

# 基于渐进式分解架构的风电时间序列预测<sup>①</sup>



丁浩<sup>1</sup>, 周成杰<sup>2</sup>, 车超<sup>1</sup>, 赵天明<sup>1</sup>, 周守亮<sup>3</sup>

<sup>1</sup>(大连大学 软件工程学院 先进设计与智能计算省部共建教育部重点实验室, 大连 116622)

<sup>2</sup>(大连理工大学 计算机科学与技术学院 社会计算与认知智能教育部重点实验室, 大连 116081)

<sup>3</sup>(大连大学 经济管理学院 (旅游学院), 大连 116622)

通信作者: 赵天明, E-mail: 1048850346@qq.com

**摘要:** 准确预测风电机组各项指标对准确管控机组和调控电网的供需有着重要意义. 预测指标任务可抽象为风电时间序列预测任务. 目前时间序列预测模型主要采用深度学习模型, 但是风电时间序列具有较强的波动性和随机性, 导致绝大部分模型不能较好挖掘风电时间序列的复杂演化特性. 为解决上述问题, 提出了一种基于渐进式分解架构的风电时间序列预测方法, 该方法首先应用神经网络池化分解方法将复杂的依赖关系简化并应用注意力机制学习长期趋势, 然后运用多变量融合捕捉模块增强了网络整体的多变量关联挖掘能力, 最后, 融合趋势项和周期项对风电时间序列做出准确的预测. 实验结果表明, 该方法在风电时间序列的多步预测中均方误差相比基线模型至高可提升 24%, 在多尺度预测长度下表现出预测性能稳定提升的同时, 计算效率显著优于同类模型.

**关键词:** 多变量时间序列预测; 神经网络; attention 机制; 时间序列分解

引用格式: 丁浩, 周成杰, 车超, 赵天明, 周守亮. 基于渐进式分解架构的风电时间序列预测. 计算机系统应用, 2024, 33(7): 112-120. <http://www.c-s-a.org.cn/1003-3254/9561.html>

## Time Series Forecasting of Wind Power Based on Progressive Decomposition Architecture

DING Hao<sup>1</sup>, ZHOU Cheng-Jie<sup>2</sup>, CHE Chao<sup>1</sup>, ZHAO Tian-Ming<sup>1</sup>, ZHOU Shou-Liang<sup>3</sup>

<sup>1</sup>(Key Laboratory of Advanced Design and Intelligent Computing Ministry of Education, School of Software Engineering, Dalian University, Dalian 116622, China)

<sup>2</sup>(Key Laboratory of Social Computing and Cognitive Intelligence Ministry of Education, School of Computer Science and Technology, Dalian University of Technology, Dalian 116081, China)

<sup>3</sup>(School of Economics and Management (School of Tourism), Dalian University, Dalian 116622, China)

**Abstract:** Accurate prediction of wind turbine metrics is important for accurate control of turbines and the regulation of grid supply and demand. The task of forecasting these indicators can be abstracted as a task of wind power time series forecasting. Currently, deep learning models are mainly used in time series prediction models, but the strong volatility and randomness of wind power time series often prevent most models from effectively capturing the complex evolutionary characteristics of the data. To address these issues, a wind power time series forecasting method based on a progressive decomposition architecture is proposed, which first applies a neural network pooling decomposition method to simplify complex dependencies and then applies an attention mechanism to learn long-term trends. Subsequently, a multivariate fusion capture module is employed to enhance the overall multivariate correlation mining ability of the network, and it fuses the trend term and the period term to make accurate forecasts of the wind power time series. Finally, the trend and period terms are fused to make accurate forecasts for wind power time series. Experimental results show that this method can achieve up to a 24% reduction in mean squared error (*MSE*) for wind power time series forecasting compared to

① 基金项目: 国家自然科学基金面上项目 (62076045); 辽宁省教育厅揭榜挂帅服务地方项目 (LJKFZ20220290); 大连大学学科交叉项目 (DLUXK-2023-YB-003, DLUXK-2023-YB-009, DLUXK-2022-ZD-003); 高等学校学科创新引智计划 (D23006)

收稿时间: 2024-01-15; 修改时间: 2024-02-26; 采用时间: 2024-03-04; csa 在线出版时间: 2024-05-31

CNKI 网络首发时间: 2024-06-04

baseline models. It also exhibits stable improvements in predictive performance across multiple forecasting lengths while significantly outperforming similar models in computational efficiency.

**Key words:** multivariate time series forecasting; neural network; attention mechanism; time series decomposition

风电的发展对于减少碳排放、改善环境污染、促进可持续发展都具有重要的意义。风电时间序列指风电机组运行过程相关的气象因素和机组产生的各项参数,如风速、气压、风电叶片转速、风电输出功率、扇叶角度等。随着风电装机容量占比的不断提高,风电并网稳定性问题日益突出,准确的风电时间序列预测成为保证电力系统稳定调度和安全运行的必要手段<sup>[1]</sup>。时间序列预测(time series forecasting, TSF)是实现准确风电机组相关参数预测的主要手段之一,它基于历史观测的时间序列数据,生成未来的预测数据,可以满足电力系统稳定调度的需求。

