

文章编号:1671-6833(2024)02-0089-08

# 基于灰狼算法和极限学习机的风速多步预测

张文煜<sup>1</sup>, 马可可<sup>2</sup>, 郭振海<sup>3</sup>, 赵晶<sup>3</sup>, 邱文智<sup>2</sup>

(1. 郑州大学 地球科学与技术学院, 河南 郑州 450001; 2. 郑州大学 计算机与人工智能学院, 河南 郑州 450001; 3. 中国科学院大气物理研究所 大气科学和地球流体力学数值模拟国家重点实验室, 北京 100029)

**摘要:** 为了提高风速的多步预测水平, 提出了一种基于数据信号分解和灰狼算法优化极限学习机的混合预测模型。首先, 使用具有自适应噪声的完全集成经验模态分解算法将原始风速时间序列分解为若干本征模态函数和一个残差序列, 并使用偏自相关函数法对模型输入进行特征选择; 其次, 在分解子序列上分别建立模型并进行预测, 构造多输入多输出策略的极限学习机神经网络, 使用灰狼优化算法求解其中的最优化隐含层权值和偏置; 最后, 对子序列进行重构并得到最终的预测结果。使用时间分辨率为 15 min 的多组实测资料开展模拟实验, 所提模型在 3 个风电场的均方根误差分别为 0.859、0.925、0.927 m/s, 均低于其他对比模型, 验证了该模型在未来 4 h 风速预测即 16 步预测中的有效性。

**关键词:** 风速预测; 多步预测; 信号分解; 特征选择; 灰狼优化算法; 极限学习机

**中图分类号:** TM614; TP183

**文献标志码:** A

**doi:** 10.13705/j.issn.1671-6833.2023.05.008

风能作为一种清洁的可再生能源, 开发潜力巨大。近年来, 中国风电装机总容量迅速增加, 但由于风具有间歇性和波动性<sup>[1-3]</sup>, 当大规模风电接入电网时, 将给电网调度带来困难, 解决这一问题的有效方式之一是对未来风速变化的准确预测。随着人工智能的发展, 人工神经网络 (artificial neural networks, ANN) 因具有强大的非线性数据拟合能力在风速时间序列的建模和预测中颇具优势<sup>[4]</sup>。

已有研究大多是对风速进行一步或少数几步预测, 而对于大步长风速预测的研究还处于初级阶段, 特别是用于风电场的 4 h、16 步预测<sup>[5]</sup>。Liu 等<sup>[6]</sup>使用一种基于谱聚类和回声状态网络的方法预测未来 16 步风速。Zhao 等<sup>[7]</sup>建立了 16 个一维卷积神经网络来预测未来 16 步风速, 提供了 16 步风速预测的一种可行思路, 但计算开销较大。

鉴于此, 本文使用极限学习机 (extreme learning machine, ELM)<sup>[8]</sup> 作为风速时间序列的预测模型, 并使用多输入-多输出策略建立 ELM。ELM 是一种特殊的单隐层前馈神经网络, 与传统的前馈网络使用梯度下降和误差反向传播进行网络更新的方式不

同, ELM 随机地生成隐含层的权值和偏置并解析地计算隐含层到输出层的权值参数。因此, ELM 具有计算速度快、泛化性能强等优点, 适用于风速预测问题<sup>[9]</sup>。一些研究表明, 使用人工智能优化算法求解最优隐含层参数可以降低 ELM 的输出误差<sup>[10]</sup>。Mirjalili 等<sup>[11]</sup>的研究结果显示, 与粒子群优化、遗传算法等经典的启发式算法相比, 灰狼优化 (grey wolf optimization, GWO) 算法<sup>[12]</sup> 在求解单峰和多峰函数问题时都更具优势。因此, 本文采用 GWO 算法搜索 ELM 隐含层参数的最优值, 并进一步计算输出层权值。

另一方面, 风速序列本身具有复杂的非线性特征, 单一模型很难取得满意的预测结果。使用信号分解方法将原始数据序列分解成若干具有不同数据特征的子序列的集合, 是改进模型预测结果的有效方法。常用的分解方法有经验模态分解 (empirical mode decomposition, EMD)<sup>[13]</sup>、变分模态分解<sup>[14]</sup> 和奇异谱分析等。其中, EMD 是应用最为广泛的方法, 它能够根据数据自身的时间尺度特征来识别信号中包含的振动模态, 并将复杂的原始信号分解为

收稿日期: 2023-04-11; 修订日期: 2023-06-02

基金项目: 国家自然科学基金资助项目 (41875085)

通信作者: 赵晶 (1988—), 女, 河南温县人, 中国科学院大气物理研究所副研究员, 博士, 主要从事数值模拟与人工智能交叉研究, E-mail: zhaojing@lasg.iap.ac.cn。

引用本文: 张文煜, 马可可, 郭振海, 等. 基于灰狼算法和极限学习机的风速多步预测[J]. 郑州大学学报(工学版), 2024, 45(2): 89-96. (ZHANG W Y, MA K K, GUO Z H, et al. Multistep prediction of wind speed based on grey wolf algorithm and extreme learning machine [J]. Journal of Zhengzhou University (Engineering Science), 2024, 45(2): 89-96.)

