

基于机器学习方法的风速偏差订正研究

贺姗姗, 王捷儒, 申彦波, 高金兵

(中国气象局公共气象服务中心, 北京 100081)

摘要: 通过风速偏差订正方法提高模式产品的预报准确性, 为风电场风电功率预测提供可靠的数据支持。以山西陵川县风电场为例, 利用循环神经网络(RNN)、非线性模型(NLinear)、Transformer这3种机器学习方法, 分别建立CMA-WSP2.0模式产品的风速偏差订正模型, 研究结果表明, 3种方法的结果均优于模式产品, 且Transformer、NLinear分别在10 m和100 m高度风速提高准确性方面表现更优。因此, 机器学习方法可有效提升风速数据的质量和可靠性, 为风能资源开发利用、功率预报等领域提供更准确的数据支持。

关键词: 机器学习; 风速; 偏差校正; 数值天气预报; 风电场; 深度学习

中图分类号: P457.5

文献标志码: A

0 引言

近年来, 中国新能源发展速度快, 年均保持两位数的增长率。其中, 风能凭借资源丰富、开发条件成熟等优势, 在新能源体系中占据愈发关键的地位。随着装机规模的不断扩张, 风电并网容量逐年提升, 其在电力系统中的占比也持续增加。截至2024年底, 中国风电累计并网容量达到约5.3亿kW, 2024年新增并网装机容量约8800万kW, 利用率保持在95%以上。在风电业务的规划建设、运行调度与功率管理等环节中, 风速预报精度直接关系到风电场运行效率与经济性, 准确的风速预测在风电场选址、发电量预测以及风电机组优化运行等环节起着关键作用。因此, 对模式产品进行可靠的偏差订正, 对提升风电场预测能力具有重要意义。基于经验公式或简单统计模型对数值模拟产品进行风速订正, 效果通常难以满足发电功率预测需求。随着人工智能技术的快速发展, 机器学习方法已在各行业得到广泛应用, 其具有的强大非线性建模能力, 也为风速偏差订正提供了新的思路和方法。

目前, 随着海量数据的不断积累和应用, 利用机器学习方法对数值模式产品开展误差订正的案例越来越多^[1]。许立兵等^[2]利用随机森林、轻量梯度提升机(light gradient boosting machine, LightGBM)、深度神经网络3种机器学习方法对天气研究与预报模式(weather research and forecasting model, WRF)预报的风场结果进行了订正, 有效改善了数值模式对大风预报的不准确问题; 朱延和等^[3]建立了一种基于U型网络(U-shaped network, U-net)深度神经网络架构的风速预报订正模型, 有效改进了原始预报存在的系统性正偏差问题, 在

天山山脉和昆仑山脉等地形复杂、原始预报误差较大的区域, 改善效果尤为显著; 毛开银等^[4]提出一种累积分布极端梯度提升(cumulative distribution XGBoost, CD-XGBoost)算法, 与传统订正算法相比, 模型的风速预报精度得到了大幅提高; 孙全德等^[5]通过对比分析最小绝对收缩与选择算子(least absolute shrinkage and selection operator, LASSO)回归、随机森林以及深度学习等3种不同机器学习算法, 发现这些方法的风速订正效果均明显优于传统模式输出统计(model output statistics, MOS)统计订正方法; 张颖超等^[6]为进一步提升短期风速预报的准确性水平, 采用极限学习机(extreme learning machine, ELM)算法对WRF模式风速预报结果订正, 有效降低了模式预报的系统性误差和随机误差。

本文以山西陵川县风电场为例, 利用循环神经网络(recurrent neural network, RNN)、非线性模型(NLinear)、Transformer这3种机器学习方法, 分别建立CMA-WSP2.0模式产品的风速偏差订正模型。检验结果表明, 3种方法均能提高模式产品的预报准确性, 其中Transformer、NLinear算法表现最优, 以期风电场风速预报优选结果提供参考思路。

