

基于时空特征挖掘的EIF-WaveNet乡村小型分布式风电集群功率预测

陈文颖^{1,2}, 吴梦达¹

(1. 华北电力大学控制与计算机工程学院, 河北保定 071003; 2. 保定市综合能源系统状态检测与优化调控重点实验室 (华北电力大学), 河北保定 071003)

摘要: 针对乡村地区小型分布式风电面临的地形复杂、风速及功率数据噪声干扰显著等问题, 为提升其超短期功率预测精度, 提出了一种融合改进孤立森林数据清洗与时空特征挖掘的WaveNet预测算法。首先, 采用改进孤立森林算法 (enhanced isolation forest, EIF) 进行数据清洗, 该算法集成了动态分位数阈值优化、路径长度窗口平滑及多模态校正策略, 有效识别并剔除异常数据点。随后, 利用最大互信息系数 (maximal information coefficient, MIC) 筛选高相关特征输入并引入时间注意力机制结合Transformer模块捕获关键时间依赖性, 同时应用空间注意力机制与图卷积网络 (graph convolutional network, GCN) 建模风机间的空间关联性, 进行时空特征融合。接着, 构建基于WaveNet架构的超短期单机功率预测模型, 以其膨胀卷积和跳跃连接精准捕捉融合特征序列的动态模式并输出单机功率预测结果。最后, 考虑各机组的位置权重与空间相关性, 聚合单机预测结果输出集群总功率。以某村庄部署的3台单机容量为10 kW的小型分布式风电场为实例, 验证结果表明所提方法具有较高预测精度。

关键词: 风功率预测; 小型分布式风电; 改进孤立森林算法; 时空特征挖掘; WaveNet

Power Prediction for Small-Scale Distributed Wind Turbine Clusters in Rural Areas Based on EIF-WaveNet with Spatiotemporal Feature Mining

CHEN Wenying^{1,2}, WU Mengda¹

(1. School of Control and Computer Engineering, North China Electric Power University, Baoding, Hebei 071003, China;

2. Baoding Key Laboratory of State Detection and Optimal Control for Integrated Energy Systems, North China Electric Power University, Baoding, Hebei 071003, China)

Abstract: To address the challenges of complex terrain, significant noise interference in wind speed and power data for small-scale distributed wind farms in rural areas, an improved WaveNet prediction algorithm integrating enhanced isolation forest data cleaning and spatiotemporal feature mining is proposed to enhance ultra-short-term power prediction accuracy. Firstly, an enhanced isolation forest (EIF) algorithm is employed for data cleaning. This algorithm incorporates dynamic quantile threshold optimization, path length window smoothing, and a multimodal correction strategy to effectively identify and eliminate anomalous data points. Subsequently, the maximal information coefficient (MIC) is utilized to screen highly correlated input features. A temporal attention mechanism combined with a Transformer module captures critical temporal dependencies, while a spatial attention mechanism integrated with a graph convolutional network (GCN) models spatial correlations among wind turbines for spatiotemporal feature fusion. Next, an ultra-short-term single-turbine power prediction model based on WaveNet architecture is constructed, which accurately captures the dynamic patterns of the fused feature sequence through its dilated convolution and skip connections, and outputs the single-turbine power prediction results. Finally, the total cluster power output is aggregated from single-turbine prediction results by considering the positional weights and spatial correlations of each turbine. A case study is conducted on three-small-turbine distributed wind farms with a single turbine capacity of 10 kW deployed in a certain rural area. The results demonstrate that the proposed method achieves high prediction accuracy.

Key words: wind power prediction; small-scale distributed wind power; enhanced isolation forest algorithm; spatiotemporal feature mining; WaveNet

0 引言

我国农村地域广阔、风能资源丰富且分布相对分散。统筹盘活农村碎片化土地资源,因地制宜开展风电就近就地开发建设,既可持续壮大村级集体经济,赋能乡村振兴战略纵深推进;亦能加速农村能源体系绿色低碳转型进程,为我国碳达峰、碳中和战略目标落地实施提供关键支撑,具备显著的经济、生态与战略多重价值^[1]。受风能自身随机性、间歇性与波动性特质所限,风功率输出稳定性欠佳,这不仅制约了风功率的稳定输出,还给能源管控与实时调度带来了很大不便^[2],因此精准的风功率预测可在很大程度上保障供电安全。

超短期风功率预测时间尺度为15 min~4 h,主要应用于风电场的实时监控与精细化管理^[3]。但部分农村地形复杂,在传感器及测风塔采集数据过程中易混入大量噪声,进而干扰功率预测精度。文献[4]采用基于密度的DBSCAN聚类方法,结合风速-功率数据的时空关联性分离异常点。文献[5]通过分层聚类结合DTW计算序列相似度,有效识别运行状态切换导致的异常段。文献[6]利用Quartile RANSAC回归分析数据分布,结合小波阈值去噪处理高比例堆叠异常。文献[7]基于风机物理特性筛选数据后再采用孤立森林(isolation forest, IF)快速识别异常点。

获取纯净数据后,精准时序建模对提升预测精度至关重要。近些年,人工智能方法在风功率预测领域得到了广泛应用。传统的预测模型多基于神经网络结构,如长短期记忆网络(long short-term memory, LSTM)^[8-9]、门控循环单元(gated recurrent unit, GRU)^[10]和卷积神经网络(convolutional neural network, CNN)^[11-13]等。文献[14]通过CNN-LSTM耦合架构优化参数空间。文献[15]利用CNN-BiLSTM-Attention混合模型增强动态权重分配,显著降低预测误差。然而,该类方法受限于

输入表征的局部感受野约束,难以对高维时序数据内部的长距离依赖关系进行有效建模。Transformer^[16]自注意力机制通过全局上下文建模突破此瓶颈,其变体Informer^[17]以概率稀疏注意力捕捉风电序列的长期依赖,而WGT-Transformer框架^[18]则通过威布尔-高斯变换与编码器-解码器结构优化计算效率。尽管自注意力机制可捕获长期依赖关系,其二次复杂度导致显存与算力瓶颈在大规模数据场景中凸显。近期研究转向膨胀因果卷积(dilated causal convolution)的WaveNet框架,以线性复杂度实现自回归序列建模,显著降低资源消耗,同时具备对风电空间异质性及高阶关联特征的建模能力^[19-21]。

分布式能源功率波动与不确定性制约了预测精度与调控效率的协同发展。现有研究已构建起预测-调控闭环框架:文献[22]将PCA-OPTICS聚类与ShapeDTW-QWGRU模型结合,建立了代表站与集群功率的精细映射机制;文献[23]通过定义空间互补特性并量化其效应,为集群优化奠定理论基础;文献[24-25]分别通过卫星云图融合和地基云图修正技术,显著提升了复杂天气场景下的预测鲁棒性;文献[26]提出的变频空调多时间尺度控制策略成功实现了分钟级波动平抑。这些研究为风电集群协同调控提供了重要方法论支撑。然而在复杂地形场景下,农村机组数据噪声诱发的特征畸变与小尺度地形时空耦合效应建模不足,成为现有研究的瓶颈。

