基于特征优选的 DBO-CNN-BiLSTM-AM 短期光伏发电预测

杜 立[1,2]，李振华[1,2]，李振兴[1]，魏 伟[3]，徐艳春[1]
(1. 三峡大学 电气与新能源学院，湖北 宜昌 443002；2. 智慧能源技术湖北省工程研究中心，湖北 宜昌 443002；3.国网湖北省电力有限公司营销服务中心 计量中心，湖北 武汉 430080)

摘 要：随着光伏装机容量逐渐增大,电力系统稳定、安全运行面临挑战,通过提高光伏发电预测精度,可以提高电力系统运行的稳定性和安全性。为此,文章提出一种基于特征优选的 DBO-CNN-BiLSTM-AM 短期光伏发电预测方法。首先,利用方差膨胀因子(Variance Inflation Factor,VIF)和极端随机树(Extremely Randomized Trees,ERT) 计算并选取 VIF 值低且重要性得分高的特征作为输入；其次，采用卷积神经网络(Convolutional Neural Network,CNN)、双向长短期神经网络(Bi-directional Long Short Term Memory,BiLSTM)和具有特征权重动态调整能力的注意力机制(Attention Mechanism,AM)建立 CNN-BiLSTM-AM 预测模型;最后,利用蜣螂优化算法(Dung Beetle Optimization algorithm,DBO)进行超参数全局寻优,以最大化预测精度。实验结果表明,文章提出的短期光伏发电功率预测方法预测精度明显高于传统方法,为优化电力调度和电力系统稳定运行提供了良好的基础。

关键词：光伏发电；双向长短期神经网络；注意力机制；蜣螂优化算法
中图分类号：TK514；TM615 **文献标志码**：A **文章编号**：1671-5292(2025)11-1458-11

0 引言

全球能源结构正逐步由传统的化石能源向可再生能源转变。2023 年,我国可再生能源新增装机容量为 2.4 亿 kW,约占全球可再生能源新增装机容量的 50%以上[1]。新能源装机容量的增长,给电力系统稳定运行和市场建设带来挑战，准确地预测光伏发电功率对提高系统的稳定性和电力市场运营有重要意义。

按预测时间的尺度，光伏发电功率预测可分为超短期预测(0~24 h)、短期预测(24~72 h)、中长期预测(1~12 月)[2]。与中长期预测相比,超短期和短期预测的时间跨度更短，预测结果往往具有更高的时效性和准确性。短期光伏功率预测广泛运用传统的机器学习方法,如随机森林、支持向量机等，这些方法能获取负荷与影响因素中的非线性关系,但在功率预测建模过程中,过分依赖相似日样本选取和聚类分析[3]。文献[4]利用卷积神经网络捕捉空间特征结合长短期记忆网络(Long Short-Term Memory,LSTM)建立组合模型,加强了梯度的稳定性,但会忽略各特征的重要程度,导致模型预测精度下降。文献[5]采用双向长短期神经网络 (Bi-directional Long Short Term Memory，BiLSTM) 和注意力机制 (Attention Mechanism，AM)进行光伏发电预测,能学习完整的上、下文信息,但无法充分捕捉到序列中的局部特征。然而，光伏发电站采集到的环境因素数据具有高度的非线性和时序性，现有方法对数据的时序特征和局部特征提取仍然不充分，尤其是不能动态地调整特征的权重。组合预测模型涉及到众多超参数,增加了调参的难度和过拟合的风险，简单地使用各子模型的默认参数会使组合模型的预测性能下降。现有方法大多只关注学习模型的组合,较少有利用先进算法对组合模型超参数进一步优化以达到模型组合之间的最优适配。

综上，本文提出一种基于特征优选的 DBO-CNN-BiLSTM-AM 光伏发电功率短期预测模型。利用基于方差膨胀因子 (Variance Inflation Factor,VIF)和极端随机树(Extremely Randomized Trees,ERT)的特征优选方法选取重要性得分高且 VIF 值低的特征子集作为模型的输入；使用卷积

收稿日期：2024-08-21。
基金项目：国家自然科学基金(52277012)；武汉强磁场学科交叉基金资助(WHMFC202202)。
通信作者：李振华(1986-),男,博士,教授,研究方向为智能电网、电力设备状态检测与故障诊断。E-mail:lizhenhua1993@163.com

神经网络（Convolutional Neural Network,CNN)和注意力机制(Attention Mechanism,AM)捕捉光伏数据的空间特征、局部特征及动态地调整不同特征的权重；通过双向长短期神经网络（Bi-directional Long Short Term Memory,BiLSTM)捕获时间序列的全局信息及长期依赖关系；利用蜣螂算法(Dung Beetle Optimization algorithm,DBO)优化模型超参数,达到组合模型之间的最优适配。实验结果证明，本文预测方法相比于传统预测方法具有更高的精度。

1　基于 VIF-ERT 的特征选取

1.1　方差膨胀因子

VIF 法[6]是一种用于评估模型输入变量之间线性关系的方法。通过计算光伏发电数据中每个特征与其他特征之间的 VIF 值,来衡量每个特征的多重共线性程度。其计算方法是将每个特征 X_i 作为因变量,将其他特征 $X_{j(j\neq i)}$作为自变量来拟合一个线性回归模型,计算式为

$$X_i=\beta_0+\sum_{j\neq i}\beta_j X_j+\varepsilon \tag{1}$$

式中:ε 为误差项;β_0 为截距项;β_j 为自变量X_j 的系数。

再计算因变量 X_i 与自变量X_j 的均方误差比值,即:

$$V_i=\frac{1}{1-R_i^2} \tag{2}$$

式中:R_i^2 为线性回归模型中，特征 X_i 与其他变量之间关系的拟合优度;V_i 为 VIF 值。

如果一个特征的 VIF 值大于某个设定阈值,特征 X_i 则能被其他自变量很好地解释,即存在严重的多重共线性问题，影响模型的解释力和预测性能。本文实验设置 VIF 值的阈值为 10,VIF 值超过 10 的特征视为存在严重的多重共线性问题,则对该特征进行剔除。为满足特征量之间分析的完整性,下一步进行极端随机树重要性分析时,仍对所有特征量进行重要性打分,但不满足 VIF 值要求的特征量不得作为模型的输入。

1.2　极端随机树

Geurts P[7]在树的集成学习方法基础上提出了 ERT 方法。ERT 通过多数投票机制(Bagging)将所有决策树的分类结果进行集成,完成最终决策。相比于传统随机森林,ERT 在构建决策树时，不仅在特征选择上具有随机性，在划分阶段的阈值选择上也引入了随机性。这一特性使得 ERT 在构建决策模型时具有更高的多样性和鲁棒性。用于光伏发电数据环境特征重要性评估的 ERT 构造步骤[8]如下。

