

基于变分模态分解的光伏功率预测方法^{*}

王世青¹ 赵许许¹ 廖俊龙¹ 詹鑫¹ 王建红¹ 邹德凡² 杨春萍²

(1. 国网新疆电力有限公司乌鲁木齐供电公司 乌鲁木齐 830001; 2. 华北电力大学电气与电子工程学院 北京 102206)

摘要:为解决光伏功率预测时不确定性强、影响因素较多等问题,提出一种基于变分模态分解(variational mode decomposition, VMD)的光伏功率预测方法。首先,对原始的光伏功率数据进行变分模态分解,分解为不同频率较平稳的模态分量;其次,计算不同模态分量的排列熵,根据排列熵将不同分量进一步合并,并在考虑不同影响因素(温度、辐射等)的条件下,分别将不同频率模态分量(intrinsic mode function, IMF)经过双向门控循环单元-自注意力(bi-directional gated recurrent unit-attention, BiGRU-Attention)模型进行预测;最后,将不同频率分量的预测结果叠加重构得到最终预测值。对中国某地区光伏功率数据进行实验测试,实验结果表明,提出的模型相较于BiGRU模型,平均绝对百分比误差(MAPE)、均方误差(RMSE)和平均绝对误差(MAE),分别降低了11.25%、8.51%和11.92%,其预测误差得到显著降低。

关键词:深度学习;光伏功率预测;VMD;GRU网络;注意力机制

中图分类号: TM615; TP183; TN9 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Photovoltaic power forecasting based on variational mode decomposition

Wang Shiqing¹ Zhao Xuxu¹ Liao Junlong¹ Zhan Xin¹ Wang Jianhong¹ Zou Defan² Yang Chunping²

(1. State Grid Xinjiang Electric Power Co. Ltd., Urumqi Power Supply Company, Urumqi 830001, China; 2. School of Electrical and Electronic Engineering, North China Electric Power University, Beijing 102206, China)

Abstract: To solve the problems of strong uncertainty and multiple influencing factors in photovoltaic power prediction, this paper proposes a photovoltaic power prediction method based on variational mode decomposition(VMD). Firstly, the original photovoltaic power data is subjected to variational mode decomposition, which decomposes it into modal components with relatively stable frequencies. Secondly, the permutation entropy of different modal components is calculated, and the different components are further merged based on the permutation entropy. Under the condition of considering different influencing factors (temperature, radiation, etc.), the different frequency modal components are predicted by the bi-directional gated recurrent unit-attention(BiGRU-Attention) model. Finally, the predicted results of different frequency components are superimposed and reconstructed to obtain the final predicted value. Experimental tests were conducted on photovoltaic power data in a certain region of China, and the results showed that compared to the BiGRU model, the model proposed in this paper reduced the mean absolute percentage error (MAPE), mean square error (RMSE), and mean absolute error (MAE) by 11.25%, 8.51%, and 11.92%, respectively, and significantly reduced its prediction error.

Keywords: deep learning; photovoltaic power forecast; VMD; GRU network; attention mechanism

0 引言

光伏发电技术^[1]已作为一种重要的新能源技术,光伏发电的输出功率会受到气象因素及自身硬件等不确定性

因素的影响,导致光伏功率具有随机性与不稳定性,这为光伏并网带来了巨大挑战。因此,提高光伏功率的预测精度至关重要^[2-4]。

目前进行光伏功率预测的方法主要是统计法^[5],统计

收稿日期:2024-08-02

^{*} 基金项目:国网新疆电力有限公司科技项目(5230WJ230007)资助

模型方法主要依赖于发电站的历史数据,通过建立与气象变量之间的统计关系,使用算法建立输入到输出的映射模型,实现光伏功率预测^[6]。比较常见的统计方法有反向传播^[7](back-propagation, BP)神经网络、极限学习机^[8](extreme learning machine, ELM)、最小二乘支持向量机^[9](least squares support vector machine, LSSVM)、长短时记忆网络^[10](long short-term memory, LSTM)等算法。文献[11]利用人工鱼群算法的全局寻优能力和内在并行计算能力优化 BP 神经网络的权值和阈值,通过 BP 神经网络对光伏进行预测。

目前光伏功率预测领域主要存在两个难点:1)不确定性强,光伏发电受到温度、风速、太阳辐射强度等影响因素,预测存在很大的不确定性并且难以精确,同时其还会受到自身硬件温度、电压等因素的影响,导致预测需要考虑多种输入特征;2)存在数据噪声与异常值,这些光伏功率数据中的噪声和异常值会影响模型预测的准确性。