基于上述研究,为克服农村风电机组数据采集噪声和地形复杂等问题,提出了融合数据降噪与时空特征挖掘的WaveNet预测模型,以数据-模型双重精简策略,将时空联合降噪与轻量级预测耦合为统一框架。首先,采用孤立森林(enhanced isolation forest, EIF)算法构建了噪声过滤机制,解决了农村机组数据质量劣化问题;随后,通过最大互信息系数(maximal information coefficient, MIC)筛选高相关特征并引入时空注意力机制与Transformer-GCN进行时空特征融合以强化小尺度地形耦合建模;接着,构建了WaveNet架构预测单机功率;最后,结合空间权重聚合输出集群功率,以某小型分布式风电场验证了所提算法的有效性。

1 乡村小型分布式风力发电

乡村小型分布式风力发电系统通过地理分散布

基金项目:中央高校基本科研基金(2023JC010);河北省省级科技计划资助项目(22567643H);河北省高等学校科学技术研究项目(CXY2023001)。

Foundation item: Supported by the Fundamental Research Funds for Central Universities (2023JC010); the Science and Technology Program of Hebei Province (22567643H); the Science and Technology Research Project of Hebei Provincial Colleges and Universities (CXY2023001).

局的风机群实现本地化供电^[27]。如图 1 所示,在山丘、山谷以及平原三地分别部署 10 kW 风机, 风机 1、2、3 向中心村庄供电。

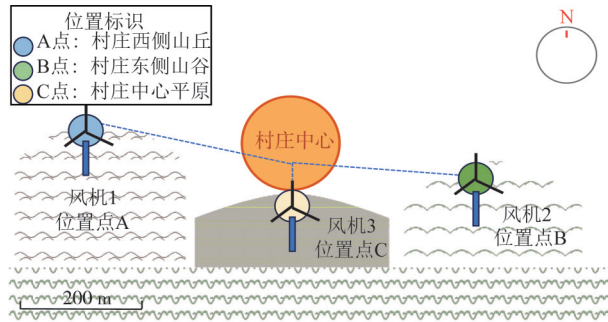


图 1 村庄小型分布式风力发电示意图

Fig. 1 Schematic diagram of small-scale distributed wind power generation in villages

虽然所属地区风资源丰富,但在复杂地形条件下,风速传感器易受局地强阵风、机械振动以及微气象突变等因素的干扰。同时,在时空异质性上,风机群跨地形分布且高差超过 200 m,造成风速剖面呈现非线性关联;山谷区午后热力环流会引发功率骤变,与平原区存在相位差。

鉴于传统集中式风电场预测方法未能充分考虑此类系统的小尺度地形耦合效应以及噪声空间关联性,本文搭建了适配复杂场景的时空联合清洗与预测框架,以提升乡村小型分布式风力发电系统的预测准确性和运行效率。

2 乡村分布式风功率预测方法

所提出的预测方法整体流程如图 2 所示,主要包括以下步骤。

1)数据清洗:采用改进孤立森林算法,通过引入动态分位数阈值优化、路径长度窗口平滑及多模态校正策略进行改进,有效识别并剔除异常样本,构建高质量训练数据集。

2)时空特征融合:采用 MIC 筛选高相关输入特征,构建面向预测任务的特征空间;进一步引入自注意力机制,在时间维度上利用 Transformer 建模长程依赖关系,在空间维度上结合 GCN 挖掘风机间的拓扑关联,实现时空特征的有效融合。

3)单机功率预测:基于 WaveNet 架构构建单机功率预测模型,通过堆叠多层膨胀卷积、门控激活单元与残差连接结构,增强模型对风电序列局部波

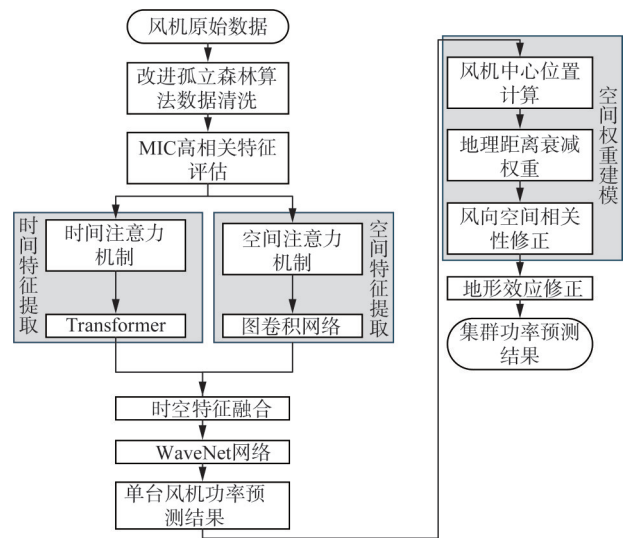


图 2 预测算法流程图

Fig. 2 Prediction algorithm flowchart

动与长期趋势的感知能力,实现对未来 15 min 各风机输出功率的高精度预测。

4)集群功率预测:综合考虑风机空间分布与地形效应,建立基于供电拓扑与地理权重的空间聚合模型,对单机预测结果进行加权融合与地形效益修正,最终输出分布式风电系统的集群功率预测值。

2.1 改进孤立森林算法

所提算法流程如图 3 所示,传统孤立森林算法的全局污染度参数无法适应不同风速区间的局部异常特征^[28],为此做了如下改进。

1)动态分位数阈值优化:输入原始风机数据,先进行算法参数初始化,接着按风速区间对数据进行分箱并对每个分箱计算动态异常阈值,有效避免全局阈值对局部异常的掩盖。

$$T_b(q) = \begin{cases} \text{Quantile}(S_b, q), & |S_b| > N_{\min} \\ \text{Quantile}(S_{\text{global}}, q), & \text{其他} \end{cases} \quad (1)$$

式中: S_b 为 b 区间的异常分数集合; S_{global} 为所有区间异常分数的联合集合; q 为分位数阈值; $N_{\min}$ 为有效样本最小数量; $T_b(q)$ 为第 b 区间的动态异常阈值;函数 $\text{Quantile}(\cdot)$ 为分位数计算函数;函数 $|\cdot|$ 为集合基数运算符。

