

CS 算法优化 VMD-BiLSTM-AM 的光伏功率预测^①



俞 敏, 王晓霞

(华北电力大学 计算机系, 保定 071003)
通信作者: 俞 敏, E-mail: whale1028@126.com

摘 要: 针对光伏发电功率的波动性与随机性对调度部门的负荷预测以及电网安全运行带来的严峻挑战, 提出了一种基于变分模态分解 (VMD) 和布谷鸟搜索 (CS) 算法优化的双向长短期记忆网络 (BiLSTM) 光伏发电功率预测方法. 首先使用 VMD 将光伏功率序列分解成不同频率的子模态, 通过皮尔逊相关性分析确定影响各模态的关键气象因子. 其次分别构建注意力机制 (AM) 和 BiLSTM 混合的光伏发电功率预测模型, 利用 CS 算法获取网络最优的权重和阈值. 最后, 将不同模态的预测结果相叠加, 得到最终的预测结果. 通过对亚利桑那州地区光伏电站输出功率进行预测, 验证了所提模型的有效性.

关键词: 双向长短期记忆网络; 变分模态分解; 布谷鸟搜索; 注意力机制; 光伏功率预测

引用格式: 俞敏, 王晓霞. CS 算法优化 VMD-BiLSTM-AM 的光伏功率预测. 计算机系统应用, 2023, 32(2): 347-355. <http://www.c-s-a.org.cn/1003-3254/8921.html>

Photovoltaic Power Prediction Based on VMD-BiLSTM-AM Optimized by CS Algorithm

YU Min, WANG Xiao-Xia

(Department of Computer, North China Electric Power University, Baoding 071003, China)

Abstract: For the severe challenges brought by the fluctuation and randomness of photovoltaic power generation to the load prediction of the dispatching department and the safe operation of the power grid, this study proposes a photovoltaic power prediction method of bidirectional long short-term memory (BiLSTM) optimized by variational modal decomposition (VMD) and cuckoo search (CS) algorithm. Firstly, VMD is employed to decompose the photovoltaic power sequence into sub-modes with different frequencies, and Pearson correlation analysis is adopted to determine the key meteorological factors affecting each mode. Secondly, the hybrid photovoltaic power prediction models of attention mechanism (AM) and BiLSTM are constructed, and the CS algorithm is utilized to obtain the optimal weight and threshold of the network. Finally, the prediction results of different modes are superimposed to obtain the final prediction results. The effectiveness of the proposed model is verified by predicting the output power of photovoltaic power stations in Arizona.

Key words: bidirectional long short-term memory (BiLSTM); variational modal decomposition (VMD); cuckoo search (CS); attention mechanism; photovoltaic power prediction

随着有限化石燃料日益枯竭和碳减排需求转趋刚性, 光伏发电等可再生能源发电技术发展迅猛^[1]. 然而, 受太阳辐照度和气象影响的光伏输出功率的变化也给

电网的运行带来了严峻的挑战. 随着光伏渗透率逐年递增, 准确预测光伏发电功率可以有效提高电网实时控制性能和经济效益^[2].

① 收稿时间: 2022-06-21; 修改时间: 2022-07-18; 采用时间: 2022-07-29; csa 在线出版时间: 2022-09-08
CNKI 网络首发时间: 2022-11-16

为提高光伏发电功率预测精度,学者们展开了大量研究.统计模型通过时间序列法^[3]、回归分析法^[4]、神经网络法^[5]等建立历史数据与输出功率的映射关系估算发电量,因建模简单,广泛应用于光伏发电功率短期预测中.其中深度神经网络因泛化能力强、能深入挖掘时间序列间的关联性,为光伏发电功率预测提供了强大支撑.文献[6]通过量化天气类型与主成分分析(PCA),使用长短期神经网络(long short-term memory, LSTM)网络提高光伏预测精度.文献[7]采用PCA对相似样本提取主成分作为神经网络的输入,简化网络结构.然而上述预测方法只考虑了历史信息对时序预测的单向影响,忽视了未来数据信息流与当前时刻的双向联系.双向长短时记忆(bidirectional long short-term memory, BiLSTM)网络通过双向LSTM提取时序数据的正反向特征来挖掘时序数据的前后联系,能有效提高光伏功率的预测精度^[8].

日前光伏序列属于非平稳序列,针对其随机性强、表现出复杂的日周期性等特性,文献[9,10]分别使用经验小波变换与变分模态分解(variational mode decomposition, VMD)将原始非平稳光伏序列分解成相对平稳的子序列,并对各子序列分别预测后再组合,有效提高预测精度.

此外,日前光伏预测作为多步预测问题,使用深度神经网络训练与预测不仅需要占用大量建模时间和计算资源,还容易陷入局部最优,因此需对神经网络的权重和阈值进行优化.文献[11]将相似日理论和蝙蝠算法优化最小二乘支持向量机LSSVM相结合,避免LSSVM陷入局部最优从而提高预测精度.文献[12]用遗传算法对BP神经网络的连接权值进行优化,以达成对网络的改进.但上述优化算法需要参数较多且耗时长,布谷鸟搜索(cuckoo search, CS)算法参数更少且具有较好的参数智能寻优性能^[13].因此,本文提出CS算法优化的基于变分模态分解和添加注意力机制(attention mechanism, AM)的BiLSTM日前光伏发电功率预测模型(VMD-BiLSTM-AM-CS).首先,针对光伏序列波动性较大,使用VMD将日前光伏序列分解为相对平稳的子模态,结合皮尔逊相关系数分析法选取影响各模态的关键气象因子共同作为预测模型的输入.其次,使用BiLSTM充分学习时序数据的正反向特征,添加注意力机制,赋予网络隐藏层单元不同的权重,使得预测模型从长时间序列输入中关注重点信息.最后,采用

CS算法对网络层的权系数与偏置进行优化,解决神经网络易陷入局部最优的问题.实验结果表明,本文所提模型有效提高日前光伏预测精度.