随着人工智能的发展和硬件水平的提升,更适用于时序预测的循环神经网络(recurrent neural network, RNN)被设计出来,与传统神经网络相比, RNN 处理时序数据时能捕捉长时间段的相关特征,但是在时间段较长时易出现梯度爆炸的问题。LSTM 通过增加记忆单元等模块,一定程度上克服了该问题并得到广泛应用。文献[12]参考风电和光伏发电之间的互补性和相关性,构建了基于时序模式注意力的 LSTM 算法的区域联合发电预测模型,对风电和光伏发电进行预测。文献[13]基于 Res-UNet 模型对短波辐照(short wave radiation, SWR)网格进行时空预测,最后构建基于编解码器的 LSTM 模型预测光伏出力,其基于卫星遥感的光伏功率预测方法在处理空间信息时具有一定优势,但在处理时间序列的非平稳性时存在局限。文献[14]采用基于包络线聚类的多模融合算法,采用 Pearson 相关系数与 XGB Feature Importance 模块分析光伏功率和各特征之间的相关关系,虽然能够处理多模态信息,提高复杂天气情况下光伏功率预测精度,但在应对噪声和异常值方面效果有限。文献[15]通过自适应信号分解和双向 LSTM 的结合,成功捕捉了光伏功率时间序列中的长短期依赖关系。然而其局限性在于双向 LSTM 结构在处理长时间序列依赖时,计算复杂度较高,且模型的训练效率相对较低。文献[16]通过能量熵优化变分模态分解(variational mode decomposition, VMD)的分解过程,解决了潮汐预测中分潮提取的难题,并且在潮汐预测中展现出较高的精度。然而,其提出的方法的局限性在于依赖潮汐数据的特定频谱特征,在处理非线性和强噪声的光伏功率数据时,可能无法完全适用。

基于上述出现的问题,本文提出了一种基于 VMD 的光伏功率预测方法。此方法相比于上述光伏预测功率的方法有 3 个独特优势,通过 VMD 分解降低了数据的非平稳性,增强了时间序列建模的准确性;通过结合双向门控循环单元(bi-directional gated recurrent unit, BiGRU)和

自注意力机制(attention),进一步增强了模型对噪声和复杂非线性特征的鲁棒性和适应性,特别是在数据噪声较大的情况下,本文所提方法优于文献[12-13]提出的 LSTM 模型,总体来说本文方法通过 VMD 分解减少了数据的非平稳性,并结合 BiGRU 和自注意力机制,既保留了对长短期依赖关系的捕捉能力,又降低了计算复杂度,提高了模型的训练效率和预测效果。

本文方法的具体步骤是首先对原始的光伏功率数据进行 VMD 分解,分解为较平稳不同频率模态分量,降低了光伏数据的非平稳性和复杂性,同时通过 VMD 分解之后可以降低光伏数据中的噪声,提高模型输入数据的质量。其次,计算不同模态分量的排列熵,根据排列熵将不同分量合并重组,并在考虑不同影响因素(温度、辐射等)的条件下,分别将不同分量分别经过 BiGRU-Attention 模型进行预测,本文引入 BiGRU-Attention 模型是因为 BiGRU 的双向信息处理能力,能够捕获更多的时序依赖关系,增强对光伏功率不确定性的建模能力,同时本文引入自注意力机制,对多种特征信息进行加权处理,突出重要特征,增强模型对关键影响因素的捕捉能力,降低了其受自身硬件温度、电压等因素的影响。最后,将不同频率分量的预测结果重构相加得到最终预测值。本文模型通过 VMD、BiGRU 和自注意力机制的有机结合,与单一模型预测的实验对比表明,在预测光伏功率数据上具有更高的预测精度,验证了该模型的有效性。

1 模型结构

1.1 VMD-BiGRU-Attention 预测模型结构

本文提出的模型如图 1 所示,预测模型结构包括输入、VMD 分解、计算排列熵并合并序列、BiGRU-Attention

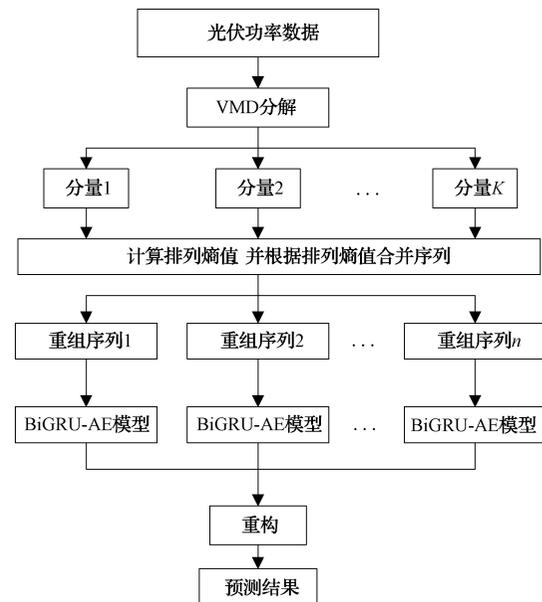


图 1 VMD-BiGRU-Attention 网络结构

Fig. 1 VMD-BiGRU-Attention network structure

预测、重构预测值并计算误差这几个部分。

1)对光伏功率的原始数据进行去除异常值等预处理,再将其输入到预测模型中。

2)通过VMD分解为K个不同频率的模态分量,计算各分量的排列熵值,将不同分量合并为n个重组序列,将重组后的序列矩阵与温度、辐照等特征信息的矩阵拼接。

3)分别将重组后的序列矩阵输入BiGRU-Attention模型中进行预测,选择ReLU函数搭建一层BiGRU网络,提取序列的时序信息与内部特征。然后经过注意力层计算输出向量的特征权重矩阵,挖掘更深层次的时序信息,突出关键信息。再通过激活函数为sigmoid的全连接层降维与提取特征向量。

