

陈鹤,蔡荣辉,陈静静,等,2022. 基于深度学习方法的温度预报技术应用与评估[J]. 气象,48(11):1373-1383. Chen H, Cai R H, Chen J J, et al, 2022. Application and evaluation of temperature forecast based on deep learning method[J]. Meteor Mon, 48 (11):1373-1383(in Chinese).

基于深度学习方法的温度预报技术应用与评估*

陈鹤^{1,2} 蔡荣辉^{1,2} 陈静静^{1,2} 傅承浩^{1,2} 周莉^{1,2} 陈龙^{1,2}

1 湖南省气象台,长沙 410118

2 气象防灾减灾湖南省重点实验室,长沙 410118

提 要: 基于 2017—2019 年欧洲中期天气预报中心的全球预报系统(ECMWF-IFS),结合对应时间段站点观测实况,采用深度学习建立了多层全连接神经网络模型(简称 DL 模型),对未来 84 h 的地面气温进行订正预报。使用湖南省 2020 年全年的预报结果进行对比检验评估,得到以下结论:从空间分布来看,DL 模型平均均方根误差(RMSE)在大部分地区为 1.5~2.0℃(全省平均 RMSE 为 1.78℃),其对 ECMWF-IFS 模式的订正效果明显,尤其是在高海拔地区,改善率随着预报站点海拔的增加而上升;数值预报的 RMSE 有明显的日变化特征,每日午后误差最大,DL 模型的改善幅度也最大,日出前时效误差最小,改善率在不同海拔高度有不同的日变化特征;DL 模型每个月相对 ECMWF-IFS 的平均 RMSE 都有明显的订正效果,其中 10 月、11 月 RMSE 改善率最高,12 月最低;从 2020 年年底一次寒潮过程的日最高、最低及逐 3 h 气温检验效果来看,DL 模型对 ECMWF-IFS 的系统误差有明显的订正效果,在单站的曲线上,也能看出 DL 模型更接近观测实况。评估结果表明:模型可以显著地减小 ECMWF-IFS 的预报误差,其输出结果基本满足日常预报业务的需求。

关键词: 温度预报,深度学习,应用与评估

中图分类号: P456, P457

文献标志码: A

DOI: 10.7519/j.issn.1000-0526.2022.070101

Application and Evaluation of Temperature Forecast Based on Deep Learning Method

CHEN He^{1,2} CAI Ronghui^{1,2} CHEN Jingjing^{1,2} FU Chenghao^{1,2} ZHOU Li^{1,2} CHEN Long^{1,2}

1 Hunan Meteorological Observatory, Changsha 410118

2 Key Laboratory of Preventing and Reducing Meteorological Disaster of Hunan Province, Changsha 410118

Abstract: By using ECMWF-IFS model data and observation data, this study established a fully connected neural network model (DL), tried to make more accurate temperature forecast based on deep learning method for the future 84 hours. The comparative evaluation of ECMWF-IFS and DL was made in 2020. The results show that DL correction can improve the ECMWF-IFS forecasts. The root mean square error (RMSE) of DL in most areas of Hunan is 1.5—2.0℃. There is a significant correction effect for ECMWF-IFS, and the improvement rate increases with the altitude of stations. The diurnal variation is also obvious for the RMSE of ECMWF-IFS/DL model. The RMSE is higher in the afternoon (with higher improvement rate), but lower before sunrise. The improvement rate has different diurnal variation characteristics at different altitudes. DL model shows higher accuracy in the whole year, of which the higher improvement rate appears in October and November (lower in December). In addition, the forecast results in a cold wave process was evaluated. In the accuracy for daily maximum/minimum temperature and the RMSE for

* 湖南省气象局第三期业务能力建设项目(NLJS01)、湖南省气象局重点项目(XQKJ17A004)共同资助

2021 年 8 月 4 日收稿; 2022 年 7 月 1 日收修定稿

第一作者:陈鹤,主要从事客观天气预报技术研发与评估工作. E-mail:chenhe@msn.cn

通讯作者:蔡荣辉,主要从事天气预报技术业务和研究工作. E-mail:mitas@163.com

3 h temperature, DL model shows obvious correction ability for the systematic deviation from ECMWF-IFS. The DL temperature curve for single station is much closer to observation than the ECMWF-IFS forecast. Thus, the model can significantly reduce numerical weather prediction error, and its products could basically meet the demands of daily forecast service.

Key words: temperature forecast, deep learning(DL), application and evaluation

引 言

现代天气预报在数值预报模型的支持下,正在快速发展,但由于资料同化方法、地理分辨率的限制,数值模式的应用仍有较大的提升空间。早期国际上常用模式输出统计(MOS)等技术,使用数值预报与观测数据,通过线性回归方法推导预报方程,评估表明,MOS方法在最高气温等要素预报中能显著减小数值预报的误差,是有效的客观预报技术(Glahn and Lowry, 1972),美国国家气象局将其进行了业务化(Klein and Glahn, 1974)。大气本身是非线性的动力系统,其对初始条件非常敏感,增加初始条件的扰动,生成多成员的集合预报方法,也能有效的优化单模型的预报结果(Leutbecher and Palmer, 2008)。根据我国气象部门智能网格预报发展要求,现已在全国范围推行逐 3 h 精细化网格预报业务。针对数值预报的气温订正,已有较多试验性的研究:智协飞等(2009)分别利用固定训练期超级集合和滑动训练期超级集合方法对北半球中纬度地面气温进行预报试验,发现均有效降低了预报误差,滑动训练期效果更优;李佰平和智协飞(2012)分别利用一元线性回归、多元线性回归、单时效消除偏差和多时效消除偏差平均等订正方法,对 ECMWF-IFS 的 2 m 气温预报结果进行订正,能有效地减小地面气温多个时效预报的误差;王敏等(2012)采用非齐次高斯回归技术对国家气象中心区域集合预报系统的 2 m 温度预报结果进行校准,预报可靠性和预报技巧均显著提高;王婧等(2015)采用平均法、双权重平均法、滑动平均法和滑动双权重平均法对 GRAPES_RAFS 系统 2 m 温度预报产品进行偏差订正,结果表明大部分地区的平均误差、均方根误差都有显著降低;冯慧敏等(2016)利用多模式集合平均、消除偏差集合平均、加权消除偏差集合及多模式超级集合方法,对地面逐 3 h 气温进行多模式集成预报试验,结果表明集成方案对预报效果有不同程度的改善。近几年,基于模式输出统计方法(吴启树