有限个本征模态函数 (intrinsic mode function, IMF)。与小波分解等方法不同, EMD 不需要预先设定任何基函数, 在处理非平稳及非线性数据时具有优势, 但 EMD 存在模态混叠现象。为了解决模态混叠现象, 具有自适应噪声的完全集成经验模态分解 (complete ensemble empirical mode decomposition with adaptive noise, CEEMDAN) [15] 算法在分解过程的每个阶段自适应地加入白噪声, 并计算唯一的余量信号, 其重构误差极低, 而且能够产生更好的模态分离结果。韩宏志等 [16] 使用 CEEMDAN 分解原始风速信号, 将分解结果作为回声状态网络的输入并建立预测模型。Ren 等 [17] 建立 CEEMDAN 与 ANN 的混合风速预测模型, 并与 EMD-ANN、EEMD-ANN 等模型进行比较, 结果表明: CEEMDAN-ANN 模型具有更高的预测精度。

基于以上分析, 本文构建了 CEEMDAN-GWO-ELM 混合模型, 用于风速时间序列的多步预测。使用 CEEMDAN 算法将原始风速分解为若干具有不同数据特征的子序列的集合, 在各子序列上建立 GWO 优化的 ELM 模型并进行多步预测, 最终的预测结果由各子序列的预测重构生成。将提出的 CEEMDAN-GWO-ELM 模型应用于实际风电场的 4 h 风速预测问题, 进行了预测步长为 16 的模拟试验和模型比较试验, 用于验证该模型在大步长的风速预测问题中的有效性和适用性。

## 1 数据预处理与特征选择

### 1.1 数据预处理

#### 1.1.1 具有自适应噪声的完全集成经验模态分解 (CEEMDAN)

CEEMDAN 通过在每个分解阶段为唯一的余量信号添加高斯白噪声后再进行 EMD 分解, 其分解过程具有完整性 [18], 能以较低的计算成本提供原始信号的精确重建和更好的模式频谱分离。定义操作符  $M_k(\cdot)$  为通过 EMD 算法得到的第  $k$  个模态分量;  $\mathbf{v}^i(t)$  为符合标准正态分布的高斯白噪声;  $\varepsilon$  为高斯白噪声的标准差。CEEMDAN 的计算过程如下。

步骤 1 设  $s(t)$  为原始信号, 对信号  $s(t) + \varepsilon_0 \mathbf{v}^i(t)$  进行  $I$  次试验, 通过 EMD 获取第一个模态分量:

$$\mathbf{IMF}_1(t) = \frac{1}{I} \sum_{i=1}^I \mathbf{IMF}_1^i(t)。 \quad (1)$$

步骤 2 在第 1 阶段 ( $k=1$ ), 计算唯一的残差项:

$$\mathbf{r}_1(t) = s(t) - \mathbf{IMF}_1(t)。 \quad (2)$$

步骤 3 以  $\mathbf{r}_1(t)$  作为新的原始信号, 进行  $I$  次试验, 每次试验使用 EMD 算法分解新构建的  $\mathbf{r}_1(t) + \varepsilon_1 \mathbf{M}_1(\mathbf{v}^i(t))$  信号, 计算第 2 个模态分量:

$$\mathbf{IMF}_2(t) = \frac{1}{I} \sum_{i=1}^I \mathbf{M}_1(\mathbf{r}_1(t) + \varepsilon_1 \mathbf{M}_1(\mathbf{v}^i(t)))。 \quad (3)$$

步骤 4 对其余各阶段即  $k=2, 3, \dots, K$ , 与步骤 2 和步骤 3 的计算过程一致。首先计算第  $k$  个残差项  $\mathbf{r}_k(t)$ , 再计算第  $k+1$  个模态分量  $\mathbf{IMF}_{k+1}(t)$ :

$$\mathbf{r}_k(t) = \mathbf{r}_{k-1}(t) - \mathbf{IMF}_k(t); \quad (4)$$

$$\mathbf{IMF}_{k+1}(t) = \frac{1}{I} \sum_{i=1}^I \mathbf{M}_1(\mathbf{r}_k(t) + \varepsilon_k \mathbf{M}_k(\mathbf{v}^i(t)))。 \quad (5)$$

步骤 5 重复执行步骤 4, 直至残差项无法满足分解条件。此时原始信号  $s(t)$  被分解为

$$s(t) = \sum_{k=1}^K \mathbf{IMF}_k(t) + \mathbf{R}(t); \quad (6)$$

$$\mathbf{R}(t) = s(t) - \sum_{k=1}^K \mathbf{IMF}_k(t)。 \quad (7)$$

式中:  $K$  为分解得到的模态个数;  $\mathbf{R}(t)$  为最终的残差项。

#### 1.1.2 归一化

使用最大最小归一化法将模态分量的值  $\mathbf{z} = [z_1, z_2, \dots, z_i, \dots, z_n]$  归一化到  $[0, 1]$  内, 如下式:

$$z_i' = \frac{z_i - z_{\min}}{z_{\max} - z_{\min}}。 \quad (8)$$

式中:  $z_i'$  为归一化后的元素值;  $z_{\max}$  和  $z_{\min}$  分别为模态分量中的最大值和最小值;  $n$  为模态分量的序列长度。

### 1.2 输入数据的特征选择

输入数据的窗口大小影响模型的预测水平, 这是因为窗口过大会引入冗余信息, 而窗口太小会丢失部分有用的信息 [19]。本文使用偏自相关函数 (partial autocorrelation function, PACF) [20] 确定模型输入数据的窗口大小。

具体地, 对于 IMF 子序列  $\mathbf{z}(t)$ , 计算其不同滞后阶数下的偏自相关函数。如果  $\mathbf{z}(t)$  滞后  $s$  阶的偏自相关函数值超出 95% 的置信区间  $[-1.96/\sqrt{n}, 1.96/\sqrt{n}]$ , 则选择  $\mathbf{z}(t-s)$  作为模型的输入数据特征 [21]。

## 2 灰狼算法优化的极限学习机模型

### 2.1 极限学习机 (ELM)

ELM 是一种单隐层前馈神经网络, 假设它的隐含层节点数为  $L$ , 输出层的节点数为  $m$ , 对于输入样本  $(\mathbf{x}_i, \mathbf{y}_i)$ ,  $i=1, 2, \dots, N$ ,  $N$  为样本总数,  $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{id}]^T$  为  $d$  维输入数据,  $\mathbf{y}_i = [y_{i1}, y_{i2}, \dots, y_{im}]^T$