4)将输出的预测值重构相加,得到最后的预测值并分析误差。

1.2 VMD 算法

变分模态分解 2014 年由 Dragomiretskiy 等^[17]提出,是一种用于信号处理的时频分析方法,可以有效提取信号的相关特征。其通过将信号分解为有限数量的本征模态函数,使得信号的分析变得更加直观和高效。VMD 在经验模态分解(empirical mode decomposition,EMD)等传统分解方法的基础上得到改进,解决了模态分量混叠和端点效应等问题。作为一种自适应、完全非递归的信号处理方法,能自适应地选择最优的中心频率和有限带宽^[18]。与EMD等分解方法相比,VMD能更好地处理噪声,避免噪声对分解结果的影响,且能有效降低复杂度和非线性强的时间序列非平稳性。VMD能够将光伏功率时间序列分解为多个模态,能够帮助提取关键特征、降低噪声、提高模型的预测能力。在短期光伏功率预测的任务中,VMD的应用能够显著增强模型对复杂时序数据的适应性和预测性能。

1)变分问题的构造

假设原信号能够分解为K个有限带宽的模态分量,在所有分量之和等于原信号的前提下,将各模态的估计带宽之和最小作为目标函数,则VMD约束变分模型如下:

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

式中: $f(t)$ 为待分解信号; $u_k(t)$ 为分解得到的第k个模态分量; $\delta(t)$ 为冲激函数; ω_k 为第k个模态分量的中心频率。

2)变分问题的求解

引入拉格朗日乘子 $\lambda(t)$ 和二次惩罚因子 α ,将该约束变分问题转化为无约束变分问题^[19]:

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

采用交替方向乘法求取式(2)的鞍点,得到最优的 u_k 和 ω_k 。

(1)初始化 $\{u_k^1\}, \{\omega_k^1\}, \lambda^1$, VMD初始迭代次数 $n=0$,在循环过程中 u_k, ω_k, λ 根据式(3)~(5)进行更新。

(2)在非负频率区间内,更新 u_k 如下:

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

式中: $\hat{u}_k^{n+1}(\omega), \hat{\lambda}^n(\omega), \hat{f}(\omega)$ 分别为相应函数的傅里叶变换。

(3)更新 ω_k 如下:

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

(4)更新 λ :

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

(5)重复步骤(2)~(5),直到满足迭代停止条件。

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

1.3 基于BiGRU-Attention的预测模型

将通过VMD分解后的各分量与收集到的特征信息进行归一化处理,并分别经过BiGRU模型进行预测,同时为了充分挖掘时序数据的关联信息,在输出侧引入自注意力机制即通过Attention层对时序信息进行进一步提取,由此得到BiGRU-Attention结构,如图2所示。

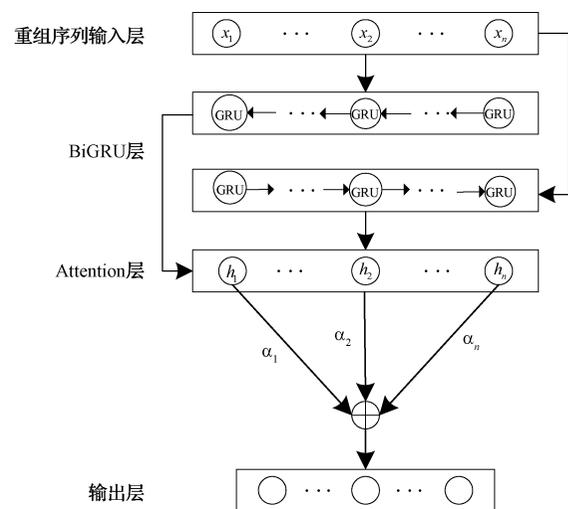


图2 BiGRU-Attention模型结构

Fig. 2 Structure of BiGRU-Attention model

1)BiGRU模型

RNN能够通过共享权重参数,挖掘时间序列中不同时间点间的信息,但随着时间序列的变长会发生梯度消失

和梯度爆炸的问题。门控循环单元(gated recurrent unit, GRU)基于RNN的结构增加了重置门和更新门,能够根据历史数据选择记忆相关时序信息或删除部分不相关信息,从而提高网络捕捉序列中较远时间点信息的能力。其中重置门是决定如何将新的输入与之前的记忆相结合,更新门则是决定保持多少过往的信息。计算公式如下:

$$R_t = \sigma(W_r \cdot [H_{t-1}, X_t]) \quad (7)$$

$$Z_t = \sigma(W_z \cdot [H_{t-1}, X_t]) \quad (8)$$

式中: R_t 表示的是重置门; Z_t 表示的是更新门; σ 是 sigmoid 激活函数,确保输出值在 0~1 之间。重置门 R_t 是由隐藏状态 H_{t-1} 与输入 X_t 拼接后与权重 W_r 相乘,再通过函数将数据维持到 0~1;更新门 Z_t 的计算方式与重置门类似,是由隐藏状态 H_{t-1} 与输入 X_t 负责决定 t 时刻更新门的输入。

