

基于特征优选与相似样本融合的LSTM-AM短期风电功率预测

吴琛¹, 崔秋实², 谢一工³, 黄润¹, 张海涛², 方斯顿², 牛涛², 陈冠宏²

(1. 云南电网有限责任公司规划建设研究中心, 昆明 650033; 2. 重庆大学电气工程学院, 重庆 400044; 3. 云南电网有限责任公司电力调度控制中心, 昆明 650011)

摘要: 随着“双碳”目标的深入推进, 近年来我国风电行业迅速发展, 如何精准有效地预测风电功率对实现风机安全并网和维持系统稳定运行至关重要。针对现有风电功率预测方法存在输入特征冗余、泛化能力不足和未能充分捕捉风电出力内在特性等问题, 提出了一种基于特征优选与相似样本融合的长短期记忆网络与注意力机制(long short term memory-attention memory, LSTM-AM)短期风电功率预测模型。首先, 利用最小绝对收缩和选择算子(least absolute shrinkage and selection operator, Lasso)回归进行输入特征优选, 减少冗余; 然后, 采用长短期记忆网络与注意力机制建立LSTM-AM融合网络模型; 最后, 通过欧氏距离计算提取相似历史样本, 与模型输出加权作为最终预测值。实验结果表明, 所提出的方法相比传统方法预测性能更优, 在风电功率预测中表现出更高的准确性, 能够为电力系统规划运行和可再生能源的深入应用提供支撑。

关键词: LSTM-AM融合模型; 风电功率预测; 相似样本提取; 电力规划运行

LSTM-AM Short-Term Wind Power Prediction Based on Feature Optimization and Similar Sample Fusion

WU Chen¹, CUI Qiushi², XIE Yigong³, HUANG Run¹, ZHANG Haitao², FANG Sidun², NIU Tao², CHEN Guanhong²

(1. Planning and Construction Research Center of Yunnan Power Grid Co., Ltd., Kunming 650033, China; 2. School of Electrical Engineering, Chongqing University, Chongqing 400044, China; 3. Power Dispatching and Control Center of Yunnan Power Grid Co., Ltd., Kunming 650011, China)

Abstract: With the further promotion of the "Double Carbon" goal, China's wind power industry has developed rapidly in recent years. How to accurately and effectively predict wind power is crucial to realize the safe grid connection of wind turbines and maintain the stable operation of the system. Aiming at the problems such as input feature redundancy, insufficient generalization ability, and insufficient capture of inherent wind power output characteristics in existing wind power forecasting methods, a short-term wind power prediction model based on feature selection and similarity based fusion using long short term memory network-attention mechanism (LSTM-AM) is proposed. Firstly, least absolute shrinkage and selection operator (Lasso) regression is used to optimize input features to reduce redundancy. Then, the long short term memory network-attention mechanism (LSTM-AM) fusion network model is established using long short term memory and attention mechanism. Finally, similar historical samples are extracted by Euclidean distance calculation and weighted with the model output as the final predicted value. The experimental results show that compared with traditional methods, the prediction accuracy of the proposed method is significantly improved, and it shows higher accuracy in wind power prediction, which can provide support for power system planning and operation and the in-depth application of renewable energy.

Key words: LSTM-AM fusion model; wind power prediction; similar sample extraction; power system planning and operation

基金项目: 国家自然科学基金资助项目(52377075); 中国南方电网有限责任公司总部管理科技项目(0500002023030301GH00107)。

Foundation item: Supported by the National Natural Science Foundation of China (52377075); the Science and Technology Project of the Headquarters of China Southern Power Grid Co., Ltd. (0500002023030301GH00107).

0 引言

能源开发利用水平、资源利用效率以及人均能源消费量是反映国民经济与社会发展状况的重要指标。目前,全球能源消费仍高度依赖煤、石油和天然气等传统化石能源,这些能源在全球总能耗中占比超过80%^[1]。然而,随着能源需求的不断增长,矿物资源的消耗速度加快,导致其储量急剧减少。由于矿物资源碳排放高且不可再生,对全球发展和能源安全带来了巨大挑战。因此,开发利用风能、太阳能、水能等零碳且可再生的清洁能源以替代传统化石能源,对助力电力系统转型和经济社会可持续发展具有重要意义^[2-5]。

风能凭借资源丰富、环境友好和易于开发等优势成为技术相对成熟并被广泛利用的绿色能源^[6]。然而,风力发电因其不稳定性和间歇性被视作一种不可控的自然资源。将风电并入主网可能导致供需不平衡,进而影响系统的安全稳定运行,还可能引发风能浪费并降低电网经济性。为提升风电在能源结构中的比例及其利用效率,支撑电力系统规划与运行,保证风电功率预测的精度至关重要^[7]。

风力发电序列的波动性较大,直接对其进行建模会导致严重的过拟合。与实际值相比,预测值偏差严重,无法准确预测风电的变化规律^[8]。当前的风电预测研究领域主要包括统计模型和数据驱动两类方法^[9]。统计方法主要包含移动平均^[10]、指数平滑^[11]、回归分析^[12]和卡尔曼滤波等^[13],这些方法虽然简单易行,但难以反映非线性数据带来的影响,无法满足风电预测的高精度要求。数据驱动方法如机器学习和神经网络等^[14-15]虽能够有效处理大量数据中的非线性问题,但在提取数据间的关联信息方面依然面临一定的困难。

随着近年来深度学习的发展,大量研究者利用深度神经网络来进行风电功率预测,并取得了显著成果。国内外相关研究表明采用神经网络的预测方法具有较高的准确性和可靠性,能够有效提高预测精度,保证风电机组的安全运行^[16]。

在国内,许多学者采用BP神经网络^[17]、自编码器(autoencoder)^[18]等模型开展相关研究。赵建利^[19]等提出一种将卷积神经网络(convolutional neural network, CNN)与长短期记忆网络(long short term memory network, LSTM)结合的方法来

进行短期风电出力预测,首先通过CNN对风电出力和气象数据进行卷积与池化处理,提取有关信息的特征,然后将其输入到LSTM模型中得到预测值。孙改平等^[20]在改进粒子群算法的基础上,设计了一种小波神经网络来实现出力预测,并以实际数据进行算例验证,该方法能够有效挖掘数据特性,使得风电预测精度提高。

国外方面,欧美等地区的研究者也积极投入于风电功率预测研究工作。近年来,随着人工智能技术的快速发展,神经网络模型逐渐被引入到风电功率预测领域中,ARIMA、SVM和ANN等模型^[21-23]被广泛应用。Nima Safari^[24]等考虑风电时序数据的混沌性质,应用多尺度奇异谱分析提出了一种局部直接预测和迭代预测相结合的短期多步预测方法,并利用加拿大某风电场的实际数据验证了所提方法在预测风力发电量方面的有效性。Vishnuthierth^[25]等将双向长短期记忆网络(bidirectional long short term memory network, BiLSTM)与双向门控架构单元(bidirectional gated architecture unit, BiGRU)相结合,提出了一种新的风电预测方法,结果具有较好的可信度和误差水平,能够进一步应用于风能管理中提高效率 and 可靠性。

