

DOI:10.19651/j.cnki.emt.2519243

基于 CW-EMA 分解与组合神经网络的风电功率预测<sup>\*</sup>王雯学<sup>1</sup> 乔燕军<sup>2</sup> 寇志伟<sup>1</sup> 崔啸鸣<sup>1</sup> 任刚<sup>1</sup>

(1. 内蒙古工业大学电力学院 呼和浩特 010080; 2. 北京京能国际控股有限公司北方分公司 呼和浩特 010000)

**摘要:** 风电功率具有强波动性与随机性,给电网调度与风电并网运行带来挑战。为提升风电功率预测的精度,本文提出一种基于通道独立指数滑动平均分解(CW-EMA)与组合神经网络的超短期风电功率预测模型(CESF-Net)。该模型首先通过噪声密度聚类算法(DBSCAN)识别并剔除风速-功率关系中的离群点,并采用线性插值进行缺失值填补;其次,利用 CW-EMA 实现多变量序列在通道维度上的趋势与季节分解;然后,利用时间序列分段机制提高模型对局部时间结构的捕捉能力,并构建快速傅里叶变换注意力与门控循环单元的双流网络,分别对季节性特征与趋势性特征进行处理;最后,将双流网络的输出进行拼接,形成 CESF-Net 模型。在风电场实测数据集上的实验表明,CESF-Net 模型在 15 min 预测任务中,MAE、RMSE、 $R^2$  较常用模型分别提升 18.78%、11.11%、0.26%;在 60 min 预测任务中,尽管各模型的预测精度均有所下降,但 CESF-Net 的各项指标仍分别提升了 2.98%、2.74% 和 0.61%。

**关键词:** 通道独立指数滑动平均分解;快速傅里叶变换注意力;时间序列分段;风电功率预测

**中图分类号:** TM614;TN91 **文献标识码:** A **国家标准学科分类代码:** 470.4047

## Wind power prediction based on CW-EMA decomposition and combined neural network

Wang Wenxue<sup>1</sup> Qiao Yanjun<sup>2</sup> Kou Zhiwei<sup>1</sup> Cui Xiaoming<sup>1</sup> Ren Gang<sup>1</sup>

(1. The College of Electric Power, Inner Mongolia University of Technology, Hohhot 010080, China;

2. Beijing Energy International Holding Co., Ltd., Northern Branch, Hohhot 010000, China)

**Abstract:** Wind power output exhibits significant volatility and randomness, posing challenges to grid scheduling and wind power integration. To enhance forecasting accuracy, this paper presents CESF-Net, an ultra-short-term wind power prediction model that combines channel-wise EMA decomposition with a hybrid neural network. First, the model employs the density-based spatial clustering of applications with noise (DBSCAN) algorithm to identify and remove outliers in the wind speed-power relationship, and uses linear interpolation to impute missing values. Second, CW-EMA is applied to decompose multivariate time series into trend and seasonal components along the channel dimension. Then, a time-series segmentation mechanism is introduced to enhance the model's ability to capture local temporal structures, and a dual-stream network based on fast fourier transform attention and gated recurrent units is constructed to separately extract seasonal and trend features. Finally, the outputs of the dual streams are concatenated to generate the prediction through CESF-Net. Experiments conducted on real wind farm datasets demonstrate that, for the 15-minute forecasting task, the CESF-Net model outperforms commonly used models by 18.78%, 11.11% and 0.26% in terms of MAE, RMSE and  $R^2$ , respectively. For the 60-minute forecasting task, although the prediction accuracy of all models decreases, CESF-Net still achieves improvements of 2.98%, 2.74% and 0.61% in the respective metrics.

**Keywords:** channel-wise exponential moving average decomposition; FFT-Attention; time series segmentation; wind power prediction

## 0 引言

在“双碳”战略目标的引领下,我国能源结构正加快向

绿色、低碳方向转型<sup>[1-2]</sup>。其中,风电作为重要的新能源发电形式之一,近年来实现了快速发展。根据国家能源局公布的数据,截至 2024 年 8 月底,全国风电累计装机容量约

收稿日期:2025-07-01

\* 基金项目:内蒙古科技计划项目(2021GG0256)资助

4.7亿千瓦,同比增长19.9%<sup>[3]</sup>。尽管风电规模持续扩张,但其输出功率存在波动性、间歇性和不确定性等特点,给电网的安全稳定运行带来了一定挑战。开展风电功率的精确预测,有助于降低风电功率波动对电力系统调度与运行的影响,因此,提升风电功率预测精度对于保障电网运行的安全性与可靠性具有重要意义。

风电功率预测按照不同的预测模型可大致分为物理模型、统计模型、深度学习模型<sup>[4]</sup>。物理模型是指以数值天气预报(numerical weather prediction, NWP)数据为基础,结合物理规律对风电场功率输出进行计算的预测方法。这种方法需要精准的气象数据、风机状态以及周围的地形数据,且模型复杂度高,模型计算量大,在实际部署过程中,计算效率难以满足实时性要求<sup>[5]</sup>。统计模型通常利用历史风电数据与气象变量之间的相关性,通过数学建模进行功率预测<sup>[6]</sup>。常见的方法有自回归模型<sup>[7]</sup>、整合移动平均自回归模型<sup>[8]</sup>、支持向量机<sup>[9]</sup>等。该方法不受地理环境的限制,计算效率高、易实现,但难以建模强非线性关系,鲁棒性差,难以满足工程的实际需要。深度学习模型相比有前两种方法在功率预测方面表现出更强的特征学习与数据挖掘能力。常见的有卷积神经网络(convolutional neural network, CNN)<sup>[10]</sup>、长短期记忆网络(long short-term memory, LSTM)<sup>[11]</sup>等。

近年来,随着深度学习的快速发展,在风电功率预测领域广大学者也取得了显著的效果。文献[12]采用改进鲸鱼优化算法优化双向长短期记忆网络(bidirectional long short-term memory, BiLSTM)超参数的方法构建光伏功率预测模型,提升了双向长短期记忆网络对光伏功率数据的挖掘能力,提升了预测的效果。文献[13]针对短期风电功率预测中存在的特征提取尺度单一问题,提出一种基于多尺度特征提取的Transformer预测模型。该模型构建多尺度特征提取网络,充分挖掘风电功率序列内部的时序特征。文献[14]将BiLSTM、注意力机制与优化算法相结合,构建了短期风电功率预测模型,使用优化算法寻找BiLSTM的最优超参数,并通过注意力机制对BiLSTM的输入特征分配不同权重,强化关键特征,实现风电功率的高精度超短期预测。文献[15]将特征选择方法与CNN-BiLSTM模型相结合,构建短期风电功率预测模型。通过筛选高相关特征并加权构建多维输入数据集,提升了模型对时序特征的提取能力。然而,上述方法主要基于原始需求进行建模,未能充分考虑风电功率数据本身具有强波动性、易受多种外部因素干扰等特点,直接构建单一预测模型往往会导致较大的预测误差。

为有效应对上述问题,研究人员提出了基于组合模型的风电功率预测策略,旨在融合多种模型的互补优势,深入挖掘数据特征,从而提升整体预测的准确性。组合方法结合深度学习与序列分解的优势,将时间序列分解为趋势分量与季节分量,分别针对趋势分量与季节分量的特点进行