①将预处理后的光伏发电数据样本集 T 用作各基分类器的训练数据,其中,样本数量为 S,特征数量为 M。

②使用分类与回归决策树(Classification And Regression Tree,CART）算法构建基分类器,针对每一个节点，从 M 个特征中选取 k 个特征,形成候选特征集,再计算各节点的基尼系数(Gini index,Gini)I_{Gin} 以及分裂后子节点的加权基尼系数之和。选择使分裂后子节点的加权基尼系数之和最小的特征及其对应的分裂点作为该节点的分裂标准,对分裂后的数据子集重复步骤②,直至生成完整的决策树。Gini 的计算式为

$$I_{\mathrm{Gin}}\left\{P=\sum_{n=1}^{N}[P_n(1-P_n)]\right\}=1-\sum_{i=1}^{n}P_n^2 \tag{3}$$

式中:P_n 为第 n 个类别的样本占比,满足 $\sum P_n=1$。

③重复执行步骤①和②,构建大量决策树,形成一个决策树的集合,即生成极端随机树。每棵决策树独立地对输入的光伏发电数据样本集进行分析，通过 Bagging 方法将所有决策树的分类结果进行集成,最后根据式(4)做出最终决策,即得到的 y 为对光伏发电数据进行特征分析的结果。

$$y=\underset{i=1,2,\cdots,n}{\operatorname{argmax}}\sum_{m=1}^{M}\lambda(y_m=i) \tag{4}$$

式中:y 为决策树最终分类决策结果;λ 为决策树的权重;n 为光伏发电数据类别总量。

本文实验对所有特征进行极端随机树训练,获得其重要性得分,筛选满足 VIF 阈值前提下重要性得分较高的特征量作为预测模型的输入。

2　DBO-CNN-BiLSTM-AM 模型的构建

2.1　卷积神经网络(CNN)

卷积神经网络是一种具备局部连接和权值共享特性的神经网络。CNN 由卷积层、池化层和全连接层构成,如图 1 所示。卷积层利用卷积核对输入的多特征复杂数据进行卷积运算，提取数据的潜在特征,该层作用最为关键;池化层采用最大池化的方式对输入数据进行降采样(Down sampling),以减少特征图的大小,从而降低计算

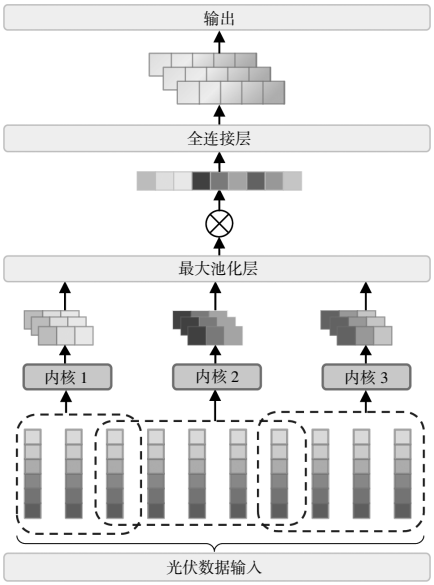


图 1 CNN 的网络结构

Fig.1 CNN network architecture

复杂度和参数量；全连接层将特征图中的高阶特征信息转换为可供输出层使用的形式，与输出层相连，用于最终的分类或回归预测。卷积层计算公式[9]为

$$x_j^l=\sigma\left(\sum_{i\in M_j}x_i^{l-1}\times\boldsymbol{W}_{ij}^l+\boldsymbol{b}_j^l\right) \tag{5}$$

式中：x_j^l为第 l 层第 j 个卷积核的输入，也是第 l–1 层的输出；σ 为激活函数；M_j 为输入映射层的集合；$\boldsymbol{W}_{ij}^l$为卷积核的权重矩阵；$\boldsymbol{b}_j^l$为偏置向量。

全连接层的正向传播输出公式为

$$x_j^{l+1}=\sum_{i=1}^{n}\boldsymbol{W}_{ij}^l x_j^l+\boldsymbol{b}_j^l \tag{6}$$

式中：$x^{l+1(j)}$为第 l+1 层第 j 个输出神经元值。

2.2 双向长短期记忆网络 BiLSTM

虽然 CNN 具有从光伏发电数据中自动提取多维度空间特征的能力，但对于时间依赖性较强的时序数据处理能力较差；LSTM 网络通过引入 3 个门控单元解决了长期依赖问题，LSTM 网络的具体计算式[10]为

$$\begin{cases}f_t=\sigma\{\boldsymbol{W}_f\cdot[h_{t-1},x_t]+\boldsymbol{b}_f\}\\ i_t=\sigma\{\boldsymbol{W}_i\cdot[h_{t-1},x_t]+\boldsymbol{b}_i\}\\ \widetilde{C}_t=\tanh\{\boldsymbol{W}_c\cdot[h_{t-1},x_t]+\boldsymbol{b}_c\}\\ \boldsymbol{C}_t=f_t\otimes\boldsymbol{C}_{t-1}\oplus i_t\otimes\boldsymbol{C}_t\\ o_t=\sigma\{\boldsymbol{W}_o\cdot[h_{t-1},x_t]+\boldsymbol{b}_o\}\\ h_t=o_t\otimes\tanh\boldsymbol{C}_t\end{cases} \tag{7}$$

式中：f_t,i_t,o_t 分别为遗忘门、输入门和输出门的输出值；$\boldsymbol{W}_f,\boldsymbol{W}_i,\boldsymbol{W}_c,\boldsymbol{W}_o$ 和 $\boldsymbol{b}_f,\boldsymbol{b}_i,\boldsymbol{b}_c,\boldsymbol{b}_o$ 分别为遗忘门、输入门、当前细胞状态和输出门的权重矩阵和偏置向量；tanh 为双曲正切激活函数。

由式(7)可知，通过对数据按照从前到后的训练得到 LSTM 网络的参数，该训练方式无法充分提取光伏发电数据在时间序列中内在的特征联系[11]。BiLSTM 由前向和后向 LSTM 网路组成，能够同时捕获数据的正向和反向时序特征，并高效挖掘当前时刻数据与过去及未来时刻数据之间的内在联系，以提高数据的有效利用率和预测精准度。BiLSTM 网络结构如图 2 所示，其计算式[12]为

$$\begin{cases}\overrightarrow{h_t}=\mathrm{LSTM}(x_t,\overrightarrow{h_{t-1}})\\ \overleftarrow{h_t}=\mathrm{LSTM}(x_t,\overleftarrow{h_{t-1}})\\ y_t=\sigma(\boldsymbol{W}_y\cdot[\overrightarrow{h_t},\overleftarrow{h_t}]+\boldsymbol{b}_y)\end{cases} \tag{8}$$

