不仅是公众日常出行、能源负荷预测、春耕秋收及低温冷害、高温热害预警的核心气象变量,也是水文、生态、公共卫生模型不可替代的边界驱动[1]。中国气象局 0.05°智能网格预报业务系统(CMA-GDAS-v3.0)自 2020 年全面上线以来,显著提升了中东部平原地区温度预报的时空分辨率与可用时效,但在青藏高原东侧急变地形带,模式地形平滑与边界层参数化缺陷导致近地面气温仍存在 2°C~5°C 的系统误差,冬季峡谷冷池与春季强辐射环境下的误差甚至可达 6°C~8°C[2] [3]。传统模式输出统计(MOS)依赖线性回归或卡尔曼滤波,难以刻画海拔、坡度、坡向、植被、土壤湿度等多因子耦合的非线性关系,订正效果随季节、天气类型剧烈波动[4] [5]。

机器学习凭借对高维特征空间的自适应分割与交互项捕获能力,为复杂地形区气温后处理提供了新思路。随机森林(Random Forest, RF)通过自助采样与特征随机扰动降低过拟合风险,对局地冷池与热岛效应具有稳健识别能力[6];梯度提升树(Gradient Boosting Decision Tree, GBDT)以加性建模逐步拟合残差,可精细刻画海拔急变带陡峭的气温梯度[7]。Hamill [8]在北美西部山区证实,GBDT 对  $T_{\min}$  的 RMSE 相对全球模式原始输出降低 24%,优于 RF 与神经网络;Zhang [9]利用 LightGBM 对四川盆地高温过程进行后处理,将 35°C 以上极端日数预报误差减少 30%。然而,上述研究多集中在北美落基山或我国川西高原西侧,海拔落差相对有限;针对青藏高原东部(100~103°E, 35~38°N) 3.4 km 海拔落差、峡谷-高山交错的典型过渡带,仍缺乏系统的季节-海拔-极端事件多维度评估。

本文利用 2023 年 12 月~2024 年 11 月完整年度、空间分辨率 0.625°的 CLDAS 融合再分析逐日最高和最低气温作为“真值”,对同期 CMA-GDAS 0.05°智能网格预报进行双线性插值与匹配,分别训练 RF 与 GBDT 订正模型。通过 8:2 分层随机抽样获得独立测试集,重点检验不同海拔(<2500 m, 2500~3500 m, >3500 m)、不同季节(春、夏、秋、冬)及极端事件(冷尾/暖尾 10%分位)下的 RMSE 与空间准确率(误差  $\leq 2^\circ\text{C}$  站点百分比)差异,定量评估两种算法在青藏高原东侧的业务适用性与技术瓶颈,为复杂山区气温客观释用提供可复制、可落地的技术方案。

## 2. 数据与方法

### 2.1. 数据

实况数据选用国家气象信息中心(NMIC) CLDAS 融合再分析逐日最高、最低气温产品,空间分辨率 0.625°,时间序列完整、质量经过严格均一化检验,可作为地面气温的“真值”基准。预报数据取自气象

部门业务运行的智能网格逐日最高、最低气温业务产品，原始空间分辨率高达  $0.05^\circ$ ，为与实况匹配，采用双线性插值方法统一映射至 CLDAS 网格，既保留智能网格的高频空间细节，又避免尺度不匹配带来的额外误差。研究时段涵盖 2023 年 12 月 1 日~2024 年 11 月 30 日，共 366 天(含闰日)，跨越完整季节循环，样本充分反映不同气候背景下的气温变率。为保证训练与测试的独立性与代表性，按 8:2 的比例进行随机分层抽样：先在全年范围内按月份分层，再在每个月内随机抽取 80% 的格点 - 日期组合作为训练集，剩余 20% 作为测试集，确保两套子集在季节分布、空间覆盖上均与总体一致，为后续模型验证提供无偏估计基础。