建模,以更好的挖掘复杂时间序列的特征,提高预测精度。文献[16]提出了将趋势季节局部加权回归分解与Informer模型相结合的方法,将光伏功率序列分解为周期项与趋势项,有效提升了中短期光伏功率预测的精度。文献[17]提出了一种改进完全自适应噪声集合经验模态分解与样本熵结合的风电功率分解的方法,将风电功率数据分解重构为高频分量与低频分量,应对风电数据强波动与随机性的问题。文献[18]提出一种基于经验模态分解的超短期风电功率预测方法,通过经验模态分解将风电序列划分为高频与低频分量,分别对不同频段的特征进行建模,以提高预测的精度。文献[19]提出一种重构式二次分解集成框架用于风电功率预测,通过自适应噪声的完全集合经验模态分解对风电序列进行分解并结合频域重构与奇异谱分析二次分解,提取不同频段的局部与全局特征,达到提升预测精度的效果。

上述方法均只对功率数据进行分解探究,未能在多变量数据建模过程中去探究不同特征变量之间表现出的异步变化趋势,且传统的统一处理方式易忽略通道间差异,限制了模型对关键趋势信息的捕捉能力。

为挖掘风电机组数据的深层耦合特征,提高预测精度,本文提出了一种基于CW-EMA分解与组合神经网络的风电功率预测模型(channel-wise EMA decomposition and segment-based fusion network, CESF-Net)。引入通道独立指数滑动平均(channel-wise exponential moving average, CW-EMA)分解方法对时间序列进行趋势分量与季节分量的分解,挖掘其长期性变化与周期性变化特征。采用时间序列分段将分量划分为固定长度片段,增强模型对局部信息的感知能力,并构建以门控循环单元与快速傅里叶变换注意力(fast Fourier transform attention, FFT-Attention)为基础的双流网络,分别用于捕捉趋势分量中的长期依赖性特征与季节分量中的周期性特征,以提高风电功率预测的精度。

## 1 数据的清洗与插补

### 1.1 异常数据识别

基于密度的空间聚类算法(density-based spatial clustering of applications with noise, DBSCAN)广泛应用于异常检测与聚类分析任务中<sup>[20]</sup>。本文采用DBSCAN算法对风电数据集中存在的异常值进行识别与剔除。该算法的显著优势在于无需预设聚类簇的数量,而是依赖样本点在数据空间中的密度分布特征,自适应地识别聚类结构。

DBSCAN算法的思想是依据数据点之间的密度连接关系,将簇定义为由密度相连的样本点所组成的最大集合,即同一簇内的各样本之间具有较强的密度连通性。该算法通过两个关键参数实现对局部密度的量化,其中邻域半径Epsilon用于确定样本点的局部邻域范围,而最小邻域内点数MinPts则用于判断样本点所在区域的密度是否满足形

成簇的要求。

根据邻域半径和最小邻域内点数这两个参数, DBSCAN 算法将数据样本点划分为核心点、边界点和噪声点这 3 类。其中, 核心点指邻域内包含的样本点数量达到或超过最小邻域内点数的点, 边界点指邻域内点数不足最小邻域内点数但位于核心点邻域范围内的点, 噪声点则是既不满足核心点条件, 也不在任何核心点邻域内的孤立点。算法的聚类过程从任意一个未访问的核心点开始, 以该点邻域内的密度结构为基础构建初始聚类簇, 再通过密度可达性原则递归地将所有密度可达的样本点纳入该簇, 从而最终形成完整的密度聚类结构。如图 1 所示, DBSCAN 能够有效识别具有任意形状的聚类结构, 并对孤立点进行异常标注, 适用于风电功率等复杂时序数据中的异常值识别任务。

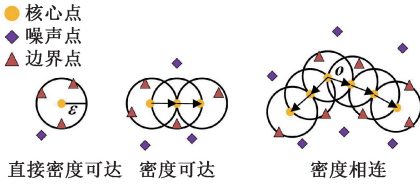


图 1 DBSCAN 算法原理

Fig. 1 Principle of the DBSCAN algorithm

### 1.2 缺失数据的插补

线性插值法是一种常用的缺失值填补方法, 适用于风电功率等时间序列数据的插补任务<sup>[21]</sup>。本文针对因传感器故障或通信中断导致的数据缺失, 采用线性插值法进行处理, 以保证数据的连续性与建模精度。计算公式为:

$$y = y_0 + \frac{(x - x_0)}{(x_1 - x_0)} \cdot (y_1 - y_0) \quad (1)$$

式中:  $x_0$  表示缺失点前一时间,  $x_1$  表示缺失点后一时间,  $y_0$  和  $y_1$  分别为对应时刻的已知数值,  $x$  为缺失值对应的时间点,  $y$  为插补后的数值。

## 2 基于 CESF-Net 的多步风电功率预测模型

### 2.1 通道独立指数滑动平均分解

指数滑动平均(exponential moving average, EMA)算法是一种指数平滑方法, 它为最近的数据点分配更大的权重, 同时平滑旧数据<sup>[22]</sup>。该方法的平滑因子  $\alpha$  为一个固定的值, 且全通道共享一个相同的平滑因子  $\alpha$ 。在多变量风电功率序列中, 各变量(如风速、风向、气压等)具有异构的时间动态特征, 直接对所有变量应用统一的指数平滑因子会削弱模型对个别通道趋势特性的感知能力。为提升趋势建模的通道适应性与建模表达能力, 本文引入了通道独立指数滑动平均分解方法, 即为每个通道  $c \in \{1, 2, 3, \dots, C\}$  分配一个独立的指数平滑因子  $\alpha^{(c)} \in (0, 1)$ , 并在训练过程中进行自动优化。

设输入长度为  $L + 1$  的多变量时间序列  $\mathbf{X} = \{x_0,$

$x_1, \dots, x_L\}, x_t \in \mathbf{R}^C$ 。对第  $c$  个通道, 趋势状态初始化为:

$$\mathbf{S}_0^{(c)} = \mathbf{x}_0^{(c)} \quad (2)$$

在  $t \geq 1$  时刻的趋势分量递推定义为:

$$\mathbf{S}_t^{(c)} = \alpha^{(c)} \mathbf{x}_t^{(c)} + (1 - \alpha^{(c)}) \mathbf{S}_{t-1}^{(c)} \quad (3)$$

将  $\mathbf{S}_t^{(c)}$  展开为关于所有历史输入的函数:

$$\mathbf{S}_t^{(c)} = \alpha^{(c)} \mathbf{x}_t^{(c)} + \dots + (1 - \alpha^{(c)})^t \mathbf{x}_0^{(c)} \quad (4)$$

定义权重系数为:

$$\omega_k^{(c)} = \alpha^{(c)} (1 - \alpha^{(c)})^k \quad (5)$$

则有:

$$\mathbf{S}_t^{(c)} = \sum_{k=0}^t \omega_k^{(c)} \mathbf{x}_{t-k}^{(c)} + (1 - \alpha^{(c)})^t \mathbf{x}_0^{(c)} \quad (6)$$

式中:  $\alpha^{(c)}$  为趋势分量的平滑因子, 范围为  $0 < \alpha^{(c)} < 1$ ;  $\mathbf{S}_t^{(c)}$  为  $t$  时刻的趋势分量。

由式(5)与(6)可知, CW-EMA 具有自适应记忆长度的能力, 即随着  $\alpha^{(c)}$  取值不同, 趋势建模可对不同通道施加差异化的平滑强度。最终, 原始多变量序列  $\mathbf{X} \in \mathbf{R}^{(L+1) \times C}$  被分解为:

$$\begin{cases} \mathbf{X}_T = CW-EMA(\mathbf{X}) \\ \mathbf{X}_S = \mathbf{X} - \mathbf{X}_T \end{cases} \quad (7)$$

式中: CW-EMA( $\cdot$ ) 为通道独立指数滑动平均;  $\mathbf{X}_T$  与  $\mathbf{X}_S$  分别为时间序列  $\mathbf{X}$  的趋势分量与季节分量。

