

钟琦,梁红丽,代刊,等,2025.降水预报机器学习订正及其在分类型降水测试的效果研究[J].气象,51(8):914-927. Zhong Q, Liang H L, Dai K, et al, 2025. Study on machine learning correction of precipitation forecast and its validation on two types of precipitation[J]. Meteor Mon, 51(8):914-927 (in Chinese).

降水预报机器学习订正及其在分类型 降水测试的效果研究^{*}

钟琦¹ 梁红丽¹ 代刊² 方祖亮¹ 申莉莉³ 侯邵禹⁴

¹ 中国气象局气象干部培训学院, 北京 100081

² 国家气象中心, 北京 100081

³ 河北省气象灾害防御和环境气象中心, 石家庄 050021

⁴ 河北省人工影响天气中心, 石家庄 050021

提 要: 强降水是对我国国计民生影响最为广泛的灾害性天气之一,其精准预报也是最具挑战的科学问题之一,湿物理过程的多尺度、非线性特征使数值预报和客观订正处理降水问题比处理一般气象要素困难得多。文章以京津冀地区 3 h 累计降水量为对象,基于站点观测和数值模式预报,通过降水样本构建和采样、降水相关物理特征输入、使用残差训练等策略,探索了机器学习算法 LightGBM 应用于降水预报订正的效果。结果显示,面对降水样本的长尾分布难题,构建数据集时综合考虑保持降水的真实分布,同时适度增大较强量级降水样本的比例,是提升强量级降水订正效果的关键一环。独立测试集的统计检验表明,LightGBM 订正方案在所有阈值降水的客观评分均较原模式预报显著提升,且提升率随降水阈值增大而增加。分类型降水的统计检验和个例评估显示,LightGBM 订正方案在不同类型降水预报均体现出强度和落区的综合调整,其中:强天气系统强迫类型降水样本数量相对占优,在各降水阈值订正效果均较显著;而弱天气系统强迫类型降水订正在 ≥ 15 mm 阈值时较难获得提升,落区的调整也更具挑战。这说明降水样本的不平衡分布仍是机器学习订正面临的挑战,但机器学习体现出对模式预报准确率更低的较大量级降水的订正潜力,即模式预报能力越低,机器学习订正提升的空间和幅度越大。特征重要性分析表明,输入与降水密切相关的动力、热力和水汽等物理特征,对提升 LightGBM 订正评分具有正贡献。

关键词: 京津冀地区,降水预报订正,机器学习,样本构建,分类降水

中图分类号: P456

文献标志码: A

DOI: 10.7519/j.issn.1000-0526.2025.052301

Study on Machine Learning Correction of Precipitation Forecast and Its Validation on Two Types of Precipitation

ZHONG Qi¹ LIANG Hongli¹ DAI Kan² FANG Zuliang¹ SHEN Lili³ HOU Shaoyu⁴

¹ CMA Training Centre, Beijing 100081

² National Meteorological Centre, Beijing 100081

³ Hebei Meteorological Disaster Prevention and Environmental Meteorology Centre, Shijiazhuang 050021

⁴ Hebei Weather Modification Centre, Shijiazhuang 050021

Abstract: Heavy precipitation is one of the most widespread hazardous weather affecting the socioeconomic stability and people's livelihoods in China. Accurately forecasting such events poses significant scientific challenges. The multi-scale nonlinear characteristics of moist physical processes make numerical weather

^{*} 国家自然科学基金项目(U2142214、42275013)、中国气象局重点创新团队智能预报技术团队项目(CMA2022ZD04)、国家重点研发计划(2021YFC3000905)和河北省重点研发计划(23375401D)共同资助

2024 年 12 月 31 日收稿; 2025 年 7 月 18 日收修定稿

第一作者:钟琦,主要从事强天气数值模拟、数值预报人工智能订正技术研究. E-mail:zhongq@cma.gov.cn

通讯作者:代刊,主要从事智能预报技术、统计后处理技术的研究. E-mail:daikan1998@163.com

prediction and objective corrections for precipitation become considerably more difficult than for other meteorological variables like wind and temperature. Utilizing station observations and numerical model forecasts, this paper explores the application effect of machine learning algorithm (LightGBM) in correcting 3 h accumulated precipitation forecasts for the Beijing-Tianjin-Hebei Region through strategies such as constructing and sampling precipitation datasets, inputting relevant physical features, and training on residuals. The results demonstrate that, to address the long-tailed distribution challenge of precipitation samples, when constructing the dataset it is crucial to comprehensively consider maintaining the true distribution of precipitation while moderately increasing the proportion of samples with stronger precipitation intensities. This is a key step in enhancing the correction effectiveness of heavy precipitation. Statistical tests on the independent test set show that the LightGBM correction scheme achieves significant improvements in skill scores for precipitation ranging from 0.1 mm to 20 mm compared to the raw model forecasts, and the increase rate ascends upward as the threshold rises. Statistical tests and individual case evaluations of precipitation by type show that the LightGBM correction scheme presents comprehensive adjustments in rainfall intensity and fall area in different types of precipitation forecasts. Among them, the number of forced precipitation samples by severe weather systems is relatively superior, and the correction effects on each precipitation threshold are more remarkable. Additionally, the evaluation of classified heavy precipitation indicates that it is more challenging to achieve improvements in the correction of convective heavy precipitation forced by weak weather systems, in particular in the cases with precipitation ≥ 15 mm. The adjustment of the fall area of precipitation is more challenging. This suggests that the unbalanced distribution of precipitation samples remains a challenge for machine learning correction. However, machine learning has shown particular promise for correcting larger magnitudes of heavy precipitation events with lower forecast accuracy from the model, that is, the lower the model's forecast ability, the greater the room and extent for the correction improvement of machine learning. The analysis of feature importance shows that the input of physical features such as dynamics, thermodynamics and water vapor, which are closely related to precipitation, has a positive contribution to enhancing the correction score of LightGBM.

Key words: Beijing-Tianjin-Hebei Region, precipitation forecast correction, machine learning, sample construction, classified precipitation

引言

强降水因其强度大、致灾性强等特点,常造成巨大的社会和经济损失,其精准预报既是业务难点,也是科学研究的重大挑战。目前客观预报在临近尺度主要采用基于观测的外推方法,观测信息通常在1~2 h后消弭;而数值模式在3~6 h前存在“spin-up”现象,预报通常需要进行订正。因此前人针对降水预报订正开展多项研究。例如,广泛应用于降水预报订正的频率匹配法可成功减小预报的系统误差(Zhu and Luo, 2015; 李俊等, 2014; 2015)。代刊等(2016)利用多模式的集合预报产品提供的预报不确定性信息提出最优百分位法,对暴雨的预报技巧表现接近或略超过预报员水平。概率匹配集合平均

法,结合具有较好空间分布的集合平均场和更高量级准确度的集合成员预报进行降水订正,获得了广泛应用(Ebert, 2001; Fang and Kuo, 2013)。其他降水客观订正方法还包括相似法(Hamill and Whitaker, 2006)、分位值映射法(Voisin et al, 2010)等。但这些方法存在一些不足,主要有:(1)多为单变量订正,即只基于模式降水预报进行订正,未考虑可能对降水订正有贡献的水汽、动力和热力学变量信息,订正中使用的信息源不完备;(2)非参数化订正,即不使用统计分布模型,因此非参数化模型推演能力较弱,通常只对数值模式的系统性偏差有一定订正能力;(3)在强量级降水小时尺度精细化预报上应用效果有限。

