

基于二次分解和 IDBO-DABiLSTM 的短期风电功率预测模型

卢苾锋, 王霄*

(贵州大学电气工程学院, 贵州 贵阳 550025)

摘要: 为提高风电功率预测精度, 针对风电的强波动性和高随机性, 提出一种基于二次分解和改进蜣螂优化算法(IDBO)-双重注意力双向长短期记忆(DABiLSTM)网络的风电功率预测模型。首先, 采用自适应噪声的完全集经验模态分解(CEEMDAN)和小波包分解(WPD)构成一种二次分解方法对历史风电功率和风速数据进行分解, 降低初始序列的随机性和非平稳性。其次, 在 BiLSTM 网络的基础上, 加入特征和时间注意力机制, 建立 DABiLSTM 模型, 充分挖掘特征间的关联性和时间序列间的长时间依赖性。最后, 采用黄金正弦算法来优化滚球蜣螂的位置, 从而增强算法在局部和全局的探索能力, 同时引入动态权重系数改进偷窃蜣螂的位置, 以平衡算法在全局和局部的探索能力, 提出 IDBO, 并用其优化 DABiLSTM 网络的超参数, 防止网络陷入局部最优解。采用贵州某风电场实际数据对所提模型进行实验, 结果表明该方法能有效提升模型的预测能力, 所提出的模型的均方根误差(RMSE)和平均绝对误差(MAE)在单步预测下分别为 0.0449 和 0.0312 MW, 与其他模型相比, 分别平均降低了 36.9% 和 31.7%, 表现出较好的预测精度和鲁棒性。

关键词: 风电功率预测; 二次分解; 双向长短期记忆网络; 改进蜣螂优化算法; 注意力机制

中图分类号: TP391

文献标志码: A

DOI: 10.19678/j.issn.1000-3428.0068365

Short-Term Wind Power Prediction Model Based on Secondary Decomposition and IDBO-DABiLSTM

LU Yifeng, WANG Xiao*

(College of Electrical Engineering, Guizhou University, Guiyang 550025, Guizhou, China)

【Abstract】 To improve the accuracy of wind power prediction, a new wind power prediction model based on secondary decomposition and Improved Dung Beetle Optimizer (IDBO)-Dual Attention Bidirectional Long Short-Term Memory (DABiLSTM) network is proposed for the high stochasticity and strong volatility of wind power. First, a secondary decomposition method based on the Complementary Ensemble Empirical Mode Decomposition of Adaptive Noise (CEEMDAN) and Wavelet Packet Decomposition (WPD) is proposed to disaggregate the raw wind power historical data and windspeed historical data, thereby reducing the randomness and volatility of the raw signals. Second, a DABiLSTM network model is established by incorporating feature and time attention mechanisms. This model fully explores the correlation between features and long-term dependency between time series, thereby improving the accuracy of wind power prediction. Finally, IDBO is proposed based on the golden sine algorithm to improve the rolling ball dung beetle position and enhance the local development and global exploration capabilities of the algorithm. In addition, a dynamic weighting factor is incorporated to improve the stealing cockroach position and balance the global exploration and local development capabilities of the algorithm. IDBO is employed for intelligent optimization of the network working hyperparameters of the DABiLSTM model to further enhance the prediction precision of the model. The proposed model is tested using real data from a Guizhou wind farm, and the findings demonstrate that the proposed approach can successfully increase the predictive power of the model. The Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) values of the proposed model are 0.0449 and 0.0312, respectively, in single-step prediction, which is a reduction by 36.9% and 31.7% on average, respectively, compared with other models in the literature, thus showing better prediction accuracy and robustness.

【Key words】 wind power prediction; secondary decomposition; Bidirectional Long Short-Term Memory (BiLSTM) network; Improved Dung Beetle Optimizer (IDBO); attention mechanism

收稿日期: 2023-09-08 修回日期: 2023-12-12

基金项目: 国家自然科学基金(61861007, 61640014); 贵州省科技计划项目(黔科合基础-ZK[2021]一般 303); 贵州省科技支撑计划(黔科合支撑[2022]一般 017, 黔科合支撑[2022]一般 264, 黔科合支撑[2023]一般 096); 中国电力建设股份有限公司科技项目(DJ-ZDXM-2022-44); 贵州省教育厅创新群体项目(黔教合 KY 字[2021]012); 贵州大学引进人才科研项目(贵大人基合字(2014)08 号)。

通信作者 E-mail: *xwang9@gzu.edu.cn

0 引言

近年来,面对环境污染、温室效应和能源危机等挑战,风能作为一种无污染可再生的新兴绿色可再生能源正得到大力发展^[1]。然而,由于自然风具有间歇性、随机性和波动性等自然属性,风力发电随之具有强随机性、间歇性和高可变性等特点,这对风电并网和电力系统调度带来巨大挑战,进而影响电能质量和系统安全运行^[2]。因此,准确的风电功率预测技术对降低系统运行成本、减少电力系统旋转备用容量以及增加系统安全可靠具有十分重要的意义^[3]。

风电功率的准确预测一直是许多研究人员的目标,他们提出了各种解决方案^[4],其中一种方法是对风电信号进行分解,减少其波动性和非平稳性^[5],该方法可以有效提高预测精度。常用的分解算法有变分模态分解(VMD)^[6]、经验模态分解(EMD)^[7]和小波变换(WT)^[8]等。文献[9]对风电功率信号采用互补集合 EMD,降低了信号的波动性和非线性。文献[10]使用集合 EMD 将风电功率序列分解为子序列,再用卷积神经网络和极端梯度提升模型进行预测,提升了预测性能。然而,部分信号子序列在一次分解后仍然有较大的噪声和较高的频率。因此,对特定的子信号进行二次分解可能会显著提高风电功率预测的准确性。

深度学习方法在提取和拟合数据特征方面的卓越表现,使其近年来在风电功率预测领域得到广泛的应用^[11],其中循环神经网络(RNN)^[12]和长短期记忆(LSTM)网络^[13]做出了显著的贡献。双向 LSTM(BiLSTM)网络^[14]是在 LSTM 网络基础上改进的网络,能够很好地学习输入序列的过去和未来时刻的双向信息,弥补了 LSTM 网络无法从后向前编码信息的缺陷。文献[15]对 LSTM 网络进行改进,提出了一种基于遗传 LSTM 网络的模型,实现了准确、可靠的预测。文献[16]提出了一种基于卷积神经网络(CNN)和 BiLSTM 网络的组合模型,提高了时间序列数据的利用率和预测精度。上述预测方法虽对风电数据进行了挖掘,但难以获取特征间的相关信息和时间信息。

深度学习模型的性能通常高度依赖于超参数的调整^[17],而手动选择网络的超参数会影响模型的准确性。为了克服这个问题,研究人员尝试采用智能优化技术来对网络进行优化。文献[18]运用布谷鸟搜索优化算法^[19]优化增强遗忘门 LSTM 网络的超参数,克服了传统预测模型的局限性,有效地提高了

预测精度。文献[20]通过鲸鱼优化算法^[21]优化核极限学习机的初始值和阈值,降低了预测的复杂性和重建误差。然而,上述方法虽然运用了优化算法,但所运用的优化算法存在全局和局部开发能力不充分平衡等问题。