平滑因子  $\alpha^{(c)}$  决定第  $c$  个通道的当前输入与历史趋势的权衡。  $\alpha^{(c)}$  越大,  $\mathbf{S}_{t-1}^{(c)}$  的权重越低、记忆更短, 分解后趋势更敏感、平滑性降低, 部分短期波动被吸收入趋势分量中;  $\alpha^{(c)}$  越小, 对  $\mathbf{S}_{t-1}^{(c)}$  依赖增强、记忆更长, 趋势更平滑, 短期扰动更多地保留在季节分量中。

为实现趋势提取的自适应调节, 本文将每个通道的平滑因子  $\alpha^{(c)}$  作为可学习参数, 利用反向传播机制与预测损失函数进行优化。

设预测误差损失为  $\mathcal{L}$ , 最终预测输出为  $\hat{y}$ , 则对任意通道的平滑因子  $\alpha^{(c)}$  的梯度可按链式法则展开为:

$$\frac{\partial \mathcal{L}}{\partial \alpha^{(c)}} = \frac{\partial \mathcal{L}}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial \mathbf{S}_t^{(c)}} \cdot \frac{\partial \mathbf{S}_t^{(c)}}{\partial \alpha^{(c)}} \quad (8)$$

其中,  $\frac{\partial \mathbf{S}_t^{(c)}}{\partial \alpha^{(c)}}$  具有的关系如式(9)所示。

$$\begin{cases} \frac{\partial \mathbf{S}_0^{(c)}}{\partial \alpha^{(c)}} = 0, & t = 0 \\ \frac{\partial \mathbf{S}_t^{(c)}}{\partial \alpha^{(c)}} = \mathbf{x}_t^{(c)} - \mathbf{S}_{t-1}^{(c)} + (1 - \alpha^{(c)}) \cdot \frac{\partial \mathbf{S}_{t-1}^{(c)}}{\partial \alpha^{(c)}}, & t \geq 1 \end{cases} \quad (9)$$

式(9)表示  $\mathbf{S}_t^{(c)}$  每一时刻对于平滑因子  $\alpha^{(c)}$  导数是由当前输入与历史趋势值的差异, 以及前一时刻的导数共同决定, 具有时间记忆性与权重衰减特性。参数更新则遵循标准梯度下降规则:

$$\alpha^{(c)} \leftarrow \alpha^{(c)} - \eta \cdot \frac{\partial \mathcal{L}}{\partial \alpha^{(c)}} \quad (10)$$

式中:  $\eta$  为学习率。为保证数值稳定性并避免指数退化, 所有  $\alpha^{(c)}$  被约束在  $[\epsilon, 1 - \epsilon]$  范围内。

在每次迭代过程中,模型根据预测误差损失函数  $L$  对平滑因子  $\alpha^{(c)}$  求导,若当前  $\alpha^{(c)}$  取值导致预测误差增大,则沿着负梯度方向调整,使  $\alpha^{(c)}$  向减小损失的方向更新。具体而言,当  $\alpha^{(c)}$  偏大时,模型对短期波动反应过强,梯度趋向正值,  $\alpha^{(c)}$  会被减小;当  $\alpha^{(c)}$  偏小时,趋势过于平滑,模型难以响应新的输入变化,梯度趋向负值,  $\alpha^{(c)}$  则被增大。通过不断迭代优化,  $\alpha^{(c)}$  在  $[\epsilon, 1-\epsilon]$  范围内逐步收敛到最优值,使各通道对长期趋势与短期波动的响应实现自适应调节。

### 2.2 时间序列分段

在时间序列预测任务当中,单一时间步往往所携带有限的语义信息,若直接输入模型,可能会遗漏部分关键的动态特征。为此,将原始序列划分为若干连续或部分重叠的片段,能够在局部时间窗口中增强模型的特征表达能力,同时提升其对周期性变化、突变行为等时间特征的识别能力<sup>[23]</sup>。此外,这种分段处理策略还在一定程度上降低了长序列输入所带来的维度负担,从而有效减少计算开销,提升模型的预测效率与响应速度。该方法对于增强模型在长期预测场景中的稳健性与泛化表现具有积极作用。因此本文采用时间序列分段的方法来处理预测模型的输入特征,帮助预测模型能够更好的学习到时间序列的特征,更好的完成所需要的预测任务,分段过程如图2所示。

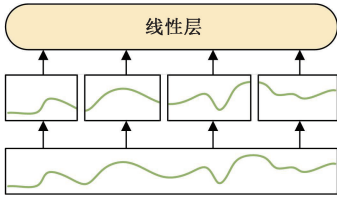


图2 时间序列分段过程

Fig.2 Time series segmentation process

对于时间序列  $\mathbf{X} = \{x_0, x_1, \dots, x_L\} \in \mathbf{R}^{1 \times (L+1)}$ , 其中  $L+1$  表示时间序列的长度,根据时间步长划分为  $N$  个不重叠的片段,每个片段的长度为  $P = L/N$ 。经过分段后得到第  $i$  个片段序列为:  $\mathbf{X}_p^{(i)} \in \mathbf{R}^{1 \times P}, i = 1, \dots, N$ 。进而得到经过分段后的时间序列为:  $\mathbf{X}' \in \mathbf{R}^{N \times P}$ 。最后,通过线性层来线性映射得到预测模型的输入为:  $\mathbf{X}'' \in \mathbf{R}^{N \times T}$ 。其中  $T$  表示每个片段线性映射后的嵌入维度。

### 2.3 FFT-Attention 注意力机制

在时间序列建模任务中,传统注意力机制往往侧重于时间域中各位置之间的全局依赖性,然而其对序列中的周期性特征挖掘能力较为有限,特别是在序列具有明显振荡结构或频率信息主导的场景下,时域注意力容易忽略隐藏在序列变化背后的频率规律性特征。若不加以处理,可能会使模型在建模复杂周期信息时面临表达受限的问题。为此,本文引入快速傅里叶变换对注意力机制进行改进<sup>[24]</sup>。如图3所示,将注意力计算过程从时间域迁移至频域。首先利用快速傅里叶变换(FFT)将输入的查询(Query)、键(Key)和值(Value)映射至频域表示,在频域中执行缩放点

积注意力操作后再通过傅里叶变换(iFFT)恢复到时间域表示。此过程能够捕捉时间序列中各频率分量的响应强度,从而提升模型对周期性特征的感知与建模能力。

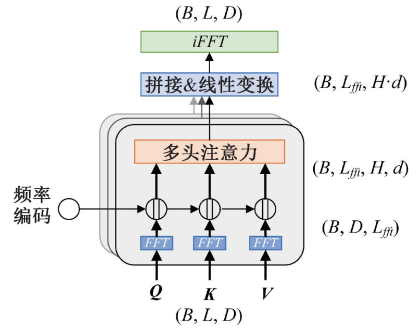


图3 快速傅里叶变换注意力机制

Fig.3 FFT-Attention mechanism

图3中,  $B$  为 batch size;  $L$  为序列长度;  $D$  为模型的输入维度;  $L_{fft}$  为对长度为  $L$  的序列做 FFT 后的频域长度;  $H$  为多头注意力中的头数;  $d$  为每头的  $Q, K, V$  维度。

FFT-Attention 注意力机制计算过程可以表示为:

$$\begin{cases} \tilde{Q} = FFT(Q) \\ \tilde{K} = FFT(K) \\ \tilde{V} = FFT(V) \end{cases} \quad (11)$$

$$FFT-Attention(\tilde{Q}, \tilde{K}, \tilde{V}) = iFFT\left(\text{softmax}\left(\frac{\tilde{Q}\tilde{K}^H}{\sqrt{d}}\right)\tilde{V}\right) \quad (12)$$

$$head_i = FFT-Attention(QW_i^Q, KW_i^K, VW_i^V) \quad (13)$$

$$MH^{FFT}(Q, K, V) = iFFT(Concat(head_1, \dots, head_h)W^O) \quad (14)$$

式中:  $Q, K, V$  表示查询、键与值;  $\tilde{Q}, \tilde{K}, \tilde{V}$  表示频域下的复数张量,包含频谱的实部与虚部信息;  $\text{softmax}(\cdot)$  作用于频率维度,计算注意力权重;  $iFFT(\cdot)$  将频域输出变换回时间域;  $Concat(\cdot)$  表示级联函数;  $Attention(\cdot)$  表示自注意力机制。