与传统降水订正方法相比,机器学习具有更多优势,如可以方便地处理时间和空间信息,快速处理

大量的预报变量,融合物理知识及从数据中自主学习挖掘信息。此外,机器学习可凭借较小的计算成本,更好地拟合出关键特征和目标间的关系。近年来,越来越多机器学习方法(Czerneckin et al, 2019; Han et al, 2015; 黄威和牛若芸, 2017)被应用于提升天气预报,尤其在临近预报领域取得了显著效果(Hernández et al, 2016; 盛杰等, 2025)。降水预报订正方面,陈浩等(2017)选择 K 近邻、多元自回归样条、支持向量机、多项对数线性模型和人工神经网络等方法,利用海拔、相对湿度、坡向、植被、风速、气温和坡度等因子订正了两种常用的青藏高原原日降水数据。张晨阳等(2019)使用基于地形的加权随机森林方法进行雷达定量降水估计,并在此基础上考虑了反射率垂直廓线的特征重要性变化和复杂地形区域降水的山岳抬升效应。在降水的小时尺度预报和订正方面,谷歌公司开发的 MetNet 系列(Sønderby et al, 2020; Espeholt et al, 2021)使用轴向注意力来提取全局的空间信息,用相对较小的计算代价实现了更大空间上的信息聚合,并使用雷达、卫星和模式预报等多源信息学习降水概率,应用于美国大陆未来 8 h 预报中的表现超过了高分辨率快速更新模式。尽管多项研究显示机器学习对降水预报有改进,但当前业务中使用的降水预报客观订正仍多基于传统方法(Clark et al, 2012; 代刊等, 2018; 罗聪等, 2021; 苏翔等, 2021; 张娇等, 2021; 张海鹏等, 2020)。降水预报机器学习订正收益多在中小雨,对 15 mm 及以上量级降水预报的提升仍收效甚微(赵华生等, 2021; 陈锦鹏等, 2021),这是由于此类降水发生频次少、样本量小,且降水样本存在严重的不平衡分布(样本数量随降水强度增加锐减),因此数据集构建时对降水样本的总量、样本分布形态的处理、采用的机器学习算法、优化目标和训练策略的平衡均有较大挑战。近年来,一些研究通过改进深度学习算法取得了降水预报订正效果的提升(Chen et al, 2023; Zheng et al, 2024; Zhou et al, 2022),但仅基于单一模式预报进行订正,且当样本量不够充分时深度学习算法通常会导致训练过拟合。机器学习相比深度学习方法更简便易用,在表格型数据上的表现通常不错(Shwartz-Ziv and Armon, 2021; Grinsztajn et al, 2022)。因此,本文采用机器学习集成算法 LightGBM,通过融合气象专家知识调整降水样本分布和选取物理特征,使用站点观测和多模式预报数据,以 3 h 累计降水量为对象开展预报

订正研究,基于京津冀地区夏季降水进行独立测试,开展统计检验和分类型降水订正效果评估。参考 Hu et al(2021)和 Peng et al(2023)对 3 h 累计降水量阈值的等级划分,及前述研究对 15 mm 及以上量级降水预报提升的不足,本文重点关注 $\geq 15 \text{ mm} \cdot 3 \text{ h}^{-1}$ 阈值降水的订正效果。

1 数据和模型

1.1 观测和模式数据

本文研究区域主要覆盖京津冀地区,经、纬度范围选取 $36^{\circ} \sim 43^{\circ} \text{N}$ 、 $113^{\circ} \sim 122^{\circ} \text{E}$ 。降水观测数据采用地面站点观测降水,包括该地区 3946 个国家级和区域级气象观测站的逐小时降水观测,该套数据已经过气象资料业务质量控制(孙卓等, 2020),站点空间位置和海拔高度如图 1 所示。降水订正目标为欧洲中期天气预报中心全球高分辨率模式预报(ECMWF)的未来 3 h 降水量预报。由于 ECMWF 全球模式采用静力平衡假设,其对对流性降水的描述存在缺陷,预报的降水强度通常偏弱,因此本文还引入了与 ECMWF 分辨率相当的区域中尺度模式(CMA-SH9)预报要素和降水量(Zhang et al, 2021)作为输入特征,其中 CMA-SH9 预报的 3 h 降水量由逐小时降水量累加获得。数据空间匹配由格点资料插值到站点。

观测和模式数据的时空分辨率信息如下:自动

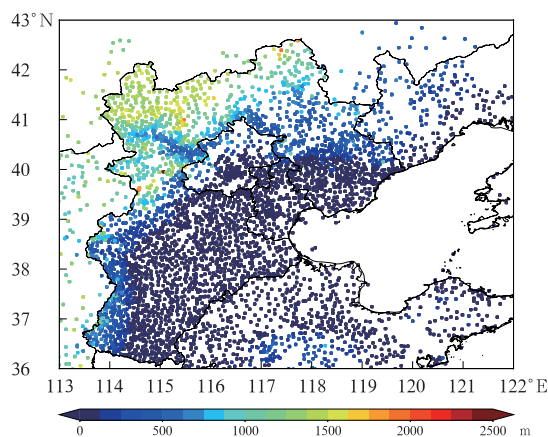


图 1 研究区域海拔高度(填色)及观测站点(圆点)分布

Fig. 1 Elevation of study area (colored) and distribution of observation stations (dot)

观测站的观测变量包括小时降水量、温度、露点温度和风等要素,模式高空预报包括各等压面的相关预报要素和物理诊断量,模式地面预报包括模式降水量、2 m 温度和湿度、10 m 风等。ECMWF 和 CMA-SH9 模式均是每天两次起报(00 时和 12 时;世界时,下同),空间分辨率为 9 km,均可提供间隔 3 h 的高空预报产品,地面预报产品则分别间隔 3 h 和 1 h。

1.2 输入特征选取

降水是多尺度系统相互作用的结果,从物理量场和形势场预报到降水预报之间存在着复杂的非线性

性过程,而由于当前数值模式在物理参数化等方面的固有缺陷,本文考虑将预报准确率相对更高的风、温度要素及相关诊断量作为特征输入,以期通过机器学习方法挖掘特征和降水间的关系进而改进降水预报。根据本研究前期试验(钟琦等,2022),除了模式降水预报外,另选取与降水密切相关的四类特征作为模型输入,包括水汽特征、动力特征、热力特征和维持时间特征,这四类特征中又选取了多个不同的模式变量,其中在大气中高层以 ECMWF 预报场输入为主,近地面层要素则同时采用了 ECMWF 和 CMA-SH9 模式的预报,以期能捕捉更丰富的中小尺度对流信息,详情如表 1 所示。

表 1 降水订正输入的 ECMWF 和 CMA-SH9 模式不同类别变量特征

Table 1 Features from ECMWF and CMA-SH9 fed to LightGBM

特征类别	描述	变量	来源
水汽特征	水汽相关物理量	整层可降水量、比湿、露点温度、水汽通量散度、相对湿度、温度露点差	ECMWF
		2 m 的露点温度和相对湿度	CMA-SH9
	暖云厚度	0°C 层高度	ECMWF
动力特征	动力抬升	850、700 和 500 hPa 垂直风速及均值,高空急流,过去 3 h 的 10 m 阵风	ECMWF
		10 m 的 u 风和 v 风	CMA-SH9
		850~500 hPa 假相当温度	ECMWF
热力特征	暖湿结构	2 m 气温	CMA-SH9, 站点观测
		K 指数	ECMWF
	对流抑制和抬升	对流有效位能	ECMWF, CMA-SH9
		对流抑制能量	CMA-SH9
维持时间特征	整体风速	环境平均风(1000~300 hPa 的 u 风和 v 风)	ECMWF
	低空急流	地表至 700 hPa 平均风速	ECMWF
	单层风速	u 风和 v 风	ECMWF

1.3 机器学习模型及参数设置

梯度提升树(Gradient Boosting Decision Tree)通过迭代拟合残差顺序地学习基学习器,并用加和的方式组合多个基学习器的结果,被认为是统计学习方法中性能最优越的方法之一(周志华,2017;李航,2019)。由于其对连续型数值进行最优特征选取时较为耗费时间,微软在 2017 年推出了 LightGBM (Light Gradient Boosting Method)(Ke et al,2017),优

化了决策树在生长时最优分支的选取策略,提高了运行效率和计算精度,增强了模型抗过拟合的能力。因此本文选取 LightGBM 开展降水预报订正试验,模型训练使用的工具为 LightGBM 在 Python 中的 sklearn 接口。为高效搜索最佳参数组合,并防止模型过拟合,LightGBM 模型训练时设置了多组参数,通过网格搜索和 5 折交叉验证的方法训练模型,选取最优参数。LightGBM 参数搜索空间及最优参数见表 2。

表 2 LightGBM 参数搜索空间及最优参数

Table 2 Parameter search space and optimal parameters of LightGBM

参数	参数解释	搜索空间	最优参数
max_depth	控制每棵决策树的最大深度	[6, 9, 15]	9
n_estimators	指定模型使用的决策树(基学习器)总数量	[100, 200, 300]	200
learning_rate	控制每棵树的贡献权重	[0.05, 0.1, 0.2]	0.1
num_leaves	单棵决策树的最大叶子节点数	[32, 100, 1024]	32