式中：LSTM 为传统 LSTM 单元的计算过程，用于更新隐藏状态；$\overrightarrow{h_t}$为 t 时刻前向 LSTM 的隐藏状态；$\overleftarrow{h_t}$为 t 时刻后向 LSTM 的隐藏状态；$\boldsymbol{W}_y$ 为用来线性组合隐藏层状态的权重矩阵；$\boldsymbol{b}_y$ 为偏置向量。

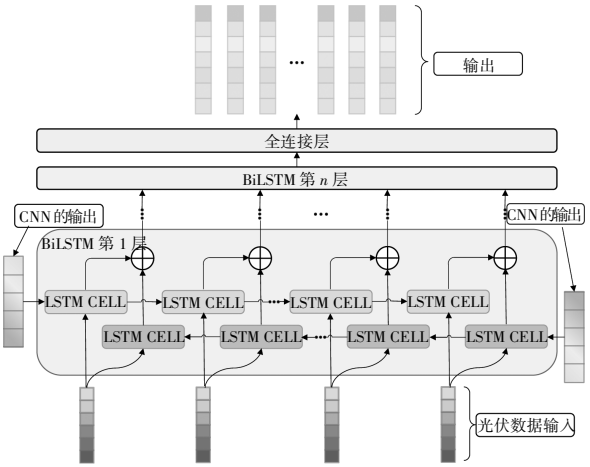


图 2 BiLSTM 网络结构

Fig.2 BiLSTM network architecture

2.3 注意力机制(AM)

由于 BiLSTM 的整合方式受限于门控机制，无法充分捕捉到全局信息之间的复杂关系；CNN 网络采用固定的权重分配方式进行特征整合，训练时会忽略各通道信息的重要程度；AM 动态地调整对不同特征信息的关注度，更灵活地集中注意力于局部重要特征以解决上述问题。AM 的实现通常包含以下步骤：将输入序列 $x=[x_1,x_2,\cdots,x_n]$

进行编码,生成查询向量 $\boldsymbol{q}$;再对输入序列之间相似度$[s_1,s_2,\cdots,s_n]$进行打分,打分函数计算式为

$$s(\boldsymbol{x}_i,\boldsymbol{q})=\boldsymbol{V}^{\mathrm{T}}\tanh(\boldsymbol{W}\boldsymbol{x}_i+\boldsymbol{U}\boldsymbol{q}) \tag{9}$$

式中:$\boldsymbol{V},\boldsymbol{W},\boldsymbol{U}$ 均为参数待学习的矩阵;$\boldsymbol{x}_i$ 为示键向量。

式(9)为加性注意力打分公式。

再使用 Softmax 函数进行归一化,得到每组输入向量的关注权重 a_i,计算式为

$$a_i=\mathrm{softmax}[s(\boldsymbol{x}_i,\boldsymbol{q})]=\frac{\exp[s(\boldsymbol{x}_i,\boldsymbol{q})]}{\sum_{i=1}^{n}\exp[s(\boldsymbol{x}_i,\boldsymbol{q})]} \tag{10}$$

最后,根据所有值向量的相关关注权重进行加权求和,得到最终光伏发电特征的分类输出 e。

$$e=\sum_{i=1}^{n}a_i \tag{11}$$

2.4　蜣螂优化算法(DBO)

蜣螂优化算法是 2022 年底提出的一种基于昆虫蜣螂行为的新型启发式优化算法。该算法模拟蜣螂的滚球、跳舞、觅食、偷窃和繁殖行为,并将这 5 种行为转化为优化算法的搜索策略,根据蜣螂的位置变化,进行局部利用和全局寻优以解决组合预测模型超参数庞杂的问题。DBO 算法各行为计算式[13]如下。

①蜣螂的滚球(无障碍向前)行为和蜣螂的跳舞(遇到障碍无法向前)行为位置计算式分别为

$$x_i^{t+1}=x_i^{t}+ak x_i^{t-1}+b\left|x_i^{t}-x_{\mathrm{worst}}^{t}\right| \tag{12}$$

$$x_i^{t+1}=x_i^{t}+\tan\theta\left|x_i^{t}-x_i^{t-1}\right| \tag{13}$$

式中:x_i^t 为第 t 次迭代第 i 个蜣螂的位置信息;a 取-1 或 1;$k\in(0,0.2]$;$b\in(0,1]$;x_{worst}^t 为全局最差位置;$\left|x_i^t-x_{\mathrm{worst}}^t\right|$ 为模拟光照强度的变化量,其值越大,光源越弱;偏转角 $\theta\in[0,\pi]$,当 θ 取 0,$\pi/2$ 或 π 时,蜣螂的位置保持不变;$\left|x_i^t-x_i^{t-1}\right|$ 为第 i 只蜣螂在第 t 次迭代与第 t-1 次迭代之间的位置偏移量。

②蜣螂的觅食行为和偷窃行为位置变化式分别为

$$x_i^{t+1}=x_i^{t}+C_1(x_i^{t}-L_{b'})+\boldsymbol{C}_2\left|x_i^{t}-U_{b'}\right| \tag{14}$$

$$x_i^{t+1}=x_{\mathrm{lbest}}^{t}+Q\boldsymbol{g}\left(\left|x_i^{t}-x_{\mathrm{gbest}}^{t}\right|+\left|x_i^{t}-x_{\mathrm{lbest}}^{t}\right|\right) \tag{15}$$

式中:$L_{b'}$和 $U_{b'}$分别为觅食区域的下限和上限;C_1 为服从正态分布的随机数;$\boldsymbol{C}_2$ 为大小为 (0,1)的 $1\times d$ 的随机向量;x_{lbest}^t 为最佳食物来源;Q 为常数值;$\boldsymbol{g}$ 为服从正态分布的 $1\times d$ 的随机向量。

③采用边界选择策略来模拟蜣螂雌性繁殖区域,其计算式为

$$x_i^{t+1}=x_{\mathrm{gbest}}^{t}+b_1(x_i^{t}-L_{b^*})+b_2\left|x_i^{t}-U_{b^*}\right| \tag{16}$$