1 数据和方法

1.1 数据来源

本文采用CMA-WSP2.0模式产品作为风速的基础预报数据(未经实测数据订正), 均方根误差在1.0~3.2 m/s之间, 空间分辨率9 km×9 km, 时间分辨率逐15 min, 共25个网格, 最终取25个网格的平均值作为样本的相关特征值。对于连续0值和异常值, 数据记录中以[-999, -888]代替。因本文

收稿日期: 2025-02-20

基金项目: 内蒙古自治区揭榜挂帅项目(2024JBGS0054)

通信作者: 贺姗姗(1982年—), 女, 博士、高工, 主要从事风能太阳能新技术研发及应用方面的研究。hess@12379.cn

采用的 CMA-WSP2.0 模式产品已经过初步的统计订正处理, 可视作为传统订正方法的代表, 本文采用的机器学习方法均被视为在此基础上的进一步优化, 因此本文中所有对比均以 CMA-WSP2.0 模式产品为基准参考。

选定山西省南部新能源转型的示范区域, 晋城市陵川县某个风电场作为研究对象, 代表山地中低风速开发特色的场站。订正的时效范围覆盖 0~72 h 的 10 m 高度以及 100 m 高度的风速预报产品。研究的时间跨度设定为 2022 年 1 月 1 日—2024 年 3 月 1 日。在数据划分方案上, 每月 1 日—25 日的数据用作模型训练集, 每月 25 日—下月 1 日的数据用作模型测试集。数据的完整性要求不低于 90%, 预测精度要求控制在 $\pm(0.1\sim0.2)$ m/s 的范围。

1.2 数据清洗

对模式数据和风电场站观测数据进行清洗, 去除异常和错误数据。同时, 根据先验知识和历史观测数据, 对模式数据进行补充, 减少系统性误差。当遇到用“999”标记的连续 0 值时, 使用 0 替换 999, 以保证特征数据在当前维度上尺度的一致性。当遇到异常数据“-999”和空数据“888”时, 优先查找相同小时和分钟的历史数据, 若该数据为有效数据, 则使用该数据填补当前缺失的数据; 若该数据仍为无效数据, 则查找上一个相同时间点的数据作为补充。考虑数据整体周期性, 原则上采用上一年相同日期和时间的数据作为补充数据, 在数据量不足时, 可酌情采用前一日或临近日期相同时间点的数据作为补充数据。

1.3 特征选取及归一化处理

机器学习特征选取是指从原始数据中选择最具代表性和相关性的特征, 用于训练机器学习模型。特征选取的目的是减少特征空间维度, 提高模型的泛化能力和效率^[7-9]。常用的特征选取方法包括过滤法、包装法和嵌入法。过滤法通过统计方法或相关性分析评估特征重要性, 然后选择最相关的特征。包装法将特征选取问题视为搜索问题, 通过尝试不同的特征子集评估模型性能, 然后选择性能最好的特征子集。嵌入法将特征选取问题融入模型训练过程, 通过正则化等方法约束特征的重要性。特征选取的关键是找到最具代表性和相关性的特征, 以提高模型的预测能力和泛化能力。

本文利用方差阈值法, 阈值设置为 0.1, 选择方差大于阈值的特征, 如小于阈值则表示数据集中变化小, 即对模型贡献较小。因此, 选取以下特征: 实际观测数据选取风速、风向、温度、湿度、气压、风力机风速、风力机功率、总功率作为特征量; 模式产品 CMA-WSP2.0 选取地表向下短波辐射 S_{down} 、地表向下直接辐射 S_{dir} 、地表散射辐射 S_{dif} 、地表向下长波辐射 L_{down} 、地面 10 m 风速 v_{10} 、10 m 风向 θ_{10} 、地面 100 m 风速 v_{100} 、100 m 风向 θ_{100} 、地面 2 m 气温 T_2 、2 m 比湿 q_2 、地表温

