

基于机器学习的复杂地形下短期数值天气预报 误差分析与订正*

任萍^{1,2} 陈明轩² 曹伟华² 王在文² 韩雷¹ 宋林焯² 杨璐²
REN Ping^{1,2} CHEN Mingxuan² CAO Weihua² WANG Zaiwen² HAN Lei¹ SONG Linye² YANG Lu²

1. 中国海洋大学, 青岛, 266100

2. 北京城市气象研究院, 北京, 100089

1. *Ocean University of China, Qingdao 266100, China*

2. *Institute of Urban Meteorology, CMA, Beijing 100089, China*

2020-01-21 收稿, 2020-06-30 改回.

任萍, 陈明轩, 曹伟华, 王在文, 韩雷, 宋林焯, 杨璐. 2020. 基于机器学习的复杂地形下短期数值天气预报误差分析与订正. 气象学报, 78(6): 1002-1020

Ren Ping, Chen Mingxuan, Cao Weihua, Wang Zaiwen, Han Lei, Song Linye, Yang Lu. 2020. Error analysis and correction of short-term numerical weather prediction under complex terrain based on machine learning. *Acta Meteorologica Sinica*, 78(6):1002-1020

Abstract A set of multi-mode integration technology of numerical prediction based on machine learning method XGBoost and consideration of the influence of topographical features has been preliminarily developed. Its integration effect was compared with that of traditional equal weight average and linear regression methods. Based on the data products of the rapid update cycle numerical prediction system in Beijing, which can provide cycle predictions including 2 m air temperature, 2 m relative humidity, 10 m wind speed and 10 m wind direction near the ground 8 times a day, three integrated models of multi-model forecast time lag integrated models were construct based on the machine learning method XGBoost, the equal weight average method and the linear regression method, respectively. Experiments were conducted to compare and analyze the effect of the integrated correction of model predictions at different times in a warm and a cold season every day. The results indicate that in the seasonal test, the integrated prediction results of 2 m air temperature and 10 m full wind speed based on the XGBoost model are significantly improved compared with the original optimal prediction results, and are much better than the results of the other two traditional methods. Using the XGBoost method, the error of 2 m air temperature integration can be reduced by 11.02%—18.09%, the error of 10 m full wind speed integration can be reduced by 31.23%—33.22%, and the error of 10 m wind direction integration can be reduced by 4.1%—8.23%. The integrated forecast error of 2 m relative humidity is close to the that from the traditional method. As a whole, the developed multi-mode integrated prediction model based on XGBoost can fully "excavate" the advantages of different modes or the rapid updating cycle prediction at different times, and therefore effectively reduces the systematic error of the mode and provides a multi-mode integrated deterministic prediction product with higher accuracy.

Key words Integration, Numerical prediction, Machine learning, XGBoost, Linear regression, Equal weight

摘要 初步研发了一套基于机器学习方法 XGBoost 且考虑地形特征影响的数值预报多模式集成技术, 并与传统的等权重平均和线性回归方法的集成效果进行了对比分析。利用北京地区快速更新循环数值预报系统每天 8 次循环预报给出的近地面 2 m 温度、2 m 相对湿度、10 m 风速、10 m 风向数据产品, 分别基于机器学习方法 XGBoost、等权重平均方法、线性回归方法构建了

* 资助课题: 国家重点研发计划项目(2018YFF0300102)、北京市自然科学基金项目(8192016)、国家自然科学基金项目(41875049、41801022)。

作者简介: 任萍, 主要从事机器学习与气象应用研究。E-mail: pren@ium.cn

通信作者: 陈明轩, 主要从事短期临近预报技术研究。E-mail: mxchen@ium.cn

3种体现地形因子影响的多模式预报时间滞后集成模型。试验对比分析了暖季、冷季每日不同时刻的模式预报集成订正效果。结果表明:分季节试验中,基于XGBoost模型对2 m温度、10 m风速的集成预报结果相对原始最优预报结果误差明显优于其他两种传统方法。XGBoost对2 m温度集成的误差可降低11.02%—18.09%,10 m风速集成误差可降低31.23%—33.22%,10 m风向集成误差可降低4.1%—8.23%。2 m相对湿度的集成预报误差与传统方法接近。基于XGBoost的多模式集成预报模型可以充分“挖掘”不同模式或不同时刻快速更新循环预报优点,有效降低模式的系统性误差,提供准确性更高的多模式集成确定性预报产品。

关键词 集成,数值预报,机器学习,XGBoost,线性回归,等权重

中图法分类号 P456

1 引言

精准的天气预报是社会发展和防灾减灾、城市安全运行、民众生活等急需的。近几十年来,随着大气科学理论、观测手段、数值计算方法和计算机技术的迅速发展,数值天气预报的准确率不断提升,成为现代天气预报最为重要的工具。但因为数值模式本身的不确定性及其非线性,数值预报还存在较为显著的误差。

数值天气预报模式不断改进完善,发展了集合预报技术用于提升天气预报的准确性(Stensrud, et al, 2000; Mylne, et al, 2002)。同时,借鉴集合数值天气预报的思路,将不同的数值模式预报或者同一数值模式系统的多个成员预报进行集成和偏差订正,以获得更为精确的预报结果(Huang, et al, 2012; Lu, et al, 2007)。在多模式集成和偏差订正方面,时间滞后集成订正预报是重要的技术之一。多元线性回归在时间滞后集成预报方面得到较好应用,根据数值预报成员的不平衡程度选择不同的权重,从而降低模式预报误差。Yuan等(2009)采用时间滞后模型进行北美加州河流域短时定量降水和概率降水预报,表明在早期初始化的预测周期中添加更多的成员可以显示出更优的预报结果。Chang等(2012)在时间滞后多模式集成预报系统的基础上,利用局域分析与预报系统对台湾地区开展了降水概率定量预报研究。Chen等(2013)对美国国家环境预测中心(NCEP)气候预测系统(CFSv2)的整体滞后季节预报进行了分析,提出预测基于滞后集成时存在最优滞后集成时间。Lu等(2007)使用时间滞后集成预报系统对比了等权重和线性回归两种集成方法,结果表明二者均能提高短期预报,线性回归方法相对于等权重更有优势。

近年来,针对数值预报结果开展了大量的后处理工作,形成了多种数值预报订正技术:(1)统计订

正方法,如基于多元线性回归的模式输出统计方法(MOS)(Glahn, et al, 1972)、完全预报法、卡尔曼滤波法、相似集合法(王在文等, 2019; Monache, et al, 2013)等;(2)机器学习方法,如人工神经网络(赵声蓉, 2006)、支持向量机(Han, et al, 2017; 杨璐等, 2018)等技术。

机器学习作为一种大数据统计方法,近年来广泛应用于数值天气预报模式结果的偏差订正方面(Whan, et al, 2018)。Yuan等(2007)利用人工神经网络试验校正定量降水预报;黄威等(2017)利用支持向量机回归模型做中期强降雨集成预报;Han等(2017)借助支持向量机回归模型进行强对流天气预报;Gagne II等(2014)使用逻辑回归、随机森林做定量降水预报;Sprenger等(2017)基于自适应增强学习算法对焚风进行临近预报。Chen等(2016)首次提出XGBoost算法,该算法是梯度增强机器学习算法(Friedman, 2001)的扩展和优化版本,可以减少模型的过拟合。过拟合是机器学习模型在训练数据过程中过度解读数据,导致训练好的模型不能很好适用其他独立样本数据。另外XGBoost算法与其他算法相比还有一个优势,即在调整模型超参数上花费的时间更少(Chen, et al, 2016)。Haberlie等(2018a, 2018b)对比了随机森林、梯度增强、XGBoost方法在雷达强度回波图像拼接识别中纬度中尺度对流系统的能力,发现XGBoost算法的优势更加明显。XGBoost算法因优良的学习效果以及高效的训练速度获得广泛的关注和应用。但是,XGBoost算法对于模式预报集成和偏差订正方面的研究尚未有太多尝试。

数值预报误差分析与订正主要是为了寻找观测与数值模式输出之间的关系模型,并利用该模型将多个模式预报转变为一个单值预报,以获得更准确的预报结果。文中的短期数值天气预报误差分析与订正是指使用机器学习方法,集成RMAPS-

ST 和 ECMWF-IFS 模式预报结果, 最终获得更接近实际的预报效果。使用机器学习的 XGBoost 算法构建了考虑地形特征的数值预报多模式时间滞后集成模型, 并与传统的等权重平均和线性回归方法的集成效果进行了对比分析, 通过对北京地区观测站点多种气象要素进行集成预报, 挖掘不同集成方法对不同气象要素的影响, 探究机器学习方法在多模式集成和偏差订正方面的优缺点和可用性。

2 方法

2.1 等权重平均方法

等权重平均方法是多模式集成预报最基本的技术方法 (Lu, et al, 2007)。假设有 n 个数值预报模式, 每一个数值预报模式的成员分别是 m_1 、 m_2 、 \dots 、 m_k , 每个预报的权重则为 $\left(\sum_{j=1}^n \sum_{i=1}^k m_{i,j}\right)^{-1}$, 对数值预报各成员集成后的预报为

$$\hat{F} = \sum_{j=1}^n \sum_{i=1}^k F_{i,j} / \sum_{j=1}^n \sum_{i=1}^k m_{i,j} \quad (1)$$

式中, $F_{i,j}$ 为预报成员, 是模式插值到站点上的预报结果, \hat{F} 为集成结果, i 为预报时效, j 为数值预报模式序号, 共 n 个模式。

2.2 基于线性回归的不等权重方法

线性回归模型 (Lu, et al, 2007) 将地形特征和数值模式各预报成员联系起来作为回归模型的输入特征, 根据各个输入特征的不平衡程度分配相应的权重, 从而进一步降低多模式集成的误差。线性回归模型表示为

$$\hat{F} = a + \sum_{k=1}^N b_k F_{i,j} + c_k H \quad (2)$$

式中, a 为随机干扰, b_k 为数值预报模式各成员的权重, $F_{i,j}$ 为预报成员, c_k 为各站点高度的权重, H 为站点高度。

2.3 XGBoost 算法

XGBoost 算法 (Chen, et al, 2016) 是由决策树衍生出的机器学习方法, 是多决策树集成算法即梯度增强机器学习算法 (Friedman, 2001) 的扩展和优化版本。决策树 (又称分类与回归树, Classification and Regression Tree) 和提升学习 (Boosting) 是 XGBoost 的基本组成部分。

2.3.1 决策树

决策树可以理解为计算机自动学习决策点的流程图。训练过程中, 在每个分支节点上, 算法会考虑一些潜在的问题来分割数据, 选择分割是为了最小化预测的误差, 可以是类别、概率或实值 (McGovern, et al, 2017)。决策树算法步骤为:

(1) 考虑训练集的所有数值预报成员和地形特征, 遍历每一个数值预报成员和地形特征下所有可能的取值或者切分点, 将数据集 $D = \{D_1, D_2, \dots, D_n\}$ 按照最优切分点 $D_i = s$ 划分成两部分 (D_{s1} 和 D_{s2})。

