

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

虞伟¹, 黄浩¹, 金晨星¹, 陈菁伟², 周玲³, 王逢浩³, 何强⁴

(1. 国网舟山供电公司电力调度控制中心, 浙江 舟山 316000; 2. 国网浙江省电力有限公司, 杭州 310007; 3. 国能日新科技股份有限公司, 北京 100096; 4. 新能源电力系统国家重点实验室(华北电力大学), 北京 102206)

摘要: 风力发电受环境影响显著, 具有高度随机性和波动性使得准确预测发电量变得困难。为此, 提出了结合多头自注意力机制、卷积神经网络和长短时记忆网络的超短期风电功率预测模型(attention mechanism-convolutional neural network-long short-term memory, AM-CNN-LSTM)。其中, 多头自注意力机制根据历史风电功率和气象数据的不同重要性分配权重, 生成关键特征的表达, 削弱干扰因素。卷积神经网络提取风电功率与气象因素的复杂空间模式和局部依赖关系。长短时记忆网络捕捉时间序列的长期依赖性和动态变化提升预测精度。结果表明, 最优时间窗口 $T=4$ h 下, 模型的平均绝对误差、均方根误差和平均绝对百分比误差分别为 13.45%、19.48% 和 3.78%, 优于对比方法 LSTM、CNN 和 CNN-LSTM; 比较不同模型最优时间窗口的未来 8 个样本点的预测误差, 发现所提模型在超短期预测准确性和稳定性方面具有显著优势; 变量分析表明, 组合使用多个气象变量显著优于单一变量, 揭示了气象变量间复杂的交互和非线性关系。

关键词: 风电功率预测; 多头自注意力机制; 卷积神经网络; 长短时记忆网络; 参数优化器; 时间窗口

Ultra-Short-Term Wind Power Forecasting Based on Multi-Head Self-Attention Feature Transformation and CNN-LSTM

YU Wei¹, HUANG Hao¹, JIN Chenxing¹, CHEN Jingwei², ZHOU Ling³, WANG Fenghao³, HE Qiang⁴

(1. Power Dispatching & Control Center of State Grid Zhoushan Power Co., Ltd., Zhoushan, Zhejiang 316000, China; 2. State Grid Zhejiang Electric Power Co., Ltd., Hangzhou 310007, China; 3. State Power Rixin Tech. Co., Ltd., Beijing 100096, China; 4. State Key Laboratory of Alternate Electrical Power System With Renewable Energy Sources, North China Electric Power University, Beijing 102206, China)

Abstract: Wind power generation is significantly influenced by environmental factors, exhibiting high randomness and volatility, making it difficult to accurately predict the amount of electricity generated. To address this, an ultra-short-term wind power forecasting model combining multi-head self-attention, convolutional neural network, and long short-term memory (AM-CNN-LSTM) is proposed. The multi-head self-attention mechanism assigns weights based on the varying importance of historical wind power and meteorological data, generating key feature representations and mitigating interference. The convolutional neural network extracts complex spatial patterns and local dependencies between wind power and meteorological factors. The long short-term memory network captures long-term dependencies and dynamic changes in time series, enhancing forecasting accuracy. The results demonstrate that under the optimal time window of $T=4$ h, the model achieves mean absolute error, root mean square error, and mean absolute percentage error of 13.45%, 19.48% and 3.78%, respectively. It outperforms comparative methods such as LSTM, CNN and CNN-LSTM. A comparison of forecasting errors for the optimal time window of different models over the next eight sample points reveals that the proposed model offers significant advantages in ultra-short-term forecasting accuracy and stability. Variable analysis indicates that the combined use of multiple meteorological variables significantly outperforms single-variable, uncovering the complex interactions and nonlinear relationships among meteorological factors.

Key words: wind power forecasting; multi-head self-attention mechanism; convolutional neural network; long short-term memory network; parameter optimizer; time window

0 引言

随着能源消耗增长和温室效应加剧,各国积极推动能源转型,可再生风电成为电力系统的重要组成部分。然而,风电的间歇性和不稳定性对电网运行带来挑战,开发准确可靠的预测模型对提高风能利用效率和优化电力调度至关重要^[1]。

风电预测分为超短期、短期和中长期预测,其中超短期和短期预测直接影响能源市场的运作和定价策略^[2-3]。目前常用方法包括物理模型、统计模型和机器学习模型。物理模型依赖气象数据,难以处理大规模数据,预测精度受限^[4-5];统计模型(如自回归模型^[6])参数简单、可解释性强,但难以捕捉非线性关系^[7]。机器学习模型,尤其是深度学习模型,如卷积神经网络(convolutional neural network, CNN)^[8]、门控循环单元(gated recurrent unit, GRU)^[9]和长短期记忆网络(long short term memory, LSTM)^[10],具有强大的非线性和非规则特征学习能力,能够克服统计模型的局限性,实现更精准的风电预测^[8-10]。

近年来,随着研究的深入,学者们发现混合预测模型相比单一预测模型具有更高的预测精度。例如,文献[11-12]研究表明,CNN与GRU结合的混合风电功率预测模型在预测精度上显著优于支持向量机、LSTM、CNN和GRU等单一模型。文献[13]研究表明,基于CNN和双向LSTM的短期风电功率预测模型在预测效果上优于单独使用CNN或双向LSTM的模型。

2017年,自注意力机制(self-attention)首次提出通过动态分配特征权重能够有效聚焦于重要信息并过滤无关噪声,增强模型在复杂特征表示中的学习能力^[14]。鉴于此,一些研究将其应用于风电功率预测,以期实现更精确的预测。例如,文献[15]提出基于自注意力机制的CNN-LSTM模型,通过为LSTM输出赋予差异化权重增强特征表示,提高预测性能。文献[16]结合CNN、双向LSTM和注意力机制,有效提取高层时序特征,显

著提升预测精度。文献[17]设计的双向GRU框架结合注意力机制优化,进一步增强模型性能。文献[18]利用注意力机制提取CNN隐藏特征,并结合双向LSTM捕捉时序信息,实现精准预测。以上研究充分证明了自注意力机制在风电功率预测中的潜力。

受上述研究启发,本文融合多头自注意力机制(multi-head self-attention mechanism, MHAM)、CNN和LSTM,提出了一种新型混合风电功率预测模型。不同于文献[15-18]中在后期特征提取或时序建模阶段引入单头自注意力机制的做法,本文创新性地将MHAM应用于原始输入特征的加权处理,从输入阶段精准提取关键信息并抑制噪声,显著提升了CNN和LSTM在后续建模中的表现。此外,多头自注意力机制通过多个注意力头的协同作用捕捉更丰富的特征信息,比单头机制表达能力更强。