度 T_{sk} 、地面气压 p_{sfc} 共 12 个变量作为特征量。

在多维时序模型中, 输入特征往往来自不同来源, 可能具有不同的量级和单位, 特征归一化可将这些不同尺度的特征统一到一个相同的标准中, 从而使模型能够更有效地利用多种特征。同时, 归一化可加快模型的收敛速度。当特征值的范围差异较大时, 优化算法(如梯度下降)的更新步骤可能会变得不稳定, 使得模型难以找到最优解。归一化通过统一特征的尺度, 帮助模型在训练过程中更加平稳地收敛, 提高训练效率。模型的输入特征若未进行归一化, 可能会导致某些特征在训练过程中对模型的影响权重过大或过小, 造成模型的偏置。归一化可确保每个特征对模型的训练有均等的影响, 提高模型的最终表现。

常见的特征归一化方法包括:

最小-最大归一化(min-max normalization): 将特征缩放至 $[0, 1]$ 的范围。Z-score 归一化: 将特征缩放为均值为 0、标准差为 1 的标准正态分布。鲁棒归一化(robust scaling): 使用中位数和四分位数进行缩放, 适合含有异常值的数据。

本文采取 Z-score 归一化, 也称为标准归一化:

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

式中: z ——标准归一化后的值; x ——原始特征值; μ ——特征的均值; σ ——特征的标准差。

2 方法与模型构建

本文利用 RNN、NLinear 和 Transformer 这 3 种机器学习方法构建风速订正模型, 所构建的监督学习框架中, 输入特征 x 为 1.3 节所述的 CMA-WSP2.0 模式产品特征(包括 S_{down} 、 S_{dir} 、 S_{dif} 、 L_{down} 、 v_{10} 、 θ_{10} 、 v_{100} 、 θ_{100} 、 T_2 、 q_2 、 T_{sk} 、 p_{sfc} 共 12 个变量, 取 25 个网格点的平均值)和实际观测的历史特征(风速、风向、温度、湿度、气压、风力机风速、风力机功率、总功率), 输出 y 为对应时刻的实际观测风速(10 或 100 m)。模型采用时序预测框架, 以历史 96 个时刻(24 h)的数据预测未来 0~72 h 的风速。具体流程如图 1 所示。

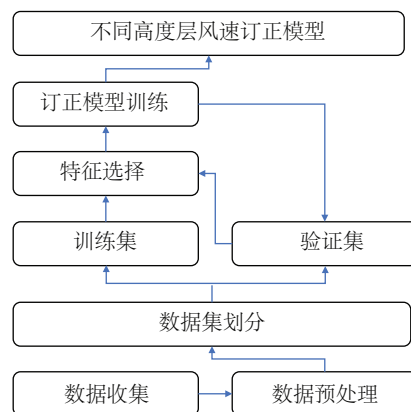


图 1 风速订正流程

Fig. 1 Wind speed correction flowchart

2.1 循环神经网络(RNN)

RNN 是专用于处理序列数据的神经网络架构,在时间序列分析领域展现出优异的性能^[10]。RNN 通过采用循环连接的网络结构设计,允许信息在网络层内的神经元之间进行传递和共享,从而有效挖掘出时序依赖关系,其网络架构如图 2 所示。RNN 的隐藏层内部各神经元节点之间存在相互连接关系,且隐藏层在每个时间步的输入不仅包含了当前时刻输入层的输出信息,还融合了前一时间步隐藏层的状态信息,其最核心特性为网络中各时间步的参数共享。RNN 及其变体在新能源发电功率预测中得到广泛应用^[11-12]。

$$h_t = f(W \cdot x_t + U \cdot h_{t-1} + b) \quad (2)$$

式中: h_t ——当前时刻的隐藏状态(即记忆); f ——激活函数,通常用 tanh 或 ReLU 等; W 、 U ——输入和隐藏状态的权重矩阵; x_t ——当前时刻的输入; h_{t-1} ——上一时刻的隐藏状态; b ——偏置项。

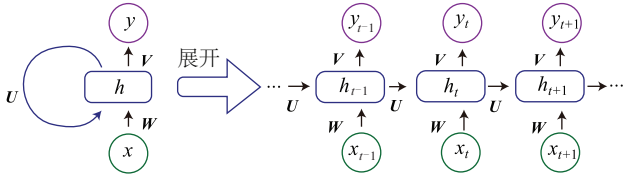


图2 RNN模型网络结构

Fig. 2 Network structure of RNN model

2.2 非线性模型(NLinear)