1 日前光伏功率预测模型构建

VMD-BiLSTM-AM-CS日前光伏功率预测模型如图1所示,步骤如下.

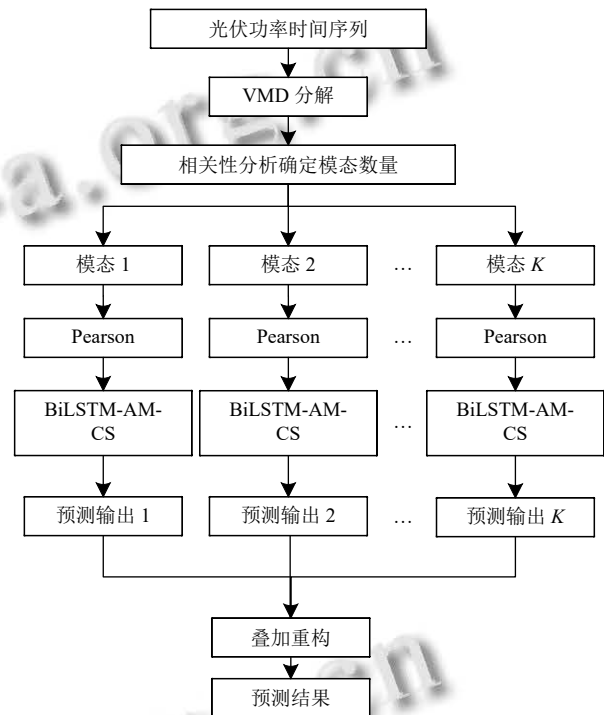


图1 VMD-BiLSTM-AM-CS 光伏功率预测模型

(1) 对原始光伏功率序列进行数据清洗,去除离群值和缺失值后利用VMD方法进行分解,结合相关性分析确定模态数量,得到 K 个模态.

(2) 将各模态光伏序列进行归一化处理,结合皮尔逊相关系数筛选出影响不同模态的关键气象因子.

(3) 针对分解后的 K 个模态与关键气象因子分别构建BiLSTM-AM预测模型,使用CS算法优化网络参数.

(4) 将各分量的预测结果叠加,重构得到待预测日前光伏功率的预测值.

1.1 VMD分析

光伏发电序列具有明显的日周期性和强烈的波动性,属于非平稳序列.VMD算法作为一种新的非递归、自适应、准正交的信号分解方法^[14],将光伏出力

数据从时域转化到频域内进行分解,可以有效捕捉光伏发电功率的非线性特征.其步骤如下.

(1) 建立变分约束问题

$$\begin{cases} \min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ \text{s.t.} \sum_{k=1}^K u_k(t) = f(t) \end{cases} \quad (1)$$

其中, * 为卷积运算符, $f(t)$ 为光伏出力序列, $u_k = \{u_1, u_2, \dots, u_k\}$ 、 $\omega_k = \{\omega_1, \omega_2, \dots, \omega_k\}$ 为分解后 K 个频带分量和对应的中心频率, $\delta(t)$ 为脉冲函数.

(2) 无约束变分求解

引入拉格朗日乘子 λ 和二次罚因子 α , 用于求解上述变分问题, 增广的拉格朗日算子如下所示:

$$\begin{aligned} L(\{u_k\}, \{\omega_k\}, \lambda) = & \alpha \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \\ & + \left\| f(t) - \sum_{k=1}^K u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^K u_k(t) \right\rangle \end{aligned} \quad (2)$$

采用交替方向乘子法得到式 (1) 的最优解, 其中模态分量 u_k 和中心频率为 ω_k 迭代过程为:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \quad (3)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega} \quad (4)$$

拉格朗日算法的算子 λ 迭代过程如式 (5):

$$\hat{\lambda}^{n+1}(\omega) \leftarrow \hat{\lambda}^n(\omega) + \tau \left(\hat{f}(\omega) - \sum_{k=1}^K \hat{u}_k^{n+1}(\omega) \right) \quad (5)$$

其中, τ 为噪声容限.

当满足迭代精度时停止, 即:

$$\sum_{k=1}^K \left\| \hat{u}_k^{n+1} - \hat{u}_k^n \right\|_2^2 / \left\| \hat{u}_k^n \right\|_2^2 < \varepsilon \quad (6)$$

1.2 气象因子选取

在光伏电站的内部技术参数基本保持不变的情况下, 从外部气象因素 (如辐照度、温度、湿度等) 中筛选出影响光伏输出功率的主要因素至关重要. 需要对这些影响变量与光伏序列各模态进行相关性分析, 利用皮尔逊相关系数 $R(X, Y)$ 计算两个变量 X 与 Y 之间的线性相关程度:

$$R(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (7)$$

由式 (7) 可知, 相关系数的绝对值越大, 相关性越强.

1.3 BiLSTM 模块

LSTM 网络是循环神经网络的变形^[15], 由 3 个存储门组成, 即输入门 (i_t)、输出门 (o_t) 和遗忘门 (f_t). 其结构如图 2 所示.