时间序列预测模型的研究可分为物理方法<sup>[2]</sup>、统计方法<sup>[3]</sup>和深度学习<sup>[4]</sup>。其中,物理方法需要大量前置假设,依赖于风电场周边的气压、气温等环境因素,并结合数值天气预报结果来预测风速,需要大量物理量测算成本高昂,因此难以进行规模化应用和推广,仅适合冷启动。机器学习方法和深度学习<sup>[4]</sup>方法则主要应用于已建设的风电场的功率预测。常见的机器学习方法有ARIMA<sup>[5]</sup>、支持向量回归<sup>[6]</sup>、随机森林<sup>[7]</sup>等,这些算法继物理方法后在时间序列预测中被广泛应用,能够取得良好的预测精度。近年来,随着风电场站大规模建设相关数据激增,为克服处理大规模数据和复杂非线性关系造成的瓶颈,深度学习<sup>[4]</sup>方法开始应用在风电场站的相关时间序列预测,进一步提高了预测精度。如考虑特征重组与改进Transformer的风电功率短期日前预测方法<sup>[4]</sup>等。这些Transformer方法可以在处理大规模数据时更好地挖掘数据中的非线性关系,提高预测的准确性和稳定性。但是由于风速时间序列复杂的非线性演化模式,以往的Transformer前馈层在将不同通道进行混合时,由于直接对特征进行通道混合无法捕捉单一通道的特征分布,忽略了风电时间序列各变量之间的动态耦合关系,导致模型的预测性能下降,这种现象被称之为通道漂移<sup>[8]</sup>。而且就风电时间序列本身特性而言,因为风速的强波动性影响,风电时间序列可能随着时间推移迅速产生复杂的变化,直接对序列进行建模会导致模型不能很好地建模这种复杂变化,

进而不能精细化预测结果。

为了改进现有深度学习模型的上述不足,本文提出一种基于序列分解和注意力机制的风电功率预测方法(progressive decomposition Transformer, PDformer),PDformer不直接对序列进行建模,而是应用渐进式分解的策略和Transformer捕捉时间序列的全局趋势和周期信息,并且创新性地设计了多变量融合捕捉模块,构建了一种新型的Transformer网络,该模型能更好地捕捉风电时间序列内部各协变量的动态耦合关系。其主要创新点如下。

1) 针对风电时间序列中错综复杂的序列依赖关系,设计了多变量融合捕捉模块(multivariate fusion capture, MFC),深度挖掘时间序列协变量之间的耦合关系,解决了现有模型受通道漂移影响导致的预测能力下降的问题。

2) 提出了PDformer模型,针对风电时间序列波动性产生的不稳定序列,应用渐进式分解架构,解决了以往模型难以建模强波动序列的问题。

在4个真实数据集上的对比实验结果表明,PDformer具有比基线模型更高的整体预测准确度,相比于同类模型拥有更高的效率。

## 1 相关工作

### 1.1 时间序列分解

风电时间序列由于受风力不稳定和强波动性的影响,直接对序列建模难以处理其复杂的非线性关系,变量间依赖关系会被纠缠的时间模式所掩盖。为了从错综复杂的时间模式中推理出依赖关系,本研究使用时间序列分解技术分离序列中的元素。目前,已涌现出众多与序列分解相关的技术,如经验模态分解(empirical mode decomposition, EMD)<sup>[9]</sup>,王涛等人使用改进的EMD方法将复杂的非线性时间序列分解为一组固有模态函数,结合支持向量机来预测风电功率,以提高模型的预测精度和稳定性<sup>[10]</sup>。变分模态分解(variational mode decomposition, VMD)和奇异谱分解(singular spectrum analysis, SSA)也是常用的风电特征提取方法。

这些方法可以将时间序列分解成一系列基函数,从而提取序列中的重要特征。例如, Lu 等人使用 VMD 方法对风电序列进行自适应时间尺度特征提取<sup>[11]</sup>。Shao 等人使用 SSA 方法提取时间序列的周期性特征,并结合卷积神经网络 (CNN) 进行风电功率预测<sup>[12]</sup>。但是以上方法都存在一定缺点,使用 EMD 进行序列分解的方法存在过拟合、噪声干扰等问题; VMD 和 SSA 方法需要人工提取时间信息,增加了模型对数据处理的依赖性。这些预处理方法受到历史序列简单分解效应的限制,忽视了时间序列底层模式之间的层次相互作用。Autoformer<sup>[13]</sup>将分解作为深度模型的内部块,在整个预测过程中逐步分解序列。受 Autoformer 启发,本研究使用神经网络中的池化操作,使用一个平均池化网络,利用渐进的分解方式去分解时间序列,将简单的分解操作堆叠达到更好的效果,针对风电时间序列过于复杂的特点,将其分解为多个简单的时间序列,对简单序列进行建模剥离纠缠时间序列内隐含的依赖关系。

## 1.2 时间序列预测深度学习方法

深度学习模型在风电时间序列预测中的应用已成为当前研究的热点之一。深度学习模型如 LSTM<sup>[14]</sup>可以从海量数据中学习得到风电时间序列数据的关键特征<sup>[15]</sup>。例如,基于 Bi-LSTM 的深度级联残差模型<sup>[16]</sup>可以获得可靠的点预测和区间预测结果;另外,目前已有多项研究利用注意力机制提高了风电时间序列预测的精度。例如,苏向敬等人在 LSTM 模型的基础上引入了特征空间和时序双重注意力机制,以提高可解释性和预测性能<sup>[17]</sup>。而 Liang 等人则构建了基于注意力机制的时序卷积网络 (TCN),用于风电功率的超短期点预测和区间预测<sup>[18]</sup>。此外,还有文献提出了采用改进自注意力机制的 TCN 结构并结合 1D-CNN 取得了较好的预测效果<sup>[19]</sup>。