等,2016)、滑动双权重平均订正法和空间误差逐步订正法综合订正技术(薛湛彬等,2019)对 ECMWF-IFS 逐日最高/最低气温进行订正,较好地改善了站点温度的预报质量并应用于业务系统中;曾晓青等(2019)使用温度格点多源融合产品和 ECMWF-IFS 的 2 m 温度预报场,采用简单误差订正、误差回归订正等 8 种方案,进行订正预报试验,对格点、站点的检验结果表明,多种方案都对 3~24 h(预报时间间隔 3 h)预报有明显的订正作用;王丹等(2019)使用递减平均法和一元线性回归法对 ECMWF-IFS 的定时最高(低)气温预报误差进行订正,订正后的预报质量有提升,但在较长预报时效不太理想,部分地区订正技巧为负;郝翠等(2019)采用最优集合预报方法对北京市各站 1~7 d 的日最高气温和日最低气温进行订正,结果表明滑动训练期、优化变量权重两种方案能改善模式固有偏差,其结果优于或与预报员主观预报相当,但在某些局地天气影响下不如主观预报;王在文等(2019)使用相似集合方法对北京快速更新循环数值预报系统的地面要素开展订正试验,并将其与支持向量机的订正效果进行了对比,在有效减小气温预报误差的同时,其计算资源要求更低;盛春岩等(2020)使用气温偏差订正和准对称混合滑动训练期 MOS 方法集成订正 ECMWF-IFS 的气温预报,取得较好的效果,但也发现从第 3 天开始客观方法高温准确率有明显下降趋势;尹珊等(2020)应用滑动平均误差订正方法和历史偏差订正方法,对 ECMWF-IFS 延伸期 2 m 温度预报进行误差订正,两种方法在全年各月都有正订正技巧,其中 1—3 月效果最明显,但无法订正模式对变温时间的预报误差;任萍等(2020)分别基于机器学习方法 XGBoost、等权重平均方法、线性回归方法且考虑地形因子影响,构建出多种模型,对北京地区快速循环同化数值预报系统输出的结果进行订正,在包括气温在内的多要素上 XGBoost 表现最优。

深度学习(deep learning, DL)方法是处理气象大数据的前沿技术。虽然受到一些客观因素的制约,现阶段 DL 技术无法完全学习到大气运动的物

理规律,但累积的数值模式历史预报数据,可处理成一项有效的数据集输入到 DL 模型中。当前,从注重数理原理解释到注重动力热力方程求解与人工智能方法相结合,已经取得了重大进展(许小峰,2018)。图像处理单元(GPU)计算技术对比传统中央处理单元(CPU)具有更高效率,给更大规模、更深层次的数据挖掘提供了可能,因此,基于深度学习方法的订正技术相比传统 MOS 方法,能从数值预报中挖掘出更多时间、空间相关信息。

目前,有学者使用 DL 方法对数值预报结果进行改进。在气温预报的应用上,Li et al(2019)基于模式输出机器学习(MOML)方案,将已有的机器学习回归算法应用到处理后的训练集、测试集中,结果比 ECMWF-IFS 模式结果和 MOS 方法更好,尤其是冬季更明显。陈昱文等(2020)基于 TIGGE 的 ECMWF-IFS 资料,建立深度学习模型,并对部分站点进行气温误差订正,结果表明机器学习在数值模式预报结果订正中具有较大的应用潜力。门晓磊等(2019)使用基于岭回归、随机森林和深度学习的 3 种后处理模型,应用多模式预报结果进行逐日气温预报,对比了多种模型的订正效果,发现 DL 方法在空间分布上效果最好。Han et al(2021)基于 U-Net(为生物医学图像分割而开发的卷积神经网络)提出深度学习模型 CU-Net,对 ECMWF-IFS 的 2 m 温度、2 m 相对湿度、10 m 风的 24~240 h 预报进行格点订正,在多个检验指标上都取得了更优的结果,并指出深度学习方法具有从 NWP 海量数据中直接进行学习从而构建 NWP 偏差特征的能力。在第二届 AI Challenger 全球挑战赛中,使用基于时间序列分析、梯度提升树、深度概率预测等 AI 模型构建的多组集合预报方法,相比之前的相似集合预报方法,显著提高了 2 m 气温的预报水平(嵇磊等,2019)。

由于对数据获取、模型运行的稳定性要求高等原因,大多数深度学习预报方法仅停留在试验阶段(大多数预报为逐日最高/最低气温预报),有的标签、验证数据使用了一日 2 次的模式分析场(或者再分析场)而非实际观测数据,或者使用较为稀疏的气象站,代表性不足,真正建立预报模型并应用到智能网格预报业务且取得较好效果的先例较少。从 2016—2018 年国家级网格指导预报、省市级订正反馈网格预报的气温要素检验来看,国家级网格统计预报模型建立后,预报误差呈现出数值模式、国家级、省市订正逐级减小的趋势,体现出智能预报的巨

大优势(韦青等,2020)。湖南省位于江南地区西部,地形上东、南、西三面环山,北部海拔稍低,东北部洞庭湖地区为其最主要的南下冷空气输入路径。目前针对湖南地区的气温预报,较少进行系统的评估,更无有效的客观订正方法,从日常预报业务检验看,ECMWF-IFS 等数值预报对湖南的气温预报结果存在明显的误差。例如,在湖南西部的雪峰山、西北部的武陵山及南部的南岭部分地区,数值预报平均偏差 3℃以上,部分海拔超过 800 m 的站点平均偏差 4℃以上,东部部分低海拔地区也出现了 3℃以上的偏差。误差并非简单随海拔变化而线性增长,它受到下垫面、坡向,以及各个层次温度平流等多个气象要素的共同影响,无法使用简单的消除偏差方法去除。本文以改善湖南全省逐 3 h 气温预报为目标,建立了精细化气温预报流程,并使用气象站观测资料进行结果验证与评估。

1 资 料

本研究使用湖南省气象信息中心通过 CMA-Cast 接收保存的北半球区域 ECMWF-IFS 高分辨率数值预报模式产品进行训练,模式地面层分辨率为 $0.125^\circ \times 0.125^\circ$,气压层分辨率为 $0.25^\circ \times 0.25^\circ$,选取了地面层所有预报要素,以及 19 个气压层的气温、高度、湿度、风场等要素;训练期为 2017 年 1 月 1 日至 2019 年 12 月 31 日,将 2020 年 1 月 1 日至 12 月 31 日的数据作为测试集进行预报评估分析,模式每日 00 时和 12 时(世界时,下同)起报。基于逐 3 h 精细化气象要素预报工作的需要,本文选择了 0~84 h,一共 27 个预报时效的要素预报数据。使用湖南省全境 421 个气象站(包含国家站和骨干站)逐 3 h(分别为 00、03、06、09、12、15、18、21 时)气温观测数据,数据来源为湖南省气象信息中心 CIMISS 系统,数据总长度为 2017 年 1 月 1 日至 2020 年 12 月 31 日,其中将 2020 年的数据作为测试集进行评估与对比分析。