1.4 检验评估方法

对机器学习订正结果进行统计评分和降水个例检验,在保证改进客观指标的同时具有天气学评估意义,通过特征重要性分析加深对所取得订正效果的物理解释。客观检验指标采取了降水预报评估通常使用的指标:TS、FAR、POD、Bias、SR。指标具体定义如下。

$$TS = \frac{TP}{TP + FP + FN}$$

$$FAR = \frac{FP}{TP + FP}$$

$$POD = \frac{TP}{TP + FN}$$

$$Bias = \frac{TP + FP}{TP + FN}$$

$$SR = 1 - FAR$$

式中:TP 为命中数,FP 为空报数,FN 为漏报数。下文客观检验计算结果均指 3 h 累计降水量 $\geq$ 设定阈值的指标评分。

2 降水预报机器学习样本构建

2.1 机器学习降水数据集制作

如上所述,基于逐小时加密站点观测、ECMWF 和 CMA-SH9 模式预报变量构造降水订正数据集,订正所用特征主要来自模式,学习目标为加密站点的观测降水。首先将特征和学习目标分别在时间、空间上进行最近邻匹配。基于站点观测的 3 h 累计降水量和表 1 所示的模式变量进行时间匹配:假定当前时刻为 t ,选取 t 时刻站点 S 的 3 h 累计降水量作为目标 Y ,ECMWF 和 CMA-SH9 模式在距 t 时刻最近的起报时间预报的 t 时刻变量,选取表 1 中的特征作为 X 。空间匹配使用双线性插值方法,将栅格分布的特征 X 插值到站点 S 上,得到站点 S 的 t_0 样本 (X, Y) 。对所有站点进行批量处理,生成样本 N 。考虑大气运动的连续性,仅选取当前时刻数据难以学习时间变化信息,本研究前期试验(图略)表明选取提前至 6 h 的历史时刻信息,对模型训练效果具有正贡献。因此,下文在制作机器学习数据集时,每个样本除当前时刻 t 的特征 X ,还增加了 $t-3$ 和 $t-6$ 时刻的相应特征。

2.2 机器学习试验数据划分

本文采用 2017—2019 年降水进行机器学习数据集构建,为避免数据泄露,同时保证测试是在独立的一段连续时间内发生的多个降水个例上进行,用 t_1 、 t_2 两个时间点作为结点,划分训练集和测试集。其中, t_1 为 2019 年 7 月 20 日, t_2 为 2019 年 9 月 1 日。以 t_1 之前时段作为训练集,包括 2017 年 5 月 1 日至 10 月 15 日、2018 年的 5 月 1 日至 10 月 15 日、2019 年 5 月 1 日至 7 月 20 日,独立测试集时段为 t_1-t_2 时间段,即 2019 年 7 月 21 日至 9 月 1 日。为考察模型的泛化性和实际应用的潜力,本文在测试时间段(t_1-t_2)进行了两个独立测试,一是基于 2.3.1 节降水日的数据集,二是未经样本处理的连续数据集。

2.3 降水样本采样对机器学习的影响

降水预报订正面临的一个关键难点在于降水样本数量随强度的增加迅速减小,例如降水相比晴天是小样本,而降水样本中小雨样本又占据了绝大多数(如图 2 所示),因此样本呈现典型的长尾分布。如果不对样本分布进行处理,会导致模型严重趋于学习无雨或小雨特征,强量级降水的预报难以提升,因此科学调整样本分布和构建训练数据集是机器学习降水订正的关键一环。为有效提升强量级降水样本的占比且同时保障数据分布的真实性,本文筛选包含短时强降水的日降水样本作为总样本(详见 2.3.1 节);下采样适当剔除小雨样本,进一步提升强量级降水占比,同时保持样本分布形态与真实情况相差不太多,以避免训练过拟合或对大雨和小雨的订正失衡(详见 2.3.2 节)。此外,采用预报偏差训练策略,即将 ECMWF 模式的预报降水量和观测降水量的偏差作为模型学习的直接目标。

2.3.1 降水日样本筛选

根据短时强降水定义及京津冀地区统计特征分析(申莉莉等,2024),综合降水强度和范围,本文降水日样本选取如下:当日有 $\geq 20 \text{ mm} \cdot \text{h}^{-1}$ 降水发生,且站次数超过当年所有短时强降水日站次 75% 分位数;或站次小于 75% 分位数但达 50% 以上,且当日有 $\geq 50 \text{ mm} \cdot \text{h}^{-1}$ 降水发生。据此挑选出 115 个降水日样本。按照 2.2 节数据集的时间划分,测试集内包含 27 个降水日样本。经过上述降水日样本筛选,去除了大量无降水样本,并在真实样本中增

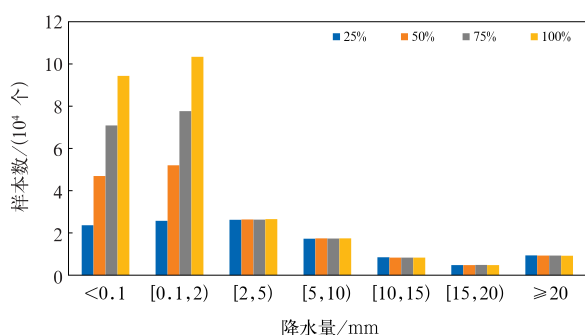


图2 <2 mm降水样本不同比例下采样后不同降水量区间的样本数分布

Fig. 2 Distribution of sample number in different precipitation intervals after different downsampling ratios of precipitation <2 mm

大了强量级降水的占比。下文机器学习模型的训练基于该降水日样本开展,测试在降水日样本和连续时间样本上分别开展。根据预报经验进一步将降水日样本区分为强天气系统强迫类型降水(strong synoptic forcing, SSF)和弱天气系统强迫类型降水(weak synoptic forcing, WSF),分类降水的判识基于再分析资料的高低空形势场由预报员主观天气分型,SSF类型降水通常形势场伴有明显的主导天气系统(如高空槽、低涡、副热带高压、台风等),WSF类型降水则通常难以定位出降水的主要影响天气系统,具体判识规则详见 Zhong et al(2022)。下文模

型训练时未进行降水类型区分,仅在测试集分别对其在这两类降水预报的订正效果进行检验评估。除此外降水日样本筛选时使用了短时强降水标准,全文其他处降水均指3 h累计降水量。

2.3.2 降水样本下采样试验

基于2.3.1节筛选出的降水日样本数(图2),可见无降水样本得到了较大控制,强量级降水占比提高,但仍呈现严重的长尾分布特征, <2 mm样本占比很高。因此进一步对其进行下采样处理,分别试验了随机抽取25%、50%、75%等不同比例 <2 mm降水样本对模型训练结果的影响。

图3a显示,对 <2 mm降水下采样比例降低确实会降低小量级降水的TS评分,但 ≥ 10 mm降水TS评分提升,即采用小量级降水样本越多(保留75%、50%、25%),在小量级降水的评分效果越好;剔除小量级降水样本越多(相应强量级降水样本比例增大), ≥ 10 mm降水的TS评分越高。而这种影响在下采样达到50%以后带来的差异却不再显著。POD、FAR指标(图3b和3d)显示出相同的变化趋势,当下采样25%时,发生较大的空报(Bias增加),同时错报率(FAR)最大。

综合考虑各检验指标的整体提升,同时尽可能兼顾样本分布的真实性,选择对 <2 mm降水下采样50%的方案开展下文试验。图4为采取该方案

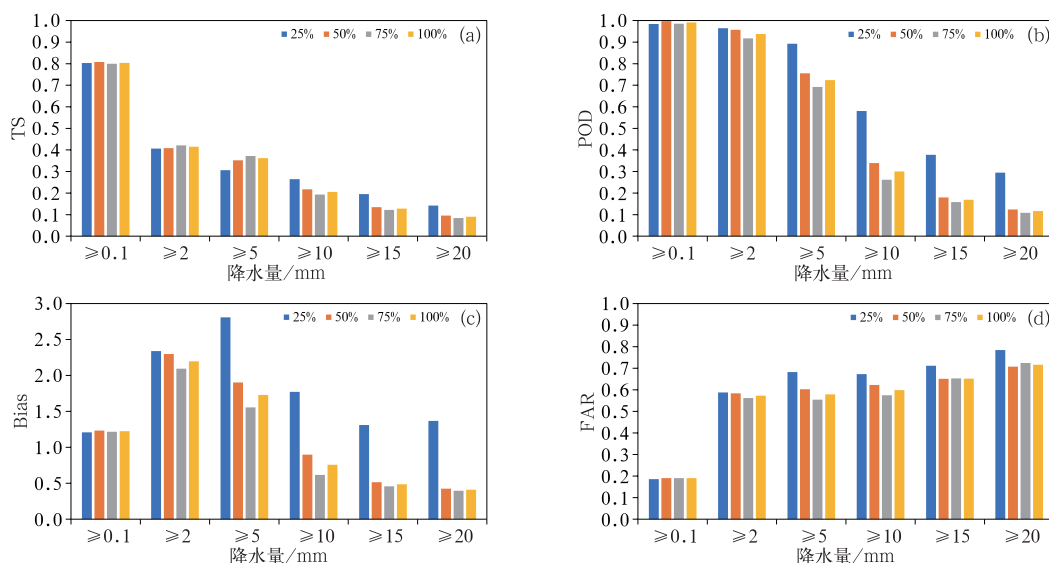


图3 <2 mm降水不同下采样比例测试集在不同降水量级的评分结果

(a) TS, (b) POD, (c) Bias, (d) FAR

Fig. 3 Results of the downsampling ratio in test set in different precipitation thresholds for precipitation <2 mm

