

基于注意力机制的VMD-BiLSTM 风力发电预测

邓文武¹, 丁咏梅^{2*}

武汉科技大学理学院, 湖北 武汉

收稿日期: 2022年12月19日; 录用日期: 2023年1月11日; 发布日期: 2023年1月30日

摘要

针对风力发电数据噪音多、波动大等特点, 本文提出了基于注意力机制的VMD-BiLSTM超短期风力发电预测模型。首先, 对数据进行变分模态分解, 降低序列的非平稳性和强非线性; 其次, 将各子模态固定在同一时间滑动窗口上, 分别使用BiLSTM进行进一步的双向序列特征提取, 然后引入注意力机制, 加强重要特征信息的影响, 提高模型的最终预测效果。最后通过NREL网站的美国西南地区的发电数据集来进行验证, 最终结果显示, 无论是从预测的准确性还是评判指标方面, 本文所提出的模型在风力发电预测上都优于其他5类模型。

关键词

风力发电, 注意力机制, 变分模态分解, 双向长短期记忆网络, 超短期预测

VMD-BiLSTM Wind Power Prediction Based on Attention Mechanism

Wenwu Deng¹, Yongmei Ding^{2*}

The School of Science, Wuhan University of Science and Technology, Wuhan Hubei

Received: Dec. 19th, 2022; accepted: Jan. 11th, 2023; published: Jan. 30th, 2023

Abstract

Aiming at the characteristics of large noise and large fluctuations in wind power generation data, an attention mechanism-based VMD-BiLSTM ultra-short-term wind power generation prediction model is proposed. Firstly, the data is decomposed using the variational mode (VMD) to reduce the

*通讯作者。

sequence non-stationarity and strong nonlinearity. Secondly, each sub-mode is fixed on the sliding window at the same time, and BiLSTM is used to further extract bidirectional sequence features, and then the attention mechanism is introduced to strengthen the influence of important feature information and improve the final performance of the model. Finally, the power generation data set in the Southwest of the United States on the NREL website was verified. The final results show that the model proposed in this paper is superior to the other 5 models in terms of wind power prediction accuracy and evaluation indicators.

Keywords

Wind Power, Attention Mechanisms, Variational Modal Decomposition, Bidirectional Long Short-Term Memory Network, Ultra-Short-Term Forecasting

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在当今错综复杂的全球背景之下, 极端天气的频发, 新冠肺炎不断肆虐, 部分国家冲突不断升级, 使得生态能源、资源也面临着巨大压力和不确定性。风能作为一种高效、清洁的可再生能源, 在世界各地得到了广泛的应用, 并且在未来的全球发展中起着至关重要的作用[1]。精准的风力发电预测, 为电网系统的顺利运行和世界社会经济的持续健康发展提供了强大的技术保证。

目前, 电力预测主要包括传统的差分整合移动平均自回归模型(ARIMA), 机器学习中的XGBoost、逻辑回归、支持向量机等模型, 以及深度学习中的循环神经网络(RNN)、长短期记忆神经网络(LSTM)、门控循环单元(GRU)和注意力机制等。对于ARIMA模型, 模型要求数据是平稳的, 且这个模型在本质上只能捕获序列中的线性关系[2]。针对电力数据波动大的特点, 许多电力负荷的预测研究一般会采用各种分解的方法来减少序列的噪音从而降低模型的复杂度。文献[3]采用了小波分解来分解数据, 虽然文中XGBoost和逻辑回归的混合模型的预测准确率较高, 但是模型在运作过程中要储存大量的特征值和对应的梯度的索引, 会占据大量内存空间, 影响模型整体性能。文献[4]将支持向量机和改进的蜻蜓算法相结合, 预测结果较好, 但是当该处理大样本时, 模型预测的准确率在一定程度上会降低。文献[5]使用变分模态分解与长短期记忆神经网络结合, 最终相比于单一LSTM模型和与其他分解结合的模型, 该方法预测效果较优。文献[6]采用了互补集合经验模态分解(CEEMD), 随后融合GRU神经网络模型, 但是该融合模型在分解时参数会使得产生的子模态数量存在差异, 最终可能导致集合平均产生误差。文献[7]和文献[8]均使用了卷积神经网络进行特征提取, 但是卷积神经网络对于数据量的需求很大, 而且在进入池化层的过程中会导致一些信息的丢失。

针对上述问题, 本文提出了基于注意力机制的VMD-BiLSTM的混合模型。针对原始风力发电数据的固有特性, 本文使用变分模态分解, 既降低了数据的复杂度也保证了其完整性, 同时也解决了模态混叠的问题, 随后将分解后的各个子模态输入BiLSTM模型中, 为了能更好的学习序列的前后规律, 再将各个输出作为注意力机制的输入, 进一步学习数据特征, 并赋予BiLSTM模型对不同序列重要性的隐性权重, 最终进行训练和预测, 并且其他单一模型和混合模型比较, 证明了该模型具有更好的预测效果, 更具有普适性。

2. 基于注意力机制的 VMD-BiLSTM 模型原理

2.1. 变分模态分解

变分模态分解(VMD)是一种能够将新型非平稳信号、序列进行自适应分解的一种模态分解方法[9]。具体理论如下:

假设将原始信号 f 分解为 k 个分量, 保证分解序列为中心频率带宽有限的模态分量, 且各模态估计带宽之和最小, 约束条件为所有模式的总和等于原始信号[10]。约束如下:

$$\min_{\{u_k, \omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{j\omega_k t} \right\|_2^2 \right\} \quad (1)$$

$$\text{s.t. } \sum_k u_k = f \quad (2)$$

其中, f 为原始信号, $u_k = \{u_1, u_2, \dots, u_k\}$ 为各模态的 $\omega_k = \{\omega_1, \omega_2, \dots, \omega_k\}$ 为各模态中心频率, $\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t)$ 为希尔伯特变换下的解析信号, $\delta(t)$ 为脉冲信号函数, $e^{j\omega_k t}$ 是对解析信号进行指数修正, $\left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{j\omega_k t} \right\|_2^2$ 是对信号进行高斯平滑调整(即 L2 范数的平方根), 从而得到每个模态函数的带宽描述。 ∂_t 是求该二范数的梯度, 这里使用二范数是防止模型过拟合[11]。

针对上述的约束最优化问题, 将约束变分问题转变为非约束变分问题, 使用了增广 Lagrangian 函数[9]:

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle \quad (3)$$

其中, α 为惩罚参数, λ 为拉格朗日乘子。

相应的模态分量 u_k 、 ω_k 中心频率的更新公式如下:

$$u_k^{n+1}(\omega) = \frac{f(\omega) - \sum_{i=1, i \neq k}^k u_i(\omega) + \frac{\lambda(x)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \quad (4)$$

$$\omega_k^{n+1} = \frac{\int_{-\infty}^{\infty} \omega |u_k^{n+1}(\omega)|^2 d\omega}{\int_{-\infty}^{\infty} |u_k^{n+1}(\omega)|^2 d\omega} \quad (5)$$

其中, ω 为频率, $f(x)$ 、 $\lambda(x)$ 、 u_k^{n+1} 分别为 $f(t)$ 、 $\lambda(t)$ 、 $u_k^n(t)$ 的傅里叶变换。

VMD 整体步骤如下:

- 1) 初始化参数, 模态数 k , 带宽限制、噪声容忍度、频域表达式 $\{u_k\}$, 中心频率 $\{\omega_k\}$, 拉格朗日乘子 λ 等;
- 2) 根据公式更新模态分量 u_k 和中心频率 ω_k ;
- 3) 利用公式 $\lambda^{n+1}(\omega) = \lambda^n(\omega) + \tau \left(f(\omega) - \sum_k^{n+1} u_k(\omega) \right)$ 更新 λ^{n+1} ;
- 4) 设置停止迭代条件, 退出循环, 然后返回步骤(2), 最终得到分解后的模态, 完成模态分解。

2.2. 双向长短期记忆网络 (BiLSTM)

长短期记忆网络(LSTM)模型将隐藏层中的神经元替换成了具有长短期记忆的神经元, 还增加了三个门控单元, 包括遗忘门、输出门和输入门, 从而有效的解决了之前模型的长期依赖问题[12]。其流程如图 1。

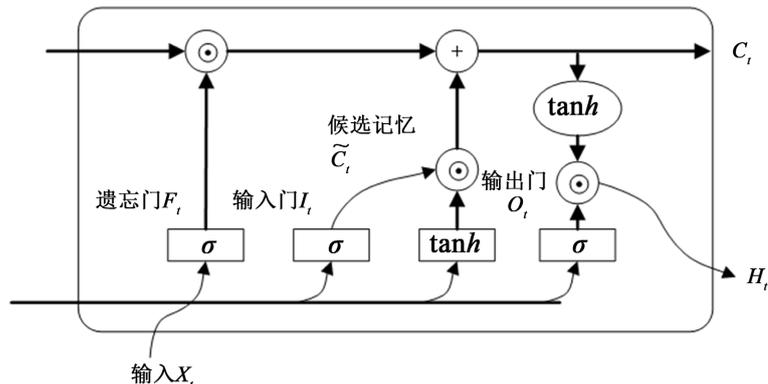


Figure 1. Long short-term memory network flowchart
图 1. 长短期记忆网络流程图

具体公式如下:

$$\text{输入门: } I_t = \text{sigmoid}(W_i [H_{t-1}, x_t] + b_i) \tag{6}$$

$$\text{遗忘门: } F_t = \text{sigmoid}(W_f [H_{t-1}, x_t] + b_f) \tag{7}$$

$$\text{候选记忆: } \tilde{C}_t = \text{sigmoid}(W_c [H_{t-1}, x_t] + b_c) \tag{8}$$

$$\text{更新状态单元: } C_t = F_t * C_{t-1} + I_t * \tilde{C}_t \tag{9}$$

$$\text{输出门: } O_t = \text{sigmoid}(W_o [H_{t-1}, x_t] + b_o) \tag{10}$$

$$\text{输出: } H_t = O_t * \tanh(C_t) \tag{11}$$

上式中的 W_i 、 W_f 、 W_c 、 W_o 分别为各个门的权重, b_i 、 b_f 、 b_c 、 b_o 分别为各个门的偏置项, σ 为激活函数[13]。

