

基于时间模式注意力机制的 BiLSTM 多风电机组超短期功率预测

王渝红¹, 史云翔¹, 周旭¹, 曾琦¹, 方飏², 毕悦²

(1. 四川大学电气工程学院, 成都 610065; 2. 国网四川综合能源服务有限公司, 成都 610031)

摘要: 针对现有预测方法难以批量处理多风机间不同特征的问题, 提出了基于时间模式注意力(temporal pattern attention, TPA)机制的双向长短时记忆(bidirectional long short-term memory, BiLSTM)网络多风电机组超短期功率预测方法。首先, 基于集合经验模态分解(ensemble empirical mode decomposition, EEMD)获得风机原始功率信号的不同模态分量, 以降低神经网络预测难度。其次, 基于 TPA 机制, 从 BiLSTM 网络得到的隐藏行向量中提取多风机之间的复杂联系, 从而使得具有不同特征的模态可以从不同时间步选择相关信息, 进而降低各模态的预测误差。最后, 将 TPA 机制与传统注意力机制应用于分散分布的 14 台风机区域功率预测任务。研究结果表明: 基于本方法的多风电机组超短期功率预测的标准均方根误差仅为 0.054 6, 证明 TPA 机制能有效提高多风电机组的超短期功率预测精度。

关键词: 超短期风电功率预测; 多风电机组; 时间模式注意力机制; 双向长短时记忆; 集合经验模态分解

Ultra-short-term Power Prediction for BiLSTM Multi Wind Turbines Based on Temporal Pattern Attention

WANG Yuhong¹, SHI Yunxiang¹, ZHOU Xu¹, ZENG Qi¹, FANG Biao², BI Yue²

(1. College of Electrical Engineering, Sichuan University, Chengdu 610065, China;

2. State Grid Sichuan Comprehensive Energy Service Co., Ltd., Chengdu 610031, China)

Abstract: Aiming at to solve the problem that the existing prediction methods are difficult to process the different characteristics of multiple wind turbines in batches, we put forward an ultra-short-term power prediction method for multi wind turbines based on bidirectional long short-term memory (BiLSTM) network and temporal pattern attention (TPA) mechanism. Firstly, based on ensemble empirical mode decomposition (EEMD), different mode components of the original power signal of the wind turbine are obtained to reduce the difficulty of neural network prediction. Secondly, based on the TPA mechanism, the complex connection between multiple turbines is extracted from the hidden row vectors obtained from the BiLSTM network, so that the modes with different characteristics can select relevant information from different time steps, thereby reducing the prediction error of each mode. Finally, the TPA mechanism and the traditional attention mechanism are applied to the distributed power prediction task of 14 wind turbines. The results show that the normalized root mean squared error of the ultra-short-term power prediction for multi-wind turbines based on this method is only 0.054 6, which proves that the TPA mechanism can effectively improve the ultra-short-term power prediction accuracy for multi-wind turbines.

Key words: ultra-short-term wind power prediction; multi-wind turbines; temporal pattern attention (TPA); bidirectional long short-term memory; ensemble empirical mode decomposition (EEMD)

0 引言

十四五规划提出“双碳”目标之后, 加快构建清洁低碳、安全高效的能源体系已成为能源电力行业的共识^[1]。然而, 风电在发电过程中具有随机性、

波动性、间歇性等特点, 且目前储能成本仍然较高, 难以大量存储电能, 导致弃风现象频发。因此, 准确的系统多风电机组功率预测成为促进新能源消纳、维护电网安全稳定运行的有效手段^[2]。

目前, 风电机组功率预测对象大多针对单台风电机组或集中分布的风电集群, 而“双碳”背景下的新型电力系统接入了多台风电机组, 各风电机组

基金资助项目: 四川省科技计划(2021YFG0026)。
Project supported by Science and Technology Program of Sichuan Province (2021YFG0026).

特征各异, 因此难以通过简单的场景迁移完成多风电机组功率的准确预测^[3-4]。文献[5]基于动态 R 藤 Copula 模型拟合区域风电集群功率特性, 简化了高维联合概率分布问题。文献[6]利用 K-means 算法将不同风电场的集群划分为趋势相近的子集群, 再利用双向长短时记忆(bidirectional long short-term memory, BiLSTM)神经网络对各子集群进行风电功率预测。文献[7]通过拟合风速-功率曲线, 利用模糊 C 均值聚类方法对功率输出特性类似的风机进行分类, 降低了多风电机组功率预测的复杂度。当小区域内多风机具有相似功率特性时, 聚类方法有较好的效果^[8]; 而当多风机的地理分布差异较大, 各风机间存在着复杂的动态相关性时, 聚类方法则难以捕捉和分析多风电场的输出功率特征。文献[9]基于注意力机制, 为门控循环(gated recurrent unit, GRU)神经网络隐含状态赋予不同的权重, 以提高重要信息的影响。文献[10-11]利用时序注意力机制提取时空关联特征, 提高了模型对时序相关性的依赖。然而, 传统的注意力机制仅能针对一台风机为输出功率分配其在时序上的权重, 不能提取多风机之间的非线性复杂关系。文献[12]提出了时间模式注意力(temporal pattern attention, TPA)机制, 通过对隐藏状态的特征进行卷积运算, 可以提取多风机之间的复杂内在联系; 同时, 注意到不同风机具有不同的特性, TPA 机制可为其在不同时间步上进行差异化权重分配。因此, 本文将引入 TPA 机制, 通过挖掘多风电机组之间的复杂内在联系, 提升功率预测的准确度。

人工智能技术具有强大的自主学习能力, 其在线性问题上求解方面的性能较高, 已被广泛地应用于风电功率预测中。早期的风电功率预测采用人工神经网络(artificial neural network, ANN)以及支持向量机(support vector machine, SVM)^[13-14]。ANN 具有较好的非线性处理能力, 但其计算量会随着维度迅速增长, 且容易陷入局部最优解^[15]。而 SVM 虽然可以获得全局最优解, 但其依赖使用者对核函数等参数的选取经验。随着循环神经网络(recurrent neural network, RNN)^[16]、长短时记忆(long short-term memory, LSTM)神经网络^[17]、GRU^[18]等针对序列数据预测的神经网络的出现, 风电功率预测准确度得到了进一步提升。文献[19]考虑到风电功率序列对历史时序的依赖, 采用 LSTM 对风电功率进行超短期功率预测, 效果优于 ANN 及 SVM。