(a) TS, (b) POD, (c) Bias, (d) FAR

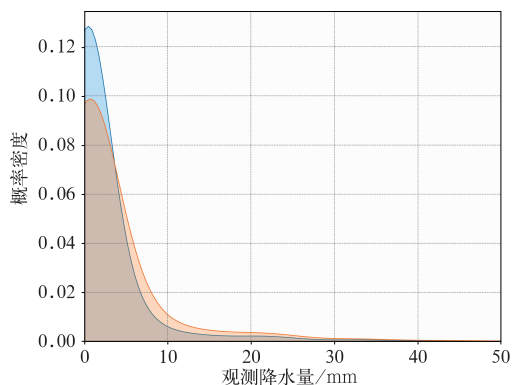


图 4 < 2 mm 降水 50% 下采样前(蓝色区域)和下采样后(橙色区域)降水样本的概率密度分布

Fig. 4 Probability density distribution of samples before resampling (blue area) and after resampling (orange area) with 50% downsampling ratio for precipitation < 2 mm

前后降水样本的分布,由图可见下采样处理明显缓解了降水样本的长尾分布形态,改善了降水样本与无降水样本、强量级降水与小量级降水样本之间的巨大差异。

表 3 给出了总降水样本和分类降水样本在训练集和测试集的样本数量,大部分样本为 SSF 类型降水,显著多于 WSF 类型降水。即在采样后的样本中还存在明显的降水类型不平衡,由于 WSF 类型降水多为局地性强、持续时间短的对流性降水,这种不平衡在强量级降水样本里更加突出,为 WSF 类型降水订正带来很大挑战。

表 3 训练集、测试集中不同类型降水的样本数量(单位:个)

Table 3 The number of two types of precipitation samples in the training set and test set (unit: number)

样本类型	样本数	训练集数量	测试集数量
总样本	135 440	118 856	16 584
SSF 类型降水	120 209	105 781	14 428
WSF 类型降水	15 231	13 075	2 156

3 降水订正结果分析

3.1 测试集统计评分

测试集上 3 h 降水量预报的 TS 评分如图 5 所示,图 5a、5b 分别为基于降水日、连续时间的测试集结果。图 5a 可见两个模式预报相比较,ECMWF 预

报技巧随降水强度增大减小, ≥ 20 mm 降水的 TS 评分低于 CMA-SH9,而 LightGBM 订正在所有降水阈值 TS 评分值均高于两个原模式。LightGBM 订正相对 ECMWF 预报的提升率随降水强度增大而增加, ≥ 2 mm 降水的提升率约为 15%, ≥ 20 mm 降水的提升率接近 50%,即原模式预报技巧越低,机器学习订正的空间和提升幅度越大。为验证模型泛化性,图 5b 给出了在连续时间测试集的效果,可见 LightGBM 订正在各量级降水预报 TS 评分均显示出较原模式的提升,与图 5a 结果一致;TS 评分提升率在 ≥ 2 mm 和 ≥ 5 mm 降水较小,但 ≥ 10 mm 降水提升率显著增加。这也一定程度印证了样本处理对较强量级降水预报效果提升的重要性。

由于天气系统带来的降水范围、大小和变化具有很大差异,下文以基于降水日测试集的结果考察 LightGBM 订正在不同降水类型取得的改进和表现。按照 2.3.1 节两大类降水类型进行区分,通常强强迫天气系统影响下的 SSF 类型降水范围大、持续时间长,弱天气系统背景下的 WSF 类型降水中心通常范围小而零散、持续时间短、强度较剧烈,WSF 类型降水的发生频次远小于 SSF 类型降水。整体降水的综合评分(图 6a)显示, ≥ 10 mm 降水 LightGBM 订正在各项指标最优,TS 评分和 POD 均较原模式取得较大提升,但仍存在漏报。图 6b、6c 显示,LightGBM 订正在两类降水均取得了显著提升。其中 WSF 类型降水 TS 评分提升率尤为明显,这与前述分析一致,当模式预报准确率较低时,机器学习订正提升的空间和幅度更大。SSF 类型降水的订正效果与整体降水测试集的分布较为相似,这是由于 SSF 类型降水在样本数上占多数(表 3)。WSF 类型降水预报订正显示出显著降低的命中率和较严重的漏报,这是由于这类降水多由小尺度系统触发,因此不确定性更强、更易漏报。LightGBM 对 WSF 类型降水的预报订正优于 ECMWF,但是 ≥ 15 mm 降水的 TS 和 POD 评分不如 CMA-SH9,这很可能是由于 WSF 类型降水样本数占比太少,LightGBM 难以有效学习其分布特征。

图 7 给出了分类型降水预报的 TS 评分及 LightGBM 订正相对 ECMWF 预报的 TS 评分增长率。可以看出,SSF 类型降水 TS 评分提升趋势与整体测试集(图 5a)表现类似,只是 TS 评分的绝对值更高(≥ 15 mm 降水 LightGBM 订正的 TS 评分大于 0.2)。图 7b 显示,WSF 类型降水的 TS 评分

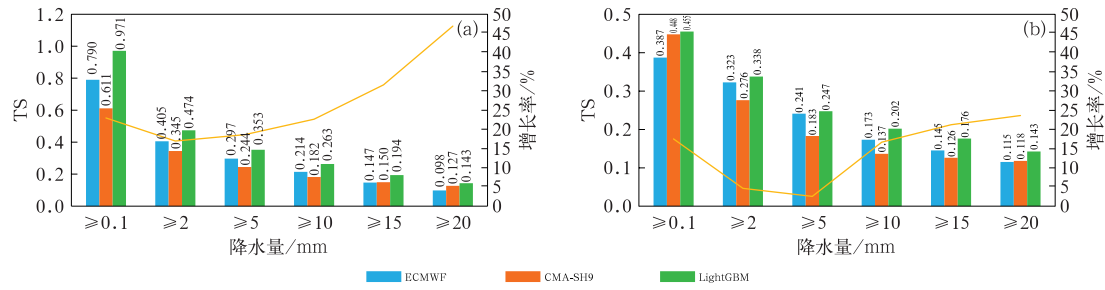


图 5 LightGBM、ECMWF 和 CMA-SH9 的 3 h 累计降水量预报在不同降水量级的 TS 评分(彩色柱),以及 LightGBM 相对 ECMWF 的增长率(黄线)

(a) 基于降水日的测试集, (b) 基于连续时间的测试集

Fig. 5 LightGBM correction, ECMWF and CMA-SH9 3 h accumulated precipitation forecasts' TS scores (colored column) and growth percentages (yellow line) of different precipitation thresholds on (a) precipitation test set and (b) continuous test set

