

DOI:10.20033/j.1003-7241.(2026)04-0045-08

# 基于时空感知图卷积注意力网络的超短期风电功率预测

徐全海<sup>1</sup>, 赵文成<sup>1</sup>, 牛海明<sup>2</sup>, 曾凡斐<sup>2</sup>, 赵陆阳<sup>3</sup>(1. 国家能源集团青海电力有限公司, 青海 西宁, 817000; 2. 国能智深控制技术有限公司, 北京 102209,  
3. 华北电力大学控制与计算机工程学院, 北京 100096)

**摘要:**针对风电机组集群超短期多步功率预测中空间相关性刻画不足、时间依赖关系建模能力有限等问题,提出了一种基于时空感知图卷积注意力网络的风电集群功率预测方法。该方法面向风电机组非规则空间分布特性,在空间信息提取模块中构建空间动态图卷积网络,并引入空间注意力机制对图邻接矩阵权重进行动态调整,从而自适应刻画不同运行工况下机组间的空间相关性;在时间信息提取模块中,设计时空感知自注意力机制,通过融合局部时间趋势信息与全局依赖关系,提高对功率波动特征的建模能力。在此基础上,构建时空协同预测框架,实现对风电机组集群超短期多步功率的联合预测。以河北省承德市某风电场33台风电机组的实际SCADA数据为算例,开展对比实验与消融分析。实验结果表明,所提模型在不同预测步长下均优于GRU、LSTM及传统Transformer等模型。研究结果验证了所提方法在挖掘风电机组集群时空关联特征方面的有效性,可为风电场运行调度提供技术支撑。

**关键词:**风功率预测;注意力机制;集群预测;图卷积网络;Transformer模型;时空相关性

中图分类号:TK8

文献标志码:A

文章编号:1003-7241(2026)04-0045-08

## Ultra-short-term wind power forecasting based on a spatiotemporal-aware graph convolutional attention network

XU Quanhai<sup>1</sup>, ZHAO Wencheng<sup>1</sup>, NIU Haiming<sup>2</sup>, ZENG Fanfei<sup>2</sup>, ZHAO Luyang<sup>3</sup>

(1. National Energy Group Qinghai Electric Power Co., Ltd., Xining 817000, Qinghai, China;

2. CHN Energy Zhishen Control Technology Co., Ltd., Beijing 102209, China;

3. School of Control and computer Engineering, North China Electric Power University, Beijing 100096, China)

**Abstract:** To address the limitations in characterizing spatial correlations and modeling temporal dependencies in ultra-short-term multi-step power forecasting for wind turbine clusters, a wind farm power forecasting method based on a spatiotemporal-aware graph convolutional attention network is proposed. Considering the irregular spatial distribution of wind turbines, a spatial dynamic graph convolutional network is constructed within the spatial information extraction module, and a spatial attention mechanism is introduced to dynamically adjust the weights of the graph adjacency matrix, thereby adaptively capturing inter-turbine spatial correlations under varying operating conditions. In the temporal information extraction module, a temporally aware self-attention mechanism is designed by integrating local temporal trend information with global dependency relationships, which enhances the modeling capability for power fluctuation characteristics. On this basis, a spatiotemporal collaborative forecasting framework is established to achieve joint ultra-short-term multi-step power prediction for wind turbine clusters. Comparative experiments and ablation studies are conducted using real SCADA data from 33 wind turbines at a wind farm in Chengde, Hebei province. Experimental results demonstrate that the proposed model consistently outperforms GRU, LSTM, and conventional Transformer models across different forecasting horizons. These findings verify the effectiveness of the proposed method in extracting spatiotemporal correlation features of wind turbine clusters and provide technical support for wind farm operation and dispatch.