2)路径长度窗口平滑:针对原始路径长度 $h(x)$ 对单点噪声敏感的问题,对路径长度进行窗口平滑并计算平滑后异常分数,以此抑制噪声引起的路径长度波动。

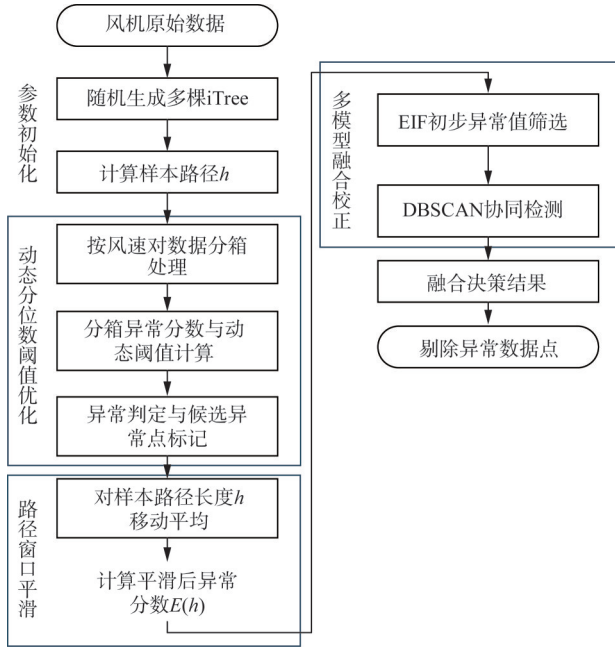


图 3 改进孤立森林算法流程

Fig. 3 Improved isolation forest algorithm procedure

$$\tilde{h}(x_i) = \frac{1}{W} \sum_{j=-K}^K h(x_{i+j}) \quad (2)$$

$$S(x_i) = 2^{-\frac{\tilde{h}(x_i)}{\bar{h}(x_i)}} \quad (3)$$

式中: $h(x_i)$ 为样本 x_i 的原始路径长度; K 为滑动窗口半径; $W = 2K + 1$ 为窗口大小; $\tilde{h}(x_i)$ 为平滑后的路径长度; $\bar{h}(x_i)$ 为所有样本路径长度的均值; $S(x_i)$ 为平滑后的异常分数。

3) 多模态融合校正: 为避免单一模型易受高维噪声干扰, 引入 DBSCAN 密度检测与决策规则^[29], 以降低误报率, 提升鲁棒性。

$$A_{\text{final}}(x_i) = \begin{cases} 1, (S_{\text{iso}}(x_i) > T_{\text{iso}}) \vee (D_{\text{dbscan}}(x_i) = 1) \\ 0, \text{其他} \end{cases} \quad (4)$$

式中: $S_{\text{iso}}(x_i)$ 为孤立森林检测到的异常分数; T_{iso} 为孤立森林的异常分数阈值; $D_{\text{dbscan}}(x_i)$ 为 DBSCAN 检测结果, 取值为 1 表示噪声点, 取值为 0 表示正常点; $A_{\text{final}}(x_i)$ 为异常点的最终被判定, 取值为 1 表示异常点, 取值为 0, 表示正常点。

2.2 时空特征融合

经 MIC 筛选的高相关性特征子集 X_{select} 分别进行时间依赖性与空间相关性建模, 具体如图 4 所示。

时间序列分析中, Transformer 通过自注意力机制 (self-attention) 捕获序列数据中的长期依赖性, 并以多头注意力 (multi-head attention) 机制进一步增

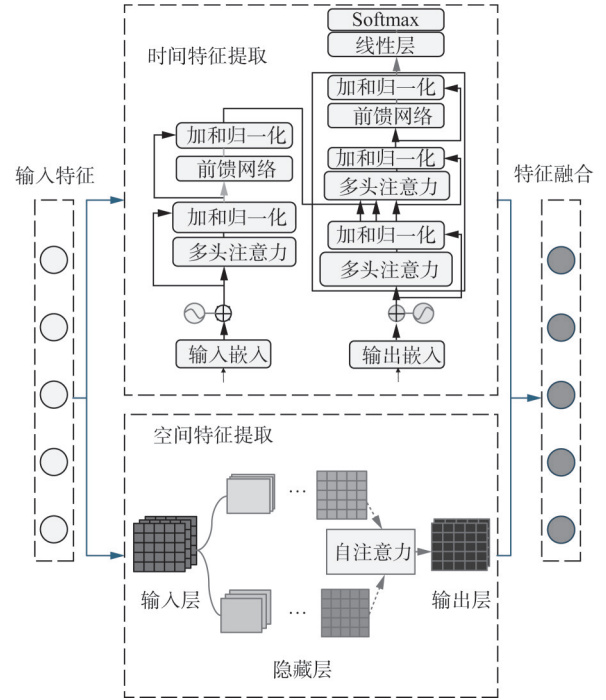


图 4 时空特征融合示意图

Fig. 4 Schematic diagram of spatiotemporal feature fusion

强对时间依赖性的捕捉能力^[30]。

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (5)$$

$$\text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Concat}(h_{\text{head}, 1}, \dots, h_{\text{head}, h})\mathbf{W}_O \quad (6)$$

$$\text{head}_i = \text{Attention}(\mathbf{Q}\mathbf{W}_{qi}, \mathbf{K}\mathbf{W}_{ki}, \mathbf{V}\mathbf{W}_{vi}) \quad (7)$$

式中: $\mathbf{Q}$ 、 $\mathbf{K}$ 、 $\mathbf{V}$ 分别为查询、键和值矩阵; d_k 为注意力头的维度; $\mathbf{W}_O$ 为输出投影矩阵; $\text{Attention}(\cdot)$ 为自注意力计算函数; $\text{softmax}(\cdot)$ 为归一化指数函数; $\text{MultiHead}(\cdot)$ 为多头注意力计算函数; $\text{Concat}(\cdot)$ 为特征拼接函数; $h_{\text{head}, i}$ 为第 i 个注意力头的输出结果; 下标 h 为多头注意力机制中注意力头的数量; $\mathbf{W}_{qi}$ 、 $\mathbf{W}_{ki}$ 、 $\mathbf{W}_{vi}$ 分别为第 i 个注意力头对应的查询、键、值线性投影权重矩阵。

同时, 利用 GCN 建模风机间的空间关联性。给定图结构 $G = (V, E)$, 其中 V 为风机节点, E 为边。GCN 层的更新规则为^[31]:

$$\mathbf{H}^{(l+1)} = \sigma(\tilde{\mathbf{D}}^{-\frac{1}{2}}\tilde{\mathbf{A}}\tilde{\mathbf{D}}^{-\frac{1}{2}}\mathbf{H}^{(l)}\mathbf{W}^{(l)}) \quad (8)$$

式中: $\mathbf{H}^{(l)}$ 为第 l 层的节点特征; $\tilde{\mathbf{A}}$ 为添加自环的邻接矩阵; $\tilde{\mathbf{D}}$ 为度矩阵; $\mathbf{W}^{(l)}$ 为可学习权重; $\sigma(\cdot)$ 为非线性激活函数。

接着, 引入空间注意力机制动态调整节点间的

权重, 最终的空间特征更新为:

$$\mathbf{h}'_i = \sigma \left(\sum_{j \in N(i)} \alpha_{ij} \mathbf{W} \mathbf{h}_j \right) \quad (9)$$

式中: $N(i)$ 为节点 i 的邻居集合; α_{ij} 为归一化后的注意力权重; $\mathbf{h}_j$ 为节点 j 的原始特征向量; $\mathbf{h}'_i$ 为节点 i 经空间注意力加权聚合后的更新特征向量。

最终, 以分配 λ 权重进行时空特征融合, 具体为:

$$\mathbf{H}_f = \lambda \mathbf{H}_t + (1 - \lambda) \mathbf{H}_s \quad (10)$$

式中: $\mathbf{H}_t$ 为时间特征; $\mathbf{H}_s$ 为空间特征; $\mathbf{H}_f$ 为经加权融合后的时空联合特征矩阵; λ 为时空特征融合权重系数, 用于调节时间特征与空间特征在最终输出中的贡献占比。

2.3 分布式集群功率预测

WaveNet 结构如图 5 所示, 融合后的特征 $\mathbf{H}_f$ 输入 WaveNet, 其能够通过膨胀卷积和门控机制高效捕获时间序列的多尺度依赖关系, 同时避免信息泄露问题^[32]。

