

引入预报时效的中国近海海面风速预报深度学习订正方法

周陈羽^{1,2,3}, 张弛^{4*}, 王喜冬^{1,2}, 魏立新⁴, 刘晓燕⁴

(1. 河海大学 自然资源部海洋灾害预报技术重点实验室, 江苏 南京 210024; 2. 河海大学 海洋学院, 江苏 南京 210024; 3. 北部湾大学海洋学院 广西海洋环境灾害过程与生态保护技术重点实验室, 广西 钦州 535011; 4. 国家海洋环境预报中心, 北京 100081)

摘要: 评估了美国国家环境预报中心全球预报系统(GFS)模式和欧洲中期天气预报中心综合预报系统(IFS)模式在中国近海海面风速的预报表现,并基于深度神经网络(DNN)模型对GFS模式的海面风速预报进行订正。结果表明:订正前,GFS模式在中国近海预报时效120 h的海面风速预报均方根误差为3.3 m/s,IFS模式为2.5 m/s,IFS模式的预报性能优于GFS模式;DNN模型可有效降低GFS模式的负系统偏差,多要素联合订正方法具有较好的订正效果。为减小预报时效延长导致的误差累积,提出一种引入预报时效的订正方法。将预报时效作为输入要素的订正模型,相较其他订正模型预报误差进一步减小。与原始GFS模式相比,该订正模型系统偏差减少1.3 m/s,均方根误差减少1.1 m/s,相对误差减少7个百分点,离散指数减少3个百分点。在最高风速区间,引入预报时效的订正模型,其预报时效96 h和120 h的预报负偏差,较IFS模式分别减少1.0 m/s和0.7 m/s,性能优于IFS模式。

关键词: 海面风; 风速; 预报时效; 深度学习; 数值模式; 订正

中图分类号: P732.1 **文献标识码:** A **文章编号:** 1003-0239(2025)05-0001-10

0 引言

随着我国海洋经济的快速发展,海上航运、渔业、海上石油天然气开采等产业日益繁荣,海面风速的准确预报愈发重要。海面风速直接影响海上活动的安全性和效率,尤其在恶劣天气条件下,准确的风速预报可帮助相关行业提前采取防范措施,避免潜在的财产损失和人员伤亡。近年来,海上风能作为可再生能源的潜力持续开发,准确的海面风速预报对于风能资源的评估和开发至关重要。提升海面风速预报准确度,将为我国海洋经济的安全可持续发展提供有力保障。

当前,海面风速的预报主要依赖于数值天气预报模式(Numerical Weather Prediction, NWP),其核心是基于大气动力学和热力学方程的物理模型^[1]。

在陆地上,因观测网络密集,数值模式可以依赖丰富的观测数据进行初始化和校准,因此目前传统物理模型的陆地风速预报精度较高^[2]。然而,海洋区域的观测资料相对稀缺,尤其是远洋地区,主要依赖浮标、船舶观测以及卫星遥感数据。近年来,人工智能(Artificial Intelligence, AI)模型在风速预报中逐步应用,欧洲中期天气预报中心(European Centre for Medium - Range Weather Forecasts, ECMWF)的全球第五代再分析数据集(ECMWF Reanalysis 5, ERA5)是大多数AI模型的主要训练数据集^[3-5]。ERA5再分析数据集采用数据同化(Data Assimilation, DA)方法,通过对数值模式和有限观测进行插值,减小大气要素的计算误差^[6]。因此,开展针对海面风速观测的独立研究仍有必要,以提高预报精度并降低对模式再分析数据的依赖。

收稿日期: 2025-02-17。

基金项目: 2023年广西北部湾海洋环境变化与灾害研究重点实验室开放课题(2023KF04)。

作者简介: 周陈羽(2002-),女,硕士在读,主要从事海洋气象与人工智能预报技术。E-mail: 241317010004@hhu.edu.cn

*通信作者: 张弛(1991-),男,工程师,硕士,主要从事海洋智能预报技术研究。E-mail: zhangchi@nmefc.cn

由于海洋观测资料的稀缺性,数值模式在海面风速预报中常存在一定误差。因此,各类订正与后处理方法被广泛应用,以提升预报精度。多模式集合预报(Multi-Model Ensemble Prediction, MME)通过结合多个预报模式的结果,提供更稳健的风速预测^[7]。以深度学习(Deep Learning, DL)为代表的机器学习(Machine Learning, ML)方法也在海面风速订正中发挥重要作用^[8]。陈德花等^[9]利用后向传播(Back Propagation, BP)人工神经网络对福建沿海风速预报进行订正,取得较好效果。张颖超等^[10]利用极限学习机(Extreme Learning Machine, ELM)算法对天气研究与预报(Weather Research and Forecasting, WRF)模式的预报风速进行订正,减小了风速预报误差。郝雨辰等^[11]基于Resnet50模型构建了江苏近海海面风场预报订正模型,结果表明该订正模型可以显著改善风速预报结果。罗江珊等^[12]对ECMWF的综合预报系统(Integrated Forecasting System, IFS)模式在青岛港区的风场预报进行评估和订正,通过随机森林(Random Forest, RF)模型有效提升了风速预报准确率。

随着机器学习和深度学习方法在风速预报订正中的广泛应用,近年来相关研究更关注训练输入要素的选取^[13]。王国松等^[14]选取21个变量作为预报因子,构建长短期记忆网络(Long Short-Term Memory, LSTM)模型,在沿海风速订正工作中取得理想效果。谢凯文等^[15]提出一种融合多要素时空特征、基于U-net和3D注意力机制的陆地风场订正模型,该模型不仅订正效果较好、且未出现订正结果过于平滑的情况。目前,采用多要素作为模型训练输入的方法已被广泛使用。

