

基于改进 Cao 算法的 SSA 与误差修正的 超短期风电功率预测*

张开伟¹ 文 中¹ 杨生鹏¹ 胡梓涵¹ 丁 剑²

(1. 三峡大学电气与新能源学院 宜昌 443002; 2. 上海勘测设计研究院有限公司 上海 200434)

摘 要: 针对风电历史信息运用不充分和未充分挖掘机器学习模型潜力的问题,提出一种特征奇异谱分析和模型误差修正的超短期功率预测。首先,利用随机森林分析不同特征对输出功率的影响程度,并利用累积贡献率进行特征提取。其次,通过改进的 Cao 算法确定奇异谱分析最佳嵌入维数,对提取的特征实现降噪处理,从而构建风电功率预测模型。最后,利用预测值与真实值的误差构建误差预测模型,通过预测的误差来修正功率预测的结果。以国内某小型风电场算例结果表明,所提方法较卷积神经网络—长短期记忆(CNN-LSTM)预测模型均方根误差(RSME)和均方误差(MSE)分别降低 45%和 53%,验证了所提模型的有效性。

关键词: 奇异谱分析;超短期功率预测;随机森林;累积贡献率;Cao 算法;误差修正

中图分类号: TN2 **文献标识码:** A **国家标准学科分类代码:** 520.604

Ultra-short-term wind power forecasting through improved Cao algorithm for SSA and error correction

Zhang Kaiwei¹ Wen Zhong¹ Yang Shengpeng¹ Hu Zihan¹ Ding Jian²

(1. School of Electrical and New Energy, Three Gorges University, Yichang 443002, China;

2. Shanghai Survey and Design Institute Co., Ltd., Shanghai 200434, China)

Abstract: In order to address the issues of underutilization of historical wind power information and insufficient exploitation of the potential of machine learning models, a method for ultra-short-term wind power forecasting has been proposed. This method is based on feature singularity spectrum analysis and model error compensation. Firstly, random forest is used to analyze the influence of different features on the output power, and the cumulative contribution rate is used to extract the features. Secondly, by improving the Cao algorithm, the optimal embedding dimension for singular spectrum analysis is determined. The extracted features are denoised and a wind power prediction model is constructed based on the denoised data. Finally, the error prediction model is constructed by using the error between the predicted value and the real value, and the result of power prediction is corrected by the predicted error. The results from a small wind farm in China confirm that the proposed method reduced RSME and MSE by 45% and 53% compared to CNN-LSTM, thus verifying its effectiveness.

Keywords: singular spectrum analysis; ultra-short-term wind power prediction; random forest; contribution rate of accumulation; cao algorithm; error correction

0 引 言

全球新能源发展正迎来巨大的变革,对环境问题的日益关注以及全球能源需求的不断增长。在这一大背景下,全球

范围内达成一个共识,应当加速以风电为代表的新能源发展,以降低对传统能源的依赖,来减缓气候变化的影响。截至 2022 年底,全球风电累计装机容量已达 906 GW^[1-3],同比增长 9%。表明风电作为一种清洁、可再生能源正逐渐

收稿日期:2024-03-09

* 基金项目:国家自然科学基金(52107108)项目资助

成为全球能源体系中的主要组成部分。然而,随着风电装机容量快速增长,也带来了一些挑战和问题。其中,最突出的问题是风电出力的随机性和波动性。在大规模并入电网的情况下会导致发电平衡难度加大^[4-7]。为了应对这一问题,迫切需要提高风电功率预测的精度,以更好地适应电网的运行需求。

提高风电预测精度是确保电力系统稳定运行的关键,而这一目标主要通过对输入数据的处理和模型预测误差补偿来实现^[8]。数据处理是一个关键的步骤,涉及从原始数据中提取和构建对模型预测有助益的特征。文献[9]采用最小支持向量机(least squares support vector machines, LSSVM)对风电功率进行预测,提高了风电功率预测的准确性,但在实际应用中参数设置不当,可能导致模型过拟合或欠拟合。文献[10]采用集成经验模态分解(ensemble empirical mode decomposition, EEMD)风电功率历史数据,并进行相空间重构,最后利用自适应布谷鸟算法优化最小二乘支持向量机进行短期风电功率预测方法。但增加了模型的复杂度,导致模型训练时间过长,计算机资源消耗大。文献[11]提出了一种变分模态分解(variational mode decomposition, VMD)和经过改进的深度学习混合核极限学习机(mixed kernel extreme learning machine, MKELM)来构建风电功率预测模型,通过优化VMD参数和MKELM的混合核参数,实现了对风电功率的短期和多步预测。通过算例验证了VMD在平稳和非平稳序列同样适用。为进一步提高风电功率预测的准确性。一些学者在在数据分析的基础上,进一步对预测模型误差进行分析。文献[5]提出一种基于高维动态藤 Copula理论的多风电场联合出力不确定性建模方法。根据风电功率预测误差概率密度特性,对误差进行精细化建模,从而提高风电功率预测的准确性。但该模型对其应用场景的研究仍有欠缺,还需要考虑更多实际因素,如风电场地理位置的多样性、风速变化的不规则性等,以进一步提高其在实际应用中的准确性和可靠性。文献[12]融合双重注意力机制改进的卷积神经网络-双向长短期记忆(CNN-BiLSTM)和LightGBM误差修正的风电功率预测方法。通过特征注意力模块和时间注意力模块自适应提取风电功率的关键特征并进行初步预测,再利用LightGBM对预测结果进行修正,相较于传统模型预测精度有显著提高。文献[13]将功率误差以风速大小分类,并训练出每类误差预测模型。由于误差分布与风速之间缺乏关联,导致较低的误差预测精度。因此需要开发一种误差修正策略,不依赖于误差分布的先验假设。这意味着需要直接从风电预测数据中挖掘误差的潜在规律,以提高误差修正的鲁棒性和适应性。解决这一挑战将有助于使风电功率预测更加准确和可靠。