BiLSTM 模型的具体结构是由两层相同的 LSTM 模型构成, 且两层的神经元数量相等。首先, 将正向的数据输入向前 LSTM 单元, 得到向前 LSTM 单元的输出, 类似的, 再将数据反向输入向后 LSTM 单元, 得到向后输出, 最后将向前和向后的输出结果按照一定的权重线性叠加, 从而得到最终的输出结果[14]。BiLSTM 神经网络模型本质上是将一个前向序列反向, 转化为 LSTM 模型的输入, 故它可以同时训练两个模型。前向 LSTM 使用过去的信息来预测未来信息, 后向 LSTM 则使用未来信息拟合过去信息[15]。模型结构如图 2 所示, 其中 $x_t (t \in [1, n])$ 表示 LSTM 单元的输入, $h_t (t \in [1, n])$ 对应于每一个输入通过 LSTM 单元得到的隐藏层输出。

2.3. 注意力机制

注意力机制可以看作模拟人脑注意力活动的一个行为[15]。注意力机制主要是基于概率来进行分配, 对关键信息给予足够的注意, 由此来突显其重要性, 从而提高模型的精度[4], 其结构如图 3 所示。其中, $x_t (t \in [1, n])$ 表示 BiLSTM 的输入, $h_t (t \in [1, n])$ 对应于每一个输入通过 BiLSTM 得到的隐藏层输出, $\alpha_t (t \in [1, n])$ 为注意力机制对隐藏层输出的注意力概率分布值, y 为注意力机制的 BiLSTM 输出值[16]。

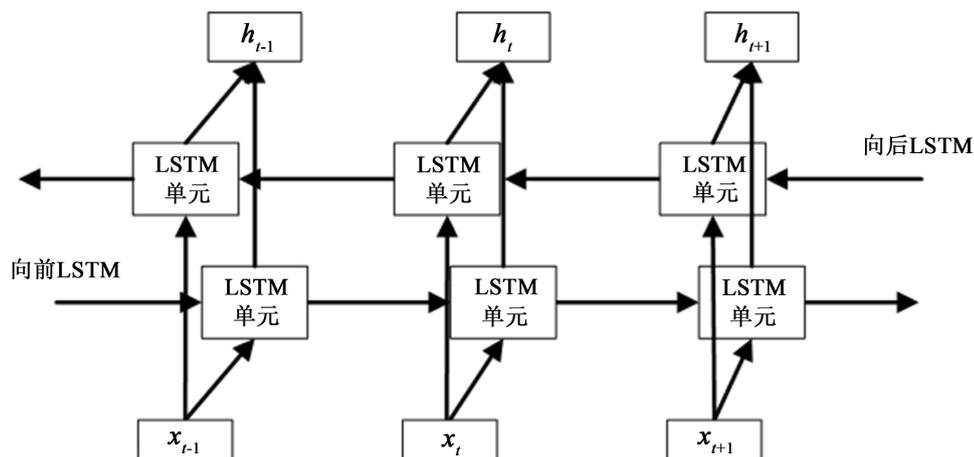


Figure 2. Two-way long short-term memory network flowchart
图 2. 双向长短期记忆网络流程图

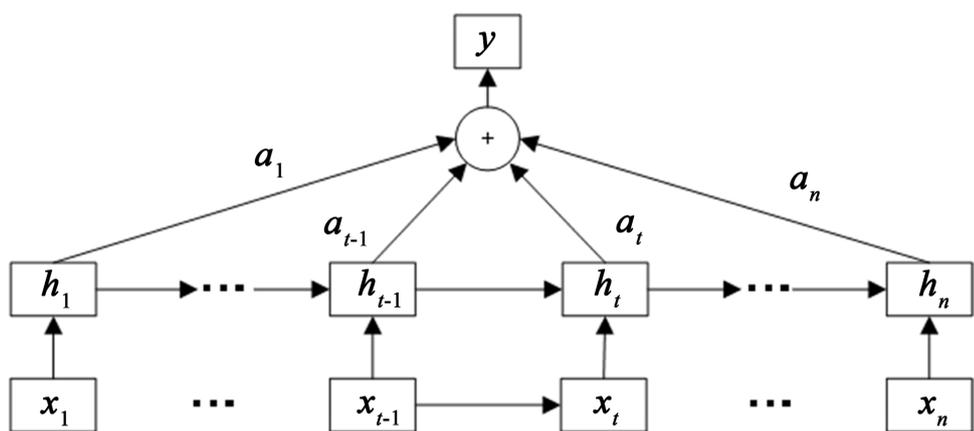


Figure 3. Attention mechanism flowchart
图 3. 注意力机制流程图

将 BiLSTM 隐藏层输出向量 $H = \{h_1, h_2, \dots, h_n\}$ 作为注意力机制的输入, 再通过公式计算出相应的权重 α_n , 如等式所示[17]:

$$e_i = \tanh(W_h h_i + b_n), e_i \in [-1, 1] \quad (13)$$

$$\alpha_i = \frac{\exp(e_i)}{\sum_{i=1}^t \exp(e_i)}, \sum_{i=1}^t \alpha_i = 1 \quad (14)$$