NLinear 模型的设计理念源自于最基础的 DMS(direct multi-step)预测框架,该模型采用轻量化的线性层结构设计,当数据集表现出分布偏移特征时,该方法首先使用序列的最后一个数值对输入序列进行差分处理,随后将处理后的序列通过一个线性变换层,在生成最终预测结果之前,将之前减去的差分值重新加回预测结果中。NLinear 模型中的加减差分操作可理解为对输入序列实施的一种简单而有效的标准化处理,具体的模型架构框架如图 3 所示。

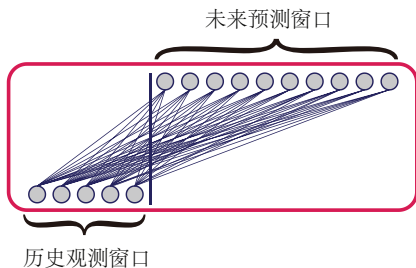


图3 NLinear模型框架

Fig. 3 Framework of NLinear model

NLinear 模型最擅长处理时序任务,能有效挖掘时间序列中长期依赖关系,解决了传统模型在长序列数据处理时面临的信息遗失问题。同时,NLinear 模型能够通过多层网络

结构,自主识别和学习不同时间尺度上的特征,进而提升预测的准确性。

2.3 Transformer模型

Transformer 模型是一种基于自注意力机制的深度学习架构,能根据输入数据的各组成部分的重要程度动态分配不同注意力权重,有效解决了 RNN 等传统序列建模方法在捕获长距离依赖关系以及实现并行化计算方面存在的固有局限性^[7,13-14]。Transformer 架构最初被设计用于自然语言处理(natural language processing, NLP)领域的任务,能够更好地处理长距离依赖关系,提高了模型的训练和推理效率,在时间序列预测方面具有显著优势,模型结构如图 4 所示。

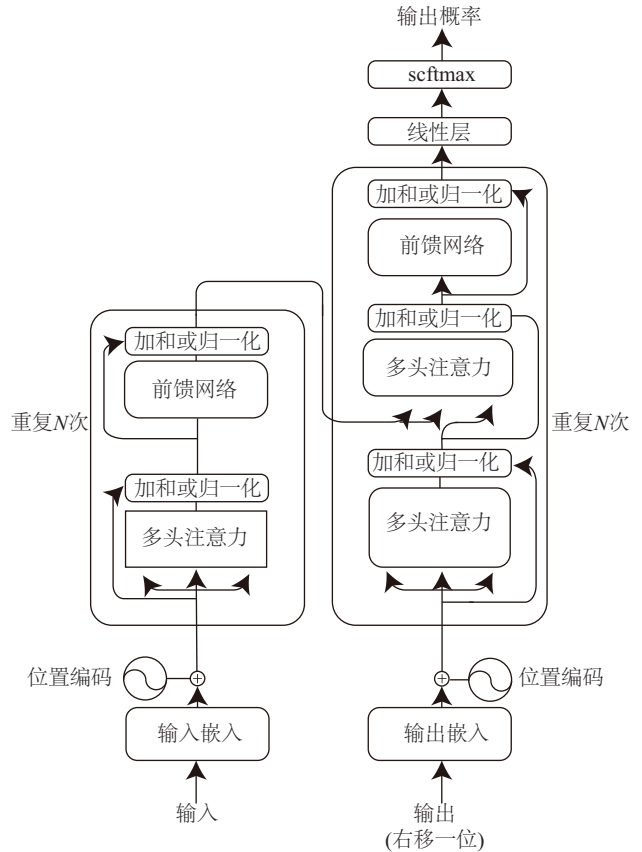


图4 Transformer模型结构

Fig. 4 Architecture of Transformer model

2.4 模型训练

2.4.1 数据分布特征

在训练模型时,分别选取地面 10 m 风速(ws10)和 100 m 风速(ws100)作为目标变量,其他特征作为观测协变量。数据集划分如下:1)训练集:2022 年 1 月 1 日 12:00—2024 年 1 月 1 日 11:45;2)验证集:2024 年 1 月 1 日 12:00—2 月 1 日 11:45;3)测试集:2024 年 2 月 1 日 12:00—3 月 1 日 11:45。