### 2.4 门控循环单元

在趋势分量建模任务中,捕捉时间序列中缓慢演化的结构性信息(如升降趋势、长期惯性)是预测模型的关键能力之一。为此,本文引入门控循环单元(gated recurrent unit, GRU)作为趋势建模模块,用于提取并记忆时间序列中的长期动态特征。

GRU 是一种循环神经网络(recurrent neural network, RNN)的改进结构<sup>[25]</sup>。该模型在保留传统 RNN 循环连接形式的基础上,引入门控机制以实现信息传递与遗忘过程的动态调节,从而提升序列建模能力。GRU 相较于标准 RNN 结构,能够有效缓解梯度消失问题,并增强对长期依赖特征的捕捉能力。与长短期记忆网络(LSTM)相比,

GRU 通过将输入门与遗忘门合并为单一的更新门,并引入重置门以控制历史状态对当前计算的影响,从而在保证建模性能的同时显著简化了结构并降低了计算复杂度。其内部主要由 3 个关键组成部分构成:更新门、重置门与候选隐藏状态,结构如图 4 所示。

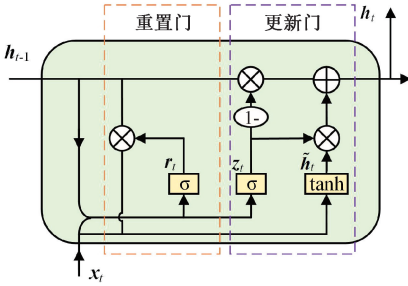


图 4 GRU 的内部结构  
Fig. 4 Internal structure of GRU

GRU 单元内部更新过程可以表示为:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (15)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad (16)$$

$$\tilde{h}_t = \tanh(W_h \cdot [r_t \odot h_{t-1}, x_t] + b_h) \quad (17)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (18)$$

式中:  $W_z$ 、 $W_r$ 、 $W_h$  为权重矩阵;  $b_z$ 、 $b_r$ 、 $b_h$  为偏置项;  $x_t$  表示  $t$  时刻输入向量;  $h_t$  表示  $t$  时刻隐藏状态;  $\tilde{h}_t$  表示  $t$  时刻候选隐藏状态;  $z_t$  表示更新门的输出;  $r_t$  表示重置门的输出;  $\odot$  是哈达玛积。

### 2.5 风电功率预测

针对风电功率序列中存在的非线性强、周期性复杂及短时波动剧烈等特点,以 CESF-Net 模型构建超短期风电功率预测模型。CESF-Net 风电功率预测模型如图 5 所示。

CESF-Net 风电功率预测的过程包括:

1) 数据预处理: 采用 DBSCAN 算法识别风速-功率关系中的离群点,通过线性插值法对缺失值进行填补;随后对特征进行归一化处理,确保不同变量在模型训练中的数值一致性;处理后将数据集按照 8:1:1 划分为训练集,验证集与测试集。

2) 序列分解: 采用 CW-EMA 分解方法对预处理后的序列进行趋势分量与季节分量的分解。其中,趋势分量反映长期变化趋势,季节分量反映周期性波动特征。

3) 时间序列分段: 分别对趋势分量与季节分量进行时间序列分段处理,将连续时间序列划分为多个固定长度的子序列片段,以增强局部时间上下文的建模能力。

4) 特征提取与建模: 将具有周期性特征的季节分量送入基于 FFT-Attention 的非线性预测模块中进行处理,用于提取周期性成分中的非线性时序特征;同时将缓慢变化的趋势分量送入基于 GRU 的线性建模模块中进行处理,用于捕捉趋势成分中的长期依赖性与动态变化规律。

5) 特征融合与预测输出: 分别由趋势流与季节流提取的线性与非线性特征进行拼接融合,随后输入至全连接层进行映射,输出最终的超短期风电功率预测值。

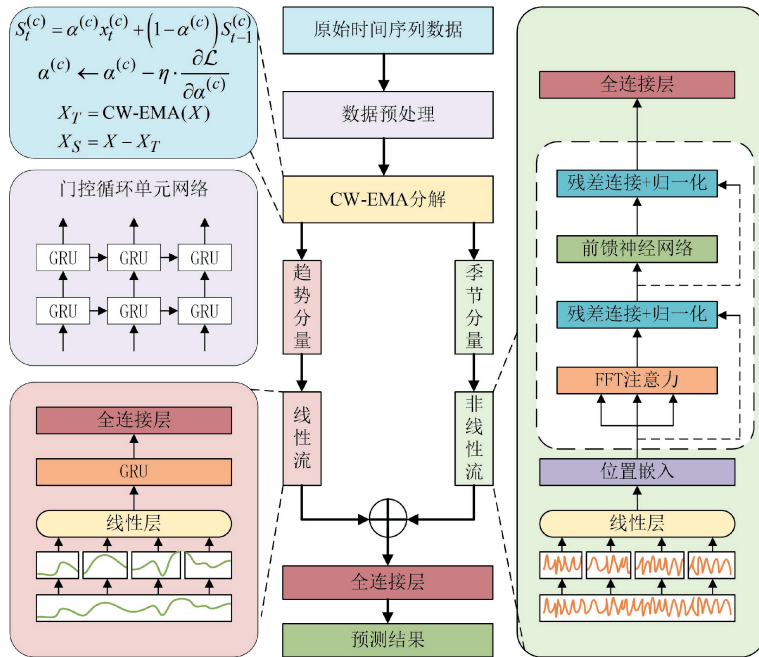


图 5 CESF-Net 的整体结构  
Fig. 5 The structure of CESF-Net

### 3 实验结果与分析

#### 3.1 实验数据

为验证 CESF-Net 模型的有效性,本文选取风电场实测数据。数据的时间跨度为 2019 年 1 月 1 日~2019 年 12 月 31 日,采样频率为 15 min/次,共有 35 040 条数据。该数据集包含测风塔 10、30、50、70 m 轮毂高度的风速与风向,温度,气压,湿度,发电功率。

#### 3.2 实验设置

实验设备采用的处理器为 Intel(R) Core(TM) i5-12600KF,显卡为 NVIDIA GeForce RTX 4060Ti,操作系统为 Ubuntu 20.04。本文利用 PyTorch 框架构建模型,CESF-Net 模型中的 GRU 共 2 层,各层神经元数量为 64,FFT-Attention 头数为 4,Transformer 编码器层数为 2 层。对于本文中所有的对比模型各层神经元数量均为 64,层数为 2。本文所涉及的所有模型的学习率为 0.000 1, batch size 为 32,选择 MSE 作为损失函数,采用 Adam 优化器,模型历史输入长度为 48,预测长度为 4。

#### 3.3 评价指标

为了验证 CESF-Net 模型的预测性能,本文采用常见的平均绝对误差(mean absolute error, MAE)、均方根误差(root mean square error, RMSE)以及判定系数(coefficient of determination,  $R^2$ )作为评价指标,其计算公式为:

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (19)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (20)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (21)$$

式中:  $n$  为样本数;  $\hat{y}_i$  表示  $t$  时刻风电功率预测值;  $y_i$  表示  $t$  时刻风电功率真实值;  $\bar{y}$  表示风电功率真实值的均值。

