



基于改进的极限学习机光伏出力短期预测

成燕 庄飞鸯 徐万万 魏婷

Short-term Prediction of Photovoltaic Output Based on Improved Extreme Learning Machines

CHENG Yan, ZHUANG Feiyang, XU Wanwan, WEI Ting

引用本文:

成燕, 庄飞鸯, 徐万万, 等. 基于改进的极限学习机光伏出力短期预测[J]. 现代电力, 2023, 40(5): 679–686. DOI: 10.19725/j.cnki.1007-2322.2022.0080

CHENG Yan, ZHUANG Feiyang, XU Wanwan, et al. Short-term Prediction of Photovoltaic Output Based on Improved Extreme Learning Machines[J]. *Modern Electric Power*, 2023, 40(5): 679–686. DOI: 10.19725/j.cnki.1007-2322.2022.0080

在线阅读 View online: <https://doi.org/10.19725/j.cnki.1007-2322.2022.0080>

您可能感兴趣的其他文章

Articles you may be interested in

基于混合麻雀算法改进反向传播神经网络的短期光伏功率预测

Short-Term Photovoltaic Power Prediction Based on Back Propagation Neural Network Improved by Hybrid Sparrow Algorithm
现代电力. 2022, 39(3): 287–294 <https://doi.org/10.19725/j.cnki.1007-2322.2021.0113>

基于混沌特征改进鲸鱼优化算法-相关向量机的超短期光伏发电输出功率预测

A Method to Forecast Ultra-Short-Term Output of Photovoltaic Power Generation Based on Chaotic Characteristic-Improved Whale Optimization Algorithm and Relevance Vector Machine
现代电力. 2021, 38(3): 268–276 <https://doi.org/10.19725/j.cnki.1007-2322.2020.0390>

基于DPK-means和ELM的日前光伏发电功率预测

Day-ahead Photovoltaic Power Generation Forecasting Based on DPK-means and ELM
现代电力. 2020, 37(4): 351–357 <https://doi.org/10.19725/j.cnki.1007-2322.2019.0929>

基于自适应噪声完备经验模态分解样本熵长短期记忆神经网络和核密度估计的短期电力负荷区间预测

Short-term Power Load Interval Prediction Based on CEEMDAN-SE-LSTM and KDE
现代电力. 2021, 38(2): 138–146 <https://doi.org/10.19725/j.cnki.1007-2322.2020.0329>

基于FCM和CG-DBN的光伏功率短期预测

Short-term Prediction of Photovoltaic Power based on Combination of FCM and CG-DBN
现代电力. 2019, 36(5): 62–67 <http://xddl.ncepujournal.com/article/Y2019/I5/62>

基于约束投票极限学习机的在线静态电压稳定评估

Online Static Voltage Stability Assessment Based on Constrained Voting Extreme Learning Machine
现代电力. 2022, 39(5): 521–528 <https://doi.org/10.19725/j.cnki.1007-2322.2021.0162>

基于改进的极限学习机光伏出力短期预测

成燕^{a,b}, 庄飞鸯^a, 徐万万^{a,b}, 魏婷^a

(武汉科技大学 a. 信息科学与工程学院; b. 冶金自动化与检测技术教育部工程研究中心, 湖北省武汉市 430081)

Short-term Prediction of Photovoltaic Output Based on Improved Extreme Learning Machines

CHENG Yan^{a,b}, ZHUANG Feiyang^a, XU Wanwan^{a,b}, WEI Ting^a

(a. School of Information Science and Engineering; b. Engineering Research Center for Metallurgical Automation and Measurement Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, Hubei Province, China)

摘要: 针对传统极限学习机易陷入局部最优解的缺点以及环境变化导致光伏出力波动的特点, 构建了一种基于自适应噪声完全集成经验模态分解 (complete ensemble empirical mode decomposition with adaptive noise, CEEMDAN) 算法, 结合黑猩猩优化算法优化极限学习机神经网络的光伏出力短期预测模型。首先利用 CEEMDAN 算法将影响光伏输出功率的关键环境因素序列进行分解, 得到数据信号在不同时间尺度的局部特征, 降低环境因素序列的非平稳性, 然后将各分解子序列和光伏历史数据序列作为黑猩猩算法优化的极限学习机预测模型输入进行预测。最后, 选用 DKASC Solar Centre 光伏电站数据集对不同预测模型进行验证对比。实例仿真结果表明, 构建的改进光伏出力预测组合模型的各项指标预测效果更好, 且适用不同环境的光伏发电预测。

关键词: 光伏短期预测; 自适应噪声完全集成经验模态分解算法; 极限学习机; 黑猩猩优化算法

Abstract: In allusion to the defect of traditional extreme learning machine easily falling into local optimal solutions and the characteristic of environment variation leading to the fluctuation of photovoltaic (abbr. PV) output, based on complete ensemble empirical mode decomposition with adaptive noise (abbr. CEEMDAN) algorithm and combining with extreme learning machine neural network optimized by chimp optimization algorithm a short-term PV output prediction model was constructed. Firstly, by use of CEEMDAN algorithm the key environment factor series impacting PV output was decomposed to obtain the local feature of data signals in different time-scales to reduce the non-stationary of environment factor series. Secondly, taking each decomposed subseries and PV historical data series as the input of extreme learning machine prediction model optimized by the chimp algorithm the prediction was

performed. Finally, the data set of DKASC Solar Centre PV station was chosen to conduct the contrast and verifying for different prediction models. Results of simulation example show that the prediction effect of each index of the constructed improved PV output prediction combined model is better and suitable to the prediction of PV generation in different environments.