新的候选状态 \tilde{H}_t 与最终状态 H_t 的计算公式如下:

$$\tilde{H}_t = \tanh(W_h [R_t \times H_{t-1}, X_t]) \quad (9)$$

$$H_t = Z_t \times H_{t-1} + (1 - Z_t) \times \tilde{H}_t \quad (10)$$

式中: H_{t-1} 表示隐藏状态; R_t 表示重置门; X_t 表示输入; W_h 表示权重矩阵; Z_t 表示新候选状态的信息; $(1 - Z_t)$ 表示保留的旧状态信息; \tanh 是激活函数, \tanh 作用是将输出值固定在 -1~1 的范围内。最终状态 H_t 通过这样加权平均的方法,GRU 能够平滑地在旧信息和新信息之间进行切换。GRU 网络结构如图 3 所示^[20]。

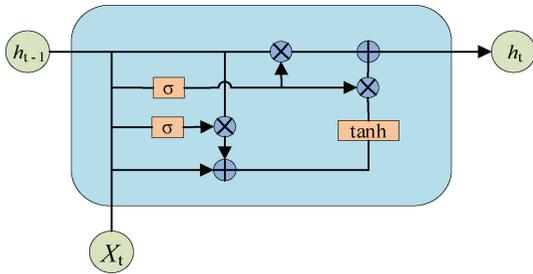


图 3 GRU 网络结构

Fig. 3 GRU network structure

GRU 模块有效的通过门控机制来控制信息的流动,此种创新方式解决了传统的 RNN 模型在长序列中常见的梯度消失和爆炸的问题。它以较少的参数和较快的计算速度应对各种序列模型任务,与 LSTM 相比较,GRU 结构更加简单。

与 GRU 模块相比,BiGRU 通过其双向结构能够同时捕获输入序列的正向和反向信息,在处理噪声数据或异常值时表现出较强的鲁棒性。即使输入序列中存在噪声或异常值,BiGRU 也能够通过门控机制对其进行过滤或调整,从而保持模型的稳定性。

BiGRU 模型结构如图 4 所示,BiGRU 是 GRU 的一种扩展,其是通过双向处理序列信息来改善模型性能。由图 4 可以看出,BiGRU 的核心思想是同时使用两个 GRU

网络,一个处理正向序列,另一个处理反向序列,这种双向结构允许网络在生成每个时间步的隐藏状态时考虑到上下文信息,从而增强模型对全局信息的理解。因此,本文采用 BiGRU 模型充分挖掘序列的时序信息^[21-22]。

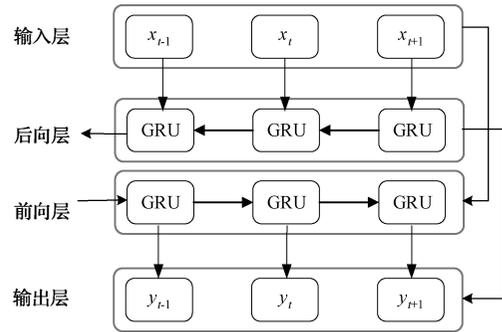


图 4 BiGRU 模型结构

Fig. 4 BiGRU model structure

2) 自注意力机制

注意力机制是模拟人脑对视觉图像的处理机制,将更多的注意力分配到一部分内容,同时减少对另一部分内容的关注甚至忽略。自注意力机制是由注意力机制改进而来,是一种在序列建模中广泛使用的机制,其核心思想是通过权重分配来让模型在处理序列中的某个位置时,充分关注序列中的其他位置,根据数据内部自身的相关性来分配注意力资源,更容易提取全局信息,从而提高了运算效率和准确率^[23-24]。自注意力机制结构如图 5 所示。

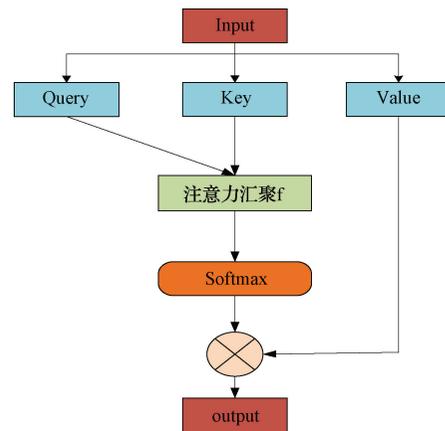


图 5 自注意力机制结构

Fig. 5 Structure diagram of self-attention mechanism

首先,将输入 X 通过 3 个不同的线性层,生成了查询(query, Q)、键(key, K)、值(value, V) 3 个矩阵,如式(11)所示。

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} w_1 = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{bmatrix} = Q$$

$$\begin{aligned} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \mathbf{w}_2 &= \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} = \mathbf{K} \\ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \mathbf{w}_3 &= \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \mathbf{V} \end{aligned} \quad (11)$$