(2) 分别计算上述两个子集的平方误差和, 选择最小平方误差对应的特征与分割点, 生成两个子节点。

(3) 对上述两个子节点递归调用步骤 (1)、(2), 直到满足停止条件。决策树停止的条件有: (a) 深度达到指定条件 (试验中深度设为 6), 节点的深度可以理解为节点与决策树根节点的距离, 如根节点的子节点深度为 1, 因为这些节点与根节点的距离为 1, 则子节点的深度要比父节点的深度大 1。决策树的深度是所有叶子节点的最大深度, 当深度到达指定的上限, 停止分裂。(b) 节点的数据量小于指定数量 (试验中节点数据量设为 1), 当节点数据量较小时, 树再分裂容易增加误差。(c) 平方误差和下降值小于指定阈值 (试验中阈值设为 0)。

2.3.2 提升学习

XGBoost 是多个决策树的集成, 集成方法为提升学习集成。提升学习是将单个决策树串联起来的一种集成学习方式, 每一个决策树都依赖前一个决策树的结果。在 XGBoost 算法内, 为了减少上一棵决策树训练结果的误差, 下一棵树的目标值不再是真实值而是上一棵树的误差。文中集成步骤为:

(1) 设有一个样本由多个数值预报成员值 (F_i) 和一个地形特征 (H) 及观测值 (O) 组成。第一棵决策树用这个样本训练得出一个预测值 (\hat{O}) (图 1a)。

(2) 第二棵决策树训练时输入样本变为多个数值预报成员值 (F_i) 和一个地形特征 (H) 及观测预报差 ($O - \hat{O}$) (图 1b), 之后的决策树输入为前一棵决策树的误差, 依此类推, 直到达到设定的树数量。例如, 当设定树数量为 5 时, 则当训练完第 5 棵树, XGBoost 停止训练。

文中基于 XGBoost 算法构建数值预报多模式时间滞后集成模型时, 考虑了不同海拔高度特征对

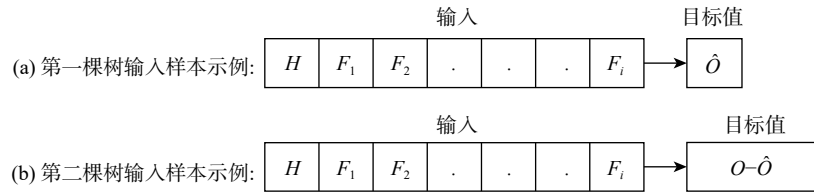


图1 输入样本示例

Fig. 1 Input samples

预报误差的影响,因而模型构建中增加了地形因素。在具体试验中将站点高度作为输入样本的成员参与训练,如图1中 H 为站点高度。

3 试验与检验

3.1 数据

数值模式预报数据分别来自区域和全球模式输出结果。区域模式数据是北京市气象局的睿图短期预报子系统(RMAPS-ST)3 km分辨率格点预报资料,全球模式数据是欧洲中期天气预报中心(ECMWF)的全球预报系统(ECMWF-IFS)0.125°分辨率格点预报数据。两套模式资料的时间均为2017年5—8月及2018年1月—2019年2月。RMAPS-ST每天3 h循环更新预报,ECMWF-IFS每天00和12时(世界时,下同)更新预报。

集成试验针对北京地区的自动气象观测站进行。为了进行对比,空间上采用反距离加权插值方法(Ware, et al, 1991)将上述模式预报数据插值到北京地区的自动气象观测站上,共295个站。以200 m地形等高线划分平原和山区,平原站179个,山区站116个(图2)。时间上,因为自动气象观测站的观测为5 min间隔,两套数值模式的预报数据间隔均为1 h,所以将自动气象观测站整点的观测数据与相应时刻模式的预报数据进行匹配。集成试验的气象要素包括:2 m气温、2 m相对湿度、10 m风速、10 m风向。

首先,进行数据质量控制和预处理。为满足建模需要,若一条样本的预报或观测数据为空或者无效,则同时剔除该样本对应的观测或者预报数据,即观测和预报同时有效才计作一组有效样本。

参考时间滞后模型预报成员的获取方式提取数值模式预报成员(Lu, et al, 2007)(图3)。图3中红色箭头表示当前时刻,黑色线代表能预报到当前时刻的起报时次。例如,RMAPS-ST模式预报时刻

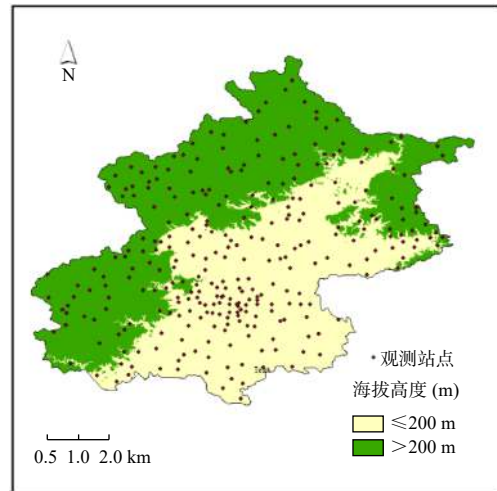


图2 集成试验所用的北京地区观测站点分布

Fig. 2 Distribution of observation sites in Beijing area used in the integration test

在21—00时的预报成员包括前一天00、03、06、09、12、15、18、21时起报的8个成员,预报时效分别为24、21、18、15、12、9、6、3 h。ECMWF-IFS模式预报时刻在00—12时的预报成员包括当天的00时、前一天的00、12时起报的3个预报成员,预报时效分别为36、24、12 h。

3.2 训练和测试集构建

采用训练集数据进行建模,测试集数据进行模型验证和对比分析。试验中训练集构建方法为:按照季节分配训练和测试数据集,将5—8月作为暖季,10—12月及1—2月作为冷季。研究中进行4批试验,分别为:2017年5—7月作为训练集,2017年8月作为测试集;2018年5—7月作为训练集,2018年8月作为测试集;2018年10—12月作为训练集,2019年1月作为测试集;2018年10—12月以及2019年1月作为训练集,2019年2月作为测试集。

3.3 检验方法

(1)均方根误差

均方根误差(RMSE)是回归问题最常用的性能

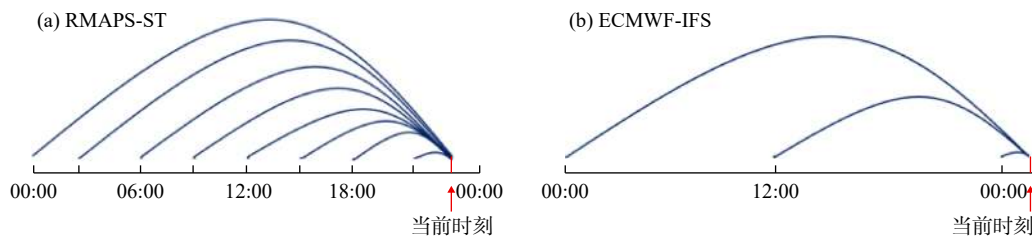


图3 数值模式预报成员的提取示意 (a. RMAPS-ST 预报成员, b. ECMWF-IFS 预报成员)

Fig. 3 Extraction diagram of members of numerical weather prediction

(a. RMAPS-ST forecast members, b. ECMWF-IFS forecast members)

评判指标之一, 公式为

$$\text{RMSE} = \sqrt{\frac{1}{N_F} \sum_{i=1}^{N_F} (F_i - O_i)^2} \quad (3)$$

式中, N_F 为预报总站数, i 为预报站标识, F_i 为第 i 站预报值, O_i 为第 i 站观测值。

(2) 平均绝对误差(MAE)

风向的检验参考中国气象局《QX/T 229—2014 风预报检验方法》。风向预报平均绝对误差为预报风向角度(或预报风向方位对应的中心角度)与观测风向角度误差绝对值的平均值。风向与其他要素不同, 预报值与观测值需要经过前期处理再使用检验公式评估。前期处理阶段, 主要是判断预报值与观测值的差值, 如果差值小于 180° 则差值不变; 如果差值大于 180° 且预报值大于观测值则预报值减 360° , 如果预报值小于观测值则观测值减 360° 。检验公式为

$$\text{MAE}_d = \frac{1}{N_{Fd}} \sum_{i=1}^{N_{Fd}} |F_{d,i} - O_{d,i}| \quad (4)$$

式中, N_{Fd} 为风向预报总站数, i 为风向预报站标识, $F_{d,i}$ 为第 i 站风向预报值, $O_{d,i}$ 为第 i 站风向观测值。

3.4 试验结果

3.4.1 2 m 气温

图4是2 m气温观测的最大与最小值、四分位数、均值。从观测和误差趋势可以看出, 暖季(图4a、b)气温越高误差越大, 冷季(图4c、d)气温越低误差越大。从2 m气温均方根误差可以看出, 原始数值预报(散点)的误差较大, 经过集成以后预报误差明显减小, 且不论哪种方法都能使集成以后预报的系统误差大幅减小。整体来看, XGBoost 优于线性回归, 线性回归优于等权重平均。XGBoost 对暖、冷季气温集成预报误差比原始各成员预报以及等

权重和线性回归集成都明显降低, 是均方根误差最小的一种集成方法, 其次是线性回归。等权重分配法对原始预报误差也有很大改善, 但其误差量级与原始模式成员的最小预报误差接近。

气温预报误差有较明显的日变化, 均方根误差曲线呈单峰分布, 2017(图4a)和2018(图4b)年暖季06时误差最大, ECMWF-IFS 的原始预报误差分别为 3.37 和 3.43°C , 采用 XGBoost 集成后降至 1.83°C ; 2019年冷季1月05时误差最小(图4c), 其中 RMAPS-ST 为 1.70°C , 采用 XGBoost、线性回归集成以后分别降至 1.48 、 1.53°C ; 冷季2月08时误差最小(图4d), RMAPS-ST 为 1.47°C , 采用 XGBoost、线性回归集成以后分别降至 1.42 和 1.40°C 。

为了阐述方便, 在后续评估过程中使用的名词缩写如表1。可以看出, 暖季3种方法对地面2 m气温的集成效果均优于冷季(表2)。XGBoost 集成预报在暖季优势最明显, 2017和2018年暖季分别提高 18.09% 和 17.64% , 2019年冷季1月提高 11.02% , 2月提高 11.39% ; 基于线性回归的集成预报结果在2017和2018年暖季分别提高 14.06% 和 9.12% , 2019年冷季1月提高 2.12% , 2月提高 3.64% ; 等权重平均在2017和2018年暖季分别提高 3.20% 和 4.00% , 而2019年冷季, 等权重平均集成后比原始预报成员的最优预报结果有所降低。

为了进一步探究气温集成误差, 分别对北京地区山区和平原进行集成试验(图5)。2017年暖季00—03时及15—23时平原的3种集成误差相对较小, 其他时刻平原与山区误差基本一致(图5a)。2018年暖季在04—08时平原的集成误差相对较小, 其他时刻山区与平原的误差基本一致, 但山区的等权重误差相对较大(图5b)。冷季2019年1月00—06时平原集成误差相对较小明显, 其他时刻平