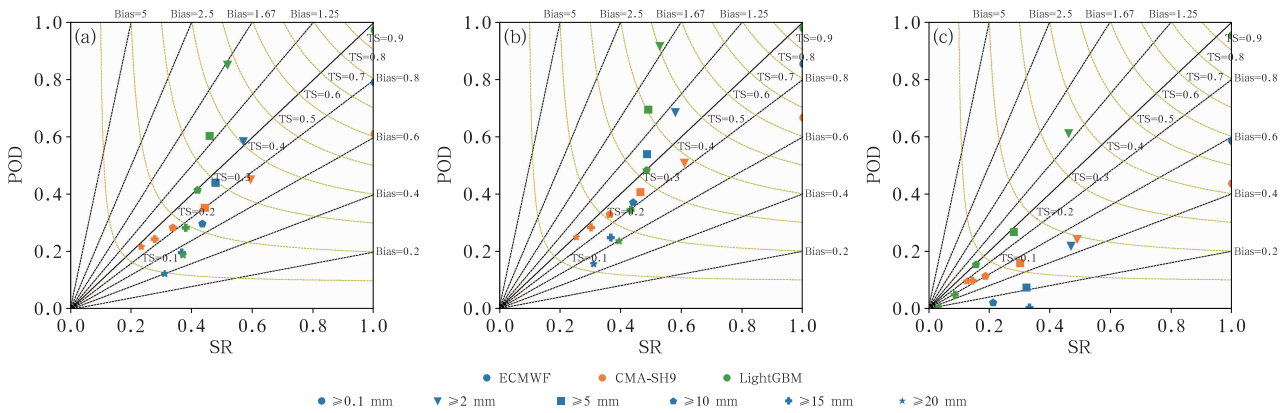


图 6 LightGBM、ECMWF 和 CMA-SH9 降水预报的综合评分

(a) 所有降水, (b) SSF 类型降水, (c) WSF 类型降水

Fig. 6 The ratings of LightGBM correction, ECMWF and CMA-SH9 forecasts (a) all precipitation, (b) SSF precipitation, (c) WSF precipitation

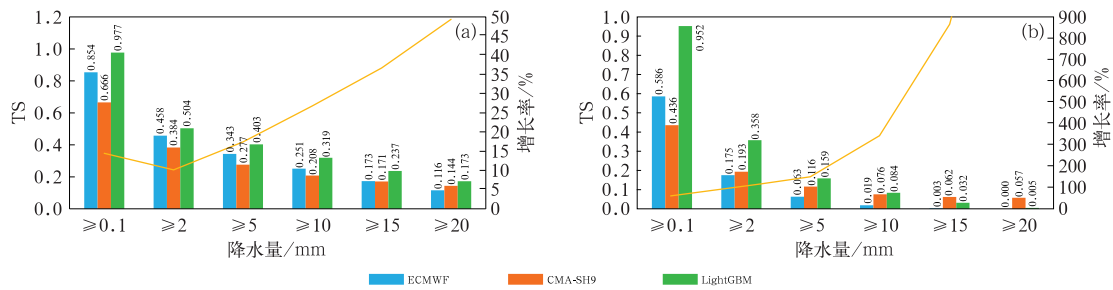


图 7 (a)SSF 类型降水和(b)WSF 类型降水的 LightGBM、ECMWF 和 CMA-SH9 预报在不同降水量级的 TS 评分(彩色柱),以及 LightGBM 相对 ECMWF 的 TS 评分增长率(黄线)

Fig. 7 LightGBM correction, ECMWF and CMA-SH9 precipitation forecasts' TS scores (colored column) and growth percentages (yellow line) of different precipitation thresholds on (a) SSF precipitation and (b) WSF precipitation

(≥ 10 mm 降水 LightGBM 订正的 TS 评分为 0.084) 显著低于 SSF 类型,且随降水强度的增大降低越为明

显,ECMWF 降幅非常明显,说明其对 WSF 类型降水的预报能力存在较大缺陷,CMA-SH9 减小幅度相对

平稳,对 ≥ 20 mm 降水仍有预报技巧(TS 评分为 0.057)。LightGBM 订正相对 ECMWF 在所有阈值均取得 TS 提升;但 ≥ 15 mm 降水 CMA-SH9 的 TS 评分最高。WSF 类型降水由于本身的可预报性较差,模式预报准确率也较低,既是模式预报的难点,也是 LightGBM 订正的难点,需要尽可能地扩充该类型降水的真实样本数量,并采用特殊的训练策略才有可能取得进一步改进。

3.2 测试集降水水平面分布评估

考虑到降水空间分布的不连续性,为检验 LightGBM 对空间分布的订正效果,将测试集各站点所在时间段的降水累加,查看其空间分布,结果如图 8 所示。从整体数值来看,降水总量的最大值和中位数的分布,地面站点观测:748.4 mm,126.1 mm;ECMWF 预报:372.7 mm,97.9 mm;CMA-SH9 预报:582.7 mm,104.8 mm。与观测相比,ECMWF 预报均显著偏小,CMA-SH9 表现优于 ECMWF,但总体仍偏小。LightGBM 订正后,两个数值分别为 541.8 mm 和 145.5 mm;虽然单站最大降水量订正

仍不够,但相对 ECMWF 取得明显改进,中位数相比两个模式有大幅提高、略有过报。这说明订正后的降水在量级上整体有提升,但在极值上存在一定程度的平滑。从图 8 空间分布来看,ECMWF 对该时段观测的四个降水中心均有体现,但覆盖范围偏小、降水强度显著偏弱,LightGBM 订正后河北的中部、东北部沿海和山东东北部沿渤海地区的降水较 ECMWF 明显提升,CMA-SH9 对这些降水中心的强度预报较 ECMWF 有改进,但 LightGBM 对降水中心的范围调整更大,尤其是河北中西部,及河北东南部沿与山东交界的强降水雨带较两个模式预报均显著提升,整体空间分布更接近实况。

3.3 分类型降水个例分析

如前测试集统计分析显示,SSF 类型降水由于天气系统信号强,降水落区和中心位置相对可预报性更高,模式预报能力更强,而在弱天气系统强迫下的局地性强的降水上预报能力相对弱得多。为进一步查看 LightGBM 对不同类型降水空间分布的订正效果,分别选取降水个例进行分析。

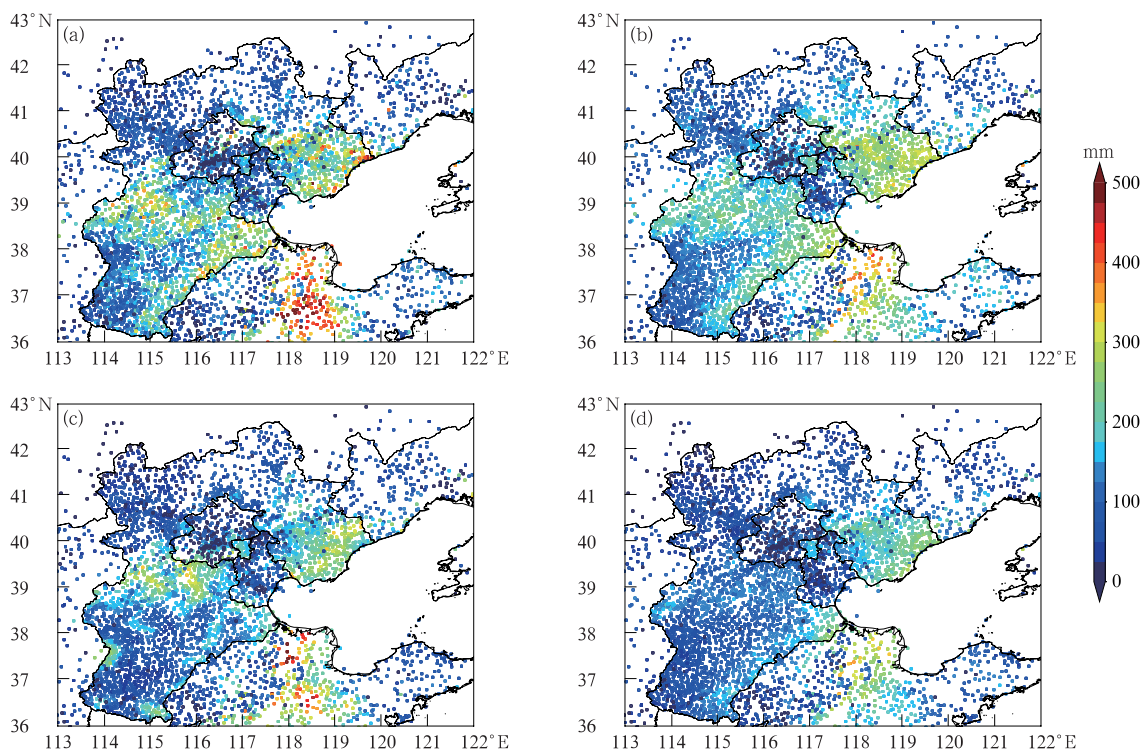


图 8 测试集(2019 年 7 月 21 日至 9 月 1 日)总降水量分布

(a) 观测, (b) LightGBM, (c) CMA-SH9, (d) ECMWF

Fig. 8 Distribution of total accumulated precipitation of test set from 21 July to 1 September 2019