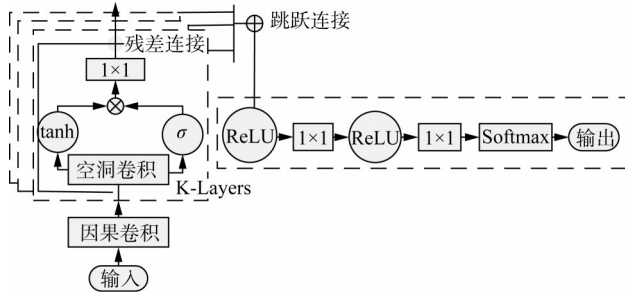


图 5 WaveNet 结构图

Fig. 5 WaveNet structure diagram

膨胀卷积: 引入膨胀率 r 来增加卷积核的跨度, 从而捕获更远时间步的信息。

$$y_t = \sum_{\tau=0}^{k-1} \mathbf{W}_\tau x_{t-\tau} \quad (11)$$

式中: y_t 为 t 时刻的膨胀卷积输出特征, t 为表征时刻的数学变量; $\mathbf{W}_\tau$ 为第 τ 个位置对应的卷积核权重矩阵; 下标 τ 为卷积核内的离散位置索引变量; k 为卷积核的尺寸大小。

残差连接与门控机制: 残差连接用于加速训练并缓解梯度消失问题, 通过门控单元控制信息流动。

$$\begin{cases} \mathbf{h}_{l+1} = \mathbf{h}_l + f(\mathbf{h}_l) \\ \mathbf{z}_l = \sigma(\mathbf{W}_z * \mathbf{h}_l + b_z) \\ \mathbf{f}_l = \tanh(\mathbf{W}_f * \mathbf{h}_l + b_f) \\ \mathbf{h}_{l+1} = \mathbf{z}_l \odot \mathbf{f}_l + (1 - \mathbf{z}_l) \odot \mathbf{h}_l \end{cases} \quad (12)$$

式中: $\mathbf{W}$ 为卷积核权重; $\odot$ 为逐元素乘法; $\mathbf{h}_l$ 、 $\mathbf{z}_l$ 、 $\mathbf{f}_l$ 分别为第 l 层的输入特征向量、更新门输出向量、候选特征向量; $*$ 为一维因果卷积运算; b_z 为更新门偏置项; b_f 为候选特征门偏置项。

跳跃连接与输出: 每一层都使用不同的膨胀率, 从而捕获不同时间尺度的特征^[33]。最终, 这些多尺度特征通过跳跃连接进行融合, 并通过全连接层生成最终输出。

$$\begin{cases} S_t = \sum_{l=1}^L \mathbf{W}_s * \mathbf{h}_l \\ O_t = \mathbf{W}_o * S_t + b_o \end{cases} \quad (13)$$

式中: S_t 为多层特征的融合结果; O_t 为最终预测输出; L 为 WaveNet 网络的卷积层数量; $\mathbf{W}_s$ 和 $\mathbf{W}_o$ 分别为跳跃连接投影权重矩阵、输出层投影权重矩阵; b_o 为输出层偏置项。

在得到 3 台风机的风功率预测结果后, 以风电场地理中心位置、地理位置衰减、风向相关性修正构建空间权重; 以草地、山谷、平原等类别映射构建地形修正函数, 最终得到集群风功率预测结果。

$$P_{\text{total}} = \sum_{i=1}^N w_i \cdot f_i(O_i, \tau_i, \Omega) \quad (14)$$

式中: P_{total} 为风电集群的总输出功率; w_i 为空间权重; $f_i(\cdot)$ 为地形修正函数; τ_i 为地形特征参数; Ω 为环境参数矩阵; N 为风电集群的风机数量。

3 算例分析

3.1 数据清洗

算例采用的数据源于我国某乡村小型分布式风电场, 包含 3 台 10 kW 风机, 采样时间间隔为 15 min, 包含测风仪处风速和风向、温度、大气压力、地形特征、风机位置等特征, 采样时间从 2024 年 1 月 1 日到 2024 年 12 月 31 日, 受到偶发采集故障影响, 存在小部分数据丢失, 每台风机有 33 665 条样本序列。

针对保存的有效数据序列, 为去除噪声干扰, 提升在复杂地形小风机数据中的鲁棒性, 设置 EIF 算法如下。

1) 动态分位数阈值: 将 0~25 m/s 风速区间以 0.5 m/s 为步长进行等分区间分箱, 逐一计算各风速分箱内异常评分的 98 % 分位数, 以此作为各区间的局部异常阈值, 有效避免全局阈值掩盖局部异常的问题。

2)路径长度窗口平滑:对单棵树模型输出的路径长度 $h(x)$ 序列采用100点滑动平均滤波,基于平滑后的路径长度计算异常分数,可显著抑制单点噪声引起的波动。

3)多模态融合校正:在孤立森林模型决策结果基础上引入DBSCAN密度复核机制,针对低、中、高3种风速区间分别设置邻域半径为0.10、0.30和0.50,固定最小核心点为20,通过核可达性约束剔除高维噪声伪异常样本。

图6展示了基于改进孤立森林算法对原始3个风机数据进行清洗的结果。

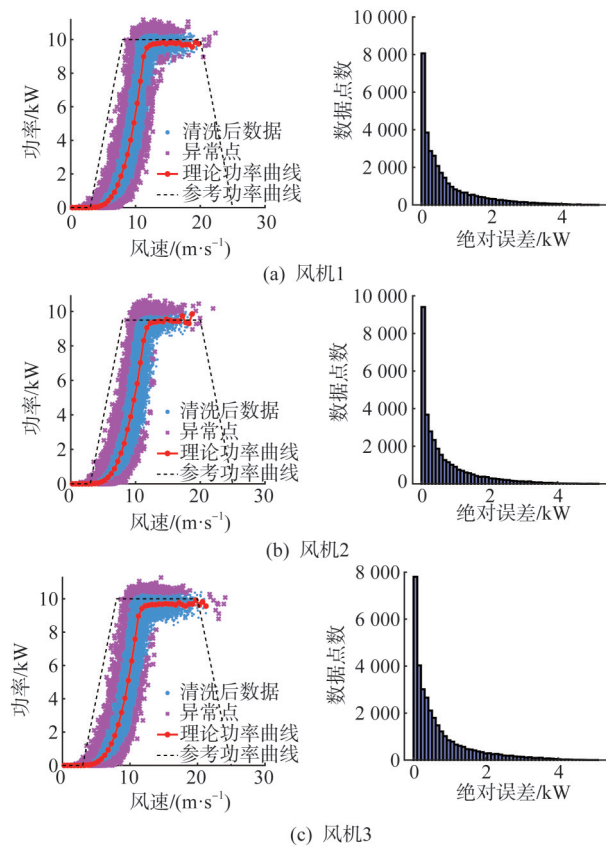


图6 EIF数据清洗结果

Fig. 6 Data cleaning results of EIF

清洗后风速-功率散点图表明,数据点分布更加集中且特征更加清晰。异常点主要集中在功率曲线的上升段、平台段及衰减段,经过清洗后,风机1、风机2和风机3的功率-风速曲线形状趋于一致,更符合风机的物理特性。