针对上述问题,本文提出了一种基于二次分解和改进蜣螂优化算法(IDBO)-双重注意力 BiLSTM(DABiLSTM)网络模型的风电功率组合预测模型。首先,采用自适应噪声的完全集合经验模态分解(CEEMDAN)对初始风速和风电功率信号进行分解,减小信号的波动性和非线性,并使用小波包分解(WPD)对 CEEMDAN 分解产生的第一分量进行分解。其次,引入特征和时间注意力机制对 BiLSTM 网络进行优化,深度挖掘特征和时间信息,以突出关键输入信息的重要性,并深入挖掘序列的整体特征和提取时间序列的长时间依赖性。然后,提出 IDBO 平衡全局搜索和局部开发能力,并用其对 DABiLSTM 模型的学习率、迭代次数以及隐藏层神经元个数等超参数进行优化,避免网络陷入局部最优解。最后,用贵州某风电场的实测数据进行仿真,验证所提方法的有效性。

1 二次分解

原始风电功率受到风速等环境因素的影响,具有波动性和随机性等特点,导致直接预测的精度低。本文采用二次分解将原始风速与风电功率数据进行分解,以提高风电功率预测的准确性和稳定性。首先使用 CEEMDAN 将原始风电功率和风速信号分解成若干子序列;其次使用 WPD 对频率较高的第一分量(IMF₁)进行二次分解,降低序列的非平稳性与非线性,进一步提高预测精度。

1.1 CEEMDAN

CEEMDAN 是 EMD 的改进^[22],其减小了重构误差,提高了分解效率。首先定义 $E_n(\cdot)$ 为 EMD 算法分解得到的第 n 阶模态分量, IMF_n 为 CEEMDAN 算法得到的第 n 个模态分量。CEEMDAN 算法的操作步骤如下:

1)将 K 次高斯白噪声序列 $\delta_k(t)$ 添加到原始信号中(其中白噪声序列的方差为 1,均值为 0),得到 K 次的待分解子序列 $h_k(t)$ ($k=1,2,\dots,K$):

$$h_k(t) = x(t) + \alpha \delta_k(t) \quad (1)$$

式中: $x(t)$ 为原始信号; α 为白噪声权值系数; $\delta_k(t)$ 为第 k 次添加的高斯白噪声序列。

2)使用 EMD 算法分解 $h_k(t)$,对得到的 K 个

第一分量求取平均值,计算出第 1 个模态分量 IMF_1 及第 1 阶残余分量 $r_1(t)$:

$$IMF_1(t) = \frac{1}{K} \sum_{k=1}^K IMF_1^k(t) \quad (2)$$

$$r_1(t) = h(t) - IMF_1(t) \quad (3)$$

3)将高斯白噪声添加到分解后得到的第 1 阶残余分量中,随后再对 $r_1(t) + \alpha_1 E_1(\delta_k(t))$ 进行 EMD,得到第二阶模态分量 IMF_2 :

$$IMF_2(t) = \frac{1}{K} \sum_{k=1}^K E_1(r_1(t) + \alpha_1 E_1(\delta_k(t))) \quad (4)$$

4)令 $n=2, \dots, N$,则第 k 阶残余分量可表示如下:

$$r_n(t) = r_{n-1}(t) - IMF_n(t) \quad (5)$$

5)依次计算,对于第 n 阶分量,与步骤 3)相似,得到第 n 阶模态分量 IMF_n :

$$IMF_n(t) = \frac{1}{K} \sum_{k=1}^K E_{n-1}(r_{n-1}(t) + \alpha_{n-1} E_{n-1}(\delta_k(t))) \quad (6)$$

6)重复执行步骤 3)的操作,直到余量信号极值点个数小于 3 个为止,最终获得 CEEMDAN 分解序列:

$$x(t) = \sum_{n=1}^N IMF_n + r_N(t) \quad (7)$$

本文中,原始风电功率和原始风速数据经过 CEEMDAN 分解成了 7 个分量,图 1 为 CEEMDAN 分解后的风电功率子序列结果。可以看出,CEEMDAN 分解风电功率信号后,降低了其非线性和随机性。

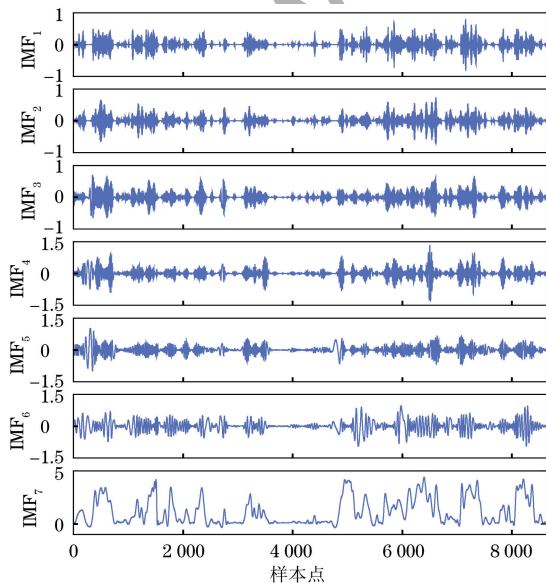


图 1 CEEMDAN 分解后的风电功率序列

Fig.1 Wind power sequence after CEEMDAN decomposition

1.2 WPD

本文使用 WPD 对频率较高、噪声较多的第一分量 IMF_1 进行进一步分解,进一步降低数据的波动性和随机性。

WPD 是一种信号时域分析方法,其在小波分解的基础上改进而来,减少了小波变换高频和低频段频率混淆的影响,具有更高的分析率,擅长分解非平稳性信号。

WPD 算法为:

$$d_i^{2n}[k] = \sum_{l \in \mathbb{Z}} h_{l-2k} d_{i+1}^n[l] \quad (8)$$

$$d_i^{2n+1}[k] = \sum_{l \in \mathbb{Z}} g_{l-2k} d_{i+1}^n[l] \quad (9)$$

小波包重构算法为:

$$d_{i+1}^n[k] = \sum_{l \in \mathbb{Z}} h_{k-2l} d_i^{2n}[l] + \sum_{l \in \mathbb{Z}} g_{k-2l} d_i^{2n+1}[l] \quad (10)$$

式中: h 是低通滤波器系数; g 是高通滤波器系数; $d_i^{2n}[k]$ 为第 n 层分解的低频系数; $d_i^{2n+1}[k]$ 为第 n 层分解的高频系数。