针对长时间的时序信息与依赖关系难以捕捉的问题, Qin 等人提出了基于两阶段注意力机制的循环神经网络多变量预测算法 DA-RNN<sup>[20]</sup>,自适应地选择相关的输入特征表示,有利于捕捉长时间的时序信息与依赖关系。Huang 等人提出了一种双重自注意力网络 (DSA-Net)<sup>[21]</sup>,通过自注意力机制可适用于动态周期或非周期序列的预测。随着对注意力机制研究的深入 Transformer<sup>[22]</sup>也被引入时间序列预测领域,来捕捉时间序列预测中的长期依赖关系。Transformer 展示了很有前景的结果,例如使用稀疏注意力机制的 Informer 模型<sup>[23]</sup>、使用渐进式分解和自动关联机制的 Autoformer 模型<sup>[13]</sup>均取得

了较好的结果,但是 Transformer 中前馈网络在挖掘风电时间序列中错综复杂的模式时极易受到通道漂移的影响,模型的效果因此受到影响。

## 2 数据填充方法与模型构造

本节将介绍风电数据填充方法、PDformer 整体结构设计,频率增强模块和多变量融合捕捉模块的结构。

### 2.1 风电数据填充方法

风电机组数据在采集过程中不可避免地会出现数据丢失和异常的情况。缺失值会扰乱预测模型对数据内信息特征的提取,因此,需要对原始采集数据中的缺失值进行处理,以提高数据质量。本文选择三次样条插值方法<sup>[24]</sup>来填充缺失值,三次样条插值 (cubic spline interpolation) 简称 Spline 插值,该方法假设原始样本点在高维空间内呈一条光滑曲线,然后在数学上通过求解方程组得出曲线函数填补缺失点。定义输入时间序列  $\mathcal{X}' = \{\mathbf{x}'_1, \dots, \mathbf{x}'_{L_x} | \mathbf{x}'_i \in \mathbb{R}^{d_x}\}$  和输出时间序列  $\mathcal{Y}' = \{\mathbf{y}'_1, \dots, \mathbf{y}'_{L_y} | \mathbf{y}'_i \in \mathbb{R}^{d_y}\}$ ,  $\mathcal{X}'$  和  $\mathcal{Y}'$  之间的函数关系为  $\mathcal{Y}' = f(\mathcal{X}')$ 。采样的曲线是区间  $\mathbf{x} \in [\mathbf{x}'_i, \mathbf{x}'_{i+1}]$  中的三次多项式函数,其三次多项式可以表示为:

$$f_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i \quad (1)$$

风电数据处理前后对比如图 1 所示,图 1(a) 为未经过处理的图,图 1(b) 为经过三次样条插值后的图,由椭圆形圈出的缺失点前后对比,可看出经过填补后缺失点数据变得连续,曲线整体过度平稳。

由于风电时间序列各变量数量级不同,直接输入模型进行训练会出现模型收敛慢,网络训练不稳定的问题。为加速模型训练和提升预测效果,本研究采用 z-score 标准化对数据进行处理,控制训练期间不同物理量梯度满足标准正态分布。z-score 标准化是一种常用的数据标准化方法,用于将样本数据转换为服从标准正态分布。其公式可表示为:

$$\hat{x}_i = \frac{x_i - \mu}{\sqrt{\sigma^2}} \quad (2)$$

$$\mu = \frac{1}{m} \sum_{i=1}^m x_i \quad (3)$$

$$\sigma^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu)^2 \quad (4)$$

其中,  $x_i$  为样本的特征,  $m$  为数据集大小,  $\mu$  为数据集平均值,  $\sigma^2$  为数据集平方差。

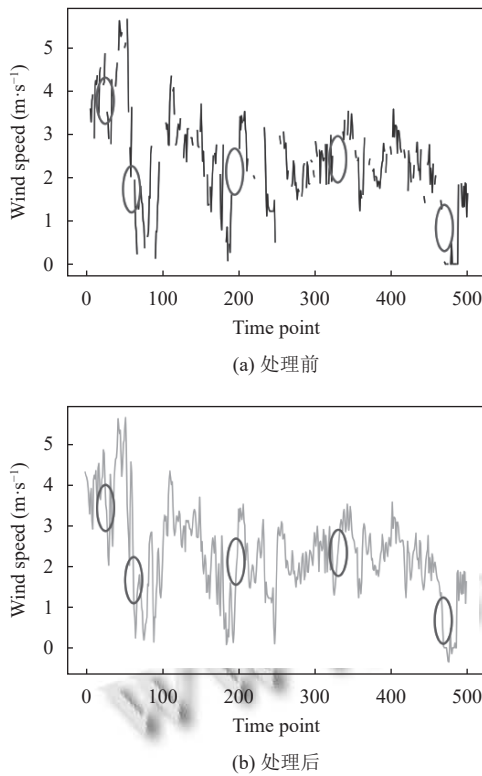


图1 风电数据处理前后对比图

### 2.2 PDformer 模型结构

本文 PDformer 模型整体结构如图2所示. 该模型基于编码器、解码器结构. 编码器和解码器由频率增强模块、分解模块、多变量融合捕捉模块堆叠构成.

模型采用与 Informer<sup>[23]</sup> 相同的一步生成方法.