也有学者采用启发式算法优化 LSTM 神经网络参数以进一步提高预测精度^[20-21]。此外, BiLSTM 通过利用时序数据的正反向信息来挖掘时序数据的前后联系, 该方法被证明能有效提高风电功率的预测精度^[6,22]。因此, 本文将基于 BiLSTM 神经网络对区域多风电机组进行超短期功率预测。

本文提出基于 TPA 机制的 BiLSTM 神经网络对多风电机组进行超短期功率预测。首先, 考虑到风电功率原始信号的波动性与随机性, 引入集合经验模态分解(ensemble empirical mode decomposition, EEMD)来提取固有模式函数(intrinsic mode function, IMF), 以降低神经网络的预测难度。其次, 采用双向 LSTM 提取风电功率数据的正反向特征, 并结合 TPA 机制进一步提取多风机间 IMF 时序数据的复杂联系, 从而实现不同时间序列特征的多风电机组功率预测。最后, 通过算例对比分析了不同模型在多风电机组超短期功率预测任务中的表现, 以验证本文所提方法的有效性。

1 基于时间模式注意力机制的 BiLSTM 模型原理

1.1 EEMD 原理

风机在发电过程中存在较强的波动性与随机性, 同时在时序上存在一定的周期性。而 EEMD 是一种自适应的时频局部化分析方法, 可以将各种类型的信号数据分解为多个不同频段的模态, 且无需设置基函数^[23]。

EEMD 的目标是将信号分解为有限个 IMF 以及频率接近于 0 的剩余函数。每个分解量 IMF 的特征为: (1)极值点和过零点数目相等, 或最多相差一个; (2)局部最大和局部最小的包络线均值为 0。EEMD 的步骤如下:

首先, 确定信号的极大值包络线 $e_{\max}(t)$ 以及极小值包络线 $e_{\min}(t)$, 并求出上下包络线的平均值 $m(t)$ 。

$$m(t) = \frac{e_{\max}(t) + e_{\min}(t)}{2} \quad (1)$$

其次, 从信号 $x(t)$ 中减去 $m(t)$, 得到 $l(t)$ 。

$$l(t) = x(t) - m(t) \quad (2)$$

然后, 判断 $l(t)$ 是否是 IMF, 若不是, 则重复上述步骤直到 $l(t)$ 满足判据为止, 此时的 $l(t)$ 即为所需提取的 IMF, 记为 $C_i(t)$ 。

最后, 每得到一阶 IMF, 就从原信号中扣除,

重复以上步骤直至剩余部分 $r_n(t)$ 为单调序列。因此，原始信号可以表示为

$$x(t) = \sum_{i=1}^N C_i(t) + r_n(t) \quad (3)$$

式中， N 表示分解得到的 IMF 个数。

为了解决模态混淆问题，将正态分布的白噪声加到原始信号中作为一个整体进行 EEMD。重复加入不同的白噪声后，将每次得到的 IMF 作集成平均处理后的结果作为最终结果，流程图如图 1 所示。

白噪声在频谱上均匀分布，使得不同时间尺度的信号自动分离到与其相适应的参考尺度上。由于白噪声均值为 0，因此集成均值的计算结果可以直接作为最终结果，也就解决了模态混淆问题。通过 EEMD 处理风电功率原始信号，将其分解为分别体现风机随机性与波动性的不同频段的 IMF，且这些 IMF 都在时序上具有一定的规律性，从而降低了风电功率原始信号的复杂度，也简化了神经网络的预测。

1.2 BiLSTM 原理结构

LSTM 于 1997 年被提出^[17]，用于处理长时间序列问题，典型 LSTM 结构如图 2 所示。

图 2 中， x_t 表示时间序列的当前输入； C_t 表示当前 LSTM 单元的细胞状态，通常只在 LSTM 内部流动，是 LSTM 的内部记忆； h_t 代表当前的编码隐藏状态向量； f_t 表示遗忘信息的程度； i_t 表示输入信息的保留程度； \tilde{C}_t 表示当前状态的处理信息。 o_t 表示输出信息的保留程度； \tanh 表示双曲正切函数。下标 $t-1$ 代表上一个时刻的 LSTM 单元所对应的状态量。

LSTM 单元有 3 道门，分别是遗忘门、输入门和输出门。遗忘门可以忘记一定比例过去信息；输入门将部分当前时刻的输入信息记录进细胞状态；输出门将编码隐藏状态向量和细胞状态有选择性地作为下一个时刻 LSTM 单元的输入。

当前时刻的输出可能不仅与过去的信息有关，而且还与未来的信息有关。但 LSTM 无法编码从后向前的信息，而 BiLSTM 通过将时间序列反向，由正反 LSTM 组成，可以更好地捕获双向序列的影响。BiLSTM 输出表达式为

$$h_t = \text{concat}(h_{ft}, h_{bt}) \quad (4)$$

式中： h_t 表示 BiLSTM 的隐藏状态向量； concat 表示在输出维度进行拼接操作； h_{ft} 、 h_{bt} 分别表示前向和后向 LSTM 的隐藏状态向量。

BiLSTM 结构示意图如图 3 所示。通过将正向序列反向后作为后向 LSTM 的输入，可以同时训练两个神经网络。前向 LSTM 利用过去的信息预测未来的信息，后向 LSTM 利用未来的信息预测过去的信息，输出结果由这两个网络的输出共同决定。BiLSTM 对于同时依赖前后信息的时间序列有着更好的预测效果，因此本文采用 BiLSTM 神经网络结合风电功率进行双向信息预测。