#### 3.4 数据预处理

本文采用 Spearman 相关系数对历史风电功率与各气象变量之间的关系进行定量分析,相关程度计算如表 1 所示。结果表明,风速与风向在不同高度之间表现出高度相关性,变量间存在冗余信息。在众多风速变量中,轮毂高度风速与风电功率的相关系数为 0.91,相关性最强,能够更准确地反映风资源强度对机组出力的驱动作用;而在风向变量中,轮毂高度风向也具有相对更高的相关性,反映其在来流风向对功率输出的影响方面具有代表性。此外,气压虽与风电功率的相关性中等,但作为表征宏观气象环境的重要特征,有助于提高模型的稳健性与泛化能力。相比之下,温度与湿度与功率的相关系数均小于  $\pm 0.15$ ,相关性较弱,说明其对风电输出的解释能力有限,且可能增

加模型训练的复杂度,故未纳入建模变量。综上所述,本文最终选取轮毂高度风速、轮毂高度风向、气压与风电功率作为模型的输入特征,以兼顾变量代表性、相关性强度与信息非冗余性。

表 1 功率与各气象因素之间的 Spearman 相关系数

Table 1 Spearman correlation coefficients between power and meteorological factors

相关系数	值	相关系数	值
$S_{10}$	0.86	$D_{50}$	-0.42
$S_{30}$	0.89	$D_{70}$	-0.43
$S_{10}$	0.91	$D_l$	-0.43
$S_{50}$	0.91	温度	-0.02
$S_l$	0.91	气压	-0.15
$D_{10}$	-0.41	湿度	-0.01
$D_{30}$	-0.40	功率	1.00

表 1 中  $S_{10} \sim S_{70}$  分别表示测风塔 10~70 m 风速与功率之间的 Spearman 相关系数;  $S_l$  表示轮毂高度风速与功率之间的 Spearman 相关系数;  $D_{10} \sim D_{70}$  分别表示测风塔 10~70 m 风向与功率之间的 Spearman 相关系数;  $D_l$  表示轮毂高度风向与功率之间的 Spearman 相关系数。

对本文选取的轮毂高度风速与功率数据使用 DBSCAN 算法进行异常数据的清洗,风电场的异常数据为图 6 中的 I~III 号簇数据。I 号簇所对应的数据点主要源于风速测量设备的异常。风速仪表发生故障时,采集到的风速信息可能失真,从而引起样本在风速-功率特性曲线中偏离其应有位置。II 号簇反映的是电网调度限制下的弃风限电现象。当风电机组具备较强发电能力但超出电网可接纳范围时,为保障电力系统的安全与稳定,需对风电出力进行强制削减,导致实际功率低于理论输出。III 号簇则包含风电功率表现出明显无规律波动的数据点,该类异常多数由极端气象条件、传感器失准或设备故障引起,造成风速与功率关系的紊乱偏离。上述多类型异常值若不加处理,可能严重干扰数据规律的学习过程,降低风电功率预测模型的拟合能力。因此,进行系统化的异常识别与剔除处理。

将清洗出的异常数据作为待填补的缺失值,并采用线性插值法进行插值,插值效果如图 7 所示,由图可见清洗插补后的数据大致符合风速-功率曲线图。

#### 3.5 时间序列分段长度 Patch length 的选择

为探究分段长度对模型性能的影响,本文构建了关于不同分段长度对模型性能影响的对比实验。实验将 CESF-Net 模型的时间序列分段长度分别设为 8、16、24 与 48 这 4 组实验。实验结果如图 8 所示。

如图 8 所示,不同的分段长度对模型是有影响的,当分段长度为 48 时(即输入序列不采用时间序列分段)在各

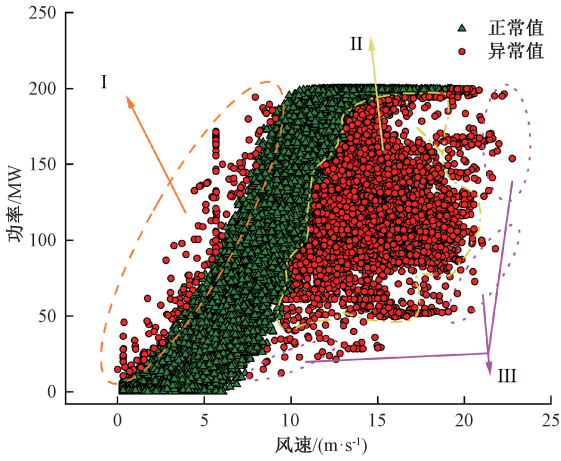


图 6 风速-功率散点图

Fig. 6 Wind speed-power scatter plot

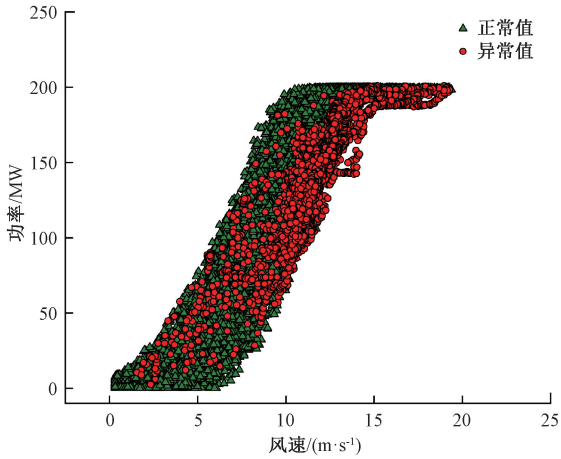


图 7 缺失值填补结果

Fig. 7 Results of missing value imputation

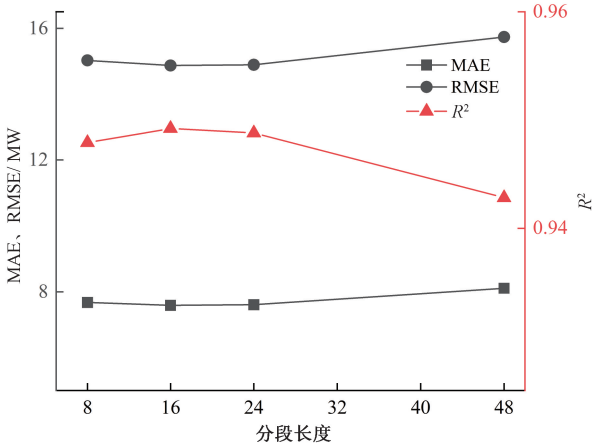


图 8 不同分段长度性能对比图

Fig. 8 Performance comparison under different patch lengths

项指标上均为最差,说明时间序列分段对模型性能是有益的。分段长度分别为 8、16、24 时,在分段长度为 16 时模型

的综合性能最优,适度的分段长度能够在时域上提供充分的上下文信息,兼顾局部动态与全局趋势,实现稳定而高效的特征建模。若分段长度过短,则信息不足、噪声敏感,模型性能下降;而分段长度过长则导致非平稳性增强,注意力难以聚焦关键变化,模型表现下降。本文后续实验 CESF-Net 模型的分段长度均设置为 16。