本文对频率较高、噪声较大的 IMF_1 再次进行了 WPD,得到了更平稳、更精细的子序列。图 2 展示了风电功率 IMF_1 信号的小波包 3 层分解结果。

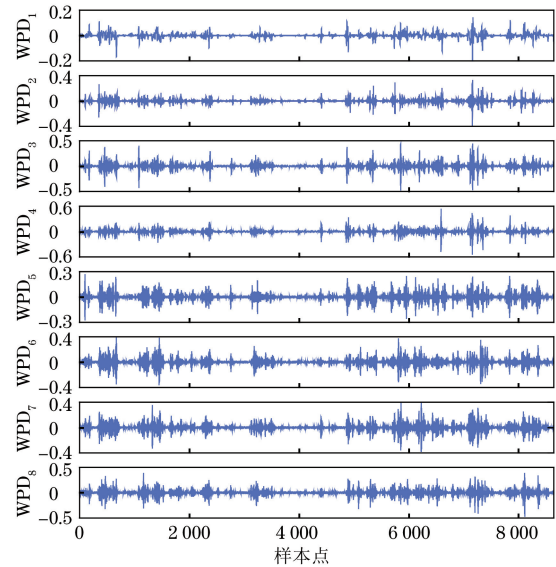


图 2 IMF_1 信号的 WPD 分解结果

Fig.2 WPD decomposition results of IMF_1 signal

2 风电功率预测模型

2.1 BiLSTM 网络

LSTM 是一种在 RNN 的基础上改进的网络,结构包括输入门、输出门和遗忘门。BiLSTM 可以学习 LSTM 所忽略的后续信息,在时间序列预测中

可以有效考虑前后数据的内在规律,捕获 LSTM 通常忽略的数据的隐藏特征和模式。BiLSTM 的网络结构如图 3 所示。

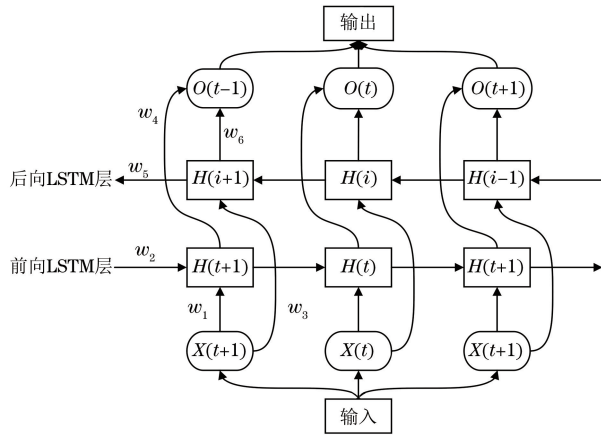


图 3 BiLSTM 网络结构

Fig. 3 The structure of BiLSTM network

由图 3 可知,当前输出受到正向层和后向层输出的共同作用。BiLSTM 在 t 时刻的输出 $O(t)$ 如下:

$$\mathbf{H}(t) = \sigma(\omega_1 x(t) + \omega_2 \mathbf{H}(t-1) + \mathbf{b}_{H(t)}) \quad (11)$$

$$\mathbf{H}(i) = \sigma(\omega_3 x(t) + \omega_5 \mathbf{H}(i-1) + \mathbf{b}_{H(i)}) \quad (12)$$

$$\mathbf{O}(t) = \sigma(\omega_4 \mathbf{H}(t) + \omega_6 \mathbf{H}(i) + \mathbf{b}_{O(t)}) \quad (13)$$

式中: ω 为权重系数; $\mathbf{H}(t)$ 为正向隐藏层输出; $\mathbf{H}(i)$ 为后向隐藏层输出; $\mathbf{O}(t)$ 为最终输出层; $\mathbf{b}_{H(t)}$ 是正向隐藏层偏置项; $\mathbf{b}_{H(i)}$ 是后向隐藏层偏置项; $\mathbf{b}_{O(t)}$ 是最终输出层偏置项。

2.2 DABiLSTM 模型

尽管 BiLSTM 网络比 LSTM 网络有所改进,但对于太长的序列,它仍然不能很好地传输序列开始时的信息。此外, BiLSTM 网络不加区分地压缩了所有信息,不能学习序列的整体特征以及输入信息的时间差异性。本文提出 DABiLSTM,该模型可以自适应地分配输入信息的注意力权重,以突出关键输入信息的重要性,并深度提取序列的整体特征和挖掘各个子序列之间的时间相关性。图 4 为 DABiLSTM 的网络结构。

如图 4 所示, DABiLSTM 在前向 LSTM 前加入特征注意力机制来学习输入序列的特征信息,在后向 LSTM 前加入时间注意力机制来挖掘风电时间序列的时间相关性。通过双重注意力机制的作用,可以提升模型对特征的学习,避免一些关联信息的丢失,并抓取风电时间序列的长时间依赖性。

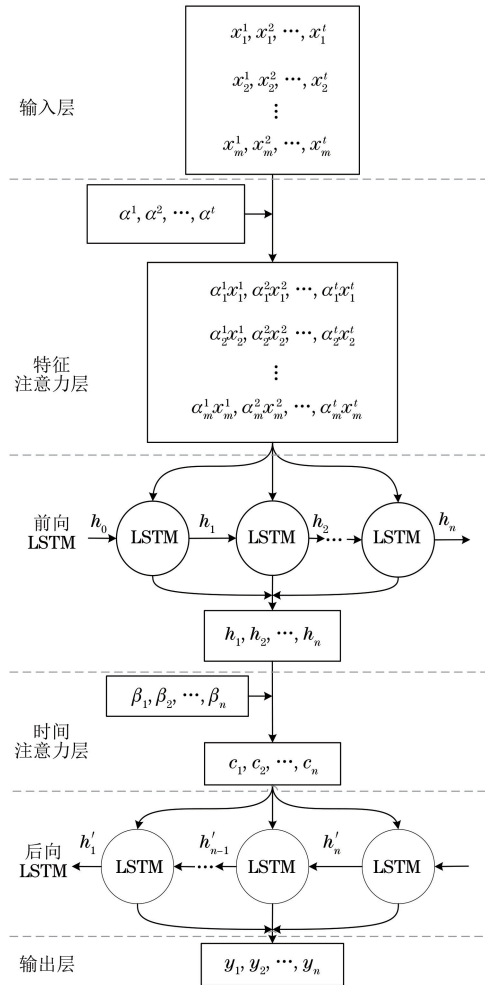


图 4 DABiLSTM 网络结构

Fig. 4 The structure of DABiLSTM network

给定含有 m 个特征的输入时间序列 $\mathbf{x}^t: \{x_1^t, x_2^t, \dots, x_m^t\}$, 通过多层感知机计算出第 k 个特征在 t 时刻的权重 $\lambda^t: \{\lambda_1^t, \lambda_2^t, \dots, \lambda_m^t\}$ 。

$$\lambda^t = \sigma(\mathbf{W}_\lambda \mathbf{x}^t + \mathbf{b}_\lambda) \quad (14)$$

式中: \mathbf{W}_λ 为训练权重矩阵; \mathbf{b}_λ 为偏置向量。然后运用 Softmax 函数对各注意力权重做归一化处理,使得权重值的和为 1。通过权重 α^t 与输入特征向量 \mathbf{x}^t 相乘得到加权向量 $\tilde{\mathbf{x}}^t$:

$$\tilde{\mathbf{x}}^t = \{\alpha_1^t x_1^t, \alpha_2^t x_2^t, \dots, \alpha_m^t x_m^t\} \quad (15)$$

将 $\tilde{\mathbf{x}}^t$ 作为前向 LSTM 的输入,得到后向 LSTM 的隐藏层状态 $\mathbf{h}^t: \{h_1^t, h_2^t, \dots, h_n^t\}$, 其中 n 为输入序列的时序长度。由于增加输入信号的长度会降低后向 LSTM 的预测能力,因此本文加入时间注意力机制,提取信号中各个时间节点之间的时间信息来捕获风电时间序列的长期依赖性,突出关键时间点信息在模型中的表达。通过后向 LSTM 得到时间注意力机制的权重 β_t , β_t 表示 t 时刻后向 LSTM 隐藏层对输出结果的影响程度。最后,得到最终的输出。

2.3 IDBO

蜚螂优化算法(DBO)是研究者在研究蜚螂的生物行为过程后提出的一种新的群体智能优化算法^[23]。该算法的灵感主要来源于蜚螂的滚球、跳舞、偷窃、觅食和繁殖等行为,算法具有全局搜索能力强、寻优精度高等特点。

考虑到原始 DBO 收敛速度慢、全局搜索和局部开发不平衡等问题,本文提出 IDBO。首先采用黄金正弦算法优化蜚螂滚球的位置计算过程,然后加入动态权重系数优化和调整蜚螂偷窃的位置更新方式。该算法不仅加快了收敛速度,而且在解空间的全局探索和局部开发过程中取得了更平衡的性能表现,在优化问题上表现出更卓越的适应性和效率。IDBO 中主要步骤如下:

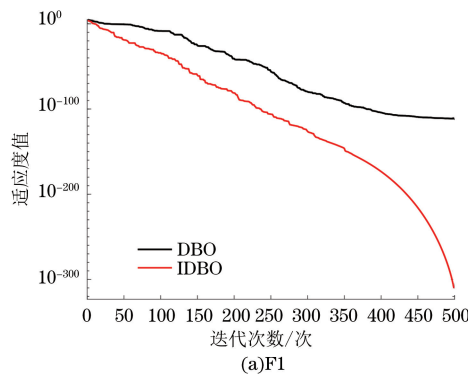
1)利用黄金正弦算法改进滚球蜚螂位置。

蜚螂利用太阳在滚动的粪球中进行导航。本文通过黄金正弦算法优化滚球蜚螂的位置。该算法模拟正弦波的动态搜索特性,在寻找全局最优解方面展现出强大的性能,同时加入黄金分割比例系数来细化算法的位置更新,促进了算法对局部区域的详尽搜索,实现了算法在全局探索和局部精细化探索之间的高效协同。位置的更新过程如下:

$$\begin{cases} \mathbf{X}_i(t+1) = \mathbf{X}_i(t) \cdot |\sin(R_1)| + \\ R_2 \cdot \sin(R_1) \cdot A \\ A = |x_1 \cdot \mathbf{X}_i^b(t) - x_2 \cdot \mathbf{X}_i(t)| \\ x_1 = -\pi + (1-\tau) \cdot 2\pi \\ x_2 = -\pi + \tau \cdot 2\pi \end{cases} \quad (16)$$

式中: t 是当前迭代次数; $\mathbf{X}_i(t)$ 表示第 i 只滚球蜚螂在第 t 次迭代时的位置信息; $\mathbf{X}_i^b(t)$ 表示第 t 代个体 i 的最优位置; R_1 和 R_2 是随机数, $R_1 \in [0, 2\pi]$, $R_2 \in [0, \pi]$; τ 是黄金分割系数; x_1 和 x_2 是根据 τ 得到的系数。

位置更新算法改进后,蜚螂个体在算法的每次迭代过程中会与当前已知的最优个体交换信



息,确保种群能够充分利用并指向当前最优,从而加速收敛过程并提升解的质量。同时,将黄金正弦算法融入算法,扩展了搜寻解空间的范围,提高了算法的全局寻优能力,有助于算法跳出局部最优解,避免算法过早收敛。为了进一步优化搜索效率,利用黄金分割系数逐步缩小搜索空间,通过选定的 R_1 和 R_2 参数精确控制了蜚螂的搜索距离和探索方向,有效地均衡了全局探索与局部开发能力,在探索大范围解空间时,能够对有潜力的区域进行深入挖掘。

2)引入动态权重系数改进偷窃蜚螂的位置。

在种群内部,有一些蜚螂会从其他蜚螂那里偷取粪球,它们被称为偷窃蜚螂。本文引入动态权重系数改进偷窃蜚螂的位置。偷窃蜚螂的位置更新过程如下:

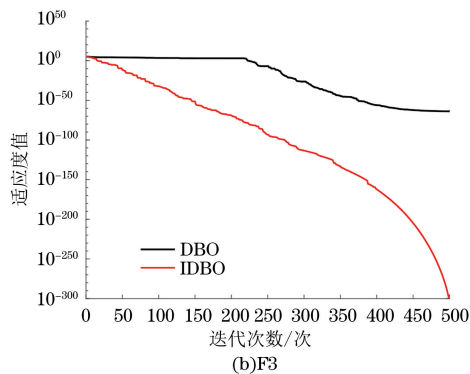
$$\begin{cases} \mathbf{X}_i(t+1) = k_1 \cdot \mathbf{X}^b + k_2 \cdot S \cdot \mathbf{g} \cdot B \\ B = |\mathbf{X}_i(t) - \mathbf{X}^*| + |\mathbf{X}_i(t) - \mathbf{X}^b| \\ k_1 = 1 - t^3 / T_{\max}^3 \\ k_2 = t^3 / T_{\max}^3 \end{cases} \quad (17)$$

式中: $\mathbf{X}_i(t)$ 为第 i 只偷窃蜚螂在第 t 次迭代时的位置信息; S 为一个常量; \mathbf{g} 表示服从正态分布的大小为 $1 \times D$ 的随机向量; T_{\max} 为最大迭代次数。

k_1 在迭代前期较大,使蜚螂在最优蜚螂附近探索更优的区域,增加算法的全局寻优能力; k_2 在迭代后期逐渐增加,使蜚螂具备跳出局部最优的能力,进一步平衡了算法全局探索和局部开发的能力。

为验证 IDBO 具有更强的寻优能力,本文分别用 IDBO 与 DBO 对 4 个单峰基准测试函数(F1、F3、F5、F7)与多峰基准测试函数(F8、F9、F11、F13)进行测试,结果如图 5 所示(彩色效果见《计算机工程》官网 HTML 版,下同)。