为  $m$  维输出数据,ELM 隐含层的输出可表示为

$$\sum_{k=1}^L \beta_k G(\mathbf{w}_k \cdot \mathbf{x}_i + b_k) = \mathbf{o}_i, i = 1, 2, \dots, N. \quad (9)$$

式中:  $G$  为隐含层节点的激活函数,通常使用 sigmoid 函数;  $\mathbf{w}_k = [w_{i1}, w_{i2}, \dots, w_{id}]^T$  为输入节点与隐含层节点间的权值向量;  $b_k$  为隐含层节点的偏置;  $\beta_k = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$  是隐含层节点与输出层节点间的权值向量。

令 ELM 的输出值零误差地逼近给定的  $N$  个样本,即

$$\sum_{i=1}^N \|\mathbf{o}_i - \mathbf{y}_i\| = 0. \quad (10)$$

也就是说,存在  $\beta_k, \mathbf{w}_k, b_k$  使得

$$\sum_{k=1}^L \beta_k G(\mathbf{w}_k \cdot \mathbf{x}_i + b_k) = \mathbf{y}_i, i = 1, 2, \dots, N. \quad (11)$$

式(11)可以简化为

$$\mathbf{H}\beta = \mathbf{Y}. \quad (12)$$

式中:  $\mathbf{H}$  为 ELM 的隐含层输出矩阵;  $\beta$  为隐含层与输出层之间的权值矩阵;  $\mathbf{Y}$  为 ELM 的理想输出,分别表示为

$$\mathbf{H} = \begin{bmatrix} G(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & G(\mathbf{w}_L \cdot \mathbf{x}_1 + b_L) \\ \vdots & \vdots & \vdots \\ G(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \cdots & G(\mathbf{w}_L \cdot \mathbf{x}_N + b_L) \end{bmatrix}_{N \times L}; \quad (13)$$

$$\beta = [\beta_1^T \cdots \beta_L^T]^T; \quad (14)$$

$$\mathbf{Y} = [\mathbf{y}_1^T \cdots \mathbf{y}_N^T]^T. \quad (15)$$

ELM 的训练过程相当于求解出  $\beta$  矩阵,其解可表示为

$$\hat{\beta} = \mathbf{H}^+ \mathbf{Y} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{Y}. \quad (16)$$

式中:  $\mathbf{H}^+$  为  $\mathbf{H}$  的摩尔-彭罗斯 (Moore-Penrose) 广义逆矩阵。

## 2.2 灰狼优化算法 (GWO)

GWO 是一种元启发式算法,其灵感来源于灰狼,本质上模仿了灰狼的领导阶层和狩猎机制。GWO 将每一匹狼作为一个解,设  $\alpha$  为全局最优解,  $\gamma$  和  $\delta$  为全局第二和第三优解,其余候选解记为  $\omega$ 。  $\omega$  狼 (搜索代理) 在  $\alpha, \gamma$  及  $\delta$  狼的带领下进行搜索和狩猎,主要过程可表示为

$$\mathbf{D} = |\mathbf{C} \cdot \mathbf{X}_p(e) - \mathbf{X}(e)|; \quad (17)$$

$$\mathbf{A} = 2ar_1 - a; \quad (18)$$

$$\mathbf{C} = 2r_2; \quad (19)$$

$$a = \frac{2(E - e)}{E}. \quad (20)$$

式中:  $\mathbf{X}$  和  $\mathbf{X}_p$  分别为搜索代理和猎物的位置向量;  $\mathbf{D}$  表示搜索代理与猎物的距离;  $e$  为当前迭代次数;  $\mathbf{A}$  和  $\mathbf{C}$  为系数向量;  $r_1$  和  $r_2$  为  $[0, 1]$  随机数;  $E$  为最

大迭代次数;  $a$  为收敛因子,在迭代过程中从 2 线性减小至 0。

搜索代理  $\mathbf{X}$  与  $\alpha, \gamma$  及  $\delta$  狼的距离确定为

$$\begin{cases} \mathbf{D}_\alpha = |\mathbf{C}_1 \cdot \mathbf{X}_\alpha(e) - \mathbf{X}(e)|; \\ \mathbf{D}_\gamma = |\mathbf{C}_2 \cdot \mathbf{X}_\gamma(e) - \mathbf{X}(e)|; \\ \mathbf{D}_\delta = |\mathbf{C}_3 \cdot \mathbf{X}_\delta(e) - \mathbf{X}(e)|. \end{cases} \quad (21)$$

则搜索代理在下一迭代的位置由下式更新。

$$\mathbf{X}(e+1) = \frac{1}{3} [(\mathbf{X}_\alpha(e) - \mathbf{A}_1 \mathbf{D}_\alpha) + (\mathbf{X}_\gamma(e) - \mathbf{A}_2 \mathbf{D}_\gamma) + (\mathbf{X}_\delta(e) - \mathbf{A}_3 \mathbf{D}_\delta)]. \quad (22)$$

## 2.3 极限学习机的参数优化

本文使用 GWO 选取使得网络输出误差最小的隐含层权值和偏置,通过搜索最优化的隐含层权值和偏置,进一步改进 ELM 模型的多步预测结果。基于 GWO 优化的 ELM 模型的具体步骤如下。

步骤 1 初始化 GWO 参数。其中,最大迭代次数  $E = 2000$ ,搜索代理的数量  $C = 5$ ,  $\alpha, \gamma, \delta$  狼的适应度值  $F_\alpha, F_\gamma, F_\delta$  设置为  $+\infty$ ,使用  $[0, 1]$  中的随机数初始化每个搜索代理的位置向量  $\mathbf{X}$ 。