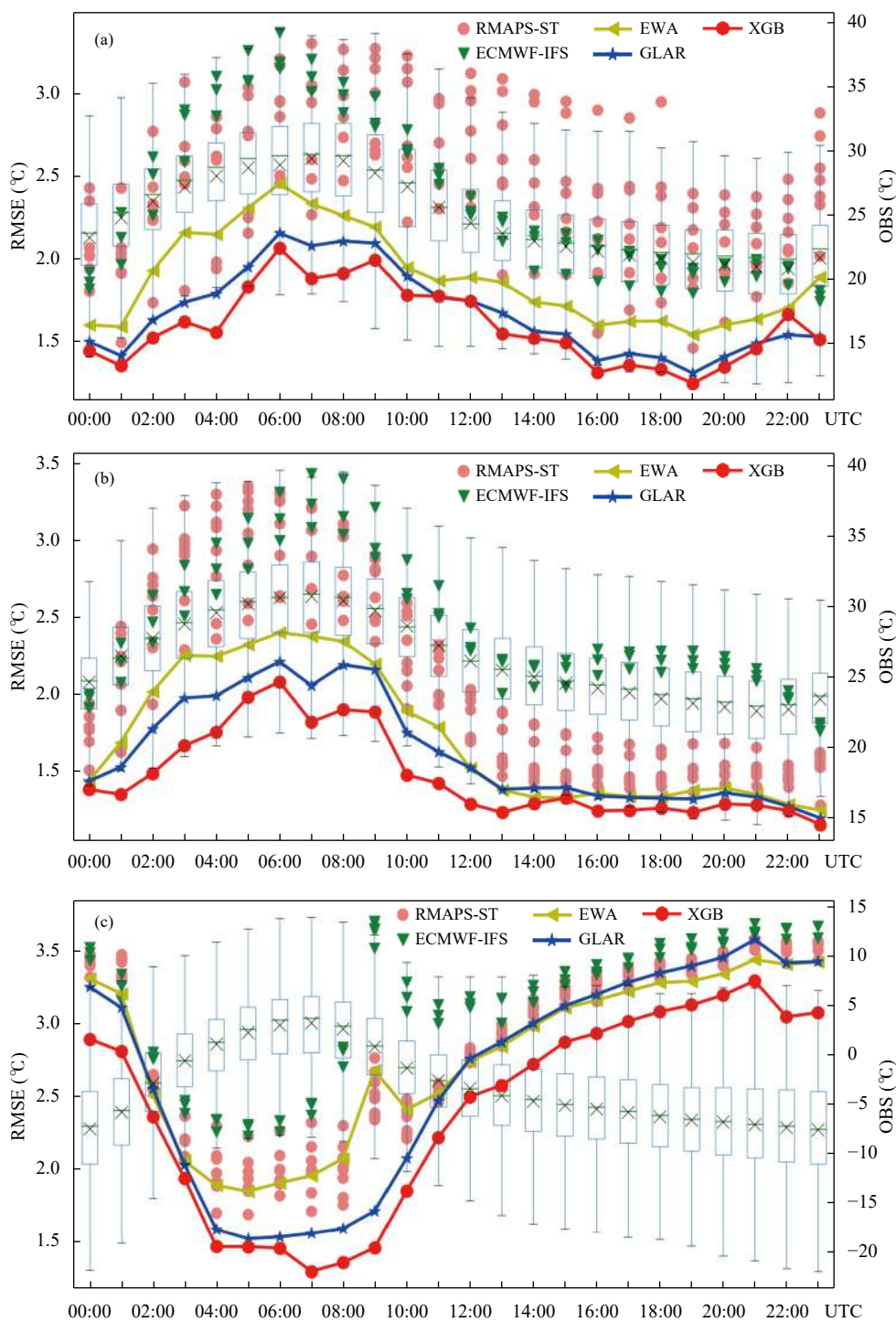
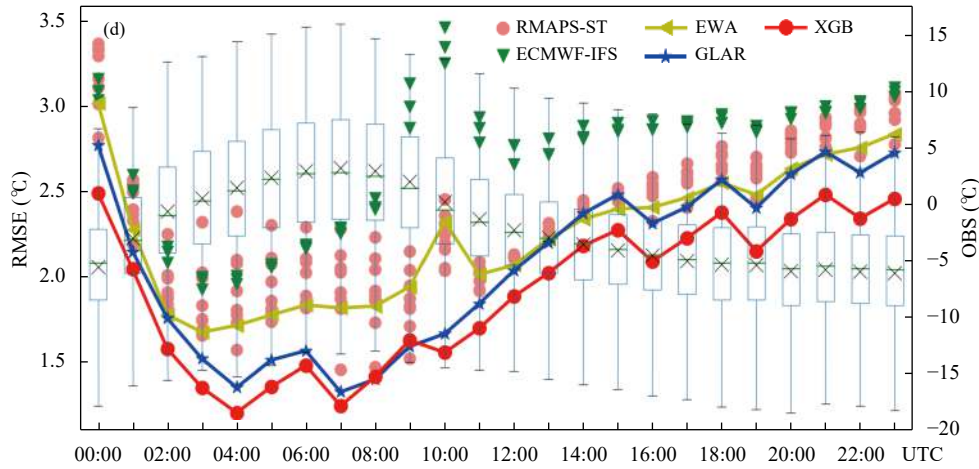


图 4 RMAPS-ST 和 ECMWF-IFS 对北京地区 2 m 气温预报集成结果 (RMAPS-ST 和 ECMWF-IFS 代表原始预报数据, EWA 代表等权重集成, GLAR 代表线性回归集成, XGB 代表机器学习方法集成; “×”表示均值; a. 2017 年 8 月, b. 2018 年 8 月, c. 2019 年 1 月, d. 2019 年 2 月)

Fig. 4 RMAPS-ST and ECMWF-IFS were used to integrate the results of 2 m temperature prediction in Beijing (The red solid round scatters show RMSEs of the original prediction data of RMAPS-ST, the green solid triangular scatter points indicate RMSEs of the original forecast data of ECMWF-IFS, the yellow solid triangular line shows the RMSE of the result of equal-weight integration, the blue star solid line is the RMSE of linear regression integration results, the solid line with red dots is the RMSE of the integration result of the XGBoost machine learning method. “×” represents the mean. a. August 2017, b. August 2018, c. January 2019, d. February 2019)



续图 4

Fig. 4 Continued

表 1 缩写词及含义

Table 1 Abbreviations and meanings

指标	缩写词	含义
等权重平均误差	EWA_AVG	24个时刻等权重误差平均值
线性回归平均误差	GLAR_AVG	24个时刻线性回归误差平均值
XGBoost平均误差	XGB_AVG	24个时刻XGBoost误差平均值
原始最优预报平均误差	ORI_MIN_AVG	24个时刻对应的原始最优成员预报误差平均值
等权重平均提高百分比	EWA_RATE	$\frac{(ORI_MIN_AVG - EWA_AVG)}{ORI_MIN_AVG} \times 100\%$
线性回归提高百分比	GLAR_RATE	$\frac{(ORI_MIN_AVG - GLAR_AVG)}{ORI_MIN_AVG} \times 100\%$
XGBoost提高百分比	XGB_RATE	$\frac{(ORI_MIN_AVG - XGB_AVG)}{ORI_MIN_AVG} \times 100\%$

表 2 三种方法对地面 2 m 气温集成预报的结果

Table 2 Results of integrated surface 2 m air temperature prediction by three methods

测试集	RMSE每时刻平均值(°C)				集成预报RMSE相对于原始RMSE最小值的变化百分比(%)		
	ORI_MIN_AVG	EWA_AVG	GLAR_AVG	XGB_AVG	EWA_RATE	GLAR_RATE	XGB_RATE
201708	1.95	1.89	1.67	1.60	3.20	14.06	18.09
201808	1.80	1.72	1.63	1.48	4.00	9.12	17.64
201901	2.73	2.78	2.67	2.43	-2.05	2.12	11.02
201902	2.16	2.25	2.08	1.91	-3.97	3.64	11.39

原集成误差无明显优势(图 5c)。冷季 2019 年 2 月 00—15 时平原误差相对较小,其他时刻平原集成误差无明显优势(图 5d)。总的来说,分地形训练后山区与平原间误差差距不大,平原误差相对较小。

3.4.2 2 m 相对湿度

图 6 的 2 m 相对湿度观测值变化趋势呈“V”型,与冷季误差(图 6c、d)变化趋势相似,暖季误差(图 6a、b)相对分散,变化不明显。整体来看,采用不同方法集成以后模式预报的 2 m 相对湿度均方

根误差明显降低。从暖季的均方根误差趋势可以看出,XGBoost 和线性回归集成对数值预报的集成订正效果非常好,且三者中 XGBoost 集成效果优于线性回归,线性回归优于等权重平均。冷季,3 种方法的集成效果基本持平,均方根误差趋势与原始预报的误差趋势相同,且冷季日变化呈明显的“V”型。冷季 1 月(图 6c)07 时为均方根误差低谷(4.42%),23 时为误差峰值(17.93%);冷季 2 月(图 6d)07 时为均方根误差低谷(7.13%),00 时为误差峰值

(23.10%)。相比之下,冷季集成以后的日变化比暖季明显。

表 3 是 2 m 相对湿度采用 3 种方法集成前后的

误差变化情况。可以看出, XGBoost 集成在 2018 年暖季的预报效果相对于原始预报数据提高 15.48%, 而线性回归在 2017 年暖季提高 8.57%; 冷季 XGBoost