1.3 TPA 机制

注意力机制模仿人脑，更加注重重要信息，而忽略相对无用的信息，已被广泛应用于自然语言处理、图像及语音识别中，近年来也被广泛应用于各

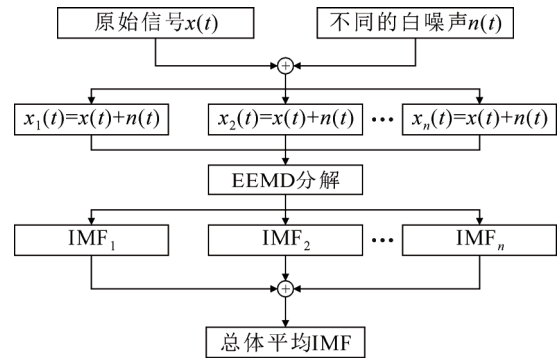


图 1 EEMD 流程图

Fig.1 EEMD process diagram

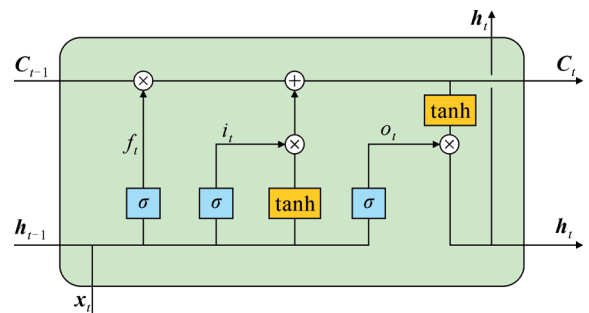


图 2 典型 LSTM 结构

Fig.2 Typical structure of LSTM

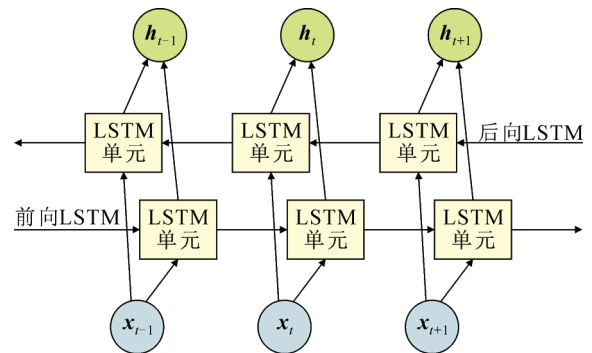


图 3 BiLSTM 结构示意图

Fig.3 Structure of BiLSTM

类预测问题^[9-11,24]。传统注意力机制注重不同时间点的权重分布, 在每个时间步只含有一个变量时有较好的效果。但对于区域内的多风电机组功率预测, 每个时间步都含有多个变量, 各个变量之间可能存在复杂的非线性内在联系, 且每个变量序列都有自己的特征和周期, 难以单独选取某个时间步作为注意重点。而 TPA 则由多个一维 CNN 滤波器从 BiLSTM 隐藏状态行向量抽取特征, 使得模型能够从不同时间步学习多变量之间的互相依赖关系。TPA 结构示意图如图 4 所示。

对原始时间序列用 BiLSTM 处理, 得到的 h_{t-w} — h_t 表示 BiLSTM 对应不同时间输入得到的隐藏状态向量, w 为时间序列长度。定义隐状态矩阵 $H=(h_{t-w}, h_{t-w+1}, \dots, h_{t-1})$, 该隐状态列向量表示同一时间步下 BiLSTM 内部神经元参数构成的变量, 行向量表示单个变量在所有时间步下的状态。

图 4 中隐状态矩阵 H 上的方框表示不同的一维卷积核, 利用一维卷积沿着 H 的 m 个特征卷积, 提取可变信号模型的时间模式矩阵 H^C

$$H_{i,j}^C = \sum_{l=1}^w H_{i,(t-w+1+l)} * C_{j,T-w+l} \quad (5)$$

式中: C_j 表示第 j 个长度为 T 的滤波器; T 表示需要注意的最大长度, 通常可取为 w ; $*$ 表示卷积运算。一维滤波器的卷积核有 k 个, 每个卷积核都沿着隐状态矩阵的行向量卷积。该时间模式矩阵包含着不同序列的复杂内在联系以及时序关系, 是不同序列复杂非线性关系的高维体现。

定义如下注意力机制函数来计算相关性:

$$f(H_i^C, h_t) = (H_i^C)^T W_a h_t \quad (6)$$

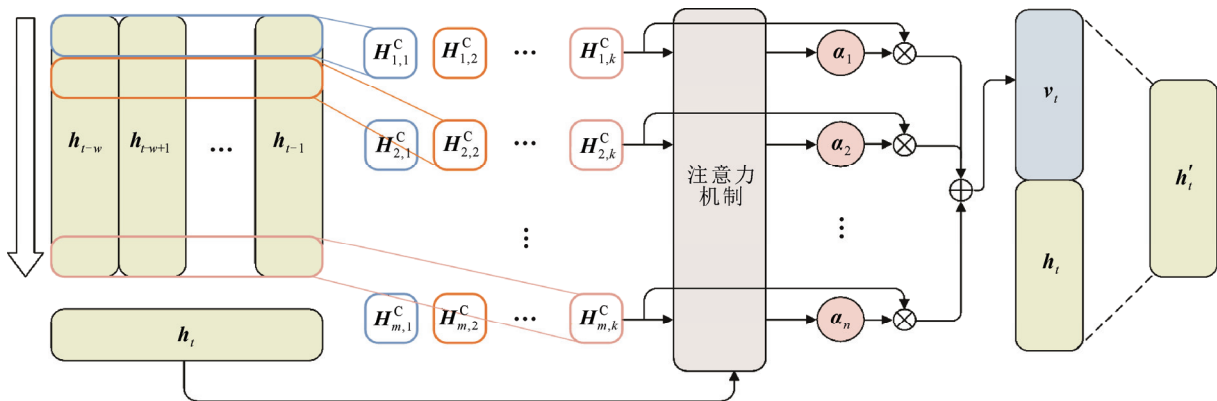


图 4 TPA 结构示意图

Fig.4 Structure of TPA

$$\alpha_i = \sigma(f(H_i^C, h_t)) \quad (7)$$

式中: H_i^C 是 H^C 的行向量; W_a 为 $m \times k$ 的权重矩阵; α_i 为注意力权重; σ 表示 Sigmoid 函数。利用得到的注意力权重 α_i 与 H^C 加权求和, 获得注意力向量 v_t

$$v_t = \sum_{i=1}^n \alpha_i H_i^C \quad (8)$$

式中, n 表示输入变量 x 的特征数。

将 v_t 与 h_t 线性映射后相加获得最终预测值

$$y_{t-1+\Delta} = W_{h'} h'_t = W_{h'} (W_h h_t + W_v v_t) \quad (9)$$

式中: $y_{t-1+\Delta}$ 表示最终预测值; h'_t 为用于生成最终值的中间变量; Δ 表示不同预测任务的预测时间尺度; $W_{h'}$, W_h 和 W_v 为对应变量的不同权重矩阵。