步骤 2 对于第  $c$  个搜索代理 ( $c = 1, 2, \dots, C$ ), 用其位置向量  $\mathbf{X}_c$  初始化 ELM 的隐含层权重和偏置,在训练集上求解隐含层与输出层之间的权值矩阵  $\beta$ 。在验证集上,计算该灰狼的适应度函数值  $F_c$ :

$$F_c = \sum_{i=1}^N |y_i - \hat{y}_i|. \quad (23)$$

式中:  $y_i$  为真实值;  $\hat{y}_i$  为预测值;  $N$  为样本总数。

如果该搜索代理的  $F_c$  满足  $F_c < F_\alpha$ ,则将  $\alpha$  狼的  $F_\alpha, \mathbf{X}_\alpha$  更新为  $F_c, \mathbf{X}_c$ ; 如果该搜索代理的  $F_c$  满足:  $F_\alpha < F_c < F_\beta$ ,则将  $\gamma$  狼的  $F_\gamma, \mathbf{X}_\gamma$  更新为  $F_c, \mathbf{X}_c$ ; 如果该个体的  $F_c$  满足  $F_\gamma < F_c < F_\delta$ ,则将  $\delta$  狼的  $F_\delta, \mathbf{X}_\delta$  更新为  $F_c, \mathbf{X}_c$ 。

步骤 3 对于第  $c$  个搜索代理 ( $c = 1, 2, \dots, C$ ), 使用式(21)和式(22)更新位置向量  $\mathbf{X}_c$ 。

步骤 4 重复执行步骤 2 和步骤 3,直到达到最大迭代次数  $E$ ,输出  $\alpha$  狼的位置向量  $\mathbf{X}_\alpha$  作为 ELM 的最优隐含层权重和偏置。

## 3 CEEMDAN-GWO-ELM 混合预测模型

本文提出的风速预测模型是由 CEEMDAN、GWO 及 ELM 组成的混合模型,记为 CEEMDAN-GWO-ELM,图 1 为混合预测模型流程图。具体步骤如下。

步骤 1 通过 CEEMDAN 将原始风速序列分解为  $K$  个 IMF 子序列和一个残差项。

步骤 2 对分解得到的子序列进行最大最小归一化。

步骤3 计算各子序列在不同滞后阶数下的PACF值,并对模型输入特征进行选择。

步骤4 在每个子序列上,分别建立GWO优化的ELM模型,模型输入由步骤3确定。对未来4 h的风速变化情况进行预测,即16步预测,并对预测结果进行反归一化。

步骤5 对各子序列的预测结果进行重构,得到原始风速序列的混合预测结果。

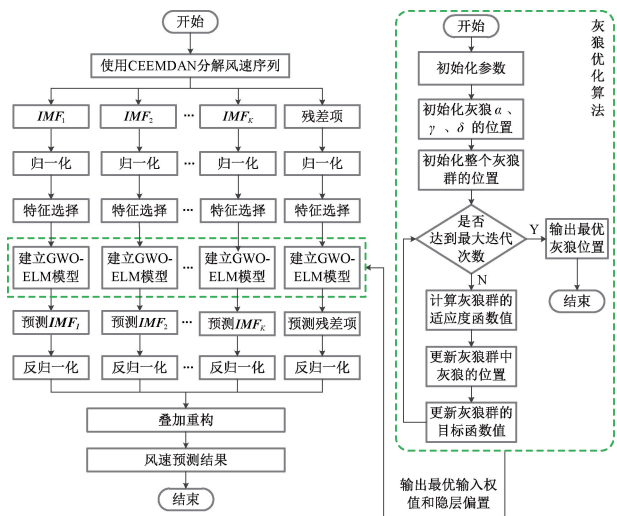


图1 CEEMDAN-GWO-ELM混合预测模型流程图

Figure 1 Flow chart of CEEMDAN-GWO-ELM hybrid forecasting model

CEEMDAN将风速序列分解为若干不同频率特征的子序列,以削弱风速序列的非线性<sup>[22]</sup>,从而提高模型的预测水平。由于ELM的隐含层权值和偏

置是随机产生的,因此可以使用GWO寻找ELM的最优隐含层权值和偏置,提高ELM预测的精度和稳定性。

## 4 算例分析

### 4.1 数据集

本文使用的数据来自中国山东省3个风电场时间间隔为15 min风速观测资料,如图2所示,它们的统计特征如表1所示。对每一个风电场,将原始数据按照7:2:1划分为训练集、验证集及测试集,其中训练集用于建立模型,测试集用于模型评价,而在GWO优化ELM参数过程中,将根据ELM在验证集上的适应度值来选择ELM的隐含层参数。

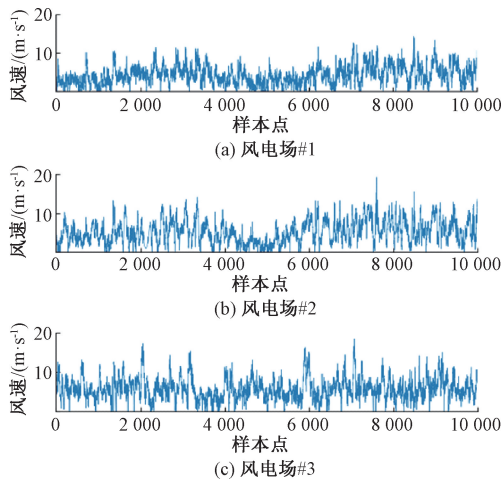


图2 3个风电场的风速序列

Figure 2 Wind speed sequence for three wind farms

表1 风速序列的统计特征

数据集	起止日期	样本量	均值/(m·s <sup>-1</sup> )	标准差/(m·s <sup>-1</sup> )	最小值/(m·s <sup>-1</sup> )	最大值/(m·s <sup>-1</sup> )
风电场#1	2019-01-01—2019-04-15	10 000	4.227	2.339	0	14.233
风电场#2	2019-01-01—2019-04-15	10 000	5.112	2.339	0	19.252
风电场#3	2021-02-01—2021-05-16	10 000	5.780	2.729	0	18.435