然而,基于神经网络模型的方法预测效果大多依赖于输入特征和数据质量,存在泛化能力不足、鲁棒性差等问题^[26]。为此,本文针对风电功率的实际特性,将LSTM与注意力机制相结合,充分利用LSTM网络在处理时序数据方面的优势,并引入注意力机制来增强模型的关键信息识别能力和适应性。在此基础上,基于Lasso回归对输入特征优化选取,并通过欧氏距离计算提取相似历史样本进行输出优化,以克服传统LSTM和AM存在的输入特征冗余、过拟合及泛化能力不足等问题,进一步提升预测性能。该方法融合相似样本来学习历史规律,能够更好地捕捉风力发电内在特性,从而提升预测性能,实现支撑电网规划运行、保障风电安全并网的目的。

1 基于Lasso回归的输入特征优选

在短期风电功率预测问题中输入特征的选择对于模型性能至关重要。传统风电功率预测方法通常依赖于原始的时序数据,如风速、风向、温度和气压等^[27]。然而,数据中往往包含大量冗余特征,这

不仅增加了计算复杂度,还可能导致模型过拟合。为了有效地选择具有预测能力的关键特征,本文采用最小绝对收缩和选择算子(least absolute shrinkage and selection operator, Lasso)回归进行特征优选。

Lasso 回归是一种线性回归方法,通过对模型系数施加 L1 正则化能够自动筛选出最重要的特征,并将冗余特征的系数收缩为 0^[28]。Lasso 回归的基本思想是通过引入 L1 范数惩罚项避免过拟合的同时进行特征选择,使得最终的回归模型更加简洁且具有较强的解释能力。在本文的研究中,本文将 Lasso 回归应用于风电功率预测模型的输入特征选择,其中输入特征因素包括风电功率、风速、风向、温度和气压等。

考虑数据集 $D = (X, Y)$, $X \in R^{n \times m}$, $Y \in R^{n \times 1}$, 通过 Lasso 回归,本文能够识别出对风电功率预测具有关键作用的特征,并去除冗余或低相关性特征,从而提高模型的训练效率和预测性能。具体而言,本文采用 Lasso 回归进行特征优选,在以准确预测风电功率为目标的基础上引入惩罚项来收缩冗余特征,通过式(1)求解获得回归系数。

$$\hat{\beta} = \arg \min \left\{ Y - X\beta^2 + \lambda \sum_{i=1}^m |\beta_i| \right\} \quad (1)$$

式中: β 为回归系数向量; $\hat{\beta}$ 为其最优估计; X 为 $n \times m$ 的输入特征矩阵; Y 为 $n \times 1$ 目标矩阵变量; λ 为惩罚系数。

通过交叉验证选择最佳的 λ 值, Lasso 回归能够有效地选择出与目标变量最相关的元素,从而避免使用 LSTM 和注意力机制进行功率预测时的输入特征冗余和过拟合,提升模型计算效率与预测性能。

2 LSTM-AM 与相似样本融合模型构建

2.1 LSTM 网络

LSTM 是 Hochreiter 等人在对 RNN 改进时提出的一种新型神经网络模型^[29],其优点在于能够有效解决梯度爆炸和消失问题。它在隐藏层中添加输入、遗忘和输出 3 种门控单元,通过这些单元来记忆历史信息 and 更新数据状态。

LSTM 由 3 个信息作为输入,包括现在时刻信息,前一时刻短期记忆信息和前一时刻长期记忆信

息。内部采用 3 种门控单元来决定历史记忆信息与现在时刻信息保留程度,分别是输入门 i_t , 遗忘门 f_t 和输出门 o_t 。模型的组成结构如图 1 所示。

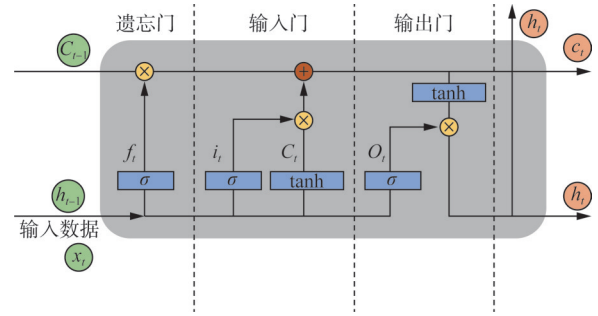


图 1 LSTM 模型结构

Fig. 1 Structure of LSTM model

具体学习步骤如下。

步骤 1: 遗忘门计算时,将前一时刻输出值和当前输入值共同作为输入,使用激活函数 sigmoid 计算传递信息值。

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2)$$

式中: W_f 为遗忘门的权重矩阵; $\sigma(\cdot)$ 为激活函数; b_f 为偏置向量; h_{t-1} 和 x_t 为遗忘门的输入。

步骤 2: 输入门计算时,同样将 h_{t-1} 和 x_t 作为输入,通过 sigmoid 激活函数计算,用 i_t 表示。

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (3)$$

式中: W_i 为输入门的权重矩阵; b_i 为偏置向量。

输入门采用 tanh 函数对单元状态进行更新,并用新的候选值替代 $\tilde{c}_t$ 传入记忆单元。

$$\tilde{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (4)$$

式中: W_c 为更新单元的权重矩阵; b_c 为偏置向量。

步骤 3: 输出门采取类似的方式进行计算,用于将单元信息输出到下一状态,用 o_t 表示。

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (5)$$

式中: W_o 为输出门的权重矩阵; b_o 为偏置向量。

步骤 4: 遗忘门与输入门结合输出新的长期记忆单元 c_t , 其中 $\otimes$ 为 Hadamard 积。

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tilde{c}_t \quad (6)$$

最后,输出门将 c_t 传入 tanh 函数,计算得到隐藏层输出值。

$$h_t = o_t \otimes \tanh(c_t) \quad (7)$$

2.2 注意力机制

当面临复杂的场景或听到一段话时人们总会不自觉地注意力集中于最重要的部分,而不是同时

处理所有信息。有关学者在这种行为的启发下提出了注意力机制 (attention mechanism, AM)^[30]。不同于传统神经网络对数据序列的处理方法, AM使得模型在生成目标序列的过程中能够动态地关注输入序列的不同部分, 而不是把整个输入压缩成固定的向量。其核心思想是: 在处理每个输入元素时, 模型可以对所有输入元素的贡献进行加权, 从而决定哪些部分更为重要。AM通过计算不同输入的权重来表征其重要程度, 使模型能够灵活选择聚焦于关键信息, 从而提升输出的准确度, 具体框架如图2所示。