针对上述数据处理与模型预测方法存在的问题。本文提出基于改进Cao算法的奇异谱分析与模型误差修正的超短期风电功率预测方法。采用随机森林(random for-

est, RF)对气象特征重要度进行分析,并基于累积贡献率进一步优化特征选择。提出一种改进Cao算法稳定性判据,减小了确定奇异谱分析(singular spectrum analysis, SSA)嵌入维数的主观性。基于SSA对非线性时间序列进行分解重构。构建风电功率预测模型,通过深入分析和修正模型的误差,挖掘模型的潜力,以提高风电功率预测的准确性和可靠性。

1 输入特征提取

风力发电受许多因素共同作用,其中包括风速、风向、温度、湿度和气压等。由于这些复杂因素之间存在非线性关系,所以需要采用灵活的建模方法,以更好地捕捉其复杂性。过多的特征输入可能对预测模型的准确性产生负面影响。因此,进行特征筛选是不可或缺的步骤。不仅有助于提高模型的准确性,还有助于增强模型的可解释性和提升计算效率。

RF不仅在预测问题上有着广泛的应用,在模型特征选择上也十分常用。RF基于Bootstrap方法生成每颗决策树的数据集,在每轮随机抽取中,未被抽中参与决策树训练的样本数据为该树的袋外数据(out-of-bag, OOB)。对于第*i*颗决策树,其袋外误分率是所有袋外数据中被错误分类的样本比例。假设总共有*m*个袋外数据样本,其中被第*i*颗决策树错误分类的样本数量为*e*,则袋外误分率如式(1)所示。然后,保持其他特征样本不变,随机改变袋外第*j*个特征值,重新计算袋外误分率。最后,对所有决策树求平均值,得到特征平均精度下降值如式(2)所示。

$$e_{rri} = \frac{e}{m} \quad (1)$$

$$M(v_j) = \frac{1}{q} \sum_{i=1}^N (e_{rri}' - e_{rri}) \quad (2)$$

式中:*m*表示错误分类的数量;*e*表示总的样本数量;*q*表示气象特征的总数;*e_{rri}*为袋外误分率;*M(v_j)*为特征平均精度下降值,*v_j*为第*j*个特征值(*j*=1,2,⋯,*n*)。

以随机森林作为分析工具,计算每个气象特征的平均精度下降值并进行排序,以获取各特征的排名。接着,计算得到每个特征的累积贡献率,这一率值反映了特征在整个模型中的相对重要性,公式如下:

$$\alpha_p = \frac{\sum_{j=1}^p M(v_j)}{\sum_{j=1}^q M(v_j)}, 1 < p < q \quad (3)$$

式中:*α_p*表示第*p*个特征的累计贡献率。

2 SSA 风电序列分解

2.1 SSA 原理

在分析复杂的分钟级风电历史数据时,由于传统分析方法存在一些问题,如模态混叠和白噪声干扰。为了克服这些问题,所以本文采用奇异谱信号分析方法,该方法结合了概率论和数理统计,特别适用于处理非线性时间序

列^[14]。首先,将长度为 N 的一维信号矩阵 $\mathbf{X} = [x_1, x_2, \dots, x_N]$ 中提取出的时间序列进行表示。接着,采用奇异值分解(SVD)对嵌入的信号矩阵进行分解,从而获得趋势、震荡和噪声等组成成分。分组和对角平均化步骤进一步增强了对这些成分的提取和理解能力。主要流程如图1所示。

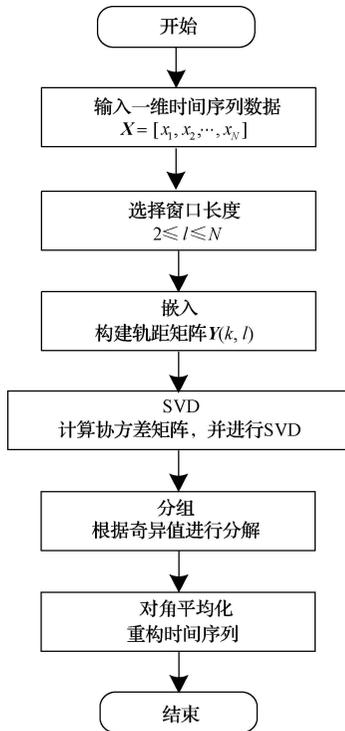


图1 奇异谱分解流程

Fig.1 Flowchart of singular spectrum analysis

2.2 SSA 嵌入维数计算

SSA 的嵌入维数 l 对信号分解至关重要,如果 l 选择较小,将会导致信号分解不彻底部分信号丢失。如果 l 选择较大,会导致不同信号混叠。所以 l 的选择对后续风电功率预测有着较大的影响。为确定最佳嵌入维数,本文采用Cao算法进行相空间重构。Cao算法对数据长度依赖性不强,可以清楚的区分出主要成分与噪声成分,以及计算效率高特点。

首先计算各嵌入维度下每个相空间矢量点最邻近点的距离变化值:

$$d(i, l) = \frac{\|\mathbf{X}_i(l+1) - \mathbf{X}_{n(i,l)}(l+1)\|}{\|\mathbf{X}_i(l) - \mathbf{X}_{n(i,l)}(l)\|} \quad (4)$$

式中: $\|\cdot\|$ 为向量 ∞ -范数; $\mathbf{X}_i(l+1)$ 为第 i 个重构相空间向量, $(l+1)$ 为嵌入维数; $\mathbf{X}_{n(i,l)}$ 为上述范数定义下离 $\mathbf{X}_i(l+1)$ 最近的向量, $n(i, l)$ 为 $1 \leq n(i, l) \leq (N-l)$ 的整数。