传统注意力机制直接对原始数据的时间序列利用 CNN 进行特征抽取, 只能对单一序列的时间特征进行提取, 无法兼顾不同序列之间的关联性。而 BiLSTM 隐状态矩阵的变量则包含了不同时间步下不同序列之间的复杂关系, 利用 CNN 对隐状态矩阵的行向量进行特征抽取, 能同时提取时间序列关系与不同变量之间的复杂联系。且注意力向量 v_t 是包含时间信息的时间模式矩阵行向量加权和, 因此模型能够从不同时间步选择相关信息。在处理多风机超短期功率预测这类时间步与不同序列存在复杂和非线性的互相依赖问题时, TPA 的先进性能展现出了独特的优势。

2 多风电机组超短期功率预测

风电机组功率是典型的时间序列, 具有连续性与随机性的特征。EEMD 在处理非平稳和非线性数

据时具有优势,且无需选取基函数,所以可以先利用 EEMD 进行特征提取。BiLSTM 网络通过双向信息能够很好地处理时间序列问题,同时兼具较好的鲁棒性。传统的注意力机制在处理单一特征的时间序列时有较好的表现,但对于多台具有不同特征且存在复杂内部联系的风机功率预测效果较差。而利用 TPA 机制可以提取多台风机之间的复杂内部联系,对不同的特征选择相关变量加权,从而更加准确地预测。

2.1 多风电机组超短期功率预测模型及方法

基于时间模式注意力机制的 BiLSTM 多风电机组超短期功率预测模型如图 5 所示。图 5 中每层具体步骤如下:

1) 输入层数据为时间序列长度为 w 的不同分段风机的历史功率数据。为了使得神经网络易于训练,同时消除量纲的影响,对各风电机组功率作如下标准化处理

$$P'_i = \frac{P_i - P_{i\min}}{P_{i\max} - P_{i\min}} \quad (10)$$

式中: P'_i 为第 i 台风电机组标准化后的功率输出; P_i 为第 i 台风电机组的实际功率输出; $P_{i\max}$ 为第 i 台风电机组的最大功率输出; $P_{i\min}$ 为第 i 台风电机组的最小功率输出。将预处理后的数据集按一定比例分为训练集和测试集,数据经过 EEMD 后得到各风机的第 j 个 IMF 被送入神经网络 BiLSTM 层。

2) BiLSTM 层将输入层训练集中所有的 IMF 分别送入不同的 BiLSTM 网络,将双向信息流进行拼接后得到的隐状态行向量 $h_{t-w}, \dots, h_{t-1}, h_t$ 送入 TPA 层。

3) TPA 层通过隐藏状态向量获得注意力向量 v_t , 与隐状态行向量分别进行线性映射相加后得到各风机第 j 个模态的预测值 P_j^{pred} 。

4) 输出层将 P_j^{pred} 作为神经网络的输出,并与各风机第 j 个模态的实际值比较得到模态损失,以该损失最小为目标优化各层参数,并将训练代数 i 加 1。若训练代数小于设定值 E ,则继续训练;反之,则输出最优模型,结合测试集测试模型性能,预测多风电机组超短期功率。

2.2 评估指标

均方根误差(root mean squared error, RMSE)以及平均绝对误差(mean absolute error, MAE)的数量级会随着风电功率输出的数量级变化。考虑到多风电机组之间可能会存在额定功率差异,均方根误差

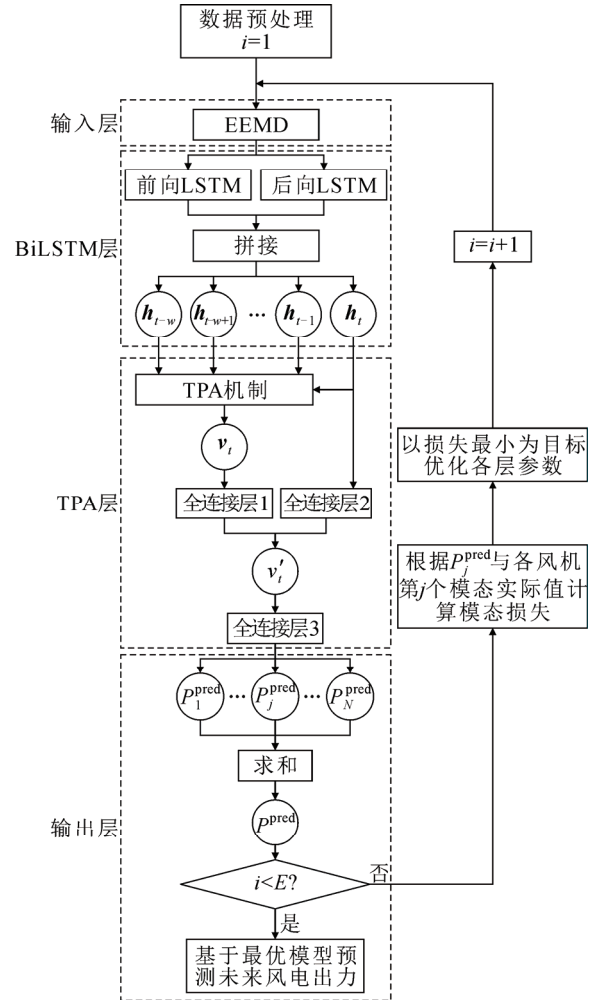


图 5 基于时间模式注意力机制的 BiLSTM 多风电机组超短期功率预测模型

Fig.5 Ultra-short-term power prediction model for multi wind turbines based on TPA of BiLSTM