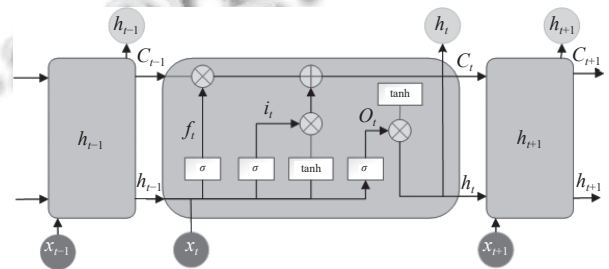


图 2 LSTM 模型结构

LSTM 不仅可以挖掘光伏功率的时序特征, 还可以学习长期依赖关系, 降低在模型训练时梯度失控的风险. x_t 、 h_t 分别表示 t 时刻的输入与输出, C_t 表示 t 时刻 LSTM 隐藏节点状态, 其更新过程如下所示:

$$\begin{cases} f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ \tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \\ C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \\ h_t = o_t * \tanh(C_t) \end{cases} \quad (8)$$

其中, W_f 、 W_i 、 W_o 、 W_c 为权重矩阵, b_f 、 b_i 、 b_o 、 b_c 为对应偏置, σ 为 ReLU 激活函数, \tanh 为双曲正切函数.

LSTM 通过挖掘历史光伏序列信息来预测下一时刻的输出, 然而某一时刻的输出往往同时依赖历史与未来时刻的状态信息. BiLSTM 结构图如图 3 所示, 其由正反向两个 LSTM 网络组成, 可以分别利用时间序列的前后变化规律进行双向预测^[16], 预测更具全局性与整体性, 有效提高光伏预测精度, 具体计算公式如下:

$$H_t = h_{tf} \oplus h_{tb} \quad (9)$$

其中, H_t 表示 BiLSTM 的隐藏状态向量; \oplus 为向量拼接操作; h_{tf} 、 h_{tb} 分别表示前向和后向 LSTM 的隐藏状态向量.

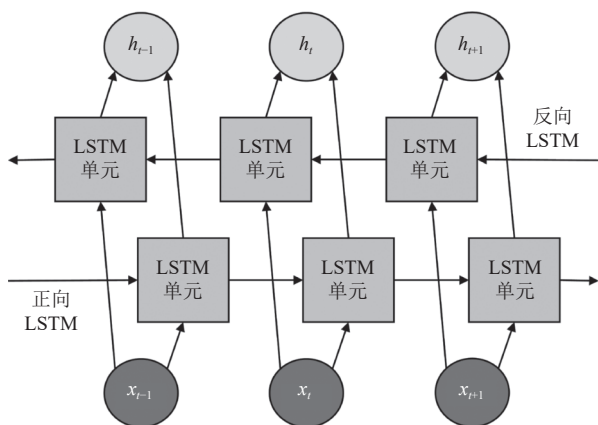


图3 BiLSTM 结构图

1.4 注意力机制

BiLSTM 虽能捕获双向序列信息, 但无法从长时间序列中突出关键信息的影响. 通过引入注意力机制, 可以动态调整神经网络隐藏层的权重, 在不增加模型计算与存储的基础上专注学习重要信息以提升预测精度, 其结构图如图4所示.

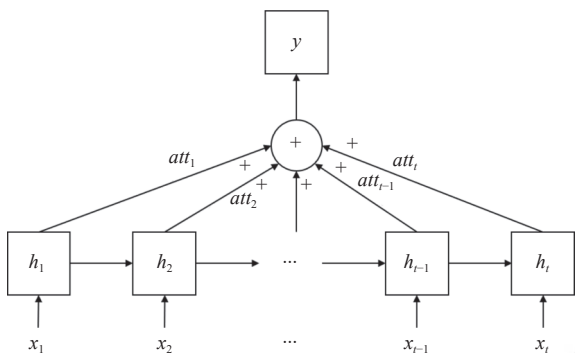


图4 注意力机制结构图

注意力机制重要计算公式如下:

$$e_t = \tanh(wh_t + b) \tag{10}$$

$$att_t = \frac{\exp(e_t)}{\sum_{j=1}^T \exp(e_j)} \tag{11}$$

$$O_t = \sum_{i=1}^T att_i h_i \tag{12}$$

其中, e_t 为 BiLSTM 网络 t 时刻输出向量 h_t 对应的注意力权重得分, w 为权重矩阵, b 为偏置向量, att_t 为注意力权重向量, O_t 为注意力机制在 t 时刻的输出, T 为输入序列维度.

1.5 CS 算法

布谷鸟搜索算法是一种模拟布谷鸟寻找产卵的鸟窝而随机游走的群体优化算法, 用以解决复杂、非线性问题^[17]. 与遗传算法、粒子群算法等其他优化算法相比, CS 算法所需参数少、操作简单、随机搜索路径更优, 并且能够收敛到全局最优解.

