

DOI:10.19651/j.cnki.emt.2518584

基于多尺度特征融合的超短期风电功率预测^{*}

高 鹏¹ 庄庆泽¹ 张 飞² 秦 岭¹ 邬锡麟³(1. 内蒙古科技大学数智产业学院 包头 014010; 2. 内蒙古科技大学自动化与电气工程学院 包头 014010;
3. 包钢钢联股份有限公司轨梁轧钢厂 包头 014010)

摘要: 鉴于风电在能源结构中的重要性及其间断性带来的挑战,本文提出了一种基于异常值处理和多尺度特征融合的端到端超短期风电功率多步预测组合模型,旨在提高超短期风电功率预测的精确度与稳定性,进而为电力系统调度与运行的准确性与稳定性提供有力支撑。首先,通过 RobustTSF 方法处理时间序列异常,为预测模型的鲁棒性提供有力的保障,减少了异常时间序列预测和噪声标签学习之间的差异。其次,融合空间金字塔匹配映射策略、Levy 飞行策略以及自适应 t 分布变异策略对蝶蛾优化算法进行改进,显著提高了全局搜索能力和收敛效率。同时,利用多策略蝶蛾优化算法优化改进的 TimeMixer 模型的超参数,以获得最优的模型性能。最后使用 CATimeMixer 模型,实现了多尺度季节特征和趋势特征的融合和预测。实验结果表明,相较于基准模型多层感知机的 MAE、RMSE、MSE 分别下降了 49.71%、41.26%、65.50%,同时 R² 提高了 4.49%,能够有效降低预测误差,为超短期风电功率的准确预测提供了一种新的方法和思路。

关键词: 超短期风电功率多步预测;异常值处理;多尺度特征融合;多策略蝶蛾优化算法

中图分类号: TP183; TN919 文献标识码: A 国家标准学科分类代码: 470.4054

Ultra-short term wind power prediction based on multi-scale feature fusion

Gao Lu¹ Zhuang Qingze¹ Zhang Fei² Qin Ling¹ Wu Xilin³(1. School of Digital and Intelligent Industry, Inner Mongolia University of Science and Technology, Baotou 014010, China;
2. School of Automation and Electrical Engineering, Inner Mongolia University of Science and Technology, Baotou 014010, China;
3. Rail and Beam Rolling Mill, Baotou Steel Joint Stock Co., Ltd., Baotou 014010, China)

Abstract: In light of the significance of wind power within the energy landscape and the challenges posed by its intermittency, this paper proposes an end-to-end, ultra-short-term wind power multi-step prediction model that integrates outlier processing and multi-scale feature fusion. The objective is to enhance the accuracy and stability of ultra-short-term wind power predictions, thereby providing robust support for the reliability of power system scheduling and operation. First, the RobustTSF method is employed to address time series anomalies, providing a strong assurance of the prediction model's robustness and minimizing the disparity between abnormal time series prediction and noise label learning. Secondly, the integration of the spatial pyramid matching mapping strategy, Levy flight strategy, and adaptive T-distribution mutation strategy enhances the dung beetle optimization algorithm, significantly improving its global search capability and convergence efficiency. Meanwhile, the multi-strategy dung beetle optimization algorithm is utilized to optimize the hyperparameters of the enhanced TimeMixer model, resulting in optimal model performance. Finally, the CATimeMixer model is employed to achieve the fusion and prediction of multi-scale seasonal features and trend features. The experimental results indicate that the MAE, RMSE, and MSE decreased by 49.71%, 41.26%, and 65.50%, respectively, compared to the benchmark model multilayer perceptron, while the R² value increased by 4.49%. This demonstrates a significant reduction in prediction error and offers a novel approach for the accurate prediction of ultra-short-term wind power.

Keywords: multi-step prediction of ultra-short term wind power; outlier processing; multi-scale feature fusion; multi-strategy dung beetle optimization algorithm

0 引言

导致了大气污染和温室效应等环境问题,对人类社会和自然环境构成了严重威胁^[1],风电作为一种清洁、可再生的能源,其开发和利用受到了广泛的关注。然而,受多种因素的

在当今全球能源转型的进程中,化石能源的过度消耗

收稿日期:2025-04-14

* 基金项目:国家自然科学基金项目(62161041)、内蒙古自然科学基金项目(2022SHZR0375)、内蒙古自治区重点研发和成果转化项目(2025SYFHH0223)资助

影响, 风电输出功率呈现出显著的波动性和不确定性, 这对电力系统的稳定运行和调度规划提出了严峻挑战。在此背景下, 风电功率的精准预测已成为当前亟待解决的关键问题。超短期风电功率预测作为风电功率预测中的重要一环, 高效、准确的预测模型有利于提升风能的利用率、降低运营成本、保障电网的安全稳定^[2]。总结目前主流的超短期风电功率预测模型, 主要可以分为物理模型、统计模型和深度学习模型^[3]。物理模型基于数值天气预报技术, 模拟风速和气象条件对风电出力的影响, 并使用流体动力学进行精确建模。然而, 物理模型计算复杂, 对初始条件敏感, 需要大量的计算资源^[4]。统计模型依据历史风电功率和气象规律建立非线性映射关系实现功率预测, 但面对较强非线性数据时模型预测效果较差^[5]。随着硬件设施和技术的进步, 深度学习模型能进一步提升预测能力, 例如长短期记忆神经网络(long-short-term memory network, LSTM)^[6], 门控循环单元(gated-recurrent-unit, GRU)^[7], 双向长短期记忆网络(bidirectional long-short-term memory network, BiLSTM)^[8], 时间卷积网络(temporal convolutional network, TCN)^[9]等, 但单一模型在复杂任务中仍存在一定的局限性。研究人员观察到, 组合模型可以进一步提高预测准确性, 从而促使各种组合方法的广泛采用^[10]。

在实际风电功率预测中, 时间序列数据常常受到多种因素的干扰, 导致各种异常值的产生。传统的存在异常数据的时间序列预测方法通常采用“检测-修复-再训练”的流程, 即先通过模型进行异常检测并对异常值进行修复, 或是简单的插补替换。程先龙等^[11]提出了一种结合变分模态分解(variational mode decomposition, VMD)、反向传播(back propagation, BP)、BiLSTM 的组合模型, 该模型首先通过计算相邻数据的平均值来检测并替换原始数据中的异常值, 接着对数据进行归一化处理, 以降低不同数据间的差异和干扰。预处理结束后, 运用 VMD 将历史功率分解为多个不同的模态分量。之后, 将这些模态分量以及对应的气象数据等输入到 BP 和 BiLSTM 的组合模型中, 对各分量进行预测后线性叠加。但这些方法在异常值处理过程中可能引入额外的误差和不确定性, 进而影响预测精度。