编码器: 如图2所示, 编码器对时间序列的周期项部分进行建模, 旨在提取长连续输入的远距离依赖性. 输入在经过嵌入表示之后, 第  $t$  个序列输入  $\mathcal{X}^t$  已被嵌入为矩阵  $\mathcal{X}_{en}^t \in \mathbb{R}^{L \times d_{model}}$ . 嵌入表示输入编码器后先经过频率增强, 然后进行周期和趋势项分解, 最后经过一个多变量融合捕捉模块. 编码器可以堆叠  $N$  层, 编码器的输出包含过去的周期信息, 编码器输出将会输入到解码器的交叉注意力层来帮助解码器精细化预测结果. 编码器用式(5)表示, 第  $l$  个编码器层的公式可总结为:

$$\begin{cases} \mathcal{X}_{en}^l = Encoder(\mathcal{X}_{en}^{l-1}) \\ \mathcal{S}_{en,-}^{l,1} = MOEDecom(\mathcal{FEB}(\mathcal{X}_{en}^{l-1}) + \mathcal{X}_{en}^{l-1}) \\ \mathcal{S}_{en,-}^{l,2} = MFC(\mathcal{S}_{en}^{l,1}) + \mathcal{S}_{en}^{l,1} \end{cases} \quad (5)$$

其中,  $MOEDecom$  表示分解模块;  $FEB$  表示频率增强块;  $MFC$  表示多变量融合捕捉模块. “ $-$ ”为被剔除的趋势部分  $\mathcal{X}_{en}^l = \mathcal{S}_{en}^{l,2}, l \in \{1, \dots, N\}$  表示第  $l$  个编码器层的输出,  $\mathcal{X}_{en}^0$  是嵌入的  $\mathcal{X}_{en}$  分别表示第  $l$  层中第  $i$  个分解块或多变量融合捕捉模块之后的周期分量. 本文将在第2.3节详细描述多变量融合捕捉模块, 它可以提高模型整体挖掘能力.  $MOEDecom$  是一个池化分解结构, 采用的是平均池化方法,  $FEB$  为频率增强块后文有详细介绍.

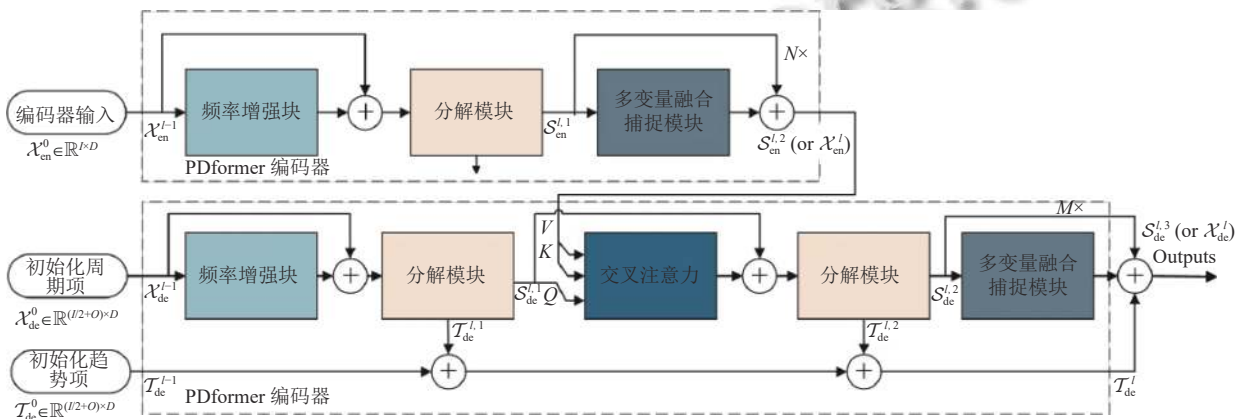


图2 PDformer 整体结构图

解码器: 解码器利用频率增强块和编码器-解码器交叉注意力块, 分别捕捉并利用现在和过去的周期信息. 在周期信息的捕捉中交替利用分解模块将原始输入分解. 每个分解模块分解出的潜在趋势与解码器输

出进行累加, 使 PDformer 能逐步细化趋势的预测. 解码器如编码器一般, 也可堆叠若干层. 假设有  $M$  个解码器层, 利用来自编码器的潜在变量  $\mathcal{X}_{en}^N$  和序列的部分输入  $\mathcal{X}_{dc}^0$ , 解码器的公式表示如下:

$$\begin{cases} S_{de}^{l,1}, \mathcal{T}_{de}^{l,1} = MOEDecomp(FEB(\mathcal{X}_{de}^{l-1}) + \mathcal{X}_{de}^{l-1}) \\ S_{de}^{l,2}, \mathcal{T}_{de}^{l,2} = MOEDecomp(MA(S_{de}^{l,1}, \mathcal{X}_{en}^N) + S_{de}^{l,1}) \\ S_{de}^{l,3} = MFC(S_{de}^{l,2}) + S_{de}^{l,2} \\ \mathcal{X}_{de}^l = S_{de}^{l,3} \\ \mathcal{T}_{de}^l = \mathcal{T}_{de}^{l-1} + \mathcal{W}_{l,1} \cdot \mathcal{T}_{de}^{l,1} + \mathcal{W}_{l,2} \cdot \mathcal{T}_{de}^{l,2} \end{cases} \quad (6)$$

其中,  $\mathcal{X}_{de}^{l-1}$  表示前一个解码器堆栈的输出, 当  $l=0$  时, 表示原始编码器初值;  $MA$  表示多头注意力计算。

### 2.3 频率增强模块

如图 2 所示编码器和解码器均采用了频率增强模块捕捉时间序列中隐藏的周期信息。图 3 展示了频率增强块的构造细节, 首先对频率增强块的输入 ( $x \in \mathbb{R}^{N \times D}$ ) 应用一个线性层进行投影 ( $w \in \mathbb{R}^{D \times D}$ ) 增加其表示能力, 然后对输入应用离散傅里叶变换将输入由时域转换到频域。下采样操作中, 由于风电时间序列采集过程中受风力不稳和扇叶变相等因素产生了一些噪声, 本文在频域中去一些高频噪声的同时从其余分量中随机挑选  $M$  个分量, 这  $M$  个分量为代表性的频域信号。  $R$  ( $R \in \mathbb{C}^{D \times D \times M}$ ) 是一个随机初始化的参数核, 用于增加原始频域空间的表现能力。将  $R$  与下采样的特征图相乘, 然后对运算结果进行零填充, 使尺寸和原大小相同, 之后应用逆傅里叶变换将特征图转换回时域。

