

基于改进 CNN-BiLSTM-Attention 的两阶段短期风电功率预测



陈艺文^{1,*}, 李昕蓉², 高春瑞¹

¹ 国电和风电开发有限公司, 辽宁沈阳 110000

² 哈尔滨工业大学机电工程学院, 黑龙江哈尔滨 150001

摘要: 为了提升短期风电功率预测的准确性, 本文研究了一种基于逐次变分模态分解 (SVMD) 方法、贝叶斯优化 (BO) 算法和 CNN-BiLSTM-Attention 模型结合的两阶段短期风电功率预测方法。该方法具体原理为: 在第一阶段, 利用 SVMD 方法对经过数据预处理后的风电功率数据进行分解, 将得到的模态分量作为 CNN-BiLSTM-Attention 短期功率预测模型的输入, 然后引入 BO 算法对预测模型的学习率、隐含层节点和正则化参数进行调优, 利用训练好的模型进行预测, 得到初始风电功率预测值; 在第二阶段, 采用预测模型对误差序列进行误差补偿预测, 得到初始误差功率预测数据, 将初始预测功率和误差预测功率求和得到最终的风电功率预测结果。结果表明: 相比于 CNN-BiLSTM-Attention 模型, 改进 CNN-BiLSTM-Attention 模型两阶段预测模型的 MAE 值、MAPE 值和 RMSE 值分别下降了 85.7%、75.2% 和 77.3%, 表明了本文研究的改进 CNN-BiLSTM-Attention 模型的两阶段短期风电功率预测方法的有效性。

关键词: 短期风电功率预测; 逐次变分模态分解; 贝叶斯优化算法; 误差补偿; 预测模型

DOI: [10.57237/j.jsts.2025.02.002](https://doi.org/10.57237/j.jsts.2025.02.002)

Two-Stage Short-Term Wind Power Prediction based on Improved CNN-BiLSTM-Attention

Chen Yiwen^{1,*}, Li Xinrong², Gao Chunrui¹

¹Guodian and Wind Power Development Co., Ltd, Shenyang 110000, China

²School of Mechanical and Electrical Engineering, Harbin Institute of Technology, Harbin 150001, China

Abstract: To enhance the accuracy of short-term wind power prediction, this paper proposes a novel two-stage forecasting framework that integrates Sequential Variational Mode Decomposition (SVMD), Bayesian Optimization (BO), and a CNN-BiLSTM-Attention model. In the first stage, the preprocessed wind power historical data is decomposed into several modal components via SVMD. These components serve as inputs to the CNN-BiLSTM-Attention model, whose hyperparameters—including the learning rate, number of hidden units, and regularization coefficient—are automatically tuned using the BO algorithm. The output of this stage is the initial power prediction. In the second stage, the prediction error sequence from the first stage is analyzed and similarly processed (decomposed and modeled) to generate an error compensation term. The final prediction is obtained by summing the initial power prediction and the predicted error

基金项目: 国家重点研发计划项目 (2017YFB0902100).

*通信作者: 陈艺文, c15831251953@163.com

收稿日期: 2025-10-18; 接受日期: 2025-12-08; 在线出版日期: 2025-12-24

<http://www.wjmst.net>

compensation. Results show that compared with the CNN-BiLSTM-Attention model, the MAE, MAPE and RMSE values of the improved CNN-BiLSTM-Attention two-stage prediction model decreased by 85.7%, 75.2% and 77.3%, respectively, demonstrating the effectiveness of the two-stage short-term wind power prediction method of the improved CNN-BiLSTM-Attention model studied in this paper.

Keywords: Short-term Wind Power Forecasting; Successive Variational Mode Decomposition; Bayesian Optimization Algorithm; Error Compensation; forecasting Model

1 引言

随着风电装机容量的持续增加,其在能源结构中的占比日益凸显。风能作为一种资源丰富的可再生能源,其有效利用对于减少化石燃料的依赖、推动绿色能源发展具有关键作用。然而由于风资源的不稳定性和风力发电的不确定性给风电并网及电网稳定运行带来了严峻挑战[1],通过对风电功率短期的准确预测,电网调度部门可以有效的制定合理的发电计划和调度策略,以缓解电力系统调峰、调频的压力[2]。因此提高对风电功率的预测精度对于优化电网调度、促进新能源消纳具有重要意义。

在风电功率预测领域,当前的研究按照预测原理可以分为三类:基于物理方法、基于统计分析方法和基于学习的方法[3]。基于物理的方法虽然建立在坚实的物理原理之上,但由于气象模型的不确定性以及环境的不稳定性导致预测误差较大,并且该方法计算模型复杂、计算效率较低,一般用来做中长期预测[4]。基于统计分析的方法主要依赖于历史数据的统计特性和模式来进行预测。然而,这种方法对于训练数据的依赖程度较高,当面对新环境或者新条件时,其预测精度可能会受到影响[5]。近年来,基于学习的方法在风电功率预测领域取得了显著成绩[6],这些方法通过机器学习、深度学习等先进技术,能够自动从大量数据中学习并提取出有用的特征,从而实现对风电功率的准确预测,因此适用于短期风电功率预测。

传统机器学习方法由于缺乏长期数据记忆能力,在风电功率预测中表现受限。而深度学习方法可以连接时间序列数据,提取更深层次的特征,在数据预测领域得到了迅速发展[7]。在风速预测领域,长短期记忆神经网络(LSTM)深度学习模型展现出了广泛的应用价值[8]。如文献[9]提出了一种基于最大信息系数与多任务学习结合的LSTM网络风电超短期预测模型,并验证了深度学习模型的性能优于传统神经网络模型。文献[10]利用卷积神经网络(CNN)和LSTM算法分

别采用直接和递归方法预测中期太阳辐照和风速。结果表明,CNN对风速的预测效果优于LSTM。文献[11]针对LSTM模型的缺陷,提出了一种基于CNN-LSTM混合神经网络的规划阶段风电功率预测模型。文献[12]为了解决CNN-LSTM模型不能量化预测的不确定性,提出了一种基于CNN-双向LSTM(BiLSTM)和深度集合策略的日前风速概率预测模型。文献[13]为了解决LSTM和BiLSTM存在对序列结构信息不敏感、难以解决长时依赖等问题,将注意力机制分别引入CNN-LSTM和CNN-BiLSTM模型中,从而实现了更高的预测精度,但CNN-BiLSTM-Attention模型在训练过程中会受到超参数影响。