本文的主要贡献包括:1)提出基于MHAM特征变换和CNN-LSTM的混合预测模型,实现高效稳定的风电功率预测;2)使用鲸鱼优化算法(whale optimization algorithm, WOA)优化模型参数,避免传统优化方法的主观性,确保性能最优;3)在不同时间窗口($T=2\text{ h}$ 、 4 h 、 6 h 、 8 h)下进行实验,验证了模型的精度与鲁棒性,并识别最优预测窗口;4)探讨了变量及其组合对预测精度的影响,识别关键气象变量及其交互作用,提升了模型的可解释性和实际应用价值。

1 风电功率预测模型

为提高风电功率预测精度,本文在CNN-LSTM深度学习模型中引入MHAM,提出了一种新的超短期风电功率预测模型(ultra-short-term wind power forecasting model combining multi-head self-attention, convolutional neural network, and long short-term memory, AM-CNN-LSTM)。

1.1 多头自注意力机制MHAM

在风电功率预测中,不同输入变量对预测结果的贡献程度各异,若对所有变量一视同仁,可能降低预测精度并增加模型训练时间。MHAM通过多个并行注意力头动态分配特征权重,筛选重要特征^[14]。本文在变量输入阶段引入MHAM,构造新的特征表示,突出关键特征并抑制无关干扰,为后

基金项目:国家自然科学基金资助项目(52007174);国网浙江省电力有限公司科技项目(5211ZS220001)。

Foundation item: Supported by the National Natural Science Foundation of China (52007174); the Science and Technology Project of State Grid Zhejiang Electric Power Co., Ltd. (5211ZS220001).

续特征提取提供更精准可靠的输入。MHAM的计算过程如图1所示。

MHAM的输入数据 X 为一个形状为 $n \times d$ 的矩阵,其中 n 为输入序列的长度, d 为特征维度,包含历史风电功率及气象数据。

步骤1:线性变换。输入矩阵 X 通过可学习权重 w_h^q, w_h^k, w_h^v 转换为查询 q_h 、键 k_h 、值 v_h (h 为注意力头的索引)。

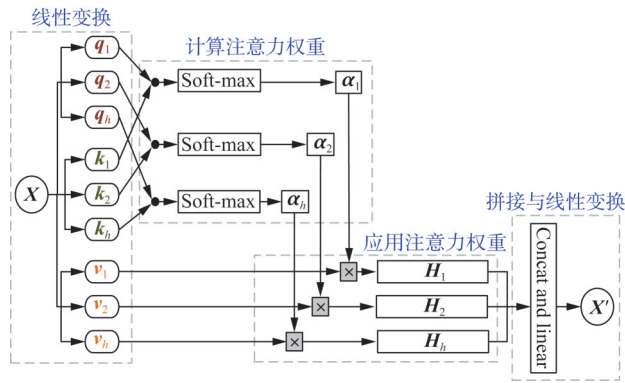


图1 多头注意力机制模块

Fig. 1 multi-head attention mechanism module

步骤2:注意力权重计算。通过 q_h 和 k_h 的点积衡量特征之间的相似性,然后,通过softmax函数对注意力分数进行归一化(式1)。如果两个特征之间的相似性很高,在计算输出时应该给予较高的权重。

$$\alpha_h = \text{softmax} \left(\frac{q_h^T \cdot k_h}{\sqrt{d_k}} \right), h = 1, 2, \dots, H \quad (1)$$

式中: α_h 为第 h 个头的注意力权重向量; d_k 为键向量 k_h 的维度,用于控制点积结果的大小,避免梯度问题; H 为注意力头的数量。

步骤3:注意力权重的应用。将注意力权重向

量 α_h 应用于值向量 v_h ,得到每个头的输出,如式(2)所示。

$$H_h = \alpha_h v_h, h = 1, 2, \dots, H \quad (2)$$

式中 H_h 为第 h 个头的输出向量。

步骤4:拼接与最终线性变换。拼接所有头的输出并通过线性变换获得最终特征表示,如式(3)所示。

$$X' = \text{Concat}(H_1, H_2, \dots, H_h) W^o \quad (3)$$

式中: X' 为多头注意力机制的最终输出矩阵; W^o 为可学习的输出权重矩阵; $\text{Concat}(\cdot)$ 为拼接操作,将各注意力头的输出向量沿特征维度拼接成一个矩阵。

1.2 卷积神经网络 CNN

风电功率预测受风速、风向、气温、湿度等环境因素及历史风电数据影响,这些因素间存在复杂的非线性关系。CNN凭借强大的特征提取和非线性映射能力,能有效捕捉这些关系,提高预测的准确性和可靠性^[8]。CNN结构如图2所示,包括输入层、卷积层、激活层、池化层、全连接层和输出层。

输入层接收多头注意力机制的最终输出矩阵 X' 。卷积层使用卷积核按一定步长提取局部特征,通过滑动窗口操作提取不同时间尺度的特征。第 k 个卷积核的特征提取过程如式(4)所示。

$$\hat{x}_k = \sigma \left(\sum_{k=1}^K w_k x_k + b_k \right) \quad (4)$$

式中: $\hat{x}_k$ 为第 k 个卷积核的输出; x_k 为第 k 个卷积核的输入; σ 为激活函数; K 为卷积核的大小; w_k 为权重矩阵; b_k 为偏置项。

激活层采用ReLU函数,其非线性特性既有助于学习复杂特征,又避免梯度消失,提升训练稳定性。池化层通过最大池化或平均池化降维,加速计算并防止过拟合。全连接层进一步映射池化层的特

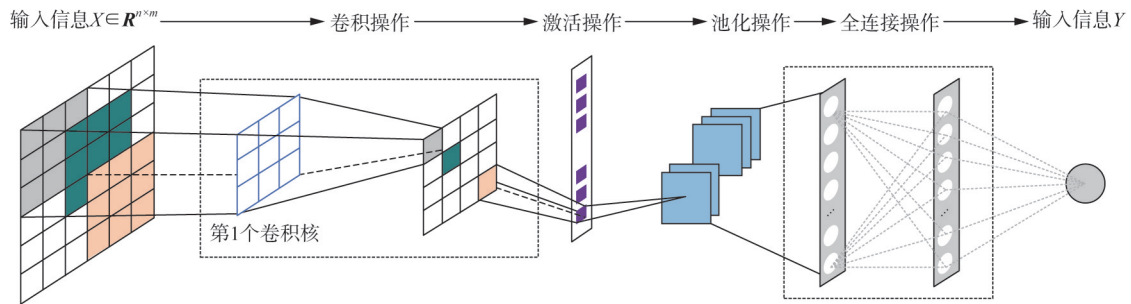


图2 卷积神经网络

Fig. 2 Convolutional neural networks

征, 最终输出预测值。

1.3 长短期记忆网络 LSTM

风电功率具有明显的时序特性。LSTM 通过其独特的门控机制(遗忘门、输入门和输出门), 能够有效捕捉长期依赖关系。相比传统 RNN, LSTM 通过门控单元解决了长序列中的梯度消失问题, 更适用于风电功率预测任务^[10]。因此, 本文在 CNN 之后引入 LSTM, 以捕捉风电功率与气象因素的时序关系。LSTM 结构如图 3 所示。