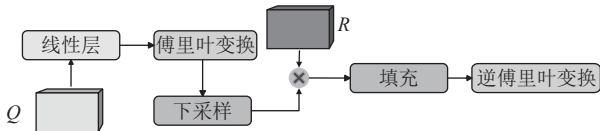


图 3 频率增强块

### 2.4 多变量融合捕捉模块

多变量融合捕捉模块在编码器和解码器中均有使用, 结构如图 4 所示, 多变量融合捕捉模块的输入 ( $x \in \mathbb{R}^{N \times D}$ )。首先使用  $w \in \mathbb{R}^{H \times N}$  进行线性投影, 因此  $A^{[1]} = w \cdot x, A^{[1]} \in \mathbb{R}^{H \times D}$ 。然后对  $A^{[1]}$  应用 ReLU 激活函数并随机屏蔽一些神经元防止过拟合。再应用一个线性层  $\psi \in \mathbb{R}^{N \times H}$  将维度映射回原向量空间内, 以便将网络堆叠多层。多变量融合捕捉模块可用公式表示如下:

$$\begin{cases} A = ReLU(w \cdot x) \\ MLP(x) = dropout[\psi(dropout(A))] \end{cases} \quad (7)$$

其中,  $dropout$  表示随机屏蔽操作;  $ReLU$  表示激活函数。

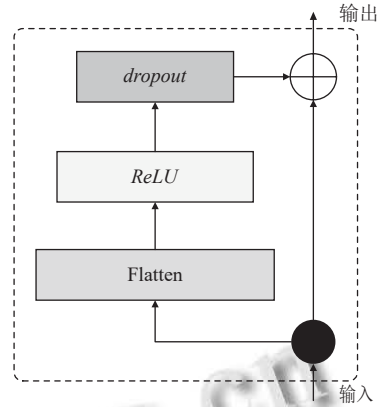


图 4 多变量融合捕捉模块

## 3 实验

### 3.1 实验数据集

为评估本文模型 PDformer 的效果, 在 4 个真实数据集上开展了实验, 数据集的基本信息如表 1 所示。

表 1 数据集简介

数据集	数据集长度	采样间隔 (min)	变量数
Wind154	12 685	15	12
Wind112	23 798	15	12
Wind170	13 418	15	12
Wind	48 673	15	7

Wind154、Wind112、Wind170 数据集收集自于某公司实际风电场站, 记录了 3 个地区风电场站的实际风电场站数据。该数据集包含丰富的变量信息如机组功率、天气信息、扇叶转速、扇叶角度等。

Wind 公开数据集 (<https://github.com/PaddlePaddle/PaddleSpatial/tree/main/paddlespatial/datasets/WindPower>), 记录了从 2020 年 1 月到 2021 年 7 月, 以 15 min 的时间间隔记录风电场的发电量及其协变量数据。

基准算法: 本文主要选取了 5 个基准算法分别是: Informer、Autoformer、FEDformer、ARIMA、GRU。

- Informer<sup>[23]</sup>: 一个基于 Transformer 的稀疏注意力模型, 它针对多变量时间序列应用了自注意力和交叉机制, 学习历史数据并利用一步生成的策略进行预测。

- ARIMA<sup>[5]</sup>: 一种自回归积分滑动平均模型 (autoregressive integrated moving average model), 是一种简单的时间序列预测模型, 通过将非平稳序列转化为平稳序列然后对滞后值和带有随机误差的现值进行建模和回归的模型。

• Autoformer<sup>[13]</sup>: 一个基于渐进式分解架构的模型, 它设计了渐进式分解结构和自相关机制, 实现了时间序列复杂的时间模式挖掘。

• FEDformer<sup>[25]</sup>: 一个用于长期时间序列预测的频率增强 Transformer 模型, 设计了频率增强注意力和专家分解机制来控制分布转移。

• GRU<sup>[26]</sup>: 一个基于循环神经网络的改进模型, 可以解决循环神经网络不能长期记忆和反向传播中的梯度问题。

评估指标: 本文选取了两种常用的评估指标来衡量模型预测性能, 分别是均方误差 (mean squared error,  $MSE$ ) 和平均绝对误差 (mean absolute error,  $MAE$ ), 其计算公式如式 (8) 和式 (9)。这两个指标表示预测值与真实值之间的误差, 因此指标越低代表模型效果越好。

$$MSE = \frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2 \quad (8)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y - \hat{y}| \quad (9)$$

### 3.2 模型参数

本文实验设置训练和评估的输入长度固定为 96, 预测长度分别固定为 48、96、192、336, 实验平台为 NVIDIA V100 16 GB, CUDA Version 11.6, 系统版本 Ubuntu 18.04.6 LTS, GRU 和 Informer、Autoformer 等深度学习模型超参数设置如表 2 所示。ARIMA 模型参数如表 3 所示。

表 2 深度学习模型超参数选取

超参数	模型取值
激活函数	GELU
损失函数	$MSE$
迭代次数	10
学习率	依次减半
优化器	Adam
Batch size	16

表 3 ARIMA 模型超参数选取

超参数	模型取值
模型参数选取准则	AIC准则
参数 $p$	$2 \leq p \leq 5$
参数 $d$	$0 \leq d \leq 2$
参数 $q$	$2 \leq q \leq 5$
是否季节性调整	是