在数值模式或AI模式天气预报中,随着预报时效延长,风速预报准确性通常会降低。主要原因是初始条件的不确定性随时间积累,导致预报误差逐渐放大^[16];大气系统本身具有高度非线性特征,即便微小的初始误差,也可能在后续预报中显著影响结果^[17]。此外,数值模式物理过程的简化、观测数据的时效性问题,以及模式误差的积累,都会加剧预报的不确定性^[18]。尽管多要素输入模型已被广泛应用,但在海面风速预报订正中,鲜有将预报时效本身作为模型训练的输入要素。深度神经网络(Deep Neural Networks, DNN)模型通过多层非线性变换

自动从数据中提取特征,具有强大的表达能力和高效率的优化能力^[19]。DNN模型包含多个隐藏层,能够通过深层次的网络结构学习到更加复杂的输入输出关系,这使其拥有更强的表达能力,适用于处理复杂的回归任务^[20]。鉴于预报时效与预报准确性存在负相关关系,可借助DNN模型深入研究不同预报时效带来的影响,以期提升风速预报的准确性。

本研究构建针对中国近海海面风速预报的订正模型,将预报时效作为输入要素纳入模型训练,利用DNN模型能够学习复杂模式、表达各要素间复杂关系的特点,减少模式误差累积以提高预报的准确性。

1 数据

1.1 模式数据

本研究采用美国国家环境预报中心(National Centers for Environmental Prediction, NCEP)全球预报系统(Global Forecast System, GFS)模式的240 h预报时效预报产品,时间覆盖2021—2023年,空间分辨率为 $0.25 \times 0.25^\circ$,时间分辨率为6 h,起报时刻为每日00时(世界时)和12时。该GFS模式预报数据集的预报要素包括10 m纬向风(U_{10})、10 m经向风(V_{10})、海平面气压(Sea Level Pressure, SLP)、100 m纬向风(U_{100})和100 m经向风(V_{100})。同时,以ECMWF的IFS模式风速数据作为对照,其时间覆盖2023年,空间分辨率为 $0.125 \times 0.125^\circ$,时间分辨率同为6 h,起报时刻同为00时和12时,预报要素为 U_{10} 和 V_{10} 。本研究选取GFS模式和IFS模式的预报区域为 $0^\circ \sim 50^\circ \text{N}$, $90^\circ \sim 180^\circ \text{E}$,覆盖中国近海海域。两个模式的海面风速均通过 U_{10} 和 V_{10} 计算得出。对预报数据开展完整性检查等数据质控后,得到不同预报时效的预报数据;各预报时效的模式预报数据均采用双线性插值方法与浮标数据进行时空匹配。

1.2 浮标数据

本研究采用中国近海浮标观测网络中23个水文气象浮标的面风速观测数据作为实况数据,时间覆盖2021—2023年,观测时间分辨率为1 h^[21]。通过对浮标数据开展质量控制以排除错误观测或

不合理观测值,质控后筛选得到 287 952 个海面风速有效观测值。

2 订正模型与评估方法

2.1 DNN 订正模型设计

为使模型训练更加稳定、提升模型性能并加速模型收敛,在构建机器学习模型前,首先对所有训练数据进行归一化处理。将数据转化为 PyTorch 张量,以适配计算需求的输入格式,同时节省内存并提升计算速度。本研究构建的 DNN 模型由若干个全连接层组成(见图 1),输入维度为实际训练输入要素的数量,激活函数采用 ReLU,用于引入非线性特征,避免梯度消失问题。模型以均方误差(Mean Squared Error, MSE, 记作 E_{MS})作为损失函数,计算公式为:

$$E_{\text{MS}} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

式中: y_i 表示模型输出; $\hat{y}_i$ 表示给定的标签数据。模型学习率设定为 0.001,选用 Adam 优化器实现自适应学习率,增强训练稳定性。

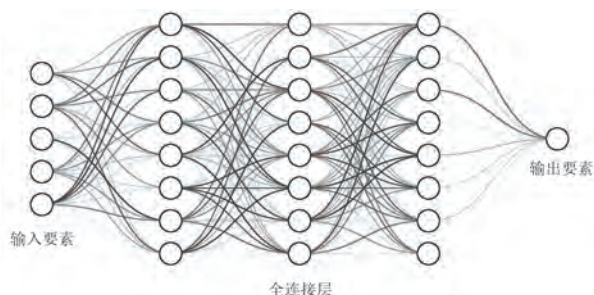


图1 DNN模型架构简略示意图

Fig.1 Schematic diagram of the DNN model

本研究选取 2021—2022 年的 GFS 模式预报数据作为训练集,2023 年 GFS 模式预报数据作为测试集;标签数据采用对应时间的浮标数据,作为训练集和测试集的真值。

本研究通过消融实验开展模型训练,即首先与海面风速最直接相关的 $U10$ 和 $V10$ 作为输入要素,训练至模型性能不再显著提升;再逐步增加训练输入要素,重新训练至模型稳定,并重复上述步骤。最终本研究得到 3 个输入要素不同的 DNN 模型,模型 1 的输入要素为 $U10$ 和 $V10$;模型 2 的输入

要素在其基础上增加 SLP、 $U100$ 和 $V100$ (输入要素共 5 个);模型 3 再纳入预报时效作为输入要素(输入要素共 6 个)。

2.2 预报时效要素处理

本研究所用 GFS 模式预报数据的文件格式均为 .nc,单个 .nc 文件涵盖 1 个初始时刻及 40 个预报时刻,共 41 个时刻;每个时刻与相邻时刻的时间间隔均为 6 h,距离初始时刻最远的预报时刻与初始时刻的时间间隔为 240 h。根据上述特征,本研究将 240 记为最大时间间隔,用于后续对预报时效要素的归一化处理。本研究重点关注的预报时效分别为 24 h、48 h、72 h、96 h 和 120 h,为提取某时刻的预报时效要素,需首先记录该预报时刻与同一文件内初始时刻的时间间隔(即 24 h、48 h、72 h、96 h 和 120 h),再用该时间间隔除以最大时间间隔 240 h,将结果记为预报时效要素的模型输入值。通过此方式处理,可确保预报时效要素的模型输入值最大值不超过 1,以更好地适配多要素训练。