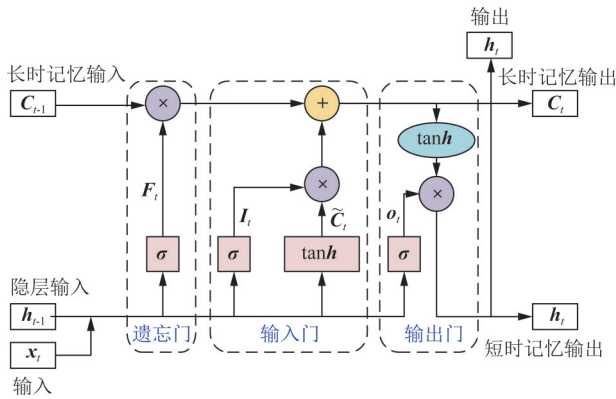


图 3 长短期记忆网络结构

Fig. 3 Long short-term memory network structure

具体来说, LSTM 门控单元包括遗忘门 F_t 、输入门 I_t 和输出门 O_t 。遗忘门 F_t 决定了从单元格状态中丢弃或保留了哪些信息; 输入门 I_t 调节新信息的添加; 而输出门 O_t 决定了单元状态信息 C_t 中的哪些部分应该被传递到隐藏状态 h_t 。单元状态 C_t 作为网络的“内存”, 保留历史信息。

遗忘门 F_t 的计算如式(5)所示。

$$F_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f) \quad (5)$$

输入门 I_t 的计算如式(6)所示。

$$I_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i) \quad (6)$$

更新候选值 $\tilde{C}_t$ 的计算如式(7)所示。

$$\tilde{C}_t = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c) \quad (7)$$

单元状态 C_t 的计算如式(8)所示。

$$C_t = F_t \cdot C_{t-1} + I_t \cdot \tilde{C}_t \quad (8)$$

输出门 O_t 的计算如式(9)所示。

$$O_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \quad (9)$$

更新隐藏状态 h_t 的计算如式(10)所示。

$$h_t = O_t \cdot \tanh(C_t) \quad (10)$$

式中: x_t 为第 t 时刻的特征向量; h_{t-1} 为 $t-1$ 时刻的输出; σ 为 sigmoid 函数; $\tanh$ 为双曲正切函数;

W_f 、 W_i 、 W_c 和 W_o 均为权重, 其中下标 f 、 i 、 c 、 o 分别表示遗忘门、输入门、候选单元状态、输出门; b_f 、 b_i 、 b_c 和 b_o 均为偏置; $[h_{t-1}, x_t]$ 表示向量 h_{t-1} 和 x_t 拼接。

1.4 基于 AM-CNN-LSTM 的短期风电功率预测流程

结合 MHAM、CNN 和 LSTM 的超短期风电功率预测模型(AM-CNN-LSTM)的预测流程如图 4 所示。具体步骤如下。

1) 数据处理: 对源数据进行清洗, 剔除无效和异常数据, 并按时间顺序排列后划分为训练集、验证集和测试集。

2) 模型训练: 原始数据 X 经过 MHAM 处理后得到新的特征表示 X' , 突出关键信息。接着 CNN 对 X' 提取空间模式和局部依赖关系, 输出矩阵 X'' 。随后, X'' 输入 LSTM 中捕捉时间序列中的长期依赖性和动态变化, 得到最后一个隐状态输出 L 。最终, L 经过全连接层融合特征, 输出预测向量 y_t 。

3) 计算模型损失: 使用验证集的平均绝对百分比误差(mean absolute percentage error, MAPE)评估模型的损失。

4) 参数优化: 通过优化器调整模型参数, 直至达到迭代次数上限后保存最终参数, 完成训练。

5) 测试模型: 将测试集输入训练好的模型, 评估其性能与预测精度。

2 实验验证

2.1 数据说明

本研究以新疆某风电场 5 号风机为实验对象, 其型号为华创 CCWE 1 500, 额定功率 1 500 kW。数据集覆盖 2020 年 3 月 25 日—6 月 5 日, 采样频率为每 15 min, 共清洗获得 6 720 条有效数据。每条数据包含时间信息(月份、日期、具体时间)、历史风电功率及 8 个气象变量。

使用 Pearson 相关系数分析 8 个气象变量与风电功率的线性相关性。结果如表 1 所示。

表 1 可知, 风速与风电功率相关性最高(0.93), 表现出极强正相关性; 其他变量相关性较弱或极弱。然而, Pearson 相关系数仅反映线性关系, 忽略了复杂的非线性和变量间交互作用^[3]。因此, 尽管大多数变量相关性较低, 本文将在实验中保留全部 8 个气象变量, 以捕捉潜在的非线性和交