式中:x_{gbest}^t 为全局最佳位置;U_{b^*} 和 L_{b^*} 分别为繁殖区域的上限和下限。

优化后的光伏功率预测模型比单一神经网络模型增强了表征能力,但存在超参数过多的问题,增加了过拟合的风险。本文采用具有强大全局搜索能力和快速收敛能力的蜣螂优化算法对 CNN-BiLSTM-AM 模型进行超参数全局寻优,对该模型中 BiLSTM 层神经元数量、CNN 层卷积核大小、正则化系数、学习率初始值、时间步长进行全局寻优,其变化数值分别设为[4,256],[1,7],[0.01,0.5],[0.000 1,1],[1,100]。经多次实例分析验证和 DBO 算法训练,最终确定 DBO-CNN-BiLSTM-AM 模型的主要超参数如表 1 所示。

表 1　DBO-CNN-BiLSTM-AM 模型超参数

Table 1 DBO-CNN-BiLSTM-AM model hyperparameters

参数	数值	参数	数值
种群数量/个	30	优化变量维度	5
滚球蜣螂/个	6	BiLSTM 层神经元数量/个	128
繁殖蜣螂/个	6	卷积核大小	3
小蜣螂/个	7	正则化系数	0.079
偷盗蜣螂/个	11	学习率	0.001 2
最大迭代次数/次	500	时间步长	10

2.5　基于 DBO 算法优化的 CNN-BiLSTM-AM 预测模型

针对提取局部特征困难、动态调整输入特征权重受限以及超参数配置繁琐等问题,在 BiLSTM 预测模型的基础上,本文提出了一种基于 DBO 算法优化 CNN-BiLSTM-AM 的光伏发电功率短期预测模型,该模型的预测流程见图 3。

由于光伏发电数据包含复杂关联且维度高的环境因素,首先,要进行预处理和归一化,并划分训练集和测试集;其次,利用方差膨胀因子和极端随机树计算特征的 VIF 值和重要性得分,筛选出

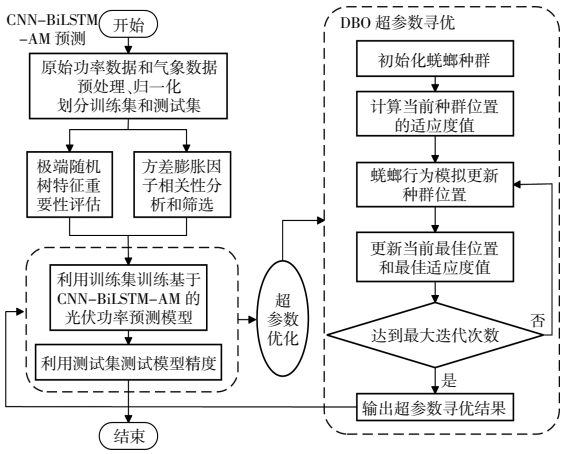


图 3 组合模型预测程序框图

Fig.3 Composite model prediction program flowchart

VIF 值低且重要性得分高的特征作为输入;再利用 DBO 算法为 CNN-BiLSTM-AM 模型筛选出最优超参数;最后,利用测试集测试模型的预测精度。

CNN-BiLSTM-AM 模型网络结构如图 4 所示。其具体预测步骤如下:利用卷积操作捕获输入数据的空间信息并提取局部特征,将提取的特征信息作为 BiLSTM 网络的初始状态嵌入;BiLSTM 网络同时从前向和后向对序列数据进行训练和建模,有效提取时序信息和全局信息,并捕捉长期依赖关系;利用注意力机制,加权整合 CNN 层和 BiLSTM 层提取的特征,动态灵活地关注输入数据中的重要特征信息,最终经 Softmax 激活函数和全连接层输出预测结果。

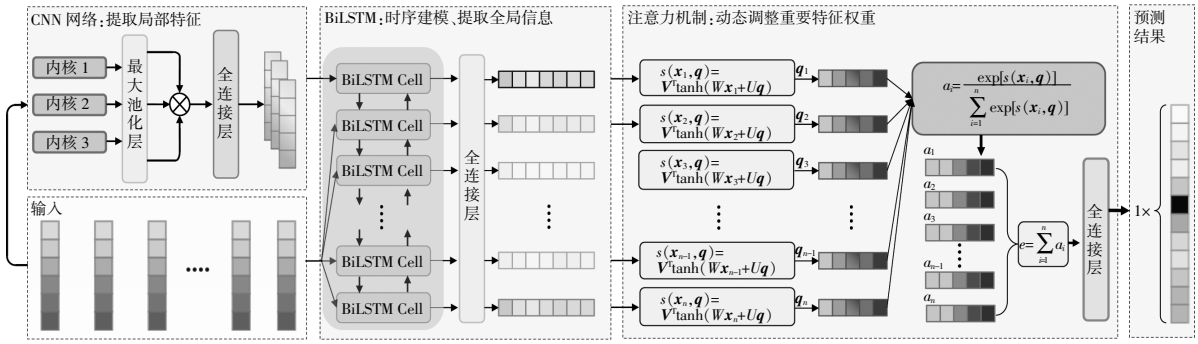


图 4 CNN-BiLSTM-AM 组合模型结构

Fig.4 The combined structure of CNN-BiLSTM-AM model

3 实验与分析

3.1 数据描述及预处理

本文实验数据取自北美加利福利亚州(California)某水平光伏发电站 2017 年 6 月-2018 年 10 月的实测数据,数据集包括辐照度、环境温度、湿度、云层高度、风速、大气压强、季节等 16 列光伏发电数据,装机容量为 1 800 kW,时间分辨率为 15 min[14]。由于光伏发电设备和数据采集设备发生故障或定期维护,光伏发电数据中存在部分缺失值和异常值,剔除缺失值所在的全天数据,共得到 462 d 的数据;再通过箱线图法计算光伏发电数据的 4 分位数和箱线的位置进行异常值检测,并将检测出的异常值替换为前后时刻数据的平均值;最后利用 Min-Max 方法将处理后的数据线性缩放到[0,1]内,并按 8:2 划分训练集和测试集,Min-Max 归一化计算式为

$$X_{\text{norm}}=\frac{X-X_{\min}}{X_{\max}-X_{\min}} \tag{17}$$

式中:X_{norm} 为标准化处理后数据;X 为待标准化数据;$X_{\min}$ 为输入数据最小值;$X_{\max}$ 为输入数据最大值。

本文实验预测任务为滑动窗口单步预测,对未来 72 h 的光伏发电功率进行预测。以预测 t 时刻的光伏发电功率为例,设置输入时间步长为 a,则模型的输入为 $t-a$,$t-a-1$,$t-a-2$,……,$t-2$,$t-1$ 时刻的环境特征和光伏发电功率数据,各时刻的输入量也随着预测时刻的变化而往后推移。经反复实验和 DBO 算法的优化,最终选择输入时间步长为 10,即将待预测时间前 30 d 的环境特征和光伏发电功率数据输入至光伏功率预测模型,在该步长下模型的预测效果最优。