CS 的主要思想基于两种策略: 布谷鸟的巢寄生和 Levy 的飞行机制. 通过 Levy 的飞行机制为寻找孵化卵子的最佳宿主巢, 设可利用的鸟窝数量为 m , 可建立如下高效优化模型:

$$\begin{cases} x_i(t+1) = x_i(t) + \alpha \oplus Levy(\lambda) \\ i = 1, 2, \dots, m; t = 1, 2, \dots \end{cases} \tag{13}$$

其中, $x_i(t)$ 是第 t 代的第 i 个鸟窝的位置, α 是步长, $Levy(\lambda)$ 为随机搜索路径, 其飞行公式为:

$$Levy(\lambda) \sim \frac{\Phi \times u}{|v|^{\frac{1}{\beta}}}, \beta \in (1, 3] \tag{14}$$

其中, u, v 都服从标准正态分布. Φ 满足:

$$\Phi = \left\{ \frac{\Gamma(1+\beta) \sin\left(\frac{\pi\beta}{2}\right)}{\Gamma\left\{\left[\frac{1+\beta}{2}\right]\beta 2^{\frac{\beta-1}{2}}\right\}} \right\}^{\frac{1}{\beta}} \tag{15}$$

2 算例分析

本算例所有仿真实验在 Python 3.6 版本、TensorFlow 1.12.0 和 Keras 2.2.0 运行环境下进行. 采用的实验数据选用美国国家能源部可再生能源实验室 (NREL) 亚利桑那州地区 2006 年 1 月到 2006 年 12 月的实测光伏功率数据与气象数据, 样本的采样周期为 30 min, 光伏电站额定装机容量为 104 MW.

使用平均绝对误差 (MAE)、均方根误差 (RMSE) 和拟合度 (R^2) 作为模型的评价指标, 其表达式如下:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - y'_i| \tag{16}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2} \tag{17}$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - y'_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \tag{18}$$

其中, y_t 为实际值, y'_t 为预测值, \bar{y} 为实际值的平均值, N 为样本数量.

2.1 数据预处理

(1) 异常值处理

原始的光伏功率与气象数据因为包含缺失值和离群值, 并不能直接用于预测, 因此需对其进行数据清洗, 去除离群值和缺失值.

(2) 数据标准化

为了消除物理维度的差异, 加快模型的收敛速度, 需要对数据进行标准化.

$$\tilde{x}(i) = \frac{x(i) - x_{\min}}{x_{\max} - x_{\min}} \quad (19)$$

其中, $x(i)$ 为原始光伏序列或气象序列样本; $\tilde{x}(i)$ 为标准化处理后的序列; x_{\min} 和 x_{\max} 为样本数据的最小值和最大值. 归一化处理, 数据的值都在 $[0, 1]$ 内.

(3) 数据集划分

训练集、验证集、测试集按照 70%、10%、20% 的比例进行划分.

2.2 基于变分模态的光伏功率分解结果

为确定模态数 K 的取值, 对不同模态数下分解出的相邻的模态分量的相关性进行分析, 求解其皮尔逊相关系数, 结果如表 1 所示. 当 $K=5$ 时, 模态 4 和模态 5 的相关性显著增强, 此时存在模态重复或产生额外的噪声, 因此 $K=4$ 时分解效果最好.

表 1 不同模态数下相邻的模态分量的相关性

模态数 K	C_{12}	C_{23}	C_{34}	C_{45}
2	0.13	—	—	—
3	0.025	0.0079	—	—
4	0.024	0.0066	0.027	—
5	0.024	0.0063	0.026	0.11

图 5 展示了真实光伏功率与不同 VMD 分解模态下的重构值间的差异, 由 MAE 及 $RMSE$ 结果可得, 当模态数 K 为 4 时, 重构误差最小, 验证了结合相关性分析确定模态数量的有效性. $K=4$ 时分解结果如图 6 所示, 其中第一个序列为光伏功率的原始信号, IMF1-IMF4 表示光伏功率分解后得到子模态, 显然 VMD 分解后的不同模态不仅平稳性更强, 且很好地保持了原光伏数据的趋势特征.

2.3 各模态气象因子选取结果

经 VMD 分解的光伏序列各模态与气象因素之间的皮尔逊相关系数如表 2 所示.

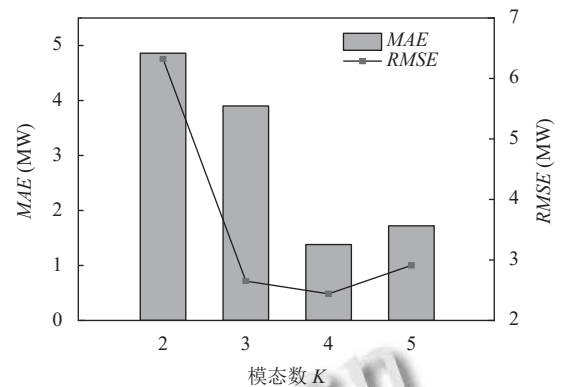


图 5 不同分解模态的重构误差

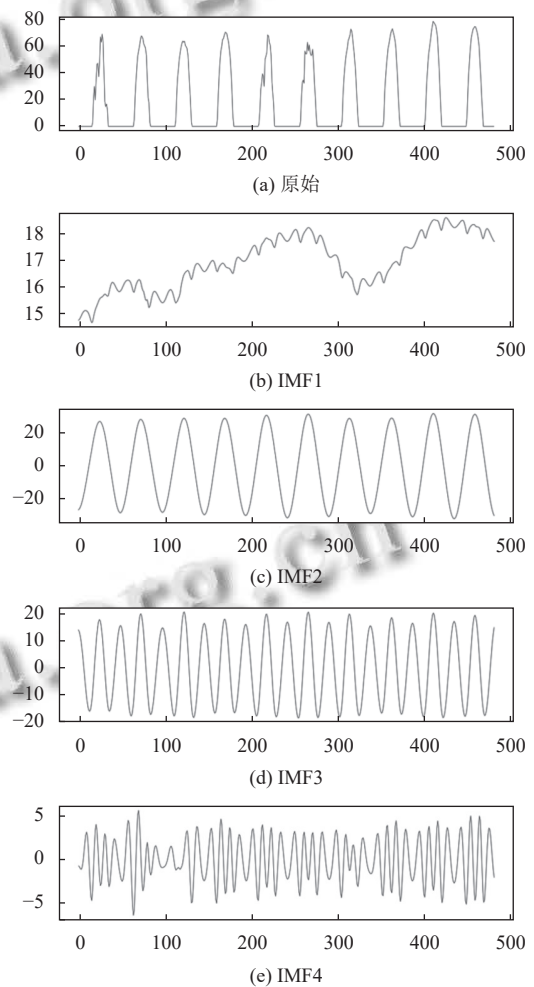


图 6 基于 VMD 的光伏功率分解结果 (横坐标: 样本点, 纵坐标: 功率 (MW))