2 模型与方法

2.1 深度学习模型

本文设计了一种适用于气象要素预测订正的深度学习模型,模型使用了多层全连接神经网络模型对数值预报气温订正函数进行拟合。

多层全连接网络模型是一种经典的神经网络模型,具有输入层、隐藏层和输出层三种结构。其中输入层和输出层有且只有一层神经元结构,隐藏层可以根据计算量的接受程度设计为多层神经元结构。相邻层之间的神经元全部相互连接,并在每个连接上赋予一个计算权重。每层神经元(除输入层)的输入均为上一层神经元的输入,每个神经元的输出均为上层各神经元的值与相应连接权重的相乘后再相加,并经过激活函数计算之后形成的值。经过前期试验,考虑到计算强度和过拟合问题,最终全连接结构使用了 4 个隐藏层,分别设置了 256、128、64 和 32 个神经元,全连接网络结构示意图(图 1)中, H_{ij} 为第 i 个隐藏层的第 j 个神经元;输入数据为某时刻经过筛选的各个气象要素的值 x_i ,输出为各个气象要素的综合提取特征 y 。为避免神经元死亡的问题,本文在初始试验阶段控制、调整了学习率。模型中使用了 Dropout 技术,在神经网络中丢弃部分隐藏神经元(图 1 中表现为虚线圆),能够避免在训练数据上产生复杂的相互适应。为了模拟天气运动的复杂多样,激活函数一般采用非线性函数,本文使用的是线性整流函数(ReLU),ReLU 具有收敛速度快、效率高、能克服梯度消失等优点。在激活函数 ReLU 示意图(图 2)中, x 轴为上一层神经网络的输入量,进行变换后, y 轴为输出的下一层神经元或者神经网络最终结果。

由于天气的演变过程复杂多变,所以模型的选择应该重点考虑提取各个气象要素的时序特征。同时天气预报具有时效性,影响天气变化的气象要素多种多样,所以模型也要考虑输入数据集的大小以及网络训练的时间成本问题。常见的模式后处理方法在对站点要素进行订正预报时,采用单点、多

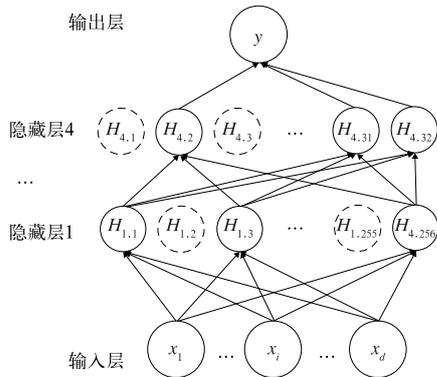


图 1 全连接网络结构示意图
Fig.1 Architecture of the Fully Connected Neural Network

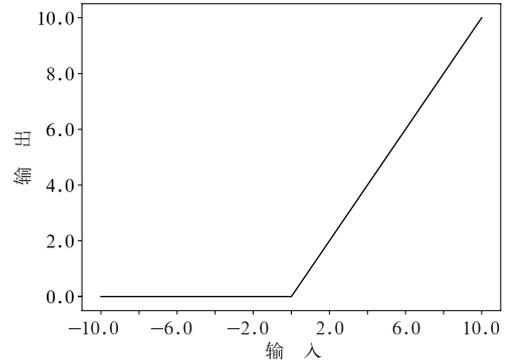


图 2 激活函数 ReLU 示意图
Fig.2 Activation function: ReLU

层、多种气象要素因子进行建模。基于这个思路,本文使用全连接神经网络建立了站点的气温预报模型(下文简称为 DL 模型),从而减少网络训练时间,以达到快速高效的预报订正效果。在建立模型时,分别按单站、区域建立了多类预报模型,以测试集的订正效果自动优选误差低的模型,尽量避免站点因为样本量过少而出现过拟合。

2.2 数据预处理

本文从数值模式训练集中提取了高空多层的位势高度、比湿、温度、风场等要素和地面单层的气温、降水等要素。先使用最近格点法插值到 421 个目标站点,再对数据集进行筛查、清洗等相关预处理,使用气象观测阈值过滤或替换奇异值。最终进行标准化处理:使用 Z-score 算法对数据集进行缩放,公式如下:

$$Z = \frac{x - u}{\sigma} \tag{1}$$

式中: x 为需被标准化的原始数据值, u 为气象要素的总体数据均值, σ 是总体数据的标准差。

然后将数据归一处理到 0~1,公式如下:

$$x^* = \frac{x - X_{\min}}{X_{\max} - X_{\min}} \tag{2}$$

式中: X_{\max} 为样本数据的最大值, X_{\min} 为样本数据的最小值。

本文中使用的每个样本包含因子 198 个(实际训练中有经过调整和筛选),2017—2019 年数据集共有样本 22 457 196 个,使用经过筛查、清洗后的数据进行建模,其中 86% 的数据作为训练集,14% 的数据作为验证集;以收集到的 2020 年实际预报业务数据作为测试集。分配如表 1 所示。

表 1 建立气温预报模型使用的数据集样本数(单位:个)

数据集	训练集	验证集	模型训练集、验证集总计	测试集总计
筛选前	19 313 188	3 144 008	22 457 196	7 463 796
筛选后	19 300 787	3 141 989	22 442 776	7 447 947

Table 1 The number of samples corresponding to temperature forecast model

2.3 检验方法

均方根误差(RMSE)是回归问题的常用性能指标,不仅可以作为深度学习模型中的损失函数,还可以明确地指示订正效果优劣,公式如下:

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

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。

改善率 I 是订正前后的均方根误差的差值相对原预报均方根误差的比值,公式如下:

$$I = \frac{RMSE_{DL} - RMSE_{EC}}{RMSE_{EC}} \quad (4)$$

式中: $RMSE_{EC}$ 、 $RMSE_{DL}$ 分别为 ECMWF-IFS、DL 模型的均方根误差。

3 订正效果

从 2020 年 1 月 1 日开始业务数据输出站点预报订正数据,同时与同期 ECMWF-IFS 输出的 2 m 气温(使用最近格点插值到 421 个站点)进行对比分析。本文重点分析气温的均方根误差以及订正预报的改善率(订正前后均方根误差的差值与订正前的均方根误差之比)。由于业务需要,使用 B-样条基函数插值得到逐 3 h 预报产品,并制作了逐日的最高、最低气温。

3.1 整体误差的空间分布

0~84 h 的 ECMWF-IFS 原始预报的全省平均 RMSE 为 2.16℃(图 3a)。结合湖南省地形可以看出,西北部的武陵山区、西南部的东北—西南向的雪峰山脉周边、南部的南岭山脉及大部分海拔超过 400 m 区域,平均 RMSE 在 2.0℃以上,海拔 100 m 以下大部分地区为 1.5~2.0℃。

DL 模型订正后的全省平均 RMSE 为 1.78℃,从全省地理分布上看,误差值分布较为均匀,大部分地区为 1.5~2.0℃,少量高海拔地区及个别平原地区站点达到 2.0℃以上(图 3b)。