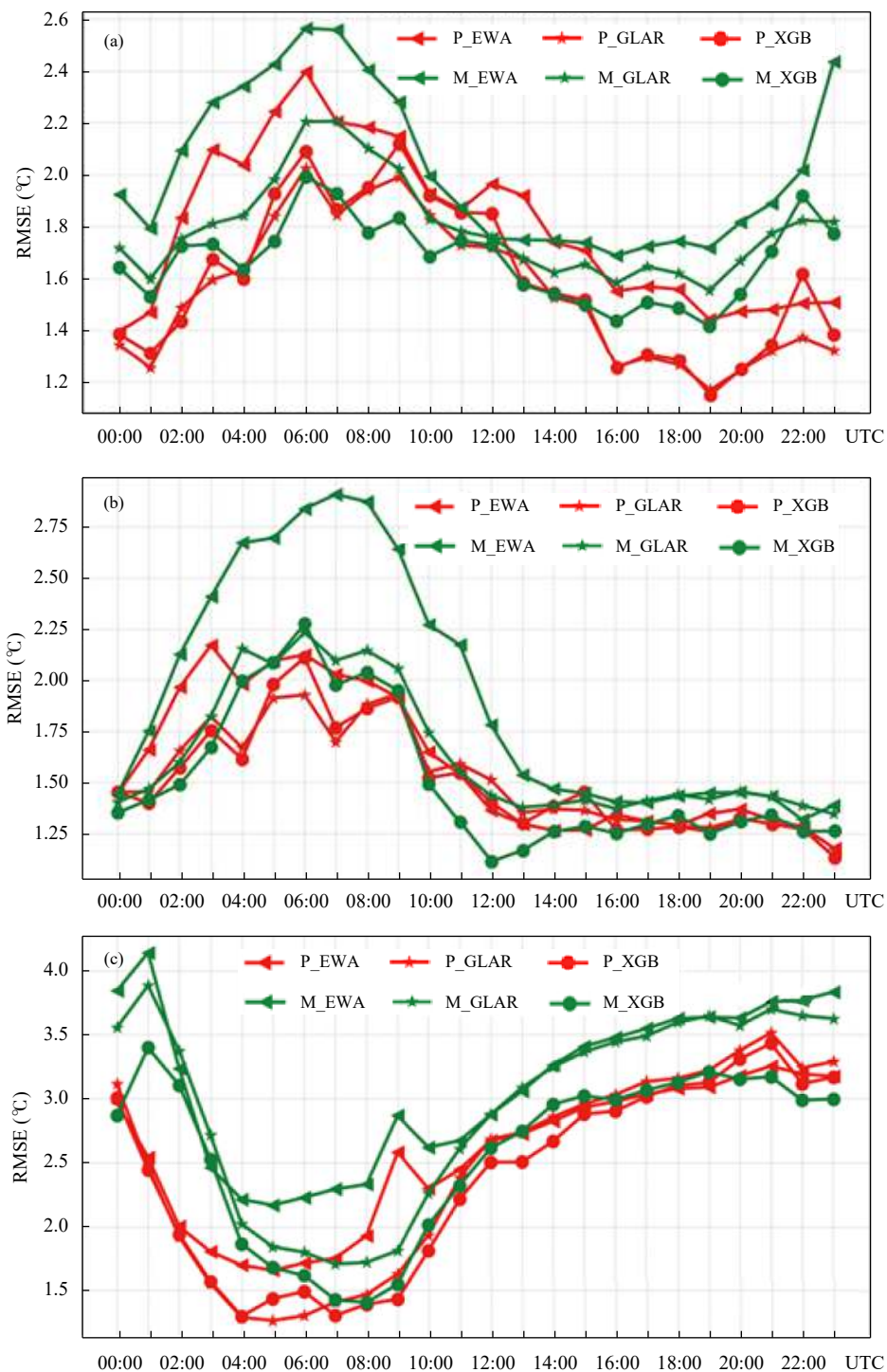
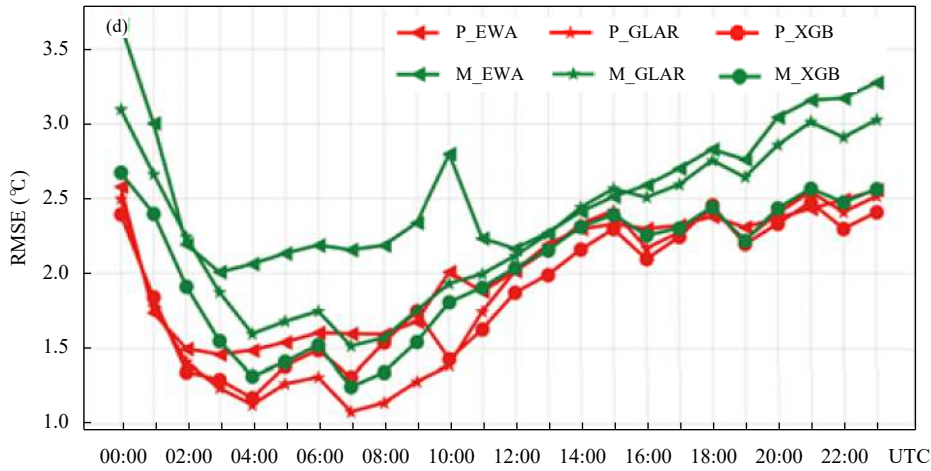


图 5 2 m 气温集成误差 (P_代表平原, M_代表山区; 其他同图 4)

Fig. 5 Integration errors of 2 m temperature training in mountain and plain (The P_ represent plains, and M_ represent mountainous; Others are the same as Fig. 4)



续图 5

Fig. 5 Continued

集成相较于传统线性回归集成方法优势较小, 1月 XGBoost 集成提高 7.15%, 线性回归集成方法提高

4.6%; 2月 XGBoost 集成提高 6.80%, 线性回归集成方法提高 5.90%。等权重平均分配集成在暖季和

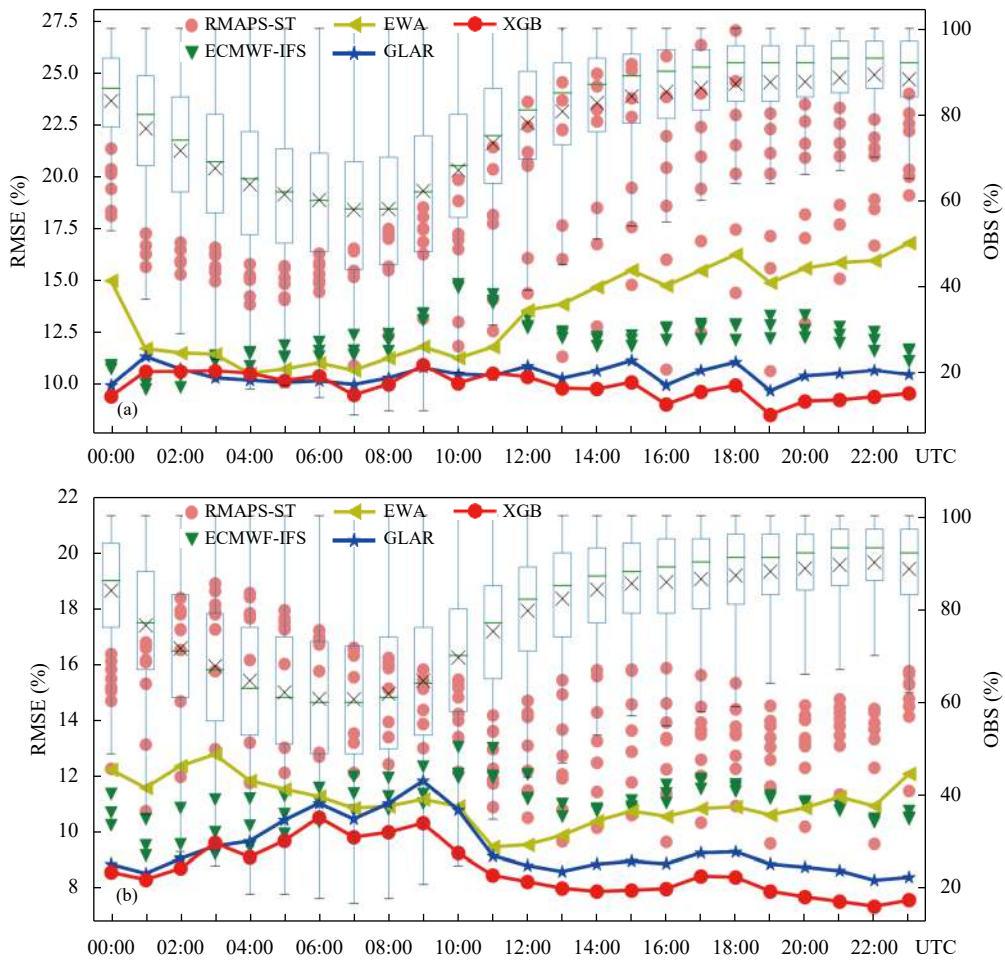
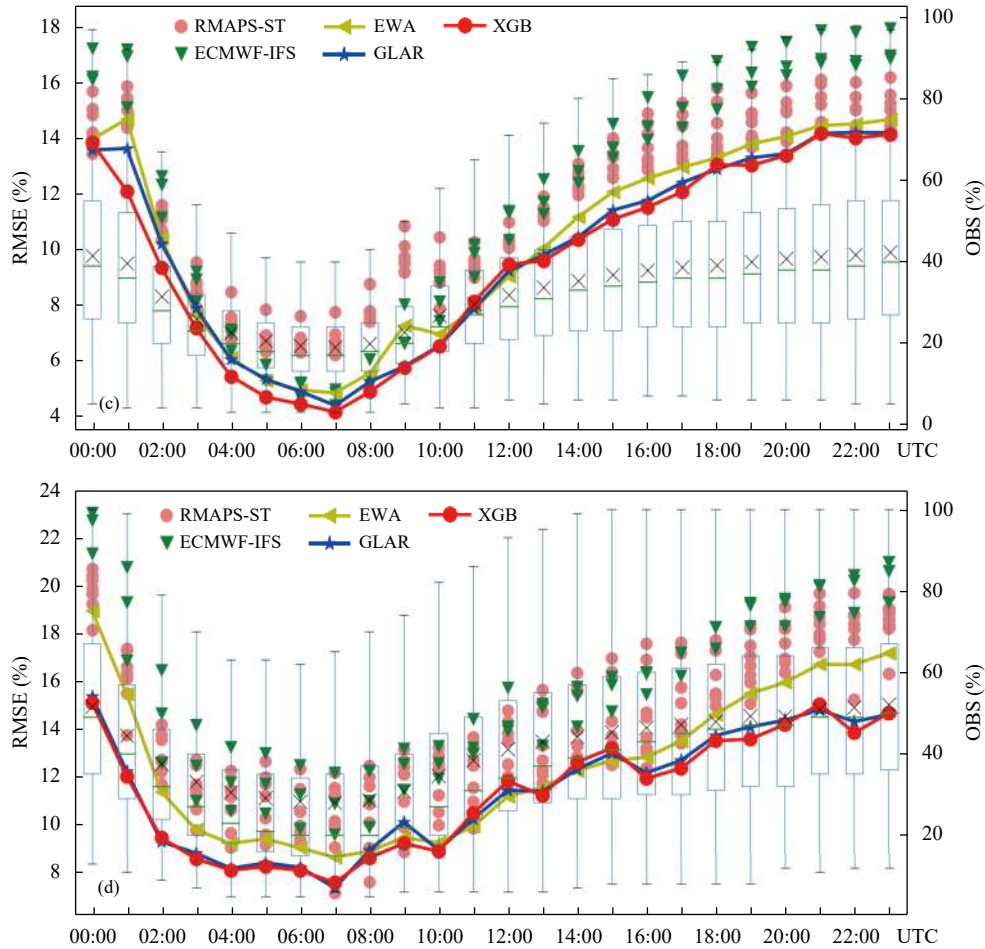