### 4.2 模型评价指标

本文使用均方根误差(RMSE)和误差方差σ来评价预测模型的准确性和稳定性。其表达式为

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}; \quad (24)$$

$$\sigma = \frac{\sum_{i=1}^n (e_i - \bar{e})^2}{n}. \quad (25)$$

式中: $y_i$ 为真实值; $\hat{y}_i$ 为预测值; $n$ 为序列长度;误差 $e_i = |y_i - \hat{y}_i|$ 。

### 4.3 试验设置

分别使用3个风电场的数据集建立图1中的预

测模型,进行4 h风速预测,即16步预测。

为了验证本文建立的CEEMDAN-GWO-ELM模型的有效性和适用性,选取了4个不同的比较模型。具体地,选取ELM和CEEMDAN-ELM作为CEEMDAN-GWO-ELM的关联模型,用于验证基于CEEMDAN的序列分解方法和基于GWO的参数优化方法对提高ELM模型预测水平的有效性。此外,选择了基于CEEMDAN分解的LSTM<sup>[23]</sup>、卷积LSTM和时域卷积网络(temporal convolutional network, TCN)<sup>[17]</sup>,分别记为CEEMDAN-LSTM、CEEMDAN-ConvLSTM和CEEMDAN-TCN,用于与CEEMDAN-GWO-ELM的对比分析。上述模型的参数设置如表2所示。



表 2 模型的参数设置

Table 2 Parameter setting of the model			
模型	模块	参数	取值
ELM	ELM	隐含层神经元数量	16
CEEMDAN-GWO-ELM	GWO	最大迭代次数	2 000
		灰狼个数	5
	ELM	隐含层神经元数量	16
CEEMDAN-LSTM	LSTM	隐含层层数	1
		隐含层神经元数量	200
CEEMDAN-ConvLSTM	Conv1D	卷积核的数量	32
		卷积核大小	3
	LSTM	隐含层层数	1
		隐含层的神经元数量	200
CEEMDAN-TCN	TCN	卷积核的数量	64
		卷积核大小	3

4.4 个例分析

以风电场#1 为例进行分析。首先使用 CEEMDAN 对原始风速序列进行分解,得到 11 个 *IMF*<sub>*s*</sub> 和一个残差项。其中,残差项的值域约为  $[-3.553 \times 10^{-15}, 3.553 \times 10^{-15}]$ ,其数量级远小于原始序列取值的数量级,对预测结果几乎没有影响,因此,不对该残差项进行建模预测。

计算原风速序列和 11 个 *IMF*<sub>*s*</sub> 的偏自相关函数值,选择偏自相关函数值超出 95% 置信区间的滞后阶数所对应的变量作为输入特征,如表 3 所示。由表 3 可以看出,对不同的 *IMF* 子序列,模型所选择的输入特征也是不同的。

表 3 输入变量选择结果

Table 3 Input variable selection results	
时间序列	输入特征
原始	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-8}, x_{t-10}, x_{t-16}$
<i>IMF</i> <sub>1</sub>	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-6}, x_{t-9}$
<i>IMF</i> <sub>2</sub>	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-7}, x_{t-8}, x_{t-10},$ $x_{t-11}, x_{t-12}, x_{t-13}, x_{t-15}, x_{t-17}, x_{t-21}$
<i>IMF</i> <sub>3</sub>	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-7}, x_{t-9}, x_{t-10},$ $x_{t-13}, x_{t-14}, x_{t-15}, x_{t-17}, x_{t-20}$
<i>IMF</i> <sub>4</sub>	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-7}, x_{t-10}, x_{t-12},$ $x_{t-13}, x_{t-14}, x_{t-18}, x_{t-19}, x_{t-21}, x_{t-22}$
<i>IMF</i> <sub>5</sub>	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-8},$ $x_{t-11}, x_{t-12}, x_{t-13}, x_{t-16}, x_{t-20}, x_{t-22}, x_{t-24}$
<i>IMF</i> <sub>6</sub>	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-9}, x_{t-11},$ $x_{t-12}, x_{t-16}, x_{t-17}, x_{t-19}, x_{t-20}, x_{t-21}, x_{t-23}$
<i>IMF</i> <sub>7</sub>	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6},$ $x_{t-7}, x_{t-16}, x_{t-18}, x_{t-19}, x_{t-23}$
<i>IMF</i> <sub>8</sub>	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-8}, x_{t-9}, x_{t-10}$
<i>IMF</i> <sub>9</sub>	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}, x_{t-13},$ $x_{t-14}, x_{t-16}, x_{t-17}, x_{t-22}, x_{t-23}, x_{t-24}$
<i>IMF</i> <sub>10</sub>	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}$
<i>IMF</i> <sub>11</sub>	$x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5}, x_{t-6}, x_{t-7}$

根据表 3 中确定的输入特征,在每一个 *IMF* 子序列上建立 GWO-ELM 模型,并对子序列的预测结果进行重构。

图 3 给出了使用 CEEMDAN-GWO-ELM 模型在风电场#1 的 4 个仿真案例,左边为未来 16 步风速的实际值与 CEEMDAN-GWO-ELM 模型的预测值,右边为每步预测值与实际值的误差。图 3 中 4 个仿真案例的预测起始时间分别为 2019-04-05T06:00:00、2019-04-05T02:00:00、2019-04-05T22:00:00 和 2019-04-06T10:00:00。由图 3 可以看出,该方法能够较好地预测未来 4 h 的风速变化情况,且随着预测步长的增加,模型的预测偏差始终保持在一定的范围内。这表明本文建立的模型在大步长的风速预测问题中能够产生可靠的预测结果。