从所有订正时效(0~84 h)进行的均方根误差统计上来看,全省范围均方根误差平均减小了 15.6%。结合地形分析改善率的空间分布(图 3c),改善最明显的地区位于湖南省西部和南部山区,部分高海拔地区达到 20%~30%,湘西南(雪峰山地区)、湘西北(武陵山地区)、南部(南岭地区)等山地站点的幅度最大,少数站点达到 50%~70%;极少数站点(10 个站)改善率为负,但控制在 10%之内。由于数值预报初始场依赖各类观测资料同化,而资料在复杂地形存在系统误差等原因,导致数值预报对某些高海拔地区的预报存在偏差。

经统计发现,随海拔高度的增加,改善率有明显的上升趋势(表 2):海拔 400 m 及以下站点共 347 个,占总站数的 82.4%,平均改善率为 13.5%,但也有 10 个站平均改善率为负(经过深入查看, RMSE 增大是因为 8 月、12 月实时数据源切换出现了单位不一致的问题);海拔 400~800 m 的站点占总站数 12.1%,平均改善率为 19.4%;海拔 1200 m 以上大部分站点改善率达 50%以上。改善率与海拔高度呈正相关,经计算得到相关系数为 0.560。改善率最高的 3 个站为大围山站(72.6%,海拔为 1378 m)、南岳站(69.6%,海拔为 1265.9 m)、雪峰山站(67.8%,海拔为 1420 m)。前文中已提到,误差的变化并非随海拔线性增长,海拔只是造成误差的众多因素之一,从误差分布上也可清晰看到 DL 预报在东部平原地区也存在一些改善不明显的站点。

3.2 不同预报时间的差异

从 00 时和 12 时起报的结果(表 3)综合来看,气温预报的平均 RMSE 有明显的日变化,每日 06 时对应的时效误差达到最大(两个起报时次平均误差 ECMWF-IFS 为 2.78℃,订正后为 2.27℃,减小了 0.51℃),每日 21 时对应的时效误差最小(两个起报时次平均误差 ECMWF-IFS 为 1.79℃,订正后为 1.46℃,减小了 0.33℃)。随着预报时效的增加(图 4),每天对应时间的 RMSE 也在缓慢增加,在每日 06 时、09 时,可以看到 ECMWF-IFS 预报的 RMSE 明显偏高,而 DL 模型 RMSE 的订正幅度也最大(达 0.50℃左右);不同起报时间(00 时和 12 时)的预报误差也有不同,但日变化规律基本一致。

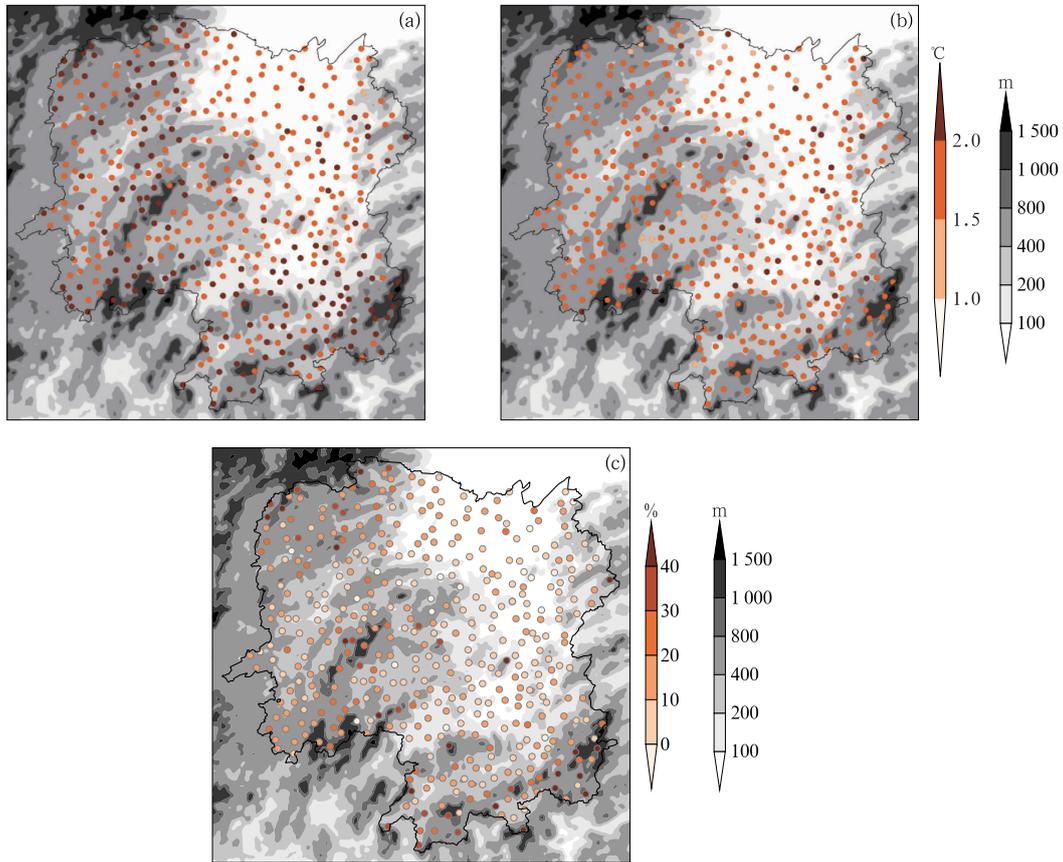


图 3 2020 年湖南不同区域站点的气温(a,b)平均均方根误差(彩色圆点)及其(c)改善率(彩色圆点)分布
(a)ECMWF-IFS 模式预报,(b)DL 模型预报
(灰色为地形海拔高度,下同)

Fig. 3 Spatial distributions of temperature (a, b) mean RMSEs (colored dots) for ECMWF-IFS, DL and (c) its improvement rate (colored dots) in 2020 for different regional stations in Hunan Province
(a) ECMWF-IFS forecast, (b) DL forecast
(Grey area is elevatin of topography, the same below)

表 2 2020 年湖南地区不同海拔高度站点的均方根误差改善率

Table 2 Statistics of RMSE improvement for stations at different elevations in Hunan Province in 2020

海拔范围/m	RMSE 改善率/%	站点数/个	站点比例/%
≤400	13.5	347	82.4
(400,800]	19.4	51	12.1
(800,1200]	27.4	13	3.1
>1200	50.0	10	2.4

表 3 2020 年湖南地区不同起报时间的逐 3 h 气温预报与实况的均方根误差(单位:℃)

Table 3 RMSEs in different starting time of 3 h temperature forecasts and observation in Hunan Province in 2020 (unit: ℃)

实况/UTC	ECMWF-IFS(00 时起报)	DL 模型(00 时起报)	ECMWF-IFS(12 时起报)	DL 模型(12 时起报)
00	1.78	1.56	1.82	1.54
03	2.24	1.86	2.29	1.93
06	2.82	2.30	2.75	2.23
09	2.64	1.99	2.69	2.17
12	2.07	1.71	2.00	1.66
15	1.89	1.57	1.84	1.56
18	1.84	1.55	1.84	1.56
21	1.81	1.53	1.78	1.38