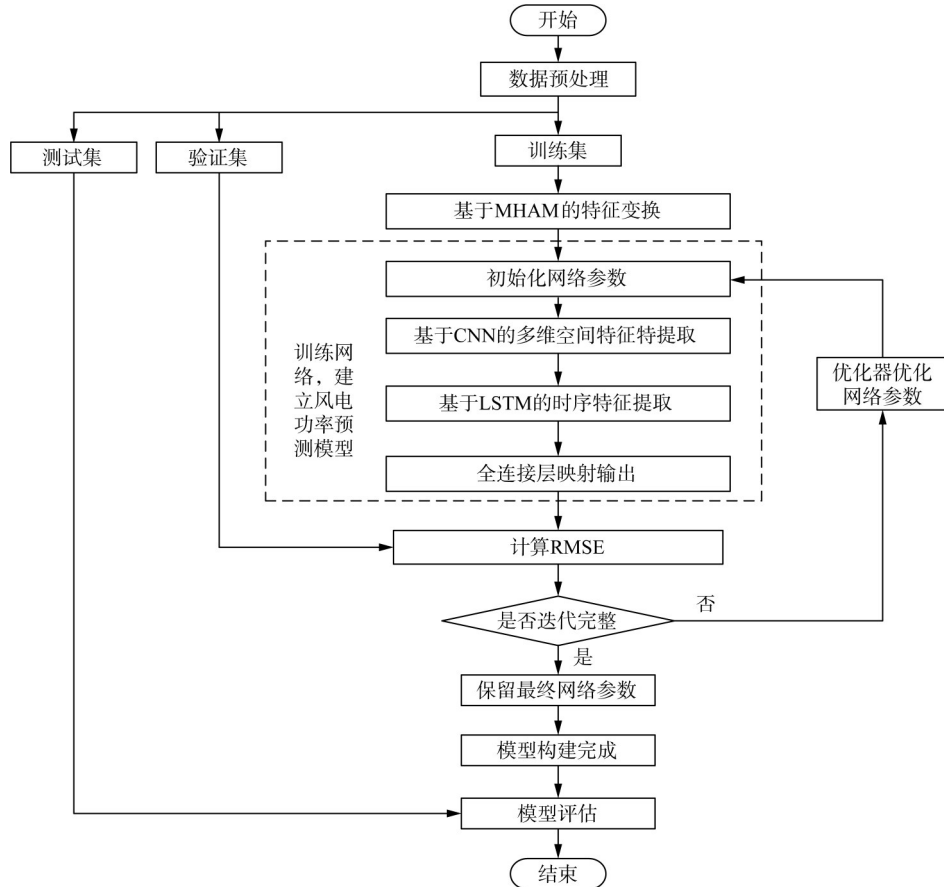


图4 基于AM-CNN-LSTM的风电功率预测模型构建与评估流程

Fig. 4 Construction and evaluation process of wind power forecasting model based on AM-CNN-LSTM

表1 气象变量与风电功率之间的相关系数

Tab. 1 Correlation coefficients between meteorological variables and wind powers

变量	与风电功率之间的相关系数
风速	0.93
风向	0.31
气温	0.15
气压	-0.22
湿度	0.12
大气稳定性	-0.13
太阳辐射	0.13
降水	0.08

互影响,进一步分析其对预测精度的影响。

2.2 评估指标

本文以平均绝对误差(mean absolute error, MAE)、均方根误差(root mean squared error, RMSE)和平均绝对百分比误差(mean absolute percentage error, MAPE)评估模型性能,3种指标

值越小,预测准确性越高。计算公式为:

$$E_{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (11)$$

$$E_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|^2} \quad (12)$$

$$E_{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \quad (13)$$

式中: E_{MAE} 、 E_{RMSE} 、 E_{MAPE} 分别为MAE值、RMSE值和MAPE值; n 为样本数量; y_i 为实际的风电功率; $\hat{y}_i$ 为模型预测的风电功率。

2.3 模型设置

为消除数据量纲和范围差异的影响,原始数据归一化至[0, 1]区间。数据集按时间顺序划分为训练集(80%)、验证集(10%)和测试集(10%)。

AM-CNN-LSTM的超参数设置为:全局学习率为0.001,批量大小为128,注意力头数量为4,卷积核大小为4, LSTM层数为3,隐层神经元数量为50,时间窗口为1 h(4个样本点)^[19-20]。为优化模型

权重与偏置，本文对比 Adam、SGD、GA、PSO、WOA 和 GWO 6 种优化器性能。设置初始种群数量为 100，迭代次数为 300，RMSE 为损失函数^[21-23]，其余参数如表 2 所示。

表 2 优化器参数设置

Tab. 2 Optimizer parameter setting

优化器	参数设置
Adam	
SGD	动量:0.9
GA	交叉概率:0.8,变异概率:0.01
PSO	惯性权重:0.5,个体学习因子:2,社会学习因子:2,速度范围:[-1,1]
WOA	包围系数:线性递减从2到0,螺旋更新系数:1,距离参数:随机在[-1,1]之间
GWO	α, β, δ 3种狼比例:0.2, 0.3, 0.5,收敛因子:线性递减从2到0

图 5 展示了各优化器的损失曲线。结果显示，Adam 和 SGD 在 300 次迭代内损失较高且波动明显，未能收敛；智能优化算法 GA、PSO、WOA 和 GWO 表现更优。对比 4 种智能优化算法，WOA 和 GWO 的损失明显低于 GA 和 PSO，且在前 250—300 次迭代内趋于稳定。其中，WOA 的损失最低，且在前 50 次迭代内实现快速收敛，后续波动极小，尤为稳定。因此，本文选用 WOA 作为模型优化器，以优化模型权重和偏置。

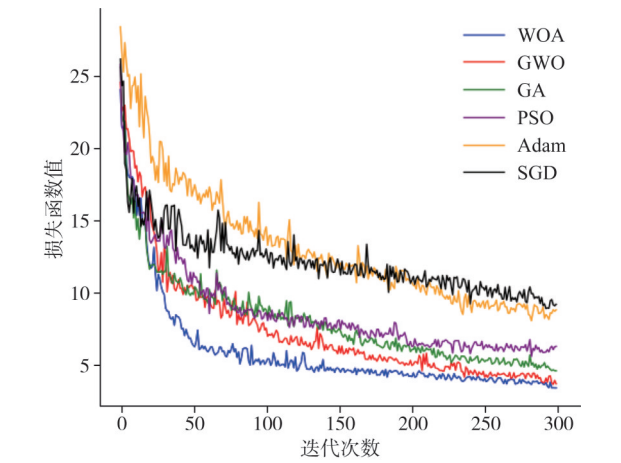


图 5 AM-CNN-LSTM 模型损失函数的变化曲线
Fig. 5 Change curves of loss function of AM-CNN-LSTM model

2.4 预测结果分析

为验证 AM-CNN-LSTM 的有效性，将其与 LSTM^[10]、CNN^[8]和 CNN-LSTM^[15]对比。时间窗口

对时序预测精度影响显著，但目前无统一标准，多依赖经验设定^[24-27]。基于本研究的数据特点，分别选用 2 h、4 h、6 h 和 8 h 历史数据预测未来 2 h 功率。

不同模型在时间窗口下的单步预测值与实际值对比曲线如图 6—9 所示。整体来看，4 种模型预测结果与实际值相近，表明基于 LSTM 和 CNN 的模型在风电功率预测中表现良好，与现有研究的结论一致。局部放大图显示，在 $T=2\text{ h}$ 、 4 h 、 6 h 时，AM-CNN-LSTM (蓝色虚线) 预测值更贴近实际值 (黑色实线)，偏差较小，表现出更高的预测精度和稳定性。这得益于其结合注意力机制、CNN 和 LSTM 的优势，能有效提取并利用时序数据中的复杂特征。然而，在 $T=8\text{ h}$ 时，AM-CNN-LSTM 与其他模型 (LSTM、CNN、CNN-LSTM) 相比均出现较大偏差，难以准确贴近实际值。这表明，随着时间窗口延长，数据复杂性和不确定性增加，模型预测精度下降。