此外, 在优化算法研究领域, 近年来涌现的新型优化方法已在诸多优化问题中展现出优异的性能。范斌等^[12]提出了一种对风电场的所有风机进行超短期功率预测的组合模型, 该模型结合了粒子群优化算法(particle swarm optimization, PSO)、支持向量机(support vector machine, SVM)和 AdaBoost 算法。首先, 他们采用基于残差孤立森林的方法对数据进行异常值检测和预处理。随后, 利用 PSO 对 SVM 的参数进行优化。接着, 将训练好的 SVM 作为 AdaBoost 算法的基础回归模型, 通过多次迭代训练多个 SVM 模型, 并优化其权重, 最终构建出一个性能更强的集成回归模型。但是, 这些优化算法在实际应用过程中仍存在陷入局部最优和收敛效率不足等亟待解决的问题。

在风电功率预测领域, 现有研究多采用分解后独立预测各分量再线性叠加的频率尺度特征融合策略, 谢金财等^[13]提出了一种风电功率预测模型 VMD-GRAU。首先运用 VMD 对风电数据进行分解; 通过设置注意门提高 GRU 对关键序列特征的提取能力, 引入误差修正模块以削减风电功率预测的随机波动性, 并在损失函数里嵌入稀疏正则化项, 以防模型过拟合风险。最后将 VMD 分解后的各分量输入到模型中再线性叠加。这种策略缺乏对其他尺度特征之间复杂交互关系的深入挖掘。除频率尺度特征融合外, 部分研究尝试从时空尺度进行特征融合, 任润虎等^[14]提出一种融合静态邻接矩阵和动态邻接矩阵的时空自适应模型。首先, 基于风电场各节点之间的空间距离和差分相似性特征, 预先构建一个静态邻接矩阵。接着, 通过时空自注意力机制提取数据的高维特征, 并生成时空嵌入矩阵, 该矩阵用于指导动态邻接矩阵的生成。最终, 将 GRU 与图神经网络相结合, 融合预先定义的静态邻接矩阵和实时生成的动态邻接矩阵, 利用时空自适应图卷积循环模块完成风电功率的预测任务。但往往面临计算复杂度高、模型可解释性差等问题, 难以在实际工程中大规模推广应用。

本文针对上述问题, 提出了一种基于异常值处理和多尺度特征融合的端到端超短期风电功率多步预测组合模型, 该模型在数据处理、优化算法以及多尺度特征融合架构这 3 个关键环节进行了创新性改进。首先, 在数据处理阶段, 模型采用 RobustTSF 方法对风电功率序列数据进行异常值处理。传统方法在处理异常值时, 往往存在流程复杂和样本选择不合理等问题。而 RobustTSF 方法对风电功率序列数据进行异常值处理, 通过结合趋势分析与样本选择, 有效识别并剔除异常数据, 从而提高模型对原始数据的鲁棒性, 且无需检测-修复-再训练过程, 弥补了异常时间序列预测(abnormal time series prediction, TSFA)和噪声标签学习(noise label learning, LNL)之间的差距。其次, 在优化算法方面, 空间金字塔匹配(spatial pyramid matching, SPM)映射策略、Levy 飞行策略以及自适应 t 分布变异策略改进蜣螂优化算法(dung beetle optimization algorithm, DBO)解决了算法容易陷入局部最优且收敛速度相对较慢的局限性。同时, 利用多策略蜣螂优化算法(multi-strategy dung beetle optimization algorithm, MSDBO)对改进的 TimeMixer 模型的超参数进行优化, 为 CATimeMixer 模型的性能优化提供了有力支持。最后, 在多尺度特征融合架构中, 风电功率预测领域多尺度特征融合方法多侧重于不同频率尺度或空间尺度简单组合, CATimeMixer 从一种新的多尺度特征融合视角来处理时间序列的复杂时间变化, 利用解耦变化和来自趋势和季节尺度特征融合序列的互补预测能力实现了超短期风电功率预测的稳定性和准确性。

1 RobustTSF

异常值是指那些显著偏离其他数据的观测值,它们可能由于测量错误、数据录入错误、系统故障或其他特殊情况而产生。由于异常值对预测模型的影响是多方面的,包括模型拟合偏差、预测精度下降、计算效率降低以及模型鲁棒性降低等,所以通常要对异常值进行处理。

RobustTSF^[15]方法借鉴 LNL,为 LNL 和 TSFA 任务之间搭建了一座桥梁。作为一种模型无关的异常值处理方法,它能够轻松适应不同的深度学习架构,但仍面临着 TSFA 任务中同时处理协变量和目标中的异常的挑战。其创新地摒弃了传统的异常修复步骤,并减少了数据的不连续性问题,通过结合趋势分析与样本选择,直接从原始输入时间序列中识别出信息样本。具体而言,首先对输入数据进行噪声添加和趋势滤波处理,从趋势滤波的结果中提取拟合后的趋势数据。随后计算输入数据和趋势数据之间的误差,并对每个样本的误差计算异常分数。接着根据异常分数筛选出符合条件的样本索引,并基于这些索引从原始数据中筛选出满足条件的样本,确定哪些样本受到异常的影响较小,从而能够更准确地反映时间序列的内在规律,有效提升了模型在面对噪声和异常值时的性能表现,且无需检测-修复-再训练过程。令 $A(\tilde{x}_n)$ 表示 (\tilde{x}_n, s_n) 的异常评分, \tilde{x}_n 为输入序列, s_n 是 x_n 的趋势, 定义如式(1)所示。

$$A(\tilde{x}_n) = \sum_{k=1}^K w(k) \cdot |\tilde{x}_n^k - s_n^k| \quad (1)$$

式中: K 为输入长度, \tilde{x}_n^k 和 s_n^k 分别为 \tilde{x}_n 和 s_n 中的第 k 个值, $|\tilde{x}_n^k - s_n^k|$ 为时间步长 k 的异常程度, $w(k)$ 为连续脉冲函数。

2 MSDBO-CATimeMixer

2.1 MSDBO

在优化算法领域, DBO^[16]作为一种新兴的优化方法,已在多种优化问题中表现出良好的性能。然而,与其他优化算法类似, DBO 在实际应用中仍面临局部最优陷阱和收敛速度较慢等挑战。为解决这些问题,本文提出了 MSDBO, 该算法结合了 SPM 映射策略、Levy 飞行策略和自适应 t 分布变异策略,以增强算法的全局搜索能力和收敛效率。