其中 e_i 注意力得分, 即上一时刻状态与本时刻输出的相似性; W_h 为权重矩阵, h_i 、 b_i 为权重和偏置[17]。注意力向量 $H' = \{h'_1, h'_2, \dots, h'_n\}$ 可以通过公式求得:

$$h'_i = \alpha_i \cdot h_i \quad (15)$$

最后将注意力向量与不同时间段输入得到的隐状态向量线性映射后相加得出最终的预测值。

2.4. 基于注意力机制的 VMD-BiLSTM 模型

相比与原始序列, 变分模态分解能够将整个序列分为多个子模态, 解决了数据中存在随机性和非线性

性等问题[18]。相比于 LSTM 模型, 虽然说 LSTM 模型解决了 RNN 的长期依赖问题, 但是没有考虑到更进一步的时间的信息, 所以针对发电数据, 无法学习到两个时间方向的关系, BiLSTM 正是针对这个问题进行构建的[7]。随后本文引入的注意力机制通过对 BiLSTM 输出的特征赋予不同的权重, 加强重要信息的影响, 减弱了其他噪音信息的影响, 从而使得模型能较好的捕获较长时间跨度的信息特征[7]。上述三种方法和模型的结合使得整体混合模型的预测结果得到进一步的提升。

本文模型流程如图 4 所示, 其算法设计步骤如下:

- 1) 将原始数据序列进行初始数据检查, 包括缺失值、重复值的查找及补全等。
- 2) 随后对序列数据进行变分模态分解, 在经过多次实验后, 最终本文发现变分模态分解超参数选择为 $k = 6$, 二次惩罚因子 $\alpha = 7000$, 将序列分解为 6 个子模态能较好的提升模型的整体预测能力。
- 3) 将变分模态分解之后得到的子模态作为 BiLSTM 的输入, 不同子模态分别输入不同的网络单元, 并通过前向和后向的网络进一步学习序列的相关性特征, 最终以隐状态的形式输出, 进一步作为注意力机制的输入。
- 4) 在注意力机制中, 隐状态行向量与注意力向量线性映射相加后再通过全连接层便可得出每个子模态的预测值[19], 最终将预测值求和便得出最终的预测结果, 随后得出相应的评价指标。

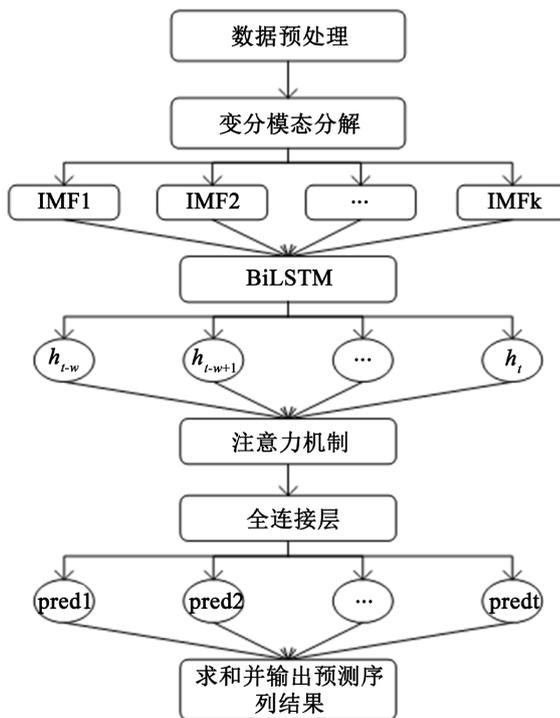


Figure 4. Model prediction flowchart
图 4. 模型预测流程图

2.5. 模型评价指标

本文评价指标为平均绝对误差(MAE)、平均绝对百分比误差(MAPE)和均方根误差(RMSE)。公式如下:

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

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (17)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (18)$$

式中, \hat{y}_i 为预测值, y_i 表示真实值, n 为预测点数目, 上述指标越小, 则表明模型预测结果越接近真实值[20]。

3. 实验分析

3.1. 数据介绍与处理

本文数据来源于 National Renewable Energy Laboratory (NREL), 为美国西南地区的风力发电数据, 本文选取的数据区间为 2020 年 1 月 1 日至 2020 年 2 月 29 日, 该数据集一共包含了三台风机, 每日数据点以每小时为单位, 每台风机包含 1440 条数据。本文针对其中风机编号为 303 的发电机组为实验对象。通过训练集和测试集的划分, 训练集有 1224 组数据, 测试集中包含 216 条数据。由于数据量庞大, 故只选取了部分数据进行展示, 如图 5 所示, 其中横轴为时间, 纵轴为发电量。随后查看数据的统计信息, 随后进行缺失值的查找和补全, 本文是运用均值对缺失值进行补全。