2.3 评价指标

本研究依据实际业务预报需求,对于实际风力等级低于 3 级的情况,不参与预报准确率计算。为比较不同模式以及订正模型对海面风速的预报效果,本研究采用以下 4 个评价指标计算预报风速与观测风速的误差。评价指标分别为系统偏差(Bias)、均方根误差(Root Mean Square Error, RMSE, 记作 E_{RMS})、相对误差(Relative Error, RE, 记作 E_{R})和离散指数(Scatter Index, SI, 记作 I_{S})。计算公式为:

$$\text{Bias} = \frac{1}{n} \sum_{i=1}^n (f_i - o_i) \quad (2)$$

$$E_{\text{RMS}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (f_i - o_i)^2} \quad (3)$$

$$E_{\text{R}} = \frac{1}{n} \sum_{i=1}^n \frac{|f_i - o_i|}{o_i} \quad (4)$$

$$I_{\text{S}} = \sqrt{\frac{\frac{1}{n} \sum_{i=1}^n \{[f_i - \bar{f}] - [o_i - \bar{o}]\}^2}{\bar{o}}} \quad (5)$$

式中: f_i 表示某组时空匹配样本中的预报风速; o_i 代表该组样本中的观测值; n 为样本总数。Bias、RMSE 和 RE 可用来评估预报值与观测值的偏差,而

SI是评估离散程度的相对指标。在评估海面风速预报准确性时,Bias和RMSE越接近0、RE和SI数值越低,则海面风速的预报表现越好。

3 模型评估

3.1 原始GFS模式海面风速预报评估

本研究首先评估原始GFS模式的海面风速预

报在中国近海海域的总体表现(观测风力 ≥ 3 级),评估时间段为2023年,各DNN模型测试集时段与之相同;同时选用ECWMF同期的IFS模式海面风速预报作为对比。结合海面风速预报在实际业务工作中的需求,本文重点评估预报时效分别为24 h、48 h、72 h、96 h和120 h的风速预报准确率。如图2和表1所示,从散点分布可见,两个模式的预报准确性均随预报时效的增加而下降,且同一预报时效的海面

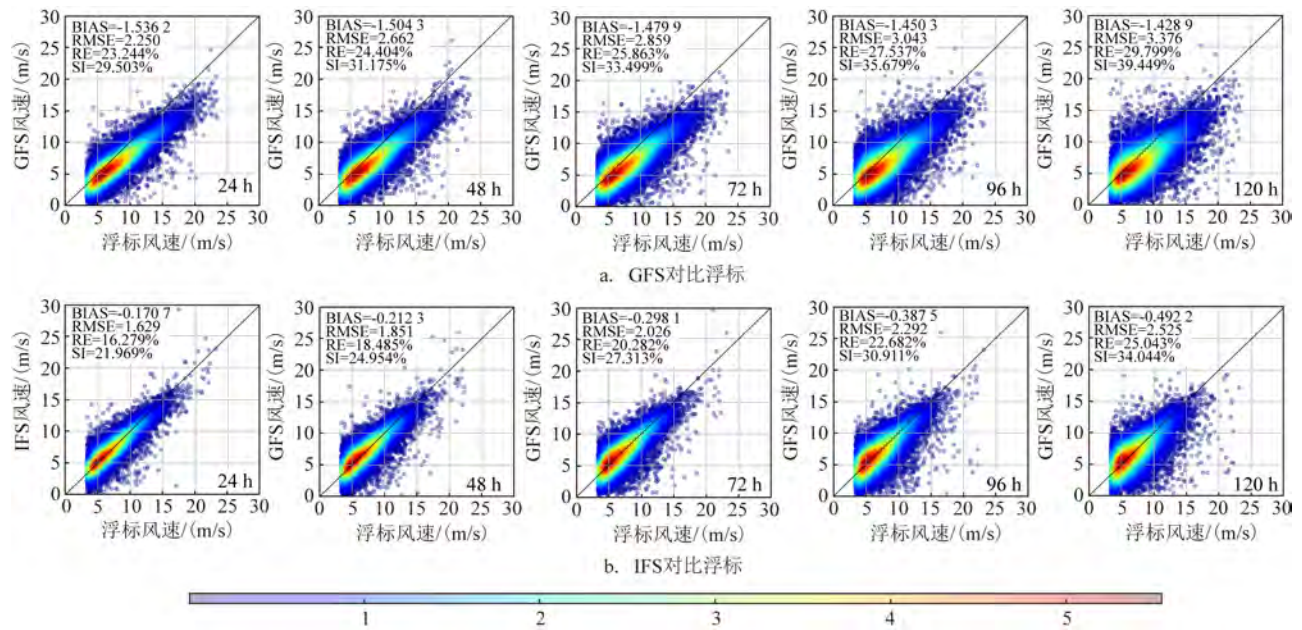


图2 2023年24~120 h GFS和IFS模式风速与实况散点密度图

Fig.2 Scatter density plots of 24~120 hour IFS and GFS wind speed forecasts vs. observations, 2023

表1 24~120 h GFS和IFS预报风速评价指标表

Tab.1 Table of evaluation metrics for 24~120 hour GFS and IFS forecast wind speeds

模式	预报时效/h	系统偏差/(m/s)	均方根误差/(m/s)	相对误差/%	离散指数/%
GFS	24	-1.536	2.520	23.244	29.503
	48	-1.504	2.662	24.404	31.175
	72	-1.480	2.859	25.863	33.499
	96	-1.450	3.043	27.537	35.679
	120	-1.429	3.376	29.799	39.449
IFS	24	-0.171	1.629	16.279	21.969
	48	-0.212	1.851	18.485	24.954
	72	-0.298	2.026	20.282	27.313
	96	-0.388	2.292	22.682	30.911
	120	-0.492	2.525	25.043	34.044