Keywords: short-term prediction of photovoltaic; CEEMDAN; extreme learning machine; chimp optimization algorithm

DOI: 10.19725/j.cnki.1007-2322.2022.0080

0 引言

由于化石能源的短缺和环境污染问题, 以及国家节能减排的压力, 太阳能光伏作为可再生能源的一种, 在中国得到快速发展^[1-2]。然而受太阳辐照度、温度等环境因素的影响, 光伏的输出功率本质上具有间歇性、波动性等特点, 并且在大规模光伏电源接入发电系统后, 影响其稳定运行以及能量管理环节中的光伏消纳率^[3-6]。因此, 要有效地利用影响光伏出力的关键环境因素, 提高光伏功率的预测精确度。

光伏出力预测的研究方法主要可以分为物理方法、统计方法和神经网络^[7]。物理预报方法不需要历史数据, 需要精确的地理、气象等实测信息和完整的光伏电池信息^[8], 但计算时间较长、实用性有待提高。统计预测方法是基于历史数据和目标预测数据之间的映射关系对未来光伏输出功率的预测^[9], 其处理动态非周期性的时间序列

数据效果不佳。人工神经网络技术对非线性处理的能力符合光伏电源的变化规律,提高了光伏输出功率的预测精度^[10-11]。与传统神经网络相比,极限学习机(extreme learning machines, ELM)神经网络计算速度快,对数据样本适应能力强。近年来,光伏出力预测研究取得了一些成果。文献[12]利用小波变换将光伏历史信号分解为多个频率序列,然后用深度卷积神经网络提取每个频率中的非线性特征进行预测;文献[13]加入注意力机制增强提取光伏出力数据时间特征,然后构建了长短期网络和卷积神经网络组合预测模型。上述方法仅依据光伏历史数据建立模型,没有考虑影响光伏出力的环境变量因素。文献[14]提出以综合指标选取相似日,采用列文伯格-马夸尔特(Levenberg-Marquardt, L-M)算法改进ELM神经网络的参数,获得最优模型;文献[15]采用主因子分析法筛选光伏发电的主要影响因子,并作为预测模型输入量,提高预测精度。然而由于天气条件的复杂性,上述预测模型可能无法完全提取相应环境因素非线性特征和时间序列特征。

因此,本文充分考虑影响光伏出力的环境因素,针对光伏出力依赖环境变化造成出力波动的特点,构建了一种基于改进的ELM的光伏输出功率短期组合预测模型。采用CEEMDAN算法对关键环境因素进行分解,降低数据的非平稳性,增加局部特征多样性,针对ELM神经网络易陷入局部最优解的缺陷,利用ChOA算法优化ELM神经网络的输入层权重和隐含层偏置。最后采用实测数据与不同预测方法进行对比分析,实现在不同天气下的光伏功率预测,验证本文提出的组合预测方法的有效性。

1 影响光伏输出功率环境因素分析

本文采用DKASC Solar Centre光伏电站中的数据集,以及光伏阵区对应的环境监测仪获取到的太阳辐照度、环境温度、环境相对湿度、风速、降雨量5种环境因素数据。

皮尔逊相关系数通常用于2个变量之间衡量其线性相关的程度,相关性的存在为相关变量作为输入变量引入模型提供了基础^[16]。因此利用皮尔逊相关系数法分析得到各环境因素信息和光伏输出功率之间的相关关系,筛选出影响光伏输出功率的关键环境因素,如公式(1)所示

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

式中: λ 为变量之间的相关性系数,在[-1, 1]之间; X_i 、 Y_i 为第*i*个数据点的2个因素的值。

表1为该数据集中各环境因素和光伏输出功率的相关性分析。从表1可以看出,太阳辐照度、环境温度、环境相对湿度、风速与光伏出力均具有较高的相关性,而降雨量表现为极弱相关性。综上,文中选择太阳辐照度、环境温度、环境相对湿度与风速作为关键环境因素。

表1 光伏功率与环境因素相关性

Table 1 Correlation between photovoltaic power and environmental factors

环境因素	太阳辐照度	环境温度	环境相对湿度	风速	降雨量
相关性系数	0.93	0.48	-0.38	0.32	0.07

2 理论基础

由于天气变化会造成光伏出力波动较大。在天气突变时,仅仅依靠光伏功率的历史数据进行预测精度不高。因此先采用CEEMDAN算法对环境因素进行分解,对非平稳数据进行平稳化处理,得到数据在不同时间尺度上的局部特征,并作为预测模型的输入值。然后采用黑猩猩算法优化ELM输入层的连接权重与隐含层神经元的偏置,并建立预测模型,优化方法改进了ELM神经网络在训练过程中,由于训练样本过多易陷入局部最优解的缺点,提高了模型预测精度。