表 3 不同模型预测精度对比及最优时间窗口选择
Tab. 3 Comparison of forecasting accuracies of different models and selection of optimal time window

模型	不同时间窗口	未来 2 h(8 个样本点)的预测误差均值			最优时间窗口
		MAE 值	RMSE 值	MAPE 值/%	
CNN	$T=2\text{ h}$	16.93	24.46	4.25	√
	$T=4\text{ h}$	16.41	23.69	4.29	
	$T=6\text{ h}$	19.43	27.3	5.04	
	$T=8\text{ h}$	21.33	31.33	5.52	
均值		18.53	26.70	4.78	
LSTM	$T=2\text{ h}$	20.89	30.98	5.57	√
	$T=4\text{ h}$	18.35	26.89	4.68	
	$T=6\text{ h}$	16.56	23.87	4.24	
	$T=8\text{ h}$	16.82	24.33	4.30	
均值		18.16	26.52	4.70	
CNN-LSTM	$T=2\text{ h}$	15.73	22.85	4.15	√
	$T=4\text{ h}$	14.92	22.15	3.97	
	$T=6\text{ h}$	14.09	20.56	3.71	
	$T=8\text{ h}$	16.08	23.38	4.17	
均值		15.21	22.24	4.00	
AM-CNN-LSTM	$T=2\text{ h}$	13.72	20.38	3.62	√
	$T=4\text{ h}$	13.45	19.48	3.78	
	$T=6\text{ h}$	14.51	21.68	3.81	
	$T=8\text{ h}$	15.00	21.85	3.85	
均值		14.17	20.85	3.77	

表3比较了不同时间窗口对未来2 h(8个样本点)预测误差的均值。

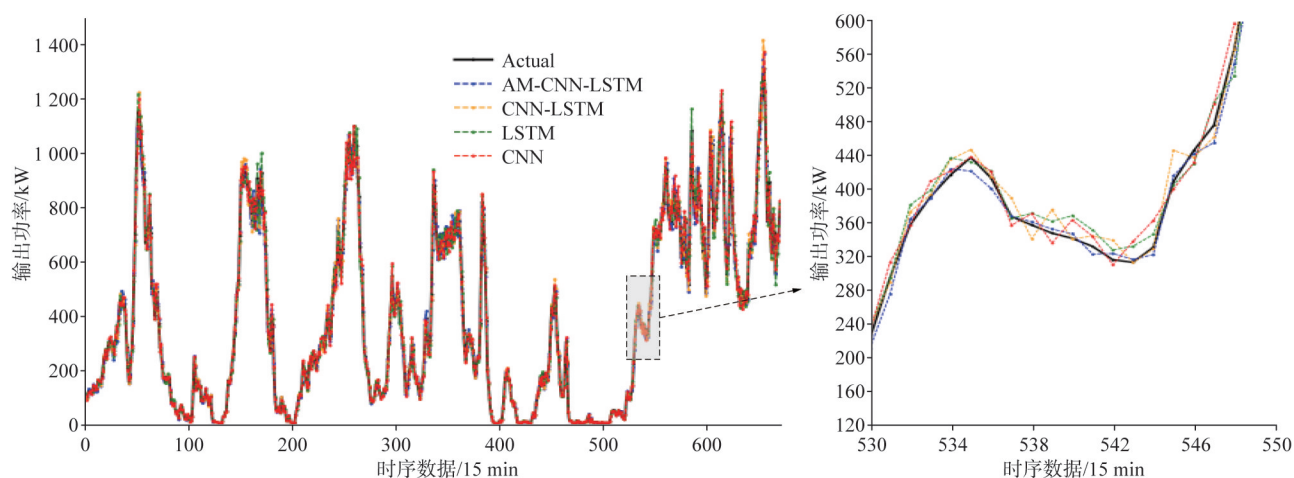


图6 时间窗口 $T=2$ h 的单步风电功率预测值与实际值对比

Fig. 6 Comparison between the predicted values and the real values of wind powers in time window $T=2$ h

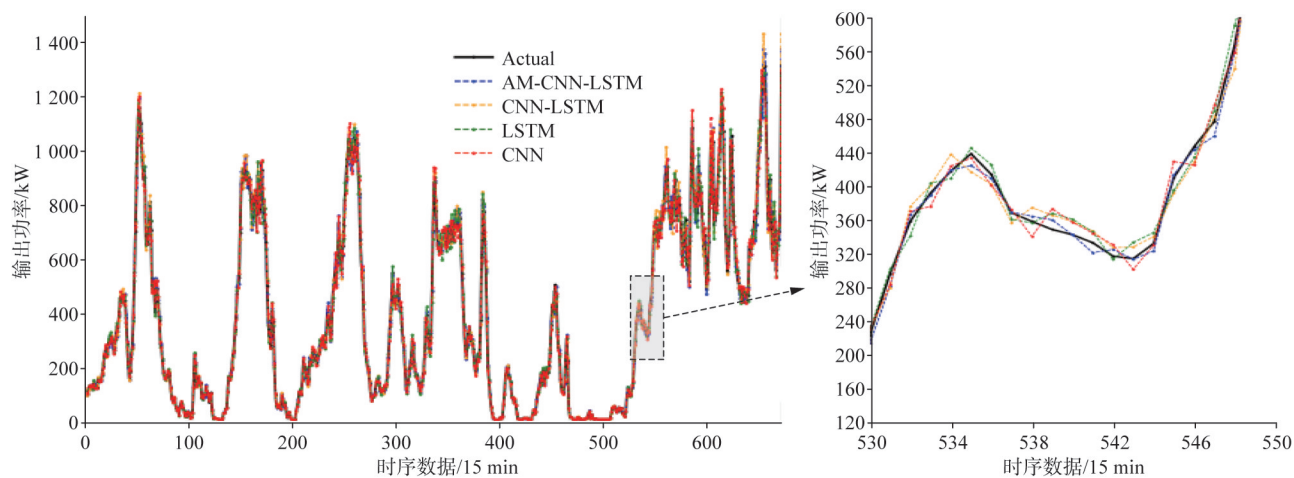


图7 时间窗口 $T=4$ h 的单步风电功率预测值与实际值对比

Fig. 7 Comparison between the predicted values and the real values of wind powers in time window $T=4$ h