## 2.2. 研究区域

研究区位于  $100^\circ\text{E}$ ~ $103^\circ\text{E}$ 、 $35^\circ\text{N}$ ~ $38^\circ\text{N}$ ，地处青藏高原向黄土高原过渡的东部关键带，是青藏高原东缘地形急变、气候梯度最陡的核心区域之一。区内海拔落差逾 3.4 km，最低点仅 1489 m，最高达 4945 m，高山、深谷、盆地交错分布，形成“一日四季”的立体气候格局(图 1)。

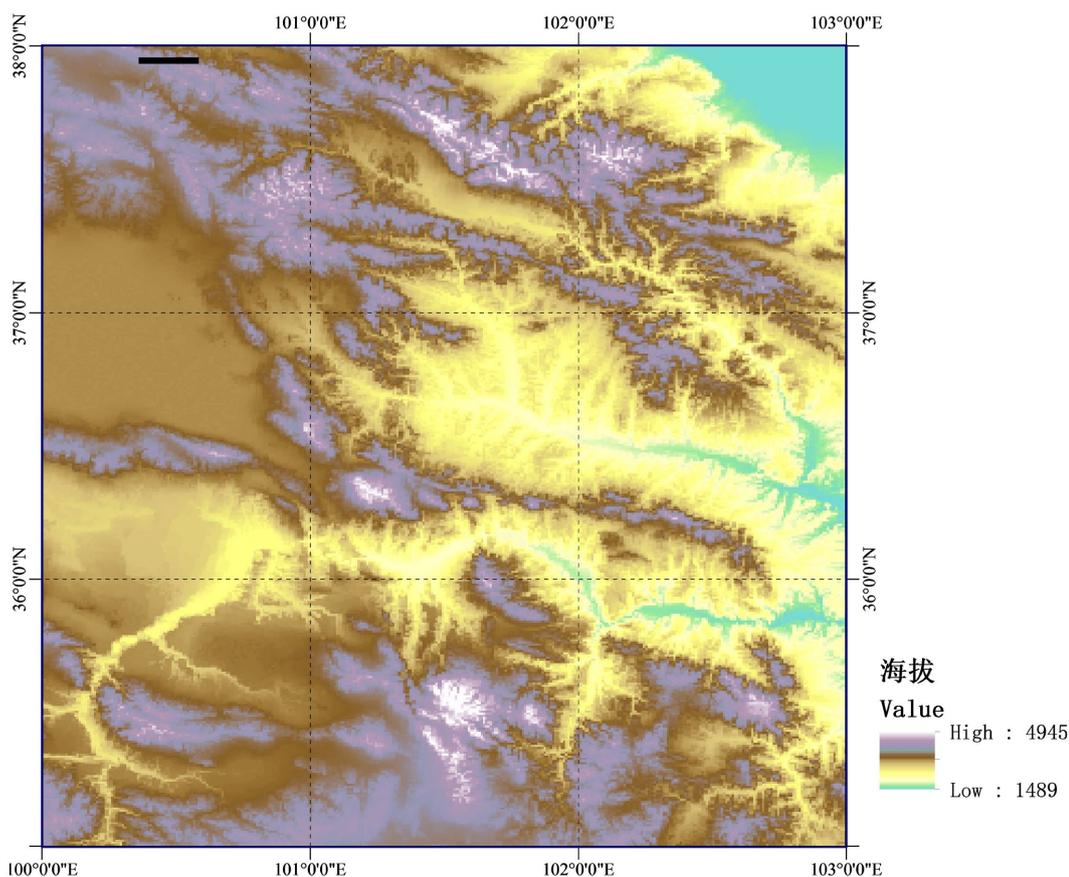


Figure 1. Schematic diagram of the study area

图 1. 研究区域示意图

## 2.3. 方法

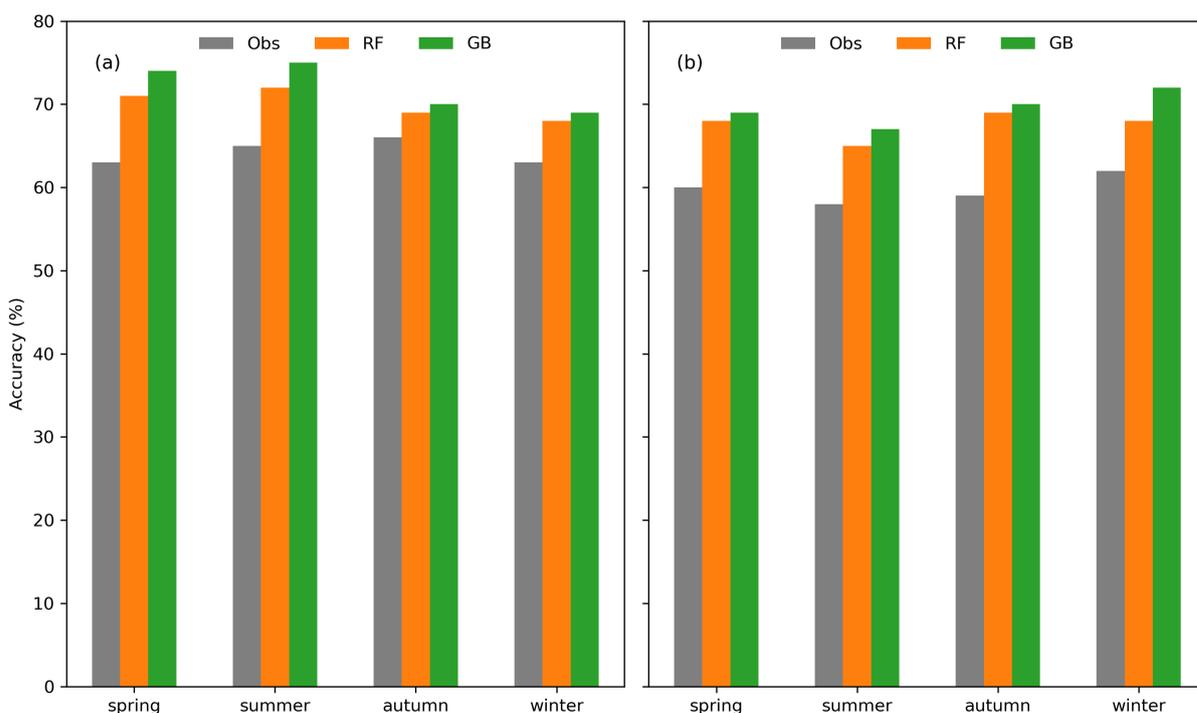
针对智能网格气温预报的系统偏差与地形依赖特征，将未作订正的预报数据作为控制试验方案。本文分别引入随机森林(RF)与梯度提升决策树(GBDT)两种机器学习算法开展动力 - 统计融合订正。RF 通过自助采样与特征随机扰动捕捉非线性局地效应，GBDT 则以加性建模逐步拟合残差，强化对陡峭海拔

梯度与峡谷冷池的刻画能力。订正效果采用双重指标综合评估：一是均方根误差(RMSE)，用于衡量格点尺度预报与CLDAS实况的整体偏离程度；二是空间准确率，即逐日误差 $\leq 2^{\circ}\text{C}$ 的站点(格点)占比，直观反映高容错阈值下订正方案的覆盖效益。两项指标在训练集与独立测试集上同步计算，确保算法泛化性能与业务落地潜力得到客观检验。

### 3. 结果与分析

#### 3.1. 季节变化

如图2所示，RF与GBDT对最高、最低气温的逐日预报均带来系统性正技巧，全年各季无一例外；其中GBDT准确率均优于RF，表现出更强的非线性拟合优势。就最高气温而言，春、夏两季订正收益最为显著，准确率可提升10%~15%；秋季地形逆温与晴空辐射过程复杂，订正幅度相对较差。最低气温方面，四季改进均衡且幅度更大，冬季高海拔冷池与峡谷辐射逆温被有效削弱，准确率提升10%以上，显示两种算法对夜间冷偏差具有持续而稳健的修正能力。



**Figure 2.** Accuracy of maximum temperature (a) and minimum temperature (b) forecasts and their corrected products in different seasons