1) SPM 映射

混沌映射是一种随机的、复杂的方法,用于替代优化算法随机初始化方法,它具有提高算法的随机性、多样性、收敛速度和跳出局部最优解的能力,已被应用于各种优化算法中。SPM 映射通过对优化问题的搜索空间进行多尺度划分,展现出高度的随机性和不可预测性。其核心机制是通过模运算在不同条件下生成均匀分布的随机数序列,从而使蜣螂优化算法中的个体初始化分布更加均匀。这种特性有助于算法更全面地探索解空间,增强全局搜索能力。

2) Levy 飞行

Levy 飞行是由法国数学家 Levy 提出的一种随机游走策略,其步长服从重尾概率分布。与正态分布等常见分布相比, Levy 飞行更频繁地产生较大的步长,从而能够探索更广阔的空间,增强解空间的多样性。通过引入 Levy 飞行策略对小蜣螂觅食的位置信息进行更新,可以使蜣螂优化算法中的个体在搜索空间中广泛分布,进而提升全局寻优能力,避免陷入局部最优, Levy 飞行如式(2)所示。

$$\text{Levy} = 0.01 \cdot \mu \cdot \eta / |v|^{(1/\alpha)} \quad (2)$$

式中: μ, η 表示 $(0, 1)$ 范围内均匀分布的随机数, α 表示 0 为均值 η 为方差的随机数。 η 的计算公式如式(3)所示。

$$\eta = \left(\frac{\Gamma(1+\alpha) \cdot \sin(\pi\alpha/2)}{\Gamma((1+\alpha)/2) \cdot \alpha \cdot 2^{(\alpha-1)/2}} \right)^{(1/\alpha)} \quad (3)$$

式中: $\Gamma(x) = (x-1)!$ 。

3) 自适应 t 分布变异

在 DBO 中,个体在迭代后期倾向于快速聚集到当前最优位置附近,导致种群的搜索能力大幅度减弱,容易陷入最优解。为了解决这种问题,引入自适应 t 分布变异策略,该策略可以根据算法的运行状态和个体的适应度动态调整变异强度,既能提高全局搜索能力,帮助算法跳出局部最优。同时又能够加强局部搜索性能,从而提高解的精度。通过自适应 t 分布变异策略对蜣螂个体最终的位置信息进行更新,t 分布概率函数为式(4)所示。

$$p(x) = \frac{\Gamma(n+1/2)}{\sqrt{n\pi} \cdot \Gamma(n/2)} \cdot (1+x^2/2)^{-(n+1)/2} \quad (4)$$

式中: $\Gamma(\cdot)$ 是为第二型欧拉积分,迭代次数 n 为参数自由度。

2.2 CATimeMixer

TimeMixer^[17]采用了创新的多尺度特征融合架构,由过去可分解融合(past decomposable mixing, PDM)和未来多预测器融合(future multipredictor mixing, FMM)构成,其中, PDM 结构用于序列解耦和多尺度特征融合,而 FMM 结构则用于未来预测。两大结构都采用多层感知机(multilayer perceptron, MLP)架构,能有效处理时序数据在不同时间尺度上的变化,从而实现了超短期预测任务上的卓越性能。

具体而言,由于不同尺度的时间序列表现出不同的特性,精细尺度主要捕捉细节模式,而粗尺度则反映宏观变化。因此,首先通过平均下采样生成多尺度序列 $X = \{X_0, X_1, \dots, X_N\}$, 模型在实现超短期预测任务中固定尺度为 4。尽管真实世界的时间序列具有高度复杂性,即使是最粗尺度的序列也呈现出复杂变化,同时包含明显的季节性和趋势性特征。因为季节性和趋势性在时间序列分析中分别对应短期和长期变化,具有不同的特征属性。所以,通过序列分解结构将多尺度风电功率序列 x^m 分解为季节分量 $T^m = \{t_0^m, t_1^m, \dots, t_N^m\}$ 和趋势分量 $S^m = \{s_0^m, s_1^m, \dots, s_N^m\}$,如式(5)所示,并提出利用堆叠的 PDM 结构,将多尺度风电

功率序列分解后的季节分量和趋势分量分别在不同尺度上进行融合,而非简单地整体混合,如式(6)所示。

$$t_n^m, s_n^m = \text{SeriesDecomp}(x_n^m), n \in \{0, 1, \dots, N\} \quad (5)$$

$$x^m = x^{m-1} + \text{FeedForward}(T - \text{Mix}(\{s_n^m\}_{n=0}^N)) + S - \text{Mix}(\{t_n^m\}_{n=0}^N) \quad (6)$$

式中, m 为第 m 个 PDM 结构, $\text{FeedForward}(\cdot)$ 包含两个线性层, 层之间的 GELU 激活函数, 即 MLP 架构, $T - \text{Mix}(\cdot)$, $S - \text{Mix}(\cdot)$ 表示季节分量和趋势分量融合。在季节分量融合中, 采用自下而上的方法从较低级别的细尺度时间序列向上整合信息, 为粗尺度的季节部分建模补充详细的信息。而对于趋势分量, 则采用自上向下的融合方法, 利用粗尺度的宏观知识来指导更细尺度的趋势建模。

在经过 PDM 结构后进入 FMM 结构, 表示为式(7):

$$\hat{x}_n = \text{predictor}_n(x_n^m), n \in \{0, 1, \dots, N\}, \hat{x} = \sum_{n=0}^N \hat{x}_n \quad (7)$$

式中: \hat{x}_n 表示来自第 m 个风电功率尺度序列的未来预测, 最终输出为 \hat{x} 。 $\text{predictor}_n(\cdot)$ 表示第 m 个尺度序列的预测器, 依然为 MLP 架构, 其中包含两个线性层, 层之间的

激活函数。FMM 是多个预测器的集合, 其中不同的预测器基于来自不同尺度序列的过去信息, 使 FMM 能够集成融合多尺度特征的互补预测能力。

为了进一步提高模型的预测能力, 本文在 TimeMixer 模型的基础上引入了一维卷积和多头自注意力机制, 构建了 CATimeMixer 模型, 其架构如图 1 所示。一维卷积操作通过对时间序列数据进行特征提取, 能够有效地捕捉时间序列中的局部特征和趋势信息。这种局部特征的提取为模型提供了对短期动态变化的敏感性, 使其能够更好地理解数据中的即时模式和周期性变化。与此同时, 多头自注意力机制并行计算多个注意力头, 从不同角度对时间序列进行建模, 从而丰富模型的特征表示能力。它能够更有效地捕捉时间序列中的关键特征和长期依赖关系, 进而提升其对复杂时间序列数据的理解。二者结合使得模型能够在捕捉局部细节的同时, 兼顾全局信息, 从而更全面地理解时间序列的整体结构, 为高精度的预测提供了坚实基础, 提升了模型在处理复杂时间序列数据时的性能和可靠性。