2.1 CEEMDAN 算法

CEEMDAN算法是在经验模态分解(empirical mode decomposition, EMD)算法的基础上改进的方法,适合分析波动性较大的非平稳信号。传统EMD算法^[17]根据数据信号本身具有的时间尺度特性将原始的复杂信号分解为有限个本征模态函数(intrinsic mode function, IMF)。每一个IMF应满足以下条件:

- 1) 在整个时间范围内,局部极点和过零点的数量必须大于等于1个;
- 2) 在任意时刻,极值构成的上下包络线平均值必须为0。

CEEMDAN算法通过在每一阶段IMF分解中

自适应地加入白噪声，求均值并逐次迭代。克服了 EMD 算法容易造成模态混叠、重构误差的缺点，在非平稳环境因素信号上更具自适应性^[18]。

定义 $E_i(\cdot)$ 为经过 EMD 算法分解后第 i 个 IMF， e_i 为符合标准正态分布的高斯白噪声，待分解的原始环境因素数据 $X(t)$ 经过 CEEMDAN 的第 i 个本征模态分量为 $\overline{C_i(t)}$ 。其分解步骤如下：

1) 在待分解环境信号 $X(t)$ 中添加高斯白噪声 $X(t) + \zeta_0 e_i(t)$ (ζ_0 为白噪声系数)。新信号通过 EMD 进行 N 次分解得到 N 个本征模态函数 C_{n1} ，对产生的 N 个本征模态函数 C_{n1} 进行集成求均值得到 CEEMDAN 分解的第 1 个 IMF 分量，即

$$\overline{C_1(t)} = \frac{1}{N} \sum_{i=1}^N C_{i1}(t) \quad (2)$$

2) 计算分解第一个 IMF 分量后的残差序列

$$R_1(t) = X(t) - \overline{C_1(t)} \quad (3)$$

3) 在 $R_1(t)$ 继续加入高斯白噪声 $R_1(t) + \zeta_1 E_1(e_i(t))$ ($e_i(t)$)，对新信号进行 N 次 EMD 分解，得到第 2 个 IMF 分量以及剩余信号序列：

$$\overline{C_2(t)} = \frac{1}{N} \sum_{i=1}^N E_1(R_1(t) + \zeta_1 E_1(e_i(t))) \quad (4)$$

4) 对于 k 阶迭代，按照步骤 2) — 3) 的计算方式得到 $k+1$ 模态分量

$$R_k(t) = R_{k-1}(t) - \overline{C_k(t)} \quad (5)$$

$$\overline{C_{k+1}(t)} = \frac{1}{N} \sum_{i=1}^N E_k(R_k(t) + \zeta_k E_k(e_k(t))) \quad (6)$$

5) 重复步骤 4) 直到剩余信号序列为单调函数不能继续分解，算法结束。

$$R(t) = X(t) - \sum_{k=1}^K \overline{C_k(t)} \quad (7)$$

则原始环境因素信号 $X(t)$ 可以表示为：

$$X(t) = R(t) + \sum_{k=1}^K \overline{C_k(t)} \quad (8)$$

2.2 黑猩猩优化算法的基本原理

黑猩猩优化算法 (chimpanzee optimization algorithm, ChOA) 是 2020 年由 Khishe 等人通过黑猩猩种群的社会行为启发而提出的一种较新的优化算法^[19]。ChOA 主要模仿黑猩猩攻击、驱动、阻挡和追逐 4 种行为捕获猎物来进行求解。其优点在于操作简单，调整参数少以及稳定性强。黑猩猩优化算

法的数学模型描述如下。

在黑猩猩驱动、追逐猎物的过程中，假设第一个攻击者、驱赶者、拦截者和追逐者发现猎物的位置，其他黑猩猩根据猎物的位置随机改变自身的位置，数学描述为

$$s = |cX_p(t) - hX_c(t)| \quad (9)$$

$$X_c(t+1) = X_p(t) - as \quad (10)$$

式中： s 为猎物与黑猩猩间的距离向量； t 为迭代次数； $X_p(t)$ 为猎物位置向量； $X_c(t)$ 为黑猩猩位置向量； h 为混沌映射产生的混沌向量 (Chaotic_value)； a 、 c 为系数向量，用下式表示

$$a = 2fr_1 - f, \quad c = 2r_2 \quad (11)$$

式中： f 在优化过程中呈现线性变化，由 2 衰减到 0； r_1 、 r_2 为 $[0, 1]$ 内的随机向量。

在攻击过程中，由攻击黑猩猩完成狩猎过程，另 3 种黑猩猩通过驱动、阻拦和追逐探查猎物位置包围猎物参与狩猎行为。4 种黑猩猩通过下式分别更新自己的位置，猎物位置由 4 个最佳个体位置估计，其他黑猩猩随机更新在 4 个最佳个体附近。数学模型为