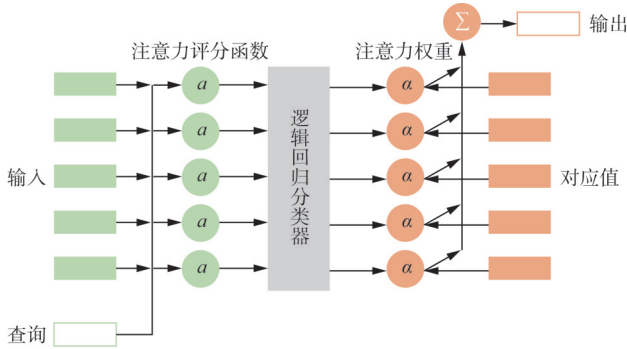


图2 注意力机制框架

Fig. 2 Attention mechanism framework

注意力聚焦函数如下。

$$f(\mathbf{q}, (k_1, v_1), \dots, (k_m, v_m)) = \sum_{i=1}^m \alpha(\mathbf{q}, k_i) v_i \quad v_i \in \mathbf{R}^v \quad (8)$$

式中: $(v_1, \dots, v_m)$ 为值序列; $(k_1, \dots, k_m)$ 为键序列; $\mathbf{q}$ 为相应的查询序列值; $\mathbf{R}^v$ 为维度为 v 的实数向量空间; m 为输入序列的长度。

a 为注意力评分, 本文采用指数形式的缩放点积函数表征, 将不同特征的注意力权重进行非线性放大, 如式(9)所示。

$$a(\mathbf{q}, k_i) = \frac{\exp(\alpha(\mathbf{q}, k_i))}{\sum_{j=1}^m \exp(\alpha(\mathbf{q}, k_j))} \quad (9)$$

式中 $a(\mathbf{q}, k_i)$ 为注意力权重, 计算公式如下。

$$\alpha(\mathbf{q}, k_i) = \frac{\mathbf{q}^T \mathbf{k}_i}{\sqrt{d}} \quad (10)$$

式中 d 为点积缩放因子, 取值为键向量维数。

2.3 LSTM-AM融合神经网络

本文将AM融入到传统LSTM网络, 以更有效地凸显输入特征信息, 进而提升模型性能。首先将时序数据输入LSTM隐藏层, 经过逐步处理后生成

输出和隐藏状态; 然后将状态信息输入AM计算相应权重, 并对原输出值加权求和, 得到聚焦于关键信息的新输出; 最后将输出与状态信息结合, 通过全连接层得到最终结果。本文将LSTM与AM融合后的模型结构如图3所示。

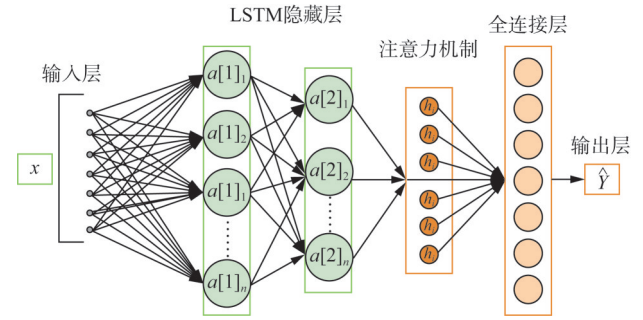


图3 LSTM-AM融合网络的模型结构

Fig. 3 Model structure of LSTM-AM fusion network

LSTM与AM融合后的网络模型包含三层关键结构, 分别为LSTM隐藏层、注意力机制层以及全连接层。LSTM层的输入、输出维度均为1, 隐藏单元数量设定为64, 逐步处理输入序列, 保留并传递数据信息。注意力机制层为线性结构, 以前一层的输出和隐藏状态作为输入, 通过评分函数计算权重。随后将这些权重与相应的原输出加权求和, 使得原输出值在更关键的特征信息上得到强化, 最后将加权后的值与状态信息输出到全连接层获得输出结果。

融合模型既拥有LSTM学习输入数据时序关系的能力, 能够较好地挖掘数据特征, 又通过AM让关键信息得到强化聚焦, 从而使模型性能和预测效果得到提升。

2.4 LSTM-AM与相似样本加权的融合模型

在短期风电功率预测任务中, 仅依赖LSTM-AM模型的输出无法充分捕捉到风电功率数据的历史规律和内在特性。因此, 为进一步提高预测精度与泛化能力, 本文将相似历史样本与LSTM-AM模型输出加权融合以优化预测性能。通过引入相似历史样本, 本文能够增强模型对历史规律的学习能力, 并在一定程度上抑制模型预测中的过拟合。

本文在测试集中融合LSTM-AM模型输出与相似历史样本, 采用欧式距离进行样本选取, 识别与当前窗口数据具有相似模式的历史样本。首先计算每个历史样本单元与当前窗口单元的相似度, 然后

根据相似度大小为每个样本赋予权重。采用指数形式的欧式距离定义相似度,将不同样本的差异进行非线性放大,计算如式(11)所示。

$$\text{sim}(x_i, z_j) = \exp\left(-\frac{\|x_i - z_j\|}{\gamma}\right) \quad (11)$$

式中: $\text{sim}(x_i, z_j)$ 为历史样本 x_i 与当前窗口单元 z_j 的相似度; γ 为缩放因子,用于将欧氏距离值缩放到合适大小便于相似度计算; x_i 为第 i 个历史样本; z_j 为测试集中第 j 个窗口单元。

选取 k 个相似度最高的样本与原输出进行融合,定义第 i 个样本的权重为 w_i , 可由式(12)得出。

$$w_i = \frac{\text{sim}(x_i, z_j)}{\sum_{i=1}^k \text{sim}(x_i, z_j)} \quad (12)$$

经过计算, k 个样本的权重得以归一化,并且相似度高的样本会获得较大权重。最后,将 LSTM-AM 模型输出与加权后的历史样本进行融合,得到本文所用方法的最终预测值,如式(13)所示。

$$\hat{y}_j = \omega_1 y_j + \sum_{i=1}^k \omega_i \times y_i \quad (13)$$

式中: $\hat{y}_j$ 为测试集中第 j 个窗口单元的最终预测值, y_j 为该窗口下 LSTM-AM 模型的输出值; ω_1 为其权重; y_i 为第 i 个相似样本单元的真实值; ω_i 为对应权重。

本文提出的风电功率预测方法原理如图 4 所示。该方法不仅可以充分利用 LSTM 模型与注意力机制对时序数据的学习能力,还能通过相似样本融合进一步提高模型的适应性和泛化能力,捕捉到更多潜在的风电功率变化特性,从而提供精准可靠的预测结果。

3 数据选取与处理

3.1 原始数据集选取

数据的采集必须准确全面,还要充分考虑距离、风速、风向、温度、气压和湿度等有关影响因素。因此,本文综合考虑这些因素进行数据集选取,实验所用数据来自中国西南某地区风电场,其提供分辨率为 15 min 的风电功率数据。本文选取 2017 一整年的实测历史数据进行实验验证,以 15 min 为间隔、20 组数据为单位来提取特征,对下一时刻点的风电输出功率值进行预测,所使用的原始