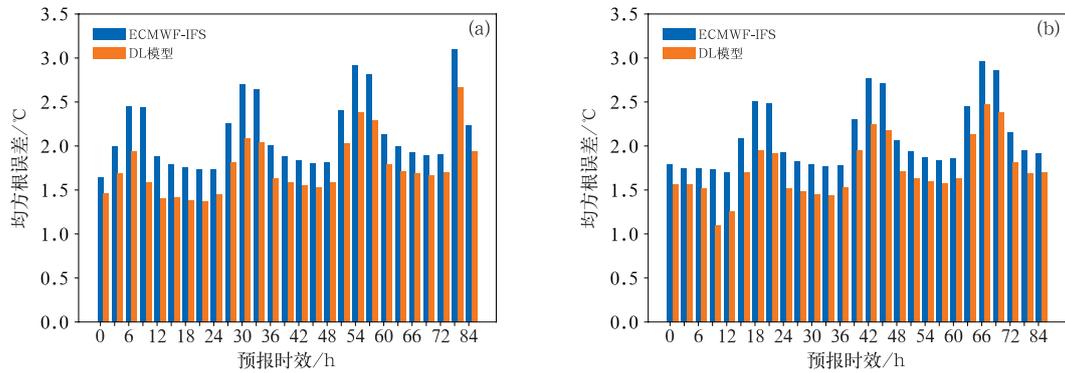


图 4 2020 年湖南地区不同起报时间的各预报时效的 ECMWF-IFS 与 DL 模型气温预报均方根误差 (a)00 时起报, (b)12 时起报

Fig. 4 RMSEs of temperature forecasts by ECMWF-IFS and DL model initiated at (a) 0000 UTC, and (b) 1200 UTC in 2020 in Hunan Province

结合海拔高度与观测时间(图 5),可以发现预报误差基本遵从海拔越高、DL 模型对数值预报的改善越明显的规律,但在不同海拔上,改善率具有不同的日变化特征。其中:海拔 1200 m 以上改善率最高的是 00、09、12、21 时,平均达 40%以上,800~1000 m 则在日落以后改善幅度最高,800 m 以下地区以每日午后(06 时、09 时)为主。

从所有站点的统计可以得出,每天 06 时和 09 时的误差改善率最高,本文选取这两个预报时间的误差改善率,分析其地理分布。午后时段(图 6a),改善最明显的是海拔 200 m 以上地区,大部分误差降低 20%以上,傍晚(图 6b)改善率更高,很多午后改善不明显的海拔 100 m 以下地区(东部平原为主),改善率达到 20%以上。

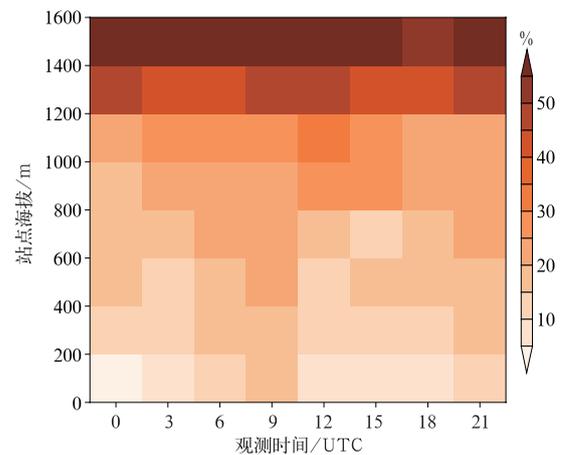


图 5 2020 年湖南地区不同海拔高度、不同观测时间的气温预报均方根误差改善率
Fig. 5 Temperature forecast RMSE improvement rate for different altitudes and observation time in 2020 in Hunan Province

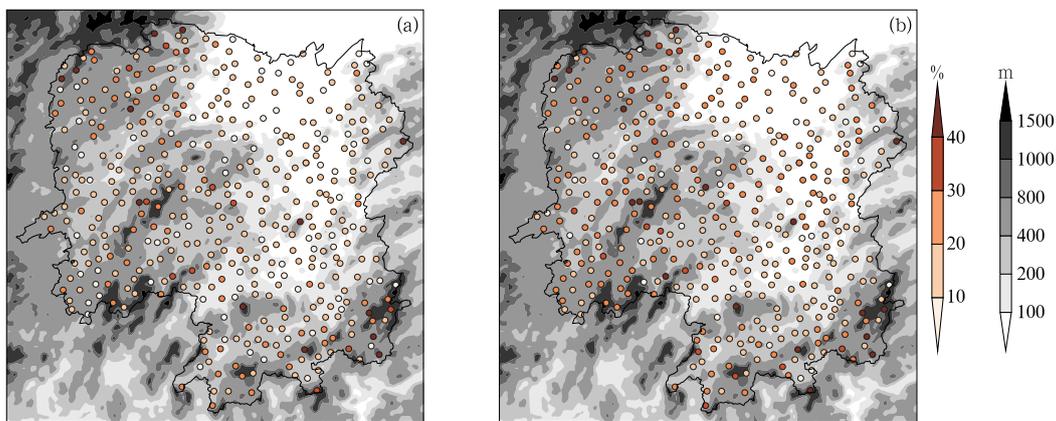


图 6 2020 年湖南地区不同海拔站点的气温预报平均均方根误差改善率(彩色圆点)分布 (a)预报时间:06 时, (b)预报时间:09 时

Fig. 6 RMSE improvement rate (colored dot) for stations at different altitudes in 2020 in Hunan Province (a) temperature forecast time: 0600 UTC, (b) temperature forecast time: 0900 UTC

在同年的不同月份,数值预报与 DL 模型的预报误差也有变化(图 7);ECMWF-IFS 只有 6、9、12 月平均 $RMSE < 2^{\circ}C$,其他月份均 $> 2^{\circ}C$,DL 模型仅 5 月平均 $RMSE > 2^{\circ}C$,其他月份均 $< 2^{\circ}C$;ECMWF-IFS 在 4 月的平均 $RMSE$ 最高为 $2.48^{\circ}C$ (DL 订正后为 $1.94^{\circ}C$),5 月 $RMSE$ 为 $2.38^{\circ}C$ (DL 订正后为 $2.04^{\circ}C$),说明 4 月、5 月预报难度较大。每个月,DL 模式都有明显的订正效果,其中 10 月、11 月(秋季)其对数值预报的改善率最高(超过 23%),其次是 3 月、4 月(春季,20%左右),12 月最低(9%)。

从全年所有日期的全省所有站 $RMSE$ 变化趋势(图略)来看,两个产品的误差变化趋势基本一致。DL 订正方法在绝大部分时间对于 ECMWF-IFS 数值预报结果进行了有效改善,特别是在 3、4、10、11 月改善较为明显。

3.3 寒潮降温过程检验

2020 年 12 月 29 日湖南省出现一次寒潮降温天气过程,此次过程为明显的转折天气:冷空气从湘东北洞庭湖地区南下影响全省,其中 29 日 00 时至 30 日 00 时,全省大部分地区 24 h 最高、最低气温相比前一日下降 $8 \sim 10^{\circ}C$,最低气温降至 $0^{\circ}C$ 左右。数值模式对降温幅度有较好的体现,但准确率不够,而