根据图 5 可以看出, IDBO 寻优精度更高、收敛速度更快,并且拥有平衡的全局探索和局部开发的能力。



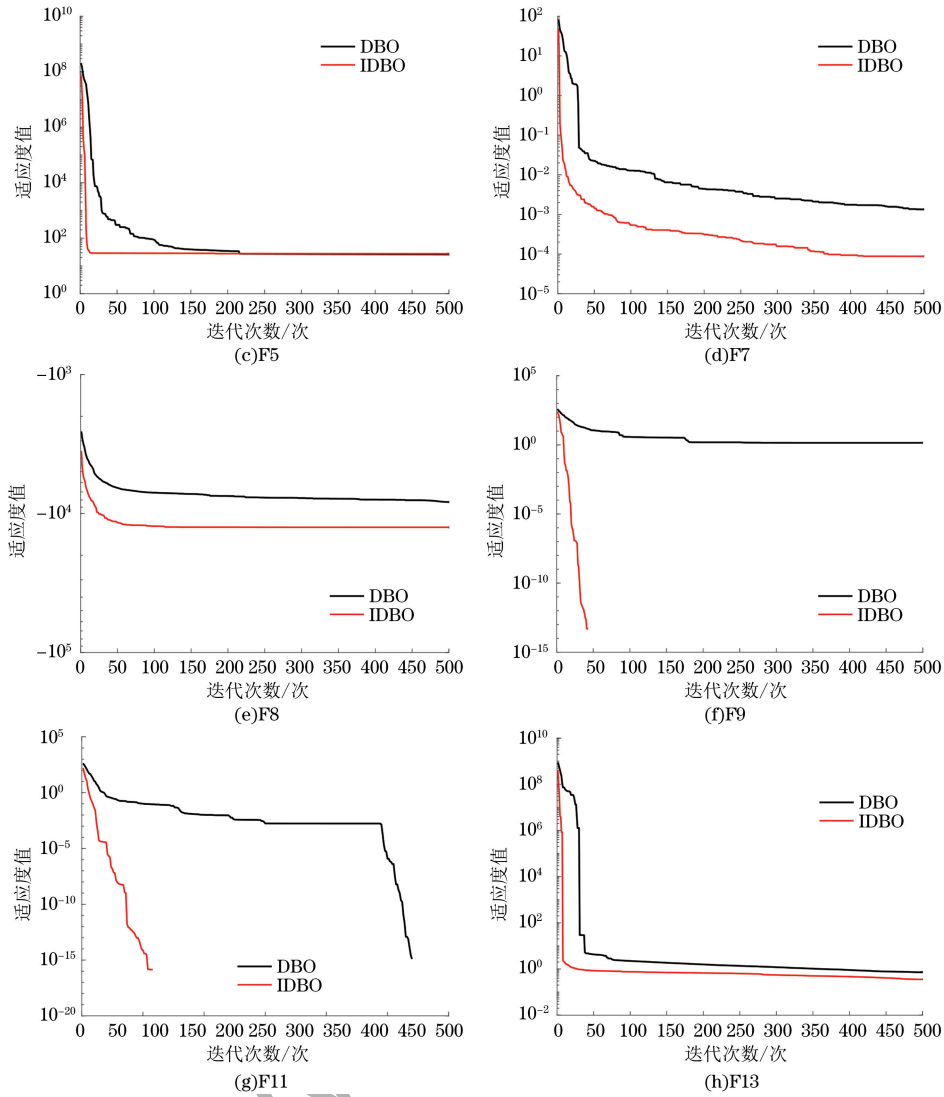


图 5 DBO 与 IDBO 的收敛过程

Fig.5 The convergence process of DBO and IDBO

2.4 基于 IDBO 优化的 DABiLSTM 时序模型

DABiLSTM 网络的预测精度主要受迭代次数、学习率和隐藏层的神经元数量的影响。本文采用 IDBO 来优化 DABiLSTM 网络。为了确定一组能使 DABiLSTM 网络误差最小的网络超参数，在本文中使用 DABiLSTM 网络的期望输出和实际输出之间的最小平均平方差作为适应度函数。优化目标包括学习率、迭代次数以及隐藏层神经元个数。IDBO-DABiLSTM 模型流程如图 6 所示，具体步骤如下：

- 1)将二次分解的风电功率信号和风速信号输入 DABiLSTM 网络。
- 2)初始化 IDBO 和 DABiLSTM 网络参数。
- 3)根据目标函数计算所有蜣螂的适应度值。
- 4)更新所有蜣螂的位置，确定当前种群最优解及其适应度值。
- 5)重复上述步骤，直到 t 满足终止准则：如果达

到，则输出全局最优解及其适应度值；否则，返回步骤 4)。

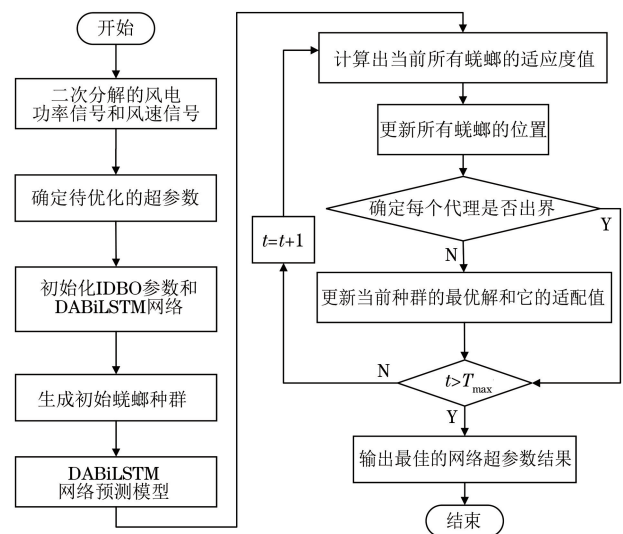


图 6 IDBO-DABiLSTM 模型的流程图

Fig.6 Flowchart of the IDBO-DABiLSTM model

6) 将得到的最优神经网络参数代入 DABiLSTM 网络进行预测。

2.5 风电功率预测模型

考虑风力发电的可变性、随机性、间歇性等特点,本文提出了一种基于二次分解和 IDBO-DABiLSTM 网络的风电预测模型。图 7 描述了风电功率预测模型的流程。

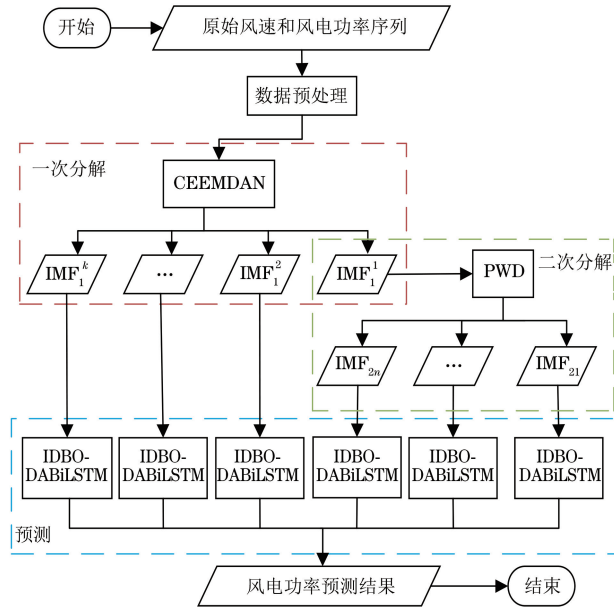


图 7 风电功率预测模型的流程图

Fig. 7 Flowchart of wind power prediction model

具体步骤如下:

1) 将原始风电功率和风速数据进行异常数据剔除和归一化等数据预处理操作。

2) 使用 CEEMDAN 对数据序列进行分解,得到 n 个不同的频率分量,然后应用 WPD 对频率最高的 IMF_1 分量进行分解,降低序列的非线性。

3) 利用 DABiLSTM 网络对二次分解的全部子序列信号进行建模。

4) 将 DABiLSTM 网络的超参数用 IDBO 进行优化,建立最佳网络结构。

5) 用确定好的模型对各子序列分别建立预测模型,然后累加所有预测值,得到最终风电功率预测结果。

3 模型参数确定

在本文模型中,CEEMDAN 的分解层数、WPD 的分解层数和时间步长会影响模型的性能。这 3 个参数将在本节中通过实验来确定。本文数据来自于贵阳某风电场的实测风电功率数据。风电场每 10 min 记录一次数据,每日记录 144 次,共统计 60 d,共计 8 640 组数据。在全部实验中,数据集中的第 1~6 480 组数据

为训练集,第 6 481~8 640 组数据为测试集。

3.1 评价指标

本文采用平均绝对误差(MAE)、均方根误差(RMSE)这 2 种误差指标来对所有模型进行定量评估和比较:

$$e_{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (18)$$

$$e_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (19)$$

式中: e_{MAE} 表示 MAE; e_{RMSE} 表示 RMSE; y_i 为预测值; \hat{y}_i 为实际值; n 为样本数。

3.2 时间步长选择

为了有效地提取特征并控制模型大小,应选择与输出功率有较强相关性的前 k 个时刻($\rho > 0.8$)进行时间序列建模^[24]。本小节计算了预测时刻 t 之前的 11 个历史时刻($t-1, t-2, \dots, t-11$)的历史风速和历史风功率的皮尔逊(Pearson)相关系数,结果如表 1 所示,可见预测时刻的前 10 个时刻与输出功率有较强相关性,因此本文时间步长确定为 10。

表 1 不同时刻的 Pearson 系数

Table 1 Pearson coefficients at different moments

历史时刻	Pearson 系数		
	历史功率/MW	历史风速/(m·s ⁻¹)	平均系数
$t-1$	0.981	0.874	0.928
$t-2$	0.958	0.853	0.906
$t-3$	0.947	0.843	0.895
$t-4$	0.927	0.831	0.879
$t-5$	0.907	0.817	0.862
$t-6$	0.898	0.807	0.853
$t-7$	0.888	0.798	0.843
$t-8$	0.876	0.786	0.831
$t-9$	0.868	0.774	0.821
$t-10$	0.857	0.765	0.811
$t-11$	0.845	0.753	0.799