其次,计算相同维度下距离变化的平均值:

$$E(l) = \frac{1}{N-l} \sum_{i=1}^{N-l} d(i, l) \quad (5)$$

式中: $E(l)$ 为 $d(i, l)$ 的平均值。

最后,观察 $E(l)$ 的变化:

$$E_1(l) = \frac{E(l+1)}{E(l)} \quad (6)$$

式中: $l > l_0$ 时 $E_1(l)$ 波动较小或停止变化,则最小嵌入维数为 $l_0 + 1$ 。

2.3 改进的Cao算法

Cao算法在确定SSA最佳嵌入维数时是 $E_1(l)$ 根据波动变化来确定的,然而对于时序数列来说,确定稳定状态极具挑战性,因此,仅通过主观判断来确定最佳嵌入维数存在一定的难度。针对这个问题,提出一种稳定性判据来确定最佳嵌入维数。

1)计算:

$$\Delta_i = |E_1(i) - E_1(i+1)|, \quad 1 < i < N-1 \quad (7)$$

2)根据 E_1 的波动范围选取阈值 ϵ , 遍历序列 E_1 找到第1个 $\Delta_i < \epsilon$ 对应的下标 u , 计算:

$$\Delta_j = \max(\Delta_i), u \leq i \leq N-1 \quad (8)$$

3)重设:

$$\epsilon = \bar{\Delta}_j, j \leq i \leq N-1 \quad (9)$$

式中: $\bar{\Delta}$ 表示平均值。

4)若 $j \leq i \leq N-2, \Delta_{i+2} < \Delta_{i+1} < \Delta_i$ 且 $\Delta_{i+1} < \epsilon$ 时,嵌入维数 $l = i + 1$ 。

3 风电功率预测模型

3.1 CNN原理

CNN是一种具有卷积计算的深度结构前馈神经网络,有强大的特征提取能力。CNN卷积层是其核心组成部分,其结构如图2所示。通过使用不同的卷积核对输入数据进行计算,从而得到包含不同特征的映射矩阵,公式如下:

$$X^{m+1} = \sigma(\sum W^m \cdot X^m + B^m) \quad (10)$$

式中: X^{m+1}, X^m 分别为该层和下一层输入; W^m 为该层权重; B^m 为偏置系数; σ 为激活函数。

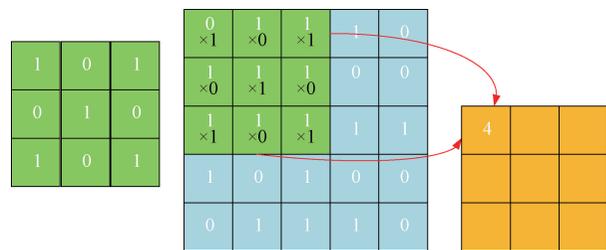


图2 卷积层工作原理

Fig.2 Convolution layer working principle diagram

卷积计算中,在输入数据上设置滤波器的大小和步长进行卷积运算。将结果用于输出特征数据,从而实现数据的降维。

3.2 长短期记忆(LSTM)原理

LSTM神经网络作为一种深度学习算法,可以灵活地

适应时序特征学习。同时,LSTM网络解决了递归神经网络在时序反向传播训练过程中的梯度消失和梯度爆炸问题,能够充分利用历史信息,对时间序列进行预测^[15],结构如图3所示。

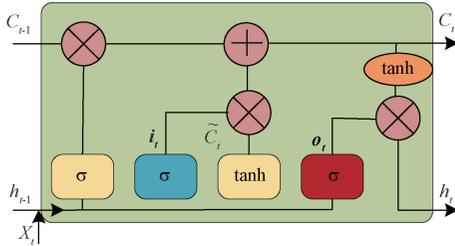


图3 LSTM原理

Fig. 3 LSTM schematic diagram

LSTM引入记忆单元、遗忘门、输入门和输出门,以实现长期和短期记忆的精确控制。遗忘门通过sigmoid激活函数调控记忆单元中哪些信息需要遗忘,输入门通过sigmoid和tanh激活函数确定哪些新信息被添加到记忆单元中。这两个门的协同作用使得LSTM能够灵活地读取和修改记忆。记忆单元的更新通过遗忘门删除一些信息并通过输入门添加新信息。输出门决定从记忆单元中提取哪些信息作为网络的输出。

3.3 RF-SSA-CNN-LSTM模型预测流程

RF-SSA-CNN-LSTM模型如图4所示。对CNN-LSTM进行改进。引入RF-SSA特征提取模块,将其置于CNN模型前端,对输入数据进行重要度计算与降噪处理。不仅有助于减小特征的维度,同时增强了数据之间的关联性。目的是提升模型的特征表达能力,使得CNN-LSTM能够更好地学习到深层次的非线性隐藏关联。这种结构设计能够在模型训练中带来更显著的性能提升,从而提高对复杂数据的学习和表达能力。