由表 2 可知, 影响不同模态的关键气象因子大不相同, 因此对各模态分别选取气象特征是有必要的, 此处选择相关系数绝对值大于 0.3 的气象因素作为输入特征来训练模型, 如在模态 2 中选取直接辐射、法向辐

射、水平散射辐射、太阳天顶角、风速、相对湿度、温度、紫外线辐照度作为关键气象因子. 考虑到模态4没有相关性强的变量, 输入特征仅为该模态历史序列.

表2 不同模态与气象因素的相关系数值表

变量	模态1	模态2	模态3	模态4
直接辐射	-0.08	0.67	0.21	0.02
法向辐射	0.26	0.84	0.25	0.09
水平散射辐射	0.21	0.86	0.40	0.08
太阳天顶角	-0.11	-0.94	0	0
风速	0.04	0.45	0.08	0.04
风向	0.07	0.2	-0.01	-0.01
相对湿度	-0.38	-0.44	-0.12	-0.03
温度	0.21	0.48	0.13	0.01
压强	0.15	0	0.03	0.08
紫外线辐照度	0.21	0.84	0.42	0.08

2.4 模型参数

预测过程中, 模型的输入数据为待预测日前一天的光伏出力与关键气象因子, 输出为待预测日的光伏发电功率.

BiLSTM-AM网络的权重系数在训练过程中使用的损失函数为MAE, 优化器为Adam^[18], 容易陷入局部最优值. 对此, 本文引入CS算法调整BiLSTM-AM内部隐含层神经元数量、学习率和dropout层系数, 设置CS算法中各参数: 种群数量 $P=150$, Levy飞行参数 $\beta=1.5$, 更换鸟巢位置概率 pa 为0.25, 最大迭代次数为300. 其训练过程如图7所示.

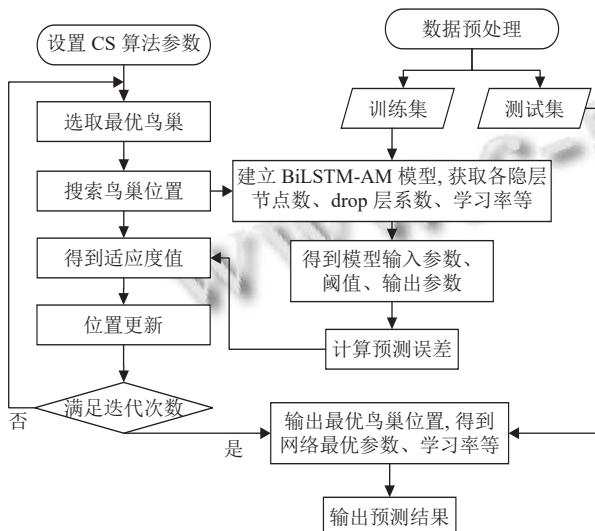


图7 BiLSTM-AM-CS 训练流程

最终 BiLSTM-AM 模型参数优化选择结果如表3.

2.5 算法复杂度分析

(1) VMD 算法

为评估 VMD 算法的性能, 分别采用经验模态分解 (empirical mode decomposition, EMD)、奇异谱分析 (singular spectral analysis, SSA) 与 VMD 算法对光伏功率数据进行分解, 每组实验重复 10 次, 比较各算法平均运行时间如表4所示. 可以看出, VMD 分解平均运行时间最短, 算法复杂度最小. 由于需要多重迭代更新, EMD 分解的平均运行时间显著高于其他算法.

(2) CS 算法

CS 算法每一代总的时间复杂度为^[19]:

$$T(n) = 3O(n + f(n)) + O(n) = O(n + f(n)) \quad (20)$$

其中, $f(n)$ 为优化 n 维问题的目标函数.

优化算法的运行时间与种群数量 P 、迭代次数 $iter$ 、维度 n 密切相关. 在迭代次数 $iter=500$ 、维度 $n=3$ 不变的前提下, 改变种群的数量以对比不同数据规模下的算法时间性能. 图8展示了粒子群优化算法 (particle swarm optimization, PSO)、遗传算法 (genetic algorithm, GA)、人工蜂群算法 (artificial bee colony algorithm, ABC) 与 CS 算法的运行时间对比图. 可以看出, CS 算法运行时间虽略高于简单、运行速度快的 PSO 算法, 但相比其他优化算法较低, 具有一定优越性.

表3 模型参数结果

参数	取值
隐藏层层数	3
隐藏层1神经元个数	85
隐藏层2神经元个数	96
隐藏层3神经元个数	48
损失函数	0.13
Epoch	100
学习率	0.0018
Batch size	64
Dropout	0.2

表4 算法平均运行时间 (s)

算法	平均运行时间
EMD	3.795
SSA	0.861
VMD	0.511

2.6 光伏功率预测对比实验

为验证上文分析建立的光伏功率预测模型的有效性, 分别与日前模型 VMD-BiLSTM、BiLSTM-AM、VMD-LSTM-AM、VMD-BiLSTM-AM 进行比较, 预测误差指标对比见表5, 表中最佳预测结果以加粗表示.