**Keywords:** wind power prediction; attention mechanism; cluster prediction; graph convolutional network; transformer; spatiotemporal correlation

风电作为一种清洁且可再生的能源,已成为应对能源危机和环境问题的有效替代方案。然而,在实际应用中,由于风况的不断变化,风力发电面临着波动性和间歇性的挑战,这些因素威胁到电力系统的稳定性。因此,解决风力发电的不稳定性问题,是风能有效利用和管理的关键。

为了防止能源波动引发电网事故,有必要对风电机组的短期功率进行准确预测<sup>[1]</sup>。

传统的风电功率预测方法包括物理方法和统计模型。物理方法主要依赖数值天气预报(numerical weather prediction, NWP)进行预测<sup>[2]</sup>,在中长期功率预测方面表

收稿日期:2025-02-14;录用日期:2025-03-25

基金项目:国家能源集团重大项目(1122422111010)

作者简介:徐全海(1967—),男,高级工程师,研究方向:新能源功率预测。

引用本文:徐全海,赵文成,牛海明,等.基于时空感知图卷积注意力网络的超短期风电功率预测[J].自动化技术与应用,2026,45(4):45-52.(XU Quanhai, ZHAO Wencheng, NIU Haiming, et al. Ultra-short-term wind power forecasting based on a spatiotemporal-aware graph convolutional attention network[J]. Techniques of Automation and Applications, 2026, 45(4): 45-52.)

现良好。但是,NWP并不适用于超短期预测,因为它需要数据同化和计算的时间,并且将天气预报数据为功率的过程中引入的额外不确定性大于超短期时间尺度的自然变化<sup>[3]</sup>。相较之下,统计模型利用历史观测数据进行风电功率预测,其在短期预测任务中的效果优于物理方法。文献[4]将经验模态分解与核极限学习机相结合,有效改善了超短期风电功率预测精度。文献[5]提出一种基于可变模式分解(variational mode decomposition, VMD)-样本熵(sample entropy, SE)的时间序列分解方法将原始风功率时间序列分解为不同带宽的模式分量,之后使用粒子群优化的贝叶斯神经网络进行预测,取得了较高的预测精度。此外高斯混合模型<sup>[6]</sup>、XGBoost<sup>[7]</sup>等统计学习方法也被应用于风电功率预测任务中。然而,统计方法需要预先定义的数学模型,无法理解复杂的非线性特征,在处理特别复杂的非线性特征或极端变化时可能遇到限制,从而影响其预测准确性。

随着人工智能的快速发展,越来越多的学者将深度学习应用于风电功率预测问题<sup>[8]</sup>。文献[9]构建卷积神经网络(convolutional neural network, CNN)、双向长短期记忆(bi-directional long short term memory, BiLSTM)和注意力机制(attention mechanism, AM)的混合深度学习模型,结合各模块的优势对风电功率进行预测。文献[10]将小波变换技术与多种深度学习算法结合,充分提取和利用了风电功率序列的时序特征。文献[11]利用注意力机制挖掘重要时刻信息,同时使用鲸鱼优化算法优化双向门控循环单元(bi-directional gate recurrent unit, BiGRU)模型的超参数,加速模型收敛,取得了较好的效果。文献[12]提出集成多尺度长短期记忆网络(long short-term memory neural network, LSTM)集成不同时间尺度的风功率预测结果得出最终预测值,提升了模型的精度与稳定性。然而,RNN系列网络在处理长序列数据时存在一个显著的缺陷,即信息在传递过程中的逐渐衰减,这被称为“长期依赖问题”。这意味着随着输入序列的增长,网络开始丢失关于早期输入的信息,这种情况可能会影响风电功率预测的准确性。近年来,Transformer模型因其在自然语言处理中的成功应用,逐渐被引入时序预测任务中。Transformer的自注意力机制使其能够捕捉长时间序列中的依赖关系,并解决了RNN类模型信息逐渐衰减的问题。因此,Transformer及其变种模型在风电功率预测等复杂时序问题中展现了较强的表现力<sup>[13]</sup>。

以上方法主要针对单台风电机组的功率进行研究。然而,由于数据采集和传输过程中不可避免的误差,仅依赖单台风电机组的运行数据,可能会影响预测结果的准确性<sup>[1]</sup>。在实际风电场中,通常部署有多台风电机组。由于风由高压区域流向低压区域,因此风电场中相邻区域的风况往往具有相似性,表现出显著的空间相关性<sup>[14]</sup>。探索并利用这些风电机组之间的时空相关性,以实现精确的风电集群功率预测,已成为当前研究领域的热点问题。