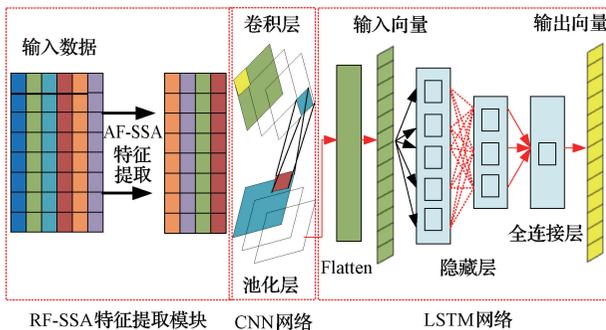


图4 RF-SSA-CNN-LSTM网络结构

Fig. 4 The network structure of RF-SSA-CNN-LSTM

本文风电功率预测流程如图5所示,实施步骤如下。

1)通过归一化处理分析不同输入对风电功率的影响,确保数据在相同尺度上进行比较,从而提高模型的性能和稳定性。

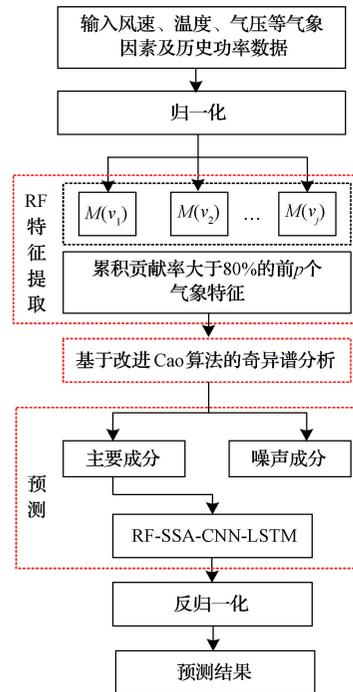


图5 RF-SSA-CNN-LSTM预测流程

Fig. 5 RF-SSA-CNN-LSTM prediction process

$$\hat{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (11)$$

$$x = \hat{x} \cdot (x_{\max} - x_{\min}) + x_{\min} \quad (12)$$

式中: \hat{x} 表示归一化值; x 表示反归一化值; x_{\min} 为样本中最小值; x_{\max} 为样本中最大值。

2)采用RF对j个特征进行重要度计算并排序,以及基于累积贡献率的计算,在能够充分考虑特征相关性的同时,减少冗余信息。

3)将多个特征与历史风电功率组合成多变量输入,并基于改进Cao算法的SSA去噪处理,有助于提取信号中的真实模式,减少噪声数据对模型的影响,为风电功率预测提供更可靠的输入。

4)建立RF-SSA-CNN-LSTM预测模型,以风电功率作为标签。

5)采用均方根误差(RMSE)、平均绝对误差(MAE)作为各预测模型的评价指标^[16]。

$$E_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (13)$$

$$R_{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (14)$$

式中:n为样本数; \hat{y}_i 为第i个预测值; y_i 为第i个真实值, $i = 1, 2, \dots, n$ 。

4 误差补偿原理与流程

在进行风电功率预测时,考虑气象条件、地理特征、风

机性能等多个因素,并通过预测模型得到初步的预测结果。然而,由于各种不确定性因素的存在,预测结果往往存在一定的误差。为了应对这些误差,引入一种补偿机制,即利用原有的影响因素和预测结果作为输入,通过特定的算法生成一系列补偿数据。目前,为了实现风电预测误差的有效补偿,主要采用了深度学习算法^[17-18]。

由本文构建的 RF-SSA-CNN-LSTM 模型在进行风电功率预测时表现出色,具有高度准确的预测能力。因此采用 RF-SSA-CNN-LSTM 作为误差补偿模型,流程如图 6 所示。

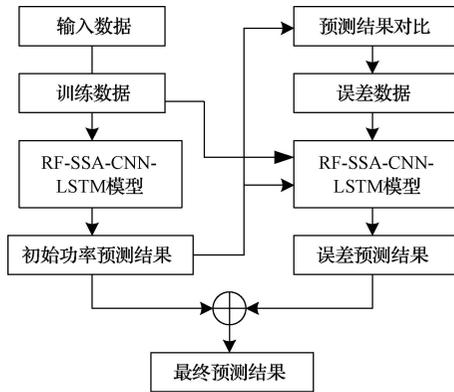


图 6 误差补偿流程

Fig. 6 Error compensation process

由图 6 可以看出,基于 RF-SSA-CNN-LSTM 模型,训练主模型 *Regressor* 为:

$$Regressor \Rightarrow \langle X_{new}^{train}, y^{train} \rangle \quad (15)$$

式中: X_{new}^{train} 表示训练特征; y^{train} 表示训练样本标签。

考虑到在实际工程中无法得到风电功率的预测误差。所以计算训练集的预测值 \hat{y}_{train} 与训练集标签的误差 e^{train} :

$$e^{train} = y^{train} - \hat{y}_{train} \quad (16)$$

构造误差预测模型 *Corrector* 预测主模型,其中误差预测模型 *Corrector* 的训练过程为:

$$Corrector \Rightarrow \langle concat(X_{new}^{train}, \hat{y}_{train}), e^{train} \rangle \quad (17)$$

式中: $concat(\cdot)$ 表示将 X_{new}^{train} 与 \hat{y}_{train} 横向拼接为训练样本集, e^{train} 为训练样本集标签。

基于 *Regressor* 训练模型、*Corrector* 误差训练模型进行测试集的预测。首先采用 *Regressor* 模型进行测试集的预测 \hat{y}_{test} 为:

$$\hat{y}_{test} = Regressor(X_{test}) \quad (18)$$

其次将 X_{new}^{test} 与 \hat{y}_{test} 进行拼接,作为输入特征,对误差 \hat{e} 进行预测:

$$\hat{e} = Corrector(concat(X_{new}^{test}, \hat{y}_{test})) \quad (19)$$

最后,对 \hat{y}_{test} 进行误差修正:

$$\hat{y}^{corr} = \hat{y}_{test} + \hat{e} \quad (20)$$

5 算例分析

5.1 数据说明

本文以国内某小型风电场 2017 年监测数据进行仿真验证。时间分辨率为 15 min,共 35 041 条数据,每条数据包含日期、风速、风向、轮毂高度、风向、温度、气压、相对湿度、功率等特征。取前 1 月 1 日~12 月 30 作为训练集,12 月 31 日作为测试集。仿真环境以 Python 为编程语言 Py-CharmCommunityEdition2022.2.3 软件中实现。模型采用 CNN-LSTM 架构,参数设置如表 1 所示。

表 1 CNN-LSTM 模型参数设置

Table 1 CNN-LSTM model parameter settings

模型参数	参数设置
卷积层	2 层
池化层	2 层
卷积核大小	2×2
池化层 stride	1
激活函数	ReLU
LSTM 层	3
LSTM 神经元个数	50, 100, 250
LSTM 激活函数	tanh