3.2 误差评估指标

为评估预测结果的有效性,本文选取绝对误差(Mean Absolute Error,MAE)e_{MA}、均方根误差(Root Mean Square Error,RMSE)e_{RMS} 和拟合系数 R^2 为误差评估指标,这 3 个评估指标[15]的公式为

$$e_{MA}=\frac{1}{n}\sum_{i=1}^{n}\left|y_i-\hat{y}_i\right| \tag{18}$$

$$e_{RMS}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2} \tag{19}$$

$$R^2=1-\frac{\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i-\frac{1}{n}\sum_{i=1}^{n}y_i\right)^2} \tag{20}$$

式中:n 为样本数量;y_i 为光伏发电功率实际值;$\hat{y}_i$ 为光伏发电功率预测值。

3.3　基于 VIF 值和极端随机树的输入变量选取

由于影响光伏出力的环境因素繁多冗杂,本文使用方差膨胀因子和极端随机树对环境因素进行降维。

表 2 为各环境变量的 VIF 值和 ERT 重要性得分。VIF 值的阈值设置为 10,超过 10 的特征存在严重多重共线性问题,如能见度、大气压强、经度、纬度和季节的 VIF 值均超过 10,不能作为训练模型的输入量; 使用 ERT 对环境变量进行训练,重要性得分越高,则在模型中具有更强的解释能力,对模型的贡献越大,辐照度和环境温度的重要性得分较高,能见度、大气压强、云高、降水量、经度、纬度和海拔的重要性得分较低,其中海拔、风速、降水量和云层高度这 4 个特征量虽满足 VIF 值的阈值,但因重要性得分过低,也不能作为模型的输入。VIF 值低且重要性得分高的特征量与光伏发电功率相关性高且不存在多重共线性问题,因此,选取辐照度、环境温度、湿度、风速以及光伏发电功率作为后续的研究对象。

表 2　光伏发电功率各环境因素评估

Table 2 Evaluation of environmental factors of photovoltaic power generation

环境变量	VIF 值	重要性得分	环境变量	VIF 值	重要性得分
辐照度	3.238	0.838 9	风速	8.596	0.114 4
海拔	5.325	0.054 1	能见度	24.531	0.009 7
湿度	3.861	0.213 8	大气压强	101.659	0.064 4
环境温度	5.164	0.440 0	云层高度	5.994	0.050 1
经度	48.392	0.010 4	纬度	48.730	0.024 6
降水量	7.676	0.007 2	季节	36.834	0.078 8

3.4　预测模型消融实验

为验证本文所提出的 DBO-CNN-BiLSTM-AM 模型的精准度和优越性, 设计了预测模型消融实验。为确保实验的可比性,每组实验分别使用相同的输入数据, 且选取同一时刻的预测结果进行对比。为探究 DBO 算法使各子模型达到最优适配后性能的提升、卷积神经网络与注意力机制对高度非线性化和时序性光伏数据特征提取的必要性、双向神经网络模型较于传统模型的优势,预测模型消融实验选用 DBO-CNN-BiLSTM-AM, CNN-BiLSTM-AM, BiLSTM 以及 GRU 模型进行对比。对 DBO-CNN-BiLSTM-AM 模型, 除使用 DBO 算法优化的 5 个超参数外, 其余参数与 CNN-BiLSTM-AM, CNN-BiLSTM-AM 和 BiLSTM 模型参数一致,GRU 模型与 BiLSTM 使用相同的 Dropout 率,GRU 中 RNN 层与 BiLSTM 层数均设置为 3,各层模型主要超参数如表 3 所示。

表 3　对比模型主要参数

Table 3 The main hyperparameters of the comparison model

名称	参数	
全局参数	时间步长 timesteps=10	损失函数 Lossfunction=MAE
	训练周期数 epochs=100	正则化系数 regularizer=0.1
	学习率 lr=0.0012	批处理参数 BatchSize=64
	优化器 Optimizer=Adam	输入特征维度 INPUT_DIMS=5
DBO 算法主要参数	种群数量 pop=30	优化变量维度 dim=5
	偏转系数 k=0.1	偏转参数 b=0.3
	蜣螂每次移动的距离 s=0.5	最大迭代次数 Max_iter=500
CNN 层主要参数	卷积核大小 kernel_size=1	激活函数 activation='relu'
	过滤器数量 filters=64	drouput=0.3
BiLSTM 层主要参数	BiLSTM 层数为 3	双向特性 return_sequences=True
	BiLSTM 层神经元数量 lstm_units=64	drouput=0.3

续表 3

名称	参数
AM 层主要参数	AM 层激活函数 activation='Softmax' Lambda 层:SINGLE_ATTENTION_VECTOR=False
GRU 主要参数	隐藏层单元数 hidden_size=20 RNN 层数 num_layers=3 drouput=0.3

预测模型对比实验具体预测误差见表 4。

表 4 各模型光伏发电功率预测误差

Table 4 Prediction errors of photovoltaic power generation for each model

模型	MAE/kW	RMSE/kW	R^2/%
DBO-CNN-BiLSTM-AM	14.149	19.447	98.96
CNN-BiLSTM-AM	16.739	22.828	98.04
BiLSTM	20.197	26.799	96.49
GRU	25.630	34.158	95.69

预测模型对比实验的光伏功率预测曲线和误差结果如图 5 和图 6 所示。

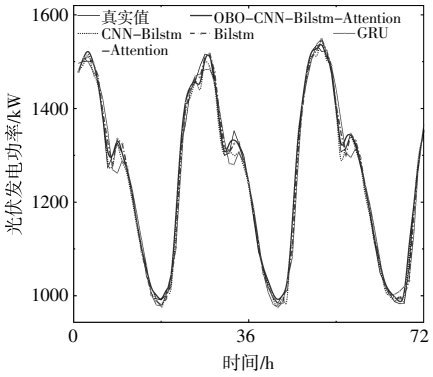


图 5 预测模型对比实验各模型光伏功率预测结果

Fig.5 Prediction model comparison experiment comparison of photo-voltaic power prediction results of each model

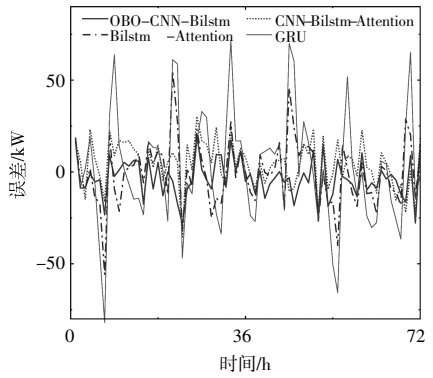


图 6 预测模型对比实验各模型光伏功率预测误差

Fig.6 Prediction model comparison experiment comparison of photo-voltaic power prediction errors of each model