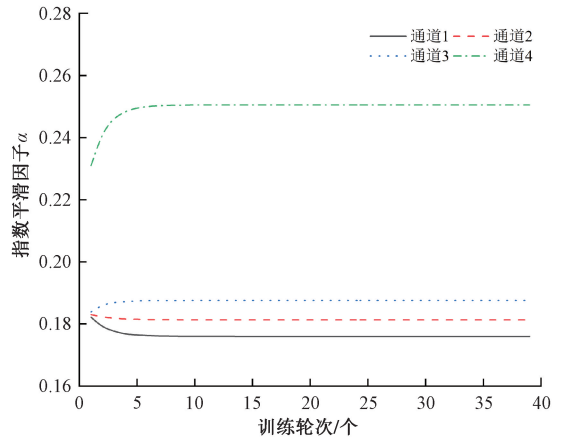
### 3.6 CW-EMA 有效性验证

为验证通道独立指数滑动平均(CW-EMA)分解方法在趋势建模中的有效性,本文构建了包含不同平滑因子设置的对比实验。具体包括以下两种设置:1)统一平滑因子 EMA(N1),所有通道共享一个固定的全局平滑因子  $\alpha$ ;2)通道独立 EMA(CESF-Net),为每个通道引入独立可学习的平滑因子  $\alpha^{(c)}$ ;初始值设定一致。结果如表 2 所示,在所有设定的  $\alpha$  初始值(0.2~0.8)下,CW-EMA 模型均在 MAE、RMSE 和  $R^2$  这 3 项指标上优于其对应的统一平滑因子版本。

表 2 CESF-Net 与 N1 在不同  $\alpha$  初始值下的功率预测结果  
Table 2 Wind power forecasting results of CESF-Net and N1 under different initial values of  $\alpha$

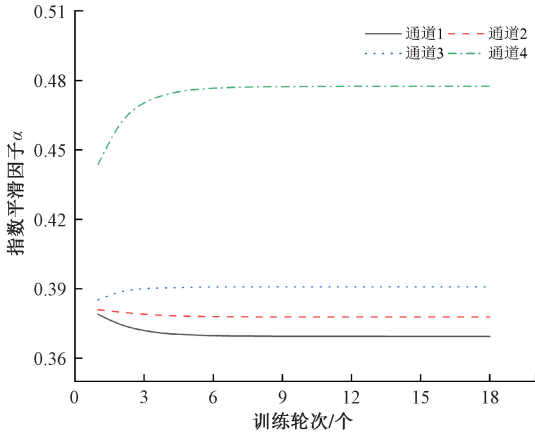
模型	$\alpha$ 初始值	MAE/MW	RMSE/MW	$R^2$
CESF-Net	0.2	7.721	15.106	0.9473
N1		7.758	15.175	0.9465
CESF-Net	0.4	7.654	14.960	0.9483
N1		7.686	15.013	0.9470
CESF-Net	0.6	7.603	14.894	0.9488
N1		7.638	14.942	0.9484
CESF-Net	0.8	7.589	14.867	0.9492
N1		7.605	14.885	0.9488

在  $\alpha$  的初始值设定为 0.2~0.8 的不同情形下,CW-EMA 各通道  $\alpha$  参数的动态演化规律如图 9 所示。实验中,通道 1 表示轮毂高度风速,通道 2 表示轮毂高度风向,通道 3 为气压,通道 4 为历史风电功率。可以观察到,尽管

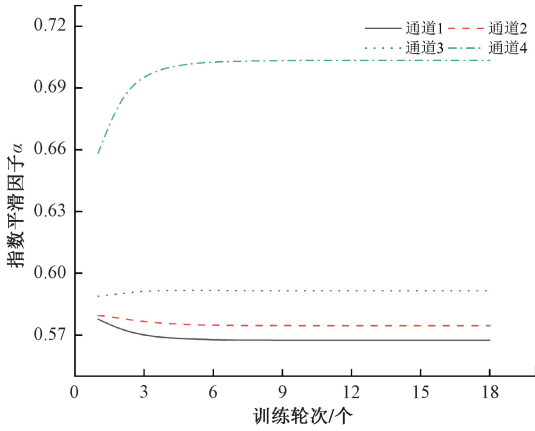


(a) 初始值  $\alpha=0.2$  时各通道  $\alpha$  变化

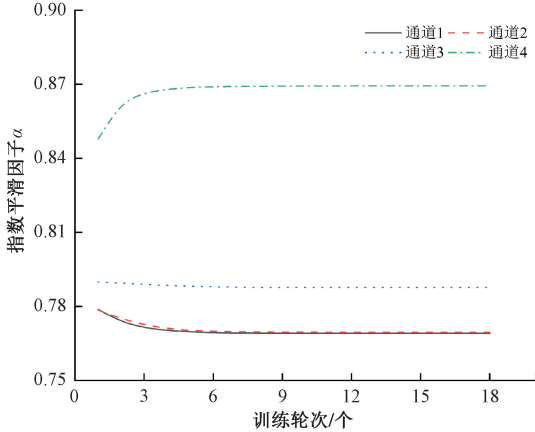
(a) Channel-wise  $\alpha$  variation with initial value  $\alpha = 0.2$



(b) 初始值 $\alpha=0.4$ 时各通道 $\alpha$ 变化  
(b) Channel-wise  $\alpha$  variation with initial value  $\alpha = 0.4$



(c) 初始值 $\alpha=0.6$ 时各通道 $\alpha$ 变化  
(c) Channel-wise  $\alpha$  variation with initial value  $\alpha = 0.6$



(d) 初始值 $\alpha=0.8$ 时各通道 $\alpha$ 变化  
(d) Channel-wise  $\alpha$  variation with initial value  $\alpha = 0.8$

图 9 不同初始值时  $\alpha$  值的动态演变规律

Fig. 9 Dynamic evolution of  $\alpha$  values under different initializations

各通道的  $\alpha$  初始值设定相同,但在训练开始后的第 1 个 epoch 后,各通道  $\alpha$  参数即发生明显调整,并最终趋于稳定。

其中,通道 1 和 2 对应的  $\alpha$  值在训练过程中逐渐下降并最终稳定,表明模型对于风速与风向变量更倾向于强调对长期趋势的平滑建模;而通道 3 与 4 的  $\alpha$  值则随训练迭

代逐步提升并最终稳定,体现出模型对于气压与历史风电功率更倾向于强调当前输入的重要性。

上述结果验证了 CW-EMA 所具备的通道自适应趋势建模能力,模型能够依据不同变量的动态特征自适应调整平滑强度,从而增强了建模过程的自适应能力。

### 3.7 消融实验

为验证模型中各部分对风电功率预测结果的影响,本文设计了 5 组消融对比实验。其中对比模型为:去掉 CW-EMA 分解部分(M1)、去掉快速傅里叶变换部分(M2)、去掉时间序列分段部分(M3)、仅线性流部分(M4)、仅非线性流部分(M5),以评估各部分对整体模型的贡献,实验结果如表 3 所示。

表 3 CESF-Net 消融实验结果

Table 3 Ablation study results of CESF-Net

模型	MAE/MW	RMSE/MW	$R^2$
CESF-Net	7.589	14.867	0.949 2
M1	7.676	15.021	0.947 5
M2	7.629	14.837	0.948 1
M3	8.108	15.734	0.942 8
M4	8.959	18.373	0.922 1
M5	7.921	15.217	0.946 6

模型 CESF-Net 在 MAE、RMSE 和  $R^2$  方面分别为 7.589 MW、14.867 MW 和 0.949 2,整体表现最佳。

当移除 CW-EMA 分解模块(M1)时,MAE 上升至 7.676 MW,  $R^2$  降至 0.947 5,表明该模块对特征解耦与长期趋势提取具有正向作用;去除快速傅里叶变换(M2)后虽然 RMSE 的性能有所提升,但 MAE 与  $R^2$  的性能都在下降,说明快速傅里叶变换在捕捉局部周期性或高频成分方面仍具有一定作用。去除时间序列分段模块(M3)后,MAE 升高至 8.108 MW,模型性能下降更加明显,说明时间序列分段有助于局部时序模式的捕捉。

在仅保留线性流模块(M4)时,MAE 和 RMSE 上升至 8.959 MW 和 18.373 MW,  $R^2$  降至 0.922 1,表明缺失非线性建模能力会严重影响预测精度;而仅保留非线性流(M5)时,尽管 MAE(7.921 MW)略优于 M4,但整体  $R^2$  仍降至 0.946 6,说明线性与非线性模块协同建模是提升整体性能的关键。