此外,绝对误差分布结果进一步表明数据的误差主要集中在较低范围,极端误差值的数量显著减少。经降噪后的数据能够更准确地反映风机的实际

运行性能,为后续风电功率预测和风机性能评估提供了更为可靠的数据基础。

3.2 实验设置与评价指标

为系统验证WaveNet模型在风电功率预测任务中的精度优越性,同时验证模型时空特征提取能力与EIF算法的有效性,本文设计两组对比实验开展分析。

1)算法对比实验:在数据清洗与时空特征挖掘一致的前提下,将WaveNet与当前主流时序模型并行训练,以预测误差为指标量化其精度增益。

2)消融实验:在固定WaveNet主干的前提下,依次构建5组对比模型:EIF-时空特征融合、时空特征融合、仅Transformer时间特征提取、仅GCN空间特征提取及基线模型,通过系统性消融实验,量化分析各功能模块对风电功率预测精度的边际贡献。

考虑到时间序列特性,所有数据严格按时间顺序划分:以2024年1月1日00:00为起点,前80%时段数据作为训练集,中间10%时段数据为验证集,最后10%时段数据为测试集。实验在配备Intel Core i9-13900HX CPU, 32 GB RAM及NVIDIA GeForce RTX 3080 GPU的工作站上完成,软件环境为Python 3.8与PyTorch框架。训练过程中,设定最大迭代轮数为10,批次大小为32。

为评估模型性能,选用误差指标平均绝对误差(mean absolute error, MAE)和均方根误差(root mean squared error, RMSE)以及拟合性能指标决定系数(R-squared, R^2)^[34]作为评价指标。其中,MAE值与RMSE值越低表明算法预测精度越高,而 R^2 值通常位于0到1之间,越接近于1,拟合效果越好。

$$E_{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (15)$$

$$E_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (16)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (17)$$

式中: E_{MAE} 、 E_{RMSE} 、 R^2 分别为平均绝对误差、均方根误差和性能指标决定系数; n 为样本数量; y_i 为真实值; $\hat{y}_i$ 为预测值。

3.3 算法对比实验

为系统评估时序建模性能,本文采用Informer、DLinear、WaveNet、Autoformer与PatchTST等代表性深度时序模型进行对比研究;对模型输入数据统一采用EIF算法开展异常清洗,并融合时空特征增强策略完成数据预处理。各模型的具体网络结构与超参数配置如表1所示。图7对比展示了各算法针对风机1、风机2、风机3及风电场集群功率的预测效果,呈现了模型在测试集首日24 h数据序列上的拟合性能。整体来看,各模型均具备较强的预测能力,其中WaveNet模型在多个场景下展现出更优的性能表现。

针对风机1的单日24 h预测结果,WaveNet模型表现出优异的短期动态追踪与响应能力,它能够精准捕捉功率在快速变化阶段的趋势,其拟合曲线与真实值高度重合。相比之下,Informer与PatchTST在部分剧烈波动点上表现出一定的滞后

性。DLinear与Autoformer虽然在趋势平滑方面表现出色,但在细节特征的还原上稍显不足。对于风机2的预测,WaveNet在周期性波动中保持了稳定的跟踪能力,尤其在剧烈波动区域对细微变化的还原度较高。其他模型在整体趋势拟合上表现良好,但在局部精度上相对较低。对于风机3,WaveNet虽在峰值处精度欠佳,但在中低值区域实现了与真实值的高度一致。在集群功率预测中,WaveNet的优势更加明显。在融合多源数据后,WaveNet仍然能够紧密跟随实际功率的变化,体现了其在复杂模式建模和多尺度特征提取方面的显著优势。WaveNet在处理多台风机数据融合后的复杂功率变化时,依然能够较好地追踪真实值的走势。

表2汇总了Informer、DLinear、WaveNet、Autoformer和PatchTST 5种模型在风机1、风机2、风机3及集群功率预测任务中的MAE值、RMSE值和 R^2 指标。整体来看,WaveNet在各类预测场景中均

表1 算法对比试验模型参数

Tab. 1 Model parameters for algorithm comparison experiments

模型	网络机制	输入/输出	嵌入维度	编码层数	解码层数	关键超参数
Informer	概率稀疏注意力	10/1	512	3	2	稀疏因子=5
DLinear	线性深度扩展	10/1	512	3	2	线形隐藏层维度=512
WaveNet	膨胀卷积残差网络	10/1	512	3	2	核大小=2,膨胀基数=2
Autoformer	自相关机制	10/1	512	3	2	移动平均窗口=25
PatchTST	时序分块嵌入	10/1	512	3	2	切片长度=16,多头注意力数=16

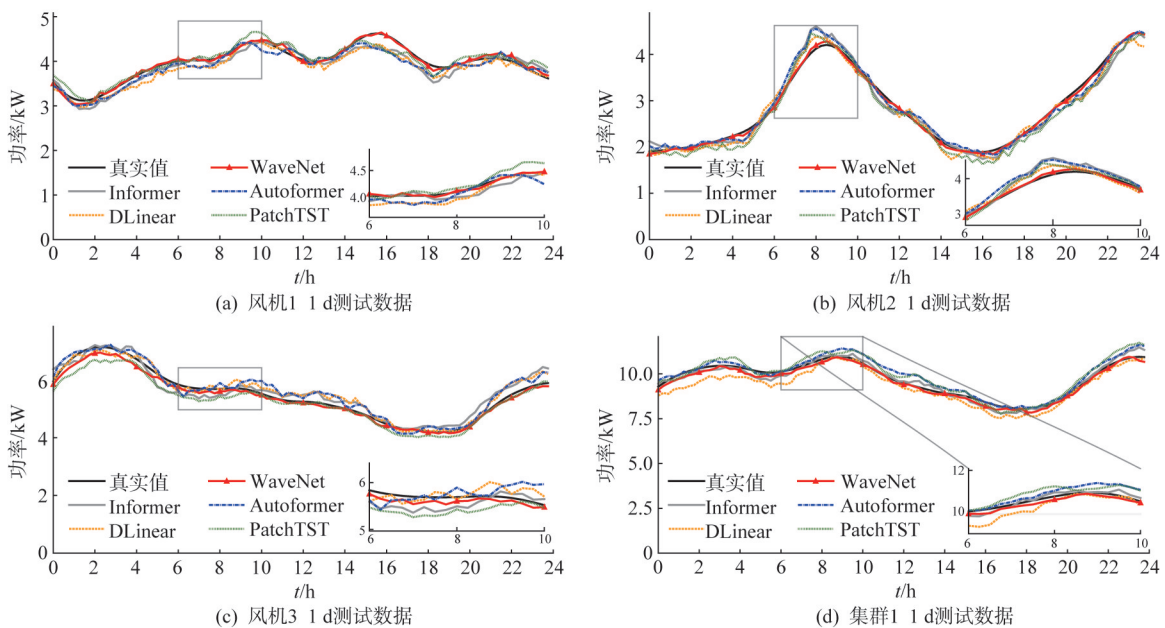


图7 算法对比拟合曲线

Fig. 7 Algorithm comparison fitting curves

表现出更优的性能, 相较其他模型具有更高的预测精度与更强的稳定性。在单台风机预测中, WaveNet 模型的预测结果 MAE 值平均降低 15.7%, RMSE 值平均下降 18.6%, R^2 平均提升 1.1%, 结果表明 WaveNet 能够更准确地刻画风速与功率之间的非线性映射关系, 在复杂动态变化下保持较低的预测偏差。在集群功率预测任务中, WaveNet 的优势更为突出, MAE 值与 RMSE 值分别平均降低 23.7% 和 23.8%, R^2 平均提高 1.4%, 体现出其在融合多源异构数据、建模多变量交互关系方面具备更强的能力, 能有效应对风场环境的不确定性与复杂性。