(a) observation, (b) LightGBM correction, (c) CMA-SH9 forecast, (d) ECMWF forecast

3.3.1 SSF 类型降水个例

选取 2019 年 7 月 29 日的一次系统性降水事件。这是一次高空槽和副热带高压共同影响下发生的暴雨,京津冀地区 57 个国家级气象观测站的日降水量达到暴雨以上等级,河北 6 个县(市、区)达到大暴雨等级,单日暴雨影响范围为近 30 年同期(7 月下旬)最大。强降水雨带呈东北—西南向影响京津冀大部分地区。图 9 给出 29 日 14—17 时观测值和两个模式预报及 LightGBM 订正结果。可以看出,ECMWF 对系统性降水的落区有很好的描述,但对河北中部雨区的降水中心预报较分散,同时强度显著偏小;CMA-SH9 预报的强度虽然较大,但落区存在较大偏差,尤其在西部山前存在显著过报,同时漏报了河北东北部的强降水中心。LightGBM(图 9b)在降水落区上更多学习了 ECMWF 的分布,强降水中心的连接性较 ECMWF 有进一步调整,但对北京南部的虚假过报仍采用了 ECMWF 的预报,同时 LightGBM 在强度预报上显著提升,订正结果相对原模式预报达到落区和强度的综合最优。

3.3.2 WSF 类型降水个例

选取 2019 年 8 月 9 日的一次局地性较强的短时强降水事件。图 10 给出 9 日 20—23 时观测值和两个模式降水预报及 LightGBM 订正结果。可以看到,该时刻强降水中心较为分散且范围较小,河北东南部与天津、山东接壤的地方零星有 ≥ 30 mm 的降水中心。ECMWF 对这些零散强降水点几乎没有反映,仅预报出 5 mm 以下的小雨,强度明显偏弱;CMA-SH9 较好地预报出降水落区,但存在显著过度预报,包括强降水中心的范围和强度,以及冀西南山前的虚假降水中心;LightGBM 订正后的结果与观测降水分布更接近,体现出了河北东南部与天津及山东接壤附近的强降水雨带,并体现出一定的零散强降水特征,强度较 ECMWF 显著提升,同时调整了 CMA-SH9 的过度预报,包括消除了其在西南山前的虚假降水中心。可见 LightGBM 能够有效综合调整该降水的强中心分布和强度,但该类型降水的订正难度大于 SSF 类型降水,尤其强度的订正仍然明显偏弱。扩充 WSF 类型降水的真实样本,或对

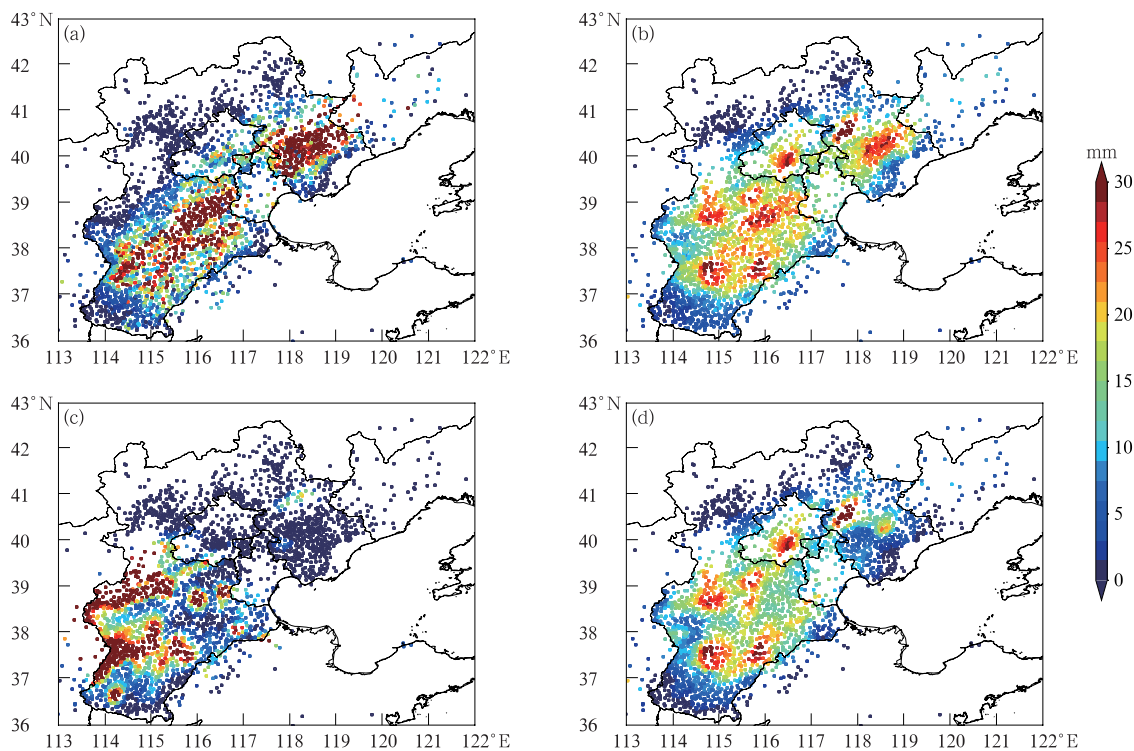


图 9 2019 年 7 月 29 日 14—17 时累计降水量分布

(a) 观测, (b) LightGBM, (c) CMA-SH9, (d) ECMWF

Fig. 9 Distribution of 3 h precipitation from 14:00 BT to 17:00 BT 29 July 2019

(a) observation, (b) LightGBM correction, (c) CMA-SH9 forecast, (d) ECMWF forecast

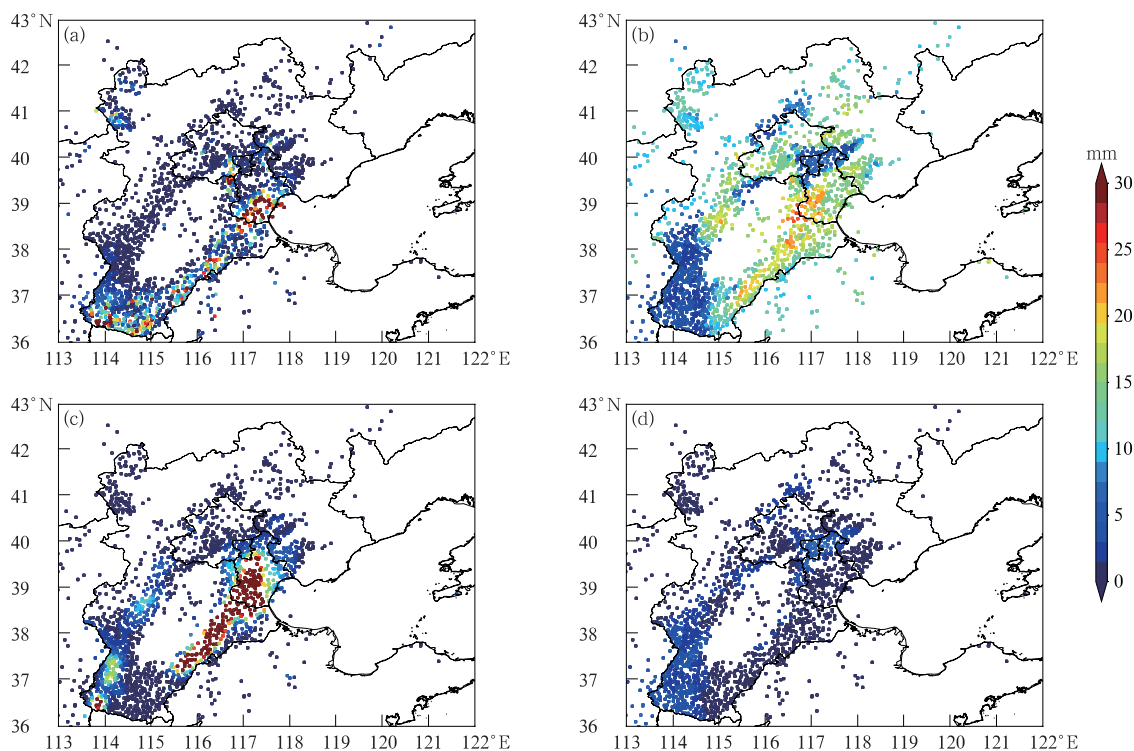


图 10 2019 年 8 月 9 日 20—23 时累计降水量分布

(b) 观测, (b) LightGBM, (c) CMA-SH9, (d) ECMWF

Fig. 10 Distribution of 3 h precipitation from 20:00 BT to 23:00 BT 9 August 2019