风速预报表现也存在差别。总体而言,IFS模式5个预报时效的预报准确性均优于GFS模式。与IFS模式相比,GFS模式各预报时效的Bias绝对值更大,且GFS模式表现出明显的负系统偏差,其5个预报时效的散点分布形状随风速增大明显偏离45°基准线。

从RMSE来看,GFS模式120 h预报RMSE为3.3 m/s,其96 h预报的RMSE也超过3 m/s;而IFS模式5个预报时效的RMSE均低于3 m/s。这表明IFS模式风速预报在中国近海海域具有一定的可靠性,而GFS模式在中国近海海域的预报能力有待提升。

3.2 DNN模型订正效果

本文构建DNN模型对GFS模式预报风速进行订正。3个不同模型的订正效果(观测风力 ≥ 3 级)见图3和表2。图3a为模型1(24~120 h)的散点密度图,与原始GFS模式相比,模型1的海面风速预报效

果更优;原始GFS模式呈现的负系统偏差得到大幅改善,订正前5个预报时效均的负偏差均超过1 m/s,经模型1订正后,5个预报时效Bias的负偏差均不超过0.2 m/s。同时,模型1各预报时效的RMSE也均有降低,体现出仅以U10和V10作为输入要素,DNN模型依然有较好的订正效果。

图3b展现了模型2的订正效果。如图所示,模型2加入了SLP、U100和V100输入要素,预报准确性进一步提升。从系统偏差来看,模型2在5个预报时效的系统偏差均小于0.1 m/s,表现优于模型1;且模型2各预报时效的RMSE、RE和SI指标较模型1也均有小幅度降低,体现出多要素订正方法相对仅用单一要素订正方法的优越性。

模型3的预报订正效果见图3c。该模型在训练中加入预报时效这一输入要素,其他输入要素均与模型2相同。与模型2相比,模型3整体各指标有小幅提升。24 h预报中,模型3的Bias较模型2减少

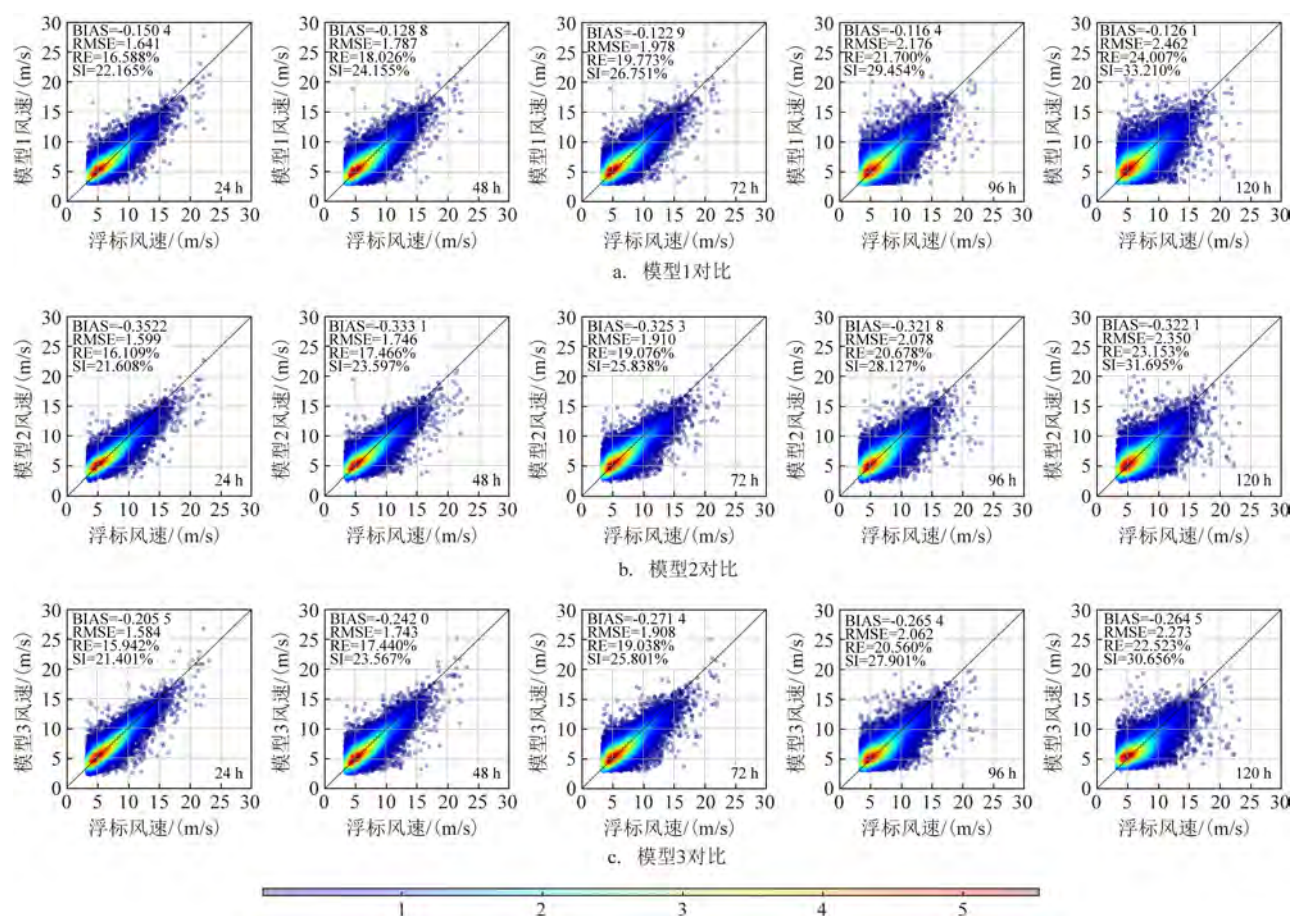


图3 2023年24~120 h 3个DNN模型预报风速与实况散点密度图

Fig.3 Scatter density plots of three DNN models' wind speed in 24~120 hours for the year of 2023