3.4 消融实验

为进一步验证 EIF 算法与时空特征挖掘在 WaveNet 中的边际贡献, 本文设计如表 3 所示的 5 种消融配置, 用于评估数据清洗和时空模块的有效性。

表 2 算法对比实验误差结果

Tab. 2 Error results of algorithms comparison experiments

预测对象	评价指标	WaveNet	Informer	DLinear	Autoformer	PatchTST
风机 1	MAE 值	0.131 2	0.132 3	0.136 8	0.142 4	0.140 6
	RMSE 值	0.163 4	0.170 0	0.174 2	0.184 7	0.181 7
	R^2	0.972 4	0.970 4	0.972 1	0.968 6	0.969 6
风机 2	MAE 值	0.102 7	0.107 8	0.118 0	0.112 1	0.130 1
	RMSE 值	0.120 8	0.136 0	0.151 2	0.140 7	0.163 2
	R^2	0.982 4	0.976 2	0.970 6	0.974 6	0.965 8
风机 3	MAE 值	0.114 3	0.182 1	0.155 6	0.184 5	0.142 0
	RMSE 值	0.145 9	0.235 0	0.196 6	0.229 6	0.181 3
	R^2	0.982 7	0.955 0	0.968 5	0.957 1	0.973 2
集群	MAE 值	0.153 8	0.217 9	0.202 2	0.209 6	0.180 2
	RMSE 值	0.192 0	0.275 0	0.251 3	0.260 7	0.226 1
	R^2	0.982 3	0.963 6	0.969 6	0.967 3	0.975 4

表 3 消融实验案例

Tab. 3 Ablation experiment cases

案例	EIF 算法	时间特征提取	空间特征提取
A: EIF-时空特征挖掘	√	√	√
B: 时空特征融合	×	√	√
C: 时间特征提取	×	√	×
D: 空间特征提取	×	×	√
E: 基线模型	×	×	×

图 8 直观展现了风机 1—3 及集群功率预测的消融实验结果。EIF-时空特征融合的预测曲线与真实

值拟合最优, 整体趋势和细节特征还原度高, 尤其在功率突变段展现出精准的跟踪能力。仅时间特征提取的曲线在突变段响应滞后, 难以快速跟随功率变化; 仅空间特征提取在局部波动处偏差显著, 对时间动态性捕捉不足; 基线模型的整体拟合效果最差, 与真实值偏离较大。案例 B 未采用 EIF 数据清洗, 其曲线在细节处的噪声还原度高于案例 A。根据表 4 的消融实验数据进一步表明了 EIF 算法在增强数据质量方面的效能, 案例 A 相较于其他案例, 在各风机预测中平均 MAE 值降低 27.8%, RMSE 值降低 29.7%, R^2 提升 2.5%; 集群预测中平均 MAE 值降低 30.4%, RMSE 值降低 31.1%, R^2 提升 2.4%。这一显著优势有力地证明了 EIF-时空特征融合不仅能有效提升预测精度, 更能精准捕捉风电功率的复杂波动模式。值得注意的是, 案例 B 在未采用 EIF 改进孤立森林算法进行数据清洗的情况下, 其性能表现逊于案例 A, 表明 EIF 算法在提升数据质量方面具备一定效果。仅时间特征提取虽能追踪长期趋势, 但却忽略了风机间的空间耦合效应, 导致预测精度受限; 而仅空间特征提取在处理瞬时功率波动时误差较大, 难以准确还原功率变化的细微特征。此外, 对比基线模型, 仅时间特征提取和仅空间特征提取虽有一定改进, 但提升效果远不及 EIF-时空特征融合显著。

综上所述, 本文通过两组关键实验对 WaveNet 模型在风电功率预测任务中的性能进行了系统评估。在算法对比实验中, 所有模型在相同的数据预处理与时空特征提取条件下并行训练, 结果表明 WaveNet 在预测精度上显著优于当前主流时序模型, 验证了其在建模风电功率复杂动态特性方面的优势。进一步开展的消融实验固定 WaveNet 主干结构, 构建了 5 类对比模型: EIF-时空特征融合、时空特征融合、仅时间特征提取、仅空间特征提取及基线模型, 以量化各模块对预测性能的边际贡献。实验结果显示, 引入 EIF 机制的时空融合模块显著提升了模型对功率波动, 尤其是在突变区间的捕捉能力, 其预测曲线与真实值高度贴合, 充分体现了该模块在挖掘风电数据深层时空相关性中的关键作用。上述结果共同验证了 WaveNet 模型及其 EIF-时空特征融合结构在风电功率预测中的有效性与优越性。