图 5 以 ws100 为例,展示了特征归一化后的数据集划分特征情况。

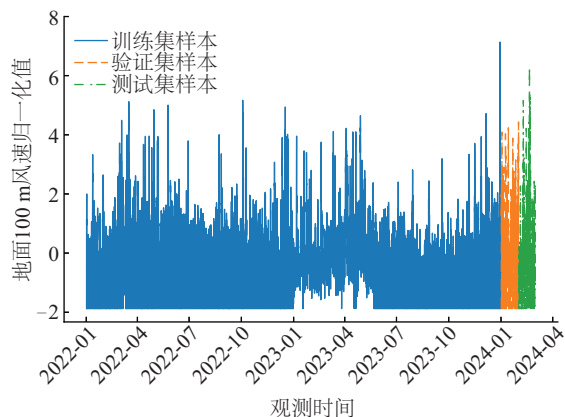


图5 ws100数据划分特征

Fig. 5 Feature map of data partitioning for ws100

由图5可见,数据在ws100维度呈现有季节性的周期性,每年3—7月份风速较高,7—11月份风速相对较低。且全年存在较随机的阵风加大情况。

2.4.2 模型训练

本文在训练过程中,均设置为连续样本输入模型,不做样本抽取,从而保留时序特征。同时,设置最大训练轮数(best epoch)为100轮,并加入训练早停机制,当模型的评估精度连续10轮不再下降时停止训练,即认为达到训练最优参数。

在调参过程中,通过调节学习率(learning rate)和批大小(batch size),比较不同模型参数下的评估精度,获取效果最优的模型。以ws10 RNN为例,表1为调参过程。

表1 ws10 RNN调参数据

Table 1 Parameter tuning data for ws10 RNN

RNN	学习率	批大小	RMSE	MAE	最大训练轮数
1	0.0001	32	1.090	0.747	9
2	0.0001	64	1.038	0.729	13
3	0.0001	128	1.107	0.763	2
4	0.0001	256	1.104	0.763	4
5	0.0001	512	1.084	0.739	17
6	0.0001	1024	1.183	0.778	12
7	0.0001	64	1.077	0.744	3
8	0.0001	32	1.111	0.767	12
9	0.0001	16	1.138	0.770	0
10	0.001	64	1.131	0.808	1
11	0.001	128	1.052	0.772	12
12	0.01	256	1.127	0.779	2

2.5 模型结果评估

本文采用均方根误差(root mean squared error, RMSE)和平均绝对误差(mean absolute error, MAE)两种指标衡量模型预测精度。

RMSE是常用的模型预测性能评价指标,用于量化模型预测值与实际观测值之间偏离程度,表示RMSE的变量 e_{RMS} 具体计算公式为:

$$e_{\text{RMS}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

式中: n ——样本的数量; y_i ——第 i 个实际值; $\hat{y}_i$ ——第 i 个预测值。

MAE是另一项重要的模型预测性能评价指标,用于衡量模型预测值与实际观测值偏差的绝对值,表示MAE的变量 e_{MA} 具体计算公式为:

$$e_{\text{MA}} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4)$$

3 结果分析

3.1 10 m 风层风速预测(ws10)

本文对比结果的测试集均采用2022年9月份的实际观测数据作为对比基线。从结果可看出,在ws10的预测表现上,3个模型的精度均高于CMA-WSP2.0,其中Transformer模型表现略优,其RMSE较CMA-WSP2.0产品降低了45%,MAE降低了42%。预测结果对比如表2、图6所示。

表2 各模型ws10预测误差评估

Table 2 Prediction error evaluation of ws10 across models

评估指标	CMA-WSP2.0 模式	Transformer 模型	非线性 模型	循环神经网络模型
RMSE	2.580	1.416	1.504	2.030
MAE	2.081	1.206	1.342	1.853