上述研究方法的重点是从预测模型的角度考虑如何提高预测精度,然而在实际过程中,如果将未经处理的风电序列直接作为输入,将会影响模型的预测精度,因此,为了提高预测精度,研究人员经常采用一些数据分解技术来处理原始数据。如文献[14]通过经验模态分解(EMD)将历史风速观测值分解为频率不同的有限本征模态函数(IMFs),并将其作为预测模型的输入,实现了更高的预测精度,但EMD会受到模态混叠的限制。文献[15]为了减少EMD出现模态混叠的现象,采用集合经验模态分解(EEMD)方法对原始数据进行处理,最终建立了EEMD和LSTM方法混合的风速预测模型。但EEMD方法需要进行多组EMD分解,计算量大。文献[16]结合变分模态分解方法(VMD)将原始数据分解为模态分量,然后分别对分解的模态向量建立了预测模型,并预测分量的结果进行累加,最终得到了风电功率的预测结果。虽然采用VMD方法计算量小,且能够提高模型的预测精度,但其参数的设定依赖于经验,若未能选择最佳的参数,将会影响预测的准确性[17]。

此外,上述方法没有考虑到预测模型的误差序列,因此本文研究了一种基于改进CNN-BiLSTM-Attention

模型的两阶段短期风电功率预测方法。该方法在第一阶段为功率预测阶段，即通过预测模型求取初始功率预测数据，第二阶段为误差补偿阶段，即通过误差数据求取初始误差预测数据，最后将初始功率预测数据和初始误差预测数据求和得到最终的风电功率预测数据。

2 风电功率预测分析

2.1 风电功率影响因素

风电功率的可预见性体现在风力发电功率是关于气象影响因素及机组自身特性的函数。风力发电的原理公式如式(1)所示[18]。

$$P = \frac{1}{2} C_p A \rho v^3 \quad (1)$$

式中： P 为风机的输出功率，kW； C_p 为风能利用系数； A 为风轮扫掠面积， m^2 ； ρ 为空气密度， kg/m^3 ； v 为风速，m/s。

式(1)表明风机的输出功率与风轮扫掠面积、空气密度和风速相关，其中风轮扫掠面积与风向相关。

式(1)中空气密度 ρ 的计算公式如式(2)所示。

$$\rho = 3.48 \frac{P_a}{T} \left(1 - 0.378 \frac{\phi P_b}{P_a} \right) \quad (2)$$

式中： P_a 为大气压力，kPa； T 为热力学温度，K； P_b 为饱和水蒸汽压力，kPa； ϕ 为空气相对湿度，%。

式(2)表明空气密度 ρ 与气压、温度和湿度相关。

由上可知，风电功率受到风速、风向、温度、气压、空气密度和湿度的影响，因此本文考虑到这六种因素对风电功率进行预测。

2.2 风电功率预测流程

风电功率的预测流程主要分为数据提取、数据预处理、预测模型选择与构建、模型训练与优化和预测结果输出和评估。

为了避免 VMD 方法中的分解层数需要依靠经验人为选取的问题，本文采用 SVMD 方法进行数据分解；同时，针对 CNN-BiLSTM-Attention 模型在训练过程中会受到超参数影响，本文采用贝叶斯优化算法实现超参数调优；此外，为了解决现有风电功率预测方法没有考虑到预测模型的误差，本文研究了一种误差补偿机制对改进 CNN-BiLSTM-Attention 模型进行补偿。本文研究的预测模型流程如图 1 所示。

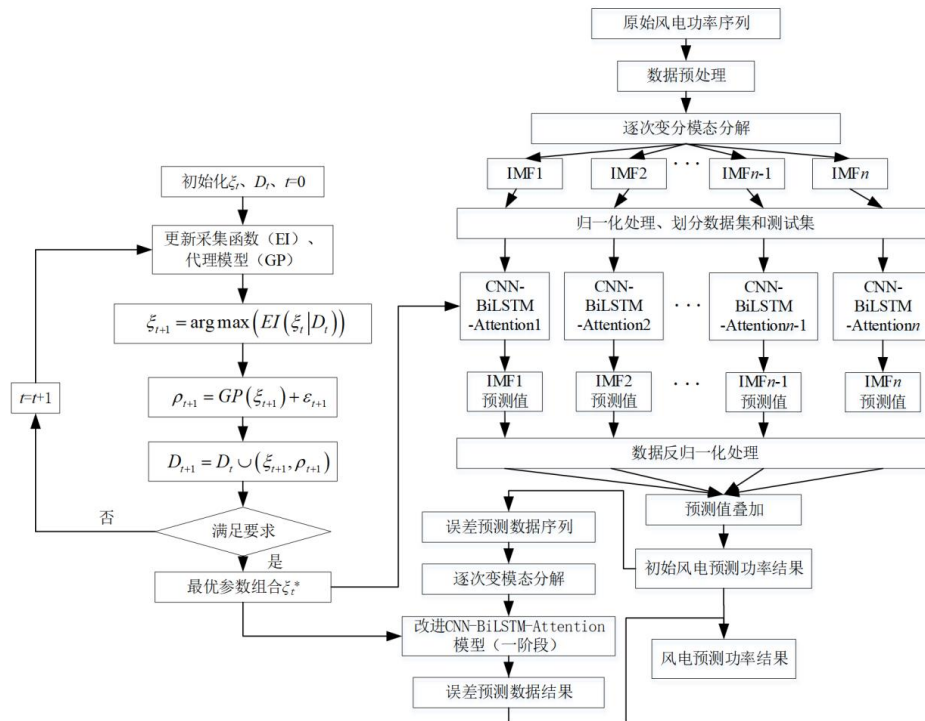


图 1 风电功率预测流程图

Figure 1 Wind power prediction flow char

具体步骤如下。

- 1) 提取数据。提取原始的历史气象数据和历史风电功率数据序列，其中历史气象数据包括风速、风向、温度、气压、空气密度和湿度。
- 2) 数据预处理。对所提取的数据进行数据预处理。数据预处理包括数据清洗、去噪和插值。
- 3) 数据分解。利用 SVMD 对原始序列进行分解，将得到的 IMFs 进行归一化处理。
- 4) 模型训练与优化。选择均方根误差 (root mean square error, RMSE) 作为模型的适应度值，利用 BO 算法对模型的超参数进行优化，得到最优的超参数，最终求取初始风电预测结果。
- 5) 误差补偿。根据初始风电预测结果计算误差预测数据序列，利用步骤 3 对风电功率误差预测序列进行分解，利用步骤 4 对改进 CNN-BiLSTM-Attention 模型进行误差预测训练，进而求取初始误差预测数据，最终将初始风电功率预测数据和误差预测数据进行求和，从而完成误差补偿。

3 逐次变分模态分解

在 VMD 分解过程中，选择合适的模态分解函数个数十分重要。传统的 VMD 方法中的分解层数需要依靠经验人为选取，仅凭经验设置这两个参数会对 VMD 分解结果带来较大误差。逐次变分模态分解 (successive variational mode decomposition, SVMD) 方法不需要在信号处理过程中指定模态分量的精确数量，并且能够在提高收敛速度的同时减少计算时间[19]，因此本文 SVMD 方法对风电功率进行分解。

假设可以将时域信号 $f(t)$ 分解为 L 阶模态 $\mu_L(t)$ 和残差信号 $f_r(t)$ 两个部分：

$$f(t) = \mu_L(t) + f_r(t) \quad (3)$$

若要满足以上的假设，则需满足下列条件：

$$J_1 = \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_L(t) \right] e^{-j\omega_L t} \right\|_2^2 \quad (4)$$