表 2 24~120 h 3 个 DNN 模型预报风速评价指标表

Tab.2 Table of evaluation metrics for 24~120 hour three DNN models forecast wind speeds

模型编号	预报时效/h	系统偏差/(m/s)	均方根误差/(m/s)	相对误差/%	离散指数/%
模型 1	24	-0.150	1.641	16.588	22.165
	48	-0.129	1.787	18.026	24.155
	72	-0.123	1.978	19.733	26.751
	96	-0.116	2.176	21.700	29.454
	120	-0.126	2.462	24.007	33.210
模型 2	24	-0.352	1.599	16.109	21.608
	48	-0.333	1.746	17.466	23.597
	72	-0.325	1.910	19.076	25.838
	96	-0.322	2.078	20.678	28.127
	120	-0.322	2.350	23.153	31.695
模型 3	24	-0.206	1.584	15.942	21.401
	48	-0.242	1.743	17.440	23.567
	72	-0.271	1.908	19.038	25.801
	96	-0.265	2.062	20.560	27.901
	120	-0.265	2.273	22.523	30.656

0.15 m/s, RMSE 减少 0.01 m/s, RE 和 SI 均减少 0.2 个百分点; 48 h 和 72 h 预报中, 模型 3 的提升不明显; 96 h 预报中, 模型 3 的 RE 减少 0.1 个百分点、SI 减少 0.2 个百分点; 在预报时效最长的 120 h 预报中, 模型 3 的 RMSE 减少 0.08 m/s, RE 减少 0.6 个百分点, SI 减少 1.0 个百分点。对比原始 GFS 模式, 模型 3 的 120 h 预报的 Bias 减少 1.2 m/s, RMSE 减少 1.1 m/s, RE 减少 7 个百分点, SI 减少 9 个百分点。由此可见加入预报时效参与训练后, 120 h 预报的准确率提升更为明显。

3.3 不同风力等级评估

将实况数据按不同风力等级划分为 6 组(见表 3), 并将 8 级及以上大风归为同一组。各误差指标对比见图 4。整体来看, GFS 模式和 3 个订正模型的预报误差均随风力等级增强而增大。

从不同的风速区间来看, 3 个订正模型在多个区间的预报误差低于 GFS 模式。低风速区间内, 4 个预报模型的差距较小, 其中 GFS 模式误差最高; 中高风速区间内, 3 个订正模型的各项误差指标均显著优于 GFS 模式。从 Bias 来看(见图 4a 的 24 h 所示), 中高风速区间内, GFS 模式的负偏差随风速增大而增加, 3 个订正模型则对该负偏差有较好改善效果。对比所有预报时刻 6 级和 7 级风区间的

RMSE(见图 4b), 3 个订正模型的 RMSE 普遍比 GFS 模式降低至少 1 m/s, 同样呈现处较好的订正效果。

观察不同预报时效的误差对比。24 h 预报中(见图 4 第一列), 模型 3 的各项误差均最小; 48 h 和 72 h 预报中(见图 4 第二列和第三列), 3 个订正模型的误差指标较为接近, 且高风速区间的预报误差较 24 h 预报有所上升, RE 和 SI 均提高约 5 个百分点; 而 GFS 模式 48 h 和 72 h 预报误差较其 24 h 预报变化幅度较小; 聚焦 96 h 和 120 h 预报(见图 4 第四列和第五列), 3 个订正模型的预报误差依然较为接近, 但在最高风速区间(风力 ≥ 8 级), 模型 3 的各项误差均最小, 其 96 h 和 120 h 的 Bias 负偏差分别为 3.8 m/s 和 4.3 m/s, 而 GFS 模式对应的负偏差分别为 4.8 m/s 和 5.5 m/s。这表明, 虽然随着预报时效的增加, 模型订正效果有所减弱, 但将预报时效作为训练输入要素的订正模型, 在长预报时效的订正效果更优。

为进一步分析长时效预报效果, 本研究选取 96 h 和 120 h 预报, 将表现最优的模型 3 与 GFS 模式、IFS 模式在不同风速区间的误差进行对比。在中低风速区间内(见图 5a 和 5b), 模型 3 的 Bias 接近或不及 IFS 模式; 而在最高风速区间, 模型 3 的 96 h 和 120 h 的负偏差较 IFS 模式分别减少 1.0 m/s 和 0.7 m/s。从 RMSE 来看(见图 5c 和 5d), 模型 3 的 RMSE 在多

表3 实况数据按风力等级划分表
Tab.3 Table of buoy data categorized by wind force levels

风力	3级	4级	5级	6级	7级	≥8级
风速/(m/s)	3.4~5.4	5.5~7.9	8.0~10.7	10.8~13.8	13.9~17.1	≥17.2