图 6 同图 4, 但为 2 m 相对湿度

Fig. 6 Same Fig. 4 but for 2 m relative humidity



续图 6

Fig. 6 Continued

表 3 三种方法对 2 m 相对湿度集成预报的结果

Table 3 Results of integrated 2 m relative humidity prediction by three methods

测试集	RMSE的每时刻平均值(%)				集成预报RMSE相对于原始RMSE最小值的变化百分比(%)		
	ORI_MIN_AVG	EWA_AVG	GLAR_AVG	XGB_AVG	EWA_RATE	GLAR_RATE	XGB_RATE
201708	11.48	13.44	10.49	9.94	-17.10	8.57	13.36
201808	10.21	11.05	9.41	8.63	-8.24	7.90	15.48
201901	10.43	10.34	9.95	9.68	0.83	4.60	7.15
201902	12.19	12.52	11.47	11.36	-2.70	5.90	6.80

冷季 2 月逊于原始预报中的最优预报。

为了进一步研究相对湿度集成误差, 分别对北京山区和平原进行集成试验(图 7)。2017 年暖季 12—23 时平原的 3 种集成误差相对较小, 其他时刻山区误差相对较小, 但差异不明显(图 7a)。2018 年暖季 11—23 时山区的集成误差相对较小, 其他时刻山区与平原的误差基本一致(图 7b)。冷季 1 月 12—23 时山区集成误差相对较小明显, 其他时刻平原与山区误差基本一致(图 7c)。冷季 2 月 07—

13 时平原误差相对较小, 其他时刻平原集成误差无明显优势(图 7d)。总的来说, 分地形训练后山区与平原误差差距不大。

3.4.3 10 m 风速

10 m 风速观测值变化趋势整体平缓, 与误差趋势基本一致; 集成误差结果显示风速的日变化没有随季节变化发生较大波动, 保持平稳(图 8)。其中, ECMWF-IFS 对 10 m 风速的原始预报明显优于 RMAPS-ST, 等权重平均分配集成效果在两种原始

数值预报之间。XGBoost 和线性回归分配集成方法效果明显优于原始预报, 并且 XGBoost 集成方法最好, 暖季误差在 1.0 m/s 以下, 2017 年的 21 时和 2018 年的 13 时分别为 0.69 (图 8a) 和 0.67 m/s

(图 8b)。

从表 4 的结果可以看出, 2017 年暖季 XGBoost 集成方法相对于原始最优预报数据的预报能力提高 33.22%, 线性回归方法提高 26.60%, 2018 年暖

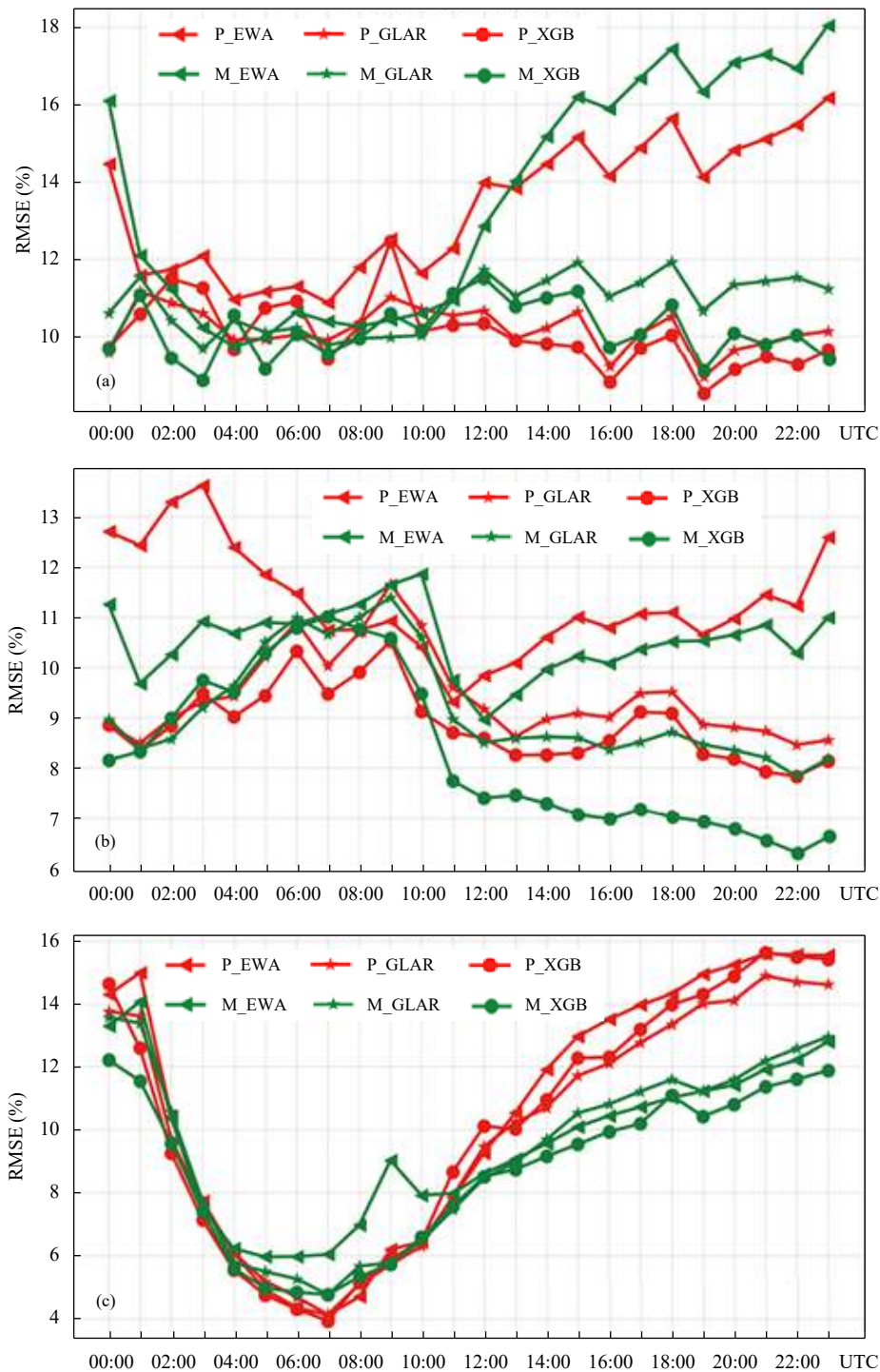
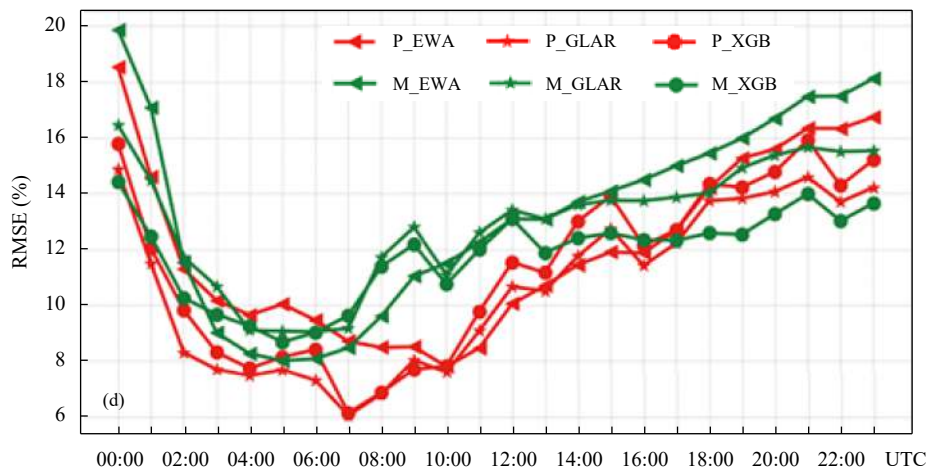


图 7 同图 5, 但为 2 m 相对湿度

Fig. 7 Same Fig. 5 but for 2 m relative humidity



续图 7

Fig. 7 Continued

季与 2017 年暖季结果基本相同。冷季的集成结果相对于暖季较差, 1 月 XGBoost 方法提高 32.19%,

线性回归方法提高 27.47%; 2 月 XGBoost 方法提高 31.23%, 线性回归方法提高 24.79%。由于两种

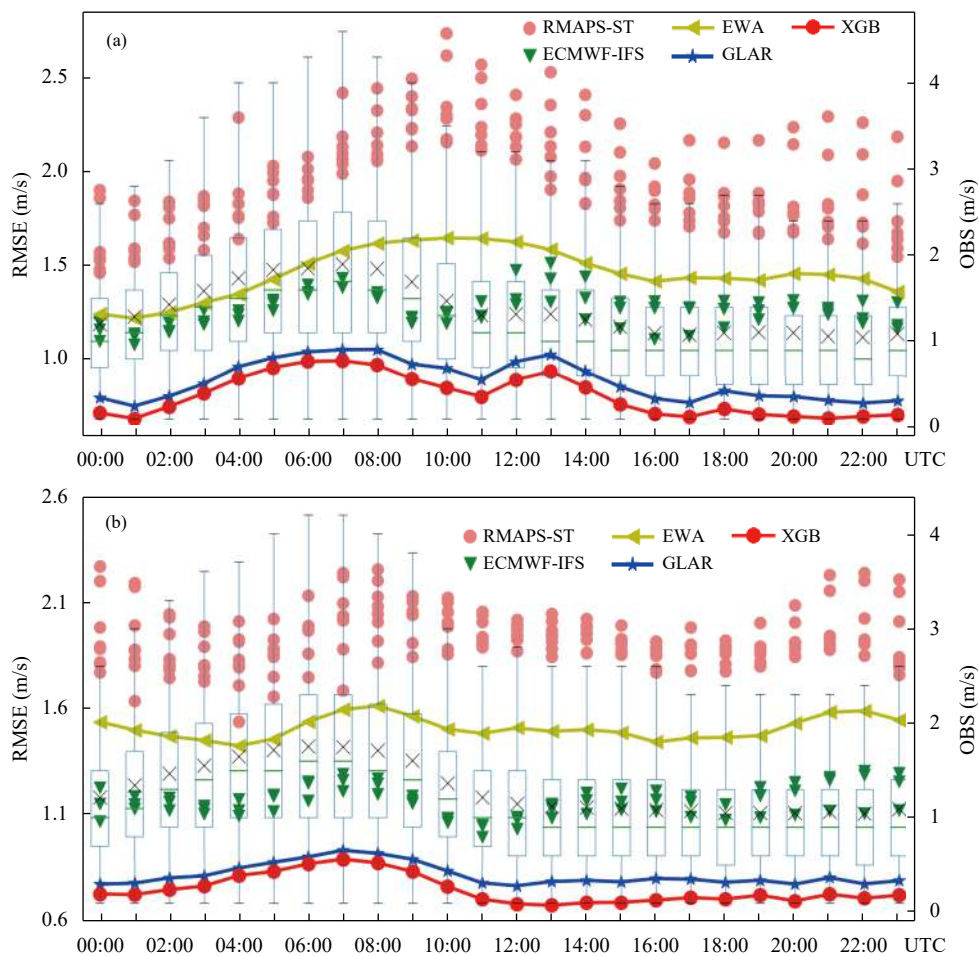
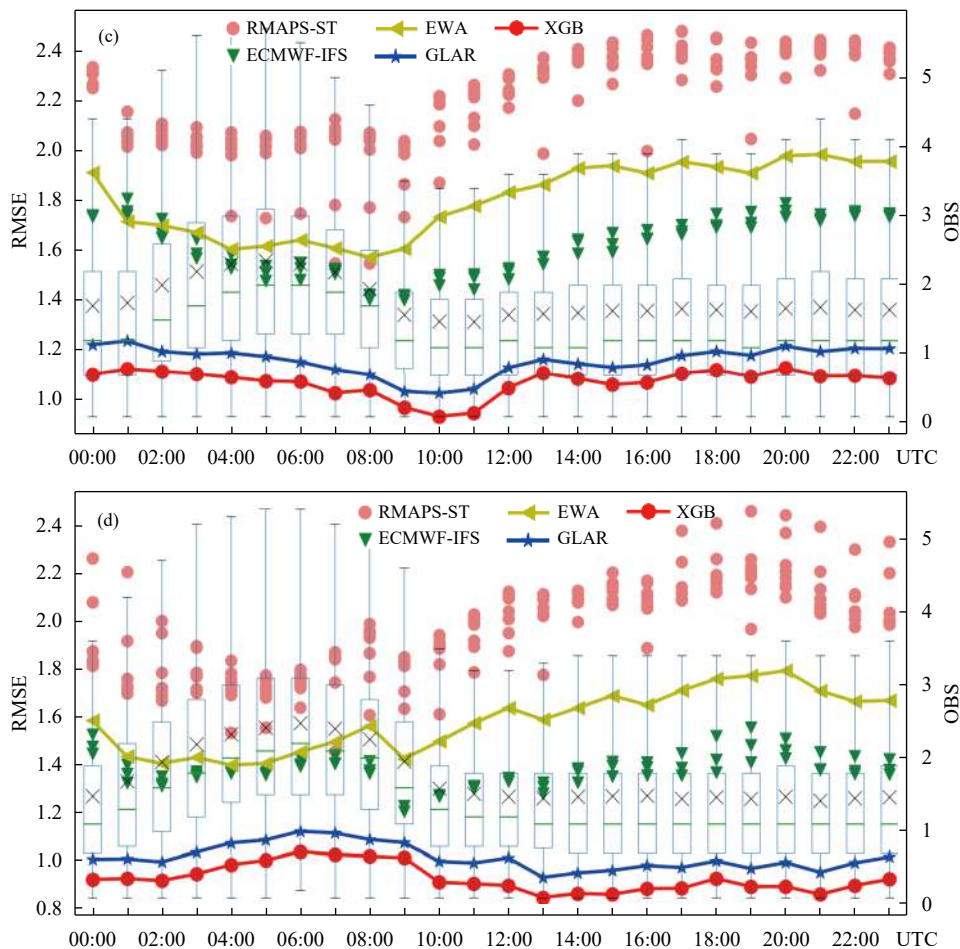