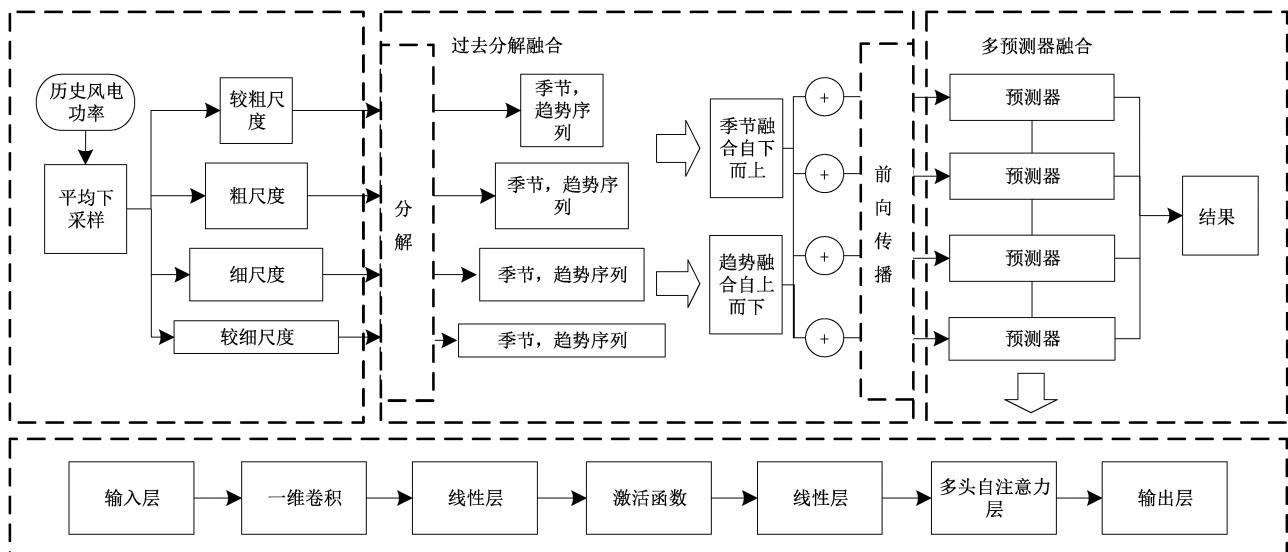


图 1 CATimeMixer 模型

Fig. 1 CATimeMixer model

3 RobustTSF-MSDBO-CATimeMixer 功率预测模型建立

基于上述提出的 RobustTSF、MSDBO 和 CATimeMixer 方法, 本文构建了一个端到端超短期风电功率多步预测组合模型, 通过整体模型直接学习输入和输出之间关系, 无需人为干预, 模型预测流程如图 2 所示, 具体实现步骤如下:

步骤 1) 划分数据集, 按照 8:1:1 的比例分为训练集, 测试集和验证集;

步骤 2) 利用 RobustTSF 方法对训练集进行异常值处理, 即通过计算输入数据与趋势滤波之间的误差来构建异常分数, 并根据异常分数筛选出符合条件的样本索引, 去除异常数据的影响;

步骤 3) 利用本文提及的 MSDBO 训练并优化 CATimeMixer 模型, 找到最佳的模型配置, 验证集用于早停机制, 避免额外开销;

步骤 4) 将测试集输入到优化后的 CATimeMixer 模型中进行超短期风电功率的多步预测, 并通过实验评价指标进行预测误差分析。

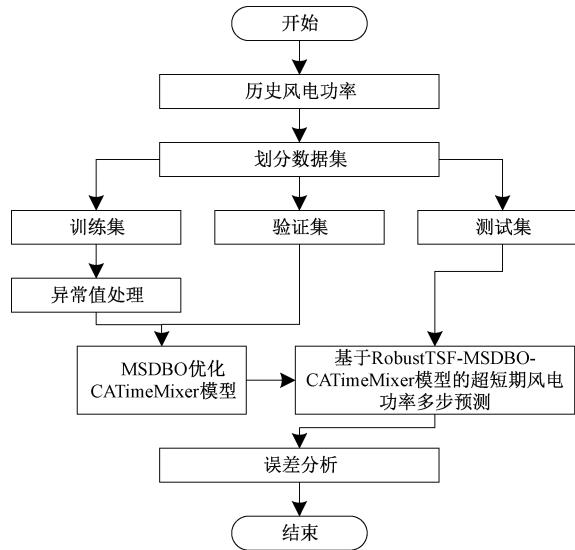


图 2 模型预测流程

Fig. 2 Model prediction process

4 实验

为验证本文提出的超短期风电功率预测模型的有效性,选取内蒙古自治区某 96 MW 风电场 2019 年 1 月 1 日~12 月 31 日的数据作为实验数据集 1。数据集由 35 040 个数据点组成,时间分辨率为 15 min,包括风机 10 m 高度处风速、30 m 高度处风速、50 m 高度处风速、70 m 高度处风速、70 m 高度处风向、风机轮毂高度处风速以及历史风电功率。由于 RobustTSF 方法结合深度学习模型在实现端到端异常时间序列预测任务中仍面临着同时处理协变量和目标中的异常的挑战,所以本实验只能以历史风电功率作为输入数据,将未来风电功率作为输出,分别构建了 1 步、2 步和 4 步的风电功率预测模型,对比实验仍以上述多变量作为输入,风电功率作为输出。模型基于 PyTorch 开发,在 NVIDIA GeForce RTX 4060 GPU 平台上完成训练与测试验证。

4.1 评价指标

为了全面评估模型的预测性能,本文采用了多种评价指标,包括 MAE、RMSE、MSE 以及 R^2 ,这些指标能够从不同的角度反映模型的预测精度和稳定性。具体而言,MAE 和 RMSE 分别能够直观地反映预测值与真实值之间的平均偏差和标准偏差,而 MSE 则通过平方运算放大了较大误差的影响,从而更敏感地捕捉异常预测值,就以上 3 个指标而言,其数值越小,表明模型的拟合程度越好,预测精度越高。 R^2 则反映了模型对数据变异的解释能力,其范围为 $[0, 1]$,越接近 1,则表示模型的解释能力越强,拟合效果越理想。以上 4 种评价指标在式(8)~(11)中定义。