式中： $[x_1 \ x_2 \ \dots \ x_n]^T$ 为输入矩阵， \mathbf{w} 为权重矩阵。

接着，使用注意力评分函数 f 来计算输入向量之间的相关性分布。然后，对这个分布进行 softmax 处理，以生成具有可解释性的概率分布值。计算公式如下：

$$\alpha = \frac{QK^T}{\sqrt{d_k}} \quad (12)$$

$$Attention(Q, K, V) = softmax(\alpha)V \quad (13)$$

在计算注意力权重时，使用查询 Q 和键 K 的点积来计算注意力权重。结果通过分子的缩放和 softmax 函数标准化为概率分布，其中 d_k 是键的维度，对其开方是防止点积维度过大，导致 softmax 的梯度消失。最后用计算出的注意力权重对值进行加权求和，得到当前输入位置的最终表示。

自注意力机制的特点主要有两方面：(1)全局依赖建模，能够直接捕捉输入序列中任意两位置键的依赖关系，而不受到相对位置的限制，适合处理长距离依赖问题；(2)可并行化，如图 5 所示，能够并行处理输入序列中不同元素的注意力全职工，提高了模型的训练速度。同时其具有更少的参数和较低的计算复杂度。这使得模型在训练时更容易优化，同时也降低了过拟合的风险。因此，为提取更多的特征信息，本文采用自注意力机制。

2 实验与结果分析

2.1 实验数据与预处理

本文实验测试所用数据集为中国某地区 2018 年真实光伏发电功率数据，共 305 000 条，记录了当地实际光伏发电功率、温度、气压和辐射等 8 个特征信息变量，时间颗粒度为 15 min，将这 8 种特征信息与分解后的模态分量拼接，以滑动窗口的方式输入 BiGRU-Attention 模型中，训练、验证、测试集比例分别为 7 : 1 : 2，如表 1 所示。实验将过去 96 个时间点作为参考的历史信息，预测未来 24 个时间点的光伏功率值，即通过过去 1 d 的历史光伏功

表 1 数据集划分时段表

Table 1 Time division table of dataset

数据类型	选取时段	时段数
数据集	2019/1/1—2019/12/31	35 040
训练集	2019/1/1—2019/9/13	24 528
验证集	2019/9/14—2019/10/19	3 504
测试集	2019/10/20—2019/12/31	7 008

率、温度、辐射等特征信息，预测未来 6 h 光伏功率值。

表 1 中数据集划分依据主要是基于 3 个方面。

1) 本文数据集选择 15 min 作为时间颗粒度，是因为光伏发电受云层变化以及太阳辐射等因素影响，如果选用 1 h 或 0.5 h 作为时间颗粒度的话，数据集的数量会减少（即计算量会减少），但是会导致模型丢失短时间内变化的关键信息，导致训练好的模型可能无法满足实际需求。如果选用 1 min 作为时间颗粒度的话，数据集中的无关信息的噪声将会大大增加，会导致模型训练困难，造成训练精度下降。

2) 本文选择一年的光伏发电数据作为本文的数据集，主要考虑到数据集完全包含了一年受季节影响的温度，太阳辐射等变化因素，这对于捕捉光伏发电的季节性规律十分重要，有助于提高模型的预测稳定性和普适性。

3) 训练集：验证集：测试集的划分比例为 7 : 1 : 2，这种数据集的划分方法是基于时间连续性考虑，确保模型不会发生数据泄露等问题，有助于提高模型的泛化能力。

BiGRU-Attention 模型参数如表 2 所示。

表 2 BiGRU-Attention 模型参数设定

Table 2 BiGRU-Attention model parameter settings

参数类型	参数设置
隐藏层层数	2
输入层神经元个数	8
隐藏层神经元个数	256
Dropout 率	0.2
激活函数	ReLU
优化器	Adam
学习率	0.001
权重衰变率	0.01

为了提高预测的准确性，采取了如下措施：对于数据集中的缺失值和异常值，我们用最近两点的平均值代替。此外，还对数据进行归一化处理，以消除不同量纲的影响，提高模型收敛速度。对光伏功率和影响功率值的特征变量归一化：

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (14)$$

式中： x 为各特征信息（光伏功率值，辐照值等）； x_{\max} 、 x_{\min} 分别为相应特征信息的最大、最小值； x^* 为归一化到 0~1 的结果。

2.2 VMD 的分解与重构

针对原始光伏功率往往存在大量噪声干扰问题，通过 VMD 分解对光伏功率数据进行处理，得到 7 组模态分量的分解结果如图 6 所示，7 组模态分量频率较稳定且未发生明显的模态混叠现象。根据图 6 可以观察到，相邻的模态分量具有相近的频率值。为了降低预测的复杂度，首先计算所有 IMF 分量的排列熵值。接着，根据这些排列熵