### 3.8 预测对比实验

为验证 CESF-Net 模型的预测性能,将本文模型与常见的预测方法 CNN<sup>[26]</sup>、LSTM<sup>[27]</sup>、GRU<sup>[28]</sup>、CNN-LSTM<sup>[29]</sup>、Transformer<sup>[30]</sup> 进行多步预测实验对比。预测步长分别为 1 步,2 步,3 步,4 步。实验结果如表 4~7、图 10 所示。

表 4~7 展示了各模型在 15、30、45、60 min 的风电功率预测结果。在 15 min 的预测步长下,CESF-Net 的模型的 MAE、RMSE、 $R^2$  分别为 3.489 MW、6.463 MW、0.990 4,

表 4 单步预测结果

Table 4 One-step forecasting results

时间	方法	MAE/ MW	RMSE/ MW	$R^2$
15 min	CESF-Net	3.489	6.463	0.990 4
	CNN	9.241	16.171	0.939 7
	LSTM	6.487	11.272	0.970 7
	GRU	4.424	7.970	0.985 3
	CNN-LSTM	5.376	9.011	0.981 3
	Transformer	4.296	7.271	0.987 8

表 5 2步预测结果

Table 5 Two-step forecasting results

时间	方法	MAE/ MW	RMSE/ MW	$R^2$
30 min	CESF-Net	6.498	12.007	0.966 7
	CNN	11.706	20.592	0.902 2
	LSTM	9.076	16.125	0.940 0
	GRU	7.562	13.753	0.956 4
	CNN-LSTM	8.069	14.250	0.953 1
	Transformer	6.964	12.555	0.963 6

表 6 3步预测结果

Table 6 Three-step forecasting results

时间	方法	MAE/ MW	RMSE/ MW	$R^2$
45 min	CESF-Net	9.056	16.612	0.936 3
	CNN	13.679	24.069	0.866 3
	LSTM	11.354	20.482	0.903 2
	GRU	9.864	17.935	0.925 8
	CNN-LSTM	10.286	18.328	0.922 4
	Transformer	9.452	17.168	0.932 0

表 7 4步预测结果

Table 7 Four-step forecasting results

时间	方法	MAE/ MW	RMSE/ MW	$R^2$
60 min	CESF-Net	11.312	20.665	0.901 5
	CNN	15.472	16.950	0.832 4
	LSTM	13.669	24.560	0.860 8
	GRU	12.132	22.044	0.887 8
	CNN-LSTM	12.497	22.384	0.884 4
	Transformer	11.660	21.220	0.896 1

表现出最高的预测性能。随着时间步长的变化,各模型的预测精度都有所变化,但 CESF-Net 模型依然保持着最优的精度。在 60 min 的预测步长下,CESF-Net 模型的各项

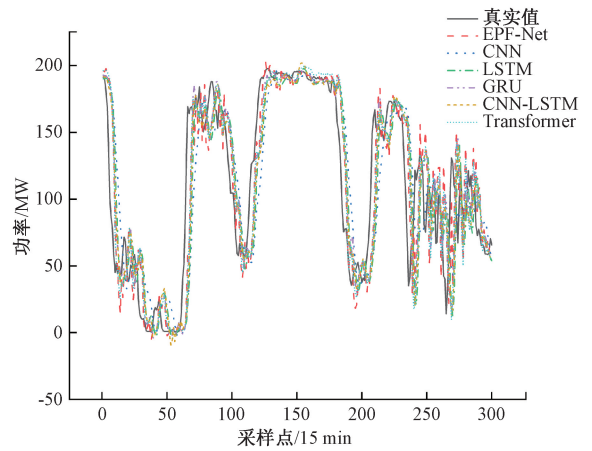


图 10 不同模型在未来 60 min 的风电功率预测结果

Fig. 10 Wind power forecasting results of different models for the next 60 minutes

评价指标比常用模型至少提高了 2.98%、2.74%、0.61%。

综合来看,本文模型能较好的适应风电功率的变化规律,能够拥有最佳的预测结果。结果表明,充分利用时间序列的趋势分量与季节分量并针对其各自的特点建模可以实现更准确的风电功率预测。

## 4 结 论

本文针对风电功率具有较强波动性及多重非线性特征的问题,提出了一种新型的风电功率预测模型—CESF-Net。通过一系列算例验证,主要得出以下结论:

本文的 CW-EMA 分解方法,为每个输入通道分配一个可学习的平滑因子,并借助反向传播机制及预测损失函数进行联合优化,从而实现对时间序列中趋势分量与季节分量的有效分离,降低数据复杂度。与传统的 EMA 分解方法相比,CW-EMA 在 MAE、RMSE 与  $R^2$  这 3 项评价指标上均表现出更优的预测性能,验证了其有效性与适应性。

在功率预测方面,本文分别构建了基于 GRU 的线性流预测模块和基于 FFT-Attention 的非线性流预测模块,分别针对 CW-EMA 分解后的趋势分量和季节分量进行建模。线性与非线性预测模块的协同融合,有效提升了模型对风电功率复杂动态特征的表达力,从而增强了对超短期风电功率的预测精度。

此外,本研究在异常值识别方面采用的 DBSCAN 算法中,领域半径与领域内最小点数参数主要通过多轮实验确定,尚未建立基于理论推导的最优参数选择机制;而在 CW-EMA 分解过程中所涉及的平滑因子  $\alpha^{(i)}$  虽可在模型训练阶段进行调优,但仍需预设初始值,且无法直接获取其全局最优值。因此,未来的研究工作可引入参数优化算法,对上述关键参数进行系统化的最优值搜索,以进一步提升风电功率预测模型的准确性。