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (8)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (9)$$

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (10)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (11)$$

式中: n 为样本数, \hat{y} 为预测值, y 为真实值, \bar{y} 为平均值。

4.2 优化算法性能测试及模型超参数调优

为了测试所提出的算法改进的有效性,5 种优化算法,即灰狼优化算法 (grey wolf optimization algorithm, GWO)^[18], 鲸鱼优化算法 (whale optimization algorithm, WOA)^[19], 北方苍鹰优化算法 (northern goshawk optimization algorithm, NGO)^[20], 金钱豹优化算法 (golden jackal algorithm, GJO)^[21] 和 DBO 被选择用于测试函数的比较优化。为了保证测试的公平性,每个算法的种群维度统一为 30,最大迭代次数为 500,并在 4 个测试函数上独立测试 30 次,以获得适应度平均值收敛曲线。如图 3 所示,其中 F1~F2 是单峰函数,F3~F4 是多峰函数,可以观察到 MSDDBO 在 4 幅图中的收敛速度上展现出了明显的优势,其收敛曲线不仅更加平滑,而且在优化过程中更不易陷入局部最优解。这一结果表明,改进后的算法在处理复杂的优化问题时,不仅效率更高,而且具有更好的全局搜索能力和稳定性。

CATimeMixer 模型的预测能力与其关键超参数紧密相关,包括模型维度、前馈神经网络维度、学习率以及多头自注意力头数。模型维度决定了输入数据的特征向量大小,影响模型能够捕捉的信息量和表达能力。前馈神经网络维度则与模型的非线性变换能力相关。学习率是控制模型训练过程中参数更新步长的关键因素。多头注意力头数则影响模型捕捉序列中不同位置依赖的能力,帮助模型从不同的角度学习数据特征。因此,本研究采用 MSDDBO 对 CATimeMixer 模型的以上超参数进行优化调参,通过智能搜索最优的参数组合,能够有效提升模型的预测精度和泛化能力。最终,当模型的维度为 128、前馈神经网络维度为 1 024、学习率为 0.000 1 和多头自注意力头数为 4 时,模型达到了其最优的预测性能。

4.3 消融实验

为了验证本文所提出的模型中各个模块的有效性,进行了消融实验。逐级去除了本文所提模型中的模块,观察 MLP、TimeMixer、CATimeMixer、MSDDBO-CATimeMixer 与之模型性能的变化,实验评价指标结果如表 1 所示。为了更直观地展示模型的预测性能,拟合图选取测试集中的 1 000 个点,如图 4 所示,可见每个模块均对模型的预测性

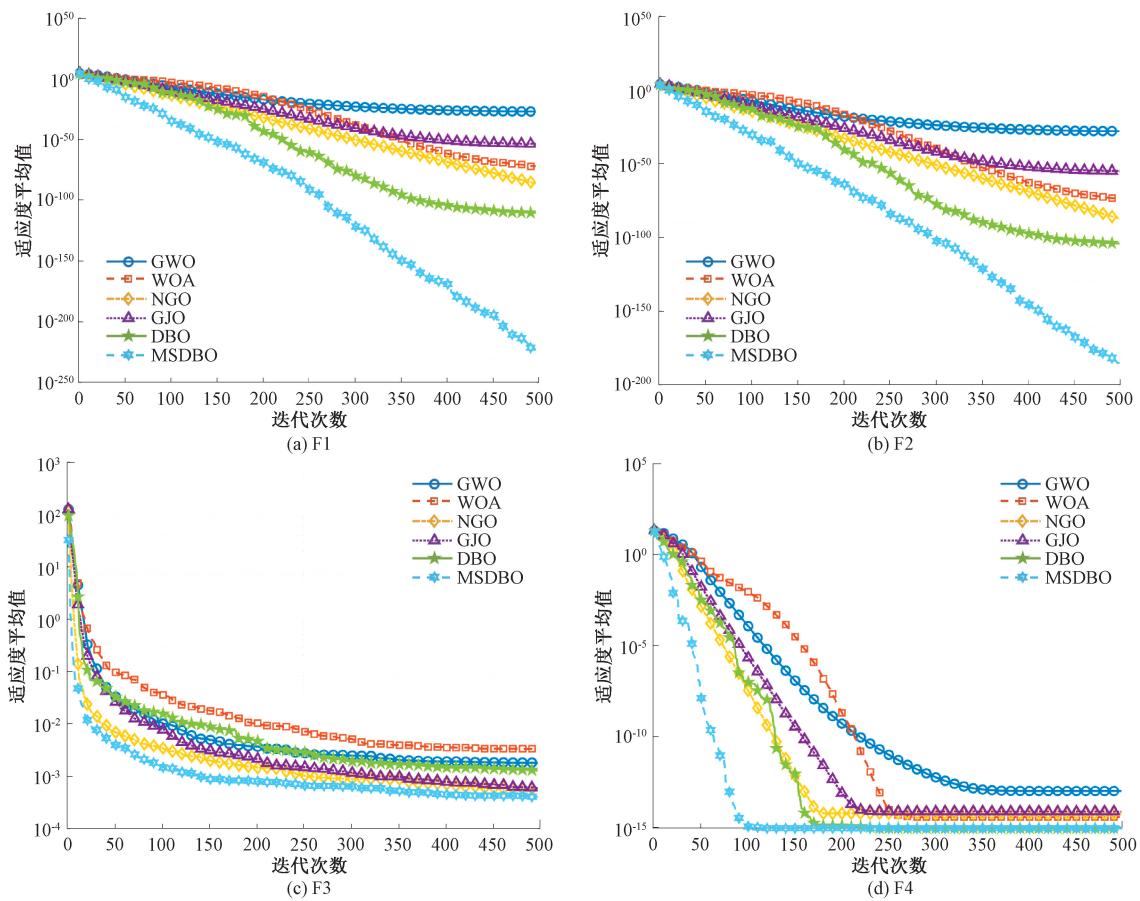


图 3 测试函数收敛曲线
Fig. 3 Test the convergence curve of the function

表 1 消融实验评价指标

Table 1 Evaluation indicators of ablation experiments

预测模型	MAE	RMSE	MSE	R^2
MLP	3.307	5.072	25.726	0.936
TimeMixer	2.107	3.568	12.730	0.968
CATimeMixer	1.857	3.249	10.559	0.973
MSDBO-CATimeMixer	1.761	3.101	9.617	0.976
本文所提模型	1.663	2.979	8.876	0.978