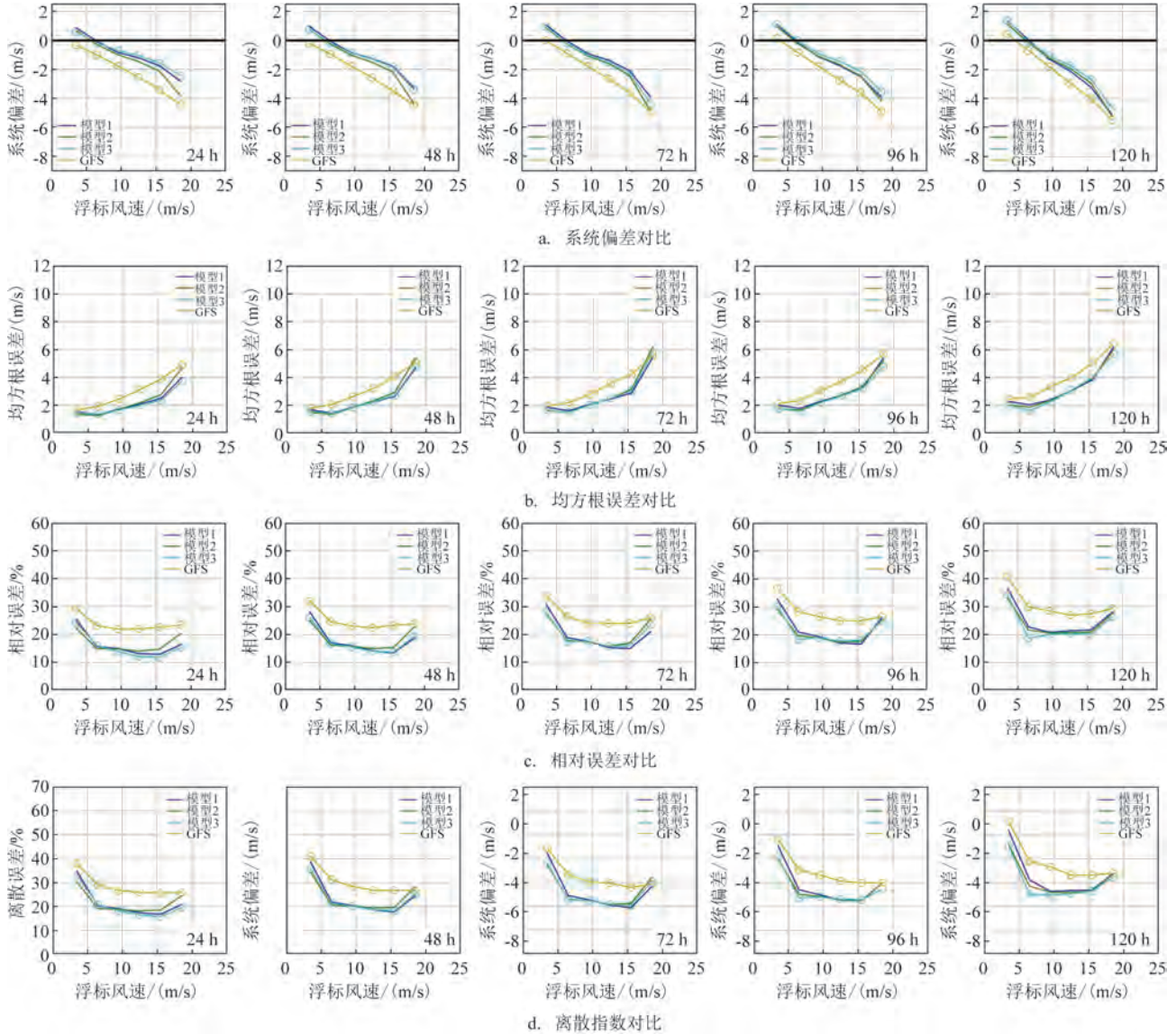


图4 GFS模式与3个DNN模型24~120 h预报的不同风力等级预报误差对比图

Fig.4 Comparison chart of forecast metrics for GFS and three DNN models categorized by wind force levels in 24~120 hours

个风速区间优于IFS模式;在最高风速区间,模型3的96 h和120 h的RMSE分别为4.8 m/s和5.8 m/s,IFS模式对应的RMSE为5.8 m/s和6.2 m/s。可见,在长预报时效的高风速区间,模型3的预报能力优于IFS模式。这进一步表明,针对高风速预报,在模

型训练中加入预报时效要素可提高预报准确率。

3.4 个例分析

为进一步展示模型3对较长时效预报的订正效果,本研究选取2023年浮标数据中时间序列较完整

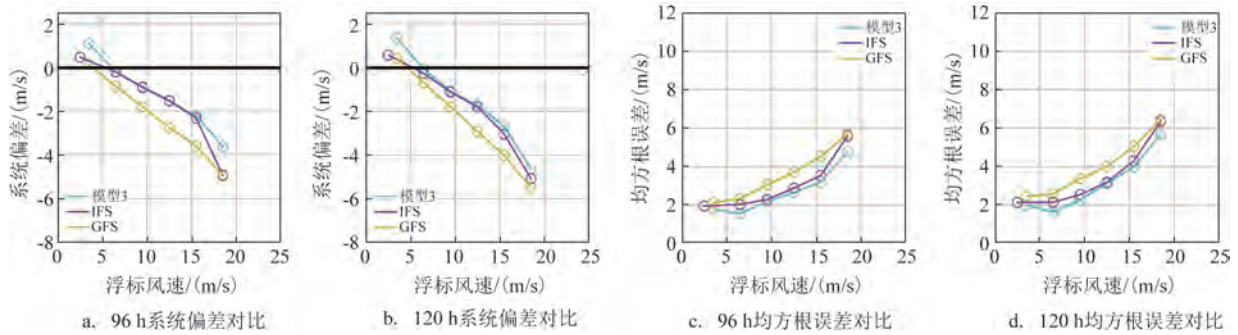


图5 模型3与GFS、IFS模式96 h、120 h预报的不同风力等级预报误差对比图

Fig.5 Comparison chart of forecast metrics for DNN model 3, GFS and IFS categorized by wind force levels in 96 and 120 hours