图 8 同图 4, 但为 10 m 风速

Fig. 8 Same Fig. 4 but for 10 m wind speed



续图 8

Fig. 8 Continued

表 4 三种方法 10 m 风速集成预报结果

Table 4 Results of integrated prediction of 10 m wind speed by three methods

测试集	RMSE的每时刻平均值(m/s)				集成预报RMSE相对于原始RMSE最小值的变化百分比(%)		
	ORI_MIN_AVG	EWA_AVG	GLAR_AVG	XGB_AVG	EWA_RATE	GLAR_RATE	XGB_RATE
201708	1.23	1.49	0.90	0.82	-21.06	26.60	33.22
201808	1.12	1.54	0.83	0.75	-36.72	26.60	33.19
201901	1.60	1.82	1.16	1.07	-13.41	27.47	32.91
201902	1.36	1.60	1.03	0.94	-17.60	24.79	31.23

模式的性能差异较大, ECMWF-IFS 的误差明显低于 RMAPS-ST, 等权重平均集成以后的误差介于两个模式原始预报误差之间。因而, 等权重集成相对于原始模式的最优预报 (ECMWF-IFS 预报) 而言, 增大了误差, 但采用线性回归和 XGBoost 可以很好地解决该问题, 集成以后的预报优于原始最优预报 (ECMWF-IFS 预报), 其中 XGBoost 集成效果更佳。

分山区和平原的集成试验得到的风速集成误差如图 9 所示。2017 和 2018 年暖季, 除个别时刻

(2017 年的 10—11 时、2018 年的 03—12 时) 外, 平原误差明显低于山区 (图 9a、b)。冷季 1 和 2 月的平原误差也明显低于山区 (图 9c、d)。

3.4.4 10 m 风向

使用订正后的 10 m 东西 (U) 和南北 (V) 风分量进行风向集成。从图 10 风向观测值可以看出, 风向在 05—09 时较集中, 风向预报误差较小, 其他时刻风向值较分散, 误差较大。参考中国气象局的风向检验标准《QX/T 229—2014 风预报检验方法》,

对风向使用平均绝对误差进行结果检验。从图 10 可见, RMAPS-ST 和 ECMWF-IFS 原始数值预报误差基本相同, 且两种原始预报数据有明显“V”型日变化。2017 年暖季 08 时的 RMAPS-ST 原始预报最小误差为 64.69° , XGBoost 可使最小误差降低到

57.90° (图 10a), 2018 年暖季 08 时的 ECMWF-IFS 原始预报最小误差为 63.97° , XGBoost 可使最小误差降低到 61.41° (图 10b), 冷季 1 月 07 时的 RMAPS-ST 原始预报最小误差为 60.0° (图 10c), 冷季 2 月 08 时的 RMAPS-ST 原始预报最小误差为 56.77° ,

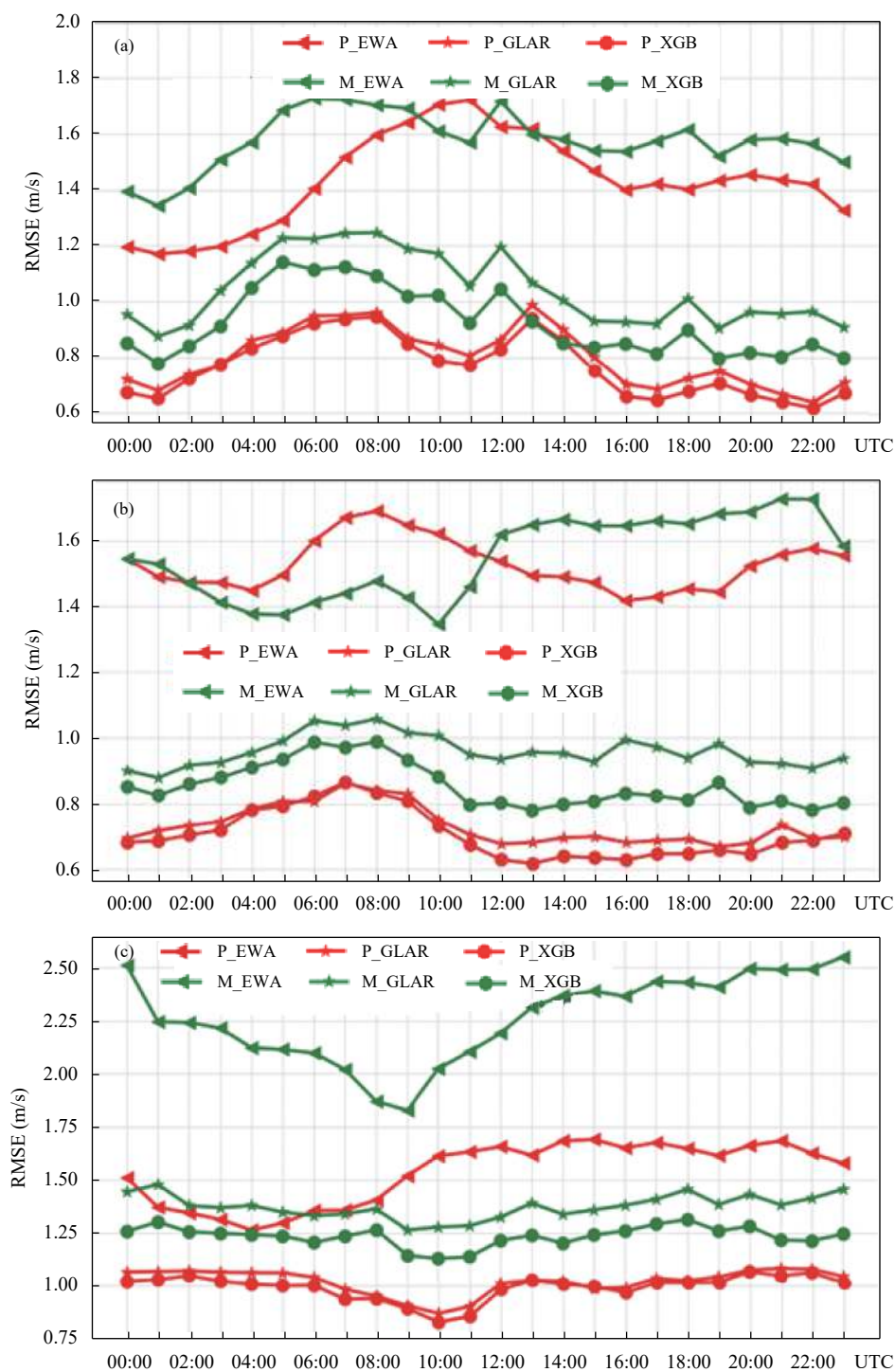
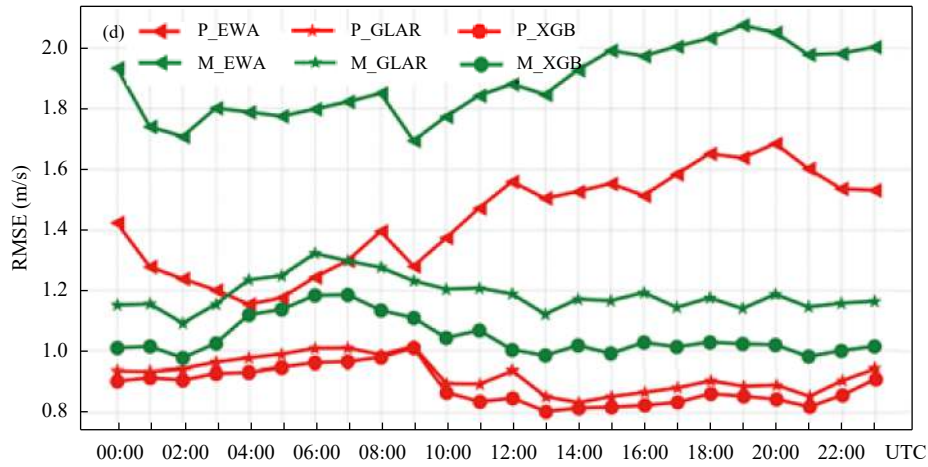