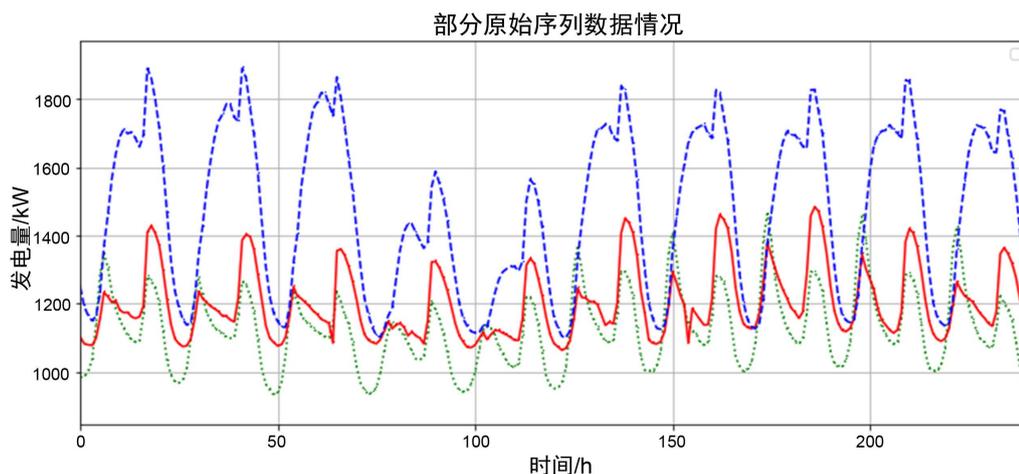


Figure 5. Partial raw series data case

图 5. 部分原始序列数据情况

3.2. 实验参数设置

由于模型超参数较多, 为了保证实验的客观性, 故将不同模型的超参数均设为同一标准, 具体设置结果如表 1 所示。

Table 1. Experimental parameter settings

表 1. 实验参数设置

超参数	超参数调试
VMD 模态分解个数 K	6
滑动窗口大小	80
训练集测试集划分比例	0.85

Continued

批量大小	32
隐藏层神经元个数	[32, 64]
训练轮数	120

3.3. 变分模态分解

在处理完数据的基础之上, 对 303 号风机的数据进行变分模态分解, 具体情况如图 6 所示。

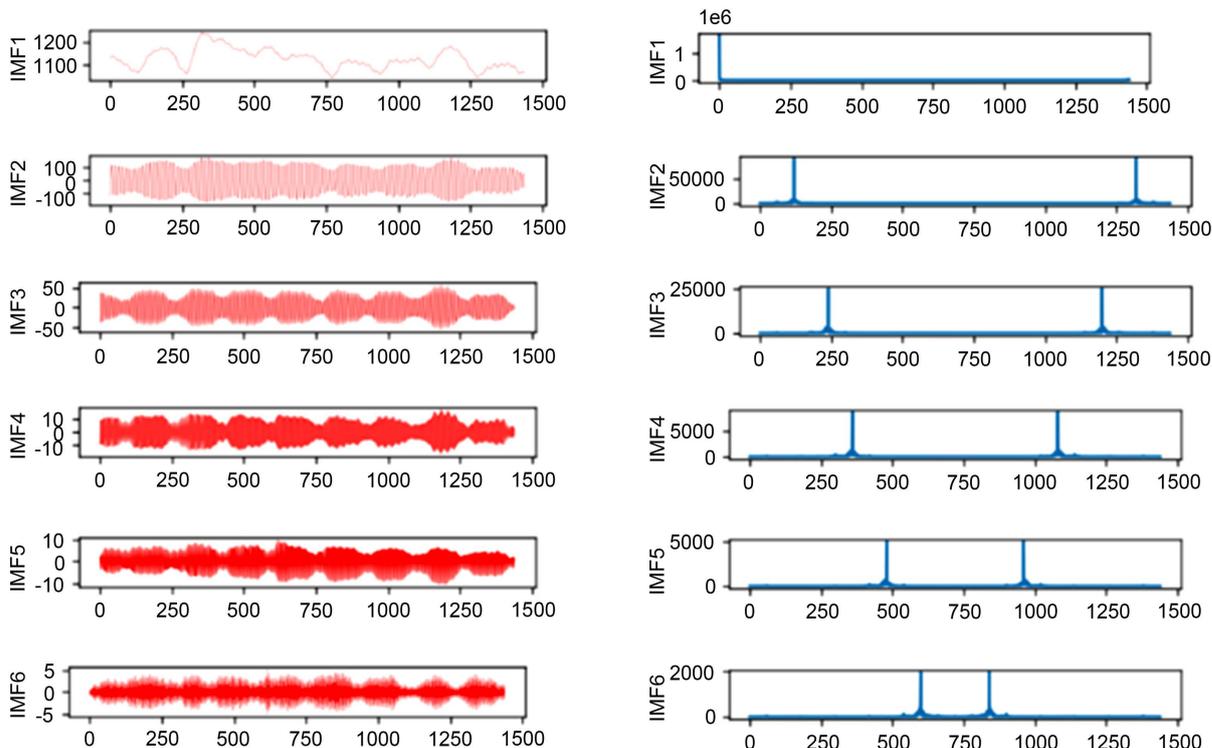


Figure 6. Variational mode decomposition of fan 303

图 6. 风机 303 的变分模态分解情况

采用 VMD 方法对数据进行降噪处理, 提取出数据中的特征规律。通过图 6 可见, 共将模态分解为 6 个 IMF, 模态 1 的波动频率相对较低, 反应出风力发电的波动性。其余模态都为高频模态, 在一定程度上体现了风力发电的随机性。从中心模态可以看出 6 个中心模态都不处于同一个频率上, 这也防止了模态分量混叠的问题。

3.4. 超短期风力发电预测结果分析

3.4.1. 分量模态预测分析

将变分模态分解后的各个分量模态进行不同模型的预测, 以 VMD-LSTM、VMD-BiLSTM-Attention 模型预测为例, VMD-BiLSTM-Attention 模型对模态 1 和模态 2 预测情况如下图 7、图 8 所示, 总体分量具体预测情况如下表。

从图 7、图 8 中可以看出, 模态 1 和模态 2 的预测果都较好, 而且从表 2、表 3 中也可以看出, VMD-BiLSTM-Attention 模型在对每个子模态下的预测结果大多优于 VMD-LSTM 模型。