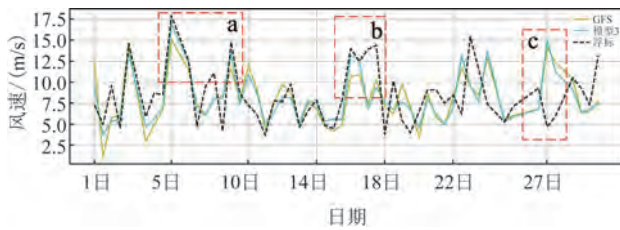


图6 MF01004浮标2023年11月的实况和预报风速时序图
Fig.6 Time series of observed and forecast wind speeds for buoy MF01004 in November 2023

的MF01004浮标,绘制11月的时间序列图。图6主要展示了2023年11月若干次冷空气引发的大风过程。在连续两次冷空气过程中(见图6虚线框a),GFS模式的预报风速均较实况风速偏小,而经优化后的模型3,其预报结果虽仍小于实况,但负偏差显著减小。以虚线框a示意的连续冷空气过程为例:11月5日实况最大风速达17.8 m/s,与GFS模式的偏差约为2.7 m/s,而模型3与实况相比仅偏低了约1.0 m/s;11月9日实况风速为14.7 m/s,GFS模式预报风速11.3 m/s,明显偏低,模型3的预报风速为13.9 m/s,更接近实况风速。

但模型3也存在一定的局限性。从图6虚线框b所示过程中可以看到,GFS模式和模型3对此次过程最大风速的预报不一致,模型3在11月15日预报出6级以上风力,与实况接近,优于GFS模式;但模型3预报11月17日风力减小至4级,与实况偏差较大,折线形态更接近GFS模式。此外,在图6虚线框c所示过程中,两个模式的预报结果与实际观测差异极大。由此说明,即使加入预报时效的订正模型

对长时效预报有一定的改善效果,但其预报表现依旧与原始模式预报紧密相关。当原始预报数据与实际偏离较大时,订正模型的预报表现依旧受限。

4 结论与讨论

本研究基于DNN模型,对GFS模式2023年海面风速预报结果与IFS模式进行对比评估,并开展订正工作。将预报时效作为输入要素加入模型训练中,验证了不同订正模型在海面风速订正方面的可行性。结论如下:

①相较于IFS模式,GFS模式在中国近海海面风速预报中的误差更大,存在较大提升空间。以GFS模式的5种气象要素和预报时效共同作为输入要素构建训练集,基于DNN模型对GFS模式的海面风速进行订正,可有效改进其预报表现。

②仅以 U_{10} 和 V_{10} 作为输入要素训练,模型也具备一定订正效果,预报表现优于原始GFS模式,展现了DNN模型的优秀性能。订正前,GFS模式5个预报时效的负偏差均超过1 m/s;经模型1订正后,5个预报时效的负偏差均不超过0.2 m/s。加入SPL、 U_{100} 和 V_{100} 气象要素作为输入后,模型2的性能优于模型1,论证了多要素联合订正方法在海面风速预报订正的适用性。

③将预报时效作为输入要素加入训练,可提升订正模型在长预报时效下的预报能力。对比原始GFS模式,模型3的120 h预报的Bias减少1.2 m/s, RMSE减少1.1 m/s, RE减少7个百分点, SI减少9个百分点;且模型3在高速区间内的预报表现也有所

提升,在最高风速区间,其96 h和120 h的负偏差较IFS模式分别减少1.0 m/s和0.7 m/s,性能优于IFS模式。由此可见,以DNN为代表的深度学习模型可以挖掘预报时效与预报准确率的负相关关系,进而自适应改进预测结果。本研究提出一种利用深度学习模型降低预报时效累积误差的订正方法,为后续同类型订正工作提供新思路。

但订正模型也存在一定的局限性,其预报表现与原始预报数据紧密相关,即使DNN模型能很好地订正GFS模式的海面风速预报,其预报准确性依旧不及IFS模式;若原始预报数据在某时刻表现较差,订正模型受其影响,订正结果也会存在一定程度的误差。后续研究拟进一步尝试其他方法处理预报时效这一输入要素(如滚动预报等方式),探究机器学习模型对预报时效的把控能力;计划搜集更多预报数据和实况观测数据,并结合AI气象大模型,开展后续研究工作。

参考文献:

- [1] BAUER P, THORPE A, BRUNET G. The quiet revolution of numerical weather prediction[J]. *Nature*, 2015, 525(7567): 47-55.
- [2] DEE D P, UPPALA S. Variational bias correction of satellite radiance data in the ERA-Interim reanalysis[J]. *Quarterly Journal of the Royal Meteorological Society*, 2009, 135(644): 1830-1841.
- [3] BI K F, XIE L X, ZHANG H H, et al. Accurate medium-range global weather forecasting with 3D neural networks[J]. *Nature*, 2023, 619(7970): 533-538.
- [4] KURTH T, SUBRAMANIAN S, HARRINGTON P, et al. Four CastNet: accelerating global high-resolution weather forecasting using adaptive Fourier neural operators[C]//*Proceedings of the Platform for Advanced Scientific Computing Conference*. Davos: Association for Computing Machinery, 2023: 13.
- [5] CHEN L, ZHONG X H, LI H, et al. A machine learning model that outperforms conventional global subseasonal forecast models[J]. *Nature Communications*, 2024, 15(1): 6425.
- [6] KALNAY E. Atmospheric modeling, data assimilation, and predictability[M]. Cambridge: Cambridge University Press, 2003.
- [7] JOSE D M, VINCENT A M, DWARAKISH G S. Improving multiple model ensemble predictions of daily precipitation and temperature through machine learning techniques[J]. *Scientific Reports*, 2022, 12(1): 4678.
- [8] MCGOVERN A, ELMORE K L, GAGNE D J, et al. Using artificial intelligence to improve real-time decision-making for high-impact weather[J]. *Bulletin of the American Meteorological Society*, 2017, 98(10): 2073-2090.
- [9] 陈德花, 刘铭, 苏卫东, 等. BP神经网络在MM5预报福建沿海大风中的释用[J]. *暴雨灾害*, 2010, 29(3): 263-267.
- [10] CHEN D H, LIU M, SU W D, et al. Interpretation and application of BP artificial neural network in MM5 model forecasting gale for coastal regions of Fujian province[J]. *Torrential Rain and Disasters*, 2010, 29(3): 263-267.
- [11] 张颖超, 肖寅, 邓华. 基于ELM的风电场短期风速订正技术研究[J]. *气象*, 2016, 42(4): 466-471.
- [12] ZHANG Y C, XIAO Y, DENG H. Modification technology research of short-term wind speed in wind farm based on ELM method[J]. *Meteorological Monthly*, 2016, 42(4): 466-471.
- [13] 郝雨辰, 杨勤胜, 霍雪松, 等. 基于Resnet50的江苏近海海面风场预报订正方法研究[J]. *海洋预报*, 2024, 41(4): 57-65.
- [14] HAO Y C, YANG Q S, HUO X S, et al. Correction method for surface wind forecasts in Jiangsu offshore area based on Resnet50 [J]. *Marine Forecasts*, 2024, 41(4): 57-65.
- [15] 罗江珊, 杨凡, 毕玮, 等. EC细网格10 m风场产品在青岛港区的预报检验和随机森林订正[J]. *海洋预报*, 2024, 41(3): 110-119.
- [16] LUO J S, YANG F, BI W, et al. Verification and random forest correction of the EC fine-grid model 10 m wind in Qingdao port [J]. *Marine Forecasts*, 2024, 41(3): 110-119.
- [17] FU S, HUANG W Y, LUO J J, et al. Deep learning improves GFS sea surface wind field forecast accuracy in the northwest pacific ocean[J]. *Journal of Geophysical Research: Atmospheres*, 2024, 129(13): e2024JD041188.
- [18] 王国松, 王喜冬, 侯敏, 等. 基于观测和再分析数据的LSTM深度神经网络沿海风速预报应用研究[J]. *海洋学报*, 2020, 42(1): 67-77.
- [19] WANG G S, WANG X D, HOU M, et al. Research on application of LSTM deep neural network on historical observation data and reanalysis data for sea surface wind speed forecasting[J]. *Haiyang Xuebao*, 2020, 42(1): 67-77.
- [20] 谢凯文, 杨昊, 邹茂扬, 等. 一种融合多要素时空特征的数值预报风场订正模型[J]. *软件导刊*, 2024, 23(11): 84-92.
- [21] XIE K W, YANG H, ZOU M Y, et al. A model for correcting numerical wind field forecast with multifactor spatiotemporal features[J]. *Software Guide*, 2024, 23(11): 84-92.
- [22] KÖHL A. Evaluating the GECCO3 1948—2018 ocean synthesis - a configuration for initializing the MPI-ESM climate model[J]. *Quarterly Journal of the Royal Meteorological Society*, 2020, 146 (730): 2250-2273.
- [23] LORENZ E N. Deterministic nonperiodic flow[J]. *Journal of the Atmospheric Sciences*, 1963, 20(2): 130-141.
- [24] DAVIES L A, BROWN A R. Assessment of which scales of orography can be credibly resolved in a numerical model[J]. *Quarterly Journal of the Royal Meteorological Society*, 2001, 127 (574): 1225-1237.
- [25] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. *Nature*, 2015, 521(7553): 436-444.
- [26] HINTON G, DENG L, YU D, et al. Deep neural networks for

acoustic modeling in speech recognition: the shared views of four research groups[J]. IEEE Signal Processing Magazine, 2012, 29 (6): 82-97.

[21] WANG J J, LI B X, GAO Z Y, et al. Comparison of ECMWF significant wave height forecasts in the China sea with buoy data [J]. Weather and Forecasting, 2019, 34(6): 1693-1704.

A study on the deep learning correction method incorporating forecast lead time for sea surface wind speed forecasts along the nearshore areas of China

ZHOU Chenyu^{1,2,3}, ZHANG Chi^{4*}, WANG Xidong^{1,2}, WEI Lixin⁴, LIU Xiaoyan⁴

(1. Key Laboratory of Marine Hazards Forecasting, Ministry of Natural Resources, Hohai University, Nanjing 210024, China; 2. College of Oceanography, Hohai University, Nanjing 210024, China; 3. Guangxi Key Laboratory of Marine Environmental Disaster Processes and Ecological Protection Technology, College of Marine Sciences, Beibu Gulf University, Qinzhou 535011, China; 4. National Marine Environmental Forecasting Center, Beijing 100081, China)

Abstract: This study evaluates the performances of the Global Forecast System (GFS) of the National Centers for Environmental Prediction and the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts in forecasting sea surface wind speed along the nearshore areas of China. Additionally, a deep neural network (DNN)-based correction method is applied to improve the GFS wind speed forecasts along the nearshore areas of China. Results show that, prior to correction, the root mean square error (RMSE) of the 120-hour sea surface wind speed forecasts is 3.3 m/s for the GFS model and 2.5 m/s for the IFS model, with the IFS model demonstrating better performance. The DNN model effectively reduces the negative systematic bias of the GFS forecasts, and the multi-variable correction approach achieves significant improvements. To mitigate error accumulation caused by increasing forecast lead time, this study proposes a correction method incorporating forecast lead time as an input variable. Compared to other correction models, the lead-time-aware correction model further reduces forecast errors. Compared to the original GFS model, this approach reduces systematic bias by 1.3 m/s, RMSE by 1.1 m/s, relative error by 7%, and scatter index by 3%. In the highest wind speed range, the lead-time-aware correction model reduces the negative bias of the 96-hour and 120-hour forecasts by 1.0 m/s and 0.7 m/s, respectively, outperforming the IFS model.

Key words: Sea surface wind; wind speed; forecast lead time; deep learning; numerical weather prediction; correction