在文献[15]中,作者采用CNN来提取风电场内部 $12 \times 12$ 规模的风电机组风速矩阵的空间相关性然后运用简单循环单元来进行多个风电机组的风速预测。文献[16]也使用了CNN捕获风电机组风速矩阵的空间相关性,但在后续阶段使用了LSTM来实现多风电机组的风速预测。文献[17]使用三维卷积神经网络(3D-CNN)通过时空特征信息解耦建立了多输出风速预测模型。显然,CNN和循环神经网络(recurrent neural networks, RNN)已经广泛应用于风能预测问题的建模中。尽管如此,风电集群功率预测仍然是一个具有挑战性的问题,因为以下问题尚未解决:

1) 上述研究主要针对规则排列的风电机组阵列。在特定地形,如山区的风电场,风电机组的安置位置往往不规则,这可能对常规空间相关性的提取方法构成挑战。这一情况要求对空间相关性的理解和利用方法进行调整,以适应复杂和不规则的风电机组布局,从而确保功率预测的准确性。

2) Transformer模型在时间序列预测领域中取得了显著进展。然而,由于其结构未针对风电集群中的空间关联特性进行专门设计,因此在多风电机组功率预测任务中可能存在局限性。

针对这些问题,本文提出了一种基于时空感知图卷积注意力网络的风电集群功率预测模型。

具体而言,本文在空间相关性提取模块中设计了空间动态图卷积网络(spatial dynamic graph convolutional network),该网络适用于风电机组分布不规则的风电集群,并通过空间注意力机制动态更新图邻接矩阵的权重,以适应不同工况。在时间相关性提取模块中,本研究设计了时态感知注意力网络。该网络具有全局感受野,解决了RNN系列网络的信息衰减问题。同时,其特别设计的时态感知注意力机制能捕捉不同时间点的功率变化趋势,从而提高了对时间信息的利用效率。以河北省承德市某风电场中33台风电机组的历史运行数据为算例进行分析,并与其他模型进行对比实验。

本研究作为综合能源仿真平台研发项目的重要组成部分,聚焦于风电机组集群时空相关性的挖掘与应用,旨在构建高精度的风电功率预测模型,为综合能源系统仿真与新能源电力系统的稳定运行提供关键技术支撑。

## 1 问题描述与知识预备

本节首先对研究问题进行陈述,然后简要回顾基础算法模块。

### 1.1 问题定义

**定义 1**(风电机组关系图):

将多台风电机组之间的相关关系定义为一张无向图 $G = (V, E)$ ,其中 $V$ 表示节点集合,每个节点对应一台风电机组,节点总数为 $N$ ;  $E$ 表示边集合,边用于刻画风电机组之间的关联关系,边的数量为 $M$ 。

**定义 2**(历史 SCADA 数据):

在时刻  $t$ , 风电机组集群的 SCADA 数据矩阵定义为  $\mathbf{X}_t = (\mathbf{x}_{t,1}, \mathbf{x}_{t,2}, \dots, \mathbf{x}_{t,N})^T \in \mathbb{R}^{N \times C}$ , 其中  $\mathbf{x}_{t,v} \in \mathbb{R}^C$  表示第  $v$  台风电机组在时刻  $t$  的特征向量,  $C$  为每个节点的特征维数。下标  $t$  和  $v$  分别表示时间索引和机组编号, 均为无量纲标识符。

**问题定义** (风电集群功率预测问题):

给定风电机组集群在过去  $T_h$  个历史时刻的 SCADA 数据序列  $\mathbf{X} = (\mathbf{X}_{t-T_h+1}, \mathbf{X}_{t-T_h+2}, \dots, \mathbf{X}_t) \in \mathbb{R}^{N \times C \times T_h}$ , 其目标是预测未来  $T_p$  个时刻的风电机组集群功率输出  $\mathbf{Y} = (\mathbf{X}_{t+1}, \mathbf{X}_{t+2}, \dots, \mathbf{X}_{t+T_p}) \in \mathbb{R}^{N \times 1 \times T_p}$ , 其中  $T_h$  和  $T_p$  分别表示历史输入序列长度和预测时间步长; 由于风电功率预测的目标变量仅为功率值, 因此模型输出张量在特征维度上的大小为 1。

## 1.2 自注意力机制

自注意力机制将一次查询和一组键值对映射至向量格式的输。输出为各值的加权总和, 而每个值的权重则由对应的键与查询共同决定。每个权重代表了查询与特定键值对之间的关系强度<sup>[18]</sup>。其具体计算过程如图 1 所示。