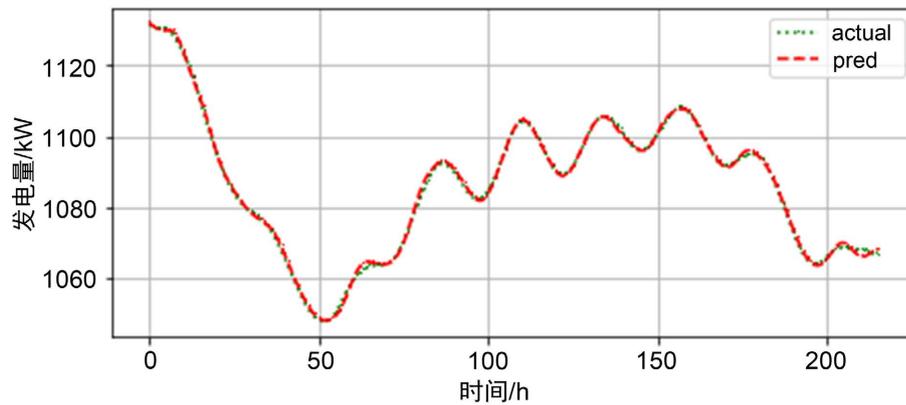


Figure 7. VMD-BiLSTM-Attention's prediction of modal 1

图 7. VMD-BiLSTM-Attention 对模态 1 的预测

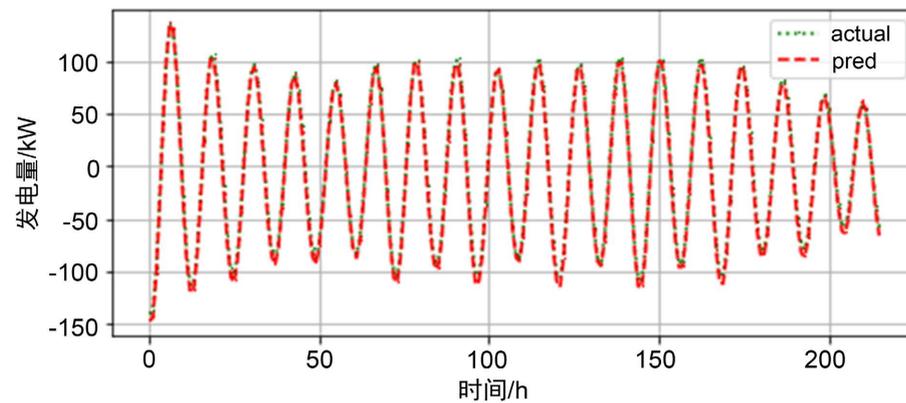


Figure 8. VMD-BiLSTM-Attention's prediction of modal 2

图 8. VMD-BiLSTM-Attention 对模态 2 的预测

Table 2. VMD-LSTM predicts individual modalities

表 2. VMD-LSTM 对各个模态预测

风机编号	模型	模态数	MAE/kW	RMSE/kW	MAPE/%
303	VMD-LSTM	IMF1	0.45	0.54	0.02
		IMF2	1.01	1.28	0.09
		IMF3	1.16	1.38	0.55
		IMF4	0.39	0.47	0.01
		IMF5	0.20	0.26	0.65
		IMF6	0.13	0.17	-1.30

Table 3. VMD-BiLSTM-Attention predicts individual modalities

表 3. VMD-BiLSTM-Attention 对各个模态预测

风机编号	模型	模态数	MAE/kW	RMSE/kW	MAPE/%
303	VMD-BiLSTM-Attention	IMF1	0.61	0.77	0.02
		IMF2	4.16	4.68	0.12
		IMF3	1.01	1.32	0.58

Continued

		IMF4	0.39	0.47	0.01
303	VMD-BiLSTM-Attention	IMF5	0.25	0.30	0.65
		IMF6	0.17	0.22	-1.28

3.4.2. 整合模态后的模型预测分析

本文首先通过变分模态分解, 再通过不同模型对各个模态分别进行预测, 最后将各个模态预测的结果进行平均, 从而得到最终的预测结果。下表 4、表 5 为编号 303 风机是否采用 VMD 分解预测的效果, 以此来证明 VMD 的可行性。最后本文针对风机 303 的数据采用了 6 种模型进行预测, 并分析其预测的具体结果, 如下表 5 所示。

Table 4. VMD's data prediction for number 303 was not used

表 4. 未采用 VMD 对编号 303 的数据预测

风机编号	模型	MAE/kW	RMSE/kW	MAPE/%
303	LSTM	50.08	58.39	0.05
	RNN	47.54	54.81	0.05
	GRU	48.81	56.02	0.04
	MLP	48.74	56.44	0.04
	BiGRU-Attention	48.59	55.86	0.7
	BiLSTM-Attention	48.53	55.89	0.04

Table 5. VMD was used to predict the number 303 data

表 5. 采用 VMD 对编号 303 数据预测

风机编号	模型	MAE/kW	RMSE/kW	MAPE/%
303	VMD-LSTM	13.69	18.35	0.11
	VMD-RNN	13.15	17.09	0.11
	VMD-GRU	15.65	19.90	0.10
	VMD-MLP	13.26	16.80	0.10
	VMD-BiGRU-Attention	15.21	19.73	0.11
	VMD-BiLSTM-Attention	12.62	16.85	0.11