5.2 RF 输入特征提取及 SSA 降噪

通过多次实验,使用 75% 的数据对 RF 模型进行训练时模型效果最好,取 10 次运行结果的平均值。计算结果如表 2 所示。

表 2 特征重要度

Table 2 Feature importance

特征	重要度	特征	重要度
风速-10 m	0.101 3	风向-750 m	0.079 3
风速-30 m	0.097 3	风向-70 m	0.074 1
风速-50 m	0.096 0	温度	0.072 6
风速-70	0.088 1	气压	0.071 4
风向-0 m	0.085 4	相对湿度	0.070 0
风向-10 m	0.084 2		
风向-30 m	0.079 8		

将特征重要度从高到低进行排序,并逐个计算其特征累积贡献度。选取贡献度超过 80% 的前 p 个特征作为输出结果。然后对时间序列进行 SSA 降噪处理。以 10 m 风速为例,采用 SSA 进行降噪处理,得到嵌入维数与 E_1 关系如图 7 所示。

由图 7 可见,当 $M = 20$ 时 E_1 开始趋于稳定,但还是有些微弱的波动,对嵌入维数的确定造成一定的影响。从波动情况来看,选取 $\epsilon = 0.000 1$,利用改进的 Cao 算法求得嵌入维数。为了对比改进后算法的效果,把 $M = 20$ 与 $M = 42$ 去噪后的效果加以对比,如图 8、9 所示。

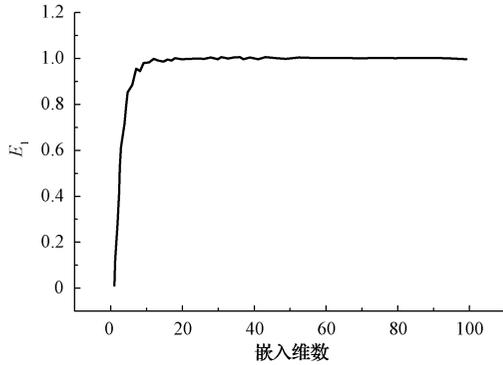


图7 嵌入维数与 E_1 的关系

Fig. 7 The relationship between embedding dimension and E_1

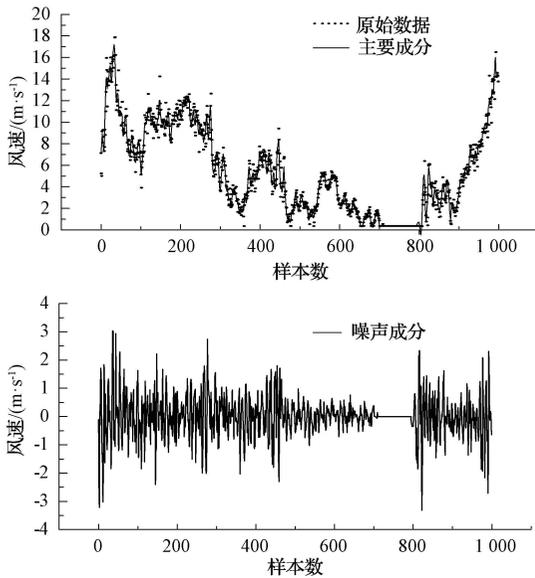


图8 $M=20$ 主要成分与噪声成分

Fig. 8 $M=20$ main component and noise component

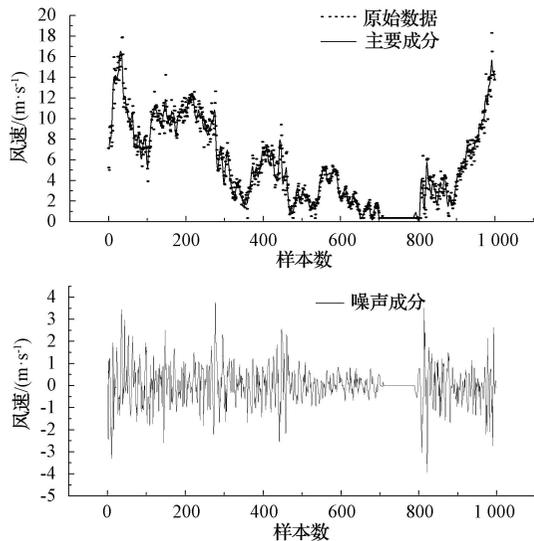


图9 $M=42$ 主要成分与噪声成分

Fig. 9 $M=42$ main component and noise component

由图9可见, $M=42$ 对于风速波峰有较高的平滑度, 去噪效果更好, 主要成分波动较缓且无尖端, 而高频噪声成分波动较大且尖端比较突出。在超短期风电功率预测中考虑随机波动性且高频成分性价比较低^[19]。因此, 本文将主要成分作为后续分析与研究的关键因素。

5.3 算例结果分析

为验证所提出方法的有效性, 对比 CNN-LSTM、RF-SSA-CNN-LSTM 以及 RF-SSA-CNN-LSTM+误差补偿模型, 如图10所示。从图10可以看出, RF-SSA-CNN-LSTM 和 RF-SSA-CNN-LSTM+误差补偿模型相较于 CNN-LSTM 模型在测试集上表现更为出色, 预测精度更高。然而, 在其基础上进行模型误差补偿后相对于实际功率的偏差程度更小, 更接近实际功率曲线。

模型性能评价结果如表3所示, RF-SSA-CNN-LSTM 模型相较于 CNN-LSTM 模型 RMSE 和 MAE 两个评价指标上分别降低了 30%、31%。RF-SSA-CNN-LSTM+误差补偿模型相较于 CNN-LSTM 模型 RMSE 和 MAE 两个评价指标上分别降低了 45%、53%。相较于 RF-SSA-CNN-LSTM 分别下降了 21%、32%。这表明 RF-SSA-CNN-LSTM 模型在功率预测方面相对于传统的 CNN-LSTM 模型有更为准确的预测精度。