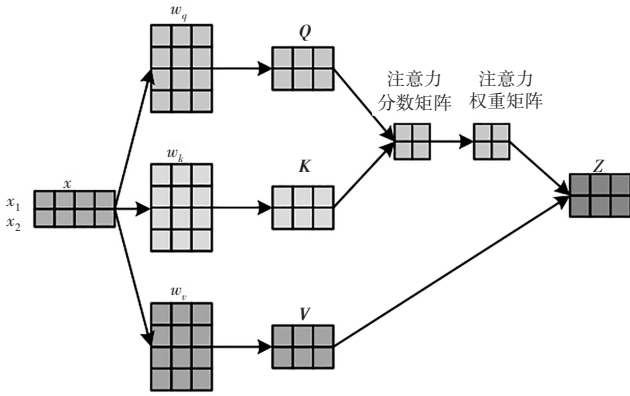


图 1 注意力机制

Fig. 1 Attention mechanism

输入序列为  $\mathbf{X} = [x_1, x_2, \dots, x_n]$ ,  $\mathbf{X}$  首先通过 3 个不同的全连接网络得到查询向量  $\mathbf{Q}$ 、键向量  $\mathbf{K}$  和值向量  $\mathbf{V}$ , 对于每个元素  $x_i$ , 其计算公式如下, 即

$$\mathbf{Q}_i = \mathbf{W}_q \cdot x_i \quad (1)$$

$$\mathbf{K}_i = \mathbf{W}_k \cdot x_i \quad (2)$$

$$\mathbf{V}_i = \mathbf{W}_v \cdot x_i \quad (3)$$

其中,  $\mathbf{Q}_i$ 、 $\mathbf{K}_i$  和  $\mathbf{V}_i$  分别表示输入序列中第  $i$  个输入元素  $x_i$  映射得到的查询 (Query)、键 (Key) 和值 (Value) 向量; 下标  $i$  表示序列中的第  $i$  个位置, 而下标  $q, k, v$  用于区分不同类型的线性映射权重, 对应查询、键和值 3 种特征空间  $\mathbf{W}_q, \mathbf{W}_k$  和  $\mathbf{W}_v$  为可学习的参数矩阵。然后使用缩放点积的方式计算每一对元素  $x_i$  和  $x_j$  之间的注意力分数, 其计算公式为

$$\text{score}(\mathbf{Q}_i, \mathbf{K}_j) = \frac{\mathbf{Q}_i \cdot \mathbf{K}_j}{\sqrt{d_k}} \quad (4)$$

其中,  $d_k$  是键向量的维度。

然后通过 softmax 函数使得注意力分数转为化 0 和 1

之间的数值且和为 1, 进而得出注意力权重  $w_{ij}$ , 最后将每个元素的值向量与其对应的注意力权重相乘, 然后求和, 得到最终的输出为

$$z_i = \sum_{j=1}^n w_{ij} V_j \quad (5)$$

通过注意力机制生成的新向量  $\mathbf{Z}$  不仅包含了当前时间点的信息, 还融合了其在整个时间序列中的上下文信息。通过这种方式, 模型能够更准确地理解每个时间点特定序列中的意义和作用, 从而提高对时间序列数据的整体理解和分析能力。

## 1.3 多头注意力机制

1.2 节详细介绍了自注意力机制的计算方法, 其可简要概括为

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right) \mathbf{V} \quad (6)$$

多头自注意力机制是实践中最广泛使用的自注意力形式, 能同时关注来自不同表示子空间的信息。其基本操作是 1.2 节中介绍的缩放点积注意力。这一机制首先对查询、键和值进行线性投影到不同的表示子空间, 随后并行执行注意力函数。最后, 将这些输出连接并进行进一步投影, 以产生最终输出, 其计算方式为

$$\text{MultiHeadSelfAttention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \bigoplus (\text{head}_1, \dots, \text{head}_h) \mathbf{W}^O \quad (7)$$

$$\text{head}_j = \text{Attention}(\mathbf{Q}\mathbf{W}_j^Q, \mathbf{K}\mathbf{W}_j^K, \mathbf{V}\mathbf{W}_j^V) \quad (8)$$

其中,  $h$  表示注意力头的数量,  $\mathbf{W}_j^Q, \mathbf{W}_j^K, \mathbf{W}_j^V$  分别是应用于  $\mathbf{Q}, \mathbf{K}, \mathbf{V}$  的投影矩阵,  $\mathbf{W}^O$  是用于最终输出的投影矩阵。