### 3.3 实验结果

$MAE$  结果较为保守,  $MSE$  为平方误差更能反映风电时间序列预测的波动性。为了更好地对比模型应对风电波动性的效果, 本文分析更侧重于  $MSE$ 。

表 4 给出了 Wind154、Wind112、Wind170 和 Wind 数据集下不同算法的  $MSE$  和  $MAE$ 。ARIMA 和 GRU 等传统算法在所有数据集上均表现较差, 随着预测长度变长效果急剧下降, PDformer 在 4 个基线数据集上与 ARIMA 相比, 336 的预测长度下  $MSE$  指标平均提升约 86%, 这是由于传统模型无法处理非平稳序列和中长期下风电时间序列各项变量间复杂的依赖关系。

表 4 实验结果

数据集	预测长度	PDformer		Autoformer		Informer		GRU		FEDformer		ARIMA	
		$MSE$	$MAE$	$MSE$	$MAE$	$MSE$	$MAE$	$MSE$	$MAE$	$MSE$	$MAE$	$MSE$	$MAE$
Wind154	48	<b>0.598</b>	0.583	0.831	0.658	0.726	0.627	0.875	0.722	<u>0.611</u>	<b>0.552</b>	0.721	<u>0.576</u>
	96	<b>0.702</b>	<b>0.599</b>	0.830	0.669	0.897	0.728	0.861	0.752	<u>0.778</u>	<u>0.628</u>	1.410	0.848
	192	<b>0.856</b>	<u>0.728</u>	<u>0.941</u>	<b>0.722</b>	1.129	0.832	1.023	0.810	0.998	0.736	3.150	1.272
	336	<b>0.956</b>	<b>0.764</b>	1.145	<u>0.811</u>	1.149	0.823	<u>1.044</u>	0.827	1.293	0.859	6.977	1.844
Wind112	48	<b>0.538</b>	<b>0.508</b>	0.676	0.582	<u>0.593</u>	0.580	0.860	0.740	0.606	<u>0.537</u>	0.852	0.575
	96	<b>0.632</b>	<b>0.553</b>	0.838	0.676	0.787	0.684	0.894	0.760	<u>0.661</u>	<u>0.564</u>	1.299	0.738
	192	<b>0.720</b>	<b>0.595</b>	0.763	0.620	0.898	0.735	0.969	0.800	<u>0.740</u>	<u>0.610</u>	2.420	0.999
	336	<b>0.741</b>	<b>0.606</b>	0.805	0.650	0.855	0.735	1.008	0.819	<u>0.790</u>	<u>0.640</u>	4.689	1.362
Wind170	48	<b>0.362</b>	<b>0.404</b>	0.903	0.657	0.674	0.607	0.817	0.686	<u>0.631</u>	0.540	0.636	<u>0.481</u>
	96	<b>0.614</b>	<b>0.532</b>	0.829	0.643	<u>0.808</u>	0.670	0.943	0.736	0.842	<u>0.633</u>	1.451	0.786
	192	<b>0.903</b>	<b>0.677</b>	1.115	0.763	1.398	0.884	1.095	0.820	<u>1.067</u>	<u>0.726</u>	3.052	1.189
	336	<b>1.131</b>	<b>0.778</b>	1.292	<u>0.835</u>	1.531	0.935	<u>1.173</u>	0.856	1.548	0.930	6.839	1.737
Wind	48	<b>0.749</b>	<b>0.579</b>	1.102	0.707	0.892	0.612	1.109	0.757	<u>0.826</u>	<u>0.602</u>	1.388	0.697
	96	<b>0.989</b>	<b>0.681</b>	1.152	0.731	1.114	0.708	1.214	0.799	<u>1.048</u>	<u>0.692</u>	2.890	1.063
	192	<b>1.186</b>	0.793	1.302	0.807	<u>1.281</u>	<b>0.776</b>	1.304	0.823	1.287	<u>0.791</u>	6.405	1.603
	336	<b>1.294</b>	<u>0.852</u>	1.487	0.883	<u>1.329</u>	<b>0.825</b>	1.335	0.854	1.561	0.886	14.636	2.372

注: 输入长度  $l=96$ , 预测长度  $O \in \{48, 96, 192, 336\}$ , 最好的结果以粗体显示, 次好结果以下划线表示

PDformer 相比 Informer、Autoformer、FEDformer 提升依旧明显,在基准数据集上至高可提升 60% 的 *MSE* 指标,平均提升可达 18%。值得注意的是,虽然 Autoformer、FEDformer 也采用了渐进式分解的设计思想,但是其传统前馈层的设计受到了通道漂移的影响,混淆了部分变量间的依赖关系,因此预测的误差高于 PDformer。PDformer 采用了多变量融合捕捉模块在建模多个变量复杂依赖的同时控制通道混合的次数,由此获得了最优的性能表现。以 *MSE* 为评价指标,在全部基准数据集上 PDformer 在所有预测长度都取得了最好的效果,这是因为 PDformer 的多变量融合捕捉模块具有强大的变量间关联挖掘能力。随着预测长度的延长 PDformer 的相对优势会更加明显,这证明了 PDformer 由于渐进式分解架构的存在,可以成功对长期风电时间序列进行精准建模,多变量融合捕捉模块在长距离条件下使得 PDformer 能更贴合风电时间序列的依赖挖掘需求。

*MAE* 指标下,PDformer 同样表现较好,与所有模型相比 PDformer 至高可提升 64%,与次优结果相比提升最高可达 16%。这再次验证了 PDformer 在风电时间序列预测中的有效性,证明了 MFC 模块增强了模型的多变量演化挖掘能力。