式中： ∂_t 表示对时间的导数； $\delta(t)$ 表示狄拉克函数； ω_L 则表示 L 阶模态的中心频率。 $*$ 表示卷积算子。

在 $\mu_L(t)$ 具有有效分量的基础上，残差信号 $f_r(t)$ 应该尽可能地最小化。为了保证改约束能够稳定运行，选

择适当的滤波器 $\beta_L(t)$ ，频率响应为：

$$\hat{\beta}_L(\omega) = \frac{1}{\alpha(\omega - \omega_L)^2} \quad (5)$$

应建立的约束为

$$J_2 = \left\| \beta_L(t) * f_r(t) \right\|_2^2 \quad (6)$$

式中： $\beta_L(t)$ 为式 (5) 滤波器的脉冲响应

通过最小化 J_1 和 J_2 约束，可能无法有效区分第 L 阶模态和前 $L-1$ 阶模态。因此，基于约束 J_2 的建立思路，所用滤波器的频率响应为：

$$\hat{\beta}_i(\omega) = \frac{1}{\alpha(\omega - \omega_i)^2}, i = 1, 2, \dots, L-1 \quad (7)$$

于是，所建立的约束为：

$$J_3 = \sum_{i=1}^{L-1} \left\| \beta_i(t) * \mu_L(t) \right\|_2^2 \quad (8)$$

式中： $\beta_i(t)$ 为式(8)中滤波器的脉冲响应。

在进行分解时，为保证信号能够完全重构，建立如下约束为：

$$f(t) = \mu_L(t) + f_r(t) + \sum_{i=1}^{L-1} \mu_i(t) \quad (9)$$

因此，提取模态分量的问题可以表示为有约束的最小化问题：

$$\begin{cases} \min_{\mu_L, \omega_L, f_r} (\alpha J_1 + J_2 + J_3) \\ \text{s.t. } \mu_L(t) + f_r(t) = f(t) \end{cases} \quad (10)$$

式中： α 为平衡 J_1 、 J_2 、 J_3 。

4 预测网络模型

4.1 CNN-BiLSTM-Attention 模型

CNN-BiLSTM-Attention 模型分为三个部分，分别为 CNN 层、BiLSTM 层和注意力机制。CNN 层对风电功率数据的空间特征进行提取；BiLSTM 网络分别结合了前向 LSTM 和后向 LSTM，通过隐藏层同时读取历史数据前向和后向的信息特征，挖掘当前时刻数据同过去和未来时刻数据的内在联系；通过添加注意力 (Attention) 机制，模型能够更加有效地识别不同信息的特征变化规律[20]。CNN-BiLSTM-Attention 模型的

结构如图 2 所示。

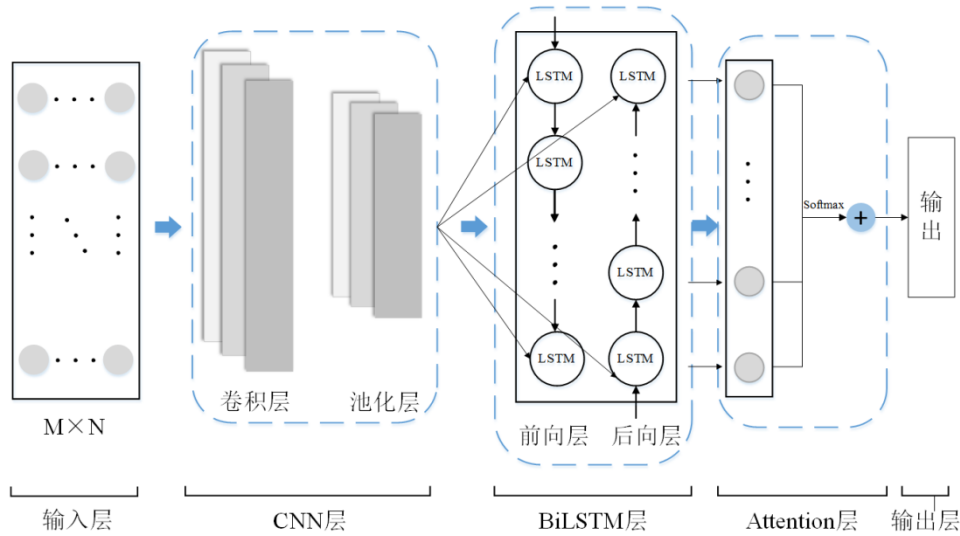


图 2 CNN-BiLSTM-Attention 模型网络结构图

Figure 2 CNN-BiLSTM-Attention Model network structure diagram

4.1.1 CNN 模型

卷积神经网络 (Convolutional Neural Network, CNN) 通常由输入层、卷积层、池化层和全连接层组成。其中从输入层输入数据后, 通过卷积层提取数据特征信息, 然后利用池化层选择这些特征, 减少特征图中变量的数量, 从而减少计算量, 提高计算效率。从池化层到全连接层, 将数据从多映射到少, 并进行降维, 以方便数据的处理[21]。在训练过程中, CNN 通过调整权重参数来最小化损失函数, 以使模型能够更好地预测目标。CNN 模型的网络结构如 3 所示。

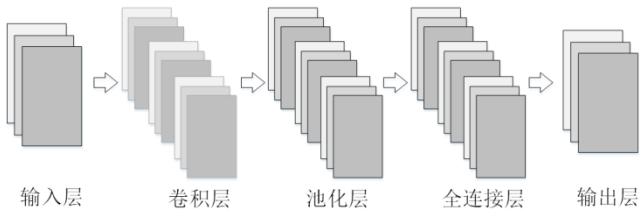


图 3 CNN 模型网络结构图

Figure 3 CNN model network structure diagram

4.1.2 BiLSTM 模型

BiLSTM 通过结合前后信息可以准确地确定当前时刻的输出, 适合处理双向依赖的长序列任务。BiLSTM 是由一个前向传播及一个后向传播的 LSTM 层堆叠起来组成。LSTM 模型主要是通过引入独特的