通过将注意力机制分割成多个头, 模型可以在不同的表示子空间中学习输入数据的不同特征。这种多头结构使得模型能够捕捉序列中更丰富的信息。

## 1.4 图卷积网络

图卷积网络 (graph convolutional network, GCN) 将传统的卷积操作从结构化数据推广到图形, 并且它们能够捕捉图中隐藏的非结构化模式。GCN 的总体思路是通过在节点之间交换信息来学习节点表示。具体来说, 给定一个节点, GCN 首先聚合其邻接表示以生成该节点的中间表示, 然后它们使用线性投影和非线性激活来转换聚合的表示。其计算公式为

$$\text{GCN}(\mathbf{Z}_i^{(l-1)}) = \sigma(\mathbf{A}\mathbf{Z}_i^{(l-1)} \mathbf{W}^{(l)}) \quad (9)$$

其中  $\mathbf{Z}_i^{(l-1)} \in \mathbb{R}^{N \times d_{\text{model}}}$ ,  $\mathbf{W}^{(l)} \in \mathbb{R}^{d_{\text{model}} \times d_{\text{model}}}$  和  $\sigma$  分别代表节点表示、投影矩阵和非线性激活函数。  $\mathbf{A} \in \mathbb{R}^{N \times N}$  表示节点之间的交互关系, 其计算公式为

$$\mathbf{A} = \tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \quad (10)$$

其中,  $\tilde{\mathbf{A}}$  代表图的邻接矩阵,  $\tilde{\mathbf{D}}$  为度矩阵,  $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$ 。

## 2 问题分析与所提模型

### 2.1 问题分析

风电集群的功率预测是一个典型的时空序列预测问

题。由于各风电机组的地理位置相近,根据地理学的第一定律<sup>[19]</sup>,它们的运行状态具有相似性,即空间相关性。对每台风电机组单独观察,其 SCADA 数据是多条时间序列数据,每条时间序列中蕴含着时间信息。如何充分运用数据中包含的时空信息,是提高风电功率预测精度的关键所在。

近年来,Transformer 系列模型在自然语言处理和计算机视觉等领域的多项任务中展现了显著的成效。其突出优势之一是在处理远程依赖关系方面的卓越能力,这一特性使得 Transformer 模型特别适合于时间序列分析任务。

本文将探究 Transformer 模型在风电集群功率预测中的应用。值得注意的是,Transformer 模型具有高度的灵活性,能够通过整合额外的空间信息模块来进行时空建模。这种增强为模型提供了空间上下文信息,从而在风电功率预测等复杂任务中发挥关键作用。通过这种方法,我们可以更准确地捕捉和分析风电集群在时间和空间上的动态变化,为风电功率预测提供更为精确的数据支持。