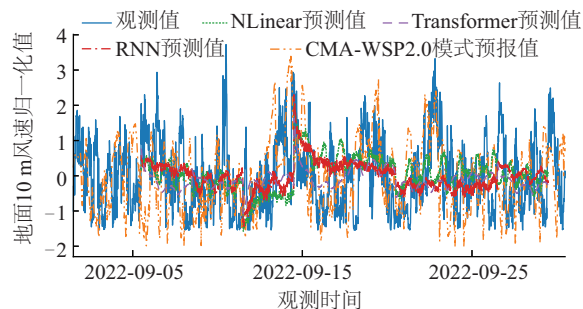


图6 各模型ws10预测值与观测值时序对比

Fig. 6 Time-series comparison of ws10 predicted and observed values across models

3.2 100 m 风层风速预测(ws100)

从结果可看出,在ws100的预测表现上,3个模型的精度

均高于 CMA-WSP2.0,其中 NLinear 模型的精度相对较高,可较准确地预测风速变化趋势,其 RMSE 较 CMA-WSP2.0 产品降低 43%,MAE 降低 33%。预测结果对比如表 3、图 7 所示。

表3 各模型ws100预测误差评估

Table 3 Prediction error evaluation of ws100 across models

评估指标	CMA-WSP2.0 模式	Transformer 模型	非线性 模型	循环神经网络模型
RMSE	3.135	2.135	1.773	2.863
MAE	2.046	1.375	1.117	1.653

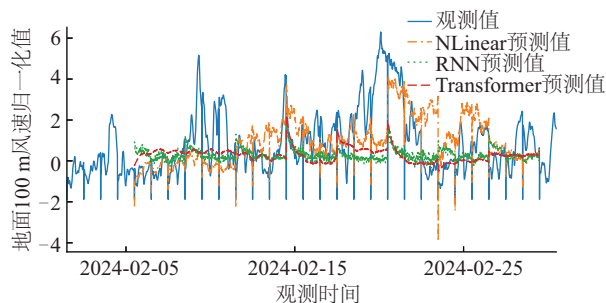


图7 各模型ws100预测值与观测值时序对比

Fig. 7 Time-series comparison of ws100 predicted and observed values across models

因此,风速偏差与气象观测数据之间通常呈现非线性关系,RNN、Transformer、Nlinear 模型能更好地反映这种非线性关系。RNN 虽能处理非线性问题,但在训练过程中易受到梯度消失和梯度爆炸的影响,可能导致模型训练时的性能下降,无法有效学习。

3.3 对比试验结果

为检验模型的移植和泛化能力,选取河北蔚县某风电场进行对比试验,时间范围为 2022 年 1 月 1 日—3 月 30 日,订正 0~72 h 的 10 m 和 100 m 层风速预报,并与实际观测结果进行对比检验。结果表明:在 ws10、ws100 的预测表现上,3 个模型的精度均高于 CMA-WSP2.0,其中 ws10 的预测 Transformer 表现较好;ws100 的预测 NLinear 表现较好,与本文示范场站结果基本一致,说明模型具有可复制性,如表 4 所示。

表4 对比试验结果

Table 4 Comparison test results

评估指标	CMA-WSP2.0 模式	Transformer 模型	非线性 模型	循环神经网络模型
ws10 RMSE	2.780	1.823	2.061	2.278
ws10 MAE	2.281	1.647	1.843	2.114
ws100 RMSE	3.302	2.159	2.019	2.701
ws100 MAE	2.587	1.873	1.741	2.228

本文选择的山西陵川和河北蔚县两个风电场分别代表了山地中低风速和平原风电场的典型特征,两个站点的一致性结果初步验证了方法的普适性。需要说明的是,由于风电场实测数据的获取受限,本文暂时无法开展更大范围的多站点验证,这也是本领域研究尤其是复杂地形风速订正的普遍挑战^[15-17]。未来随着数据共享机制的完善,将进一步在更多类型的风电场进行验证,以全面评估方法的适用范围和局限性。