“门”结构来提取时间特征。其中包含三个门: 输入门、遗忘门和输出门。LSTM 的单元结构如图 4 所示,

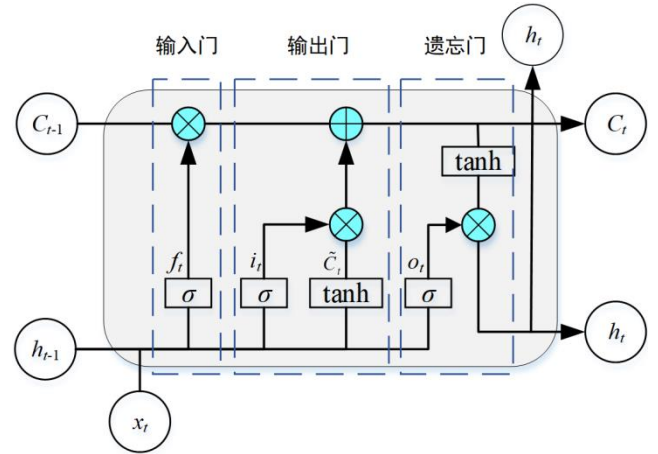


图 4 LSTM 单元结构图

Figure 4 LSTM unit structure diagram

图中 x_t 为 t 时刻输入, h_t 为 t 时刻隐藏输出; C_t 为 t 时刻记忆单元, $C_t^{(\sim)}$ 为 t 时刻临时记忆单元; $\odot$ 表示逐元素相乘, $\oplus$ 表示逐元素相加。

输入门获取外部信息; 遗忘门决定是否选择性地遗忘神经元细胞中的信息; 输出门负责输出当前状态信息。计算过程如下:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (11)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (12)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (13)$$

$$C_t = f_t \otimes C_{t-1} \oplus i_t \otimes \tilde{C}_t \quad (14)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (15)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (16)$$

式中: W_f 、 W_i 、 W_c 、 W_o 为各个模块对应的权重矩

阵, b_f 、 b_i 、 b_c 、 b_o 为偏置项, σ 为 sigmoid 激活函数, $\tanh$ 为双曲正切激活函数。

由上可知 LSTM 的网络参数是对数据按照从前到后的训练得到的, 这种训练方式对数据的利用率较低, 无法充分提取数据在时间序列中的内在特征联系。而双向 LSTM 是由两个算法相同的单向 LSTM 组成, 一个用来从前到后提取数据特征, 一个用来从后到前提取数据特征, 通过结合前后信息可以准确的确定当前时刻的输出, BiLSTM 的神经网络结构如图 5 所示。

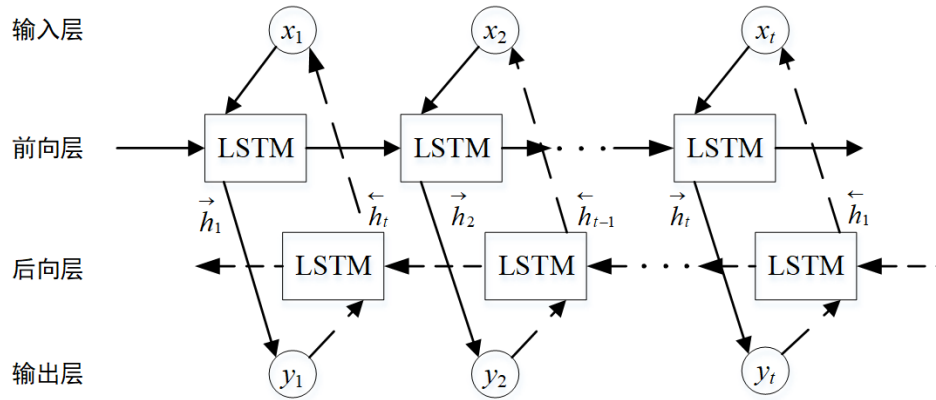


图 5 BiLSTM 神经网络结构图

Figure 5 BiLSTM neural network structure diagram

假设 BiLSTM 在 t 时刻的输入特征为 x_t , BiLSTM 的更新过程如下:

$$\vec{h}_t = LSTM(n_t, x_t) \quad (17)$$

$$\overleftarrow{h}_t = LSTM(n_t, x_t) \quad (18)$$

$$y_t = \sigma(W_y \cdot [\vec{h}_t, \overleftarrow{h}_t] + b_y) \quad (19)$$

式中, $h^{(\rightarrow)}_t$ 为 LSTM 在 t 时刻的正向输出特征, $h^{(\leftarrow)}_t$ 为 LSTM 在 t 时刻的反向输出特征; W_y 、 b_y 分别为权重矩阵和偏置项; y_t 为 BiLSTM 在 t 时刻的输出特征。

4.1.3 注意力机制

虽然 BiLSTM 可以从数据中提取数据的相关性, 但在提取数据特征时会忽略不同信息的权重[22], 然而不同的信息对输出结果的影响是不同的, 因此本文在 BiLSTM 层之后引入了 SE 注意力机制。注意力机制的基本思想是识别不同信息的重要性, 对权重较大的信息给予更多的关注, 本质上是权重因子的分布。其步骤

分为

1) 计算注意力权重。将 BiLSTM 输出的特征向量 y_t 作为网络的输入向量, 然后使用打分函数 s_t 计算每个隐层向量的得分, 用来评估其对输出的影响。如式 (20) 所示

$$s_t = \tanh(y_t \cdot W_y + b_y) \quad (20)$$

2) 归一化处理。通过利用 Softmax 函数对其进行归一化得到各特征的注意力权重系数 a_t 。如式(21)所示。

$$a_t = \text{Softmax}\left(\frac{\exp(s_t)}{\sum \exp(s_t)}\right) \quad (21)$$

3) 加权求和。对每个时间步长的局部信息进行加权求和, 得到全局特征输出值 s 。如式(22)所示

$$s = \sum a_t y_t \quad (22)$$

从 s 的计算结果可以看出, 每个 s 的结果值都与整个输入向量 y_t 相关。

4.2 贝叶斯优化

CNN-BiLSTM-Attention 模型在训练过程中, 会受到超参数影响, 当多种超参数耦合在一起时, 它们之间的相互作用会使调参过程变得复杂。而贝叶斯优化 (Bayesian Optimization, BO) 是一种有效的全局优化策略, 适用于那些难以获得闭式表达式的复杂目标函数[23]。该方法通过建立目标函数的概率模型, 能够在少次参数评价的前提下, 有效的求解这些复杂函数的最优解。

BO 方法对风电功率预测模型的寻优过程表示为

$$\xi_{t+1} = \arg \max_{\xi \in \Psi} f(\xi_t) \quad (23)$$

式中, f 为用于评价性能的目标函数; ξ 为模型的超参数组合; Ψ 为超参数空间。贝叶斯优化源于贝叶斯定理, 它使用贝叶斯公式来建立优化过程的概率分布。

$$p(f|D) = \frac{p(D|f)p(f)}{p(D)} \quad (24)$$

$$D = \{(\xi_0, \rho_0), (\xi_1, \rho_1), \dots, (\xi_t, \rho_t)\} \quad (25)$$

$$\rho_t = f(\xi_t) + \varepsilon_t \quad (26)$$

式中, D 为超参数观测值的集合; ρ 为超参数组合的评价, ε_t 为观测噪声; $p(f|D)$ 为 f 的后验概率分布; $p(f)$ 为 f 的先验概率分布; $p(D|f)$ 为观测值的似然分布; $p(D)$ 为边际似然分布。