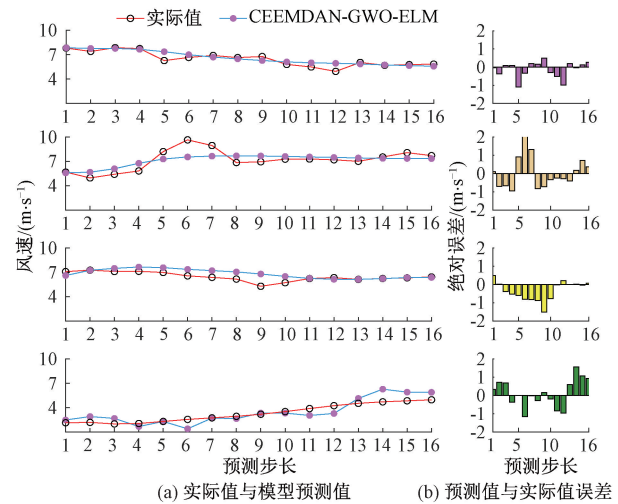


图 3 CEEMDAN-GWO-ELM 预测风电场#1 的仿真案例  
Figure 3 Simulation case of wind farm #1 prediction of CEEMDAN-GWO-ELM

4.5 模型比较分析

表 4 对比了不同模型对 3 个风电场进行预测的误差指标情况。首先,比较 ELM 和 CEEMDAN-ELM 模型的预测结果。可以看出,使用 CEEMDAN 对原始风速序列进行分解能够显著提高模型的预测水平。具体地,相比于 ELM,CEEMDAN-ELM 在风电场#1 的 *RMSE* 下降了 37.6%,在其他风电场上的预测精度也有相当大的提升。同时,相比于 ELM 模型,CEEMDAN-ELM 在风电场#1~#3 的误差方差分别降低了 52.7%、75.1% 和 72.9%,这表明 CEEMDAN 分解方法能够在降低模型预测误差的同时提高预测的稳定性。这是因为风速时间序列具有复杂的非线性特征,使用单一模型时通常难以取得满意的预测结果。CEEMDAN 将原始风速序列分解为若干具有不同频率特征的子序列,使得预测模型能够更好地描述各数据序列的变化特征,从而提高模型

预测水平。

在对 CEEMDAN-GWO-ELM 和 CEEMDAN-ELM 的比较中能够发现,通过搜索 ELM 的最优隐含层权值和偏置,模型的预测水平得到了进一步提升。结果显示,基于 GWO 的参数优化使得 3 个风电场的

$RMSE$  分别下降了 13.8%、9.3%和 5.5%, $\sigma$  分别下降了 26.9%、18.8%和 8.9%。这表明与随机初始化 ELM 隐含层的权值和偏置相比,优化选取的隐含层参数值能够改进 ELM 的网络性能。本文使用的 GWO 算法提供了搜索上述最优化参数的有效途径。

表 4 不同模型对 3 个风电场的预测结果

Table 4 Forecasting result of different models at three wind farms

模型	风电场#1		风电场#2		风电场#3	
	$RMSE/(m \cdot s^{-1})$	$\sigma$	$RMSE/(m \cdot s^{-1})$	$\sigma$	$RMSE/(m \cdot s^{-1})$	$\sigma$
ELM	1.597	0.928	2.059	1.820	1.792	1.492
CEEMDAN-LSTM	1.153	0.565	1.290	0.760	1.115	0.525
CEEMDAN-ConvLSTM	0.937	0.370	1.100	0.500	1.001	0.429
CEEMDAN-TCN	0.890	0.311	1.070	0.519	1.005	0.433
CEEMDAN-ELM	0.996	0.439	1.020	0.453	0.981	0.404
CEEMDAN-GWO-ELM	0.859	0.321	0.925	0.368	0.927	0.368

进一步分析基于 CEEMDAN 分解的 LSTM、ConvLSTM 及 TCN 在 16 步风速预测中的表现。表 4 显示,CEEMDAN-ConvLSTM 和 CEEMDAN-TCN 的预测结果整体上优于 CEEMADN-LSTM,这表明在处理风速时间序列的多步预测问题时,卷积网络具有一定的优势,这与文献[5]和[7]的结果一致。与本文提出的模型相比,CEEMDAN-TCN 在风电场#1 的误差方差略低于 CEEMDAN-GWO-ELM,表现出更强的稳定性。除此之外,本文提出的 CEEMDAN-GWO-ELM 模型对 3 个风电场的全部评价指标均优于所有的比较模型,表明该混合模型在风速时间序列的多步预测问题中具有较强的适用性。

4.6 模型的预测性能随预测步长的变化分析

使用统计模型进行多步预测时,随着预测步长的增加,模型的预测精度和稳定性通常会逐渐下降。这是因为统计模型是基于历史数据集的内在统计规律建立的,并对未来变化进行预估。在多步预测过程中,模型所产生的预测值与真实值之间的偏差会随预测步长的增加不断累积,导致预测误差逐渐增大甚至发散。分析误差随预测步长的变化情况,对讨论多步预测模型在实际应用中的有效性和适用性是颇具意义的。图 4 直观展示了不同预测模型对 3 个风电场的  $RMSE$  和  $\sigma$  随预测步长的变化情况。