及平均绝对误差指标难以比较不同风电机组之间的预测精度。标准化平均绝对误差(normalized mean absolute percentage error, NMAPE)由于被归一化,因此可能会出现极大值从而失去意义。故采用标准绝对平均误差(normalized mean absolute error, NMAE)和标准均方根误差(normalized root mean squared error, NRMSE)作为多风电机组功率预测准确性的评价指标,其表达式分别为:

$$e_{\text{NMAE}} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |P_{i,j}^{\text{true}} - P_{i,j}^{\text{pred}}| \quad (11)$$

$$e_{\text{NRMSE}} = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (P_{i,j}^{\text{true}} - P_{i,j}^{\text{pred}})^2} \quad (12)$$

式中: e_{NMAE} 表示所有风机功率预测的标准化绝对平均误差; e_{NRMSE} 表示所有风机功率预测的标准均

方根误差; N 表示测试样本数; M 为风机台数; $P_{i,j}^{true}$ 表示第 i 台风机第 j 个样本的标准化实际功率值; $P_{i,j}^{pred}$ 表示第 i 台风机第 j 个样本的标准化预测功率值。

3 算例分析

3.1 实验设计

本文以某区域分散分布的 14 台风电机组功率数据进行实验, 采样间隔为 15 min。训练集有 4 277 组数据, 测试集有 1 000 组, 每组数据输入不同风电机组各 IMF 的时间序列, 输出为下一个时间点的对应功率值。时间序列过短可能会导致神经网络不能有效利用过去的信息进行学习; 而时间序列过长可能会导致历史信息过多神经网络无法选择重点关注。因此, 综合考虑将时间序列长度设为 24。利用 L2 正则化以及 dropout 层预防过拟合。所有实验均在 NVIDIA GeForce RTX 2080 Ti 上完成。

3.2 EEMD 结果分析

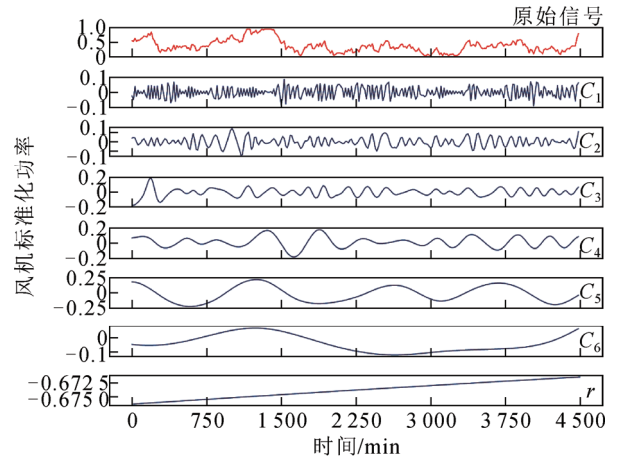
将生成的 IMF 个数设置为 12, 对某一段时间的风机输出功率使用 EEMD。由于模态 7—12 与模态 6 相似, 故此处省略。第 4 台风机与第 10 台风机的 EEMD 部分结果如图 6 所示。

图 6 中, 第一个序列表示风电功率的原始信号, 其余序列表示风电功率 EEMD 后得到的固有模态 $C_1—C_6$ 和剩余函数 r , 图中纵坐标表示风机原始功率经过式(10)变换后的标准化功率。模态 $C_1、C_2$ 频率较高, 变化相对剧烈, 可以在一定程度上反映风电功率的随机性。模态 $C_3、C_4$ 波动性频率相对较低, 但是周期性明显, 能在一定程度上反映风电功率的波动性。其余模态以及剩余函数变化缓慢, 波动性弱, 体现了风机一段时间内的基本发电情况。每个 IMF 都能在一定程度上凸显风电功率的特征, 且低频 IMF 波动性弱, 预测精度很高。因此, 风机功率总预测误差主要由前两个高频模态预测误差所决定, 较少受到多次预测值叠加而导致误差成倍累积。相比于直接采用风机原始功率进行预测, EEMD 能够减少风电数据的非平稳性和波动性对预测精度的影响, 可以降低待预测风电功率数据的复杂度, 从而提升 BiLSTM 神经网络的预测准确度。

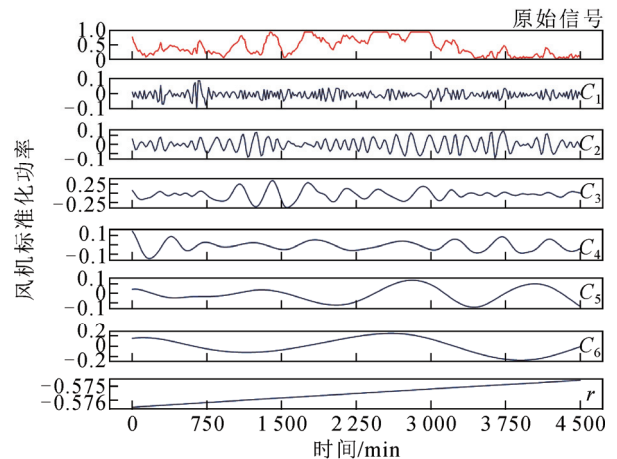
3.3 各风机模态预测结果分析

为了验证 TPA 机制在多风机功率预测任务上的提升, 分别对基于时间模式注意力机制的 TPA-BiLSTM 模型以及基于传统注意力机制的 AM-BiLSTM 模型进行训练。输入为时间序列长度

为 24 的 14 台风机各自的 IMF, 在其余条件相同的情况下各自训练 20 代, 损失函数采用 MSE。模态 3 与模态 6 在传统注意力机制和 TPA 机制下的各风机损失贡献列于表 1。