可以看出, VMD-BiLSTM-AM-CS 模型具有最佳的性能, 各指标结果均最优. 通过对比模型 VMD-LSTM-AM 与 VMD-BiLSTM-AM, 后者预测误差更低, 表明 BiLSTM 通过将前向和后向特征信息结合, 相比单向 LSTM 可以挖掘更多有效信息; 通过对比本模型与 VMD-BiLSTM、BiLSTM-AM 模型的预测误差, 验证了注意力机制与 VMD 分解的有效性; 此外, VMD-BiLSTM-AM-CS 显然较未使用 CS 优化的模型预测效果更佳, 验证了 CS 参数优化有效改善了网络性能, 进一步提高了预测精度.

为进一步验证本文模型在不同天气类型下的预测有效性, 本文选取 10 月 22 日 (晴天), 10 月 28 日 (阴天), 11 月 14 日 (雨天), 11 月 3 日 (多云) 作为预测日. 分别与日前模型 VMD-BiLSTM、BiLSTM-AM、VMD-LSTM-AM、VMD-BiLSTM-AM 进行比较, 预测结果如图 9 所示, 预测误差见表 6, 表中每种气象类型下的最佳预测结果以加粗表示.

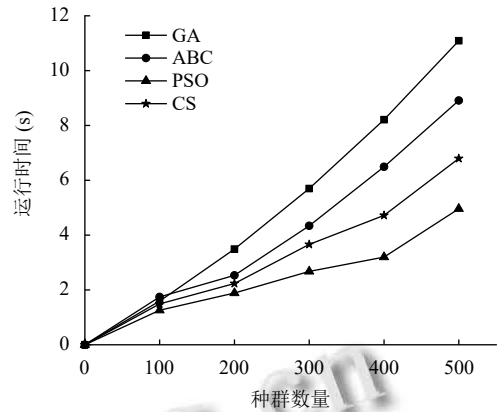
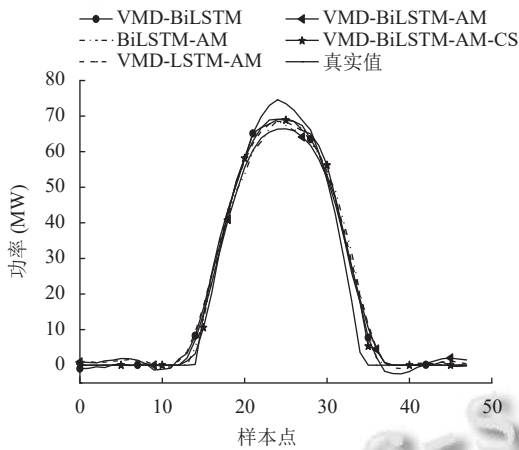


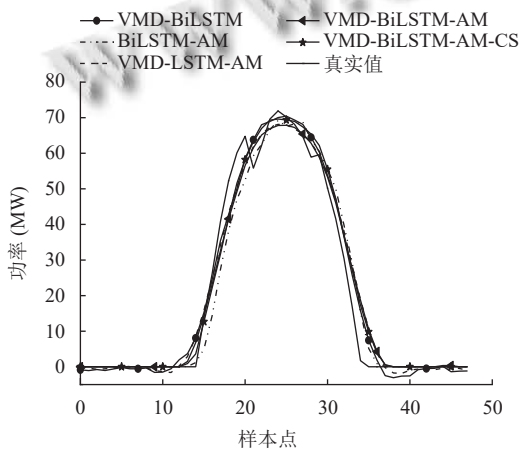
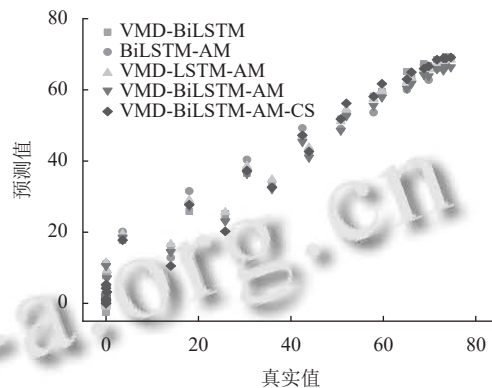
图 8 不同优化算法运行时间对比

表 5 不同模型预测误差指标对比

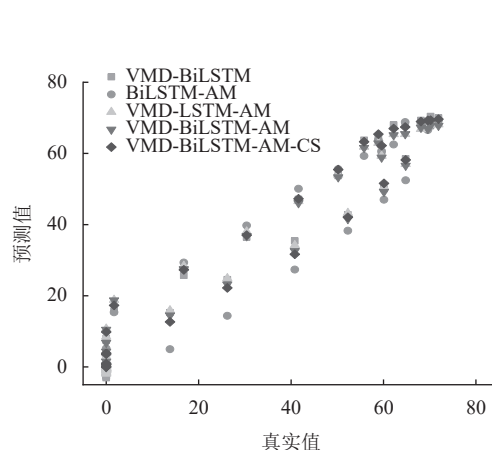
方法	MAE (MW·h)	RMSE (MW·h)	R ²
VMD-BiLSTM	3.86	6.44	0.924
BiLSTM-AM	4.35	8.82	0.876
VMD-LSTM-AM	4.42	7.76	0.904
VMD-BiLSTM-AM	3.58	6.63	0.930
VMD-BiLSTM-AM-CS	2.94	5.77	0.947