图 9 同图 5, 但为 10 m 风速

Fig. 9 Same Fig. 5 but for 10 m wind speed



续图 9

Fig. 9 Continued

XGBoost可使最小误差降低到 52.14°(图 10d)。

相对于原始风向最优预报在 2018 年暖季和冷季

从表 5 风向的误差统计可以看出,线性回归法

2 月误差增大。XGBoost 相对于线性回归和等权重

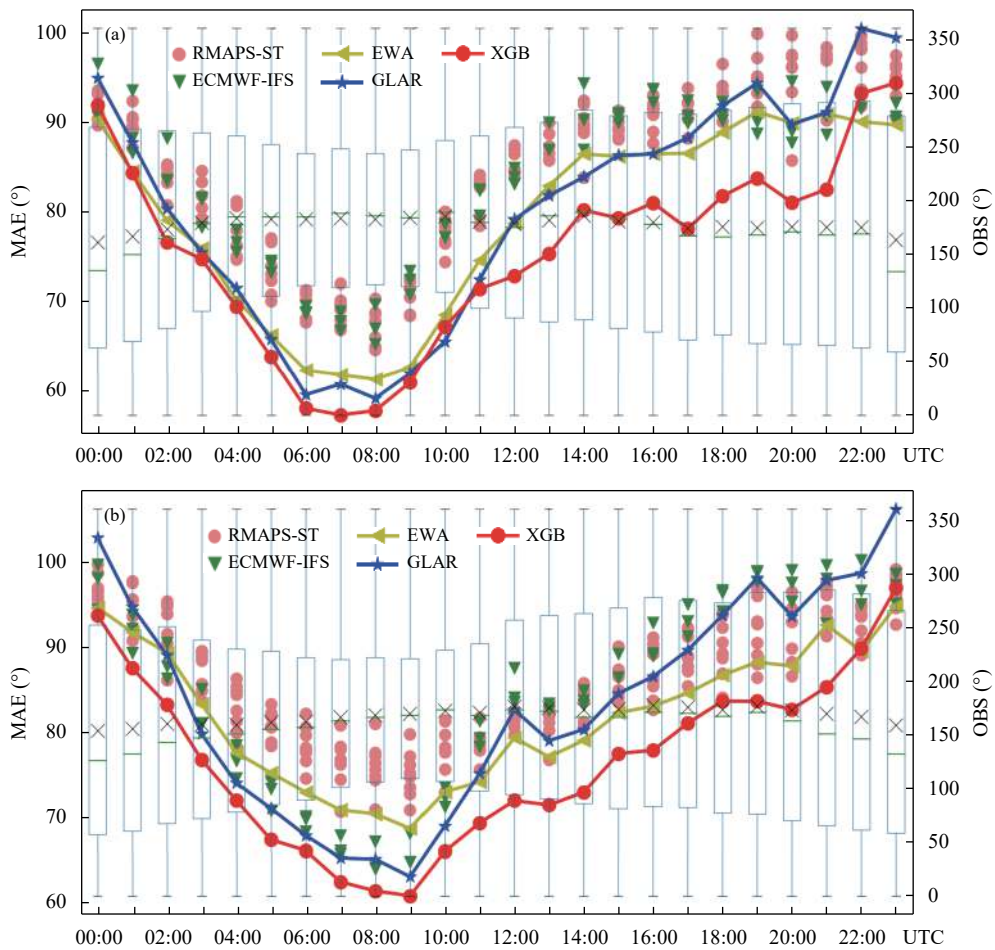
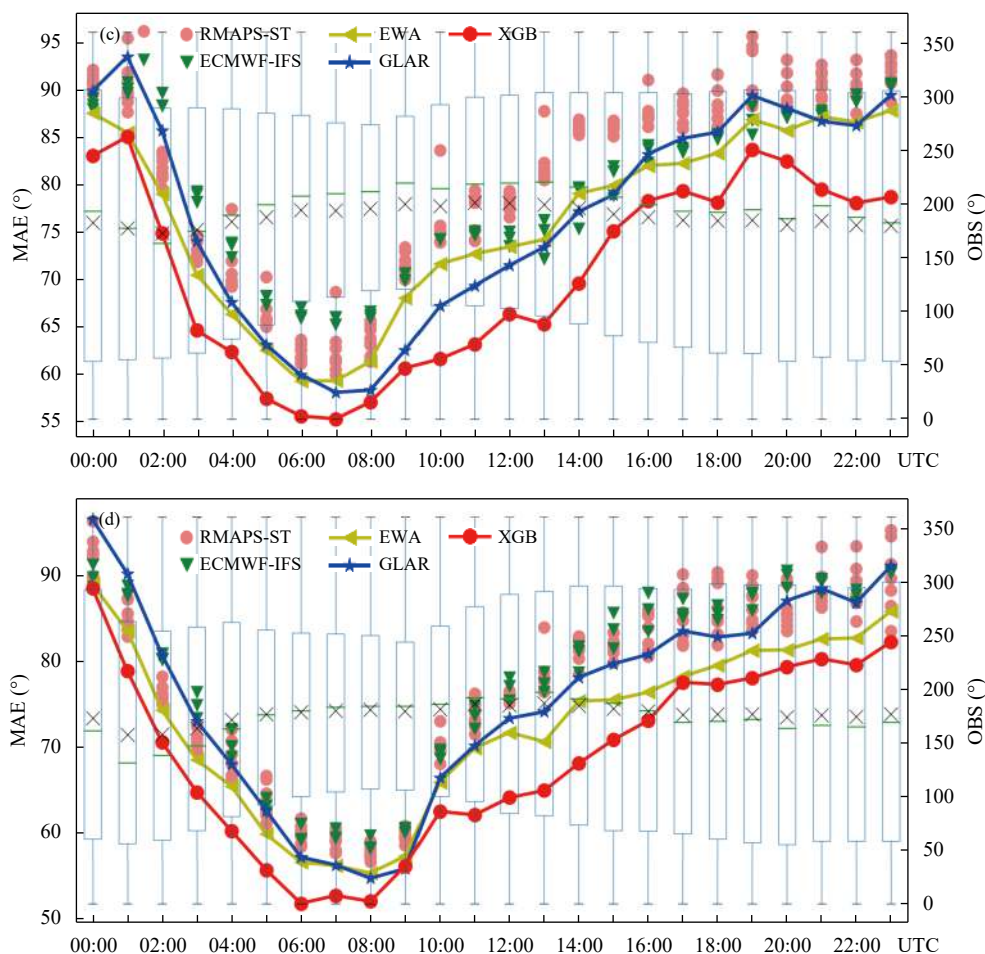


图 10 同图 4, 但为 10 m 风向

Fig. 10 Same Fig. 4 but for 10 m wind direction



续图 10

Fig. 10 Continued

表 5 三种方法对 10 m 风向集成预报的结果

Table 5 Results of integrated 10 m wind direction prediction by three methods

测试集	MAE的每时刻平均值(°)				集成预报MAE相对于原始MAE最小值的变化百分比(%)		
	ORI_MIN_AVG	EWA_AVG	GLAR_AVG	XGB_AVG	EWA_RATE	GLAR_RATE	XGB_RATE
201708	81.19	79.41	80.36	75.74	2.18	1.02	6.71
201808	80.03	81.97	83.67	76.74	-2.43	-4.56	4.10
201901	77.03	76.37	76.85	70.69	0.85	0.23	8.23
201902	74.67	72.69	75.93	68.88	2.65	-1.69	7.75

法有明显优势, XGBoost 在 2017 年暖季预报效果提高 6.71%, 等权重提高 2.18%。2018 年暖季 XGBoost 提高 4.1%, 等权重下降了 2.43%。冷季 1 月, XGBoost 提高 8.23%, 等权重提高 0.85%。冷季 2 月, XGBoost 提高 7.75%, 等权重提高 2.65%。总的来说 XGBoost 集成效果是最好的。

分山区和平原的集成试验得到的风向集成误差结果如图 11 所示。可以看出, 风向集成以后误差在平原有明显“V”型日变化, 每日 08 时前后误差

最小, 但山区“V”型日变化不再明显, 每日 12—23 时误差变化比较平缓。平原和山区对比来看, 2017 年暖季(图 11a)04—11 时平原集成误差相对较小, 13—23 时山区误差相对较小, 其他时刻平原与山区误差基本一致。2018 年暖季(图 11b)16—23 时山区集成误差相对较小, 与平原误差差距不大。冷季 1 月(图 11c)04—13 时平原的集成误差明显较小, 16—23 时山区等权重和 XGBoost 集成误差相对较小, 其他时刻平原与山区误差无明显差距。