(a) 第4台风机功率EEMD部分结果



(b) 第10台风机功率EEMD部分结果

图6 风机功率 EEMD 部分结果

Fig.6 EEMD partial results of wind power

表1 模态3与模态6的各风机损失贡献

Table 1 Loss of each wind turbine for mode 3 and mode 6

风机编号	模态 3		模态 6	
	TPA	AM	TPA	AM
1	0.000 81	0.004 06	0.000 17	0.000 47
2	0.001 00	0.003 89	0.000 10	0.000 39
3	0.000 78	0.006 96	0.000 10	0.000 32
4	0.000 88	0.005 16	0.000 06	0.000 38
5	0.000 85	0.005 82	0.000 19	0.000 40
6	0.000 73	0.003 96	0.000 32	0.000 36
7	0.001 03	0.006 08	0.000 18	0.000 56
8	0.000 47	0.005 72	0.000 08	0.000 65
9	0.000 66	0.005 57	0.000 06	0.000 42
10	0.000 69	0.006 26	0.000 08	0.000 55
11	0.000 55	0.003 82	0.000 11	0.001 02
12	0.000 76	0.004 31	0.000 10	0.000 63
13	0.000 57	0.004 56	0.000 10	0.000 38
14	0.000 75	0.004 17	0.000 15	0.000 49
平均值	0.000 75	0.005 02	0.000 13	0.000 50

表 1 的结果表明, TPA 机制下各风机的模态损失值相比于传统注意力机制而言更小。这是由于不同的风机功率序列具有不同的特征, 传统注意力机制对每台风机分配相同的时间注意力权重, 无法为不同的序列回顾配置不同的时间步数。而 TPA 机制通过卷积层对隐藏行状态矩阵进行特征提取, 每个卷积核都包含了不同风机之间的复杂关联信息以及时序上的信息。通过对每个卷积核包含的多个时间步长的信息加权求和得到注意力向量 v_t , 其包含着多风机间的关联以及时序上的权重。因此, TPA 机制可以为每一个风机序列从不同时间步选择相关信息。将 TPA-BiLSTM 模型各模态损失列于表 2。

从表 2 可以看出, 越靠后的模态其频率越低, 模态损失也越低。这是因为频率低的模态变化幅度小, 神经网络更好预测。因此, TPA-BiLSTM 模型的预测误差主要来自前面频率较高的模态, 在神经网络训练的过程中, 可以提高前 5 个模态损失的权重, 从而进一步降低误差, 提高预测精度。

3.4 多风电机组超短期功率预测结果分析

为进一步验证 TPA-BiLSTM 模型的性能, 本文将 ARMA 模型, AM-BiLSTM 模型和基于传统注意力机制的 AM-CNN-GRU 模型^[9]作为本文所提模型对比。其中模型 1—4 分别为: 输入数据经过 EEMD 的 TPA-BiLSTM 模型、TPA-CNN-GRU 模型、AM-BiLSTM 模型以及 AM-CNN-GRU 模型; 模型 5 为输入数据没有经过 EEMD 的 TPA-BiLSTM 模型; 模型 6 为 ARMA 模型。

图 7 为只建立单个模型对第 4 台风机和第 10 台风机单点功率预测的各算法比较。从图 7 中可以看出, 即使风机功率具有较强的波动性, 模型 1、2 的预测曲线也更加接近于真实值, 而模型 3 与模型 4 则仅在部分时段预测结果尚可, 但存在某些与真实值偏差较大的时段, 这是由于传统注意力机制对所有风机分配了相同的时序注意力权重。如果风电机组本身特性符合该权重分配, 那么预测效果就较好。但是, 如果风机自身特性与该权重分配不契合的话, 那么就会产生较大误差, 如图 4(a)所示, 基于传统注意力机制的两条曲线预测都相对真实值偏离较大。TPA 机制则通过一维卷积抽取隐状态矩阵行向量的特征, 提取了风机之间的复杂内在联系, 为时间模式矩阵分配权重的同时, 兼顾了各风机的不同特征, 因此其能在各风机各时段预测上均表现出较好的效果。此外, 从模型 1 与模型 5 的对比可

表 2 基于 TPA-BiLSTM 模型各模态损失

Table 2 Loss of each mode based on the TPA-BiLSTM model

模态	模态损失	模态	模态损失
C ₁	0.001 78	C ₈	0.000 09
C ₂	0.000 88	C ₉	0.000 10
C ₃	0.000 75	C ₁₀	0.000 13
C ₄	0.000 66	C ₁₁	0.000 04
C ₅	0.000 40	C ₁₂	0.000 02
C ₆	0.000 13	Re	0.000 03
C ₇	0.000 14	总损失	0.000 63

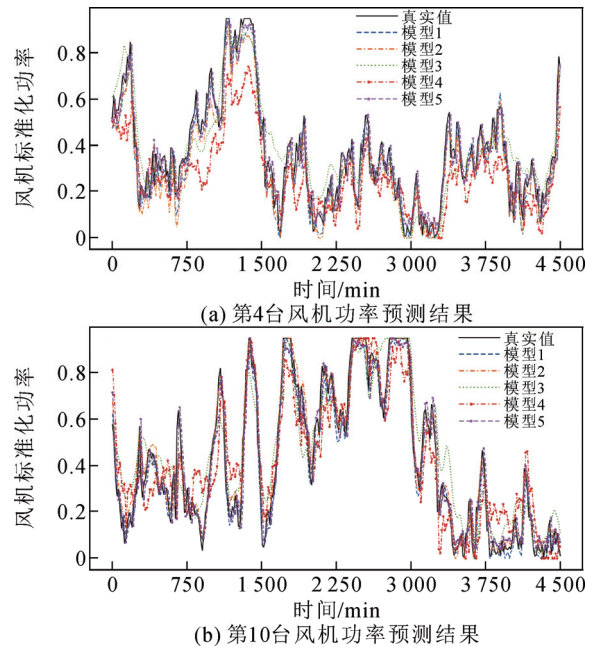


图 7 风机功率预测结果对比

Fig.7 Comparison of prediction results of the wind turbines

以看出, 经过 EEMD 后的模型预测精度更高, 这是因为 EEMD 将原始波动性强的数据分解为多个波动性相对更弱的数据, 降低了神经网络的预测难度, 从而提高了预测精度。

为了进一步比较 TPA-BiLSTM 模型性能, 将各模型不同预测时长的 NMAE、NRMSE 与训练模型所用时间列于表 3。表 3 中“单模型”为建立单个模型对 14 台风机同时进行功率预测; “多模型”则表示建立 14 个模型分别对每台风机功率独立预测。