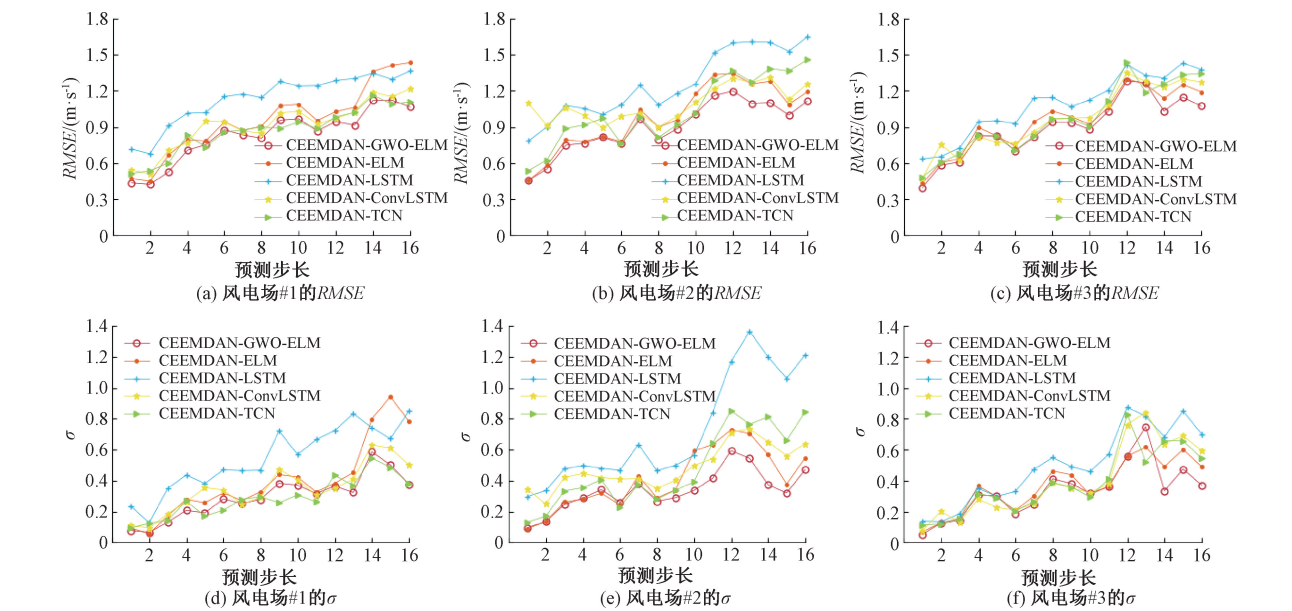


图 4 不同模型在 3 个风电场的误差指标随着预测步长的变化趋势

Figure 4 Variation trend of the error indicators of different models in three wind farms with the predicted step

整体来看,在 3 个风电场中,所有模型的  $RMSE$  和  $\sigma$  都随着预测步长的增加而逐渐增大。以  $RMSE$  为例,在风电场#1 的预测结果中,各模型的  $RMSE$  误差曲线较为集中,且 CEEMDAN-TCN 和 CEEMDAN-GWO-ELM 的  $\sigma$  更小。特别地,在预测步长大于 12 步的情况下(即大于 3 h 的预测),本文提出的 CEEMDAN-GWO-ELM 模型在  $RMSE$  和  $\sigma$  的比较中都更具优势。在风电场#2 和#3 的预测结果中,各模型之间具有明显差异。例如在风电场#2 中,CEEMDAN-GWO-ELM 对未来第 4 h 风速预测的  $RMSE$  和  $\sigma$  分别为 1.116 4 和 0.470 7,分别比排名第二的 CEEMDAN-ELM 降低了 6.4% 和 14.1%;而在风电场#3 中,CEEMDAN-GWO-ELM 对未来第 4 h 风速预测的  $RMSE$  和  $\sigma$  分别为 1.073 8 和 0.371 1,分别比排名第二的 CEEMDAN-ELM 降低了 9.8% 和 24.2%。从图 4(b)~4(f) 也可以看出,在大于 3 h 的预测中,CEEMDAN-GWO-ELM 模型几乎在所有的预测步长上都有更低的  $RMSE$  和  $\sigma$ 。这表明,所提出的 CEEMDAN-GWO-ELM 模型能够更好地描述风速时间序列的复杂变化特征,能够在大步长的预测问题中产生可靠的结果。

5 结论

- (1)利用 CEEMDAN 将原始风速序列分解为若干具有不同频率特征的子序列能够降低模型的预测难度。
  - (2)利用 GWO 搜索 ELM 的最优隐含层权值和偏置能够提高 ELM 的预测精度和稳定性。
  - (3)在 3 个风电场进行模型比较实验中,提出的 CEEMDAN-GWO-ELM 混合模型都具有最高的预测水平,验证了该模型具有一定的适用性。
- 在本文的研究中,CEEMDAN 的分解结果存在高频子序列,一定程度上影响模型的预测精度,因此,下一步将在降低高频子序不规则性的基础上改进预测模型。

参考文献:

[1] ZHAO W G, WEI Y M, SU Z Y. One day ahead wind speed forecasting: a resampling-based approach[J]. Applied Energy, 2016, 178: 886-901.

[2] XU Y Y, YANG G K, LUO J L, et al. A multi-location short-term wind speed prediction model based on spatio-temporal joint learning[J]. Renewable Energy, 2022, 183: 148-159.

[3] XU W F, LIU P, CHENG L, et al. Multi-step wind speed prediction by combining a WRF simulation and an

error correction strategy[J]. Renewable Energy, 2021, 163: 772-782.

[4] WANG Y, ZOU R M, LIU F, et al. A review of wind speed and wind power forecasting with deep neural networks[J]. Applied Energy, 2021, 304: 117766.

[5] CHEN G G, LI L J, ZHANG Z Z, et al. Short-term wind speed forecasting with principle-subordinate predictor based on Conv-LSTM and improved BPNN[J]. IEEE Access, 2020, 8: 67955-67973.

[6] LIU D, WANG J L, WANG H. Short-term wind speed forecasting based on spectral clustering and optimised echo state networks[J]. Renewable Energy, 2015, 78: 599-608.

[7] ZHAO X Y, JIANG N, LIU J F, et al. Short-term average wind speed and turbulent standard deviation forecasts based on one-dimensional convolutional neural network and the integrate method for probabilistic framework[J]. Energy Conversion and Management, 2020, 203:112239.

[8] HUANG G B, WANG D H, LAN Y. Extreme learning machines: a survey[J]. International Journal of Machine Learning and Cybernetics, 2011, 2(2): 107-122.

[9] 朱抗, 杨洪明, 孟科. 基于极限学习机的短期风力发电预测[J]. 电力科学与技术学报, 2019, 34(2): 106-111.