表3 各模型预测结果评价

Table 3 Evaluation of prediction results of each model

模型	R_{RMSE}	E_{MAE}
CNN-LSTM	0.647 2	0.436 4
RF-SSA-CNN-LSTM	0.450 7	0.300 6
RF-SSA-CNN-LSTM+误差补偿	0.353 8	0.203 5

5.4 功率剧烈波动情况分析

在风电场气象因素剧烈变化的情况下, 准确预测风电功率是当前研究中的难点。这一难点涉及到对气象条件变化的高度敏感以及模型对这些变化的适应能力。本文选取两个风电功率剧烈波动时段(图10(b)、(c))进行评估, 结果如表4所示。

在功率波动剧烈的情况下, CNN-LSTM 模型难以快速适应其影响因素的变化。这表明其对于瞬时功率波动响应较为迟缓, 可能无法捕捉到快速变化的气象条件对功率的影响。相比之下, RF-SSA-CNN-LSTM 模型更贴近实际功率的变化趋势。经过误差补偿后, RF-SSA-CNN-LSTM 模型的变化幅度进一步提升, 且与实际功率曲线的变化趋势高度吻合。表明误差补偿机制在提升模型性能方面起到了关键作用, 使得模型更准确地反映功率曲线的实际变化情况。

由表4可知, CNN-LSTM 模型的预测准确性相对于总体预测结果略有下降, 这两个时段的 RMSE 分别增加了 6.6% 和 1.2%。RF-SSA-CNN-LSTM 模型与 RF-SSA-CNN-LSTM 模型+误差补偿在这两个时段内的

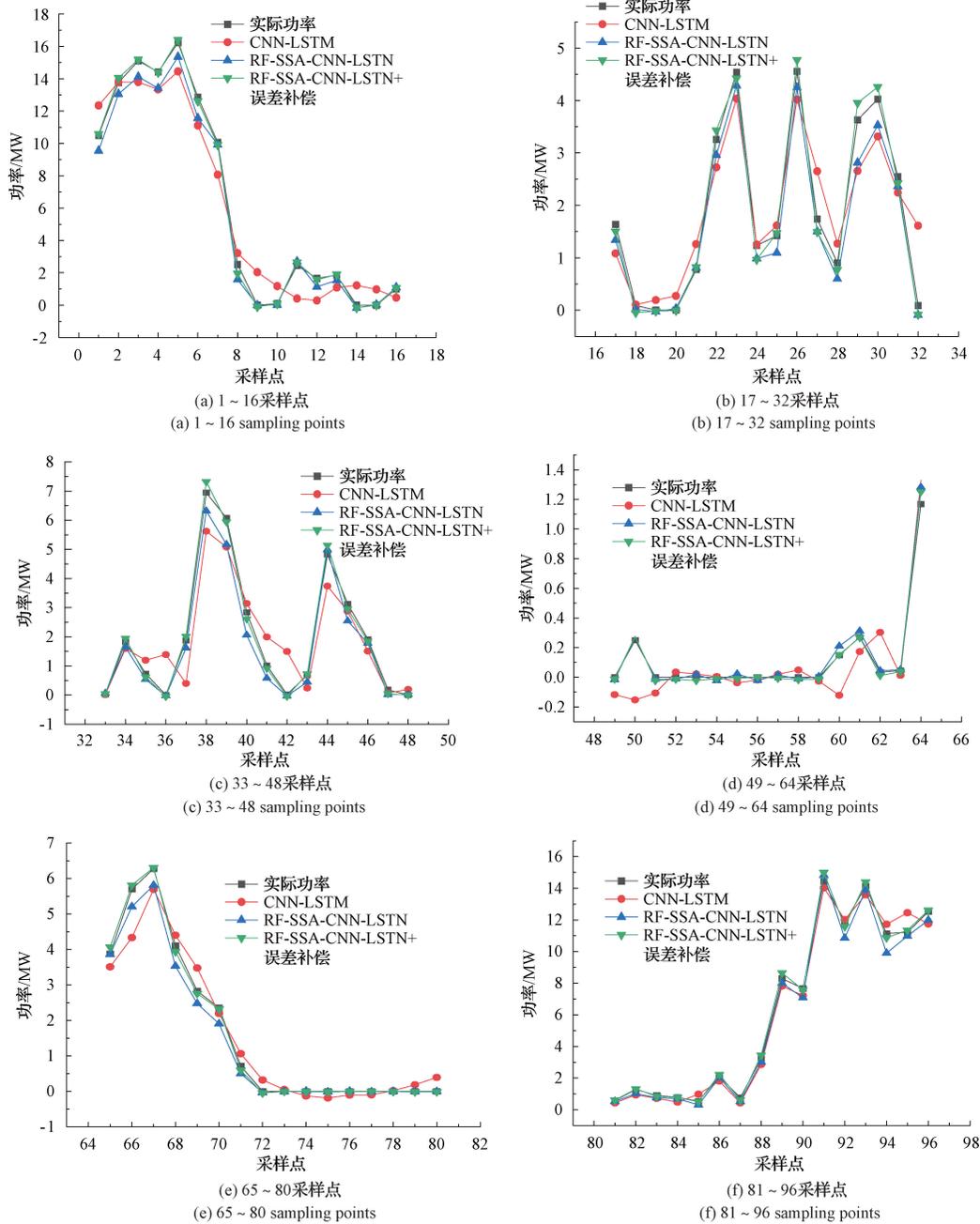


图 10 各模型预测结果

Fig. 10 Prediction results of each model

表 4 功率剧烈波动时评估结果

Table 4 Assessment of results during intense power fluctuations

模型	评价指标	17~32 采样点	33~48 采样点
CNN-LSTM	RMSE	0.690 1	0.655 4
	MAE	0.556 5	0.495 7
RF-SSA-CNN-LSTM	RMSE	0.543 6	0.405 1
	MAE	0.428 0	0.301 1
RF-SSA-CNN-LSTM+误差补偿	RMSE	0.387 2	0.363 7
	MAE	0.236 0	0.210 1