$$\begin{cases} s_a = |c_1 X_a - h_1 X| \\ s_b = |c_2 X_b - h_2 X| \\ s_c = |c_3 X_c - h_3 X| \\ s_d = |c_4 X_d - h_4 X| \end{cases} \quad (12)$$

$$\begin{cases} X_1 = X_a - a_1 s_a \\ X_2 = X_b - a_2 s_b \\ X_3 = X_c - a_3 s_c \\ X_4 = X_d - a_4 s_d \end{cases} \quad (13)$$

$$X(t+1) = \frac{X_1 + X_2 + X_3 + X_4}{4} \quad (14)$$

式中： s_a 、 s_b 、 s_c 、 s_d 分别为当前 4 类最佳黑猩猩与猎物的间距； X_a 、 X_b 、 X_c 、 X_d 分别为当前 4 类黑猩猩相对于猎物的位置向量； a_1 — a_4 、 h_1 — h_4 、 c_1 — c_4 分别为 4 类黑猩猩的系数向量； X_1 、 X_2 、 X_3 、 X_4 分别为更新后的黑猩猩位置向量； X 为其他黑猩猩位置向量。

种群在狩猎的最后阶段，随后的社会动机 (性动机) 会使黑猩猩不顾个人职责而陷入混乱的抢食状态。这一混乱行为有助于 ChOA 在迭代后期避免陷入局部最优问题。在迭代过程中，假设黑猩猩以同等概率选择黑猩猩正常位置或通过混沌模型生成位置。数学模型表示为

$$X_c(t+1) = \begin{cases} X_p(t) - as, & \mu \leq 0.5 \\ Chaotic_value, & \mu > 0.5 \end{cases} \quad (15)$$

式中： μ 为 $[0, 1]$ 内随机值。

2.3 ELM 的基本原理

ELM 神经网络随机生成网络的输入权值与隐含层偏置，在训练中只需要设置隐含层神经元的个数便得到唯一的全局最优解，克服了传统单隐层前馈神经网络训练速度慢、易陷入局部极值等缺陷^[20-21]。ELM 的网络结构如图 1 所示。

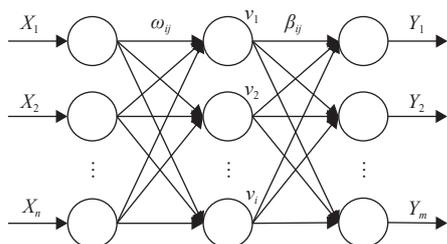


图 1 ELM 结构图

Fig. 1 Structure chart of ELM

由图 1 可以看出，假设有 N 个样本 (X_i, T_i) ，其中 $X_i = [X_{i1}, X_{i2}, \dots, X_{im}]^T$ ， $T_i = [T_{i1}, T_{i2}, \dots, T_{im}]^T$ 。则对于一个含有 i 个隐含层的 ELM 神经网络的输出为

$$Y_j = \sum_{i=1}^n \beta_i G(\omega_i X_j + v_i), \quad j = 1, 2, \dots, n \quad (16)$$

式中： ω_i 、 β_j 分别为输入层、输出层与隐含层间的权重； v_i 为隐含层偏置； $G(x)$ 为激活函数。

$$H_{\omega, v, X} \beta = T \quad (17)$$

式中： T 为期望输出向量； $H_{\omega, v, X}$ 为隐含层的输出矩阵，可表示为

$$H_{\omega, v, X} = \begin{bmatrix} G(\omega_1 X_1 + v_1) & \cdots & G(\omega_i X_1 + v_i) \\ \vdots & \ddots & \vdots \\ G(\omega_1 X_n + v_1) & \cdots & G(\omega_i X_n + v_i) \end{bmatrix} \quad (18)$$

应用下式求解输出权重 β^*

$$\beta^* = H^+ T \quad (19)$$

式中： H^+ 是矩阵 H 的广义摩尔逆 (Moore-Penrose)。

3 基于 CEEMDAN-ChOA-ELM 的光伏功率预测模型

3.1 预测步骤

由于环境因素，比如：辐照度、温度等对光

伏出力有很大的影响，因此本文首先使用 CEEMDAN 算法对原始环境序列数据进行分解得到局部特征，然后用由 ChOA 优化后的 ELM 神经网络进行预测，具体步骤如下：