能具有重要贡献, 缺少任一模块都会导致预测精度显著下降。具体来说, 本文所提模型相较于基准模型 MLP 的 MAE、RMSE、MSE 分别下降了 49.71%、41.26%、65.50%, 同时 R^2 提高了 4.49%, 这一结果充分证明了本文所提出模型结构的合理性和有效性, 表明各模块在提升模型整体性能中均发挥了不可替代的作用。

为了全面评估所提出的模型的鲁棒性和泛化能力, 在消融实验的基础上, 进一步选取了中国南方某 96 MW 风电场 2019 年 1 月 1 日~12 月 31 日的数据作为实验数据集 2。通过在这一全新的数据集上进行验证, 能够更客观

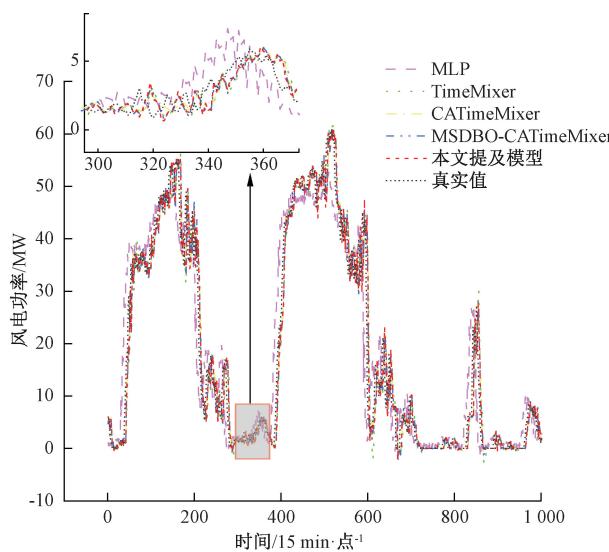


图 4 消融实验结果拟合图
Fig. 4 Fitting graph of ablation experiment results

地评估模型的泛化能力, 确保其在面对未知数据时仍能保

持良好的预测性能和鲁棒性表现。从表 2 可以看出,尽管两个数据集在地理位置、气候条件上存在部分差异,但所提出的模型在两个数据集上的 R^2 差异仅为 0.613%。这一微小的差异表明,模型在不同数据集上的性能基本一致,证明了其较强的鲁棒性和泛化能力,为模型在更广泛的实际应用中提供了有力支持。

表 2 不同数据集所提模型评价指标

Table 2 Model evaluation metrics proposed for different datasets

数据集	MAE	RMSE	MSE	R^2
数据集 1	1. 663	2. 979	8. 876	0. 978
数据集 2	3. 294	5. 337	28. 486	0. 972

4.4 多步预测对比实验

为了进一步在数据集 1 中验证本文模型的优势,将本文所提模型与其他当前广泛应用于超短期风电功率预测领域的主流预测模型进行了多步预测对比实验,这些模型包括 BiLSTM^[22]、TCN^[23]、Nons-Transformer^[24] 以及逐步加入分解算法和优化算法的 CNN-LSTM^[25] (VMD-CNN-LSTM、VMD-DBO-CNN-LSTM),实验评价指标结果如表 3 所示。为了更直观地展示模型的预测性能,拟合图取测试集中的

1 000 个点,如图 5、6、7 所示,可以看出随着预测步长的增加,所有模型在处理复杂且波动较大的风电功率数据时性能有所下降,但本文所提模型在 1 步、2 步和 4 步中风电功率预测中均表现出最高的预测精度且真实值与预测值拟合效果更好。具体而言,在 4 步预测中,相较次优网络 VMD-DBO-CNN-LSTM 的 MAE、RMSE、MSE 分别降低了 36.64%、18.25%、33.16%,同时 R^2 提升了 1.14%;在 2 步预测中,相较次优网络非平稳 Transformer 的 MAE、RMSE、MSE 分别降低了 36.58%、26.67%、46.24%,同时 R^2 提升了 0.92%;在 1 步预测中,相较次优网络 DBO-CNN-LSTM 的 MAE、RMSE、MSE 分别降低了 23.92%、19.10%、34.71%,同时 R^2 提升了 0.3%,这充分证明了本文模型在超短期风电功率预测中的优越性。本文所提模型的优势主要体现在以下几个方面:RobustTSF 有效处理异常数据,为预测模型提供了可靠的数据基础。MSDBO 通过高效的超参数优化,提升了模型的收敛速度和精度。CATimeMixer 从一种新的多尺度特征融合视角来处理不同尺度序列特征的复杂变化,结合卷积和注意力机制,充分捕捉了时间序列的局部特征和全局依赖。综上所述,RobustTSF-MSDBO-CATimeMixer 模型通过以上 3 种优势的结合,在超短期风电功率预测中表现出较高的精度和稳定性,为解决实际应用中的风电功率预测问题提供了一种潜在的有效方案。

表 3 对比实验评价指标

Table 3 Evaluation indicators of comparative experiments

预测步长	评价指标	BiLSTM	TCN	DBO-CNN-LSTM	VMD-DBO-CNN-LSTM	非平稳 Transformer	本文所提模型
1 步	MAE	1. 426	1. 583	1. 176	2. 076	1. 103	0. 895
	RMSE	2. 225	2. 436	1. 901	2. 949	1. 917	1. 538
	MSE	4. 953	5. 932	3. 615	8. 700	3. 625	2. 360
	R^2	0. 988	0. 985	0. 991	0. 974	0. 990	0. 994
2 步	MAE	2. 604	2. 554	1. 848	2. 205	1. 282	1. 172
	RMSE	3. 908	3. 902	2. 812	2. 937	2. 238	2. 062
	MSE	15. 269	15. 229	7. 907	8. 627	5. 009	4. 251
	R^2	0. 962	0. 963	0. 980	0. 979	0. 987	0. 989
4 步	MAE	2. 951	3. 152	2. 914	2. 625	2. 194	1. 663
	RMSE	4. 581	4. 714	4. 414	3. 644	3. 696	2. 979
	MSE	20. 984	22. 221	19. 486	13. 279	13. 659	8. 876
	R^2	0. 948	0. 945	0. 952	0. 967	0. 966	0. 978