(a) 晴天



(b) 阴天



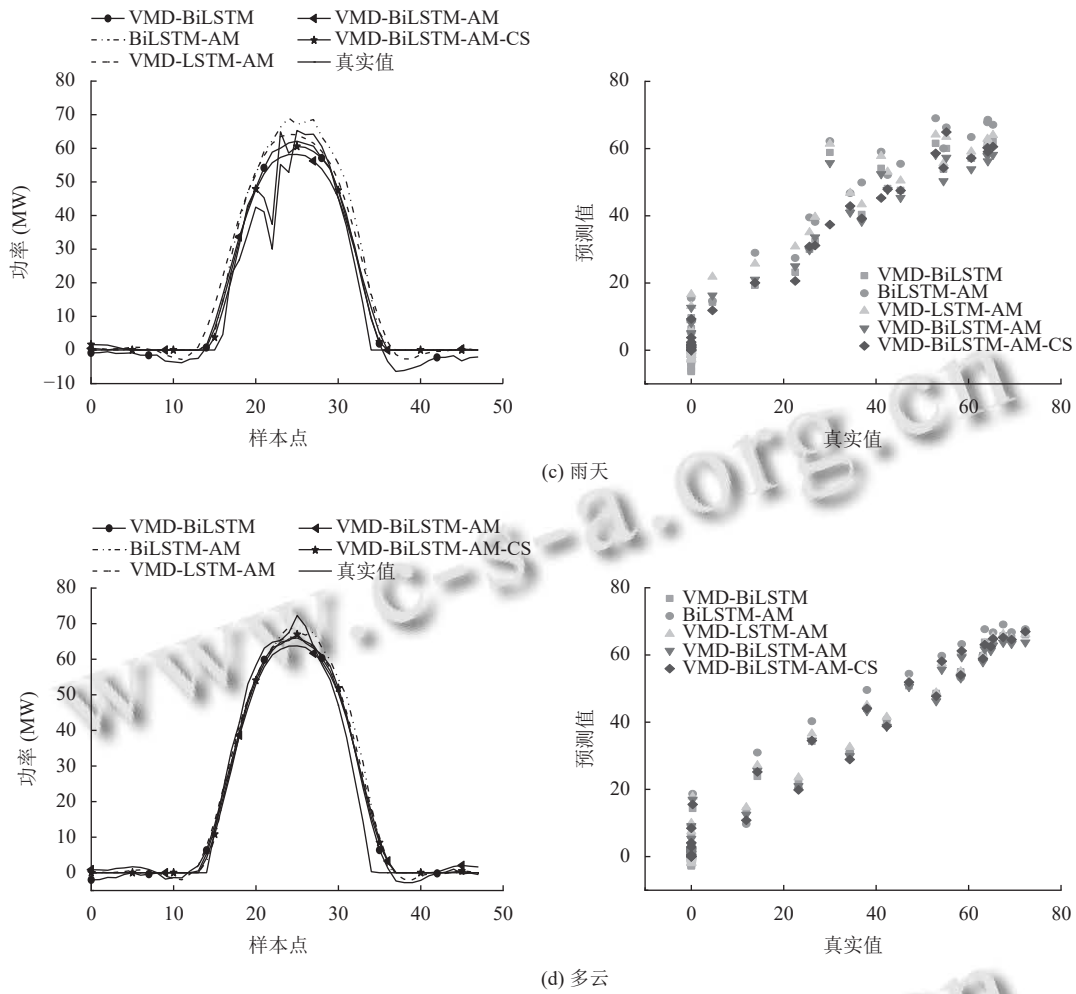


图9 不同气象类型的预测曲线

表6 不同气象类型的评价指标

气象类型	方法	MAE (MW·h)	RMSE (MW·h)	R ²
晴天	VMD-BiLSTM	2.42	3.84	0.984
	BiLSTM-AM	2.56	4.64	0.973
	VMD-LSTM-AM	2.32	4.17	0.977
	VMD-BiLSTM-AM	2.17	3.76	0.976
	VMD-BiLSTM-AM-CS	1.98	3.41	0.985
阴天	VMD-BiLSTM	3.10	4.67	0.971
	BiLSTM-AM	3.18	5.69	0.957
	VMD-LSTM-AM	3.07	4.84	0.969
	VMD-BiLSTM-AM	2.74	4.82	0.969
	VMD-BiLSTM-AM-CS	2.71	4.66	0.974
雨天	VMD-BiLSTM	4.19	6.15	0.926
	BiLSTM-AM	4.69	8.31	0.866
	VMD-LSTM-AM	5.01	8.15	0.871
	VMD-BiLSTM-AM	3.16	5.82	0.934
	VMD-BiLSTM-AM-CS	2.28	3.66	0.945
多云	VMD-BiLSTM	2.56	3.72	0.978
	BiLSTM-AM	2.71	5.14	0.962
	VMD-LSTM-AM	2.64	4.48	0.972
	VMD-BiLSTM-AM	2.88	4.43	0.972
	VMD-BiLSTM-AM-CS	2.37	4.08	0.980

由各模型不同气象类型下的仿真结果对比分析可得: 本文所提模型的预测曲线更加贴近实际光伏发电功率曲线, 在不同的天气条件下, VMD-BiLSTM-AM-CS模型的MAE均低于其他模型, R²均高于其他模型, 其中晴天的预测误差最低, 阴天预测误差最高. 不同气象类型下, VMD-BiLSTM、BiLSTM-AM模型的预测误差均高于本模型, 尤其是雨天, 本模型的MAE分别降低了45.6%与51.4%; 与模型VMD-LSTM-AM相比, VMD-BiLSTM-AM的MAE、RMSE均更低, 验证了不同气象类型下BiLSTM特征提取效果更优秀; 此外, VMD-BiLSTM-AM-CS较未使用CS优化的模型精度更高, 在雨天和多云天气下MAE分别降低了27.8%、17.7%, 提高了模型的泛化能力. 综上所述, 该模型在不同天气下运行具有较强鲁棒性, 提高了日前光伏预测精度.