由图 5、图 6 和表 4 可知，对比 BiLSTM 和 GRU 模型,GRU 预测曲线拟合效果相对较差,尤其是在拐点处,BiLSTM 模型的 MAE 降低了 21.20%,RMSE 降低了 21.54%,R^2 提高了 0.8%,这说明，面对高度非线性和时序性的光伏发电数据,BiLSTM 具有更强的表征能力，同时也减少了梯度消失与梯度爆炸问题，更好地解决了拐点波动大的问题,BiLSTM 能双向学习，具有更强的记忆能力与适应性。对于 CNN-BiLSTM-AM 模型与 BiLSTM 模型的预测结果,单一预测模型面对复杂的光伏数据的拟合效果较差，对光伏特征的提取不充分。加入卷积神经网络和注意力机制后，CNN-BiLSTM-AM 模型预测结果的精准性得到明显提升，误差曲线中的极差较大幅度下降,MAE 降低了17.12%,RMSE 降低了 14.81%,R^2 提高了 1.55%。CNN 网络提取输入特征的局部特征并压缩,使BiLSTM 层能更好地获取上、下文信息,注意力机制能灵活地关注隐层状态，更深层次地挖掘光伏数据的时间序列属性，同时注意力机制能有效地减少历史信息的流失并突出权重较高的特征信息，以减少冗余信息对光伏发电功率预测的影响。与CNN-BiLSTM-AM 模型的结果相比,本文所提出DBO-CNN-BiLSTM-AM 模型拟合效果与真实结果非常相近，在拐点处与梯度不变的位置均精度极佳,其 MAE 降低了 15.47%,RMSE 降低了 14.81%,R^2 提高了 0.92%。DBO 算法在迭代初期就能在全局范围搜索到目标区域并避免许多局部最优解,迭代后期在目标区域快速收敛至最优解,获得子模型间相互适配的最优超参数。在预测模型消融实验中，本文提出的 DBO-CNN-BiLSTM-AM 模型预测效果最佳，误差折线波动最小,误差值最低,相比于传统方法更具有优越性。

3.5 优化算法对比实验

为了衡量和评估 DBO 算法的优化性能与泛化能力,优化算法对比实验将 DBO 算法对照目前主流算法，如麻雀优化算法（Sparrow Search

Algorithm,SSA)、灰狼优化算法 (Grey Wolf Optimizer,GWO) 和粒子群算法 (Particle Swarm Optimization,PSO),对比实验各算法主要参数见表 5。

表 5　优化算法对比实验各算法主要参数

Table 5 The optimization algorithm compares the main parameters of each algorithm in the experiment

算法名称	主要参数
蜣螂优化算法 DBO	种群数量 pop=30　偏转系数 k=0.1 偏转参数 b=0.3 蜣螂每次移动的距离 s=0.5
麻雀优化算法 SSA	安全阈值 ST=0.6　发现者比例 PD=0.7 警戒者比例 SD=0.2　种群规模 PS=30 探索因子 EF=1.0　学习因子 LF=1.5
灰狼优化算法 GWO	灰狼种群数量 PS=30　收敛因子 $C\in[0,2]$ 领导者比例 LR=0.2
粒子群算法 PSO	粒子数量 PS=30　位置更新因子 c_1=1.5 社会更新因子 c_2=1.5　惯性权重 w=0.9

为保证优化算法对比实验的可行性,在其他条件保持一致的情况下,选用 9 个测试函数进行验证试验,测试函数均为 CEC2005 基准测试函数,维度均为 5,基准测试函数如表 6 所示。

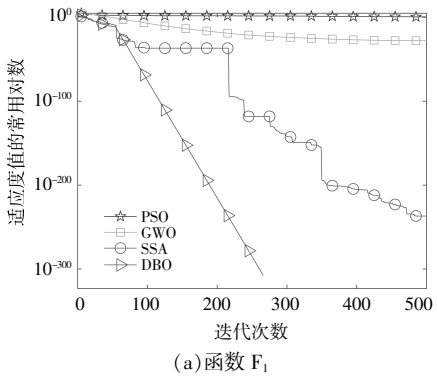
各算法在不同单峰测试函数下的迭代变化趋势如图 7 所示,同时对各算法的参数进行合理设置,最大迭代次数均为 500,保持各算法种群数量均为 30。

由图 7 可见,对比 SSA,GWO 和 PSO 算法,本文中的 DBO 算法对 9 个基准测试函数的最优适应度明显更低,并且收敛速度最快,具有更好的超参数寻优性能。DBO 算法是一个基于 SI 的无梯度优化器,根据蜣螂不同的行为特征,衍生出了不同的搜索模式;利用不同搜索模式中不同时间段的信息,深入探索目标空间、蜣螂产卵区和最佳觅食区等不同区域的搜索策略。相比其他算法具

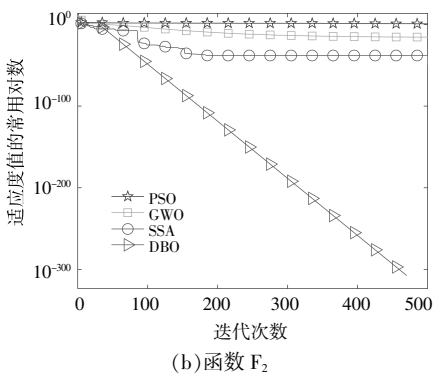
表 6　单峰基准测试函数

Table 6 Single peak benchmark test function

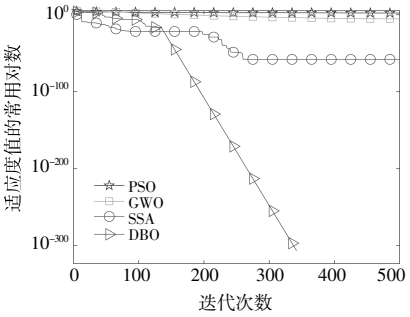
函数名称	基准测试函数表达式	维度	搜索范围	全局最优解
F_1	$f(x)=\sum_{i=1}^{n}x_i^2$	5	[-100,100]	0
F_2	$f(x)=\sum_{i=1}^{n}\lvert x_i\rvert+\prod_{i=1}^{n}\lvert x_i\rvert$	5	[-10,10]	0
F_3	$f(x)=\max\{\lvert x_i\rvert,1\leqslant i\leqslant n\}$	5	[-100,100]	0
F_4	$f(x)=\sum_{i=1}^{n}[(x_i+0.5)]^2$	5	[-100,100]	0
F_5	$f(x)=\sum_{i=1}^{n}ix_i^4+\mathrm{random}[0,1)$	5	[-1.28,1.28]	0
F_6	$f(x)=\sum_{i=1}^{n}\left[100\left(x_{i+1}-x_i^2\right)^2+(x_i-1)^2\right]$	5	[-30,30]	0
F_7	$f(x)=\sum_{i=1}^{n}\left[x_i^2-10\cos(2\pi x_i)+10\right]$	5	[-5.12,5.12]	0
F_8	$f(x)=-20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2}\right)-\exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos 2\pi x_i\right)+20+e$	5	[-32,32]	0
F_9	$f(x)=\frac{1}{4\,000}\sum_{i=1}^{n}x_i^2+\prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right)+1$	5	[-600,600]	0