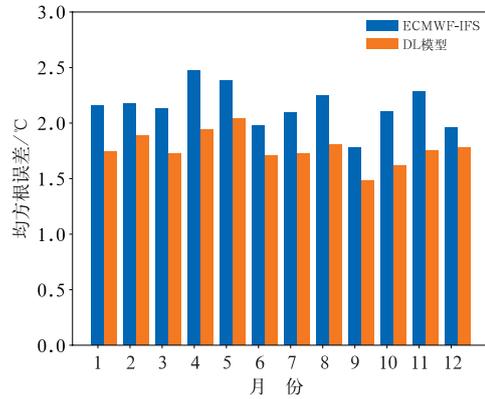


图 7 2020 年不同月份的 ECMWF-IFS 模式与 DL 模型订正后的气温预报均方根误差

Fig. 7 Temperature RMSEs for ECMWF-IFS and DL model in different months in 2020

基于深度学习方法订正后的预报,准确率有明显的提升。

从此次过程统计上(表 4)看,过程 3 d 全省站点高、低温准确率(误差 $2^{\circ}C$ 以内)和逐 3 h 均方根误差都是 DL 模型占优势(除了第 3 天低温准确率 ECMWF-IFS 表现稍好);其中:第 2 天、3 天的高温准确率,DL 提高了近 20 个百分点,第 1 天低温准确率,DL 模型从 49.0% 提升至 80.7%,优势非常明显。

表 4 2020 年 12 月 29 日 00 时开始的湖南寒潮过程 DL 模型与 ECMWF-IFS 的气温预报评估

Table 4 Forecast evaluation of cold wave weather process starting from 0000 UTC 29 December 2020 in Hunan Province

预报时效 评分项	0~24 h(第 1 天)			24~48 h(第 2 天)			48~72 h(第 3 天)		
	高温准确率	低温准确率	RMSE/°C	高温准确率	低温准确率	RMSE/°C	高温准确率	低温准确率	RMSE/°C
ECMWF-IFS	88.6%	49.0%	2.09	41.2%	61.7%	2.59	40.9%	71.7%	2.42
DL 模型	95.0%	80.7%	1.86	59.5%	65.7%	2.41	69.5%	65.5%	2.33

以第 1 天误差分布(图 8)为例:ECMWF-IFS 在湘西、湘东南地区(山地为主)部分站点高温预报偏低 $2 \sim 3^{\circ}C$,DL 模型在西部地区订正较好,将大部分地区的偏差控制在 $\pm 2^{\circ}C$ 以内,但在东南部山区的预报的订正提升效果不明显,在东北部平原地区甚至比 ECMWF-IFS 预报偏高;从低温误差分布上看,ECMWF-IFS 在全省大部分地区预报偏低 $2^{\circ}C$,局地偏低 $3^{\circ}C$ 以上,DL 模型将大部分地区的误差控制在 $\pm 2^{\circ}C$ 以内,效果较明显。

DL 订正在减小系统性误差的同时,因为受“学习到的经验”影响,对转折天气可能也有负向订正,对其极端性考虑不够(例如第 2 天降温最显著的东部平原地区气温预报偏高,第 3 天低温准确率低于 ECMWF-IFS)。此外,经过分析同时段 ECMWF 集

合预报的 50 个扰动成员,发现每个单独站都存在理论上最优成员,其对逐 3 h 气温、日高低温的准确率接近 100%,订正方法对原始预报的改善幅度仍不及理想中的最优集合成员(但不同条件下最优成员的选取值得研究)。

岳阳国家基本气象站位于湖南省东北部洞庭湖平原地区,是冷空气直接影响、降温显著的地区。从单站逐 3 h 预报、观测时序中(图 9)可以看出,相对实况观测(红色折线),ECMWF-IFS(绿色折线)的预报偏低,特别是每天日出前的最低气温偏差明显,大部分时次,DL 模型(蓝色)预报值比 ECMWF-IFS 更接近观测值,改善幅度(黑色虚线)为正且大多在 50% 以上。也有个别时次 DL 模型误差稍大,但在可接受的范围内。

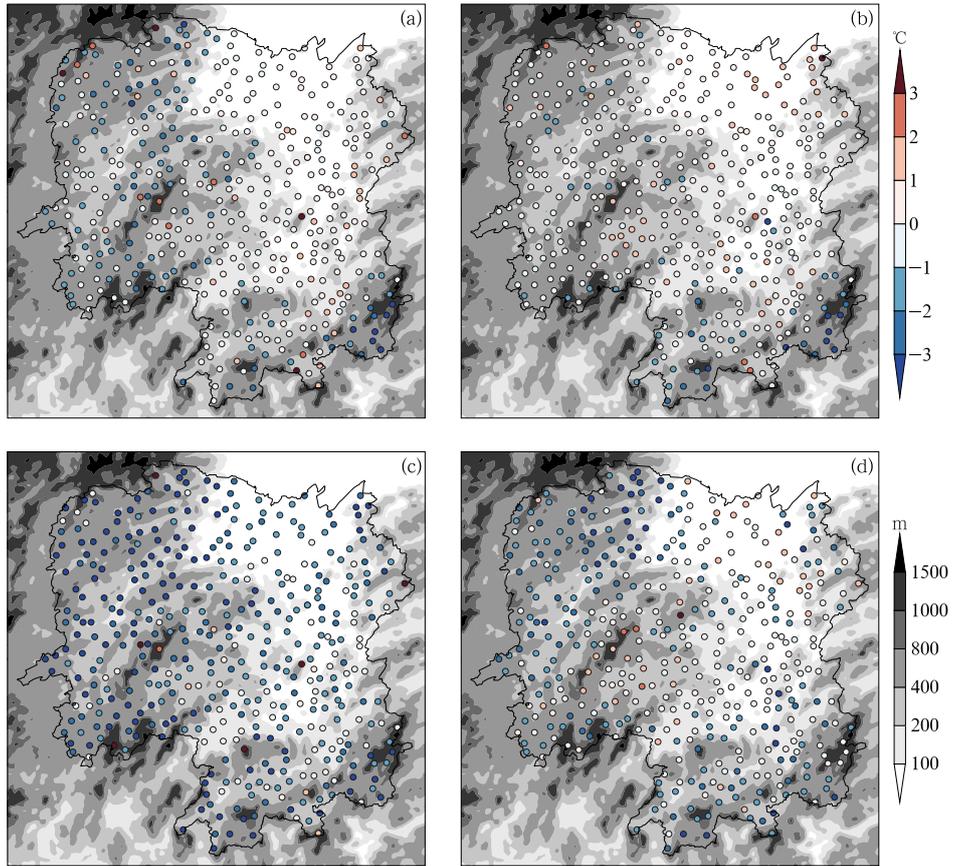


图 8 2020 年湖南地区站点(a,b)最高、(c,d)最低气温预报的偏差(彩色圆点)分布
(a,c)ECMWF-IFS 模式,(b,d)DL 模型

Fig. 8 Spatial distributions of errors (colored dot) for daily (a, b) maximum and
(c, d) minimum temperatures in Hunan Province in 2020
(a, c) ECMWF-IFS model, (b, d) DL model