(a) observation, (b) LightGBM correction, (c) CMA-SH9 forecast, (d) ECMWF forecast

其单独建模训练是进一步提升其订正效果的可能途径,但单独建模订正在实际应用中还需考虑客观分类的问题。

3.4 输入特征重要性分析

为检验所选降水相关物理特征对 LightGBM 订正效果的贡献度,将表 2 中的四类输入特征去除,只保留模式降水和站点经(纬)度、高程等静态特征,

测试对 LightGBM 订正效果的影响。

图 11a 显示,不论是否采用四类输入特征,LightGBM 订正后所有阈值降水的 TS 评分均相对原数值模式明显提升。四类输入特征对 LightGBM 具有正贡献,如 ≥ 5 mm 降水,当只输入模式降水和地理静态特征时 TS 评分为 0.28,而输入要素特征时评分可提升至 0.35。如果不采用四类输入特征, ≥ 2 mm 降水 TS 评分显著减小,但仍高于两个原模

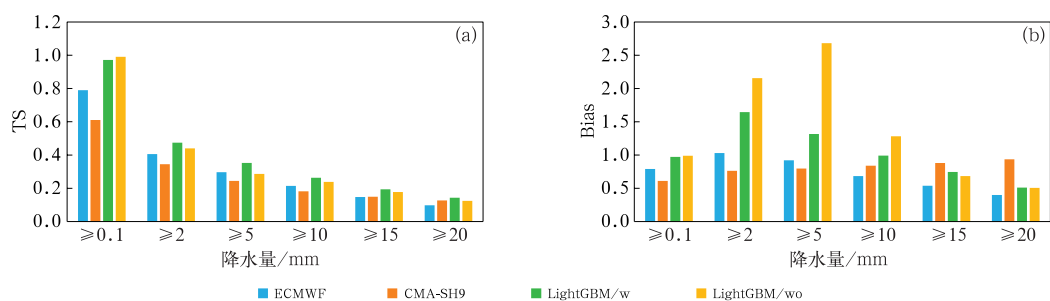


图 11 采用(LightGBM/w)和去除(LightGBM/wo)四类输入特征的 LightGBM、ECMWF、CMA-SH9 降水预报在测试集的(a)TS 和(b)Bias 评分

Fig. 11 (a) TS and (b) Bias scores with (LightGBM/w) and without (LightGBM/wo) four types of input features on test set

式。这种表现可能是由于在没有其他物理特征参考学习的情况下,LightGBM对两种模式降水预报进行了简单融合使其接近观测,因此订正后在中小雨更倾向于向ECMWF预报调整,而在强量级降水上向CMA-SH9调整。此外,图11b显示,采用LightGBM订正后,中小雨Bias显著增大呈现空报,而采用四类输入特征辅助订正可显著减小空报(图11b)。特征重要性分析表明,LightGBM降水订正在所有阈值的TS评分均较原模式预报得到提升,但在中小雨存在较严重空报;通过输入动力、热力和水汽等物理特征可明显改进Bias,同时有效提升模型在各降水量级的预报效果。

4 结论和讨论

降水发生和分布的复杂性和稀有性,使数值预报和客观订正在降水问题极具挑战。本研究通过降水机器学习样本构建和采样、降水相关物理特征输入、残差训练等策略,对京津冀地区累计3h降水量预报进行LightGBM订正,并在独立测试集开展了统计检验、分类型降水的个例评估和特征重要性分析。研究主要结论如下:

(1)面向降水样本长尾分布问题,尽量保持其真实分布和适度的样本处理是LightGBM在强量级降水预报取得改进的关键一环。本文通过降水日筛选和下采样 $<2\text{ mm}$ 降水样本,有效调整了晴雨样本分布和提升了强量级降水的样本比例。下采样结果显示,剔除 $<2\text{ mm}$ 降水样本越多, $\geq 10\text{ mm}$ 降水的TS评分越高,但这种影响在下采样达到50%以后带来的差异不再显著,反而会带来较大的空报和错报。

(2)LightGBM订正在所有阈值降水TS评分均取得较原模式预报显著提升,且取得的提升率随阈值增大而增加, $\geq 2\text{ mm}$ 降水的提升率约为15%, $\geq 20\text{ mm}$ 降水的提升率近50%,也即原模式预报技巧越低情况下,LightGBM订正的空间和提升幅度越大。

(3)分类型降水预报的订正效果表明,SSF类型降水由于占据了降水样本绝大部分,订正效果更优;而WSF类型降水订正相对更难,尤其是对 $\geq 15\text{ mm}$ 降水。WSF类型降水的可预报性较低,相应的模式预报准确率相对更低,有可能为机器学习发挥作用提供了潜在空间,未来如何从模型设计、针对性的物

理特征选取和卫星、雷达等多源观测中提取有益信息是进一步提升WSF类型降水预报的可能途径。

(4)输入与降水密切相关的大气动力、热力和水汽等物理特征,对提升模型在各降水量级的预报有正贡献,同时可明显改进LightGBM订正在中小雨的空报。

本文探索了应用LightGBM订正降水预报的潜力,在测试集统计指标和分类型降水应用均取得较好提升效果。研究也显示,针对降水这一样本极度不平衡的学习任务,以及考虑到降水预报偏差是时间、空间和强度预报误差的综合结果,未来除了本文探索的样本集构建和特征输入外,在模型算法的选择和损失函数调优等很多方面还值得探索。此外,本文分类型降水尚依靠主观判识。针对WSF类型降水这类预报难点问题,对其单独建模是提升预报的可能途径,但仍可能较大依赖模式预报能力和精细观测特征的输入,如果模式预报和观测均不能提前捕获,如何对这种零散的对流性降水“无中生有”的学习,亦是目前尚在探索的前沿领域。

参考文献

- 陈浩,宁忱,南卓铜,等,2017. 基于机器学习模型的青藏高原日降水数据的订正研究[J]. 冰川冻土,39(3):583-592. Chen H, Ning C, Nan Z T, et al, 2017. Correction of the daily precipitation data over the Tibetan Plateau with machine learning models[J]. J Glaciol Geocryol, 39(3):583-592(in Chinese).
- 陈锦鹏,冯业荣,蒙伟光,等,2021. 基于卷积神经网络的逐时降水预报订正方法研究[J]. 气象,47(1):60-70. Chen J P, Feng Y R, Meng W G, et al, 2021. A correction method of hourly precipitation forecast based on convolutional neural network[J]. Meteor Mon, 47(1):60-70(in Chinese).
- 代刊,曹勇,钱奇峰,等,2016. 中短期数字化天气预报技术现状及趋势[J]. 气象,42(12):1445-1455. Dai K, Cao Y, Qian Q F, et al, 2016. Situation and tendency of operational technologies in short- and medium-range weather forecast[J]. Meteor Mon, 42(12):1445-1455(in Chinese).
- 代刊,朱跃建,毕宝贵,2018. 集合模式定量降水预报的统计后处理技术研究综述[J]. 气象学报,76(4):493-510. Dai K, Zhu Y J, Bi B G, 2018. The review of statistical post-process technologies for quantitative precipitation forecast of ensemble prediction system[J]. Acta Meteor Sin, 76(4):493-510(in Chinese).
- 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[J]. 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[J]. Meteor Mon, 43(9):1110-1116(in Chinese).
- 李航,2019. 统计学习方法:第2版[M]. 北京:清华大学出版社. Li H, 2019. Statistical Learning Methods[M]. 2nd ed. Beijing:

Tsinghua University Press(in Chinese).