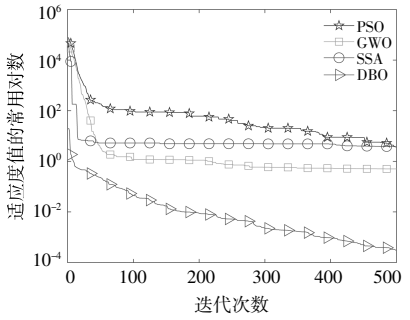
(a)函数 F_1



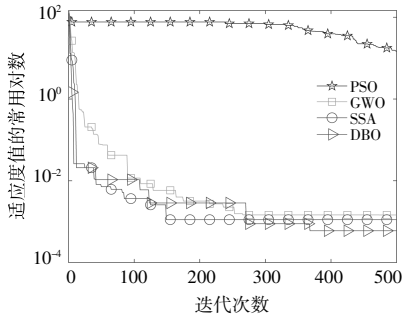
(b)函数 F_2



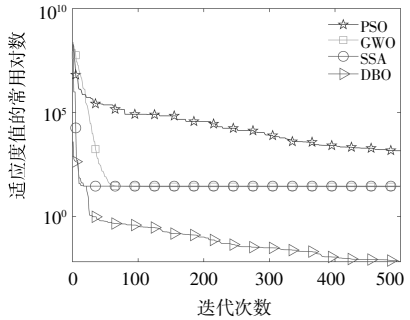
(c)函数 F_3



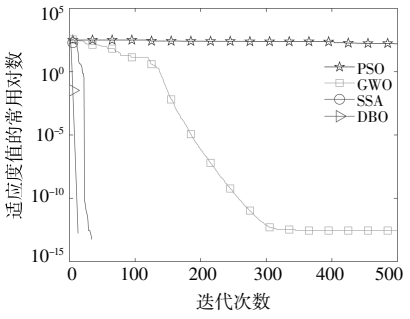
(d)函数 F_4



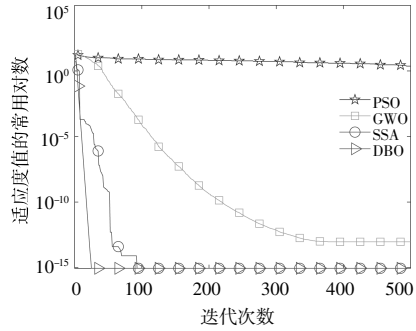
(e)函数 F_5



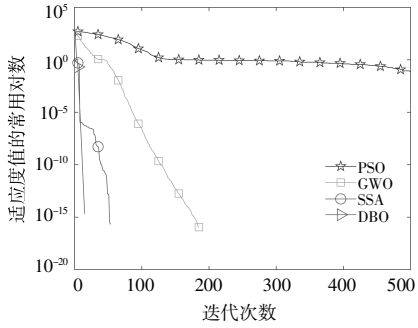
(f)函数 F_6



(g)函数 F_7



(h)函数 F_8



(i)函数 F_9

图 7 不同优化算法的最优适应度迭代曲线

Fig.7 Optimal fitness iterative curves of different optimization algorithms

有更强的搜索能力、更快的收敛速度,且能避免陷入局部最优[16]。

为进一步验证 DBO 算法对优化本文提出模型超参数的有效性与优越性,各算法统一在CNN-BiLSTM-AM 模型的基础上,使用上述算法对 BiLSTM 层神经元数量、CNN 层卷积核大小、正则化系数、学习率、步长这 5 个超参数进行优化,并将优化后的超参数结果输入模型进行训练,再与测试集结果进行对比,预测曲线如图 8 所示。

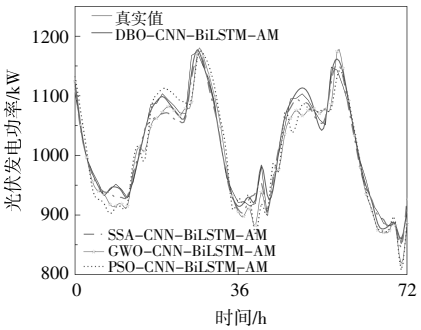


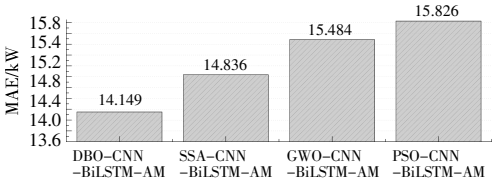
图 8 各种优化算法对比实验结果

Fig.8 Comparison of experimental results of different prediction models

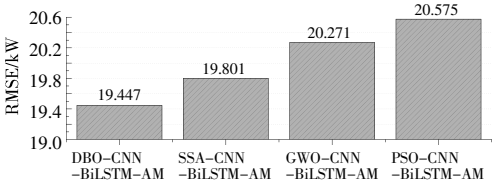
由图 8 可见,通过模拟蜣螂的觅食、繁殖等行

为，利用不同搜索模式进行更强的全局搜索,经 DBO 算法优化参数后的模型预测曲线更接近真实值,拟合曲线更加平滑,使模型超参数之间的适配更加合理，本文所提方法表现出更好的鲁棒性和可靠性。GWO 和 PSO 对预测模型超参数优化效果相对较差，在收敛速度上明显慢于 DBO,寻优性能相对较差,SSA 优化模型超参数后的效果得到一定提升，在部分测试函数上收敛速度略次于 DBO,但 SSA 设置初始参数较为繁琐,对参数设置也更为敏感，运用在超参数繁多的预测模型上的效果不如 DBO,DBO 算法对预测模型超参数寻优效果相较于其他主流算法更具优越性。