值,对 IMF 分量进行合并与重组(表 3)。由表 3 可以看出,IMF 分量的排列熵值能够表明该序列的随机程度,一般地随着频率的提高,排列熵值也在不断增加。

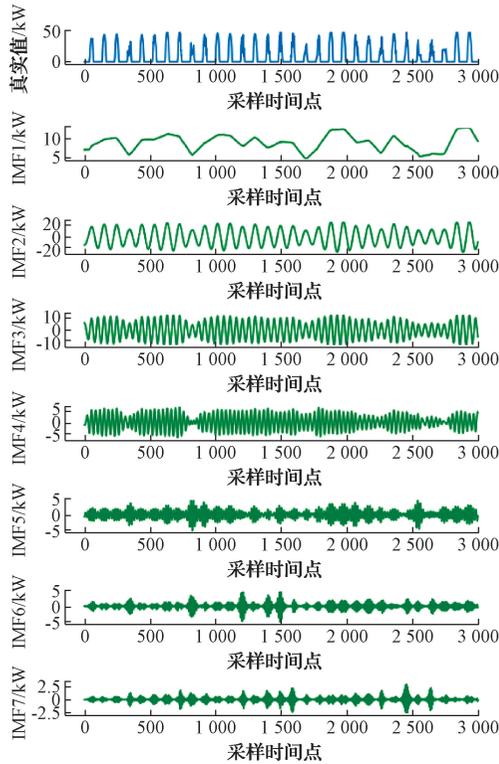


图 6 VMD 分解结果

Fig. 6 VMD decomposition results

表 3 各分量的排列熵值

Table 3 Arrangement entropy values of each component

分量	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7
排列熵	0.515	0.451	0.499	0.549	0.621	0.707	0.831

根据排列熵值反映出的各 IMF 分量的随机程度,将相邻样本熵差值在 0.07 以内的 IMF 分量合并以降低模型复杂度。如表 3 所示,IMF1~IMF4 分量样本熵相近,将其合并为 H1,IMF5、IMF6 和 IMF7 排列熵差别过大,因此将它们作为单独的序列,分别为 H2、H3、H4,再将其分别输入预测模型中。

2.3 实验环境

本文实验环境如下:操作系统 Win10,处理器 AMD Ryzen 7 4800H with Radeon Graphics 2.90 GHz,显卡 NVIDIA GeForce RTX 2060,内存 16 GB,开发语言 Python3.9,框架为 Pytorch。

2.4 预测结果分析

为验证本文所提出的模型有效性,与 ELM、SVR^[25]、GRU-Attention、LSTM、SSA-BiLSTM^[15]、VMD-GRU^[16]、TPA-LSTM^[12] 7 种模型进行对比实验,GRU 隐藏层大小

为 256,注意力层采用自注意力机制。对所有模型分别训练 100 轮,并使用早停法防止过拟合,停止条件为连续 20 轮在验证集上误差未下降。设置学习率为 0.001 的 Adam 优化器。为了确保实验的有效性及其可靠性,所有实验的条件和环境均相同的。

实验结果如表 4 所示,通过与其他模型的预测结果对比可知,本文提出模型的平均绝对百分比误差(MAPE)、均方根误差(RMSE)和平均绝对误差(MAE)更低,有更高的预测准确率。同时,加入了自注意力机制的 GRU-Attention 模型预测误差相对更低,说明注意力机制能够帮助模型更好地挖掘光伏数据中的时序信息。同时,从预测结果可以看出,在加入 VMD 分解以后,预测准确率也得到提升,因为 VMD 分解能够将不平稳的噪声多的原始光伏数据转化为较平稳的不同频率分量,降低了预测的难度和误差。

表 4 不同模型预测误差

Table 4 Prediction errors of different models

模型	MAPE	RMSE/kW	MAE/kW
ELM	0.175	5.658	4.289
SVR	0.158	5.326	3.532
GRU-Attention	0.149	5.106	3.288
LSTM	0.187	6.032	4.756
SSA-BiLSTM ^[15]	0.175	5.304	3.457
VMD-GRU ^[16]	0.143	5.036	3.420
TPA-LSTM ^[12]	0.162	5.374	3.589
VMD-BiGRU-Attention	0.142	4.893	3.132

为更进一步观察不同模型的预测结果,国内外研究现状中提到的 3 种代表性算法 SSA-BiLSTM^[15]、VMD-GRU^[16]、TPA-LSTM^[12] 与本文所提算法在 2 d 的光伏预测结果对比。如图 7 所示,能够看出所有模型都有较好的拟合结果,其中本文提出的 VMD-BiGRU-Attention 模型预测误差更低。

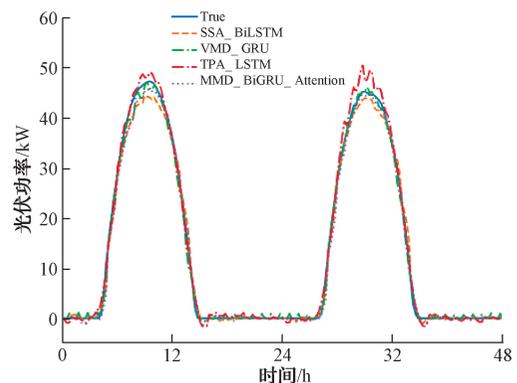


图 7 不同模型 2 d 预测结果

Fig. 7 Two day prediction results of different models

为了进一步验证模型的有效性,对该组合模型进行消融实验,如表5所示,在加入VMD分解或注意力机制后,模型的预测误差都有所降低,而在同时加入VMD与注意力机制后,预测精度进一步提高。

表5 消融对比实验

Table 5 Comparison experiment of ablation

模型	MAPE	RMSE/kW	MAE/kW
BiGRU	0.160	5.348	3.556
VMD-BiGRU	0.152	5.245	3.487
BiGRU-Attention	0.146	4.956	3.145
VMD-BiGRU-Attention	0.142	4.893	3.132