1) 对光伏功率和环境因素数据进行数据清洗，以天为单位剔除由于故障等原因而导致数据为 0 的“坏数据”；

2) 通过 CEEMDAN 算法对原始环境序列数据进行分解，得到每种环境因素的 IMF 分量，以此来突显原始环境序列中不同时间尺度特征；

3) 使用 ChOA 优化 ELM 神经网络的输入层连接权重和隐含层偏置参数，构造预测模型；

4) 将 CEEMDAN 分解后的环境数据以及光伏功率数据进行归一化，转化成适用于改进后的 ELM 神经网络数据集，并分别进行预测；

5) 叠加子序列的预测结果并验证预测模型的准确性。组合预测方法流程如图 2 所示。

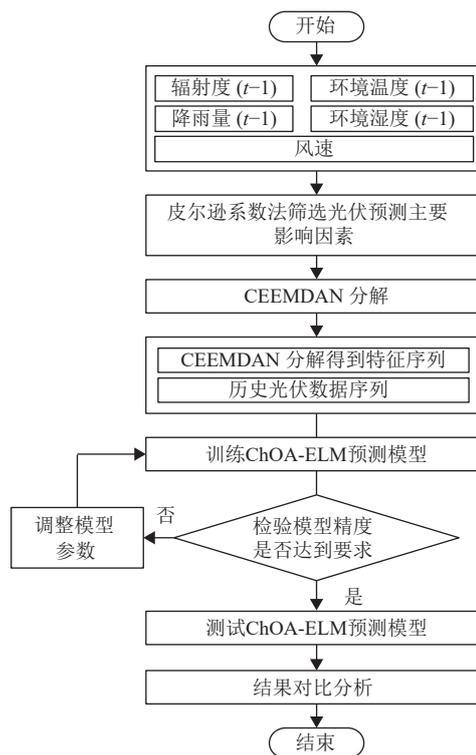


图 2 预测方法流程图

Fig. 2 The flowchart of prediction idea

3.2 模型参数设定

CEEMDAN 参数设置及分解结果：白噪声系数 ζ_0 和每次迭代 EMD 分解次数 N 是非常重要的参数，根据试验验证，当 $\zeta_0 = 0.2$ ， $N = 500$ 时，预测模型效果最佳。图 3 是环境因素中选取部分

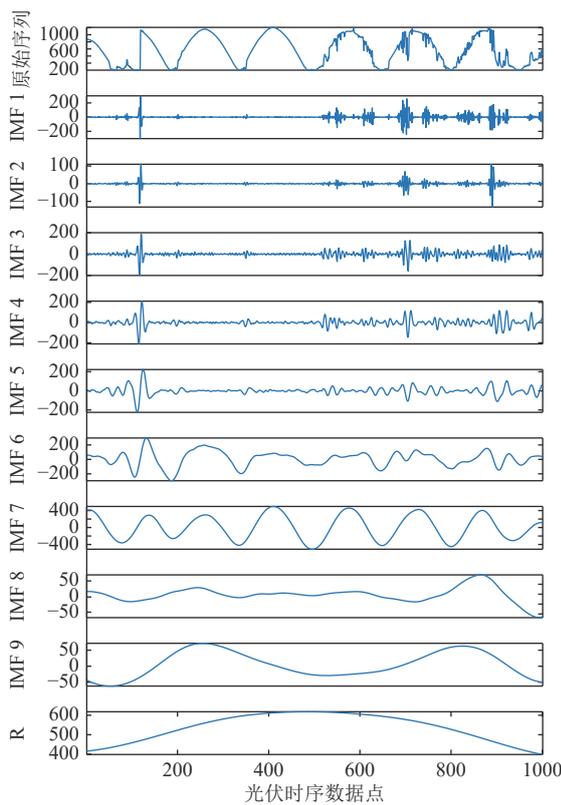


图 3 部分太阳辐照度样本分解结果

Fig. 3 Decomposition results of partial solar irradiance samples

太阳辐照度样本数据的分解结果，可以看出该分解算法能有效地获得信号的频率分量。

表 2 中数据为经 CEEMDAN 算法对 4 种环境数据分解后，得到的每种环境因素的 IMF 分量，共计 61 个和 4 个剩余分量 R，共计 65 维特征序列作为新的输入序列集合。

ChOA 参数设定：黑猩猩初始种群规模为 30，最大迭代次数为 100。

ELM 网络结构：ELM 随机初始化输入层与隐含层连接权重，范围为 $(-1, 1)$ 和隐含层偏置，范围为 $(0, 1)$ 。经过迭代，最佳隐含层为 29。

3.3 数据预处理

本文以光伏电站实际功率和环境数据作为预测模型的原始数据，其中选取太阳辐照度、环境温度、环境相对湿度、风速作为环境因素。为了保证数据量纲一致，处理方便，使程序运行收敛加快，在环境数据输入预测模型前，采用 min-max 标准化方法通过式 (20) 将数据归一化处理。

$$X_i' = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (20)$$

表 2 分解环境数据得到的 IMF 分量及剩余分量数量
Table 2 The number of IMF components and remaining components obtained by decomposing environmental data

环境因素数据	IMF 分量	剩余分量 R
太阳辐照度	15	1
环境温度	16	1
环境相对湿度	16	1
风速	14	1

式中： X_i' 为标准化后的数据； X_{\min} 、 X_{\max} 分别为原始数据的最小值和最大值。

3.4 预测模型评价

本文选取 3 个指标用于评估预测模型的预测性能和数据特征，包括平均绝对值误差 (E_{MAE})，均方根误差 (E_{RMSE}) 和拟合优度 (R^2)。其中 E_{MAE} 和 E_{RMSE} 越小，表明预测结果的精确度越高， R^2 的值越接近于 1，预测结果越准确。其定义如下：

$$E_{MAE} = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \quad (21)$$

$$E_{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (22)$$

$$R^2 = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m \left(y_i - \frac{\sum_{i=1}^m y_i}{m} \right)^2} \times 100\% \quad (23)$$