## 参考文献

- [1] 张智刚,康重庆.碳中和目标下构建新型电力系统的挑战与展望[J].中国电机工程学报,2022,42(8):2806-2819.  
ZHANG ZH G, KANG CH Q. Challenges and prospects for constructing the new-type power system towards a carbon neutrality future[J]. Proceedings of the CSEE, 2022, 42(8): 2806-2819.
- [2] LI ZH M, 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: 120282.
- [3] 国家能源局发布2024年1-8月份全国电力工业统计数据[J].电力勘测设计,2024(9):39.  
National Energy Administration releases the statistics of the national electric power industry from January to August 2024[J]. Electric Power Survey & Design, 2024(9): 39.
- [4] 孙荣富,张涛,和青,等.风电功率预测关键技术及应用综述[J].高压技术,2021,47(4):1129-1143.  
SUN R F, ZHANG T, HE Q, et al. Review on key technologies and applications in wind power forecasting[J]. High Voltage Engineering, 2021, 47(4): 1129-1143.
- [5] 冯俊磊,吕卫东,段雪艳,等.基于模态分解和TCN-BiLSTM的风电功率预测[J].电子测量技术,2024,47(14):49-56.  
FENG J L, LYU W D, DUAN X Y, et al. Wind power prediction based on mode decomposition and TCN-BiLSTM[J]. Electronic Measurement Technology, 2024, 47(14): 49-56.
- [6] SI ZH Y, YANG M, YU Y X, et al. Photovoltaic power forecast based on satellite images considering effects of solar position[J]. Applied Energy, 2021, 302: 117514.
- [7] JIANG Y, CHEN X Y, YU K, et al. Short-term wind power forecasting using hybrid method based on enhanced boosting algorithm[J]. Journal of Modern Power Systems and Clean Energy, 2017, 5(1): 126-133.
- [8] 李炜聪,潘福全,胡盼,等.基于季节性差分整合移动平均自回归模型的城市公交短期客流预测[J].济南大学学报(自然科学版),2022,36(3):308-314.  
LI W C, PAN F Q, HU P, et al. Short term passenger flow forecast of urban public transport based on seasonal autoregressive integrated moving average model[J]. Journal of University of Jinan(Science and Technology), 2022, 36(3): 308-314.
- [9] 王涛,高靖,王优胤,等.基于改进经验模态分解和支持向量机的风电功率预测研究[J].电测与仪表,2021,58(6):49-54.  
WANG T, GAO J, WANG Y Y, et al. Wind power prediction based on improved empirical mode decomposition and support vector machine [J]. Electrical Measurement & Instrumentation, 2021, 58(6): 49-54.
- [10] 任东方,马家庆,何志琴,等.基于AVMD-CNN-GRU-Attention的超短期风功率预测研究[J].太阳能学报,2024,45(6):436-443.  
REN D F, MA J Q, HE ZH Q, et al. Ultra-short-term wind power forecasting based on AVMD-CNN-GRU-Attention model [J]. Acta Energiæ Solaris Sinica, 2024, 45(6): 436-443.
- [11] 辛峰悦,寇志伟,刘焱晨,等.基于双流特征增强网络的超短期光伏功率预测[J].电子测量技术,2025,48(5):65-73.  
XIN F Y, KOU ZH W, LIU Y CH, et al. Ultra-short-term photovoltaic power prediction based on two-stream feature enhancement network [J]. Electronic Measurement Technology, 2025, 48(5): 65-73.
- [12] 吴仕宏,张璧臣,吴佳文,等.基于AWOA-Bi-LSTM的光伏发电功率预测[J].沈阳农业大学学报,2025,56(2):131-143.  
WU SH H, ZHANG B CH, WU J W, et al. Photovoltaic power generation prediction based on AWOA-Bi-LSTM[J]. Journal of Shenyang Agricultural University, 2025, 56(2): 131-143.
- [13] 徐武,范鑫豪,沈智方,等.多尺度特征提取的Transformer短期风电功率预测[J].太阳能学报,2025,46(2):640-648.  
XU W, FAN X H, SHEN ZH F, et al. Short-term wind power forecasting based on Transformer with multi-scale feature extraction [J]. Acta Energiæ Solaris Sinica, 2025, 46(2): 640-648.
- [14] 章志晃,徐启峰,林穿.基于SSA-BiLSTM-AM的短期风电功率预测[J].国外电子测量技术,2023,42(3):46-51.  
ZHANG ZH H, XU Q F, LIN CH. Short-term wind power prediction based on SSA-BiLSTM-AM [J]. Foreign Electronic Measurement Technology, 2023, 42(3): 46-51.
- [15] CHEN Y F, ZHAO H, ZHOU R, et al. CNN-BiLSTM short-term wind power forecasting method based on feature selection[J]. IEEE Journal of Radio Frequency Identification, 2022, 6: 922-927.
- [16] 李灯熬,白晓东,夏鸿伟,等.基于STL-Former的中短期光伏功率预测[J].热能动力工程,2024,39(4):

- 140-146,153.
- LI D AO, BAI X D, XIA H W, et al. Short-and medium-term photovoltaic power prediction based on STL-Former[J]. *Journal of Engineering for Thermal Energy and Power*, 2024, 39(4): 140-146,153.
- [17] 高正中,况逸,张经龙. 基于模态分解与组合神经网络的超短期风电功率预测[J]. *热力发电*, 2025,54(10): 82-92.
- GAO ZH ZH, KUAN Y, ZHANG J L. Ultra-short-term wind power prediction based on modal decomposition and combined neural network [J]. *Thermal Power Generation*, 2025,54(10):82-92.
- [18] SUN Y, YANG J J, ZHANG X T, et al. An ultra-short-term wind power forecasting model based on EMD-EncoderForest-TCN[J]. *IEEE Access*, 2024, 12: 60058-60069.
- [19] CHENG R K, YANG D, LIU D, et al. A reconstruction-based secondary decomposition-ensemble framework for wind power forecasting[J]. *Energy*, 2024, 308: 132895.
- [20] 刘畅,郑涛,王志华,等. 基于 DBSCAN 的智能变电站交流采样异常实时识别算法[J]. *电力系统保护与控制*, 2024, 52(24): 140-148.
- LIU CH, ZHENG T, WANG ZH H, et al. Algorithm for real-time identification of sampling abnormalities in smart substations based on DBSCAN[J]. *Power System Protection and Control*, 2024, 52(24): 140-148.
- [21] SOMAN S S, ZAREIPOUR H, MALIK O, et al. A review of wind power and wind speed forecasting methods with different time horizons [C]. *North American Power Symposium*, 2010: 1-8.
- [22] STITSYUK A, CHOI J. XPatch: Dual-stream time series forecasting with exponential seasonal-trend decomposition [J]. *ArXiv preprint arXiv: 2412.17323*,2024.
- [23] 杨信廷,郭向阳,韩佳伟,等. 基于 TiDE-PatchTST 模型的柑橘冷藏效率时序预测模型优化[J]. *农业机械学报*, 2024, 55(7): 396-404.
- YANG X T, GUO X Y, HAN J W, et al. Optimization of citrus cold storage efficiency time-series prediction model based on TiDE-PatchTST[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(7): 396-404.
- [24] BENTSEN L Ø, WARAKAGODA N D, STENBRO R, et al. Spatio-temporal wind speed forecasting using graph networks and novel Transformer architectures [J]. *Applied Energy*, 2023, 333: 120565.
- [25] 张心成,李翔晟,李藏龙,等. 基于改进 GRU 模型算法的锂电池 SOC 估计[J]. *电子测量技术*, 2023, 46(11): 57-65.
- ZHANG X CH, LI X SH, LI C L, et al. SOC estimation of lithium battery based on improved gated recurrent unit model algorithm [J]. *Electronic Measurement Technology*, 2023, 46(11): 57-65.
- [26] DOU J L, LIU CH CH, WANG B. Short-term wind power forecasting based on convolutional neural networks [C]. *IOP Conference Series: Earth and Environmental Science*, 2018, 170(4): 042023.
- [27] LIU X, ZHOU J. Short-term wind power forecasting based on multivariate/multi-step LSTM with temporal feature attention mechanism [J]. *Applied Soft Computing*, 2024, 150: 111050.
- [28] 陈庆明,廖鸿飞,孙颖楷,等. 基于 GWO-GRU 的光伏发电功率预测[J]. *太阳能学报*, 2024, 45(7): 438-444.
- CHEN Q M, LIAO H F, SUN Y K, et al. Photovoltaic power forecasting based on GWO-GRU [J]. *Acta Energiæ Solaris Sinica*, 2024, 45(7): 438-444.
- [29] ZHANG H T, ZHAO L X, DU ZH P. Wind power prediction based on CNN-LSTM[C]. *2021 IEEE 5th Conference on Energy Internet and Energy System Integration(EI2)*, 2021: 3097-3102.
- [30] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. *ArXiv preprint arXiv: 1706.0376*,2017.

## 作者简介

王雯学,硕士研究生,主要研究方向为新能源出力预测。

E-mail:wxw10059@163.com

乔燕军(通信作者),硕士,高级工程师,主要研究方向为新能源运维。

E-mail:yanjun.qiao@bjei.com

寇志伟,博士,副教授,主要研究方向为模式识别与智能系统、电工理论与新技术。

E-mail:kouzhiwei@imut.edu.cn

崔啸鸣,硕士,副教授,主要研究方向为导航与测控技术、模式识别与智能系统、检测技术与自动化装置。

E-mail:cuixiaoming@imut.edu.cn

任刚,硕士,主要研究方向为目标检测。

E-mail:2271446926@qq.com