3 结论

针对光伏数据非线性强、影响因素较多和预测复杂等问题,本文提出了一种基于VMD分解的短期光伏功率预测方法。首先对原始光伏功率数据进行VMD分解,然后计算分解出的不同模态分量的排列熵,再根据排列熵对不同模态分量合并重组,以降低预测模型复杂度。再构建基于自注意力机制的BiGRU-Attention模型,最后将重组后的分量分别输入该模型,充分挖掘模态分量与温度、辐射等特征间的相关信息,将预测得到的不同分量的叠加,得到最终预测结果。实验结果表明,与ELM、SVR、GRU-Attention这3种不同模型相比,本文提出的VMD-BiGRU-Attention模型的预测结果的MAPE分别降低了18.65%、10.13%、4.69%,RMSE分别降低了13.52%、8.13%、4.17%,MAE分别降低了26.96%、11.33%、4.74%。通过在真实光伏数据集的实验说明,本文提出的预测模型有一定的应用潜力。接下来计划在该方向上,增加光伏预测时参考的历史信息及影响因素,以提高预测精度并提升预测效率。

参考文献

[1] LI Z, WU L, XU Y, et al. Distributed tri-layer risk-averse stochastic game approach for energy trading among multi-energy microgrids[J]. Applied Energy, 2023, 331.

[2] 赵书强, 张婷婷, 李志伟, 等. 基于数值特性聚类的日前光伏出力预测误差分布模型[J]. 电力系统自动化, 2019, 43(13): 36-45.

ZHAO SH Q, ZHANG T T, LI ZH W, et al. Error distribution model for predicting solar power output based on numerical clustering [J]. Power System Automation, 2019, 43(13): 36-45.

[3] 王诚良, 朱凌志, 党东升, 等. 云团移动对光伏电站出力特性及系统调频的影响[J]. 可再生能源, 2017,

35(11): 1626-1631.

WANG CH L, ZHU L ZH, DANG D SH, et al. Cloud clusters move towards the photovoltaic power station the influence of force characteristics and system frequency modulation[J]. Renewable Energy, 2017, 35(11): 1626-1631.

[4] 张倩, 马愿, 李国丽, 等. 频域分解和深度学习算法在短期负荷及光伏功率预测中的应用[J]. 中国电机工程学报, 2019, 39(8): 2221-2230,5.

ZHANG Q, MA Y, LI G L, et al. Application of frequency domain decomposition and deep learning algorithms in short term load and photovoltaic power prediction [J]. Chinese Journal of Electrical Engineering, 2019, 39(8): 2221-2230,5.

[5] WANG K, QI X, LIU H. Photovoltaic power forecasting based LSTM-convolutional network [J]. Energy, 2019, 189: 116225.

[6] 王瑞, 杨源浩, 逯静. 基于CDTW相似日和CSSA-LSTM的短期光伏功率预测[J/OL]. 河海大学学报(自然科学版), 1-15[2024-08-01].

WANG R, YANG Y H, LU J. Short term photovoltaic power prediction based on CDTW similarity day and CSSA-LSTM[J/OL]. Journal of Hehai University (Natural Science Edition), 1-15 [2024-08-01].

[7] ZHU C, WANG Z, LONG F, et al. Prediction of ground settlement of subway shield based on ABC-BP neural network [J]. Journal of Hehai University (Natural Sciences), 2023, 51(4): 72-80.

[8] LI Q, ZHANG X, MA T, et al. A multi-step ahead photovoltaic power prediction model based on similar day, enhanced colliding bodies optimization, variational mode decomposition, and deep extreme learning machine [J]. Energy, 2021, 224 (2): 120094.

[9] 张涛, 朱瑞金, 扎西顿珠. 基于改进骨干差分进化算法优化LSSVM的短期光伏发电功率预测[J]. 热力发电, 2021, 50(5): 102-107.

ZHANG T, ZHU R J, ZHAXI D ZH. Short term photovoltaic power prediction based on improved backbone differential evolution algorithm optimized LSSVM [J]. Thermal Power Generation, 2021, 50(5): 102-107.