### 2.2 模型总体架构

本文所提出的时空感知图卷积注意力模型是一种时空预测模型<sup>[20]</sup>,示意图如图 2 所示。

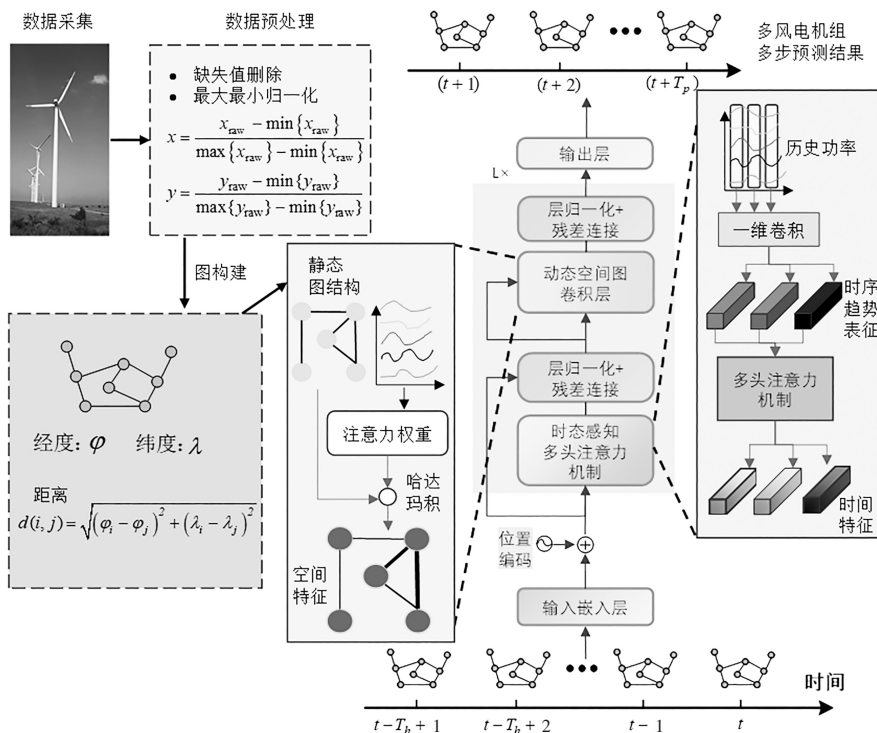


图 2 所提方法示意图

Fig. 2 Schematic diagram of the proposed method

如图 2 所示,首先对采集到的 SCADA 数据进行数据预处理,删除缺失数据,之后通过最大-最小归一化方法进行标准化,确保所有特征处于统一的尺度上,从而有助于模型的学习过程。

接下来进行图构建,以初步建模不同风电机组之间的空间依赖关系。

在预测模型中,使用长度为  $H$  历史时间步长的时空序列数据作为模型输入,模型的输出为未来  $P$  个时间步的风电集群功率预测结果。输入数据首先经过嵌入层的处理,以便将其转换为高维度的表示,同时融合了可学习的位置编码,增强了输入数据的空间信息特征。在此基础上,数据进入时态多头自注意力机制模块,用于深度挖掘其时间序列的依赖关系。随后,通过残差连接和层归一化操作,不仅保留了原始数据的重要特征,还显著提高了模型训练的稳定性。紧接着,处理后的数据被传递至动态空间图卷积层,以提炼风机之间的复杂空间相关性。最终,类似的残差连接与层归一化操作同样用于空间模块,确保

了模型的收敛性及输出结果的精准性。

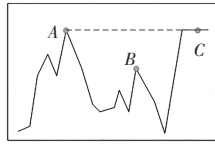
值得注意的是,本模型的核心创新在于时态感知多头自注意力机制和动态空间图卷积层。这两大模块专为风电集群功率预测任务中的关键挑战而设计。接下来,我们将详细阐述这两个模块的原理与应用。

### 2.3 时态感知多头自注意力机制

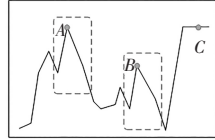
传统 Transformer 模型通过多头自注意力机制,对每个时间点的数值进行两两计算以获取注意力权重。这种方法赋予了 Transformer 在长时间序列建模中的显著优势。然而,由于其注意力计算仅依赖于单个时间点的数值,该模型难以有效捕捉时间序列中的相似趋势。图 3 对此问题进行了具体阐释。

如图 3 所示,  $A$ 、 $B$  和  $C$  代表了不同时间段的 3 个数据点。使用传统自注意力机制时,可能会错误地将  $A$  点和  $C$  点相匹配,原因是它们在逐点数值上看起来相同。但实际上,  $A$  点和  $C$  点的局部趋势是不同的:  $A$  点正处于波动的高峰,而  $C$  点的功率值较为平稳。简而言之,  $A$  点和  $C$  点的近

期历史趋势存在明显差异。因此,当在这些数值数据上使用多头自注意力时,可能会错误地判定数据对之间的关联性。这种误判可能导致序列表示不准确,从而影响到最终预测的效果。



(a) 传统注意力机制



(b) 时间趋势感知注意力机制

图3 不同注意力计算方式应用于时间序列数据效果对比

Fig. 3 Comparison of different attention calculation methods applied to time series data

为了克服传统多头注意力机制中无法识别局部趋势的问题,本文设计了时态感知多头注意力机制。此机制通过一维卷积操作提取原有多头注意力机制的查询和键的时间特征,融合每个点与其邻近时间点的信息。接着,使用多头注意力机制处理这些融合了邻域信息的数据,使模型能够正确匹配图中的A点和B点。具体计算公式为

$$\text{TPAM}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \oplus (\text{TPhead}_1, \dots, \text{TPhead}_h) \mathbf{W}^o \quad (11)$$

$$\text{TPhead}_j = \text{Attention}(\Phi_j^q \times \mathbf{Q}, \Phi_j^k \times \mathbf{K}, \mathbf{V} \mathbf{W}_j^v) \quad (12)$$