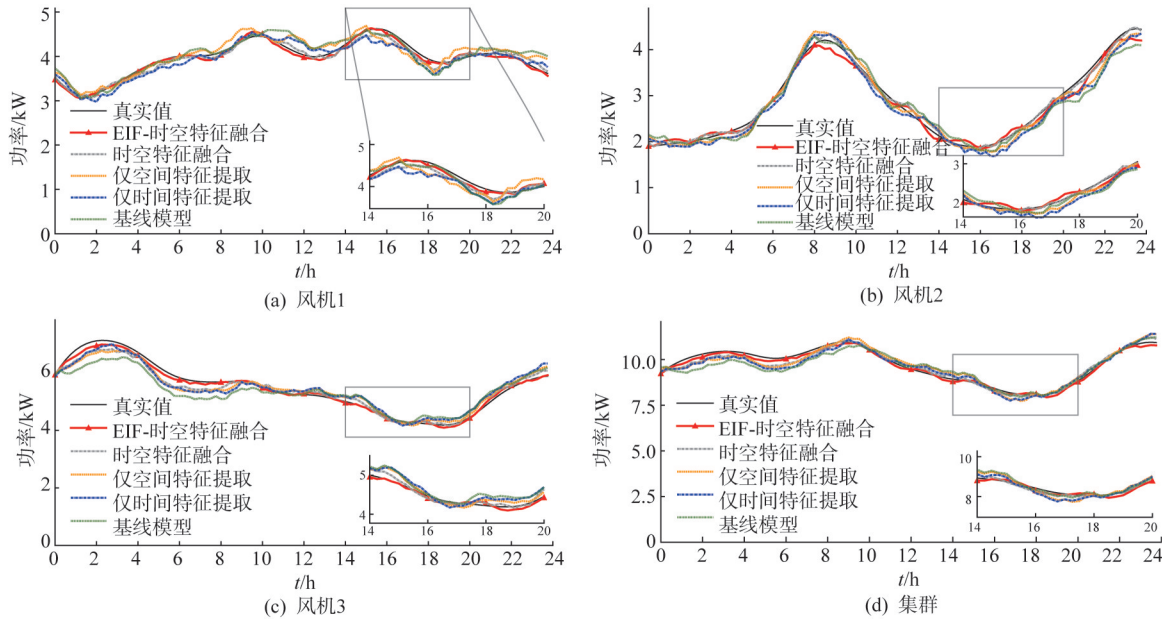


图8 消融实验拟合曲线

Fig. 8 Fitting curves of the ablation study

表4 消融实验误差结果

Tab. 4 Prediction errors of ablation experiment

预测对象	评价指标	案例A	案例B	案例C	案例D	案例E
风机1	MAE 值	0.131 2	0.159 3	0.155 1	0.142 8	0.208 8
	RMSE 值	0.163 4	0.194 6	0.198 1	0.183 0	0.264 1
	R^2	0.972 4	0.969 2	0.963 9	0.969 2	0.935 9
风机2	MAE 值	0.102 7	0.138 8	0.183 1	0.140 0	0.189 8
	RMSE 值	0.120 8	0.178 5	0.223 1	0.179 7	0.238 1
	R^2	0.982 4	0.970 7	0.936 2	0.958 6	0.927 4
风机3	MAE 值	0.114 3	0.117 1	0.158 7	0.174 0	0.220 1
	RMSE 值	0.145 9	0.149 7	0.203 9	0.223 3	0.279 3
	R^2	0.982 7	0.979 1	0.966 1	0.959 4	0.936 5
集群	MAE 值	0.153 8	0.183 4	0.203 0	0.235 0	0.288 6
	RMSE 值	0.192 0	0.234 1	0.256 0	0.294 9	0.358 3
	R^2	0.982 3	0.973 6	0.968 4	0.958 1	0.938 2

4 结论

针对乡村地区分布式风电面临的地形复杂、数据噪声干扰显著等问题,本文提出了一种基于改进孤立森林数据清洗与时空特征挖掘的WaveNet预测算法,以某乡村分布式风电实例得到以下结论。

1)改进孤立森林算法通过引入动态分位数阈值机制增强局部异常识别能力、结合路径长度窗口平滑策略抑制单点噪声引起的路径长度波动、融合DBSCAN密度聚类结果进行多模式决策校正,显著

降低了误报率,提高了算法在高维噪声环境下的稳定性与判别精度。2)通过MIC方法筛选高相关特征,结合Transformer建模时间长期依赖与GCN挖掘风机空间拓扑关联实现时空特征的深度协同融合,显著提升了模型在小尺度地形耦合情景的预测精度。3)基于WaveNet架构的单机功率预测模型通过多层膨胀卷积、门控激活与残差结构的协同设计,有效捕捉风电时间序列的局部波动与长期趋势,实现单台风机超短期功率的高精度预测;进一步结合供电拓扑与地理权重构建空间聚合模型并对单机预测结果进行地形效应修正与加权融合,提升了集群功率预测的准确性与空间一致性,实现了在单机与集群尺度上的优异预测性能。

参考文献

- [1] 国家发展改革委,国家能源局.国家发展改革委 国家能源局农业农村部关于组织开展“千乡万村驭风行动”的通知[EB/OL]. (2024-03-25).https://zfxgk.nea.gov.cn/2024-03/25/c_1310769840.html.
- [2] 杨茂,闫琦,苏欣,等.季节分型下一种面向风电功率日前预测的深度自适应滤波框架[J].南方电网技术,2023,17(6): 62-71.
YANG Mao, YAN Qi, SU Xin, et al. A depth adaptive filtering framework for wind power day-ahead prediction under seasonal classification [J]. Southern Power System Technology, 2023, 17(6): 62-71.
- [3] 康田雨,覃智君.基于超参数优化和双重注意力机制的超短期风电功率预测[J].南方电网技术,2022,16(5): 44-53.