RMSE 较 CNN-LSTM 模型分别下降了 21%、38% 和 43%、44%。与表 4 总体预测结果中的提升(30%、45%)相比,RF-SSA-CNN-LSTM 模型+误差补偿的 RMSE 提升更为显著。

原因在于风电场所处环境较为恶劣。除了对风速变化的关注,还需要监测风向、温度、气压和湿度等弱相关特征的微小变化,如图 11 所示,如果没有关注到这些弱相关特征的微弱变化,模型就很难将这些变化反映到功率预测上。所以本文采用 RF-SSA-CNN-LSTM 模型通过 RF-SSA 对弱相关特征进行降噪处理增加特征的表达能力,从而挖掘更加深层的隐含信息。使得模型在预测时能够捕捉到这些微弱变化,从而提高模型预测的准确性。

5.5 预测功率为负值情况分析

由图 10(d)和(e)可见,当实际风电功率接近 0,但预测功率出现了负值,评估结果如表 5 所示。

由表 5 可知,在这两个时段中 RF-SSA-CNN-LSTM+误差补偿模型预测误差最低,其次是 RF-SSA-CNN-LSTM 模型,CNN-LSTM 模型预测误差最大。原因在于,

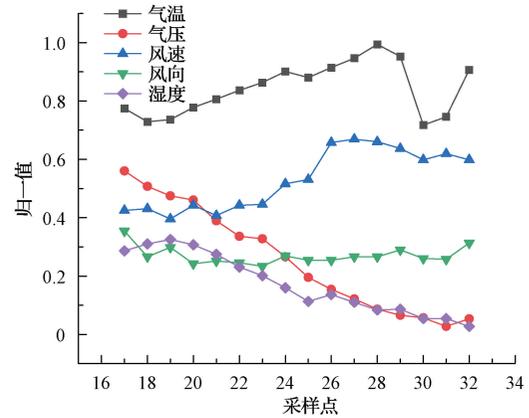


图 11 17~32 采样点风电机组特征变化
Fig. 11 Characteristic changes of wind turbines at 17~32 sampling points

风速在影响风电功率方面起主导作用,当风速波动较小时,风电输出功率也趋于稳定,而其他弱相关特征还在随时间发生波动,如图 12 所示。

表 5 预测为负值时评估结果

Table 5 The evaluation results when the prediction is negative

模型	评价指标	49~64 采样点	65~81 采样点
CNN-LSTM	RMSE	0.153 4	0.457 2
	MAE	0.104 5	0.327 8
RF-SSA-CNN-LSTM	RMSE	0.037 0	0.270 0
	MAE	0.026 0	0.165 7
RF-SSA-CNN-LSTM+误差补偿	RMSE	0.025 4	0.081 0
	MAE	0.017 2	0.052 8

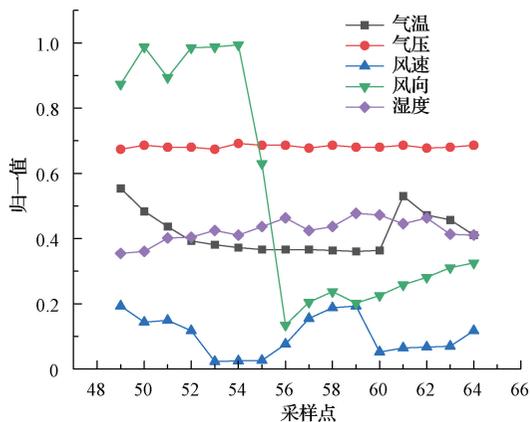


图 12 48~64 采样点风电机组特征变化
Fig. 12 Characteristic changes of wind turbines at 48~64 sampling points

当外界环境风速小于风电机组运行最小风速时,理论上风电机组停止发电,但是预测模型仍然预测出一个非零

的数值。原因在于,当风速波动非常小接近于 0 时,而一些弱相关特征还在随着时间变化而变化,通过 RF-SSA 降噪处理增强了这些特征微弱变化能力,导致对风电功率贡献度变大。从而预测出非零的数值。对于这个问题可以通过设定风速下限来解决。

6 结论

为提高风电功率预测的准确性,本文提出基于改进 Cao 算法的 SSA 特征降噪与模型误差修正的风电功率预测方法,得到如下结论。

1) 本文提出改进 Cao 算法稳定性判据,极大的降低了确定 SSA 嵌入维数的主观性。显著的提高了 SSA 算法的计算效率与准确性。

2) 使用 RF 进行特征重要度分析和 SSA 降噪处理有助于模型挖掘特征之间的潜在信息,提升模型的非线性学习能力。使得 RF-SSA-CNN-LSTM 模型的预测精度比 CNN-LSTM 模型提高了 30%。

3) 在 RF-SSA-CNN-LSTM 模型预测的基础上,在对

模型进行误差的修正,使得补偿后模型预测精度较未补偿时提高了21%。

4)风电功率剧烈波动时段,风电功率波动不仅仅与风速有关,一些弱相关特征波动同样需要关注。RF-SSA-CNN-LSTM模型在适应动态环境方面有着强大的学习能力,可有效适应非线性特征,使其在风电系统中更灵活应对其变化。

5)在风电功率预测值接近0的时段,此时风速波动接近于0,从而一些弱相关特征对风电功率预测占主要地位。RF-SSA-CNN-LSTM加强了对弱相关特征的学习导致预测功率出现非零值。