从表 4 和表 5 的对比可以看出, 对于 6 个模型来说, 在使用变分模态分解之后, 无论是 MAP 还是 RMSE 都有很大的下降, 在对数据使用变分模态分解之后, 6 个预测模型的 MAP 分别下降了 36.39、34.39、33.16、35.48、33.38、35.91, MAPE 分别下降了 40.03、37.72、36.12、39.64、36.12、39.03。可见在运用变分模态分解之后能够较大的减少模型预测的误差, 同时也提升了模型的预测精度。由上述分析我们能够清楚发现, 本文所提出的 VMD-BiLSTM-Attention 预测模型的实际预期效果一般都要好于文中其他模型, 其次预测效果较好的是 VMD-GRU、VMD-BiGRU-Attention、VMD-RNN、VMD-MPL, 最后是 VMD-LSTM。

为了更加直观的展示结果, 下图 9 将 6 种模型对于风机 303 的预测结果进行比较, 从下图 9 可以看出, 6 种模型都对实际数据进行了预测, 虽然风力发电数据的波动较大, 噪音多, 但是通过训练后的模型仍然能够很好的抓取数据中的大部分特征, 而且从图中还可以看出, 在波峰处模型预测的效果要好于波谷处。

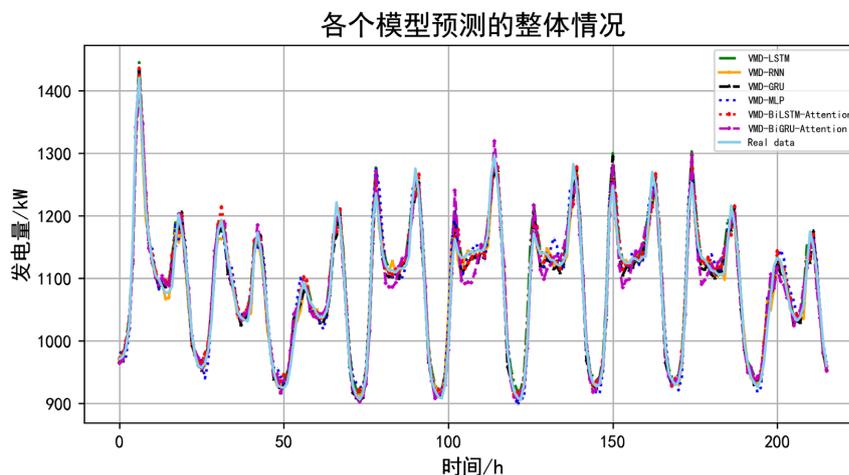


Figure 9. The overall forecast of the fan 303

图 9. 风机 303 整体预测情况

最后, 为了独立论证本文提出的 VMD-BiLSTM-Attention 模型的预测效果, 下图 10 为 303 号风机的原始发电数据与本文所提出的模型预测的数据对比。

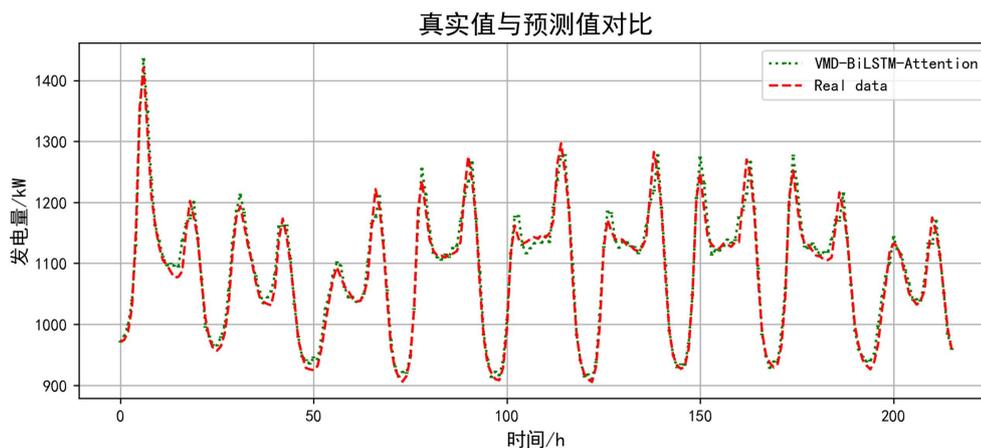


Figure 10. Prediction of the model VMD-BiLSTM-Attention

图 10. 模型 VMD-BiLSTM-Attention 的预测情况

从图 10 可以看出, VMD-BiLSTM-Attention 模型较为精准的预测了风力发电数据, 而且能够较好的捕获数据规律特征, 从上表 5 的误差指标中可知, 无论是预测精度还是误差方面, 本文提出的模型都优于其他模型, 最后可得, 本文提出的 VMD-BiLSTM-Attention 模型在风力发电领域有较大的应用潜力。

4. 结论

本文提出的 VMD-BiLSTM-Attention 模型, 通过 NREL 网站的美国西南部地区的电力数据集进行验证, 得到如下结论:

通过实验可知, 同一数据、同一模型在使用变分模态分解之后, 其 MAE 和 MAPE 分别能够下降 30%~40%。这一方法不但降低了数据和原始模型的复杂度, 同时也证明了变分模态分解在本文的可行性, 随后输入 BiLSTM 和注意力机制, 通过不断的实验, 提出了最终的预测模型, 最终发现, 在同一数据和实验参数下, 无论是与单一的神经网络模型还是本文的混合模型相比, 本文使用的 VMD-BiLSTM-Attention