ZHU K, YANG H M, MENG K. Short-term wind power forecast using extreme learning machine[J]. Journal of Electric Power Science and Technology, 2019, 34(2): 106-111.

[10] 杨锡运, 关文渊, 刘玉奇, 等. 基于粒子群优化的核极限学习机模型的风电功率区间预测方法[J]. 中国电机工程学报, 2015, 35(增刊 1): 146-153.

YANG X Y, GUAN W Y, LIU Y Q, et al. Prediction intervals forecasts of wind power based on PSO-KELM[J]. Proceedings of the CSEE, 2015, 35(S1): 146-153.

[11] MIRJALILI S, MIRJALILI S M, LEWIS A. Grey wolf optimizer[J]. Advances in Engineering Software, 2014, 69: 46-61.

[12] FARIS H, ALJARAH I, AL-BETAR M A, et al. Grey wolf optimizer: a review of recent variants and applications[J]. Neural Computing and Applications, 2018, 30(2): 413-435.

[13] HUANG N E, SHEN Z, LONG S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[EB/OL]. (1998-03-08)[2023-02-16]. <https://doi.org/10.1098/rspa.1998.0193>.

[14] DRAGOMIRETSKIY K, ZOSSO D. Variational mode decomposition[J]. IEEE Transactions on Signal Process-

ing, 2014, 62(3): 531–544.

[15] TORRES M E, COLOMINAS M A, SCHLOTTHAUER G, et al. A complete ensemble empirical mode decomposition with adaptive noise [C]//2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway:IEEE, 2011: 4144–4147.

[16] 韩宏志,唐振浩. 基于 CEEMDAN 与回声状态网络的风速预测方法[J]. 电力系统保护与控制, 2020, 48(12): 90–96.

HAN H Z, TANG Z H. Wind speed prediction method based on CEEMDAN and echo state network[J]. Power System Protection and Control, 2020, 48(12): 90–96.

[17] REN Y, SUGANTHAN P N, SRIKANTH N. A comparative study of empirical mode decomposition-based short-term wind speed forecasting methods[J]. IEEE Transactions on Sustainable Energy, 2015, 6(1): 236–244.

[18] 舒畅,金潇,李自品,等. 基于 CEEMDAN 的配电变压器放电故障噪声诊断方法[J]. 高电压技术, 2018, 44(8): 2603–2611.

SHU C, JIN X, LI Z P, et al. Noise diagnosis method of distribution transformer discharge fault based on CEEMDAN[J]. High Voltage Engineering, 2018, 44(8): 2603–2611.

[19] WANG H F, XIU C B, LI Y Q, et al. Hysteretic neural network and its application in the prediction of the wind speed series[C]//The 26th Chinese Control and Decision Conference (2014 CCDC). Piscataway: IEEE, 2014: 762–765.

[20] SUN S Z, FU J Q, LI A. A compound wind power forecasting strategy based on clustering, two-stage decomposition, parameter optimization, and optimal combination of multiple machine learning approaches [J]. Energies, 2019, 12(18): 3586.

[21] WANG J J, ZHANG W Y, LI Y N, et al. Forecasting wind speed using empirical mode decomposition and Elman neural network[J]. Applied Soft Computing, 2014, 23: 452–459.

[22] 董雪,赵宏伟,赵生校,等. 基于二次分解和多目标优化的超短期海上风电功率预测[J]. 高电压技术, 2022, 48(8): 3260–3270.

DONG X, ZHAO H W, ZHAO S X, et al. Ultra-short-term offshore wind power forecasting based on secondary decomposition and multi-objective optimization[J]. High Voltage Engineering, 2022, 48(8): 3260–3270.

[23] HE F F, ZHOU J Z, FENG Z K, et al. A hybrid short-term load forecasting model based on variational mode decomposition and long short-term memory networks considering relevant factors with Bayesian optimization algorithm [J]. Applied Energy, 2019, 237: 103–116.

[24] BAI S J, KOLTER J Z, KOLTUN V. Trellis networks for sequence modeling[EB/OL]. (2018–10–15) [2023–02–16]. <https://arxiv.org/abs/1810.06682>.

Multistep Prediction of Wind Speed Based on Grey Wolf Algorithm and Extreme Learning Machine

ZHANG Wenyu<sup>1</sup>, MA Keke<sup>2</sup>, GUO Zhenhai<sup>3</sup>, ZHAO Jing<sup>3</sup>, QIU Wenzhi<sup>2</sup>

(1. School of Earth Sciences and Technology, Zhengzhou University, Zhengzhou 450001, China; 2. School of Computer and Artificial Intelligence, Zhengzhou University, Zhengzhou 450001, China; 3. State Key Laboratory of Numerical Modeling of Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China)

**Abstract:** In order to improve the multi-step prediction of wind speed, a hybrid prediction model based on data signal decomposition and grey wolf optimization algorithm was proposed to optimize extreme learning machine. Firstly, the original wind speed time series was decomposed into several intrinsic mode functions and a residual sequence using the complete ensemble empirical mode decomposition with adaptive noise, and the partial autocorrelation function model input. Then, the model was built and the prediction was made on the decomposition subsequence. An extreme learning machine neural network with multi-input-multi-output strategy was constructed, and grey wolf algorithm was used to solve the weight and bias of the optimal hidden layer. Finally, the subsequence was reconstructed and the final prediction result was obtained. Simulation experiments were conducted using multiple sets of measured data with a time resolution of 15 minutes. The root mean square errors of the proposed model in the three wind farms were 0.859, 0.925, and 0.927, respectively, which were lower than other comparative models, verifying the effectiveness of the model in predicting wind speed in the next four hours, i.e. 16 steps prediction.

**Keywords:** wind speed prediction; multi-step prediction; signal decomposition; selection of features; grey wolf optimization; extreme learning machine