参考文献

- [1] WU C, ZHANG X P, STERLING M. Wind power generation variations and aggregations [J]. CSEE Journal of Power and Energy Systems, 2021, 8(1): 17-38.
- [2] XIANG L, ZHU H W, ZHANG Y, et al. Impact of wind power penetration on wind-thermal-bundled transmission system [J]. IEEE Transactions on Power Electronics, 2022, 37(12):15616-15625.
- [3] ZHOU Y, SUN Y H, WANG S, et al. Performance improvement of very short-term prediction intervals for regional wind power based on composite conditional nonlinear quantile regression[J]. Journal of Modern Power Systems and Clean Energy, 2021, 10(1): 60-70.
- [4] 阮前途,陈赞,潘智俊.新型电力系统对实现沿海城市双碳目标贡献的探讨[J].电力系统自动化,2022,46(14):11-18.
- RUAN Q T, CHEN Y, PAN ZH J. Discussion on the contribution of a novel power system to achieving dual carbon goals in coastal cities[J]. Power System Automation, 2022, 46(14): 11-18.
- [5] 段偲默,苗世洪,李力行,等.计及预测误差动态相关性的多风电场联合出力不确定性模型[J].电力系统自动化,2019,43(22):31-37.
- DUAN S M, MIAO SH H, LI L X, et al. Joint output uncertainty model for multiple wind farms considering dynamic correlation of prediction errors[J]. Power System Automation, 2019, 43(22): 31-37.
- [6] WANG S, SUN Y H, ZHANG S M, et al. Very short-term probabilistic prediction of PV based on multi-period error distribution [J]. Electric Power Systems Research, 2023, 214: 108817.
- [7] CATALÁN P, WANG Y B, ARZA J, et al. A comprehensive overview of power converter applied in high-power wind turbine: Key challenges and potential solutions[J]. IEEE Transactions on Power Electronics, 2023, 38(5): 6169-6195.
- [8] YANG M, ZHANG L B, CUI Y, et al. Investigating the wind power smoothing effect using set pair analysis [J]. IEEE Transactions on Sustainable Energy, 2020, 11(3): 1161-1172.
- [9] 庞传军,尚学伟,张波,等.基于改进梯度提升算法的短期风电功率概率预测[J].电力系统自动化,2022,46(16):198-206.
- PANG CH J, SHANG X W, ZHANG B, et al. Short-term wind power probability prediction based on improved gradient boosting algorithm[J]. Power System Automation, 2022, 46(16): 198-206.
- [10] 姜贵敏,陈志军,李笑竹,等.基于EEMD-ACS-LSSVM的短期风电功率预测[J].太阳能学报,2020,41(5):77-84.
- JIANG G M, CHEN ZH J, LI X ZH, et al. Short-term wind power forecasting based on EEMD-ACS-LSSVM[J]. Acta Energetica Solaris Sinica, 2020, 41(5): 77-84.
- [11] KRISHNA RAYI V, MISHRA S P, NAIK J, et al. Adaptive VMD based optimized deep learning mixed kernel ELM autoencoder for single and multistep wind power forecasting [J]. Energy, 2022, 244: 122585.
- [12] 龙铖,余成波,何铖,等.基于双重注意力机制CNN-BiLSTM与LightGBM误差修正的超短期风电功率预测[J].电气工程学报,2024,19(2):138-145.
- LONG CH, YU CH B, HE CH, et al. Ultra-short-term wind power forecasting based on dual attention mechanism CNN-BiLSTM and LightGBM error correction [J]. Journal of Electrical Engineering, 2024, 19(2): 138-145.
- [13] 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. Integrated short-term wind power forecasting method based on error correction [J]. High Voltage Engineering, 2022, 48(2): 488-496.
- [14] AFSHAR K, BIGDELI N. Data analysis and short term load forecasting in Iran electricity market using singular spectral analysis(SSA)[J]. Energy, 2011, 36: 2620-2627.
- [15] 陈泽慧,李博,李博.考虑加权理论的风电场集群风速预测方法[J].国外电子测量技术,2021,40(10):34-39.
- CHEN Z H, LI B, LI B. Wind speed forecasting

- method for wind farm clusters considering weighted theory[J]. *Foreign Electronic Measurement Technology*, 2021, 40(10): 34-39.
- [16] 王慧莹,吴亮红,梅盼盼,等. 果蝇优化广义神经网络的风电功率短期预测[J]. *电子测量与仪器学报*, 2019,33(6):177-183.
WANG H Y, WU L H, MEI P P, et al. Short-term wind power forecasting based on fruit fly optimization generalized neural network[J]. *Journal of Electronic Measurement and Instrumentation*, 2019, 33(6): 177-183.
- [17] TAO C, LU J J, LANG J X, et al. Short-term forecasting of photovoltaic power generation based on feature selection and bias compensation-LSTM network[J]. *Energies*, 2021, 14(11): 3086.
- [18] 边春元,邢海洋,李晓霞,等. 基于速度变化率的无位置传感器无刷直流电机风力发电系统换相误差补偿策略[J]. *电工技术学报*, 2021,36(11):2374-2382.
BIAN CH Y, XING H Y, LI X X, et al. Commutation error compensation strategy for sensorless brushless DC motor wind power generation system based on speed change rate[J]. *Transactions of China Electrotechnical Society*, 2021, 36(11): 2374-2382.
- [19] 梁志峰,王峥,冯双磊,等. 基于波动规律挖掘的风电功率超短期预测方法[J]. *电网技术*, 2020, 44(11), 4096-4104.
LIANG ZH F, WANG ZH, FENG SH L, et al. Ultra-short-term wind power prediction method based on fluctuation pattern mining [J]. *Power System Technology*, 2020, 44(11): 4096-4104.

作者简介

张开伟,硕士研究生,主要研究方向为风电功率曲线与整机状态检测。

E-mail:202208540021048@ctgu.edu.cn

文中,副教授,主要研究方向为风电功率预测、电力系统运行与控制。

E-mail:2791415163@qq.com

杨生鹏,硕士研究生,主要研究方向为新能源发电出力预测。

E-mail:860295415@qq.com

胡梓涵,硕士研究生,主要研究方向为配电网优化调度。

E-mail:806377503@qq.com

丁剑,工程师,主要研究方向为输电线路设计。

E-mail:1748633501@qq.com