冷季2月(图11d)04—15时平原集成误差相对于山区明显变小,17—23时山区的XGBoost集成误差相对较小。总的来说,在原始预报误差相对较

小的时刻,即“V”型谷底时,平原的集成误差相对较小,原始预报误差相对较大时,山区误差相对较小。

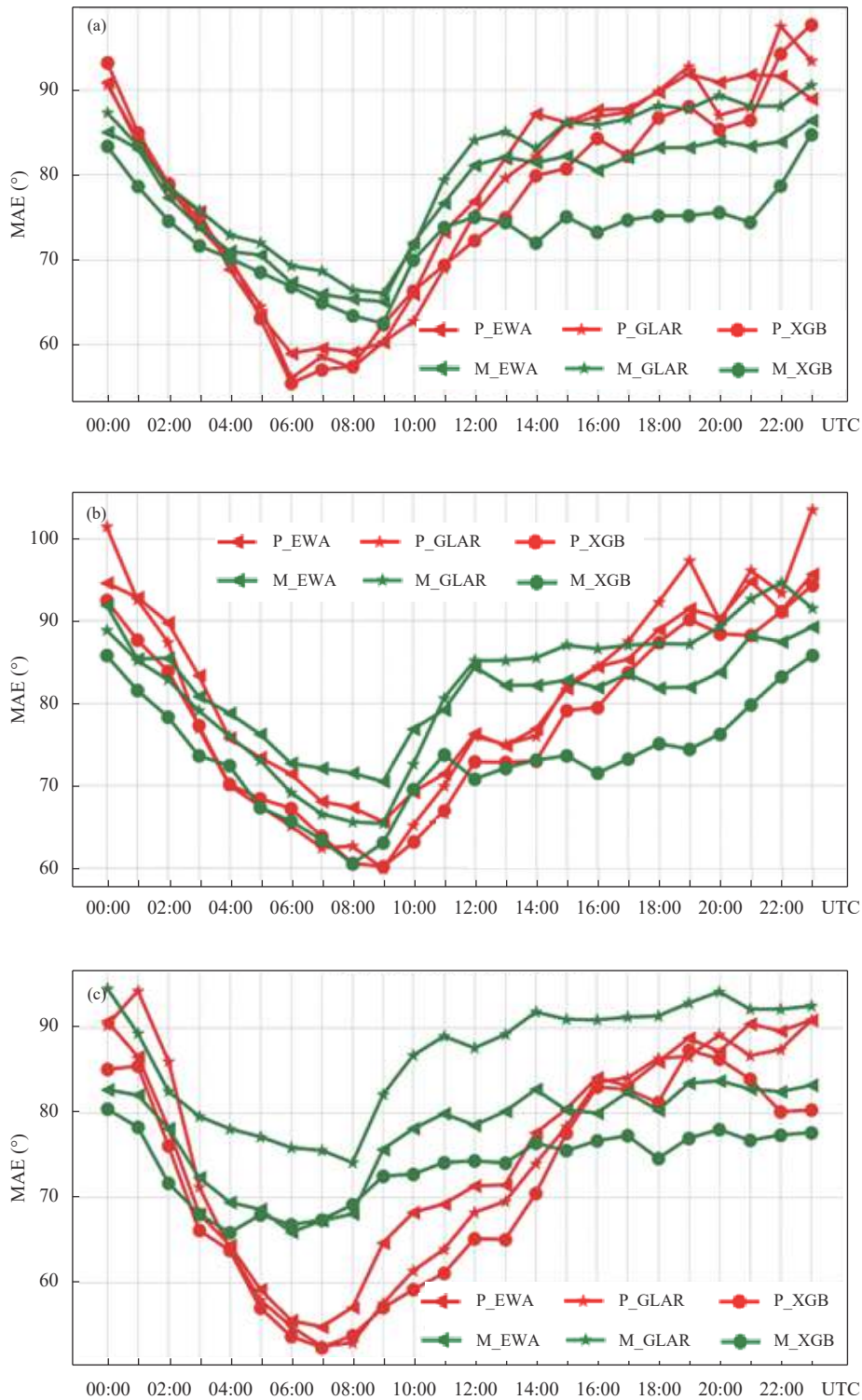
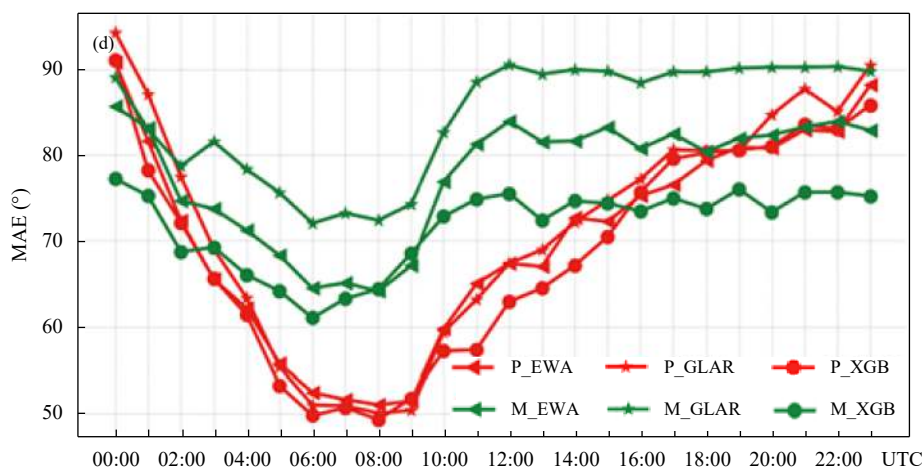


图 11 同图 5, 但为 10 m 风向
 Fig. 11 Same Fig. 5 but for 10 m wind direction



续图 11

Fig. 11 Continued

4 结论与讨论

采用传统统计集成方法(等权重平均法、线性回归法)和机器学习方法(XGBoost)对 2 m 气温、2 m 相对湿度、10 m 风速、10 m 风向进行了集成预报试验,结果表明:

(1)传统集成方法中等权重集成能较好地拟合数值预报成员,但是基本不会超过最优的单一数值预报成员。对于 10 m 风速来说,等权重集成的误差最大,平均误差为 1.49—1.82 m/s,订正效果相对不佳。

(2)机器学习方法 XGBoost 对常规气象要素(2 m 气温、10 m 风速)来说有较好的集成预报效果,其结果明显优于两种传统集成方法。对 2 m 气温集成的误差相对于原始预报数据减小 11.02%—18.09%, 10 m 风速集成误差相对于原始预报数据减小 31.23%—33.22%。对于冷季 2 m 相对湿度的集成, XGBoost 与等权重和线性回归方法相当。使用 XGBoost 对订正后的 U 、 V 风分量合成风向预报后,结果明显优于传统方法,误差减小 4.1%—8.23%。

(3)在分地形试验时, 2 m 气温和相对湿度在山区和平原误差差距不大,平原集成误差相对较小。10 m 风速的平原集成误差明显低于山区的集成误差。当 10 m 风向的原始预报误差相对较小时,平原集成误差明显小于山区,原始预报误差相对较大时,山区的集成误差相较于平原小。

XGBoost 之所以表现相对最优,是因为本质上 XGBoost 是树形算法的集合,是在特征维度上对样

本进行处理,训练过程中对每一个特征进行划分比较,最终训练出最优的切分特征,训练好的树形算法更接近于人类的思维方式,而传统等权重和线性回归法是在样本维度上对所有特征分配一个权重从而得到最终结果。

ECMWF-IFS 原始预报误差常小于 RMAPS-ST 的原始预报误差,如果单独使用 ECMWF-IFS 预报进行集成可能会更优,但为了增加空间和时间上的样本量(RMAPS-ST 每 3 h 更新循环,ECMWF-IFS 每天 00、12 时循环预报),也为了增加不同数据源(不同模式)以符合机器学习对大数据多样性特征(离散特征)的基本要求,以构建更为科学合理、误差特征更稳定的数值预报集成订正模型,所以采用了两个完全独立的数值预报模式(全球模式和区域模式),并采集了每个模式在有效预报时间内的全部预报,从而构建出用于集成试验的原始数值预报数据集。

通过常规气象要素模式集成试验可以看出, XGBoost 体现了机器学习的优势,对于多数气象要素而言, XGBoost 误差比传统集成方法明显降低。为了进一步验证试验结果的可靠性,在今后的试验中,应采用细致划分地形以及进行更全面的滚动集成模型构建和对比试验。

参考文献

- 黄威,牛若芸. 2017. 基于集合预报和支持向量机的中期强降雨集成预报试验. *气象*, 43(9): 1110-1116. Huang W, Niu R Y. 2017. The medium-term multi-model integration forecast experimentation for heavy rain

- based on support vector machine. *Meteor Mon*, 43(9): 1110-1116 (in Chinese)
- 王在文, 陈敏, Monache L D 等. 2019. 相似集合预报方法在北京区域地面气温和风速预报中的应用. *气象学报*, 77(5): 869-884. Wang Z W, Chen M, Monache L D, et al. 2019. Application of analog ensemble method to surface temperature and wind speed prediction in Beijing area. *Acta Meteor Sinica*, 77(5): 869-884 (in Chinese)
- 杨璐, 韩丰, 陈明轩等. 2018. 基于支持向量机的雷暴大风识别方法. *应用气象学报*, 29(6): 680-689. Yang L, Han F, Chen M X, et al. 2018. Thunderstorm gale identification method based on support vector machine. *J Appl Meteor Sci*, 29(6): 680-689 (in Chinese)
- 赵声蓉. 2006. 多模式温度集成预报. *应用气象学报*, 17(1): 52-58. Zhao S R. 2006. Multi-model consensus forecast for temperature. *J Appl Meteor Sci*, 17(1): 52-58 (in Chinese)
- Chang H L, Yuan H L, Lin P L. 2012. Short-range (0-12 h) PQPFs from time-lagged multimodel ensembles using LAPS. *Mon Wea Rev*, 140(5): 1496-1516
- Chen M Y, Wang W Q, Kumar A. 2013. Lagged ensembles, forecast configuration, and seasonal predictions. *Mon Wea Rev*, 141(10): 3477-3497
- Chen T Q, Guestrin C. 2016. XGBoost: A scalable tree boosting system//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, California: ACM
- Friedman J H. 2001. Greedy function approximation: A gradient boosting machine. *Ann Statist*, 29(5): 1189-1232
- Gagne II D J, McGovern A, Xue M. 2014. Machine learning enhancement of storm-scale ensemble probabilistic quantitative precipitation forecasts. *Wea Forecasting*, 29(4): 1024-1043
- Glahn H R, Lowry D A. 1972. The use of model output statistics (MOS) in objective weather forecasting. *J Appl Meteor*, 11(8): 1203-1211
- Haberlie A M, Ashley W S. 2018a. A method for identifying midlatitude mesoscale convective systems in radar mosaics. Part I: Segmentation and classification. *J Appl Meteor Climatol*, 57(7): 1575-1598
- Haberlie A M, Ashley W S. 2018b. A method for identifying midlatitude mesoscale convective systems in radar mosaics. Part II: Tracking. *J Appl Meteor Climatol*, 57(7): 1599-1621
- Han L, Sun J Z, Zhang W, et al. 2017. A machine learning nowcasting method based on real-time reanalysis data. *J Geophys Res: Atmos*, 122(7): 4038-4051
- Huang L X, Isaac G A, Sheng G. 2012. Integrating NWP forecasts and observation data to improve nowcasting accuracy. *Wea Forecasting*, 27(4): 938-953
- Lu C G, Yuan H L, Schwartz B E, et al. 2007. Short-range numerical weather prediction using time-lagged ensembles. *Wea Forecasting*, 22(3): 580-595
- McGovern A, Elmore K L, Gagne II D J, et al. 2017. Using artificial intelligence to improve real-time decision-making for high-impact weather. *Bull Amer Meteor Soc*, 98(10): 2073-2090
- Monache D L, Eckel F A, Rife D L, et al. 2013. Probabilistic weather prediction with an analog ensemble. *Mon Wea Rev*, 141(10): 3498-3516
- Mylne K R, Evans R E, Clark R T. 2002. Multi-model multi-analysis ensembles in quasi-operational medium-range forecasting. *Quart J Roy Meteor Soc*, 128(579): 361-384
- Sprenger M, Schemm S, Oechslein R, et al. 2017. Nowcasting foehn wind events using the AdaBoost machine learning algorithm. *Wea Forecasting*, 32(3): 1079-1099
- Stensrud D J, Bao J W, Warner T T. 2000. Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Mon Wea Rev*, 128(7): 2077-2107
- Ware C, Knight W, Wells D. 1991. Memory intensive statistical algorithms for multibeam bathymetric data. *Comput Geosci*, 17(7): 985-993
- Whan K, Schmeits M. 2018. Comparing area probability forecasts of (extreme) local precipitation using parametric and machine learning statistical postprocessing methods. *Mon Wea Rev*, 146(11): 3651-3673
- Yuan H L, Gao X G, Mullen S L, et al. 2007. Calibration of probabilistic quantitative precipitation forecasts with an artificial neural network. *Wea Forecasting*, 22(6): 1287-1303
- Yuan H L, Lu C G, McGinley J A, et al. 2009. Evaluation of short-range quantitative precipitation forecasts from a time-lagged multimodel ensemble. *Wea Forecasting*, 24(1): 18-38