贝叶斯优化 BiLSTM 网络模型超参数由两大核心

部分组成, 分别是概率代理模型和采集函数表达。

概率代理模型可以对函数分布进行估计, 常采用高斯过程, TPE 等。本文采用高斯过程 (Gaussian Process, GP) 为代理模型, 对变量 $x=[x_1, x_2, \dots, x_t]$, 有

$$g(x) \square GP(\mu(x), k(x, x')) \quad (27)$$

式中: $\mu(x)$ 为均值, 通常设置为 0; $k(x, x')$ 为协方差函数;

采集函数用来衡量不同取值的效能, 常采用概率提升法, 期望提升法, 置信度上界法等。本文采用期望改进 (Expectation Improvement, EI) 作为采集函数。

利用 BO 算法对 CNN-BiLSTM-Attention 模型进行超参数调优 (学习率、隐含层节点和正则化参数), 可以提高风电功率预测的性能。

5 实验验证

为验证所提模型的有效性, 本文以内蒙古某层高为 90m 的风电场为研究对象。采集 2020 年 9 月 18 日至 2020 年 10 月 2 日, 共 15 天的风电功率数据。每 15 分钟进行一次数据采集, 共 1440 个采样点。经过数据预处理后的原始风电功率序列如图 6 所示。本文将前 14 天的数据作为训练集, 去预测第 15 天的数据。其中第 15 天数据共 96 个样本点, 按每 6 小时时刻划分, 分别对应第 0 个样本点, 第 24 个样本点, 第 48 个样本点, 第 72 个样本点和第 96 个样本点。

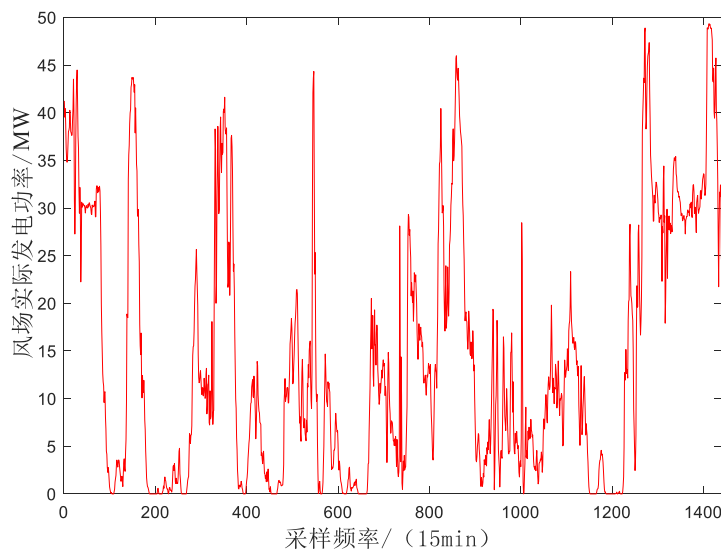


图 6 原始风电功率序列数据

Figure 6 Original wind power sequence data

为衡量风电预测性能, 本文将平均绝对误差 MAE、平均绝对百分误差 MAPE 和均方根误差 RMSE 作为风电功率预测的误差评价指标, 其公式分别为

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (28)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left(\frac{|\hat{y}_i - y_i|}{\hat{y}_i} \right) * 100\% \quad (29)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \quad (30)$$