深度学习模型在 4 个数据集上的平均运行时间如图 5 所示。PDformer 由于减少了堆叠层数和优化了频域采样方法,与同样应用频域采样技术的 FEDformer 相比运行速度平均可提升 83%,并且短期预测长度下运行时间低于 Autoformer 模型。全部预测长度下优于 Informer 模型。PDformer 在中长期预测长度下,由于傅里叶变换为平方时间复杂度,速度会慢于 Autoformer,但是速度的降低换取了平均 20% 的 *MSE* 指标提升。

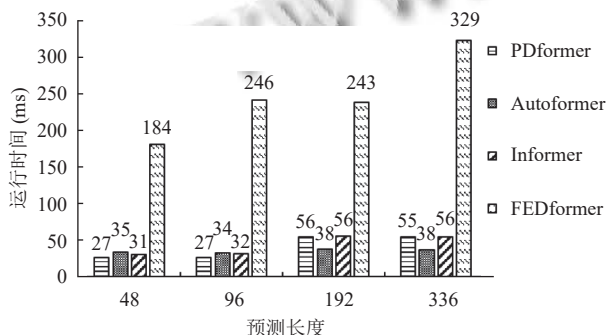


图 5 平均运行时间对比图

总体而言,PDformer 的改进在不同的预测长度下表现是一致的,因为 MFC 模块和渐进式分解架构赋予

了模型挖掘复杂依赖的能力,相对于其他模型没有因为风电的波动性产生预测效果的下降。实验结果证明了其在风电时间序列预测中可平稳化强波动性的风电时间序列,在多个预测尺度下均具有稳定的效率和准确率优势。

### 3.4 消融实验

为了证明 MFC 模块可以帮助模型更好地学习了当前通道值和历史值的潜在关系,本节进行了消融实验,将完整 PDformer 的结果作为基准和去除 MFC 模块的 PDformer 算法进行比较。

从表 5 可以看出,在 4 个基准数据集上引入多变量融合捕捉模块后,预测结果显著提升,在 Wind170 长度为 48 的预测窗口下,最高可以提升 *MSE* 约 77%。在不同长度的预测窗口下,带有 MFC 的模型都达到了最小的 *MSE*,证明了多变量融合捕捉模块的有效性。且长预测窗口相对短预测窗口提升更大,这是因为 MFC 模块能够对输入数据进行较长跨度的考量和特征提取,从而捕捉到通道值和历史值之间的复杂关联,避免了随时间演进产生的通道依赖变化的风险,进而提高了对风电时间序列数据的建模和预测能力。

表 5 消融实验结果

数据集	预测长度	PDformer		w/o MFC	
		<i>MSE</i>	<i>MAE</i>	<i>MSE</i>	<i>MAE</i>
Wind154	48	<b>0.598</b>	0.583	0.607	<b>0.549</b>
	96	<b>0.702</b>	<b>0.599</b>	0.741	0.616
	192	<b>0.901</b>	<b>0.703</b>	0.926	0.704
	336	<b>0.956</b>	<b>0.764</b>	1.149	0.802
Wind112	48	<b>0.538</b>	<b>0.508</b>	1.010	0.875
	96	<b>0.632</b>	<b>0.553</b>	1.046	0.861
	192	<b>0.720</b>	<b>0.595</b>	1.019	0.882
	336	<b>0.741</b>	<b>0.606</b>	1.041	0.893
Wind170	48	<b>0.362</b>	<b>0.404</b>	1.622	1.074
	96	<b>0.614</b>	<b>0.532</b>	1.651	1.076
	192	<b>0.903</b>	<b>0.677</b>	1.708	1.098
	336	<b>1.131</b>	<b>0.778</b>	1.626	1.055
Wind	48	<b>0.749</b>	<b>0.579</b>	1.068	0.750
	96	<b>0.989</b>	<b>0.681</b>	1.205	0.774
	192	<b>1.186</b>	0.793	1.218	<b>0.782</b>
	336	<b>1.294</b>	<b>0.852</b>	1.451	0.869

注:基准数据集多元长期序列预测结果,输入长度  $I=96$ ,预测长度  $O \in \{48, 96, 192, 336\}$ 。最佳结果以粗体突出显示,w/o MFC 表示去除 MFC 模块的 PDformer 模型

## 4 结论与展望

本研究针对风电时间序列强波动性的特性,设计了一种新的深度学习模型结构,改进了现有算法面对

风电时间序列多变量演化特性挖掘能力的不足. 本文提出的 PDformer 模型应用改进型渐进式分解结构, 分离原始序列趋势项和周期项, 使不平稳序列变成平稳序列进而可以更好地学习序列演化规律. 此外, 利用 MFC 模块替换捕捉到了当前通道值和历史值之间的复杂关联, 以及时间序列值的动态时序特征, 增强了现有模型的预测能力. 在 4 组风电数据集上进行的详尽实证研究表明, 基于渐进式分解架构的风电时间序列预测模型对于多变量演化特性的挖掘优于现有的时间序列预测算法. 实现了时间序列建模方法对风电时间序列波动性的针对性优化.

虽然在不同预测长度下 PDformer 均能达到最好效果, 但 PDformer 在超短期的预测窗口下较之长期相对提升变少, 超短期的预测表现和次优模型相差较小, 因此在超短期预测上仍有提升空间. 在未来的工作中, 会在本算法的基础上研究如何提升超短期预测的效果, 进一步优化风电时间序列的预测精度和模型泛化能力.