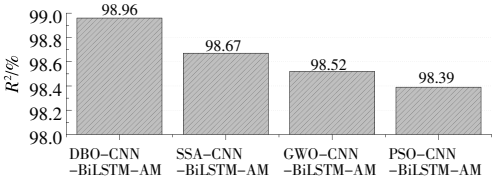
测试集中 MAE,RMSE,R^2 如图 9 所示。



(a)不同优化算法模型的 MAE



(b)不同优化算法模型的 RMSE



(c)不同优化算法模型的 R^2

图 9　各优化算法模型预测误差对比图

Fig.9 The comparison chart of prediction errors for various optimization algorithm models

由图 9 可知,4 种不同算法优化模型的 MAE 分别是 14.149,14.836,15.484,15.826 kW,RMSE 分别是 19.447,19.801,20.271,20.575 kW,DBO 算法优化模型超参数后的 MAE 和RMSE 最低。

4　结论

本文提出了一种基于特征优选的 DBO-CNN-BiLSTM-AM 短期光伏发电预测方法。在 BiLSTM 模型的基础上,引入了卷积神经网络和注意力机制,并利用 DBO 算法对组合模型超参数进行寻优,以公开数据集进行了对比实验分析,得出以下结论。

①利用方差膨胀因子和极端随机树方法对输入环境特征进行降维,筛选出 VIF 值低且重要性得分高的特征,降低了模型输入变量的维度,解决了环境特征之间的冗余性问题和传统特征分析方法无法评估非线性关系和多重共线性关系等问题。

②提出了 CNN-BiLSTM-AM 组合预测模型,利用 CNN 提取空间信息和局部特征，并引入注意力机制以动态地调整输入特征的权重,充分利用对预测重要性影响较大的特征。经实验证明,该模型在短期光伏发电预测上具有良好的预测性能。

③使用 DBO 算法对 CNN-BiLSTM-AM 预测模型的超参数进行寻优，建立模型消融实验与不同优化算法对比实验。将 DBO 与 SSA,GWO 和 PSO 算法进行对比，实验结果表明,DBO 算法优化模型超参数方面效果最好,全局搜索能力最强,收敛性最快,算法参数配置简单,解决了传统算法易陷入局部最优、收敛速度慢、超参数配置繁琐等问题。

参考文献:

[1] 胡臻达,黄夏楠,曾振松,等.基于碳夹点分析法的可再生能源优化调度方法研究[J].可再生能源,2023,41(8):1095-1099.

[2] 陈骏嚎,张娜,刘广忱,等.基于 EOSSA-ELM 的光伏短期输出功率预测[J].可再生能源,2022,40(7):890-898.

[3] 任建吉,位慧慧,邹卓霖,等.基于 CNN-BiLSTM-Attention 的超短期电力负荷预测 [J]. 电力系统保护与控制,2022,50(8):108-116.

[4] Zhou X,Sheil B,Suryasentana S,et al.Multi-fidelity fusion for soil classification via LSTM and multi-head self-attention CNN model [J].Advanced Engineering Informatics,2024,62:102655-102655.

[5] 张家安,郝峰,董存,等.基于两阶段不确定性量化的光伏发电超短期功率预测 [J]. 太阳能学报,2023,44(1):69-77.

[6] 马良玉,程东炎,梁书源,等.基于 LightGBM-VIF-MIC-SFS 的风电机组故障诊断输入特征选择方法[J].热力发电,2024,53(1):154-164.

[7] Geurts P,Ernst D,Wehenkel L.Extremely randomized trees[J].Machine Learning,2006,63(1):3-42.

[8] 徐先峰,李芷菡,刘状壮,等.基于半监督学习标签传播-极端随机树算法的光伏阵列故障诊断及定位[J].电网技术,2023,47(3):1038-1047.
[9] 张红颖,贾一超,汪江志,等.考虑注意力机制的 CNN-LSTM 高渗透风电并网暂态电压稳定性研究[J].可再生能源,2024,42(3):398-406.
[10] 方鹏,高亚栋,潘国兵,等.基于 LSTM 神经网络的中长期光伏电站发电量预测方法研究 [J]. 可再生能源,2022,40(1):48-54.
[11] 孟宏宇, 张建良, 蔡兆龙, 等. 基于 CNN-BiLSTM-Attention 的直流微电网故障诊断研究[J].中国电机工程学报,2025,45(4):1369-1380.
[12] 薛阳,张宁,俞志程,等.基于 BiLSTM 和 Bootstrap 方法的风电功率区间预测[J].可再生能源,2020,38(8):1059-1064.
[13] 汪茹康,王健,徐钢,等.基于 SG-IDDBO-VMD 的风光互补复合波动平抑策略[J].中国电机工程学报,2025,45(16):6314-6325.
[14] Pasion C,Wagner T,Koschnick C, et al.Machine learning modeling of horizontal photovoltaics using weather and location data [J].Energies,2020,13(10):2570.
[15] 武新章,王泽宇,代伟,等.基于异质聚类与 Stacking 的双集成光伏发电功率预测[J].电网技术,2023,47(1):275-284.
[16] Jiankai X,Bo S .Dung beetle optimizer:a new meta-heuristic algorithm for global optimization[J].The Journal of Supercomputing,2022,79(7):7305-7336.

DBO-CNN-BiLSTM-AM short-term photovoltaic power generation prediction based on feature optimization

Du Li[1,2], Li Zhenhua[1,2], Li Zhenxing[1], Wei Wei[3], Xu Yanchun[1]
(1.College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443002, China; 2.Hubei Provincial Engineering Research Center of Intelligent Energy Technology, Yichang 443002, China; 3.State Grid Hubei Electric Power Co.,Ltd., Marketing Service & Measurement Center, Wuhan 430080, Chian)

Abstract: With the gradual increase of photovoltaic installed capacity, the stable and safe operation of power system is facing challenges. By improving the prediction accuracy of photovoltaic power generation, the operation stability and safety of power system can be improved. Therefore, this paper proposes a DBO-CNN-BiLSTM-AM short-term photovoltaic power generation prediction method based on feature optimization.Firstly, Variance Inflation Factor (VIF) and Extremely Randomized Trees (ERT) are used to calculate and select the features with low VIF value and high importance score as input. Secondly, a CNN-BiLSTM-AM prediction model is established by using Convolutional Neural Network (CNN), Bi-directional Long Short Term Memory (BiLSTM) and Attention Mechanism (AM) with dynamic adjustment ability of feature weights. Finally, the Dung Beetle Optimization algorithm (DBO) is used to perform global optimization of hyperparameters to maximize the prediction accuracy.The experimental results show that the prediction accuracy of the short-term photovoltaic power prediction method proposed in this paper is significantly higher than that of the traditional method, which provides a good basis for optimizing power dis-patching and stable operation of power system.

Keywords: photovoltaic power generation; BiLSTM; attention mechanism; dung beetle optimization algorithm