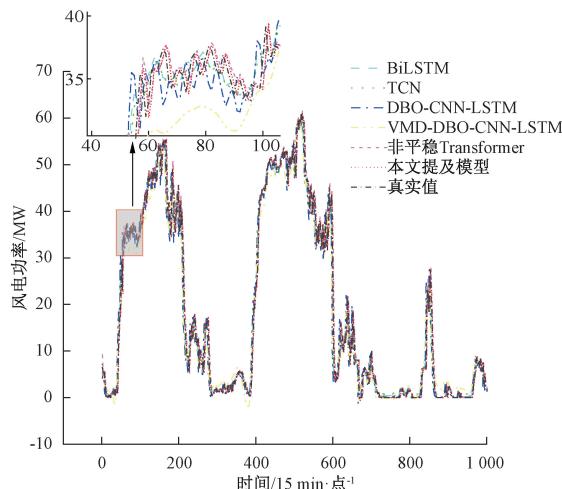


图 5 1 步预测结果拟合图

Fig. 5 One-step prediction result fitting graph

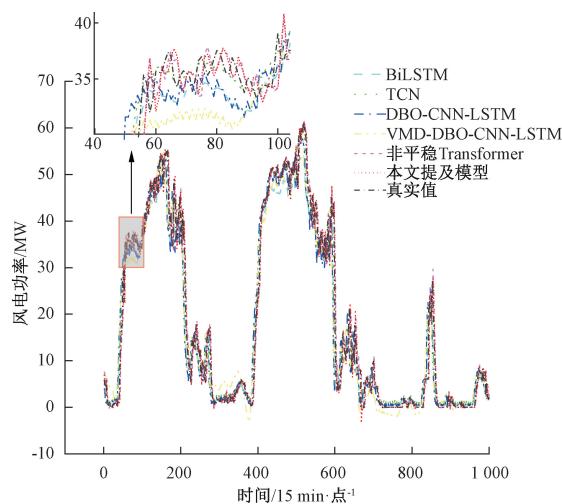


图 6 2 步预测结果拟合图

Fig. 6 Two-step prediction result fitting graph

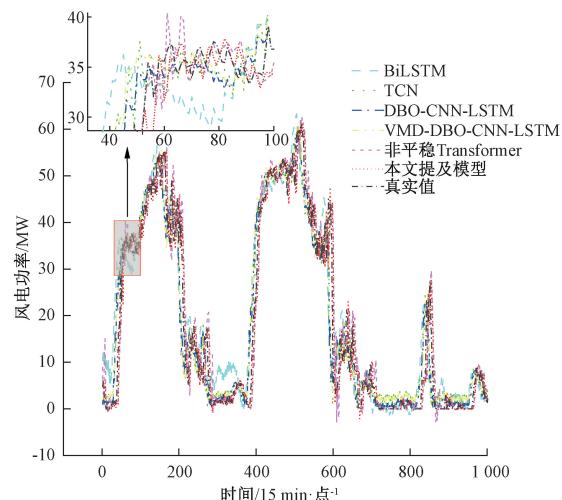


图 7 4 步预测结果拟合图

Fig. 7 Four-step prediction result fitting graph

5 结 论

本文提出了一种基于异常值处理和多尺度特征融合的端到端超短期风电功率多步预测组合模型 RobustTSF-MSDBO-CATimeMixer, 模型创新性地结合数据处理、优化算法和多尺度特征融合架构三大核心方法, 实现了在复杂风电功率时间序列数据中的准确预测。RobustTSF 方法的引入, 不仅为模型在异常数据环境下的鲁棒性提供了有力保障, 还简化了数据处理流程, 为风电功率预测奠定了坚实的数据基础。同时, MSDBO 凭借其高效的优化策略, 为模型超参数的优化提供了强大支持, 进一步提升了模型的性能。此外, TimeMixer 模型中的一维卷积与多头自注意力机制的结合, 使得模型在处理复杂时间序列数据时能够更好地捕捉局部与全局特征, 显著提高了预测的稳定性和准确性。在实际数据集上的应用验证了该模型的高效性、鲁棒性和泛化能力, 为电力系统的调度和运行提供了有力的技术支持。

然而, 尽管 RobustTSF-MSDBO-CATimeMixer 模型在当前研究中取得了显著进展, 但仍存在一些有待进一步完善的地方。当前模型主要关注目标变量中的异常值处理, 对于协变量中的异常问题尚未解决。在未来的研究中, 应进一步优化 RobustTSF 方法, 以实现对协变量和目标变量中异常问题的协同处理, 从而提升模型在更复杂环境下的适应性和预测精度。

参 考 文 献

- [1] ZHUANG Q Z, GAO L, ZHANG F, et al. MIVNDN: Ultra-short-term wind power prediction method with MSDBO-ICEEMDAN-VMD-Nons-DCTransformer net[J]. Electronics, 2024, 13(23): 4829.
- [2] 王颖, 刘兴杰, 梁英, 等. 一种基于 MGWO-Informer 的超短期风电功率预测方法[J]. 太阳能学报, 2024, 45(11): 477-485.
- [3] WANG J, LIU X J, LIANG Y, et al. An ultra-short-term wind power prediction method based on MGWO-Informer[J]. Acta Energiae Solaris Sinica, 2024, 45(11): 477-485.
- [4] LI C, XIAO B J, YUAN Q P. Ultra-short-term wind power prediction algorithm based on bidirectional neural controlled differential equations[J]. Electric Power Systems Research, 2025, 243: 111479.
- [5] LIU Z X, GUO H, ZHANG Y SH, et al. A comprehensive review of wind power prediction based on machine learning: Models, applications, and challenges[J]. Energies, 2025, 18(2): 350.
- [6] 蒲晓云, 杨靖, 杨兴, 等. 基于分解技术的 IZOA-Transformer-BiGRU 短期风电功率预测[J]. 电子测量技术, 2025, 48(2): 39-48.

PU X Y, YANG J, YANG X, et al. Short-term wind power prediction of IZOA-Transformer-BiGRU based on decomposition technology [J]. Electronic Measurement Technology, 2025, 48(2): 39-48.

[6] WANG J ZH, KOU M G, LI R Z, et al. Ultra-short-term wind power forecasting jointly driven by anomaly detection, clustering and graph convolutional recurrent neural networks [J]. Advanced Engineering Informatics, 2025, 65: 103137.

[7] 柳天虹,齐胜利,裔扬,等.基于分位数回归的改进权重GRU风电功率区间预测[J].太阳能学报,2024,45(12):291-298.

LIU T H, QI SH L, YI Y, et al. Improved weighted GRU wind power range prediction based on quantile regression[J]. Acta Energiae Solaris Sinica, 2024, 45(12): 291-298.

[8] 冯俊磊,吕卫东,段雪艳,等.基于模态分解和TCN-BiLSTM的风电功率预测[J].电子测量技术,2024,47(14):49-56.

FENG J L, LYU W D, DUAN X Y, et al. Wind power prediction based on modal decomposition and TCN-BiLSTM[J]. Electronic Measurement Technology, 2024, 47(14): 49-56.