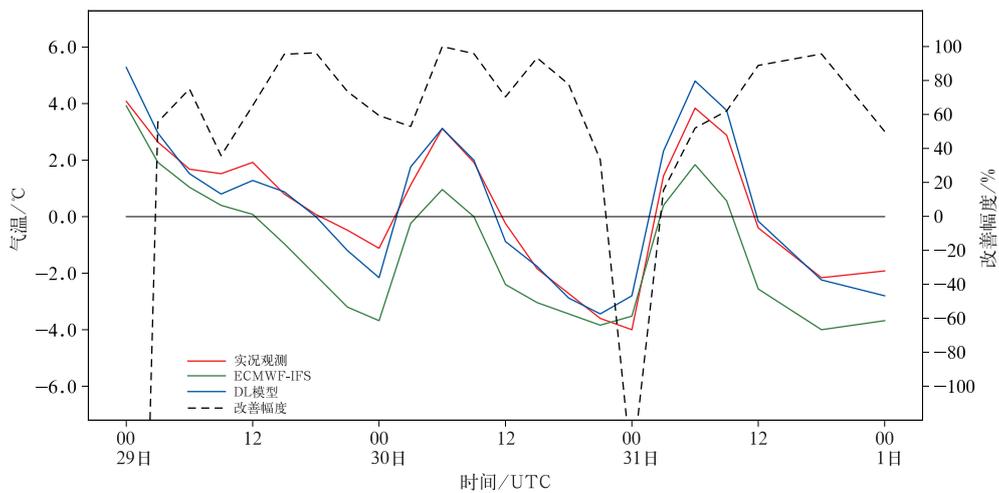


图 9 2020 年 12 月 29 日 00 时至 2021 年 1 月 1 日 00 时岳阳站逐 3 h 气温预报与观测对比
Fig. 9 Temperature forecasts and observation of Yueyang Station
from 0000 UTC 29 December 2020 to 0000 UTC 1 January 2021

此次过程,DL 模型在逐 3 h 预报上, RMSE 减少了 0.2°C 左右,说明逐 3 h 预报也是气温预报的难点。在实际业务应用中,预报员将偏差明显的个别站点进行了手工订正,起到了最终把关作用。

4 结论与讨论

本文从站点预报实际应用的角度,评估了基于深度学习算法的气温订正方法相对数值预报性能的改进,并简要介绍了数据处理、训练模型、评估检验到实际应用的方法。检验结果表明:

(1)从空间分布来看,DL 模型预报的气温 RMSE 在大部分地区控制在 2°C 以下,其对 ECMWF-IFS 模式的订正效果明显,尤其是在湖南省西部和南部的雪峰山、武陵山、南岭地区有显著的改善,从海拔分布上看,改善率随着预报站点海拔的增加而上升,特别是海拔 1200 m 以上的站点,改善率达到 50% 左右。

(2)模型预报的 RMSE 有明显的日变化,每日 06 时、09 时,ECMWF-IFS 误差最大(达 2.7°C 左右),DL 模型对其订正的幅度也最大(RMSE 平均减小 0.5°C),二者在每日 21 时对应的时效误差最小,在不同海拔高度有不同的日变化特征。

(3)DL 模型每个月相对 ECMWF-IFS 模式的平均 RMSE 都有明显的订正效果,二者在 4 月、5 月(春季)误差最大,其中 10 月、11 月(秋季)RMSE 改善率最高(超过 23%),其次是 3 月、4 月(春季,20% 左右),12 月最低(9%);DL 订正产品在 5 月的平均 RMSE $> 2^{\circ}\text{C}$,其他月份 $< 2^{\circ}\text{C}$ 。

(4)单次寒潮过程的检验效果表明,DL 模型对 ECMWF-IFS 模式的系统误差有明显的订正效果:降温开始当天,日最低气温准确率提升 30 个百分点,第 2 天、第 3 天日最高气温准确率提升 20 个百分点,大部分时间的最低气温准确率也有明显的改善,在逐 3 h 预报上, RMSE 平均减小 0.2°C ;在单站的曲线上,也能看出 DL 模型比 ECMWF-IFS 预报更接近观测实况。

湖南省气象台原有业务中使用的多模式最优集成,并没有针对数值模式带来误差的原因进行消除,其效果仅优于单个数值模式,但在一些地区(特别是山地)存在较明显的系统性误差,需要人工逐站订正。人工订正往往基于经验,一般参考前 1~3 d 数值预报的气温要素预报,假定短时期内预报偏差固定不变,并直接减去固定的偏差,这种方式无法定量

考虑多要素对气温预报的影响,容易受到干扰和误导。

对比由他人发展的订正方法(如滑动加权重订正方法等),本文中使用的方案更接近的是依据数值模式对当前天气形势场各要素的预报,来对不同下垫面不同地区的气温进行重新预报,而不是从原始预报值减去推导出的偏差值。在不同的地形,气温预报偏差受海拔、气压、云量、风场、湿度等因素影响,基于多气象因子的深度学习偏差订正更具有物理意义。

本文中采用的深度学习订正方法针对单站、单点建模,输出站点预报,在实际业务中设计出使用站点预报对背景预报场进行误差订正的方案,可以直接输出成各类服务产品,并作为第一优选客观预报源加入智能网格预报系统。在业务实践中,此产品逐 3 h 准确率评分高于其他所有参考预报来源、每日业务检验的高低温也高于其他来源。预报员大部分时间已不再进行手动订正(手动订正结果评分经常低于客观输出结果),特别是在极端天气高发时期,节省了大量的时间和精力,体现了研究型业务成果在实际工作中发挥的巨大作用。

研究初期,统计了本地 ECMWF-IFS、CMA-GFS 全球预报以及华东、华南等区域模式的气温预报准确率及均方根误差,发现 ECMWF-IFS 综合误差最小,因此优先采用其进行建模训练。

因为时间关系,本文模型暂未就日最高/最低气温单独建模,仅使用拟合日变化曲线计算高/低温,未进行针对性评估(针对日最高/最低气温的预报模型也在研究、改进中),基于我国高分辨率区域数值预报、CLDAS 气温分析场的逐小时智能网格预报模型也在研究中,但因数据量较大,每次建模需要较长时间。但随着模型的逐步优化、算力的提升,后期将得到更加精准的预报结果。