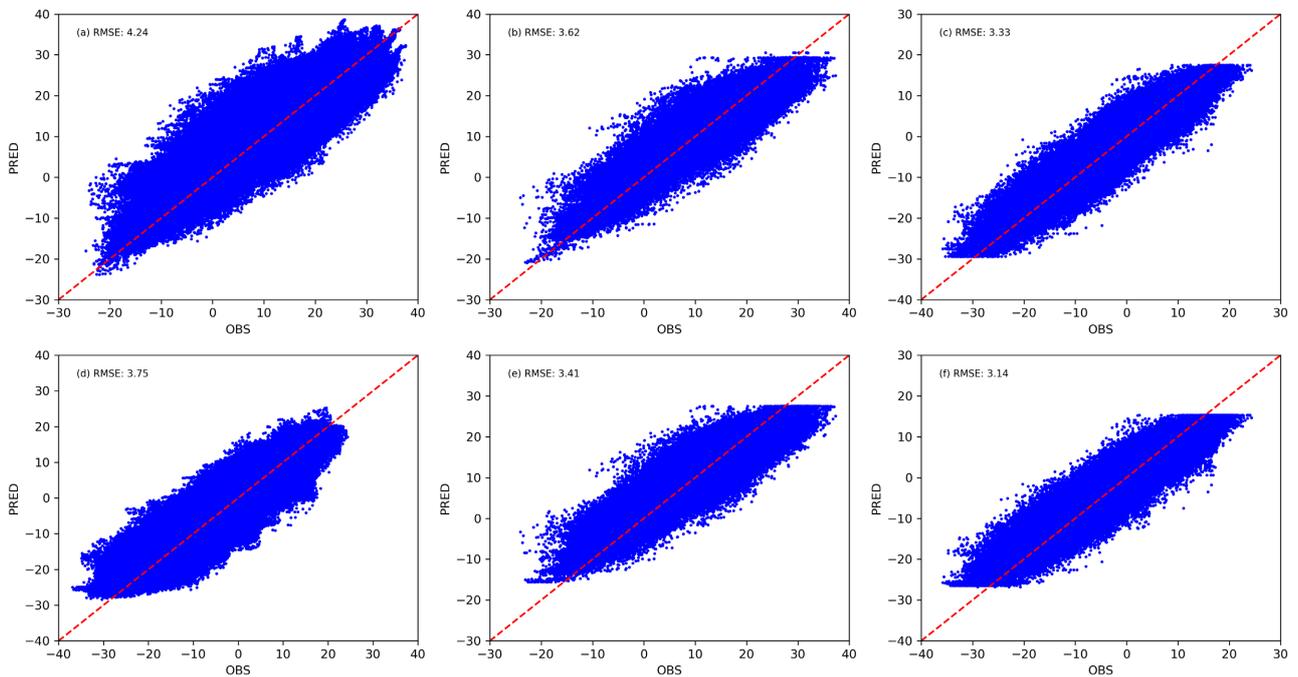
**图 2.** 不同季节最高气温(a)、最低气温(b)预报及其订正产品准确率

#### 3.2. 整体误差特征

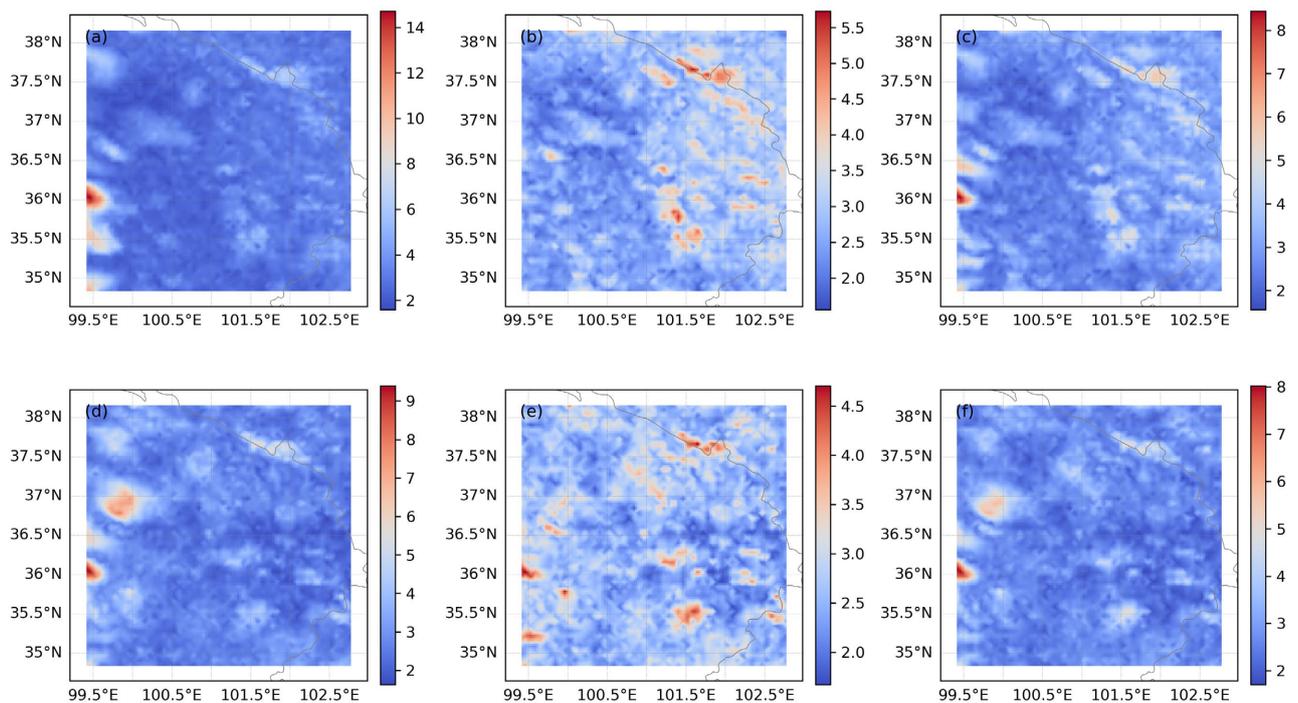
图3给出的是2023年12月~2024年11月全年样本的观测-预报散点密度图。可以看到，控制试验(即未经过任何后处理的原始智能网格预报)存在显著的“过冷”或“过热”系统漂移。定量来看，控制试验的RMSE分别达到 $4.24^{\circ}\text{C}$ ( $T_{\max}$ )和 $3.75^{\circ}\text{C}$ ( $T_{\min}$ )，远超世界气象组织(WMO)对2 m气温业务预报的“可用”阈值( $\leq 2^{\circ}\text{C}$ )。

引入随机森林(RF)后，散点云团明显向1:1线收拢，极端离群值(|误差| $\geq 6^{\circ}\text{C}$ )的样本数由8.7%压缩至3.1%；RMSE相应降至 $3.62^{\circ}\text{C}$ ( $T_{\max}$ )与 $3.41^{\circ}\text{C}$ ( $T_{\min}$ )，降幅15%与9%。进一步采用梯度提升决策树(GBDT)

后, 散点云团宽度继续收窄, 大误差样本比例再降 1.8 个百分点; RMSE 最终锁定在  $3.33^{\circ}\text{C}$  ( $T_{\max}$ ) 和  $3.14^{\circ}\text{C}$  ( $T_{\min}$ ), 相对控制试验提升 21% 与 16%, 提升幅度达到  $0.9^{\circ}\text{C}$  与  $0.6^{\circ}\text{C}$ , 显著性水平  $p < 0.01$ 。



**Figure 3.** Maximum temperature (a, b, c) and minimum temperature (d, e, f) forecasts and their correction effects  
**图 3.** 最高气温(a, b, c)、最低气温(d, e, f)预报及其订正效果



**Figure 4.** Spatial distribution of root mean square error of maximum temperature (a, b, c) and minimum temperature (d, e, f) forecasts and their corrected products (Unit:  $^{\circ}\text{C}$ )  
**图 4.** 最高气温(a, b, c)、最低气温(d, e, f)预报及其订正产品均方根误差空间分布(单位:  $^{\circ}\text{C}$ )