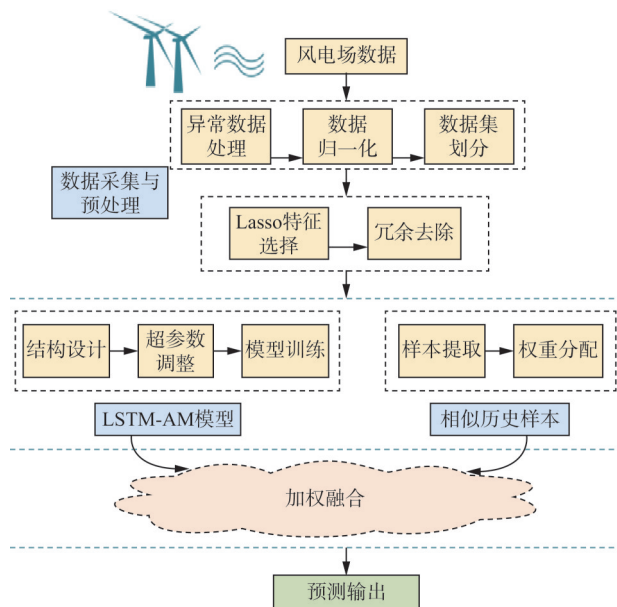


图 4 提出的风电功率预测方法原理图

Fig. 4 Schematic diagram of proposed wind power prediction method

数据样本如图 5 所示。

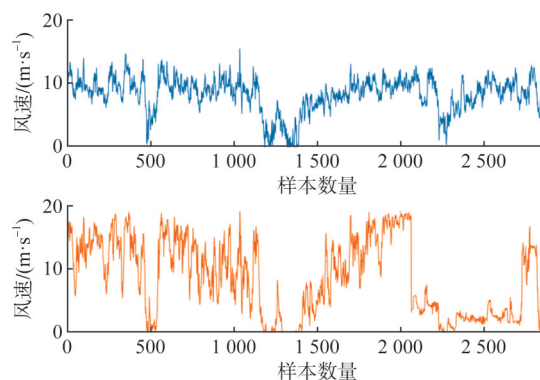


图 5 原始数据集样本

Fig. 5 Samples of raw data set

3.2 数据预处理

原始数据集存在部分异常值或缺失信息,这将严重破坏数据质量,降低模型性能与预测效果,进而影响实验结论与方法评估。为此,本文在开展算例实验前对有关数据进行了预处理。

数据缺失值一般包括两种类型,一种为单个数据缺失,本文通过 Lagrange 插值来生成缺失信息;另一种为整个时刻数据缺失,针对这类缺失,本文采取分段拟合方法对单个或多个时刻的行数据进行拟合重构。而对于原始数据中存在的异常数据信息,本文采用滤波将其剔除,再使用相应方法视作

缺失信息处理。

为缩短训练时间和提高拟合精度,本文对原始数据采取归一化处理,以避免因输入数据单位量纲不同而造成收敛速度慢和训练误差大等问题。处理后的训练样本转换至 $[0, 1]$ 区间,具体如下。

$$x_i^* = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (14)$$

式中: x_i^* 、 x_i 、 $x_{\max}$ 、 $x_{\min}$ 分别为处理后的数据、原始数据、样本集的最大值和最小值。

在得到模型最终输出后,还需将其反归一化为真实量纲,公式如式(15)所示。

$$x_i = x_i^* \cdot (x_{\max} - x_{\min}) + x_{\min} \quad (15)$$

3.3 实验评价指标

本实验采用决定系数(R^2)、均方根误差(root mean square error, RMSE, 记作 e_{RMSE})和绝对误差(mean absolute error, MAE, 记作 e_{MAE})作为指标来评价模型的预测性能,以反映预测模型的误差水平和拟合效果,它的值介于 $[0, 1]$,越大则模型效果越好。3个指标的具体计算公式分别为:

$$R^2 = 1 - \frac{\sum_{i=1}^{n_{\text{test}}} (y_i - \hat{y}_j)^2}{\sum_{i=1}^{n_{\text{test}}} (y_i - \bar{y})^2} \quad (16)$$

$$e_{\text{RMSE}} = \sqrt{\frac{1}{n_{\text{test}}} \sum_{i=1}^{n_{\text{test}}} (y_i - \hat{y}_j)^2} \quad (17)$$

$$e_{\text{MAE}} = \frac{1}{n_{\text{test}}} \sum_{i=1}^{n_{\text{test}}} |y_i - \hat{y}_j| \quad (18)$$

式中: n_{test} 为测试样本数量; y_i 和 $\hat{y}_j$ 分别为真实值与预测值; $\bar{y}$ 为样本平均值。

4 实验与分析

4.1 超参数调优

在进行广泛测试后,发现最佳学习率在0.001~0.01之间。采用控制变量法进行实验,不同学习率下的预测效果如表1所示。

图6—7分别展示了不同学习率下的决定系数和损失迭代情况,综合结果表明学习率为0.005时验证集效果达到最佳,为此将其设置为最终学习率。

初次训练时设定迭代次数为500轮,实验结果表明在200轮以后训练集和验证集损失不再下降,如图8所示。因此将训练的迭代次数设定为200,

表1 不同学习率下的测试集 R^2 值

Tab. 1 R^2 of test set with different learning rates

训练/测试	隐藏单元规模	学习率	决定系数 R^2
8:2	32	0.001	0.892
8:2	32	0.002	0.923
8:2	32	0.003	0.927
8:2	32	0.004	0.922
8:2	32	0.005	0.934
8:2	32	0.006	0.930
8:2	32	0.007	0.919
8:2	32	0.008	0.921
8:2	32	0.009	0.916

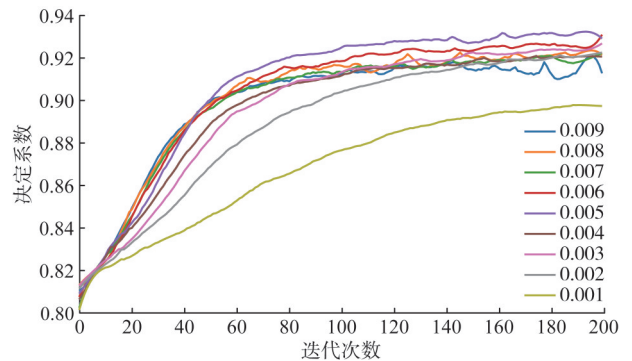


图6 不同学习率下的决定系数 R^2 迭代情况

Fig. 6 Determination coefficient R^2 iteration under different learning rates

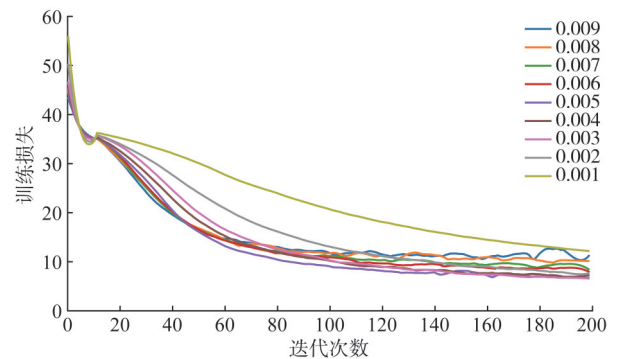


图7 不同学习率下的损失迭代情况

Fig. 7 Iterations of losses under different learning rates

以加快模型训练速度。

4.2 网络结构及AM处理

设置学习率为0.005、迭代次数为200轮后,开展实验探索网络深度和隐藏单元大小对训练结果的影响,如图9所示。结果表明,网络深度及隐藏单元大小并未对预测效果产生较大影响。如表2所示,综合比较不同隐藏单元及网络深度下的预测性