- KANG Tianyu, QIN Zhijun. An ultra-short-term wind power forecasting method based on hyperparameter optimization and dual-stage attention mechanism [J]. Southern Power System Technology, 2022, 16(5): 44 – 53.
- [4] WANG Y, ZHANG Y, LIU H, et al. Wind turbine abnormal data cleaning method considering multi-scene parameter adaptation [C]//2022 Asian Conference on Frontiers of Power and Energy (ACFPE), October 21 – 23, 2022, Chengdu, China. New York: IEEE, 2022: 292 – 297.
- [5] WU Y, LIU J, REN Z, et al. Cleaning and reconstruction of wind power anomaly data based on hierarchical clustering [C]//2023 3rd New Energy and Energy Storage System Control Summit Forum (NEESSC), September 26 – 28, 2023, Mianyang, China. New York: IEEE, 2023: 327 – 332.
- [6] ZHANG F, ZHANG X, XU Z, et al. Cleaning of abnormal wind speed power data based on quartile RANSAC regression [J]. Energies, 2024, 17(22): 5697.1 – 5697.14.
- [7] LAI G, LIAO L, ZHANG L, et al. Wind speed power data cleaning method for wind turbines based on fan characteristics and isolated forests [J]. Journal of Physics: Conference Series, 2023, 2427(1): 012001.1 – 012001.10.
- [8] PAN Guangxu, ZHANG Haijing, JUN Wenjie, et al. A prediction method for ultra short-term wind power prediction basing on long short-term memory network and extreme learning machine [C]//2020 Chinese Automation Congress (CAC), November 6 – 8, 2020, Shanghai, China. New York: IEEE, 2020: 7608 – 7612.
- [9] 匡洪海, 郭茜. 基于多特征提取-卷积神经网络-长短期记忆网络的短期风电功率预测方法 [J]. 发电技术, 2025, 46(1): 93 – 102.
- KUANG Honghai, GUO Qian. Short-term wind power prediction method based on multimodal feature extraction-convolutional neural network-long-short term memory network [J]. Power Generation Technology, 2025, 46(1): 93 – 102.
- [10] CHEN Y. Research on application of neural networks to wind power prediction [J]. Journal of Physics: Conference Series, 2020, 1533(2): 022005.1 – 022005.11.
- [11] 任东方, 马家庆, 何志琴, 等. 基于 AVMD-CNN-GRU-Attention 的超短期风功率预测研究 [J]. 太阳能学报, 2024, 45(6): 436 – 443.
- REN Dongfang, MA Jiaqing, HE Zhiqin, et al. Research on ultra-short-term wind power forecast based on AVMD-CNN-GRU-Attention [J]. Acta Energaie Solaris Sinica, 2024, 45(6): 436 – 443.
- [12] 张程, 林锦平. 基于双重注意力卷积双向长短期记忆神经网络的电力系统状态估计 [J]. 南方电网技术, 2026, 20(5): 59 – 70.
- ZHANG Cheng, LIN Jinping. Power system state estimation based on dual attention convolutional bidirectional long and short-term memory neural networks [J]. Southern Power System Technology, 2026, 20(5): 59 – 70.
- [13] 邵宜祥, 刘剑, 胡丽萍, 等. 一种改进组合神经网络的超短期风速预测方法研究 [J]. 发电技术, 2024, 45(2): 323 – 330.
- SHAO Yixiang, LIU Jian, HU Liping, et al. Research on an ultra-short-term wind speed prediction method based on improved combined neural networks [J]. Power Generation Technology, 2024, 45(2): 323 – 330.
- [14] AHILAN D T, GANITHA S, YARRAPRAGADA K R. Wind turbine power prediction via deep neural network using hybrid approach [J]. Proceedings of the Institution of Mechanical Engineers Part A Journal of Power and Energy, 2022, 237(3): 484 – 494.
- [15] BASHIR T, WANG H, TAHIR M, et al. Wind and solar power forecasting based on hybrid CNN-ABiLSTM, CNN-transformer-MLP Models [J]. Renewable Energy, 2025(239): 122055.1 – 122055.10.
- [16] FOURNIER Q, CARON G M, ALOISE D. A practical survey on faster and lighter transformers [J]. ACM Computing Surveys, 2023, 55(14): 1 – 40.
- [17] WEI H, WANG W, KAO X. A novel approach to ultra-short-term wind power prediction based on feature engineering and In-former [J]. Energy Reports, 2023(9): 1236 – 1250.
- [18] SHI Z, LI J, JIANG Z, et al. WGformer: a Weibull-Gaussian informer based model for wind speed prediction [J]. Engineering Applications of Artificial Intelligence, 2024(131): 107891.1 – 107891.20.
- [19] 李青, 张新燕, 摆志俊, 等. 基于 MQ-WaveNet 的风电集群发电功率多步概率预测 [J]. 电力系统自动化, 2023, 47(8): 156 – 168.
- LI Qing, ZHANG Xinyan, BAI Zhijun, et al. Multi-step probability prediction of power generation for wind power clusters based on multi-horizon Quantile-WaveNet [J]. Automation of Electric Power Systems, 2023, 47(8): 156 – 168.
- [20] ZHANG Y, ZHANG Z, HU Y, et al. GMD-Linear: integrating graph neural networks and wavenet for enhanced wind power forecasting [C]//2024 12th International Conference on Information Systems and Computing Technology (ISCTech), November 8 – 11, 2024, Xi'an, China. New York: IEEE, 2024: 1 – 5.
- [21] LIU X, ZOU L, ZHANG L, et al. A multi-channel spatiotemporal SegNet model for short term wind power prediction with s-equence decomposition and reconstruction [J]. Results in Engineering, 2025(27): 105797.1 – 105797.17.
- [22] 欧阳静, 秦龙, 王坚锋, 等. 基于 PCA-ShapeDTW-QWGRU 的分布式光伏集群短期功率预测 [J]. 太阳能学报, 2024, 45(5): 458 – 467.
- OUYANG Jing, QIN Long, WANG Jianfeng, et al. Short-term power prediction for distributed PV clusters based on PCA-SHapeDTW-QWGRU [J]. Acta Energaie Solaris Sinica, 2024, 45(5): 458 – 467.
- [23] 阮呈隆, 李康平, 李正辉, 等. 分布式光伏集群功率短期预测的空间互补特性初探 [J]. 电力系统自动化, 2024, 48(3): 42 – 50.
- RUAN Chenglong, LI Kangping, LI Zhenghui, et al. Preliminary study on spatial complementarity characteristics of short-term power prediction for distributed photovoltaic clusters [J]. Automation of Electric Power Systems, 2024, 48(3): 42 – 50.
- [24] 王于波, 郝玲, 徐飞, 等. 分布式光伏集群发电功率波动模式识别与超短期概率预测 [J]. 上海交通大学学报, 2024, 58(9): 1334 – 1343.
- WANG Yubo, HAO Ling, XU Fei, et al. Pattern recognition and ultra-short-term probabilistic forecasting of power fluctuating in aggregated distributed photovoltaics clusters [J]. Journal of Shanghai Jiaotong University, 2024, 58(9): 1334 – 1343.
- [25] 张翔颖. 海量分布式光伏集群动态划分与短期功率预测技术研究 [D]. 南京: 东南大学, 2023.
- [26] 杨炜晨, 苗世洪, 刘志伟, 等. 面向分布式电源功率波动平抑

- 的变频空调集群多时间尺度模型预测控制策略[J]. 电工技术学报, 2022, 37(19): 4848 – 4861.
- YANG Weichen, MIAO Shihong, LIU Zhiwei, et al. Multi-time-scale model predictive control of inverter air conditioner cluster for distributed power fluctuation stabilization [J]. Transactions of China Electrotechnical Society, 2022, 37(19): 4848 – 4861.
- [27] 李丽旻. 乡村风电开发正当时[N]. 中国能源报, 2024 – 06 – 17(014).
- [28] WANG W, YANG S, YANG Y. An improved data-efficiency algorithm based on combining isolation forest and mean shift for anomaly data filtering in wind power curve [J]. Energies, 2022, 15(13): 4918.1 – 4918.12.
- [29] ZHANG X, SHEN X, OUYANG T. Extension of DBSCAN in online clustering: an approach based on three-layer granular models [J]. Applied Sciences, 2022, 12(19): 9402.1 – 9402.22.
- [30] ZHANG Z, SUN Z, GUO X, et al. Short-term multi-step wind power prediction model based on Pt-Transformer neural network integrating spatio-temporal feature and sparse attention [J]. Electric Power Systems Research, 2025(248): 111970.1 – 111970.9.
- [31] WU M, FENG W, LI X, et al. Short-Term power load forecasting using an improved model integrating GCN and transformer [J]. Applied Sciences, 2025, 15(13): 7003.1 – 7003.34.
- [32] WANG H, PENG C, LIAO B, et al. Wind power forecasting based on wavenet and multitask learning [J]. Sustainability, 2023, 15(14): 10816.1 – 10816.22.
- [33] 王家坤, 司化涛, 王希转, 等. 基于MQ-WaveNet的智慧新能源大规模风力发电智能控制[J]. 可再生能源, 2024, 42(10): 1341 – 1347.
- WANG Jiakun, SI Huatao, WANG Xizhuan, et al. The intelligent control of smart new energy large scale wind power generation based on MQ-WaveNet [J]. Renewable Energy Resources, 2024, 42(10): 1341 – 1347.
- [34] 陈海鹏, 李赫, 阚天洋, 等. 考虑风电时序特性的深度小波-时序卷积网络超短期风功率预测[J]. 电网技术, 2023, 47(4): 1653 – 1665.
- CHEN Haipeng, LI He, KAN Tianyang, et al. DWT-DTCNA ultra-short-term wind power prediction considering wind power timing characteristics [J]. Power System Technology, 2023, 47(4): 1653 – 1665.
-
- 收稿日期: 2025-07-07; 网络首发日期: 2025-10-24
- 作者简介:
- 陈文颖(1980), 女, 通信作者, 副教授, 博士, 研究方向为风力发电系统优化与控制, 51651854@ncepu.edu.cn;
- 吴梦达(2001), 男, 硕士研究生, 研究方向为风电功率预测技术, wmd201247@163.com。

广 告

电网仿真重点实验室 封二
宁国市裕华电器有限公司 封三
南网知识星球 A1

积成电子股份有限公司 A2
南方电网科技创新中心 A3
北京四方继保自动化股份有限公司 A4