式中: y_i^0 表示 i 时刻功率预测的实际值, y_i 代表 i 时刻功率的预测值, N 代表测试集中数据样本的数量。

5.1 分解结果

分别采用 EMD、EEMD、VMD 和 SVMD 算法分解风电功率数据的结果如图 7, 8, 9, 10 所示。

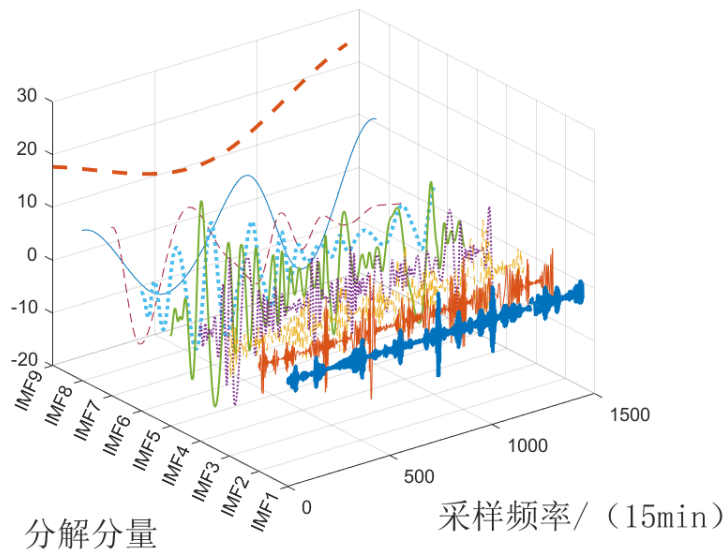


图 7 EMD 分解

Figure 7 EMD Decomposition

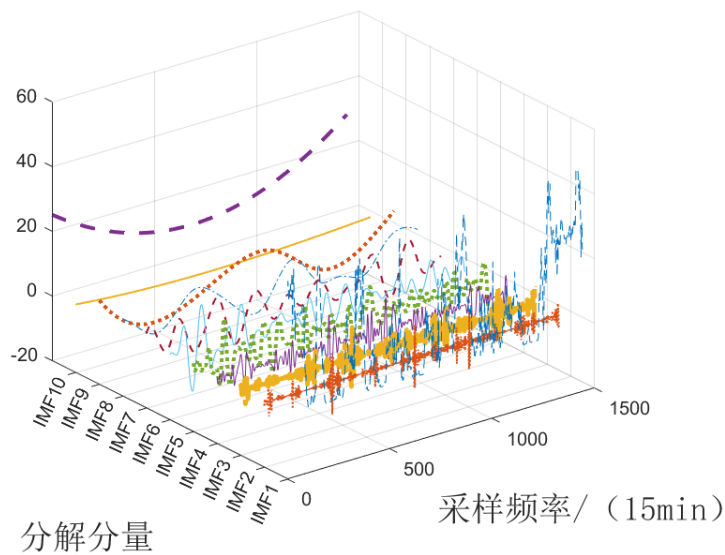


图 8 EEMD 分解

Figure 8 EEMD Decomposition

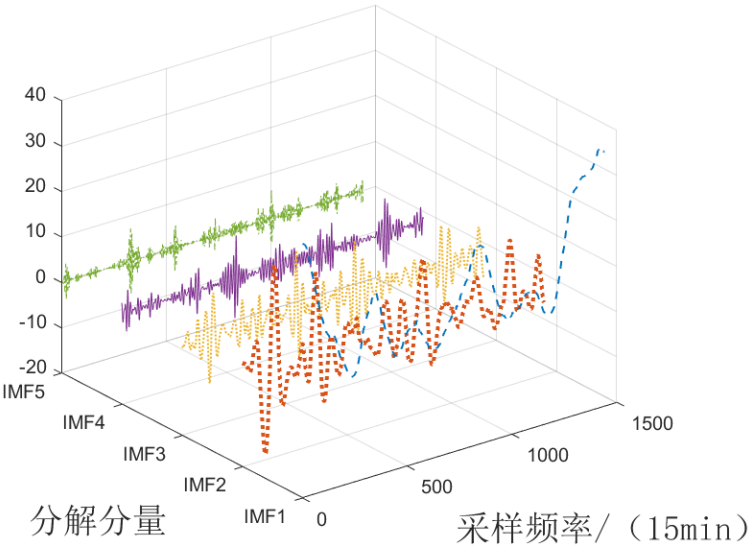


图 9 VMD 分解

Figure 9 VMD Decomposition

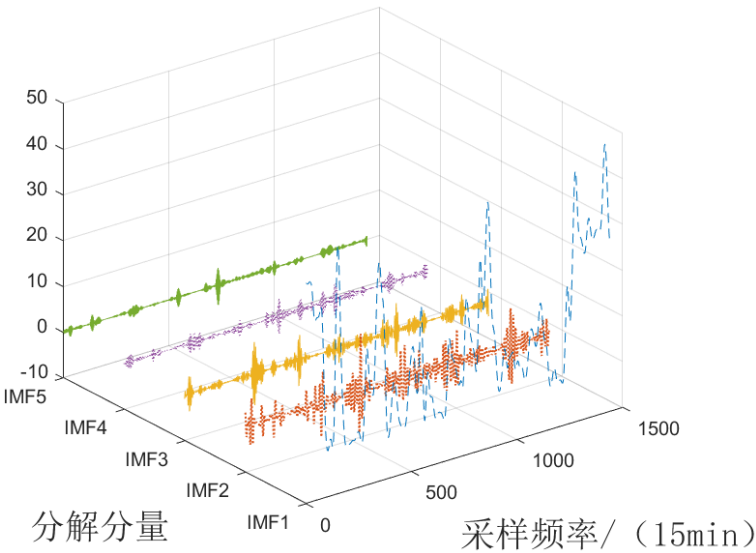


图 10 SVMD 分解

Figure 10 SVMD Decomposition

结合三维分解图可知经过分解之后的风电功率子序列，IMFs 频率由高到低排列，可以发现高频序列波动比较大，无论是 EMD、EEMD、VMD，还是 SVMD 模型均能展现出风电功率原始数据的波动特征，但 EMD 和 EEMD 分解得到的模态分量数量多于 VMD 和 SVMD，这是受到 EMD 分解的彻底性和 EEMD 白噪声的辅助作用和参数设置的影响。为了量化分解算法的性能，本文引入正交性指数、误差和运行时长三个参数对其进行进一步分析，如表 1 所示。

表 1 分解评价结果

Table 1 Decomposition evaluation results

指标	EMD	EEMD	VMD	SVMD
正交性指数	0.164	0.138	0.043	0.005
误差	1.674	1.445	0.573	0.474
运行时长/s	6.029	6.309	5.204	4.627

从正交性指数上来看，SVMD 正交性指数最小，其中正交性指数越小，表明分解性能越好；从重构后的误差角度分析，SVMD 的误差最小，这表明 SVMD

的分解准确度高；从运行时间来看，SVMD 的运行时长也最短。综上分析，SVMD 在性能上优于本文其他三种分解方法。

5.2 贝叶斯优化结果

得到分解结果后，采用 BO 算法优化改进 CNN-BiLSTM-Attention 功率模型的结果如图 11 所示。

其中各项参数范围设置如下：学习率范围为 $(10, 100)$ ，隐藏层节点为 $(10^{-2}, 1)$ ，正则化系数为 $(10^{-1}, 10^{-3})$ ，鉴于贝叶斯优化的高效性，并经过初步试验，发现在迭代 10 次后目标函数（RMSE）已无明显下降，故设置最大迭代次数为 10，以平衡优化效果与计算成本。

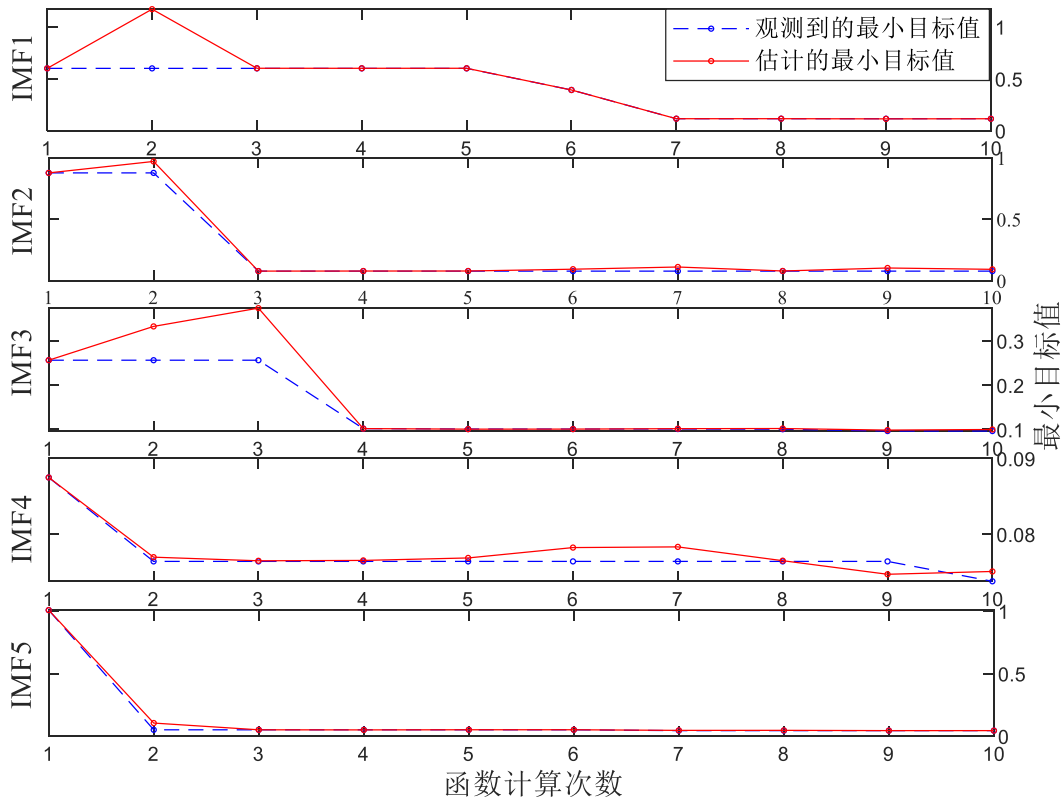


图 11 贝叶斯优化训练结果

Figure 11 Bayesian optimization training results

从图中可以看出，估计的最小目标值与观测到的最小目标值曲线除 IMF4 外，其他四个分量曲线贴合。为量化贝叶斯优化算法的效果并得到模型的三个超参数（学习率、隐含层节点和正则化参数），将贝叶斯优化算法的观测目标和估计目标的各项参数表 2 如下。