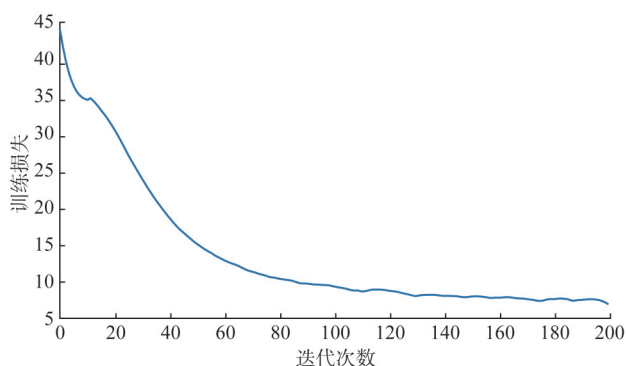


图8 模型训练下的损失迭代情况

Fig. 8 Iterations of losses under model trainings

能,并结合训练速度与预测效果,选定隐藏单元规模为64,网络深度为1层。

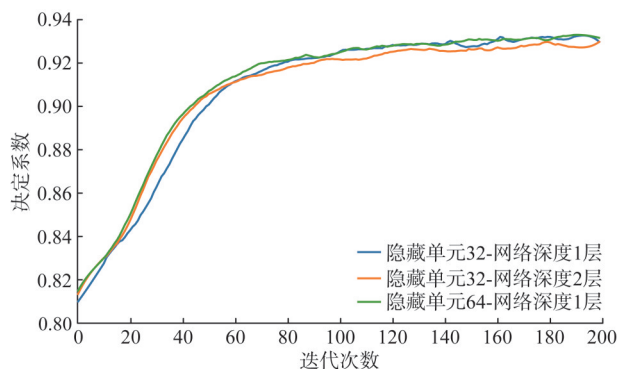
图9 不同网络结构下的决定系数 R^2 Fig. 9 Coefficients of determination R^2 for different network structures

表2 不同隐藏单元及网络深度下的预测性能比较

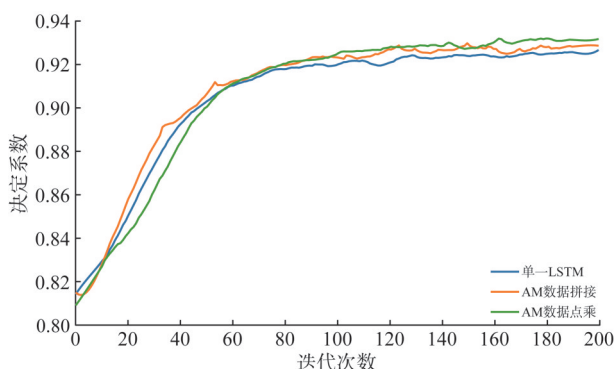
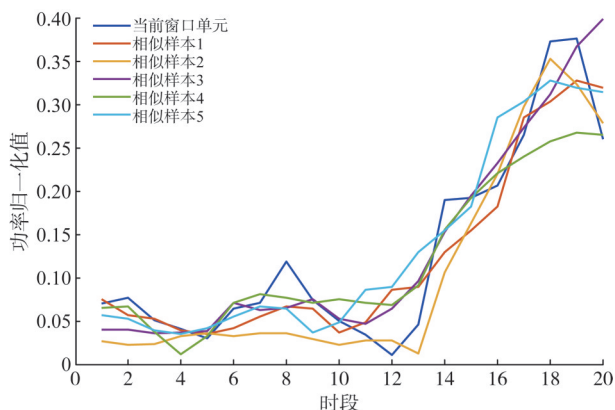
Tab. 2 Comparison of prediction performances under different hidden cells and network depths

隐藏单元规模	网络深度	预测性能
32	1	较好
32	2	较差
64	1	最好

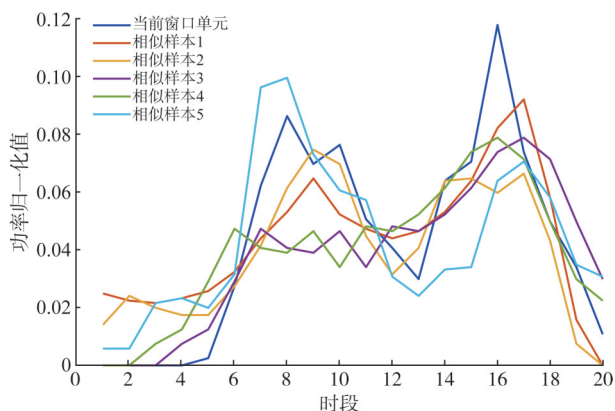
加入注意力机制后模型在训练过程中损失下降更快,学习能力更强。如图10所示,验证集上的预测效果因不同的处理方式而不同。结果表明,将AM的结果作为权重与LSTM输出结果进行点乘后效果最好。

4.3 相似样本提取

在测试集中,采用前文所述方法提取与当前窗口单元相似的历史样本,部分相似样本与原始窗口单元比较示例如图11所示。

图10 不同AM处理方式下的决定系数 R^2 Fig. 10 Coefficients of determination R^2 under different AM processing modes

(a) 相似样本提取示例1



(b) 相似样本提取示例2

图11 滑动窗口单元与相似样本

Fig. 10 Sliding window units and similar samples

结果表明,本文采取基于欧式距离的相似样本提取方法能够在不同时段滑动窗口选取相似度较高的历史样本,所提取的样本与当前窗口单元具有类似的变化特性。

4.4 预测结果

本文使用前5 h的数据预测下一时刻风电功率

值。为验证所提方法的预测性能, 引入 LSTM、CNN-LSTM、Transformer 和时空卷积网络 (time convolutional network, TCN) 等具有代表性的深度学习算法, 以及业内学者提出的熵权综合关联度数据挖掘方法^[31] (entropy weight comprehensive association data mining, EWCADM)、改进灰狼算法-Informer 模型^[32] (modified grey wolf optimizer-informer, MGWO-Informer) 与本文方法进行比较。此外, 将所提方法与加入注意力机制的 LSTM-AM、融合相似样本的 LSTM-AM-SS、进行特征优选的 Lasso-LSTM-AM 等进行比较, 验证了本文方法的优势。对比算法中各层模型主要参数设置如表 3 所示。