式中： y_i 为真实光伏功率值； \hat{y}_i 为预测值； m 为测试样本集的数量。

4 实验仿真分析

4.1 数据来源

文中选取 DKASC Solar Centre 中某光伏电站 2020 年 1 月 1 日至 2020 年 10 月 30 日每天 7:00—18:00 光伏历史输出功率和对映的环境信息作为本研究的数据集，时间间隔为 5 min，并把后 20% 的数据集作为测试集。

4.2 实验结果分析

为了验证文中提出的 CEEMDAN-ChOA-ELM 组合预测模型方法的优越性以及 CEEMDAN 算法的有效性，进而说明所构建模型对光伏出力预测精度有提升作用，分别与 CEEMDAN-ELM 神经网络预测模型、ChOA-ELM 神经网络预测模

型进行预测对比。同时,采用传统反向传播(back propagation, BP)神经网络预测模型与单一 ELM 神经网络预测模型进行对比分析。

表 3 为上述各模型在测试样本数据集中的预测效果对比,由于 ELM 神经网络自适应输入层的权重与隐含层的偏置参数,对样本的适应能力强,在输入变量一致的情况下,与传统 BP 神经网络和单一 ELM 神经网络预测模型相比,采用 ChOA 优化的 ELM 神经网络预测模型各项指标均好于单一 ELM 神经网络和 BP 神经网络预测模型。CEEMDAN-ELM 预测模型随着测试样本的增多,特征过多,ELM 容易产生过拟合,预测效果下降,但是当采用 CEEMDAN-ChOA-ELM 组合预测模型后, E_{MAE} 和 E_{RMSE} 与单一 ELM 神经网络预测模型相比分别降低了 3.14% 和 2.74%。与 BP 神经网络预测模型相比降低了 5.12% 和 4.84%。与 ChOA-ELM 神经网络预测模型相比降低了 2.15% 和 1.88%。 R^2 则分别增加了 0.59%、1.08%、0.36%。由此可以看出,本文选用的 CEEMDAN-ChOA-ELM 神经网络组合预测模型有较好的预测效果。

为了更清楚地展示文章中选用的预测模型的预测效果,从测试数据中抽取 2020 年 10 月 1 号(晴天)和 2020 年 9 月 21 号(多云)的数据进行预测展示,分别采用上文提出的 4 种预测模型与本文提出的 CEEMDAN-ChOA-ELM 预测模型验证在晴天和多云天气下的预测精度。图 4(a)与(b)分别为晴天和多云天气下的 5 个模型预测对比图。表 4 与表 5 分别为晴天和多云天气下的

表 3 各模型测试集预测指标对比

Table 3 Comparison of predictive indicators of each model in test set

预测模型	$E_{MAE}/\%$	$E_{RMSE}/\%$	$R^2/\%$
BP	23.61	31.36	96.21
ELM	21.63	29.26	96.70
ChOA-ELM	20.64	28.40	96.93
CEEMDAN-ELM	53.26	75.82	77.82
CEEMDAN-ChOA-ELM	18.49	26.52	97.29

各项指标对比。由图 4 可以看出,在晴天天气下,由于环境波动不大,光伏出力较为稳定,5 个预测模型预测效果都较好,但是本文采用的 CEEMDAN-ChOA-ELM 神经网络预测模型 E_{MAE} 和 E_{RMSE} 为 6.55% 和 7.91%,拟合优度达到 99.72%,预测精度和各项指标更好。在多云天气下,由于太阳受到云层遮挡,太阳辐照度有较大变化,5 个预测模型各项指标均有所下降,但是由于本文采用的预测模型,对环境因素进行特征分解提取,预测精度和各项预测指标较其他 4 种预测模型更好,适合环境多变的光伏出力预测。

5 结论

1) 本文利用 CEEMDAN 算法将 4 种关键环境因素序列进行分解,克服了传统 EMD 算法重构误差的缺陷,将原始环境序列分解为不同特征的模态分量和剩余分量,得到环境序列数据中不同时间尺度特征。