[9] 张中丹,李加笑,冯智慧,等.基于RF-TCN-SA及误差修正的风电功率超短期预测[J].电网与清洁能源,2025,41(2):113-119.

ZHANG ZH D, LI J X, FENG ZH H, et al. Ultra-short-term prediction of wind power based on RF-TCN-SA and error correction [J]. Power Grid and Clean Energy, 2025, 41(2): 113-119.

[10] ZHU G P, JIA W Q, XING ZH T, et al. CMLLM: A novel cross-modal large language model for wind power forecasting [J]. Energy Conversion and Management, 2025, 330: 119673.

[11] 程先龙,张杰,李思莹,等.基于VMD-BP-BiLSTM的短期风电功率预测[J].综合智慧能源,2025,47(6):20-29.

CHENG X L, ZHANG J, LI S Y, et al. Short-term wind power prediction based on VMD-BP-BiLSTM [J]. Integrated Smart Energy, 2025, 47(6):20-29.

[12] 范斌,李豫明,郭强强,等.基于PSO-SVM-AdaBoost的超短期风电功率预测[J].内蒙古农业大学学报(自然科学版),2025,46(3):54-61.

FAN B, LI Y M, GUO Q Q, et al. Ultra-short-term wind power prediction based on PSO-SVM-AdaBoost[J]. Journal of Inner Mongolia Agricultural University (Natural Science Edition), 2025, 46(3): 54-61.

[13] 谢金财,于千城,王之赐,等.基于VMD-GRAU的短期风电功率预测[J].广西科学,2024,31(4): 773-780,787.

XIE J C, YU Q CH, WANG ZH C, et al. Short-term wind power prediction based on VMD-GRAU [J]. Guangxi Science, 2024, 31(4): 773-780,787.

[14] 任润虎,戴天乐,郭巨新.基于时空自适应图卷积循环网络的风电功率预测[J].电力系统及其自动化学报,2024,36(12):93-100.

REN R H, DAI T L, GUO J X. Wind power prediction based on spatio-temporal adaptive graph convolutional cyclic network[J]. Journal of Electric Power System and Its Automation, 2024, 36(12): 93-100.

[15] CHENG H, WEN Q S, LIU Y, et al. RobustTSF: Towards theory and design of robust time series forecasting with anomalies[J]. ArXiv preprint arXiv: 2402.02032, 2024.

[16] XUE J K, SHEN B. Dung beetle optimizer: A new meta-heuristic algorithm for global optimization[J]. The Journal of Supercomputing, 2022, 79 (7): 7305-7336.

[17] WANG S Y, WU H X, SHI X M, et al. TimeMixer: Decomposable multiscale mixing for time series forecasting [J]. ArXiv preprint arXiv: 2405.14616, 2024.

[18] 梁昌侯,龙华,李帅,等.基于MIC-VMD-GWO-LSTM的短期风电功率预测[J].现代电子技术,2023,46(22):115-120.

LIANG CH H, LONG H, LI SH, et al. Short-term wind power prediction based on MIC-VMD-GWO-LSTM [J]. Modern Electronic Technology, 2023, 46(22): 115-120.

[19] ZHAO Y H, PENG X H, TU T, et al. WOA-VMD-SCINet: Hybrid model for accurate prediction of ultra-short-term photovoltaic generation power considering seasonal variations [J]. Energy Reports, 2024, 12: 3470-3487.

[20] 宋江涛,崔双喜,刘洪广.基于二次分解NGO-VMD残差项与长短时记忆神经网络的超短期风功率预测[J].科学技术与工程,2023,23(6):2428-2437.

SONG J T, CUI SH X, LIU H G. Ultra-short-term wind power prediction based on quadratic decomposition NGO-VMD residual term and long short-term memory neural network [J]. Science, Technology and Engineering, 2023, 23 (6): 2428-2437.

[21] CHITHRA S, ARUNACHALAPERUMAL C, RAJAGOPAL R, et al. Hybrid optimization for

energy management in smart grids using golden jackal algorithm and deep convolutional neural networks[J]. *Electrical Engineering*, 2025, 107: 8201-8215.

[22] 王进峰,吴盛威,花广如,等.基于Bi-LSTM和改进残差学习的风电功率超短期预测方法[J].华北电力大学学报(自然科学版),2025,52(1):56-65.
WANG J F, WU SH W, HUA G R, et al. Ultra-short-term prediction method of wind power based on Bi-LSTM and improved residual learning[J]. *Journal of North China Electric Power University (Natural Science Edition)*, 2025, 52(1): 56-65.

[23] LI Q, REN X Y, ZHANG F, et al. A novel ultra-short-term wind power forecasting method based on TCN and Informer models [J]. *Computers and Electrical Engineering*, 2024, 120: 109632.

[24] 张亚丽,王聪,张宏立,等.基于非平稳Transformer的超短期风电功率多步预测[J].智慧电力,2024,52(1):108-115.
ZHANG Y L, WANG C, ZHANG H L, et al. Multi-step prediction of ultra-short-term wind power based on non-stationary transformer [J]. *Smart Power*, 2024, 52(1): 108-115.

[25] 虞伟,黄浩,金晨星,等.基于多头自注意力特征变换和CNN-LSTM的超短期风电功率预测[J/OL].南方电网技术,2025,1-10 [2025-05-20]. <https://link.cnki.net/urlid/44.1643.tk.20241205.0934.002>.

YU W, HUANG H, JIN CH X, et al. Ultra-short-term wind power prediction based on multi-head self-attention feature transformation and CNN-LSTM[J/OL]. *China Southern Power Grid Technology*, 2025, 1-10 [2025-05-20]. <https://link.cnki.net/urlid/44.1643.tk.20241205.0934.002>.

作者简介

高 鹰,硕士研究生,副教授,主要研究方向为社会网络隐私保护、深度学习算法在风力发电中的应用。

E-mail:glzly@126.com

庄庆泽(通信作者),硕士研究生,主要研究方向为机器学习、风电功率预测。

E-mail:hxlxl1025@163.com

张 飞,博士,副教授,主要研究方向为新能源发电系统优化控制、工业互联网数据融合、风电场风功率预测及优化控制。

E-mail:rollrocket@imust.edu.cn

秦 岭,博士,教授,主要研究方向为光通信技术、光定位系统,基于人工智能的定位算法、先进的调制方式,光信息处理。

E-mail:qinling1979@imust.edu.cn

邬 锡麟,硕士研究生,高级工程师,主要研究方向为自动化电力电子技术在冶金行业中的智能制造应用。

E-mail:bt_wxl@sina.com