采用不同模型在所选数据集中进行短期风电功率预测, 对比结果如表 4 所示。

从表 4 可以看出, 本文所提出的基于特征优选与相似样本融合的 LSTM-AM 风电功率预测方法性能优于其他模型。根据表 3 的结果, 单一 LSTM 模型具有良好的预测效果, 但由于未进行输入特征优选和关键信息捕捉能力强化, 精度相对较低。CNN-LSTM 通过卷积从局部特征中提取关键信息, 但其对历史规律和内在特性学习不足, 因此优势有限。Transformer 擅长捕捉复杂数据集的全局信息, 但在处理序列数据时缺乏对局部特征的专注。TCN 能够更好地捕捉序列数据中远距离依赖关系, 但对局部信息的捕捉不够灵活。EWCADM 能够有效降低数据冗余度, 但未充分捕获时序数据的依赖关系与历史模式。MGWO-Informer 显著提高了算法的全局搜索精度, 但未对数据特征提取能力进行优

化。LSTM-AM 引入注意力机制, 使得模型性能得到显著提升, 但未有效降低输入特征冗余。LSTM-AM-SS 和 Lasso-LSTM-AM 分别进行了相似历史样本融合、冗余特征收缩, 预测精度比 LSTM-AM 有所提升。本文所提方法在 LSTM-AM 基础上引入特征优选并结合相似历史样本, 使得模型预测性能显著提高, R^2 值达到 0.971 3, 且 e_{RMSE} 和 e_{MAE} 均为最低, 体现出在风电功率预测任务中的优越性。

表 4 不同模型的预测性能比较

Tab. 4 Comparison of prediction performance under different models

模型	R^2	$e_{\text{RMSE}}/\text{MW}$	e_{MAE}/MW
LSTM	0.926 7	6.327 8	4.812 7
CNN-LSTM	0.943 3	5.571 5	4.280 3
Transformer	0.958 4	4.769 7	3.649 6
TCN	0.959 1	4.735 1	3.598 5
EWCADM ^[31]	0.968 5	4.174 3	3.182 6
MGWO-Informer ^[32]	0.970 2	4.062 7	3.094 2
LSTM-AM	0.964 6	4.403 1	3.395 4
LSTM-AM-SS	0.967 2	4.342 6	3.314 9
Lasso-LSTM-AM	0.968 7	4.258 3	3.250 6
本文方法	0.971 3	4.095 9	3.125 8

如图 12 所示, 加入注意力机制后 LSTM 模型的性能得到显著增强。

如图 13 所示, 输入特征优选、融合相似历史样本后的 LSTM-AM 模型预测结果与真实值偏差较小, 表明本文采用的方法具有良好的预测性能。

实验结果表明, 本文所提方法在短期风电功率预测任务中拥有较高的准确性, R^2 值达到 0.97 以

表 3 对比算法的主要参数设置

Tab. 3 Main parameter settings of comparison algorithms

名称	参数	
全局参数	时间步长 timesteps=20 训练周期数 epochs=200 优化器 Optimizer=Adam	损失函数 Lossfunction= R_2 正则化系数 regularizer=0.1
LSTM 主要参数	学习率 l_r =0.005	隐藏单元数 hidden_size=64
CNN 主要参数	卷积核大小 kernel_size=1	过滤器数量 filters=64
AM 主要参数	激活函数 activation=Softmax	Lambda 层 SINGLE_ATTENTION_VECTOR=False
Transformer 主要参数	键维度 key_dim=64	编码器层数 encoder_layers=3
TCN 主要参数	卷积核大小 kernel_size=3	膨胀系数 dilations=[1~2, 4]
EWCADM ^[31] 关联度指标	Pearson 相关系数 Kendall 相关系数	Spearman 相关系数 互信息相关系数
MGWO-Informer ^[32] 主要参数	学习率 l_r =0.008	丢弃率 dropout=0.008

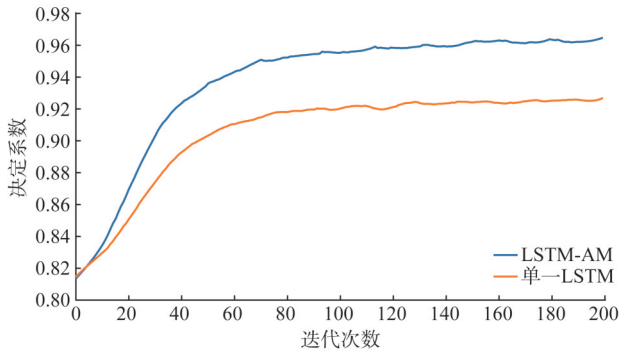


图 12 加入 AM 前后的 LSTM 模型预测效果

Fig. 12 Prediction effects of LSTM model before and after adding AM

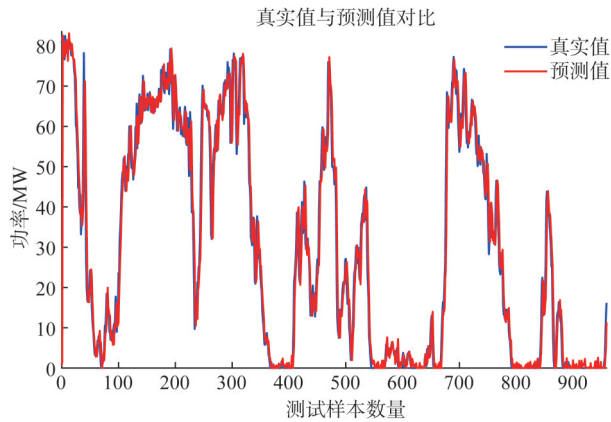


图 13 预测结果与真实值对比

Fig. 13 Comparison of prediction results with real values

上, 相比于传统模型有明显提升。可以看出, 输入特征优选以及融合相似样本后的 LSTM-AM 模型能够更有效地捕捉时间序列数据的长期特性与内在联系, 从而提升预测性能。

4.5 泛化能力验证

为验证本文所提方法的泛化能力, 本节采用国外某风电场 2011 年的数据进行研究^[33]。该场站的风机额定功率为 16 MW, 数据采样颗粒度为 15 min。实验中采取 8:2 的方式划分训练集与测试集。表 5 对该数据集下不同模型的预测结果进行了比较。由结果可知, 本文所提方法的 R^2 值为 0.958 9, e_{RMSE} 和 e_{MAE} 都低于其他模型, 在该数据集下仍表现最优, 证明其具有较好的泛化能力。

5 结语

本文针对现有短期风电功率预测方法泛化性差、精度不足, 以及风电功率数据时序关联强、不确定性高等问题, 采用 Lasso 回归降低模型输入特

表 5 公开数据集下不同模型的预测性能比较

Tab. 5 Comparison of prediction performances of different models under public data sets

模型	R^2	$e_{\text{RMSE}}/\text{MW}$	e_{MAE}/MW
LSTM	0.918 4	6.672 9	4.982 6
CNN-LSTM	0.932 5	6.083 1	4.613 2
Transformer	0.946 3	5.413 2	4.117 2
TCN	0.948 6	5.297 5	4.054 6
EWCAD ^[31]	0.955 7	4.917 3	3.774 9
MGWO-Informer ^[32]	0.957 2	4.834 0	3.682 7
LSTM-AM	0.951 5	5.149 3	3.962 5
LSTM-AM-SS	0.954 2	4.9957	3.826 9
Lasso-LSTM-AM	0.956 1	4.897 7	3.761 4
本文方法	0.958 9	4.724 6	3.633 7