参考文献

- 陈昱文,黄小猛,李熠,等,2020. 基于 ECMWF 产品的站点气温预报集成学习误差订正[J]. 应用气象学报,31(4):494-503. Chen Y W, Huang X M, Li Y, et al, 2020. Ensemble learning for bias correction of station temperature forecast based on ECMWF products[J]. J Appl Meteor Sci, 31(4):494-503(in Chinese).
- 冯慧敏,智协飞,崔慧慧,等,2016. 基于多模式集成技术的地面气温精细化预报[J]. 气象与环境科学,39(4):73-79. Feng H M, Zhi X F, Cui H H, et al, 2016. Refined forecasting of surface temperature based on multi-model ensemble technology[J]. Meteor Environ Sci, 39(4):73-79(in Chinese).
- 郝翠,张迎新,王在文,等,2019. 最优集合预报订正方法在客观温度

- 预报中的应用[J]. 气象, 45(8):1085-1092. Hao C, Zhang Y X, Wang Z W, et al, 2019. Application of analog ensemble rectifying method in objective temperature prediction. Meteor Mon, 45(8):1085-1092(in Chinese).
- 嵇磊, 王在文, 陈敏, 等, 2019. 人工智能技术能否提高地面气温预报的精度——记 AI Challenger 2018 全球天气预报挑战赛[J]. 气象学报, 77(5):960-964. Ji L, Wang Z W, Chen M, et al, 2019. How much can AI techniques improve surface air temperature forecast; a report from AI Challenger 2018 Global Weather Forecast Contest[J]. Acta Meteor Sin, 77(5):960-964(in Chinese).
- 李佰平, 智协飞, 2012. ECMWF 模式地面气温预报的四种误差订正方法的比较研究[J]. 气象, 38(8):897-902. Li B P, Zhi X F, 2012. Comparative study of four correction schemes of the ECMWF surface temperature forecasts[J]. Meteor Mon, 38(8):897-902(in Chinese).
- 门晓磊, 焦瑞莉, 王鼎, 等, 2019. 基于机器学习的华北气温多模式集合预报的订正方法[J]. 气候与环境研究, 24(1):116-124. Men X L, Jiao R L, Wang D, et al, 2019. A temperature correction method for multi-model ensemble forecast in North China based on machine learning[J]. Clim Environ Res, 24(1):116-124(in Chinese).
- 任萍, 陈明轩, 曹伟华, 等, 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[J]. 气象学报, 78(6):1002-1020. Ren P, Chen M X, Cao W H, et al, 2020. Error analysis and correction of short-term numerical weather prediction under complex terrain based on machine learning[J]. Acta Meteor Sin, 78(6):1002-1020(in Chinese).
- 盛春岩, 范苏丹, 荣艳敏, 等, 2020. 几种气温客观预报方法对比及最优集成预报研究[J]. 气象, 46(10):1351-1361. Sheng C Y, Fan S D, Rong Y M, et al, 2020. Comparison of several objective methods and optimal consensus forecast study of temperature[J]. Meteor Mon, 46(10):1351-1361(in Chinese).
- 王丹, 王建鹏, 白庆梅, 等, 2019. 递减平均法与一元线性回归法对 ECMWF 温度预报订正能力对比[J]. 气象, 45(9):1310-1321. Wang D, Wang J P, Bai Q M, et al, 2019. Comparative correction of air temperature forecast from ECMWF model by the decaying averaging and the simple linear regression methods[J]. Meteor Mon, 45(9):1310-1321(in Chinese).
- 王婧, 徐枝芳, 范广洲, 等, 2015. GRAPES_RAFS 系统 2 m 温度偏差订正方法研究[J]. 气象, 41(6):719-726. Wang J, Xu Z F, Fan G Z, et al, 2015. Study on bias correction for the 2 m temperature forecast of GRAPES_RAFS[J]. Meteor Mon, 41(6):719-726(in Chinese).
- 王敏, 李晓莉, 范广洲, 等, 2012. 区域集合预报系统 2 m 温度预报的校准技术[J]. 应用气象学报, 23(4):395-401. Wang M, Li X L, Fan G Z, et al, 2012. Calibrating 2 m temperature forecast for the regional ensemble prediction system at NMC[J]. J Appl Meteor Sci, 23(4):395-401(in Chinese).
- 王在文, 陈敏, Luca Delle Monache, 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[J]. 气象学报, 77(5):869-884. Wang Z W, Chen M, Luca D M, et al, 2019. Application of analog ensemble method to surface temperature and wind speed prediction in Beijing Area[J]. Acta Meteor Sin, 77(5):869-884(in Chinese).
- 韦青, 代刊, 林建, 等, 2020. 2016—2018 年全国智能网格降水及温度预报检验评估[J]. 气象, 46(10):1272-1285. Wei Q, Dai K, Lin J, et al, 2020. Evaluation on the 2016—2018 fine gridded precipitation and temperature forecasting[J]. Meteor Mon, 46(10):1272-1285(in Chinese).
- 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, 2016. The optimal training period scheme of MOS temperature forecast[J]. J Appl Meteor Sci, 27(4):426-434(in Chinese).
- 许小峰, 2018. 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, 44(3):341-350. Xu X F, 2018. From physical model to intelligent analysis; a new exploration to reduce the uncertainty of weather forecast[J]. Meteor Mon, 44(3):341-350(in Chinese).
- 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, Zhang Y, et al, 2019. Bias correction method for the 2 m temperature forecast of ECMWF high resolution model[J]. Meteor Mon, 45(6):831-842(in Chinese).
- 尹姗, 李勇, 马杰, 等, 2020. 延伸期温度预报误差订正技术初探[J]. 气象, 46(3):412-419. Yin S, Li Y, Ma J, et al, 2020. Preliminary study on bias correction for the extended-range temperature forecast[J]. Meteor Mon, 46(3):412-419(in Chinese).
- 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 R X, et al, 2019. Comparison study on several grid temperature rolling correction forecasting schemes[J]. Meteor Mon, 45(7):1009-1018(in Chinese).
- 智协飞, 林春泽, 白永清, 等, 2009. 北半球中纬度地区地面气温的超级集合预报[J]. 气象科学, 29(5):569-574. Zhi X F, Lin C Z, Bai Y Q, et al, 2019. Superensemble forecasts of the surface temperature in Northern Hemisphere middle latitudes[J]. Sci Meteor Sin, 29(5):569-574(in Chinese).
- Glahn H R, Lowry D A, 1972. The use of Model Output Statistics (MOS) in objective weather forecasting[J]. J Appl Meteor Climatol, 11(8):1203-1211.
- Han L, Chen M X, Chen K K, et al, 2021. A deep learning method for bias correction of ECMWF 24—240 h forecasts[J]. Adv Atmos Sci, 38(9):1444-1459.
- Klein W H, Glahn H R, 1974. Forecasting local weather by means of model output statistics[J]. Bull Am Meteor Soc, 55(10):1217-1227.
- Leutbecher M, Palmer T N, 2008. Ensemble forecasting[J]. J Comput Phys, 227(7):3515-3539.
- Li H C, Chen Y, Xia J J, et al, 2019. A model output machine learning method for grid temperature forecasts in the Beijing Area[J]. Adv Atmos Sci, 36(10):1156-1170.