表 2 Bo 算法优化参数结果

Table 2 Algorithm optimization parameter results

指标	IMF1	IMF2	IMF3	IMF4	IMF5
观测目标值	0.114	0.078	0.096	0.073	0.044
估计目标值	0.114	0.078	0.100	0.075	0.044
观测隐含层节点	53	21	26	23	94
估计隐含层节点	53	21	26	23	94
观测学习率	0.008	0.002	0.002	0.003	0.007

指标	IMF1	IMF2	IMF3	IMF4	IMF5
估计学习率	0.008	0.002	0.002	0.003	0.007
观测正则化参数	4.623 (e-10)	1.412 (e-07)	2.315 (e-08)	1.082 (e-10)	1.202 (e-10)
估计正则化参数	4.623 (e-10)	1.412 (e-07)	2.317 (e-08)	1.083 (e-10)	1.202 (e-10)

从表 2 可以看出，贝叶斯优化对功率模型的五个 IMF 分量的估计目标值与观测值几乎吻合，同时估计得到的三个超参数也与观测值几乎吻合，进一步表明了 BO 算法可以通过准确的获取改进 CNN-BiLSTM-Attention 模型的学习率、隐含层节点和正则化参数，从而对 SVMD-CNN-BiLSTM-Attention 模型进行训练优化。

5.3 风电功率预测结果

在进行 SVMD 分解和 BO 算法优化之后，可以得到本文训练的风电功率预测模型。为了验证所提改进 CNN-BiLSTM-Attention 模型的有效性，本文比较 CNN-BiLSTM-Attention 模型、粒子群优化（PSO）算法、灰狼优化（GWO）算法和贝叶斯优化算法下的 SVMD-CNN-BiLSTM-Attention 模型预测结果，所得预测结果见图 12 所示。

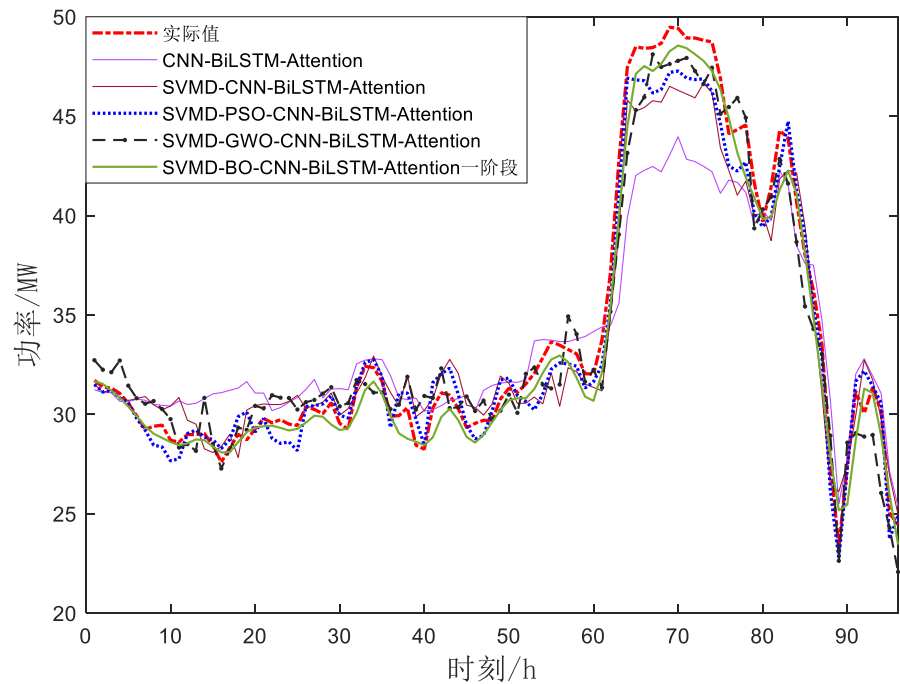


图 12 不同模型预测的结果对比图

Figure 12 Comparison of results predicted by different models

由图 12 中可知，在第 60-70 个样本点附近，实际功率出现剧烈波动。对比其他模型，基于 BO 优化的模型（BO-SVMD-CNN-BiLSTM-Attention）的预测曲线最能贴合这一变化趋势，这表明其超参数组合使模型具备了更强的非线性动态捕捉能力。

5.4 误差补偿结果

在得到功率预测数据结果后，进行二阶段预测，

其方法是将功率预测数据和真实值之间的误差序列作为改进 CNN-BiLSTM-Attention 模型的训练集输入，利用 SVMD 分解方法、BO 算法对风电功率进行误差预测，将误差预测结果与功率预测结果求和，进而完成对风电功率的预测补偿，所得功率预测补偿前后结果见图 13 所示。

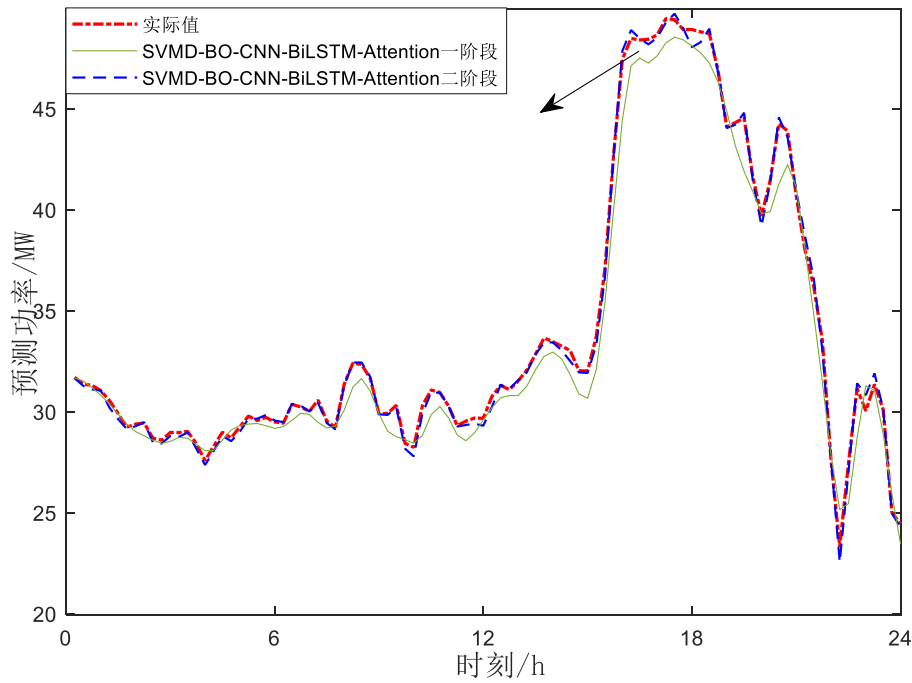


图 13 功率预测补偿前后对比图

Figure 13 Power prediction before and after compensation

从图 13 所示，在进行误差补偿后，各个时刻下模型的风电功率预测值与实际值更加贴合。为进一步评估预测的效果，利用 MAE、MAPE、RMSE 三种评价指标对误差补偿前后结果进行评价，评价结果如表 3 所示。