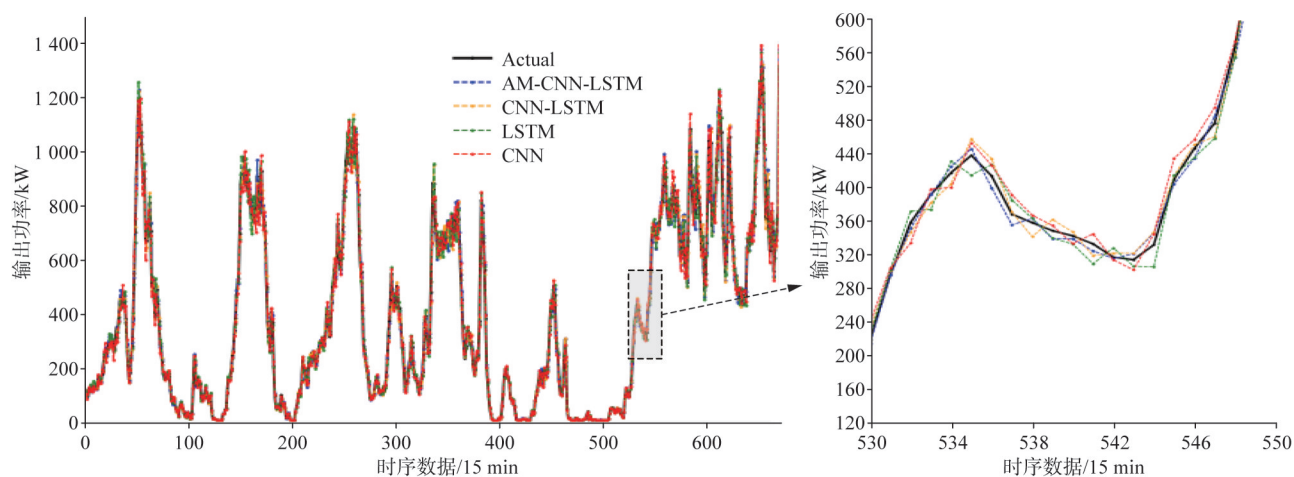


图8 时间窗口 $T=6$ h 的单步风电功率预测值与实际值对比

Fig. 8 Comparison between the predicted values and the real values of wind powers in time window $T=6$ h

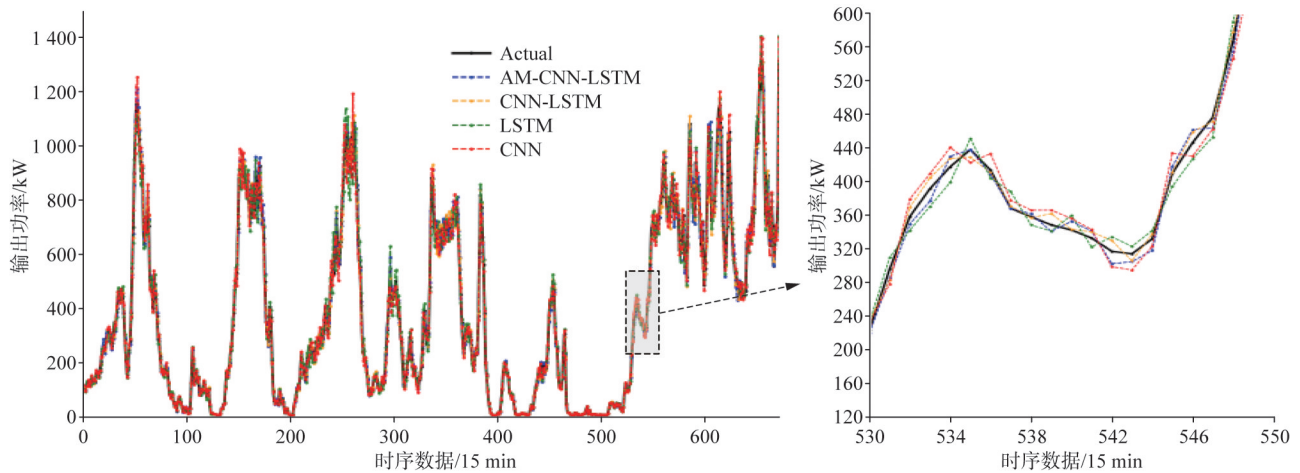


图9 时间窗口 $T=8$ h 的单步风电功率预测值与实际值对比

Fig. 9 Comparison between the predicted values and the real values of wind powers in time window $T=8$ h

结果显示, CNN 和 AM-CNN-LSTM 在 $T=4$ h 时精度最高, LSTM 和 CNN-LSTM 在 $T=6$ h 时表现最佳。此外, LSTM 在 $T=2$ h 时的误差最大, 而其他 3 种模型在 $T=8$ h 时的误差最大。这反映了模型对特征提取和长序列处理能力的差异: CNN 擅长提取局部特征, 在较短时间窗口 ($T=4$ h) 内捕捉风电数据短期变化效果良好, 但在 $T=8$ h 时因缺乏处理长时间依赖的能力, 预测精度下降。LSTM 能够捕捉长时间序列依赖关系, 在中等时间窗口 ($T=6$ h) 时表现优异, 但在 $T=2$ h 时无法充分发挥其优势, 误差显著增加。CNN-LSTM 结合了局部特征提取与长期依赖捕捉能力, 在 $T=6$ h 时表现较好, 但在 $T=8$ h 时因噪声和冗余信息累积导致预测精度下降。AM-CNN-LSTM 模型通过引入注意力机制, 增强了对关键特征的提取能力, 在时间窗口为 $T=4$ h 时表现突出, 在处理更长时间序列 ($T=8$ h) 时, 虽然精度下降, 但优于对比模型。当时间窗口较长 ($T=8$ h) 时, 大多数模型的预测误差增加, 这可能是由于长时间序列中的噪声和不相关信息的累积所致。另一方面, 长时间尺度内还会受到气候因素之外的一些其他外部因素 (如风电设备故障、电网波动等) 的影响, 从而进一步增加了预测误差。

进一步分析表 3 发现, AM-CNN-LSTM 在最优时间窗口的误差和平均误差均低于其他模型。例如, 其 MAE 值、RMSE 值和 MAPE 值相比 CNN 分别降低了 4.36 %、5.85 % 和 1.01 %; 相比 LSTM 分别降低了 3.99 %、5.67 % 和 0.93 %; 相比

CNN-LSTM 分别降低了 1.04 %、1.39 % 和 0.23 %。这得益于 AM-CNN-LSTM 结合了 MHAM、CNN 和 LSTM 的优势, 能够更好地学习关键特征和多尺度特征, 提高预测精度。在多个误差指标上, AM-CNN-LSTM 均展现了明显的性能优势。