4 结论与展望

本文采用 RNN、NLinear、Transformer 这 3 种机器学习方法对 CMA-WSP2.0 模式产品的风速偏差进行订正,结果表明,3 种方法的结果均优于模式产品,Transformer 在 10 m 风速表现更优,NLinear 在 100 m 风速表现更优。10 m 和 100 m 模型的预测结果不同,表明模型各自的优势不同,Transformer 在低空小尺度通过自注意力机制能够有效捕捉局部依赖关系,结果则表现更优;而高空大尺度环流,简单的 NLinear 比 Transformer 等复杂模型更具有鲁棒性,结果则更优。在模型复杂度方面,需要指出的是,尽管其自注意力机制能够有效捕捉多个气象要素间的复杂依赖关系,但模型复杂度较高,NLinear 为轻量级模型,RNN 则介于两者之间,在实际业务应用中,应根据计算资源和精度要求选择合适的模型。

因此,合理利用机器学习方法能够有效提升风速数据的质量和可靠性,为风能资源开发利用、功率预报等领域提供了更准确的数据支持。但本文选取场站实测数据仅为 1~2 a,对于订正模型的训练和优化存在一定的影响,偏差订正的结果还有优化调整的空间。未来的研究可考虑融合更多的数据源,如雷达、卫星等数据,以获取更全面的数据信息支持。在机器学习方法方面,未来可探索研究机器学习集合算法,根据不同方法的特点选取最优组合,进一步提高风速预报准确率,为风电场和电网调度提供更高服务质量。

[参考文献]

- [1] 冉靖,张智刚,梁志峰,等.风电场风速和发电功率预测方法综述[J].数理统计与管理,2020,39(6):1045-1059.
RAN J, ZHANG Z G, LIANG Z F, et al. Review of wind speed and wind power prediction methods [J]. Journal of applied statistics and management, 2020, 39(6): 1045-1059.
- [2] 许立兵,孔扬,周峥,等.基于机器学习的风场预报订正方法研究[J].陕西气象,2023(1):15-20.
XU L B, KONG Y, ZHOU Z, et al. Research on prediction and correction method of wind field based on machine learning [J]. Journal of Shaanxi meteorology, 2023(1): 15-20.