3.3 CEEMDAN 分解层数选择

为了减少风力发电和风速时间序列的随机性和不稳定性,本文在一次分解中使用了 CEEMDAN。在本小节中,CEEMDAN-IDBO-DABiLSTM 组合预测模型被用来研究一次分解的最佳分解层数。实验中的分解数被设定为 2~9 层。本文将进行风力发电的多步骤预测的实验,基于已知的数据,预测未来的风电功率 $P(t+l\Delta t)$, $l=1,2,\dots,L$,其中 L 是多步骤预测的步骤数。

图 8 为 CEEMDAN-IDBO-DABiLSTM 模型在单步、两步和三步预测中不同分解层数的误差展示。由图可知,当一次分解层数设置为 7 层时,预测精度达到最高。因此,接下来的案例研究中,CEEMDAN 的

分解层数设置为 7 层。

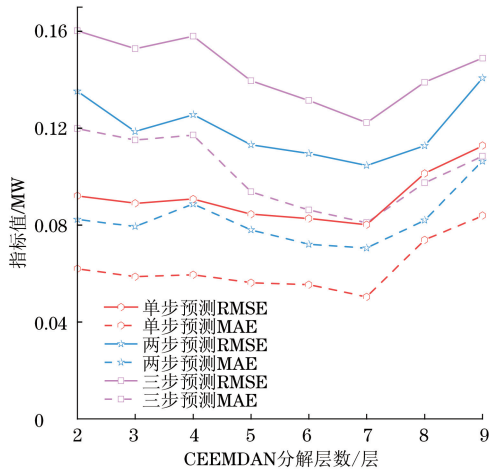


图 8 CEEMDAN 分解后的多步预测误差

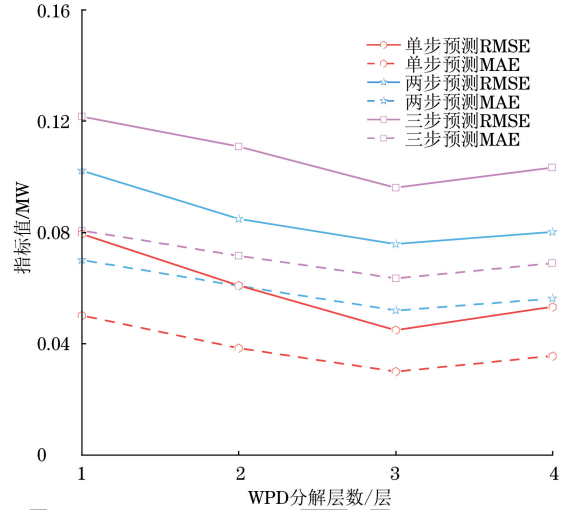


图 9 二次分解后的多步预测误差

Fig.8 Multi-step prediction error after CEEMDAN decomposition

Fig.9 Multi-step prediction error after secondary decomposition

3.4 WPD 分解层数选择

由于 CEEMDAN 产生的 IMF_1 信号具有较多的残余噪声且频率较高,难以准确预测。因此,本文在此基础上使用二次分解,通过 WPD 将 IMF_1 信号进一步分解成更多的模式分量,以提高预测效果。本小节采用 SD-IDBO-DABiLSTM 组合预测模型来评估理想的二次分解层数。实验中的分解层数被设定为 1~4 层。图 9 展示了 SD-IDBO-DABiLSTM 模型在单步、两步和三步预测中不同二次分解层数 F 的误差。

根据图 9 可知,当 WPD 的分解层数为 3 层时,预测误差最小。所以,WPD 在后续研究中的分解层数设定为 3 层。

4 实验分析

在本节中,将进行一系列相关的实验来验证 SD-IDBO-DABiLSTM 模型的准确性和有效性,并进行全面的分析。

4.1 二次分解有效性验证

为了验证二次分解模型的有效性,本文将使用一次分解(CEEMDAN)的支持向量机(SVM)^[25]、反向传播神经网络(BP)^[26]和 IDBO-DABiLSTM 预测模型分别与使用二次分解的 SVM、BPNN 和 IDBO-DABiLSTM 预测模型进行对比,以证明二次分解能有效提高预测精度,实验结果如表 2 所示。

表 2 不同分解模型的多步预测误差 1

Table 2 Multi-step prediction error 1 of different decomposition models

单位:MW

分解模型	预测模型	步长	RMSE	MAE
一次分解	CEEMDAN-SVM	单步	0.148 1	0.084 1
		两步	0.199 6	0.120 3
		三步	0.244 7	0.135 8
	CEEMDAN-BPNN	单步	0.137 2	0.079 2
		两步	0.184 5	0.105 8
		三步	0.221 2	0.122 3
CEEMDAN-IDBO-DABiLSTM	单步	0.080 2	0.050 4	
	两步	0.104 6	0.070 5	
	三步	0.127 3	0.081 2	
二次分解	SD-SVM	单步	0.081 5	0.051 3
		两步	0.143 7	0.087 8
		三步	0.181 1	0.104 6
SD-BPNN	单步	0.078 2	0.049 5	
	两步	0.135 4	0.079 6	
	三步	0.168 6	0.097 1	
SD-IDBO-DABiLSTM	单步	0.044 9	0.031 2	
	两步	0.075 9	0.052 1	
	三步	0.096 1	0.063 5	

根据表 2 可知,无论基准模型是 SVM、BPNN 还是 IDBO-DABiLSTM,使用了二次分解的组合模型预测精度都更高。例如,在单步、两步和三步的预测中,SD-BPNN 的 RMSE 分别比 CEEMDAN-BPNN 低 43.0%、26.6%和 23.8%,同时 MAE 降低了 37.5%、24.8%和 20.6%。虽然二次分解后的子序列仍然会有高频分量,但是 WPD 也会将带有较多噪声、频率较高的 IMF₁ 信号分解成为更平稳的中低频信号,这些二次分解后的子序列降低了风电功率预测难度。实验结果表明,对 IMF₁ 信号进行二次分解能够有效地提高预测模型的性能,即二次分解有效。

4.2 IDBO-DABiLSTM 有效性验证

为了验证 IDBO-DABiLSTM 模型的有效性,本文将 SD-IDBO-DABiLSTM 与 SD-SVM、SD-BPNN、SD-BiLSEM 和 SD-DABiLSTM 进行了对比实验,结果如表 3 所示。此外,图 10 为测试集中 100 个点的多步预测结果,其中对部分样本点的结果进行了放大。

表 3 不同分解模型的多步预测误差 2

Table 3 Multi-step prediction error 2 of different

decomposition method		单位:MW	
预测模型	步长	RMSE	MAE
SD-SVM	单步	0.081 5	0.051 3
	两步	0.143 7	0.087 8
	三步	0.181 1	0.104 6
SD-BPNN	单步	0.078 2	0.049 5
	两步	0.135 4	0.079 6
	三步	0.168 6	0.097 1
SD-BiLSTM	单步	0.077 1	0.049 1
	两步	0.118 6	0.070 7
	三步	0.141 3	0.087 8
SD-DABiLSTM	单步	0.054 8	0.036 3
	两步	0.088 6	0.057 2
	三步	0.113 1	0.072 4
SD-IDBO-DABiLSTM	单步	0.044 9	0.031 2
	两步	0.075 9	0.052 7
	三步	0.096 1	0.063 5