征冗余, 并将相似历史样本与注意力机制融入到 LSTM 网络模型来进行风电功率预测。通过对原始数据进行 Lasso 特征优选, 减少了样本冗余的影响。在此基础上将 AM 加入到 LSTM 网络中用于聚焦重要特征, 提升模型的关键信息提取能力。最后, 结合相似历史样本对输出值进行加权融合, 得到最终预测结果。与单一 LSTM、LSTM-AM 和 Transformer 等模型的对比实验结果表明, 本文提出的基于特征优选与相似历史样本融合的 LSTM-AM 方法在风电功率预测任务中表现更优, 拥有较强的预测性能, 能够为电力系统规划运行提供支撑。

参考文献

- [1] 杨占红, 吕连宏, 曹宝, 等. 国际能源消费比较特征分析及中国发展建议[J]. 地球科学进展, 2016, 3(1): 94–102.
YANG Zhanhong, LÜ Lianhong, CAO Bao, et al. Comparative characteristics of international energy consumption and China's development proposals [J]. Advance in Earth Science, 2016, 3(1): 94–102.
- [2] 朱向东. 目前中国风电弃风现状及对策[J]. 能源与节能, 2012(10): 30–67.
ZHU Xiangdong. Current situation and countermeasures of wind power abandonment in China [J]. Energy & Energy Conservation, 2012(10): 30–67.
- [3] 宋延涛, 范雪峰, 吴庆范, 等. 基于 MMC 的海上风电柔性直流送出控制策略研究[J]. 高压电器, 2023, 59(12): 63–74.
SONG Yantao, FAN Xuefeng, WU Qingfan, et al. Research on flexible DC output control strategy for offshore wind power based on MMC [J]. High Voltage Apparatus, 2023, 59(12): 63–74.
- [4] 黄婧杰, 李金成, 刘科明, 等. 含 CVaR 及增广 ϵ -约束法的多目标光伏储能电站容量优化配置[J]. 南方电网技术, 2023, 17(10): 94–103.
HUANG Jingjie, LI Jincheng, LIU Keming, et al. Optimal allocation of multi-objective photovoltaic energy storage charg-

- ing station capacity with CVaR and augmented ε -constraint method [J]. Southern Power System Technology, 2023, 17(10):94-103.
- [5] 黄碧月, 辛清明, 赵晓斌, 等. 大规模光伏电站孤岛经柔性直流送出的电压特性[J]. 南方电网技术, 2024, 18(3):3-13. HUANG Biye, XIN Qingming, ZHAO Xiaobin, et al. Voltage characteristics of large-scale islanded PV station transmission via VSC-HVDC[J]. Southern Power System Technology, 2024, 18(3):3-13.
- [6] 段学伟. 基于误差分析修正的超短期风速及短期风电功率预测研究[D]. 济南:山东大学, 2016.
- [7] LIU Hui, CHEN Chao, LÜ Xinwei, et al. Deterministic wind energy forecasting: a review of intelligent predictors and auxiliary methods [J]. Energy Conversion and Management, 2019(195):328-345.
- [8] 钱政, 裴岩, 曹利霄, 等. 风电功率预测方法综述[J]. 高电压技术, 2016, 42(4):1047-1060. QIAN Zheng, PEI Yan, CAO Lixiao, et al. Review of wind power forecasting method [J]. High Voltage Engineering, 2016, 42(4):1047-1060.
- [9] 薛禹胜, 郁琛, 赵俊华, 等. 关于短期及超短期风电功率预测的评述[J]. 电力系统自动化, 2015, 39(6):141-151. XUE Yusheng, YU Chen, ZHAO Junhua, et al. A review on short-term and ultra-short-term wind power prediction [J]. Automation of Electric Power Systems, 2015, 39(6):141-151.
- [10] 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 Jing, ZHANG Zhigang, LIANG Zhifeng, et al. Review of wind speed and wind power prediction methods [J]. Journal of Applied Statistics and Management, 2020, 39(6):1045-1059.
- [11] 田波, 朴在林, 王慧. 基于时间序列建模在风力发电功率短期预测中的研究[J]. 电网与清洁能源, 2016, 32(3):115-119, 126. TIAN Bo, PIAO Zailin, WANG Hui. Short term prediction of wind power based on time series modeling [J]. Power System and Clean Energy, 2016, 32(3):115-119, 126.
- [12] LI Lingling, CEN Zeyao, TSENG Minglang, et al. Improving short-term wind power prediction using hybrid improved cuckoo search arithmetic-support vector regression machine [J]. Journal of Cleaner Production, 2021(279):123739. 1-123739. 11.
- [13] ZHOU Yong, LIU Yanfeng, WANG Dengjia, et al. A novel combined multi-task learning and Gaussian process regression model for the prediction of multi-timescale and multi-component of solar radiation [J]. Journal of Cleaner Production, 2021(284):124710. 1-124710. 9.
- [14] 梁超, 刘永前, 周家慷, 等. 基于卷积循环神经网络的风电场内多点位风速预测方法[J]. 电网技术, 2021, 45(2):534-542. LIANG Chao, LIU Yongqian, ZHOU Jiakang, et al. Wind speed prediction at multi-locations based on combination of recurrent and convolutional neural networks [J]. Power System Technology, 2021, 45(2):534-542.
- [15] HU Shuai, XIANG Yue, HUO Da, et al. An improved deep belief network based hybrid forecasting method for wind power [J]. Energy, 2021(224):120185. 1-120185. 8.
- [16] 孙荣富, 张涛, 和青, 等. 风电功率预测关键技术及应用综述 [J]. 高电压技术, 2021, 47(4):1129-1143. SUN Rongfu, ZHANG Tao, HE Qing, et al. Review on key technologies and applications in wind power forecasting [J]. High Voltage Engineering, 2021, 47(4):1129-1143.
- [17] 师洪涛, 杨静玲, 丁茂生, 等. 基于小波-BP神经网络的短期风电功率预测方法[J]. 电力系统自动化, 2011, 35(16):44-48. SHI Hongtao, YANG Jingling, DING Maosheng, et al. A short-term wind power prediction method based on wavelet decomposition and BP neural network [J]. Automation of Electric Power Systems, 2011, 35(16):44-48.
- [18] 周海, 刘建峰, 周健, 等. 基于改进降噪自编码器和多元时序聚类的海上风电功率预测[J]. 太阳能学报, 2023, 44(3):129-138. ZHOU Hai, LIU Jianfeng, ZHOU Jian, et al. Offshore wind power prediction based on improved denoising autoencoder and multivariate time series clustering [J]. Acta Energetica Sinica, 2023, 44(3):129-138.
- [19] 赵建利, 白格平, 李英俊, 等. 基于CNN-LSTM的短期风电功率预测[J]. 自动化仪表, 2020, 41(5):37-41. ZHAO Jianli, BAI Geping, LI Yingjun, et al. Short-term wind power Prediction based on CNN-LSTM [J]. Automatic Instrumentation, 2019, 41(5):37-41.
- [20] 孙改平, 蒋传文. 基于两级级联聚类的神经网络风电功率预测[J]. 太阳能学报, 2021, 42(3):56-62. SUN Gaiping, JIANG Chuanwen. Wind power prediction by neural network based on two-level cascade clustering [J]. Acta Solar Energy Sinica, 2021, 42(3):56-62.
- [21] KAVASSERI R G, SEETHARAMAN K. Day-ahead wind speed forecasting using F-ARIMA models [J]. Renewable energy, 2009, 34(5):1388-1393.
- [22] ZENG J, QIAO W. Short-term wind power prediction using a wavelet support vector machine [J]. IEEE transactions on sustainable energy, 2012, 3(2):255-264.
- [23] MABEL M C, FERNANDEZ E. Analysis of wind power generation and prediction using ANN: a case study [J]. Renewable energy, 2008, 33(5):986-992.
- [24] SAFARI N, CHUNG C Y, PRICE G C D. Novel multi-step short-term wind power prediction framework based on chaotic time series analysis and singular spectrum analysis [J]. IEEE Transactions on Power Systems, 2018, 33(1):590-601.
- [25] VISHNUTHEERTH E P, VIJAY V, SATHEESH R, et al. A comprehensive approach to wind power forecasting using advanced hybrid neural networks [J]. IEEE Access, 2024(12):124790-124800.
- [26] XU Q, HE D, ZHANG N, et al. A short-term wind power forecasting approach with adjustment of numerical weather prediction input by data mining [J]. IEEE Transactions on sustainable energy, 2015, 6(4):1283-1291.
- [27] 叶瑞丽, 郭志忠, 刘瑞叶, 等. 基于小波包分解和改进Elman神经网络的风电场风速和风电功率预测[J]. 电工技术学报, 2017, 32(21):107-115. YE Ruili, GUO Zhizhong, LIU Ruiye, et al. Prediction of wind farm wind speed and wind power based on wavelet packet decomposition and improved Elman neural network [J]. Transactions of China Electrotechnical Society, 2017, 32(21):107-115.
- [28] 石立贤, 金怀平, 杨彪, 等. 基于局部学习和多目标优化的选择性异质集成超短期风电功率预测方法[J]. 电网技术,