表 3 中, 当基于单模型进行多风机预测时, 模型 1、2 由于采用了更复杂的注意力机制, 导致神经网络参数数量更大且训练时间更长。但是, 可以看出模型 1、2 在不同预测时长下的 NRMSE 以及 NMAE 都远小于模型 3 以及模型 4, 且模型 1 的预测精度略高于模型 2。这是由于传统注意力机制不能对不同风电功率差异化分配时序权重, 导致各风机预测

表3 不同模型下风电功率预测评估指标值
Table 3 Evaluation index values of wind power prediction under different models

模型类型	预测时长	e_{NMAE}	e_{NRMSE}	耗时/h	
单模型	模型 1	15 min	0.043 0	0.054 6	11.1
		1 h	0.071 2	0.094 5	
		4 h	0.0995	0.124 5	
	模型 2	15 min	0.049 1	0.060 9	10.6
		1 h	0.087 2	0.108 1	
		4 h	0.105 2	0.135 1	
	模型 3	15 min	0.106 4	0.134 5	9.1
		1 h	0.158 7	0.203 8	
		4 h	0.185 9	0.232 6	
	模型 4	15 min	0.098 4	0.124 5	6.6
		1 h	0.129 3	0.175 1	
		4 h	0.179 5	0.226 5	
模型 5	15 min	0.065 5	0.084 3	5.0	
	1 h	0.110 3	0.137 3		
	4 h	0.160 4	0.197 2		
多模型	模型 3	15 min	0.054 1	0.068 9	49.7
		1 h	0.093 6	0.119 4	
		4 h	0.140 6	0.168 0	
	模型 4	15 min	0.044 2	0.056 2	38.1
		1 h	0.079 3	0.101 1	
		4 h	0.122 9	0.144 6	
	模型 6	15 min	0.073 0	0.094 9	3.9
		1 h	0.115 5	0.150 3	
		4 h	0.194 8	0.251 6	

误差都较大。此外, BiLSTM 神经网络由于提取了时序数据双向信息, 因此在 TPA 机制下的预测精度更高。模型 5 由于风机原始功率数据没有经过 EEMD 处理, 导致神经网络的参数量大大降低, 从而训练消耗时间最短, 但各预测时长下的 NMAE 和 NRMSE 都比模型 1 更大。上述结果说明, EEMD 虽然增加了神经网络复杂度, 但能有效提高风电功率的预测精度。当采用另外两种算法分别对每台风机建立神经网络进行单风机功率预测时, 传统注意力机制的优势则有所体现。通过分别对每台风机分配契合各自特性的时序权重, 相比于单模型能有效提高预测精度。但是, 由于多模型方法对多风机分别预测, 忽略了多风机之间的复杂内在联系, 因此误差仍大于 TPA-BiLSTM。多模型方法针对每台风机分布建立模型独立预测, 使得模型参数量激增, 需要耗费大量的训练时间。此外, 传统的 ARMA 模型不仅 NMAE 和 NRMSE 都比神经网络模型大, 而且消耗时间也会随着预测集的增大而变长。而神经网络模型则可以采用训练好的模型快速预测风电功率, 在处理预测多风机功率预测任务时优势更加

明显。综上所述, 与建立多模型分别预测各风机功率相比, TPA-BiLSTM 算法基于单个模型预测多风机功率能够极大程度地降低训练时间, 并在一定程度上提高预测精度。

4 结论

1) 将原始风电功率信号进行 EEMD 后得到具有一定规律性的不同频段 IMF, 降低了神经网络的预测难度。通过 BiLSTM 结合风电功率的正反向信息预测, 充分提取数据的时序信息, 提高了预测精度。

2) TPA 机制可以从 BiLSTM 的隐藏行向量中提取多风机之间的复杂联系, 同时针对具有不同特征的风机从不同时间步选择相关信息, 多风机功率预测的精度高于基于传统注意力机制的模型。

3) TPA-BiLSTM 模型相比于针对单风机建立的模型而言, 不仅训练时间大大降低, 而且在不同时间尺度下均能有效提高多风电机组功率超短期预测的精度, 相比于传统注意力机制, 其在处理多风机功率预测任务上有着更优异的性能。

参考文献 References

- [1] 谭显东, 刘俊, 徐志成, 等. “双碳”目标下“十四五”电力供需形势[J]. 中国电力, 2021, 54(5): 1-6.
TAN Xiandong, LIU Jun, XU Zhicheng, et al. Power supply and demand balance during the 14th five-year plan period under the goal of carbon emission peak and carbon neutrality[J]. Electric Power, 2021, 54(5): 1-6.
- [2] 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.
- [3] 刘燕华, 刘冲, 李伟花, 等. 基于出力模式匹配的风电集群点多时间尺度功率预测[J]. 中国电机工程学报, 2014, 34(25): 4350-4358.
LIU Yanhua, LIU Chong, LI Weihua, et al. Multi-time scale power prediction of wind farm cluster based on profile pattern matching[J]. Proceedings of the CSEE, 2014, 34(25): 4350-4358.
- [4] 王晨, 寇鹏. 基于卷积神经网络和简单循环单元集成模型的风电场多风机风速预测[J]. 电工技术学报, 2020, 35(13): 2723-2735.
WANG Chen, KOU Peng. Wind speed forecasts of multiple wind turbines in a wind farm based on integration model built by convolutional neural network and simple recurrent unit[J]. Transactions of China Electrotechnical Society, 2020, 35(13): 2723-2735.
- [5] 涂青宇, 苗世洪, 林毓军, 等. 基于动态 R 藤 Copula 模型的区域风电集群超短期功率区间预测方法[J]. 高电压技术, 2022, 48(2): 456-470.
TU Qingyu, MIAO Shihong, LIN Yujun, et al. Ultra-short-term interval forecasting method for regional wind farms based on dynamic R-vine Copula model[J]. High Voltage Engineering, 2022, 48(2): 456-470.
- [6] 杨子民, 彭小圣, 郎建勋, 等. 基于集群动态划分与 BLSTM 深度