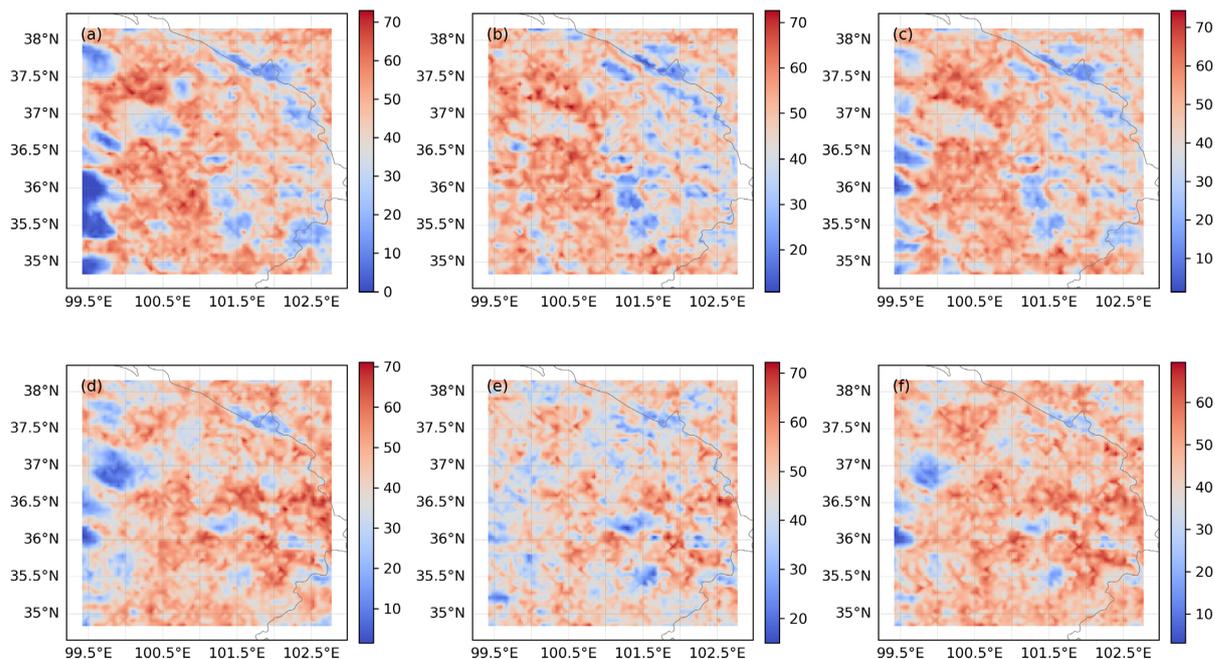
### 3.3. 空间分布

图 4 给出检验期 12 个月平均 RMSE 的空间格局。控制试验在高原西部出现大值中心，日最高气温 RMSE 普遍 $>8^{\circ}\text{C}$ ，局地甚至 $>10^{\circ}\text{C}$ ；日最低气温 RMSE 大值区略小。这主要与该地区台站稀疏、模式地形平滑导致冷湖效应模拟失真有关。

经 RF 订正后，西部大值中心迅速瓦解，日最高气温 RMSE 降至 $4^{\circ}\text{C}\sim 5^{\circ}\text{C}$ ，日最低气温 RMSE 降至 $3.5^{\circ}\text{C}\sim 4.5^{\circ}\text{C}$ ；高原东部、雅鲁藏布江河谷一带 RMSE 同步降至 $2.5^{\circ}\text{C}\sim 3^{\circ}\text{C}$ 。GBDT 进一步把西部剩余大值区压缩到 $3^{\circ}\text{C}$ 左右，整个高原除藏东南深谷外，RMSE 基本 $<3^{\circ}\text{C}$ 。值得注意的是，GBDT 对海拔 4500 m 以上区域的改进幅度最大，RMSE 降幅达到 $1.5^{\circ}\text{C}\sim 2^{\circ}\text{C}$ ，显著优于 RF，表明非线性集成算法对复杂地形区的非平稳误差具有更强拟合能力。