在预测精度方面都高于其他模型, 表明该模型在风力发电预测方面有较好的效果, 同时也为风力发电预测领域提供了相应的理论参考价值。

在接下来的研究中, 可以选取特征变量丰富的数据, 例如包含气象条件的特征变量数据; 对超参数的选择, 通过网格搜索等方法来选择超参数, 进一步优化模型。

基金项目

湖北省教育厅科学研究计划指导性项目(B2022001)。

参考文献

- [1] 张爱枫, 段新宇, 何泉峰. 基于 CNN 和 LightGBM 的新型风电功率预测模型[J]. 电测与仪表, 2021, 58(11): 121-127.
- [2] Nepal, B., Yamaha, M., Yokoe, A., *et al.* (2020) Electricity Load Forecasting Using Clustering and ARIMA Model for Energy Management in Buildings. *Japan Architectural Review*, **3**, 62-76. <https://doi.org/10.1002/2475-8876.12135>
- [3] 佟佳弘, 武志刚, 谢钰. 缺少气象数据场景下的超短期风电预测[J]. 电力系统及其自动化学报, 2022, 34(3): 142-150.
- [4] Li, L.L., Zhao, X., Tseng, M., *et al.* (2020) Short-Term Wind Power Forecasting Based on Support Vector Machine with Improved Dragonfly Algorithm. *Journal of Cleaner Production*, **242**, Article ID: 118447. <https://doi.org/10.1016/j.jclepro.2019.118447>
- [5] 阳曾, 丁施尹, 叶萌, 李晶, 薛书倩, 吴昊天. 基于变分模态分解和深度学习的短期电力负荷预测模型[J]. 电测与仪表, 2022, 1-8.
- [6] 朱伟, 孙运全, 钱尧, 金浩, 杨海晶. 基于 CEEMD-GRU 模型的短期电力负荷预测方法[J]. 电测与仪表, 2023, 60(1): 16-22.
- [7] 赵兵, 王增平, 纪维佳, 高欣, 李晓兵. 基于注意力机制的 CNN-GRU 短期电力负荷预测方法[J]. 电网技术, 2019, 43(12): 4370-4376.
- [8] Meng, Y., Chang, C., Huo, J., *et al.* (2022) Research on Ultra-Short-Term Prediction Model of Wind Power Based on Attention Mechanism and CNN-BiGRU Combined. *Frontiers in Energy Research*, **10**, Article ID: 920835. <https://doi.org/10.3389/fenrg.2022.920835>
- [9] 梁智, 孙国强, 李虎成, 卫志农, 臧海祥, 周亦洲, 陈霜. 基于 VMD 与 PSO 优化深度信念网络的短期负荷预测[J]. 电网技术, 2018, 42(2): 598-606.
- [10] Dragomiretskiy, K. and Zosso, D. (2014) Variational Mode Decomposition. *IEEE Transactions on Signal Processing*, **62**, 531-544. <https://doi.org/10.1109/TSP.2013.2288675>
- [11] Wang, S., Liu, C., Liang, K., *et al.* (2022) Wind Speed Prediction Model Based on Improved VMD and Sudden Change of Wind Speed. *Sustainability*, **14**, 8705. <https://doi.org/10.3390/su14148705>
- [12] Tan, M., Yuan, S., Li, S., *et al.* (2022) Ultra-Short-Term Industrial Power Demand Forecasting Using LSTM Based Hybrid Ensemble Learning. *IEEE Transactions on Power Systems*, **35**, 2937-2948. <https://doi.org/10.1109/TPWRS.2019.2963109>
- [13] 肖浩逸, 何晓霞, 梁佳佳, 李春丽. 一种基于模态分解和机器学习的锂电池寿命预测方法[J/OL]. 储能科学与技术, 2022, 11(12): 3999-4009.
- [14] 潘洛华. 基于 SSA-CNN-BILSTM 的电力短期负荷数据预测方法[D]: [硕士学位论文]. 杭州: 浙江大学, 2022.
- [15] Tian, C., Niu, T. and Wei, W. (2022) Developing a Wind Power Forecasting System Based on Deep Learning with Attention Mechanism. *Energy*, **257**, Article ID: 124750. <https://doi.org/10.1016/j.energy.2022.124750>
- [16] 刘国海, 孙文卿, 吴振飞, 等. 基于 Attention-GRU 的短期光伏发电功率预测[J]. 太阳能学报, 2022, 43(2): 226-232.
- [17] Zhou, H., Zhang, Y., Yang, L., *et al.* (2019) Short-Term Photovoltaic Power Forecasting Based on Long Short Term Memory Neural Network and Attention Mechanism. *IEEE Access*, **7**, 78063-78074. <https://doi.org/10.1109/ACCESS.2019.2923006>
- [18] 刘杰, 金勇杰, 田明. 基于 VMD 和 TCN 的多尺度短期电力负荷预测[J]. 电子科技大学学报, 2022, 51(4): 550-557.

- [19] 王金玉, 金宏哲, 王海生, 张忠伟. ISSA 优化 Attention 双向 LSTM 的短期电力负荷预测[J]. 电力系统及其自动化学报, 2022, 34(5): 111-117.
- [20] 姜旭初, 许宇澄, 宋超. 短期风力发电负荷预测的新方法[J]. 北京师范大学学报(自然科学的版), 2022, 58(1): 39-46.