#### 参考文献

- 米阳, 卢长坤, 申杰, 等. 基于条件生成对抗网络的风电功率极端场景生成. 高电压技术, 2023, 49(6): 2253–2263.
- 冯双磊, 王伟胜, 刘纯, 等. 风电场功率预测物理方法研究. 中国电机工程学报, 2010, 30(2): 1–6.
- 孙荣富, 张涛, 和青, 等. 风电功率预测关键技术及应用综述. 高电压技术, 2021, 47(4): 1129–1143.
- 李练兵, 高国强, 吴伟强, 等. 考虑特征重组与改进 Transformer 的风电功率短期日前预测方法. 电网技术, 2023: 1–13. [doi: 10.13335/j.1000-3673.pst.2023.0656]
- Hodge BM, Zeiler A, Brooks D, *et al.* Improved wind power forecasting with ARIMA models. Computer Aided Chemical Engineering, 2011, 29: 1789–1793.
- Dhiman HS, Deb D, Guerrero JM. Hybrid machine intelligent SVR variants for wind forecasting and ramp events. Renewable and Sustainable Energy Reviews, 2019, 108: 369–379. [doi: 10.1016/j.rser.2019.04.002]
- Ye L, Li YL, Pei M, *et al.* A novel integrated method for short-term wind power forecasting based on fluctuation clustering and history matching. Applied Energy, 2022, 327: 120131. [doi: 10.1016/j.apenergy.2022.120131]
- Zeng AL, Chen MX, Zhang L, *et al.* Are transformers effective for time series forecasting? Proceedings of the 37th AAAI Conference on Artificial Intelligence. Washington: AAAI, 2023. 11121–11128.
- 谢丽蓉, 王斌, 包洪印, 等. 基于 EEMD-WOA-LSSVM 的超短期风电功率预测. 太阳能学报, 2021, 42(7): 290–296.
- 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. 电测与仪表, 2021, 58(6): 49–54.
- Lu P, Ye L, Pei M, *et al.* Short-term wind power forecasting based on meteorological feature extraction and optimization strategy. Renewable Energy, 2022, 184: 642–661. [doi: 10.1016/j.renene.2021.11.072]
- Shao Z, Han J, Zhao W, *et al.* Hybrid model for short-term wind power forecasting based on singular spectrum analysis and a temporal convolutional attention network with an adaptive receptive field. Energy Conversion and Management, 2022, 269: 116138. [doi: 10.1016/j.enconman.2022.116138]
- Wu HX, Xu JH, Wang JM, *et al.* Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. Proceedings of the 35th Conference on Neural Information Processing Systems. 2021. 22419–22430.
- Ewees AA, Al-Qaness MAA, Abualigah L, *et al.* HBO-LSTM: Optimized long short term memory with heap-based optimizer for wind power forecasting. Energy Conversion and Management, 2022, 268: 116022. [doi: 10.1016/j.enconman.2022.116022]
- 杨京渝, 罗隆福, 阳同光, 等. 基于气象特征挖掘和改进深度学习模型的风电功率短期预测. 电力自动化设备, 2023, 43(3): 110–116.
- Ko MS, Lee K, Kim JK, *et al.* Deep concatenated residual network with bidirectional LSTM for one-hour-ahead wind power forecasting. IEEE Transactions on Sustainable Energy, 2021, 12(2): 1321–1335. [doi: 10.1109/TSTE.2020.3043884]
- 苏向敬, 周汶鑫, 李超杰, 等. 基于双重注意力 LSTM 神经网络的可解释海上风电出力预测. 电力系统自动化, 2022, 46(7): 141–151.
- Liang JK, Tang WY. Ultra-short-term spatiotemporal forecasting of renewable resources: An attention temporal convolutional network-based approach. IEEE Transactions on Smart Grid, 2022, 13(5): 3798–3812. [doi: 10.1109/TSG.2022.3175451]
- Ma ZJ, Mei G. A hybrid attention-based deep learning approach for wind power prediction. Applied Energy, 2022, 323: 119608. [doi: 10.1016/j.apenergy.2022.119608]
- Qin Y, Song DJ, Chen HF, *et al.* A dual-stage attention-based recurrent neural network for time series prediction. Proceedings of the 26th International Joint Conference on



- Artificial Intelligence. Melbourne: IJCAI.org, 2017. 2627–2633.
- 21 Huang ST, Wang DL, Wu XH, *et al.* DSANet: Dual self-attention network for multivariate time series forecasting. Proceedings of the 28th ACM International Conference on Information and Knowledge Management. Beijing: ACM, 2019. 2129–2132.
- 22 Vaswani A, Shazeer N, Parmar N, *et al.* Attention is all you need. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 6000–6010.
- 23 Zhou HY, Zhang SH, Peng JQ, *et al.* Informer: Beyond efficient Transformer for long sequence time-series forecasting. Proceedings of the 35th AAAI Conference on Artificial Intelligence. AAAI, 2021. 11106–11115.
- 24 Sun M, Lan L, Zhu CG, *et al.* Cubic spline interpolation with optimal end conditions. Journal of Computational and Applied Mathematics, 2023, 425: 115039. [doi: [10.1016/j.cam.2022.115039](https://doi.org/10.1016/j.cam.2022.115039)]
- 25 Zhou T, Ma ZQ, Wen QS, *et al.* FEDformer: Frequency enhanced decomposed Transformer for long-term series forecasting. Proceedings of the 39th International Conference on Machine Learning (ICML). Baltimore: PMLR, 2022. 27268–27286.
- 26 Fu R, Zhang Z, Li L. Using LSTM and GRU neural network methods for traffic flow prediction. Proceedings of the 31st Youth Academic Annual Conference of Chinese Association of Automation. Wuhan: IEEE, 2016. 324–328.

(校对责编: 孙君艳)