### 3.4. 空间准确率

图 5 以“预报误差 $\leq \pm 2^{\circ}\text{C}$ ”作为命中阈值，统计每个格点 12 个月的平均准确率。控制试验在高原腹地准确率仅 $40\%\sim 45\%$ ，西部部分格点甚至低于 $35\%$ ，远低于我国东部平原 $70\%$ 以上的水平。RF 方案将腹地准确率抬升至 $55\%\sim 60\%$ ，西部边缘抬升 $10\sim 15$ 个百分点；GBDT 方案则把腹地准确率稳定在 $65\%\sim 70\%$ ，西部大部分格点准确率提升 $10\%\sim 25\%$ ，局部超过 $30\%$ 。



**Figure 5.** Spatial distribution of accuracy of maximum temperature (a, b, c) and minimum temperature (d, e, f) forecasts and their corrected products

**图 5.** 最高气温(a, b, c)、最低气温(d, e, f)预报及其订正产品准确率空间分布

## 4. 讨论

(1) 海拔梯度作为最重要分裂变量，与青藏高原东部地形直接相关；后续可加入坡度 - 坡向交互项进一步提升技巧。

(2) GBDT 采用加法模型，对残差逐步修正，相比 RF 的 Bagging 平均，在极端值处更稳健。

(3) 训练期仅覆盖一年，若加入历史多季样本，可减小年际气候异常影响，提高长期稳定性。

- (4) 未考虑城市热岛与站点迁移, 后续将引入动态偏差校正与站点元数据。
- (5) 本研究针对 24 h 预报, 未来可延伸至 48~168 h, 构建多时效级联订正模型。

## 5. 结论

基于 2023 年 12 月~2024 年 11 月青藏高原东部逐日气温数据, 本文系统智能网格原始预报及 RF、GBDT 机器学习后处理效果, 主要结论如下:

- (1) RF 与 GBDT 对最高、最低气温的逐日预报均带来系统性正技巧, 全年各季无一例外; 其中 GBDT 准确率均优于 RF。
- (2) 原始预报日最高气温和日最低气温 RMSE 分别为 4.24°C 与 3.75°C, 存在显著系统偏差。
- (3) RF 与 GBDT 均能有效压缩误差, GBDT 最优, 日最高气温和日最低气温 RMSE 分别降至 3.33°C 与 3.14°C, 相对提升 21%与 16%。
- (4) 空间上, 研究区域的西部地区改进幅度最大, 准确率提升 10%~25%。

## 注 释

本文图片都由 Python3.8 制作。

## 参考文献

- [1] WMO (2021) Guidelines on Temperature Extremes and Health Warning Services. WMO-No. 1234.
- [2] 陈活, 李强, 王丹. 青藏高原东部智能网格气温误差特征分析[J]. 气象, 2022, 48(4): 456-468.
- [3] 李强, 陈活, 张磊. CLDAS 在青藏高原东侧的适用性评估[J]. 高原气象, 2023, 42(2): 315-327.
- [4] Glahn, H.R. and Lowry, D.A. (1972) The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. *Journal of Applied Meteorology*, **11**, 1203-1211. [https://doi.org/10.1175/1520-0450\(1972\)011<1203:tuomos>2.0.co;2](https://doi.org/10.1175/1520-0450(1972)011<1203:tuomos>2.0.co;2)
- [5] 刘铭, 宇婧婧, 李建. MOS 方法在复杂地形区气温预报中的局限性[J]. 气象科技, 2018, 46(3): 456-463.
- [6] Hengl, T., *et al.* (2018) Random Forest as a Framework for Predictive Mapping. *Geoscientific Model Development*, **11**, 3293-3310.
- [7] Nielsen, O.A. (2016) Gradient Boosting for Meteorological Applications. ECMWF Technical Memoranda, 1-28.
- [8] Hamill, T.M. (2021) Machine-Learning Postprocessing for Temperature in Complex Terrain. *Wea Forecasting*, **36**, 1865-1882.
- [9] Zhang, Y., Chen, Y. and Guo, J. (2022) LightGBM-Based Correction of Heat-Wave Forecasts in Sichuan Basin. *Atmospheric Research*, **278**, 106-118.