2022, 46(2):568 – 577.

SHI Lixian, JIN Huaiping, YANG Biao, et al. Selective Heterogeneous integration ultra-short term wind power prediction method based on local learning and multi-objective optimization [J]. Power System Technology, 2022, 46(2): 568 – 577.

- [29] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735 – 1780.
- [30] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate [C]//2015 International Conference on Learning Representations, May 7 – 9, 2015, San Diego, CA, USA. San Diego: ICLR, 2015: 112 – 131.
- [31] 杨茂, 马剑, 李大勇, 等. 基于熵权综合关联度数据挖掘的风电功率实时预测研究[J]. 太阳能学报, 2019, 40(12):3578 – 3585.

- [32] 王颀, 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.

- [33] DRAXL C, CLIFTON A, HODGE B M, et al. The wind integration national dataset (wind) toolkit[J]. Applied Energy, 2015(151): 355 – 366.

收稿日期: 2025-03-01; 网络首发日期:2025-06-09

作者简介:

吴琛 (1974), 女, 高级工程师 (教授级), 硕士, 从事电力系统规划与运行工作;

崔秋实 (1986), 男, 通信作者, 副教授, 博士, 研究方向为电网大数据与人工智能, qcui@cqu.edu.cn;

谢一工 (1972), 男, 高级工程师, 硕士, 从事电力系统及自动化工作。

(上接第 151 页 Continued from Page 151)

- [19] 贺才郡, 李开成, 杨王旺, 等. 基于双通道 GAF 和深度残差网络的电能质量复合扰动识别[J]. 电网技术, 2023, 47(1): 369 – 379.

HE Caijun, LI Kaicheng, YANG Wangwang, et al. Power quality compound disturbance identification based on dual channel GAF and depth residual network [J]. Power System Technology, 2023, 47(1): 369 – 379.

- [20] 郑炜, 林瑞全, 王俊, 等. 基于 GAF 与卷积神经网络的电能质量扰动分类[J]. 电力系统保护与控制, 2021, 49(11): 97 – 104.

ZHENG Wei, LIN Ruiquan, WANG Jun, et al. Power quality disturbance classification based on GAF and a convolutional neural network [J]. Power System Protection and Control, 2021, 49(11):97 – 104.

- [21] 李贝奥, 李开成, 肖贤贵, 等. 基于多尺度卷积融合时间序列 Transformer 的复合电能质量扰动识别[J/OL]. 电网技术, 1 – 12[2024 – 06 – 12].<https://doi.org/10.13335/j.1000-3673.pst.2023.1764>.

LI Beiao, LI Kaichen, XIAO Xiangui, et al. Compound power quality disturbances identification based on multi-scale fusion time series transformer[J]. Power System Technology, 1 – 12 [2024 – 06 – 12]. <https://doi.org/10.13335/j.1000-3673.pst.2023.1764>.

- [22] CUI C, DUAN Y, HU H, et al. Detection and classification of multiple power quality disturbances using Stockwell transform and deep learning [J]. IEEE Transactions on Instrumentation and Measurement, 2022(71): 1 – 12.

- [23] MISHRA M. Power quality disturbance detection and classification using signal processing and soft computing techniques: a comprehensive review [J]. International Transactions on Electrical Energy Systems, 2019, 29(8): 12008.1 – 12008.9.

- [24] 王飞, 王立辉, 周少武, 等. 基于双分支并联的特征融合电能质量扰动分类方法[J]. 电力系统保护与控制, 2024, 52(5): 178 – 187.

WANG Fei, WANG Lihui, ZHOU Shaowu, et al. A dual-branch parallel-based feature fusion approach for power quality disturbance classification [J]. Power System Protection and Control, 2024, 52(5): 178 – 187.

- [25] 覃日升, 况华, 于辉, 等. 自适应 VMD 与新型 K-M 互卷积窗的非稳态谐波参数检测方法[J]. 南方电网技术, 2024, 18(8): 70 – 79.

QIN Risheng, KUANG Hua, YU Hui, et al. Non-stationary harmonic parameter detection method based on adaptive VMD and novel K-M convolution window [J]. Southern Power System Technology, 2024, 18(8): 70 – 79.

收稿日期: 2024-01-05; 网络首发日期:2024-08-20

作者简介:

刘禹彤 (1998), 女, 硕士研究生, 研究方向为电能质量监测, lytnineeight123@163.com;

苏小玲 (1986) 女, 副教授, 博士, 研究方向为电力电子化电力系统建模及稳定性分析, m18287284953@163.com;

刘可 (1988), 男, 高级工程师, 硕士, 从事电压管理与电能质量监测研究, 397553383@qq.com。