[10] 吉铎格, 李慧, 刘思嘉, 等. 基于MIE-LSTM的短期光伏功率预测[J]. 电力系统保护与控制, 2020,

- 48(7): 50-57.
- JI X G, LI H, LIU S J, et al. Short term photovoltaic power prediction based on MIE-LSTM[J]. Power System Protection and Control, 2020, 48(7): 50-57.
- [11] 陈文进, 朱峰, 张童彦, 等. 基于 AFSA-BP 神经网络的光伏功率预测方法[J]. 浙江电力, 2022, 41(4): 7-13.
- CHEN W J, ZHU F, ZHANG T Y, et al. Photovoltaic power prediction method based on AFSA-BP neural network [J]. Zhejiang Electric Power, 2022, 41(4): 7-13.
- [12] CHEN Y J, XIAO J W, WANG Y W, et al. Regional wind-photovoltaic combined power generation forecasting based on a novel multi-task learning framework and TPA-LSTM[J]. Energy Conversion and Management, 2023, 297.
- [13] 刘晓艳, 王珏, 姚铁锤, 等. 基于卫星遥感的超短期分布式光伏功率预测[J]. 电工技术学报, 2022, 37(7): 1800-1809.
- LIU X Y, WANG Y, YAO T CH, et al. Ultra short term distributed photovoltaic power prediction based on satellite remote sensing[J]. Journal of Electrical Engineering, 2022, 37(7): 1800-1809.
- [14] 杨国清, 张凯, 王德意, 等. 基于包络线聚类的多模融合超短期光伏功率预测算法[J]. 电力自动化设备, 2021, 41(2): 39-46.
- YANG G Q, ZHANG K, WANG D Y, et al. Multi mode fusion ultra short term photovoltaic power prediction algorithm based on envelope clustering [J]. Power Automation Equipment, 2021, 41(2): 39-46.
- [15] 袁建华, 蒋文军, 李洪强, 等. 基于 SSA-BiLSTM 非线性组合方法的光伏功率预测[J]. 电子测量技术, 2023, 46(21): 63-71.
- YUAN J H, JIANG W J, LI H Q, et al. Photovoltaic power prediction based on SSA BiLSTM nonlinear combination method [J]. Electronic Measurement Technology, 2023, 46(21): 63-71.
- [16] 赵杰, 解则晓, 刘世萱. 基于能量熵 VMD 最优分解与 GRU 循环神经网络的潮汐预测精度提升方法研究[J]. 仪器仪表学报, 2023, 44(12): 79-87.
- ZHAO J, JIE Z X, LIU SH X. Research on the method of improving tidal prediction accuracy based on energy entropy VMD optimal decomposition and GRU recurrent neural network[J]. Chinese Journal of Scientific Instrument, 2023, 44(12): 79-87.
- [17] DRAGOMIRETSKIY K, ZOSSO D. Variational mode decomposition[J]. IEEE Transactions on Signal Processing, 2014, 62(3): 531-544.
- [18] LYU L L, WU Z Y, ZHANG J H, et al. A VMD and LSTM based hybrid model of load forecasting for power grid security [J]. IEEE Transactions on Industrial Informatics, 2022, 18(9): 6474-6482.
- [19] 韩莹, 朱宏宇, 李琨. 融合聚类及随机配置网络的短期光伏功率预测[J]. 电子测量与仪器学报, 2023, 37(11): 205-216.
- HAN Y, ZHU H Y, LI K. Short term photovoltaic power prediction using fusion clustering and random configuration network [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(11): 205-216.
- [20] 朱正林, 张冕. 基于 AO 优化 VMD-CE-BiGRU 的光伏发电功率预测[J]. 国外电子测量技术, 2022, 41(10): 56-61.
- ZHU ZH L, ZHANG M. Photovoltaic power prediction based on AO optimization of VMD-CE BiGRU [J]. Foreign Electronic Measurement Technology, 2022, 41(10): 56-61.
- [21] 李善寿, 马泉杰, 潘璐茜, 等. 基于 VMD 和 LSTM-CNN 的短期负荷预测模型研究[J]. 控制工程, 2023, 30(3): 469-478.
- LI SH SH, MA X J, PAN L Q, et al. Research on short term load forecasting model based on VMD and LSTM-CNN[J]. Control Engineering, 2023, 30(3): 469-478.
- [22] 侯恩科, 夏冰冰, 吴章涛, 等. 基于 CEEMDAN-BO-BiGRU 的矿井涌水量预测研究[J]. 科学技术与工程, 2023, 23(28): 12012-12019.
- HOU EN K, XIA B B, WU ZH T, et al. Research on mine water inflow prediction based on CEEMDAN-BO-BiGRU[J]. Science Technology and Engineering, 2023, 23(28): 12012-12019.
- [23] 穆晨宇, 薛文斌, 穆羨瑛, 等. 基于 VMD-LSTM-Attention 模型的短期负荷预测研究[J]. 现代电子技术, 2023, 46(17): 174-178.
- MU CH Y, XUE W B, MU X Y, et al. Research on short-term load forecasting based on VMD-LSTM-Attention model[J]. Modern Electronics Technique, 2023, 46(17): 174-178.
- [24] 朱继忠, 苗雨旺, 董朝阳, 等. 基于 Attention-LSTM

与多模型集成的短期负荷预测方法[J]. 电力工程技术, 2023, 42(5): 138-147.

ZHU J ZH, MIAO Y W, DONG CH Y, et al. Short term load forecasting method based on attention LSTM and multi model integration [J]. Electric Power Engineering Technology, 2023, 42(5): 138-147.

[25] SU X H, HE X L, ZHANG G, et al. Research on SVR water quality prediction model based on improved

sparrow search algorithm [J]. Computational Intelligence and Neuroscience, 2022, 2022: 7327072.

作者简介

王世青, 副高级工程师, 主要研究方向为负荷预测。

E-mail: waitroad@sina.com

杨春萍(通信作者), 博士, 副教授, 主要研究方向为人工智能、目标识别、负荷预测。

E-mail: ychp@ncepu.edu.cn