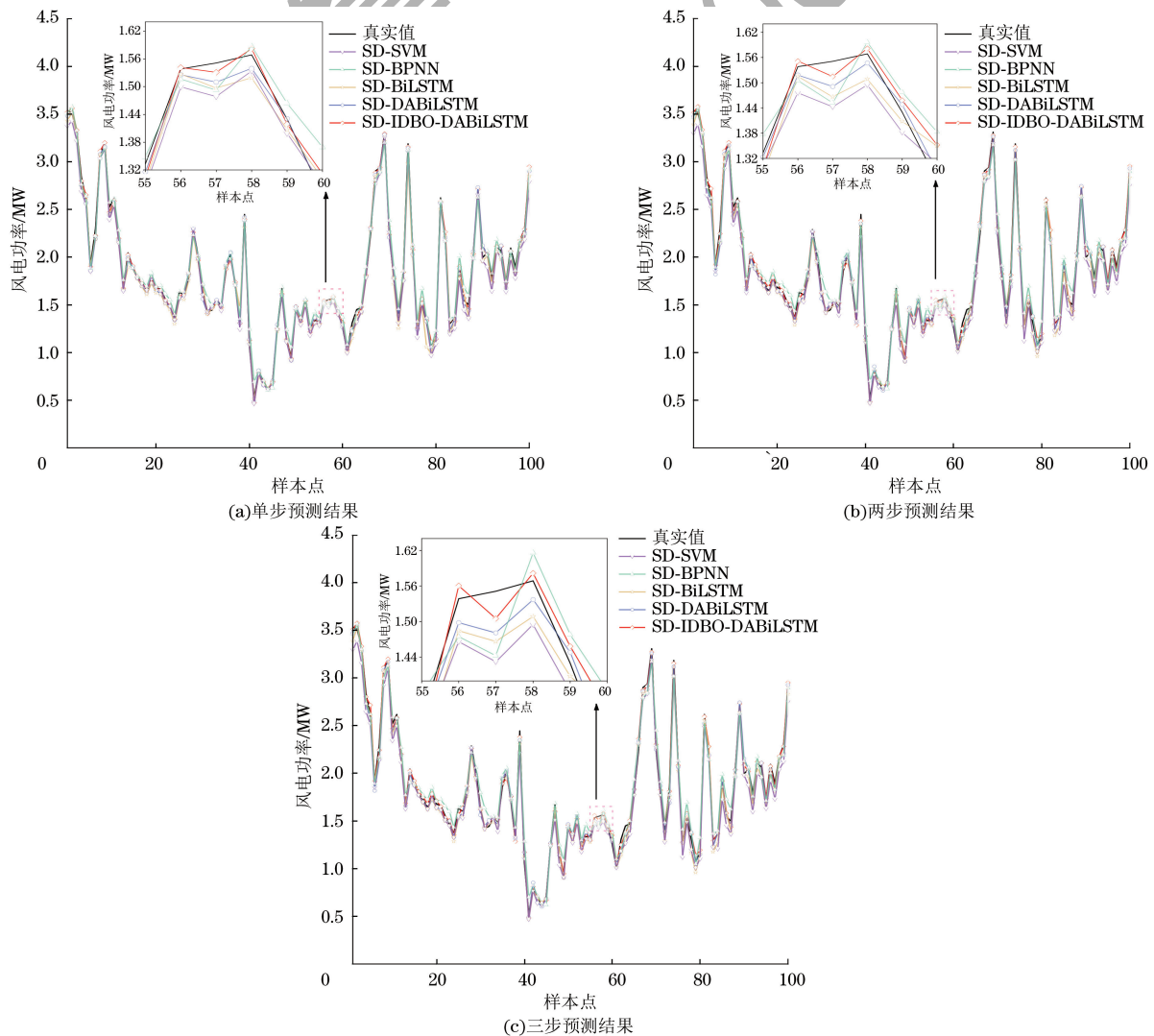


图 10 多步预测结果图

Fig. 10 Chart of multi-step prediction results

从表 3 和图 10 可知:

1)SD-DABiLSTM 模型能够有效地提取时间序列的整体特征,突出关键输入信息的重要性,获取时间序列的长时间依赖性,提高预测精度。与 SD-BiLSTM 相比,SD-DABiLSTM 在单步、两步和三步预测中的 RMSE 分别降低了 28.9%、25.3% 和 19.9%。此外,SD-DABiLSTM 的 MAE 下降了 26.1%、19.1%和 17.5%。在前向 LSTM 前加入特征注意力机制,有效地提取了输出与输入变量间的隐藏关联信息;在后向 LSTM 后加入时间注意力机制,深入抓取预测序列的时间相关性,提高了风电功率预测精度。

2)SD-IDBO-DABiLSTM 模型可以有效地克服人工选择网络超参数带来的误差,提高模型的预测精度。与 SD-DABiLSTM 相比,在单步、两步和三步预测中,SD-IDBO-DABiLSTM 的 RMSE 分别降低了 18.1%、14.3%和 15.0%。同时,MAE 也分别下降了 14.0%、7.9%和 12.3%。加入 IDBO 算法后,避免了模型陷入局部最优解,最小化预测误差,有效地提升了网络的整体性能。

3)SD-IDBO-DABiLSTM 模型的预测性能优于其他模型,其单步、两步和三步预测的误差均低于其他模型。对比实验结果可以看出,二次分解、双重注意力机制和改进蜣螂优化算法对网络的预测效果均有提升。将二次分解、DABiLSTM 和 IDBO 结合后,可以得到最佳的性能。以 SD-SVM 和 SD-BPNN 为例,SD-IDBO-DABiLSTM 的 RMSE 分别在一步、两步与三步的预测中降低了 44.9%和 42.6%、44.9%和 43.9%、46.9%和 43.0%。同时,MAE 减少了 39.2%和 37.0%、40.0%和 33.8%、39.3%和 34.6%。

4)SD-IDBO-DABiLSTM 模型整体预测性能较好,但是在波峰、波谷等处的预测误差较大,对于转折性时段的预测能力不足,在后续应深入对降低转折性时刻预测误差的研究。

5 结束语

本文提出了一种基于二次分解、DABiLSTM 和 IDBO 的新型组合短期风电功率预测模型。首先,使用 CEEMDAD 分解原始风速和风电功率序列,并对高频特征模式函数 IMF_1 使用 WPD 分解为更平稳的子序列,降低了原始信号的波动性与随机性。其次,为了提取时间序列的关键重要信息和整体特征,挖掘时间序列间的时间信息,引入双重注意力机制优化 BiLSTM 网络。此外,提出了一种 IDBO 算

法,增加了模型的收敛能力,提高并平衡了全局搜索和局部挖掘能力,并用其智能确定了 DABiLSTM 网络的最佳超参数,避免网络陷入局部最优。实验结果表明,本文模型具备有效性,相比其他模型,该模型具有更优的预测性能,为时间序列的研究提供了一定的思路。下一步将对风电场集群的波动性进行研究,并探索预测结果的可靠性,定量地评估预测结果的可靠性。