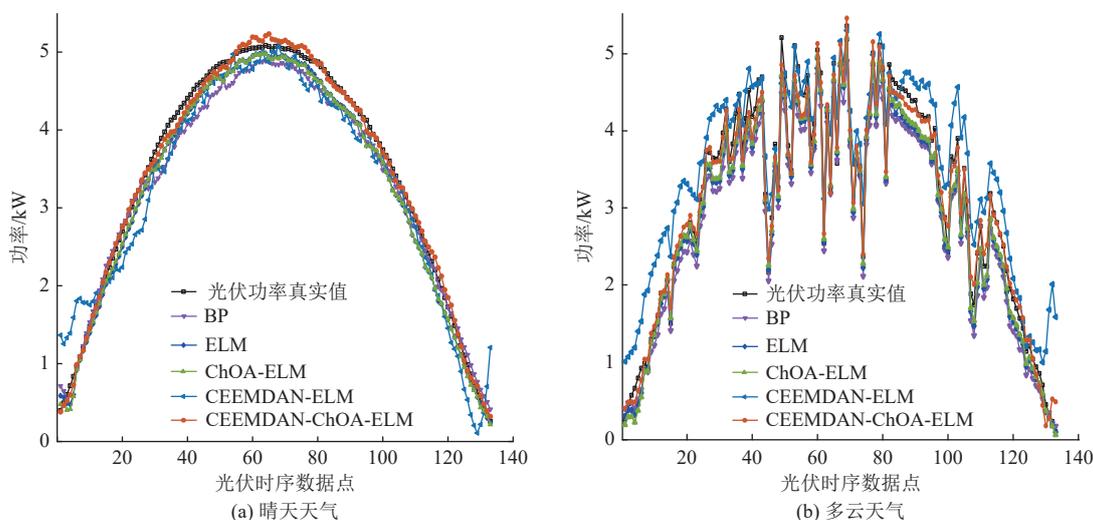


图 4 晴天和多云天气下各种预测模型结果对比

Fig. 4 Comparison of results from various prediction models under sunny and cloudy weather conditions

表 4 晴天天气下各模型预测指标对比

Table 4 Comparison of predictive indicators of various models in sunny weather

预测模型	$E_{MAE}/\%$	$E_{RMSE}/\%$	$R^2/\%$
BP	19.36	22.52	97.73
ELM	16.42	17.51	98.63
ChOA-ELM	15.42	16.64	99.01
CEEMDAN-ELM	28.59	34.28	94.74
CEEMDAN-ChOA-ELM	6.55	7.91	99.72

表 5 多云天气下各模型预测指标对比

Table 5 Comparison of predictive indicators of various models in cloudy weather

预测模型	$E_{MAE}/\%$	$E_{RMSE}/\%$	$R^2/\%$
BP	31.4	35.16	93.37
ELM	21.76	25.58	96.49
ChOA-ELM	19.34	22.75	97.23
CEEMDAN-ELM	45.6	54.87	83.86
CEEMDAN-ChOA-ELM	12.63	15.53	98.71

2) CEEMDAN-ChOA-ELN 神经网络预测模型在测试数据集中的预测精度最高, 预测效果最好。并且在晴天天气下拟合优度达到 99.72%, 在环境波动大的多云天气下各预测模型中预测效果最好。该组合预测模型适合环境多变的天气。

致谢

本文得到了冶金自动化检测技术教育部工程研究中心开放基金的资助, 在此表示感谢。

参考文献

- [1] LIU Hongfei, GAO Qian, MA Pengcheng. Photovoltaic generation power prediction research based on high quality context ontology and gated recurrent neural network[J]. Sustainable Energy Technologies and Assessments, 2021, 45(16): 2213–1388.
- [2] 龚莺飞, 鲁宗相, 乔颖, 等. 光伏功率预测技术[J]. 电力系统自动化, 2016, 40(4): 140–151.
GONG Yinfei, LU Zongxiang, QIAO Ying, *et al.* Photovoltaic power prediction technology[J]. Automation of Electric Power Systems, 2016, 40(4): 140–151 (in Chinese).
- [3] BERMUDEZ-GARCIA A, VOARINO P, RACCURT O *et al.* Environments, needs and opportunities for future space photovoltaic power generation: A review[J]. Applied Energy, 2021, 290(15): 2906–2916.
- [4] 管霖, 赵琦, 周保荣, 等. 基于多尺度聚类分析的光伏功率特性建模及预测应用[J]. 电力系统自动化, 2018, 42(15): 24–30.
GUAN Lin, ZHAO Qi, ZHOU Baorong, *et al.* Multi-scale clustering analysis based modeling of photovoltaic power characteristics and its application in prediction[J]. Automation of Electric Power Systems, 2018, 42(15): 24–30 (in Chinese).
- [5] BRI-MATHIAS H, MARTINEZ-ANIDO B, CARLO, *et al.* The combined value of wind and solar power forecasting improvements and electricity storage[J]. Applied Energy, 2018, 214(15): 1–15.
- [6] 黄予春, 曹成涛, 顾海. 基于IKFCM与多模式SSO优化SVR的光伏发电短期预测[J]. 电力系统保护与控制, 2018, 46(24): 102–109.
HUANG Yuchun, CAO Chengtao, GU Hai. Short-term photovoltaic power generation prediction based on IKFCM and multi-modes SSO optimization SVR[J]. Power System Protection and Control, 2018, 46(24): 102–109 (in Chinese).
- [7] WANG Kejun, QI Xiaoxia, LIU Hongda. A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network[J]. Applied Energy, 2019, 251(1): 1–14.
- [8] GU Wenbo, MA Tao, SONG Aotian, *et al.* Mathematical modelling and performance evaluation of a hybrid photovoltaic-thermoelectric system[J]. Energy Conversion and Management, 2019, 198(15): 1–13.
- [9] SHARADGA H, HAJIMIRZA S, BALOG R S. Time series forecasting of solar power generation for large-scale photovoltaic plants[J]. Renewable Energy, 2020, 150(12): 797–807.
- [10] CERVONE G, HARDING L C, ALESSANDRINI S, *et al.* Short-term photovoltaic power forecasting using artificial neural networks and an analog ensemble[J]. Renewable Energy, 2017, 108(1): 274–286.
- [11] 王新普, 周想凌, 邢杰, 等. 一种基于改进灰色BP神经网络组合的光伏发电预测方法[J]. 电力系统保护与控制, 2016, 44(18): 81–87.
WANG Xinpu, ZHOU Xiangling, XING Jie, *et al.* A prediction method of PV output power based on the combination of improved grey back propagation neural network[J]. Power System Protection and Control, 2016, 44(18): 81–87 (in Chinese).
- [12] WANG Huaizhi, YI Haiyan, PENG Jianchun, *et al.* Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network[J]. Energy Conversion and Management, 2017, 153(1):