表 3 误差补偿前后结果

Table 3 Results before and after error compensation

模型名称	MAE	MAPE	RMSE
模型一阶段	0.778	2.456%	1.061
模型二阶段	0.111	0.608%	0.280

由表 3 可知，本文研究的改进 CNN-BiLSTM-Attention 的二阶段风电功率预测方法，在 MAE、MAPE、RMSE 的指标上均小于一阶段预测方法模型。与一阶段模型相比，二阶段模型中 MAE 值分别下降了 85.7%，MAPE 值下降了 75.2%，RMSE 值下降了 77.3%。

通过最终的预测结果对比，本文研究的改进 CNN-BiLSTM-Attention 模型的两阶段风电功率预测方法更好地实现了对风电功率数据的预测，证明了本文研究方法的有效性及其对短期风电功率预测精度的提升。

6 结论

本文在考虑预测模型的误差序列基础上，将 BO 优化算法和 SVMD 模态分解以及 CNN-BiLSTM-Attention 模型结合，建立了一种基于改进 CNN-BiLSTM-Attention 模型的两阶段短期风电功率预测方法。主要得到以下结论：

- 1) 采用 SVMD 分解能够将风电功率分解为特征性较为稳定的子序列，相比本文其他模态分解方法，避免了需要凭借经验选取分解层数导致精度下降的问题，从而提高了后续风电功率预测的精度。
- 2) 相比于 PSO 算法、GWO 算法而言，BO 算法在改进 CNN-BiLSTM-Attention 模型中能够更准确的预测风电功率，并且能够在实际值剧烈变化处也能较好地预测出其变化趋势，从而有效的解决模型的超参数的求取问题，避免了人工选取参数困难，造成的风电功率预测模型精度下降的问题。
- 3) 与单一阶段的预测模型相比，引入误差补偿机制后，模型的 MAE、MAPE 和 RMSE 值分别

显著下降了 85.7%、75.2%和 77.3%。表明了所提出的两阶段预测可有效修正初始预测的偏差，提升预测精度。

参考文献

- [1] 曾娜梅, 霍志红, 许昌, 等. 基于改进 HHT 的分钟级超短期风速预测 [J]. 动力工程学报, 2021, 41(04): 309-315+329.
- [2] 朱菊萍, 魏霞, 谢丽蓉, 等. 基于 VMD 和改进 BiLSTM 的短期风电功率预测 [J]. 太阳能学报, 2024, 45(06): 422-428.
- [3] 刘金朋, 邓嘉明, 高鹏宇, 等. 基于 VMD-IMPA-SVM 的超短期风电功率预测 [J]. 智慧电力, 2024, 52(07): 24-31+79.
- [4] 罗潇远, 刘杰, 杨斌, 等. 基于改进鱼鹰优化算法与 VMD-LSTM 的超短期风电功率预测 [J]. 太阳能学报, 2025, 46(03): 652-660.
- [5] 李文鑫. 基于深度神经网络的超短期风电功率预测研究 [D]. 保定: 华北电力大学控制与计算机学院, 2022.
- [6] 王继东, 于俊源, 孔祥玉. 基于双重分解和双向长短时记忆网络的中长期负荷预测模型 [J]. 电网技术, 2024, 48(08): 3418-3426.
- [7] 王霞. 基于数据驱动的风电特征提取和功率预测方法研究 [D]. 北京: 中国矿业大学电气工程学院, 2023.
- [8] 谷学静, 陈洪磊, 孙泽贤, 等. 基于 VMD-RL-LSTM 的短期风功率预测 [J]. 计算机仿真, 2023, 40(04): 89-93+309.
- [9] Junqiang W; Xuejie, W; Tianming Y, et al. Ultra-short-term forecasting of wind power based on multi-task learning and LSTM [J]. International Journal of Electrical Power & Energy Systems. 2023, 149, 109073.
- [10] Blazakis K, Katsigiannis Y, Stavrakakis G. One-day-ahead solar irradiation and windspeed forecasting with advanced deep learning techniques [J]. Energies, 2022; 15: 4361.
- [11] 唐清苇, 向月, 代佳琨, 等. 基于 CNN-LSTM 的风电场发电功率迁移预测方法 [J]. 工程科学与技术, 2024, 56(02): 91-99.
- [12] Zhang Y M, Wang H. Multi-head attention-based probabilistic CNN-BiLSTM for day-ahead wind speed forecasting [J]. Energy, 2023, 278, 127865.
- [13] 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(05): 44-53.
- [14] Wang Z J, Wang L, Revanesh M, et al. Short-Term Wind Speed and Power Forecasting for Smart City Power Grid With a Hybrid Machine Learning Framework [J]. IEEE Internet of Things Journal, 2023, 10(21): 18754-18765.
- [15] Yan Y, Wang X R, Ren F, et al. Wind speed prediction using a hybrid model of EEMD and LSTM considering seasonal features [J]. Energy Reports, 2022, 8: 8965-8980.
- [16] Zhang Y J, Zhang L, Sun D, et al. Short-Term Wind Power Forecasting Based on VMD and a Hybrid SSA-TCN-BiGRU Network [J]. Applied Sciences, 2023, 13(17): 988.
- [17] 马志侠, 张林鎔, 巴音塔娜, 等. 基于自适应二次分解与 CNN-BiLSTM 的超短期风电功率预测 [J]. 太阳能学报, 2024, 45(06): 429-435.
- [18] 何柳青. 基于深度学习的短期风电功率区间预测 [D]. 西安: 西安理工大学水利水电学院, 2023.
- [19] Reziwan K, Wumaier T. Yan F, et al. Hospital Outpatient Volume Prediction Model Based on Gated Recurrent Unit Optimized by the Modified Cheetah Optimizer [J]. IEEE Access, 2023, 11: 139993-140006.
- [20] 雷柯松, 吐松江·卡日, 伊力哈木·亚尔买买提, 等. 基于 WGAN-GP 和 CNN-LSTM-Attention 的短期光伏功率预测 [J]. 电力系统保护与控制, 2023, 51(09): 108-118.
- [21] 常青松, 杨昭, 杨熠辉, 等. 基于相似日聚类的超短期光伏功率组合预测模型 [J]. 热力发电, 2023, 52(11): 123-131.
- [22] 刘益豪, 石宇强. 基于注意力机制的 CNN-BiLSTM 复杂设备剩余寿命预测 [J]. 机械设计, 2024, 41(S1): 58-63.
- [23] 荆志宇, 李培强, 林文婷. 结合贝叶斯优化及通道注意力的双端优化时序式风功率预测模型 [J]. 电力系统及其自动化学报, 2024, 36(08): 39-47+59.