为进一步探讨模型在超短期预测中的表现, 表 4 展示了各模型在其最优时间窗口下对未来 2 h (8 个样本点) 的单点预测误差 (RMSE 值)。结果表明, AM-CNN-LSTM 在 $t+15$ min 至 $t+120$ min 的预测误差均低于其他模型, 尤其在 $t+75$ min 至 $t+120$ min 范围内, 其 RMSE 值显著低于 CNN、LSTM 和 CNN-LSTM。此外, AM-CNN-LSTM 的 RMSE 在各时间点之间波动较小, 显示出更好的预测稳定性。这一结果表明, AM-CNN-LSTM 在超短期风电功率预测中具有显著优势, 能够更准确且稳定地捕捉短时间内的变化, 适应复杂时序数据特性, 为实际应用提供可靠的预测支持。

2.5 变量对模型的精度影响分析

本节通过在历史风电功率数据的基础上逐一添加或组合 8 个气象变量, 分析其对模型 MAE 值、RMSE 值和 MAPE 值的影响, 实验结果如表 5 所示。

由表 5 可知, 单独使用风速时, 模型 MAE 值为 19.78, RMSE 值为 27.67, MAPE 值为 5.09 %。使用风向、气压等弱相关变量时, 预测误差较高, 尤其是单独使用极弱相关变量时。这表明风速对风电功率影响最大, 其次是风向和

表 4 不同模型在最优时间窗口下的超短期风电功率预测误差比较

Tab. 4 Comparison of ultra-short-term wind power forecasting errors of different models under the optimal time windows										
模型	最优时间窗口	未来 2 h(8 个样本点)的风电功率预测误差(RMSE)								均值
		$t+15\text{ min}$	$t+30\text{ min}$	$t+45\text{ min}$	$t+60\text{ min}$	$t+75\text{ min}$	$t+90\text{ min}$	$t+105\text{ min}$	$t+120\text{ min}$	
CNN	$T=4\text{ h}$	22.83	22.74	22.82	23.69	23.81	23.92	24.82	24.89	23.69
LSTM	$T=6\text{ h}$	23.00	23.07	22.92	23.22	23.59	24.72	24.92	25.49	23.87
CNN-LSTM	$T=4\text{ h}$	19.78	19.90	19.73	19.81	20.71	21.24	21.58	21.72	20.56
AM-CNN-LSTM	$T=6\text{ h}$	19.12	19.03	19.19	19.40	19.59	19.86	19.78	19.90	19.48

表 5 气象变量对风电功率预测精度的影响
Tab. 5 Influence of climate variables on wind power forecasting accuracies

变量	MAE 值	RMSE 值	MAPE 值/%
历史风电功率+风速(极强相关变量)	19.78	27.67	5.09
历史风电功率+风向(弱相关变量)	45.15	62.69	11.87
历史风电功率+气压(弱相关变量)	43.85	60.74	11.68
历史风电功率+气温(极弱相关变量)	84.41	120.17	21.56
历史风电功率+太阳辐射(极弱相关变量)	81.46	113.77	21.25
历史风电功率+大气稳定性(极弱相关变量)	87.32	125.34	22.20
历史风电功率+湿度(极弱相关变量)	86.85	124.65	21.91
历史风电功率+降水(极弱相关变量)	89.54	125.44	23.42
历史风电功率+风向+气压	39.85	53.29	10.52
历史风电功率+气温+大气稳定性+太阳辐射+湿度+降水	71.98	99.29	18.75
历史风电功率+风速+风向+气压	14.31	21.13	3.88
历史风电功率+全部气象变量	13.45	19.48	3.78

气压,与 Pearson 相关系数分析一致。组合多个变量显著提升了预测精度。例如,风向与气压组合使 MAE 值降至 39.85, RMSE 值为 53.2, MAPE 值为 10.52%, 优于单一变量。同样地,气温、大气稳定性等 5 个变量组合使用的预测误差也明显低于单独使用其中一个的误差。这表明气温、大气稳定性、太阳辐射、湿度和降水这些变量虽然单独来看与风电功率的线性相关性较弱,但它们可以通过复杂的非线性关系共同影响风电功率。与单独使用风速(极强相关变量)的模型相比,组合使用风速、风向、气压的模型 MAE 值降低至 14.31, RMSE 值降至 21.13, MAPE 值降至 3.88%。最终,当使用全部 8 个气象变量时,模型误差进一步降低: MAE 值为 13.45, RMSE 值为 19.48, MAPE 值为 3.78%。

这表明,结合多个气象变量显著提升了风电功率预测精度,并证明 AM-CNN-LSTM 模型能够充分挖掘变量间复杂交互和非线性关系,实现更精准的预测。

3 结论

本文基于历史风电功率和气象数据,结合多头自注意力机制、卷积神经网络和长短期记忆网络,提出了一种新型超短期风电功率预测模型(AM-CNN-LSTM)。研究结果表明如下。

1) WOA 优化器在模型参数优化中表现最佳,显著提升了预测精度,相较 Adam、SGD 及其他智能优化算法(GA、PSO、GWO)具有更突出的优越性。

2) AM-CNN-LSTM 结合 HMAM 的关键特征识别、CNN 的非线性特征提取和 LSTM 的时序捕捉能力,在 4 种时间窗口下的预测精(MAE)度较 CNN 和 LSTM 平均提升 4.36 和 3.99,较 CNN-LSTM 提升 1.04,展现了卓越的超短期预测性能。

3) 风速对风电功率影响最大,其次为风向和气压。组合多个气象变量显著提高预测精度,反映了变量间复杂的非线性关系,验证了 AM-CNN-LSTM 捕捉交互关系的能力。

未来将探索区间预测方法,研究不同时间窗口对预测周期的影响,并与更多模型对比,以提升预测可靠性与实用性。

参考文献

[1] WANG G, JIA R, LIU J, et al. A hybrid wind power forecasting approach based on Bayesian model averaging and ensemble learning[J]. Renew Energy, 2020(145): 2426 – 2434.
[2] YANG M, WANG D, ZHANG W. A short-term wind power prediction method based on dynamic and static feature fusion mining[J]. Energy, 2023(280): 128226.1 – 128226.12.
[3] 刘雨佳,樊艳芳,白雪岩,等. 基于特征交叉机制和误差补偿