- [3] 朱延和, 朱寿鹏, 夏平, 等. 基于U-net神经网络的新疆10 m 风速预报订正研究[J]. 大气科学, 2024, 48(6): 2300-2315.
ZHU Y H, ZHU S P, XIA P, et al. Forecast calibration of 10-m wind speed over Xinjiang based on a U-net neural network [J]. Chinese journal of atmospheric sciences, 2024, 48(6): 2300-2315.
- [4] 毛开银, 赵长名, 何嘉. 基于XGBoost的10 m 风速订正研究[J]. 成都信息工程大学学报, 2020, 35(6): 604-609.
MAO K Y, ZHAO C M, HE J. A research for 10 m wind speed prediction based on XGBoost [J]. Journal of Chengdu University of Information Technology, 2020, 35(6): 604-609.
- [5] 孙全德, 焦瑞莉, 夏江江, 等. 基于机器学习的数值天气预报风速订正研究[J]. 气象, 2019, 45(3): 426-436.
SUN Q D, JIAO R L, XIA J J, et al. Adjusting wind speed prediction of numerical weather forecast model based on machine learning methods [J]. Meteorological monthly, 2019, 45(3): 426-436.
- [6] 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.
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.
- [7] 李练兵, 高国强, 吴伟强, 等. 考虑特征重组与改进Transformer的风电功率短期日前预测方法[J]. 电网技术, 2024, 48(4): 1466-1480.
LI L B, GAO G Q, WU W Q, et al. Short-term day-ahead wind power prediction considering feature recombination and improved transformer [J]. Power system technology, 2024, 48(4): 1466-1480.
- [8] 张淑清, 杨振宁, 姜安琦, 等. 基于EN-SKPCA降维和FPA优化LSTMNN的短期风电功率预测[J]. 太阳能学报, 2022, 43(6): 204-211.
ZHANG S Q, YANG Z N, JIANG A Q, et al. Short term wind power prediction based on EN-SKPCA dimensionality reduction and FPA optimizing LSTMNN [J]. Acta energiae solaris sinica, 2022, 43(6): 204-211.
- [9] 李嘉文, 盛德仁, 李蔚, 等. 基于多目标优化和误差修正的短期风速预测[J]. 太阳能学报, 2022, 43(8): 273-280.
LI J W, SHENG D R, LI W, et al. Short-term wind speed prediction based on multi-objective optimization and error correction [J]. Acta energiae solaris sinica, 2022, 43(8): 273-280.
- [10] 范彤昕, 全洁, 曾美洁, 等. 基于RNN模型的新能源发电功率预测研究[J]. 建模与仿真, 2025, 14(1): 379-387.
FAN T X, QUAN J, ZENG M J, et al. Research on power prediction of new energy power generation based on RNN model [J]. Modeling and simulation, 2025, 14(1): 379-387.
- [11] WANG D, XU M, ZHU G M, et al. Enhancing wind power forecasting accuracy through LSTM with adaptive wind speed calibration (C-LSTM) [J]. Scientific reports, 2025, 15: 5352.
- [12] JOSEPH L P, DEO R C, PRASAD R, et al. Near real-time wind speed forecast model with bidirectional LSTM networks [J]. Renewable energy, 2023, 204: 39-58.
- [13] WU H J, MENG K, FAN D, et al. Multistep short-term wind speed forecasting using transformer [J]. Energy, 2022, 261: 125231.
- [14] 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.
- [15] 许沛华, 陈正洪, 孙延维, 等. 湖北山区复杂地形条件下风电功率预报算法研究[J]. 干旱气象, 2021, 39(3): 524-532.
XU P H, CHEN Z H, SUN Y W, et al. Research on wind power prediction algorithm under complicated terrain in mountainous area of Hubei Province [J]. Journal of arid meteorology, 2021, 39(3): 524-532.
- [16] 杨璐, 宋林烨, 荆浩, 等. 复杂地形下高精度风场融合预报订正技术在冬奥会赛区风速预报中的应用研究[J]. 气象, 2022, 48(2): 162-176.
YANG L, SONG L Y, JING H, et al. Fusion prediction and correction technique for high-resolution wind field in winter Olympic games area under complex terrain [J]. Meteorological monthly, 2022, 48(2): 162-176.
- [17] 庄文兵, 章涵, 王建, 等. 基于中尺度和微尺度的复杂地形大风预报方法研究[J]. 气象科技进展, 2017, 7(2): 13-19.
ZHUANG W B, ZHANG H, WANG J, et al. Study of wind forecasts based on numerical modeling and downscale diagnostics [J]. Advances in meteorological science and technology, 2017, 7(2): 13-19.

RESEARCH ON WIND SPEED DEVIATION CORRECTION BASED ON MACHINE LEARNING METHODS

He Shanshan, Wang Jieru, Shen Yanbo, Gao Jinbing

(Public Meteorological Service Centre, China Meteorological Administration, Beijing 100081, China)

Abstract: By improving the forecasting accuracy of model products through wind speed deviation correction methods, reliable data support is provided for wind power generation forecasting in wind farms. Taking the Lingchuan County Wind Farm in Shanxi Province as an example, three machine learning methods recurrent neural network (RNN), nonlinear model (NLinear), and Transformer were employed to establish wind speed deviation correction models for CMA-WSP2.0 model products. The results show that the results of three methods are better than that of the original model products, and Transformer and NLinear perform better in improving the accuracy of wind speed at heights of 10 m meters and 100 m meters, respectively. Therefore, machine learning methods can effectively improve the quality and reliability of wind speed data, offering a more accurate data foundation for wind energy resource development and utilization, power forecasting and other fields.

Keywords: machine learning; wind speed; bias correction; numerical weather prediction; wind farm; deep learning