- 学习的风电集群短期功率预测[J]. 高电压技术, 2021, 47(4): 1195-1203.
- YANG Zimin, PENG Xiaosheng, LANG Jianxun, et al. Short-term wind power prediction based on dynamic cluster division and BLSTM deep learning method[J]. High Voltage Engineering, 2021, 47(4): 1195-1203.
- [7] YANG M, SHI C Y, LIU H Y. Day-ahead wind power forecasting based on the clustering of equivalent power curves[J]. Energy, 2021, 218: 119515.
- [8] 王 勃, 刘 纯, 冯磊磊, 等. 基于集群划分的短期风电功率预测方法[J]. 高电压技术, 2018, 44(4): 1254-1260.
- WANG Bo, LIU Chun, FENG Shuanglei, et al. Prediction method for short-term wind power based on wind farm clusters[J]. High Voltage Engineering, 2018, 44(4): 1254-1260.
- [9] 赵 兵, 王增平, 纪维佳, 等. 基于注意力机制的 CNN-GRU 短期电力负荷预测方法[J]. 电网技术, 2019, 43(12): 4370-4376.
- ZHAO Bing, WANG Zengping, JI Weijia, et al. A short-term power load forecasting method based on attention mechanism of CNN-GRU[J]. Power System Technology, 2019, 43(12): 4370-4376.
- [10] 孟安波, 陈 顺, 王陈恩, 等. 基于混沌 CSO 优化时序注意力 GRU 模型的超短期风电功率预测[J]. 电网技术, 2021, 45(12): 4692-4700.
- MENG Anbo, CHEN Shun, WANG Chen'en, et al. Ultra-short-term wind power prediction based on chaotic CSO optimized temporal attention GRU model[J]. Power System Technology, 2021, 45(12): 4692-4700.
- [11] 李 卓, 叶 林, 戴斌华, 等. 基于 IDSCNN-AM-LSTM 组合神经网络超短期风电功率预测方法[J]. 高电压技术, 2021: 1-13 [2021-10-24]. <https://doi.org/10.13336/j.1003-6520.hve.20210557>.
- LI Zhuo, YE Lin, DAI Binhua, et al. Ultra-short-term wind power prediction method based on IDSCNN-AM-LSTM combination neural network[J]. High Voltage Engineering: 1-13[2021-10-24]. <https://doi.org/10.13336/j.1003-6520.hve.20210557>.
- [12] SHIH S Y, SUN F K, LEE H Y. Temporal pattern attention for multivariate time series forecasting[J]. Machine Learning, 2019, 108(8/9): 1421-1441.
- [13] KARINIOTAKIS G N, STAVRAKAKIS G S, NOGARET E F. Wind power forecasting using advanced neural networks models[J]. IEEE Transactions on Energy Conversion, 1996, 11(4): 762-767.
- [14] 李 智, 韩学山, 韩 力, 等. 地区电网风电场功率超短期预测方法[J]. 电力系统自动化, 2010, 34(7): 90-94.
- LI Zhi, HAN Xueshan, HAN Li, et al. An ultra-short-term wind power forecasting method in regional grids[J]. Automation of Electric Power Systems, 2010, 34(7): 90-94.
- [15] 韩自奋, 景乾明, 张彦凯, 等. 风电预测方法与新趋势综述[J]. 电力系统保护与控制, 2019, 47(24): 178-187.
- HAN Zifen, JING Qianming, ZHANG Yankai, et al. Review of wind power forecasting methods and new trends[J]. Power System Protection and Control, 2019, 47(24): 178-187.
- [16] WILLIAMS R J, ZIPSER D. A learning algorithm for continually running fully recurrent neural networks[J]. Neural Computation, 1989, 1(2): 270-280.
- [17] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [18] CHO K, VAN MERRIËNBOER B, BAHDANAU D, et al. On the properties of neural machine translation: encoder-decoder approaches[C] // Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation. Doha: Association for Computational Linguistics, 2014: 103-111.
- [19] 朱乔木, 李弘毅, 王子琪, 等. 基于长短期记忆网络的风电场发电功率超短期预测[J]. 电网技术, 2017, 41(12): 3797-3802.
- ZHU Qiaomu, LI Hongyi, WANG Ziqi, et al. Short-term wind power forecasting based on LSTM[J]. Power System Technology, 2017, 41(12): 3797-3802.
- [20] SHAHID F, ZAMEER A, MUNEEB M. A novel genetic LSTM model for wind power forecast[J]. Energy, 2021, 223: 120069.
- [21] LIU H, CHEN D H, LIN F, et al. Wind power short-term forecasting based on LSTM neural network with dragonfly algorithm[J]. Journal of Physics: Conference Series, 2021, 1748: 032015.
- [22] 谢小瑜, 周俊煌, 张勇军, 等. 基于 W-BiLSTM 的可再生能源超短期发电功率预测方法[J]. 电力系统自动化, 2021, 45(8): 175-184.
- XIE Xiaoyu, ZHOU Junhuang, ZHANG Yongjun, et al. W-BiLSTM based ultra-short-term generation power prediction method of renewable energy[J]. Automation of Electric Power Systems, 2021, 45(8): 175-184.
- [23] WU Z H, HUANG N E. Ensemble empirical mode decomposition: a noise-assisted data analysis method[J]. Advances in Adaptive Data Analysis, 2009, 1(1): 1-41.
- [24] 王 红, 史金钊, 张志伟. 基于注意力机制的 LSTM 的语义关系抽取[J]. 计算机应用研究, 2018, 35(5): 1417-1420, 1440.
- WANG Hong, SHI Jinchuan, ZHANG Zhiwei. Text semantic relation extraction of LSTM based on attention mechanism[J]. Application Research of Computers, 2018, 35(5): 1417-1420, 1440.



WANG Yuhong
Ph.D., Professor



SHI Yunxiang



ZENG Qi
Ph.D.
Associate professor
Corresponding author

王渝红

1971—, 女, 博士, 教授, 博士生导师
研究方向为高压直流输电
E-mail: yuhongwang@scu.edu.cn

史云翔

1998—, 男, 硕士生
主要从事电力系统自动化、新能源发电技术方面的研究
E-mail: shiyunxiang@stu.scu.edu.cn

曾琦(通信作者)

1977—, 女, 博士, 副教授
研究方向为电力系统稳定和控制、高压直流输电、柔性直流输电
E-mail: zengqi@scu.edu.cn