- 李俊,杜钧,陈超君,2014.降水偏差订正的频率(或面积)匹配方法介绍和分析[J].气象,40(5):580-588. Li J, Du J, Chen C J, 2014. Introduction and analysis to frequency or area matching method applied to precipitation forecast Bias correction [J]. Meteor Mon, 40(5):580-588(in Chinese).
- 李俊,杜钧,陈超君,2015.“频率匹配法”在集合降水预报中的应用研究[J].气象,41(6):674-684. Li J, Du J, Chen C J, 2015. Applications of “frequency-matching” method to ensemble precipitation forecasts[J]. Meteor Mon, 41(6):674-684(in Chinese).
- 罗聪,时洋,吴乃庚,等,2021.基于邻域最优概率的 GRAPES 快速更新循环同化预报系统逐时降水订正方法[J].热带气象学报,37(4):569-578. Luo C, Shi Y, Wu N G, et al, 2021. Correction method based on neighborhood optimal probability for hourly precipitation forecast from GRAPES rapid updating cycle assimilation and forecasting system[J]. J Trop Meteor, 37(4):569-578(in Chinese).
- 申莉莉,李江波,王秀明,等,2024.京津冀暖季短时强降水环境特征对比分析[J].气象与环境学报,40(1):37-46. Shen L L, Li J B, Wang X M, et al, 2024. Comparative analysis of the environmental characteristics of short-duration heavy rainfall in the warm season in the Beijing-Tianjin-Hebei Region[J]. J Meteor Environ, 40(1):37-46(in Chinese).
- 盛杰,金荣花,张小雯,等,2025.临近气象预报大模型“风雷”V1 版本检验及个例评估[J].气象,51(4):389-399. Sheng J, Jin R H, Zhang X W, et al, 2025. Verification and case evaluation of the “Fenglei” V1 meteorological nowcasting model [J]. Meteor Mon, 51(4):389-399(in Chinese).
- 苏翔,袁慧玲,朱跃建,2021.四种定量降水预报客观订正方法对比研究.气象学报,79(1):132-149. Su X, Yuan H L, Zhu Y J, 2021. A comparative study of four objective quantitative precipitation forecast calibration methods[J]. Acta Meteor Sin, 79(1):132-149(in Chinese).
- 孙卓,高嵩,李江波,等,2020.基于 MICAPS4 的灾害性天气个例检索与训练系统设计及应用[J].暴雨灾害,39(2):207-212. Sun Z, Gao S, Li J B, et al, 2020. Design and application of disastrous weather case retrieval and training system based on MICAPS4 [J]. Torr Rain Dis, 39(2):207-212(in Chinese).
- 张晨阳,杨雪冰,张文生,2019.气象大数据超短临精准降水机器学习与典型应用[J].农业大数据学报,1(1):78-87. Zhang C Y, Yang X B, Zhang W S, 2019. Accurate precipitation nowcasting with meteorological big data: machine learning method and application[J]. J Agric Big Data, 1(1):78-87(in Chinese).
- 张海鹏,智协飞,吉璐莹,2020.中国区域降水偏差订正的初步研究[J].气象科学,40(4):467-474. Zhang H P, Zhi X F, Ji L Y, 2020. A preliminary study on calibration in precipitation forecast over China[J]. J Meteor Sci, 40(4):467-474(in Chinese).
- 张娇,王东勇,郑淋淋,等,2021. ECMWF 模式强降水预报偏差订正方法研究及应用[J].暴雨灾害,40(4):430-436. Zhang J, Wang D Y, Zheng L L, et al, 2021. Research and application of Bias correction method for heavy rainfall forecast of ECMWF model [J]. Torr Rain Dis, 40(4):430-436(in Chinese).
- 赵华生,金龙,黄小燕,等,2021.基于 CNN 和 RF 算法的 ECMWF 降水分级订正预报方法[J].气象科技,49(3):419-426. Zhao H S, Jin L, Huang X Y, et al, 2021. ECMWF precipitation grading correction forecast method based on CNN and RF algorithm[J]. Meteor Sci Technol, 49(3):419-426(in Chinese).
- 钟琦,方祖亮,孙卓,等,2022.基于特征工程的强降水物理要素提取及分析[J].计算机科学与应用,12(1):147-157. Zhong Q, Fang Z L, Sun Z, et al, 2022. Using feature engineering to extract important physical parameters for heavy rain [J]. Comput Sci Appl, 12(1):147-157(in Chinese).
- 周志华,2017.机器学习[M].北京:清华大学出版社. Zhou Z H, 2017. Machine Learning[M]. Beijing: Tsinghua University Press (in Chinese).
- Chen Y T, Huang G, Wang Y, et al, 2023. Improving the heavy rainfall forecasting using a weighted deep learning model[J]. Front Environ Sci, 11:1116672.
- Clark A J, Weiss S J, Kain J S, et al, 2012. An overview of the 2010 hazardous weather testbed experimental forecast program spring experiment[J]. Bull Amer Meteor Soc, 93(1):55-74.
- Czernecki B, Taszarek M, Marosz M, et al, 2019. Application of machine learning to large hail prediction-The importance of radar reflectivity, lightning occurrence and convective parameters derived from ERA5[J]. Atmos Res, 227:249-262.
- Ebert E E, 2001. Ability of a poor man's ensemble to predict the probability and distribution of precipitation[J]. Mon Wea Rev, 129(10):2461-2480.
- Espeholt L, Agrawal S, Sønderby C, et al, 2021. Skillful twelve hour precipitation forecasts using large context neural networks[J]. arXiv preprint arXiv:2111.07470.
- Fang X Q, Kuo Y H, 2013. Improving ensemble-based quantitative precipitation forecasts for topography-enhanced typhoon heavy rainfall over Taiwan with a modified probability-matching technique[J]. Mon Wea Rev, 141(11):3908-3932.
- Grinstajn L, Oyallon E, Varoquaux G, 2022. Why do tree-based models still outperform deep learning on tabular data? [J]. arXiv preprint arXiv:2207.08815.
- Hamill T M, Whitaker J S, 2006. Probabilistic quantitative precipitation forecasts based on reforecast analogs: theory and application [J]. Mon Wea Rev, 134(11):3209-3229.
- Han H, Lee S, Im J, et al, 2015. Detection of convective initiation using meteorological imager onboard communication, ocean, and meteorological satellite based on machine learning approaches [J]. Remote Sens, 7(7):9184-9204.
- Hernández E, Sanchez-Anguix V, Julian V, et al, 2016. Rainfall prediction: a deep learning approach[C]//Proceedings of the 11th International Conference on Hybrid Artificial Intelligence Systems. Seville, Spain: Springer.
- Hu Y, Yin F, W Z, 2021. Deep learning-based precipitation bias correction approach for Yin-He global spectral model[J]. Meteor Appl, 28(5):e2032.

- Ke G L, Meng Q, Finley T, et al, 2017. LightGBM: a highly efficient gradient boosting decision tree[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc.
- Peng X, Li Q, Chen L, et al, 2023. A Structured Graph Neural Network for improving the numerical weather prediction of rainfall[J]. *J Geophys Res: Atmos*, 2023, 128(22): e2023JD039011.
- Shwartz-Ziv R, Armon A, 2022. Tabular data: deep learning is not all you need[J]. *Inf Fusion*, 81: 84-90.
- Sønderby C K, Espeholt L, Heek J, et al, 2020. MetNet: a neural weather model for precipitation forecasting[J]. *arXiv preprint arXiv:2003.12140*.
- Voisin N, Schaake J C, Lettenmaier D P, 2010. Calibration and downscaling methods for quantitative ensemble precipitation forecasts[J]. *Wea Forecasting*, 25(6): 1603-1627.
- Zhang X, Yang Y, Chen B, et al, 2021. Operational precipitation forecast over China using the weather research and forecasting (WRF) model at a gray-zone resolution: impact of convection parameterization[J]. *Wea Forecasting*, 36(3): 915-928.
- Zheng J Q, Ling Q, Li J, et al, 2024. Improving the short-range precipitation forecast of numerical weather prediction through a deep learning-based mask approach[J]. *Adv Atmos Sci*, 41(8): 1601-1613.
- Zhong Q, Sun Z, Chen H M, et al, 2022. Multi model forecast Biases of the diurnal variations of intense rainfall in the Beijing-Tianjin-Hebei region[J]. *Sci China Earth Sci*, 65(8): 1490-1509.
- Zhou K H, Sun J S, Zheng Y G, et al, 2022. Quantitative precipitation forecast experiment based on basic NWP variables using deep learning[J]. *Adv Atmos Sci*, 39(9): 1472-1486.
- Zhu Y J, Luo Y, 2015. Precipitation calibration based on the frequency-matching method[J]. *Wea Forecasting*, 30(5): 1109-1124.

(本文责编:何晓欢 戴洋)