参考文献

- [1] 钱政,裴岩,曹利宵,等. 风电功率预测方法综述[J]. 高电压技术, 2016, 42(4): 1047-1060.
QIAN Z, PEI Y, CAO L X, et al. Review of wind power forecasting method[J]. High Voltage Engineering, 2016, 42(4): 1047-1060. (in Chinese)
- [2] 廖雪超,伍杰平,陈才圣. 结合注意力机制与 LSTM 的短期风电功率预测模型[J]. 计算机工程, 2022, 48(9): 286-297, 304.
LIAO X C, WU J P, CHEN C S. Short-term wind power prediction model combining attention mechanism and LSTM [J]. Computer Engineering, 2022, 48(9): 286-297, 304. (in Chinese)
- [3] WANG G, JIA R, LIU J H, et al. A hybrid wind power forecasting approach based on Bayesian model averaging and ensemble learning[J]. Renewable Energy, 2020, 145: 2426-2434.
- [4] 唐新姿,顾能伟,黄轩晴,等. 风电功率短期预测技术研究进展[J]. 机械工程学报, 2022, 58(12): 213-236.
TANG X Z, GU N W, HUANG X Q, et al. Progress on short term wind power forecasting technology[J]. Journal of Mechanical Engineering, 2022, 58 (12): 213-236. (in Chinese)
- [5] 赵凌云,刘友波,沈晓东,等. 基于 CEEMDAN 和改进时间卷积网络的短期风电功率预测模型[J]. 电力系统保护与控制, 2022, 50(1): 42-50.
ZHAO L Y, LIU Y B, SHEN X D, et al. Short-term wind power prediction model based on CEEMDAN and an improved time convolutional network [J]. Power System Protection and Control, 2022, 50(1): 42-50. (in Chinese)
- [6] 殷豪,董朕,孟安波. 基于 VMD-SE-IPSO-BNN 的超短期风电功率预测[J]. 电测与仪表, 2018, 55(2): 45-51.
YIN H, DONG Z, MENG A B. Ultra short-term wind power forecasting based on VMD-SE-IPSO-BNN [J]. Electrical Measurement & Instrumentation, 2018, 55 (2): 45-51. (in Chinese)
- [7] 王晓兰,李辉. 基于 EMD 分解的风电场风速和输出功率年度预测[J]. 太阳能学报, 2011, 32(3): 301-306.
WANG X L, LI H. Annual forecasting of wind speed and power in wind farm based on EMD[J]. Acta Energeiae Solaris Sinica, 2011, 32(3): 301-306. (in Chinese)
- [8] 王世谦,苏娟,杜松怀. 基于小波变换和神经网络的短期风电功率预测方法[J]. 农业工程学报, 2010, 26(S2): 125-129.
WANG S Q, SU J, DU S H. A method of short-term wind power forecast based on wavelet transform and neural network [J]. Transactions of the Chinese Society of Agricultural Engineering, 2010, 26 (S2): 125-129. (in Chinese)
- [9] 赵征,汪向硕. 基于 CEEMD 和改进时间序列模型的超短期风功率多步预测[J]. 太阳能学报, 2020, 41(7): 352-358.
ZHAO Z, WANG X S. Ultra-short-term multi-step wind power prediction based on ceemd and improved time series

- model[J]. *Acta Energaie Solaris Sinica*, 2020, 41(7): 352-358. (in Chinese)
- [10] 周盛山, 汤占军, 王金轩, 等. EEMD 和 CNN-XGBoost 在风电功率短期预测的应用研究[J]. *电子测量技术*, 2020, 43(22): 55-61.
ZHOU S S, TANG Z J, WANG J X, et al. Application of EEMD and CNN-XGBoost in short-term wind power prediction[J]. *Electronic Measurement Technology*, 2020, 43(22): 55-61. (in Chinese)
- [11] 丁婷婷, 杨明, 于一潇, 等. 基于误差修正的短期风电功率集成预测方法[J]. *高电压技术*, 2022, 48(2): 488-496.
DING T T, YANG M, YU Y X, et al. Short-term wind power integration prediction method based on error correction [J]. *High Voltage Engineering*, 2022, 48(2): 488-496. (in Chinese)
- [12] 周洪煜, 曾济贫, 王照阳, 等. 基于混沌 DNA 遗传算法与 PSO 组合优化的 RNN 短期风电功率预测[J]. *电力系统保护与控制*, 2013, 41(2): 144-149.
ZHOU H Y, ZENG J P, WANG Z Y, et al. Ridgelet neural network model for short-term wind power forecasting based on the combination of chaos DNA genetic and particle swarm optimization algorithm [J]. *Power System Protection and Control*, 2013, 41(2): 144-149. (in Chinese)
- [13] 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 Q M, LI H Y, WANG Z Q, et al. Short-term wind power forecasting based on LSTM [J]. *Power System Technology*, 2017, 41(12): 3797-3802. (in Chinese)
- [14] 高鹭, 孔繁苗, 张飞, 等. 基于 IPSO-BiLSTM-AM 模型的超短期风电功率预测方法[J]. *智慧电力*, 2022, 50(4): 27-34.
GAO L, KONG F M, ZHANG F, et al. Ultra short-term wind power prediction method based on IPSO-BiLSTM-AM model[J]. *Smart Power*, 2022, 50(4): 27-34. (in Chinese)
- [15] SHAHID F, ZAMEER A, MUNEEB M. A novel genetic LSTM model for wind power forecast [J]. *Energy*, 2021, 223: 120069.
- [16] CHEN Y F, ZHAO H, ZHOU R, et al. CNN-BiLSTM short-term wind power forecasting method based on feature selection [J]. *IEEE Journal of Radio Frequency Identification*, 2022, 6: 922-927.
- [17] 陈申, 叶小岭, 熊雄, 等. 基于天鹰优化算法的短期风电功率区间预测[J]. *重庆理工大学学报*, 2023, 37(8): 304-314.
- [18] DEVI A S, MARAGATHAM G, BOOPATHI K, et al. RETRACTED ARTICLE: hourly day-ahead wind powerforecasting with the EEMD-CSO-LSTM-EFG deep learning technique [J]. *Soft Computing*, 2020, 24 (16): 12391-12411.
- [19] YANG X S, DEB S. Cuckoo search via Lévy flights[C]// *Proceedings of the World Congress on Nature & Biologically Inspired Computing (NaBIC)*. Washington D. C., USA: IEEE Press, 2009: 210-214.
- [20] DING Y F, CHEN Z J, ZHANG H W, et al. A short-term wind power prediction model based on CEEMD and WOA-KELM[J]. *Renewable Energy*, 2022, 189: 188-198.
- [21] MIRJALILI S, LEWIS A. The whale optimization algorithm [J]. *Advances in Engineering Software*, 2016, 95: 51-67.
- [22] 武新章, 梁祥宇, 朱虹谕, 等. 基于 CEEMDAN-GRA-PCC-ATCN 的短期风电功率预测[J]. *山东大学学报(工学版)*, 2022, 52(6): 146-156.
WU X Z, LIANG X Y, ZHU H Y, et al. Short-term wind power prediction based on CEEMDAN-GRA-PCC-ATCN[J]. *Journal of Shandong University (Engineering Science)*, 2022, 52(6): 146-156. (in Chinese)
- [23] XUE J K, SHEN B. Dung beetle optimizer: a new meta-heuristic algorithm for global optimization[J]. *The Journal of Supercomputing*, 2023, 79(7): 7305-7336.
- [24] LI J T, GENG D, ZHANG P, et al. Ultra-short term wind power forecasting based on LSTM neural network [C]// *Proceedings of the IEEE 3rd International Electrical and Energy Conference (CIEEC)*. Washington D. C., USA: IEEE Press, 2019: 1815-1818.
- [25] 魏鹏飞, 樊小朝, 史瑞静, 等. 基于改进麻雀搜索算法优化支持向量机的短期光伏发电功率预测[J]. *热力发电*, 2021, 50(12): 74-79.
WEI P F, FAN X C, SHI R J, et al. Short-term photovoltaic power generation forecast based on improved sparrow search algorithm optimized support vector machine [J]. *Thermal Power Generation*, 2021, 50(12): 74-79. (in Chinese)
- [26] 王金锋, 杨宇琦, 温栋, 等. 基于 GA-BP 和 RBF 的风力发电时间序列混沌预测组合模型[J]. *电网与清洁能源*, 2022, 38(11): 117-125.
WANG J F, YANG Y Q, WEN D, et al. A combined model of chaos prediction of wind power generation time series based on GA-BP and RBF [J]. *Power System and Clean Energy*, 2022, 38(11): 117-125. (in Chinese)