- 409–422.
- [13] 崔佳豪, 毕利. 基于混合神经网络的光伏电量预测模型的研究[J]. *电力系统保护与控制*, 2021, 49(13): 142–149.
CUI Jiahao, BI Li. Research on photovoltaic power forecasting model based on hybrid neural network[J]. *Power System Protection and Control*, 2021, 49(13): 142–149(in Chinese).
- [14] 王琦, 季顺祥, 钱子伟, 等. 基于熵理论和改进ELM的光伏发电功率预测[J]. *太阳能学报*, 2020, 41(10): 151–158.
WANG Qi, JI Shunxiang, QIAN Ziwei, *et al.* Photovoltaic power prediction based on entropy theory and improved ELM[J]. *Acta Energetica Solaris Sinica*, 2020, 41(10): 151–158(in Chinese).
- [15] 左远龙, 黄玉水, 杨晓辉, 等. 基于PFA-MBAS-BP神经网络模型的光伏发电短期预测[J]. *电力系统保护与控制*, 2020, 48(15): 84–91.
ZUO Yuanlong, HUANG Yushui, YANG Xiaohui, *et al.* Short-term prediction of photovoltaic power generation based on a PFA-MBAS-BP neural network model[J]. *Power System Protection and Control*, 2020, 48(15): 84–91(in Chinese).
- [16] 张雲钦, 程起泽, 蒋文杰, 等. 基于EMD-PCA-LSTM的光伏功率预测模型[J]. *太阳能学报*, 2021, 42(9): 62–69.
ZHANG Yunqin, CHENG Qize, JIANG Wenjie, *et al.* Photovoltaic power prediction model based on EMD-PCA-LSTM[J]. *Acta Energetica Solaris Sinica*, 2021, 42(9): 62–69(in Chinese).
- [17] HUANG N E, SHEN Zheng, LONG S R, *et al.* The empirical mode decomposition and the Hilbert spectrum for non-linear and non-stationary time series analysis[J]. *Proceedings Mathematical Physical & Engineering Sciences*, 1998, 454(1971): 903–995.
- [18] COLOMINAS M A, SCHLOTTHAUER G, TORRES M E, *et al.* Noise-assisted EMD methods in action[J]. *Advances in Adaptive Data Analysis*, 2012, 4(4): 1–11.
- [19] KHISHE M, MOSAVI M R. Chimp optimization algorithm[J]. *Expert Systems with Applications*, 2020, 149(1): 1–14.
- [20] HUANG, Guangbin, ZHU Qinyu, SIEW C K. Extreme learning machine: Theory and applications[J]. *Neurocomputing*, 2006, 70(1-3): 489–501.
- [21] 田宏伟, 李志鹏, 王煜伟, 等. CEEMDAN-WOA-ELM模型风机振动趋势预测[J]. *中国测试*, 2020, 46(7): 146–152.
TIAN Hongwei, LI Zhipeng, WANG Yuwei, *et al.* Fan vibration trend prediction based on CEEMDAN-WOA-ELM model[J]. *China Measurement & Testing Technology*, 2020, 46(7): 146–152(in Chinese).

收稿日期: 2022–03–16

作者简介:

成燕(1972), 女, 博士, 副教授, 研究方向为微电网运行与控制, E-mail: chengyan100@163.com;

庄飞鸢(1996), 女, 硕士研究生, 通信作者, 研究方向为微电网运行与控制, E-mail: 1209035682@qq.com;

徐万万(1989), 女, 博士研究生, 研究方向为新能源、分布式发电和微电网运行与控制, E-mail: rosywan0114@163.com;

魏婷(1996), 女, 硕士研究生, 研究方向为用户用电行为分析, E-mail: 1987048448@qq.com。