3 结论

本文提出了CS算法优化的VMD-BiLSTM-AM日前光伏发电功率预测方法,该方法通过VMD将原始序列分解为不同模态较为平稳的序列,降低了不同光伏序列间的相互影响;再为各序列分别选取关键气象因子,使用BiLSTM网络充分挖掘光伏序列与气象因子的双向时序特征,结合注意力机制突出关键信息对预测结果的影响;最后利用CS算法进行网络参数智能寻优,防止陷入局部最优。实例分析表明,该方法有效提高了日前光伏预测精度,在不同气象条件下均具有更优异的性能。

参考文献

- 1 Feng L, Zhang JN, Li GJ, *et al.* Cost reduction of a hybrid energy storage system considering correlation between wind and PV power. *Protection and Control of Modern Power Systems*, 2016, 1(1): 11. [doi: 10.1186/s41601-016-0021-1]
- 2 Das UK, Tey KS, Seyedmahmoudian M, *et al.* Forecasting of photovoltaic power generation and model optimization: A review. *Renewable and Sustainable Energy Reviews*, 2018, 81: 912–928. [doi: 10.1016/j.rser.2017.08.017]
- 3 丁明, 鲍玉莹, 毕锐. 应用改进马尔科夫链的光伏出力时间序列模拟. *电网技术*, 2016, 40(2): 459–464. [doi: 10.13335/j.1000-3673.pst.2016.02.018]
- 4 Li YT, He Y, Su Y, *et al.* Forecasting the daily power output of a grid-connected photovoltaic system based on multivariate adaptive regression splines. *Applied Energy*, 2016, 180: 392–401. [doi: 10.1016/j.apenergy.2016.07.052]
- 5 余光正, 陆柳, 汤波, 等. 基于云图特征提取的改进混合神经网络超短期光伏功率预测方法. *中国电机工程学报*, 2021, 41(20): 6989–7002. [doi: 10.13334/j.0258-8013.pcsee.201929]
- 6 宋绍剑, 李博涵. 基于LSTM网络的光伏发电功率短期预测方法的研究. *可再生能源*, 2021, 39(5): 594–602. [doi: 10.13941/j.cnki.21-1469/tk.2021.05.006]
- 7 廖卫强, 张认成, 俞万能, 等. 基于相似样本及PCA的光伏输出功率预测. *太阳能学报*, 2016, 37(9): 2377–2385. [doi: 10.3969/j.issn.0254-0096.2016.09.031]
- 8 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. *电力系统自动化*, 2021, 45(8): 175–184. [doi: 10.7500/AEPS20200718002]
- 9 陈涛, 王艳, 纪志成. 基于经验小波变换的光伏功率组合预测模型. *系统仿真学报*, 2021, 33(11): 2627–2635. [doi: 10.16182/j.issn1004731x.joss.21-FZ0709]
- 10 王粟, 江鑫, 曾亮, 等. 基于VMD-DESN-MSGP模型的超短期光伏功率预测. *电网技术*, 2020, 44(3): 917–925. [doi: 10.13335/j.1000-3673.pst.2019.0958]
- 11 张彩庆, 郑强. SKBA-LSSVM短期光伏发电功率预测模型. *电力系统及其自动化学报*, 2019, 31(8): 86–93. [doi: 10.19635/j.cnki.csu-epsa.000104]
- 12 刘沛汉, 袁铁江, 梅生伟, 等. 基于遗传算法优化神经网络的光伏电站短期功率预测. *水电能源科学*, 2016, 34(1): 211–214.
- 13 江友华, 帅禄玮, 吴琦娜, 等. 基于布谷鸟算法的双向直流电子变压器谐振参数优化. *太阳能学报*, 2021, 42(11): 485–494. [doi: 10.19912/j.0254-0096.tynxb.2019-1246]
- 14 杨晶显, 张帅, 刘继春, 等. 基于VMD和双重注意力机制LSTM的短期光伏功率预测. *电力系统自动化*, 2021, 45(3): 174–182. [doi: 10.7500/AEPS20200226011]
- 15 Gao MM, Li JJ, Hong F, *et al.* Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM. *Energy*, 2019, 187: 115838. [doi: 10.1016/j.energy.2019.07.168]
- 16 王渝红, 史云翔, 周旭, 等. 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. *高电压技术*, 2022, 48(5): 1884–1892. [doi: 10.13336/j.1003-6520.hve.20211561]
- 17 Deb S, Yang XS. Cuckoo search via lévy flights. *Proceedings of 2009 World Congress on Nature & Biologically Inspired Computing*. Coimbatore: IEEE, 2009. 210–214.
- 18 Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv:1412.6980, 2014.
- 19 张永韡, 汪镭, 吴启迪. 动态适应布谷鸟搜索算法. *控制与决策*, 2014, 29(4): 617–622. [doi: 10.13195/j.kzyjc.2012.1815]

(校对责编: 孙君艳)