- 的风力发电功率短期预测[J]. 电工技术学报, 2023, 38(12): 3277 – 3288.
- LIU Yujia, FAN Yanfang, BAI Xueyan, et al. Short-term wind power prediction based on feature crossover mechanism and error compensation[J]. Transactions of China Electrotechnical Society, 2023, 38(12): 3277 – 3288.
- [4] ZHANG J H, YAN J, INFELD D, et al. Short-term forecasting and uncertainty analysis of wind turbine power based on long short-term memory network and Gaussian mixture model[J]. Applied Energy, 2019(241): 229 – 244.
- [5] 沈小军, 周冲成, 付雪娇. 基于机联网-空间相关性权重的风电机组风速预测研究[J]. 电工技术学报, 2021, 36(9): 1782 – 1790, 1817.
- SHEN Xiaojun, ZHOU Chongcheng, FU Xuejiao. Wind speed prediction of wind turbine based on the internet of machines and spatial correlation weight[J]. Transactions of China Electrotechnical Society, 2021, 36(9): 1782 – 1790, 1817.
- [6] POGGI P, MUSELLI M, NOTTON G, et al. Forecasting and simulating wind speed in Corsica by using an autoregressive model[J]. Energy Conversion and Management, 2003, 44(20): 3177 – 3196.
- [7] WANG J, QIN S, ZHOU Q, et al. Medium-term wind speeds forecasting utilizing hybrid models for three different sites in Xinjiang, China[J]. Renew Energy, 2015(76): 91 – 101.
- [8] LIU J, SHI Q, HAN R, et al. A hybrid GA-PSO-CNN model for ultra-short-term wind power forecasting[J]. Energies, 2021, 14(20): 6500.1 – 6500.22.
- [9] LIN Y, ZHANG H, LIU J Y, et al. Research on short-term wind power Prediction of GRU based on similar days[C]// Journal of Physics: Conference Series, September 24 – 26, 2021, Jilin, China. Beijing: IOP Publishing, 2021, 2087(1): 012089.1 – 012089.9.
- [10] TUERXUN W, XU C, GUO H, et al. A wind power forecasting model using LSTM optimized by the modified bald eagle search algorithm[J]. Energies, 2022, 15(6): 2031.1 – 2031.10.
- [11] 薛阳, 王琳, 王舒, 等. 一种结合 CNN 和 GRU 网络的超短期风电预测模型[J]. 可再生能源, 2019, 37(3): 456 – 462.
- XUE Yang, WANG Lin, WANG Shu, et al. An ultra-short-term wind power forecasting model combined with CNN and GRU networks[J]. Renewable Energy Resources, 2019, 37(3): 456 – 462.
- [12] ZHAO Z, YUN S, JIA L, et al. Hybrid VMD-CNN-GRU-based model for short-term forecasting of wind power considering spatio-temporal features[J]. Engineering Applications of Artificial Intelligence, 2023(121): 105982.1 – 105982.10.
- [13] 程杰, 陈鼎, 李春, 等. 基于 GWO-CNN-BiLSTM 的超短期风电预测[J]. 科学技术与工程, 2023, 23(35): 15091 – 15099.
- CHENG Jie, CHEN Ding, LI Chun, et al. Ultra-short-term wind power prediction based on GWO-CNN-BiLSTM[J]. Science Technology and Engineering, 2023, 23(35): 15091 – 15099.
- [14] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing System, December 4, 2017, Long Beach, USA. Long Beach: IONIP, 2017: 6000 – 6010.
- [15] 姚越, 刘达. 基于注意力机制的卷积神经网络-长短期记忆网络的短期风电功率预测[J]. 现代电力, 2022, 39(2): 212 – 218.
- YAO Yue, LIU Da. Short-term wind power forecasting based on attention mechanism of CNN-LSTM[J]. Modern Electric Power, 2022, 39(2): 212 – 218.
- [16] MA Z, MEI G. A hybrid attention-based deep learning approach for wind power prediction[J]. Applied Energy, 2022(323): 119608.1 – 119608.33.
- [17] WANG S, SHI J, YANG W, et al. High and low frequency wind power prediction based on transformer and BiGRU-Attention[J]. Energy, 2024(288): 129753.1 – 129753.14.
- [18] 龙铖, 余成波, 何铖, 等. 基于双重注意力机制 CNN-BiLSTM 与 LightGBM 误差修正的超短期风电功率预测[J]. 电气工程学报, 2024, 19(2): 138 – 145.
- LONG Cheng, YU Chengbo, HE Cheng, et al. Ultra-short-term wind power prediction based on dual attention mechanism CNN-BiLSTM and LightGBM error correction[J]. Journal of Electrical Engineering, 2024, 19(2): 138 – 145.
- [19] 廖雪超, 陈才圣, 伍杰平. 基于 CNN-LSTM 及深度学习的风电场时空组合预测模型[J]. 信息与控制, 2022, 51(4): 498 – 512.
- LIAO Xuechao, CHEN Caisheng, WU Jieping. Combined spatiotemporal wind farm prediction model based on CNN-LSTM and deep learning[J]. Information and Control, 2022, 51(4): 498 – 512.
- [20] 张子华, 李琰, 徐天奇, 等. 基于麻雀算法优化的 VMD-CNN-LSTM 的短期风电功率研究[J]. 电气传动, 2023, 53(5): 77 – 83.
- ZHANG Zihua, LI Yan, XU Tianqi, et al. Research on short-term wind power forecasting based on VMD-CNN-LSTM optimized by sparrow Algorithm[J]. Electric Drive, 2023, 53(5): 77 – 83.
- [21] KASSA Y, ZHANG J H, ZHENG D H, et al. A GA-BP hybrid algorithm based ANN model for wind power prediction[C]//2016 IEEE Smart Energy Grid Engineering (SEGE), August 21 – 24, 2016, Oshawa, ON, Canada. New York: IEEE, 2016: 158 – 163.
- [22] YU M, NIU D, GAO T, et al. A novel framework for ultra-short-term interval wind power prediction based on RF-WOA-VMD and BiGRU optimized by the attention mechanism[J]. Energy, 2023(269): 126738.1 – 126738.17.
- [23] ZHANG J, ZHAO Z, YAN J, et al. Ultra-short-term wind power forecasting based on CGAN-CNN-LSTM model supported by lidar[J]. Sensors, 2023, 23(9): 4369.1 – 4369.20.
- [24] 常雨芳, 杨子潇, 潘风, 等. 基于 CEEMDAN-PE-WPD 和多目标优化的超短期风电功率预测方法[J]. 电网技术, 2023, 47(12): 5015 – 5026.
- CHANG Yufang, YANG Zixiao, PAN Feng, et al. Ultra-short-term wind power prediction based on CEEMDAN-PE-WPD and multi-objective optimization[J]. Power System Technology, 2023, 47(12): 5015 – 5026.
- [25] 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.