其多头注意力计算方式与式(7)相同,区别在于式(9)对单个注意力头使用注意力机制时,对查询和键进行一维卷积操作,该操作可以对数据进行滤波去噪<sup>[21-23]</sup>并融合每个时间点的上下文信息生成新的表示。 $\Phi_j^q, \Phi_j^k$ 代表卷积核参数。

#### 2.4 动态空间图卷积层

为捕捉多台风电机组的空间相关性,本文设计了动态空间图卷积层。图卷积网络将传统卷积运算从结构化数据扩展至图形数据,有效地捕捉图中的非结构化模式。其主要方法是通过节点间的信息交换来学习节点表示。1.4节对其基本原理已进行了简要介绍。

风电集群的空间相关性包括静态相关性和动态相关性两部分。静态相关性由风电机组间的地理位置距离决定,它是构建图结构的基础。在大多数应用中,邻接矩阵中的元素是使用基于距离的函数来计算的,具体地,使用如下函数计算风电集群中各机组的带权邻接矩阵<sup>[24]</sup>,即

$$\mathbf{A}_N(i, j) = \begin{cases} e^{-d_{ij}}, & e^{-d_{ij}} > \frac{1}{N} \sum_{i=1, j=1}^N e^{-d_{ij}} \\ 0, & \text{其他} \end{cases} \quad (13)$$

其中,  $d_{ij}$  表示第  $i$  台风电机组与第  $j$  台风电机组之间的欧式距离;  $\frac{1}{N} \sum_{i=1, j=1}^N e^{-d_{ij}}$  代表距离权重的平均值,当两台风电机组的距离权重大于风电机组集群平均距离权重时,视为

两台机组间存在边。

上文根据风电机组间的欧氏距离建立了带权邻接矩阵,然而,传统图卷积运算是时不变的,即对于给定图  $G$ ,相应的权重矩阵  $\mathbf{A}$  是一个常数。这对于风电功率预测而言是个限制,因为风电机组间的相关性可能会随着时间变化,受风速和风向的影响。如果简单地将图卷积网络应用于风电集群功率预测,它无法捕捉这种动态相关性。为解决这个问题,本文提出了动态空间图卷积层,它能自适应地调整带权邻接矩阵中节点间的连接强度。

动态空间图卷积层的主要思想是使用自注意力机制来动态计算不同风电机组之间的空间相关强度。将时态感知多头自注意力机制模块的输出记为  $\mathbf{Z}_t^{(l-1)} \in \mathbb{R}^{N \times d_{\text{model}}}$ ,在  $t$  时刻空间注意力权重矩阵  $\mathbf{S}_t$  的计算公式如下,即

$$\mathbf{S}_t = \text{softmax} \left( \frac{\mathbf{Z}_t^{(l-1)} \mathbf{Z}_t^{(l-1)\top}}{\sqrt{d_{\text{model}}}} \right) \in \mathbb{R}^{N \times N} \quad (14)$$

$\mathbf{S}_t$  的元素  $S_{ij}$  代表第  $i$  台风电机组和第  $j$  台风电机组的相关性强度,值越大表示相关性越强。之后利用  $\mathbf{S}_t$  使用逐元素点积运算调整静态权重矩阵  $\mathbf{A}$ ,具体计算方式如下,即

$$\mathbf{X}_t^{(l)} = \sigma((\mathbf{A} \odot \mathbf{S}_t) \mathbf{Z}_t^{(l-1)} \mathbf{W}^{(l)}) \quad (15)$$

本文使用的动态空间图卷积层,基于输入  $\mathbf{Z}_t^{(l-1)}$  计算动态的带权邻接矩阵,并通过聚合邻居信息捕捉不同风电机组间的动态空间相关性。

本文中图结构的构建方法结合了先验知识和数据驱动方法,既考虑了由风电机组间地理距离决定的静态相关性,也考虑了在不同运行工况下的时变特性。

## 3 算例分析

### 3.1 数据集介绍

本文选取中国承德市某风电场 33 台风电机组作为研究对象,其空间分布如图 4 所示。

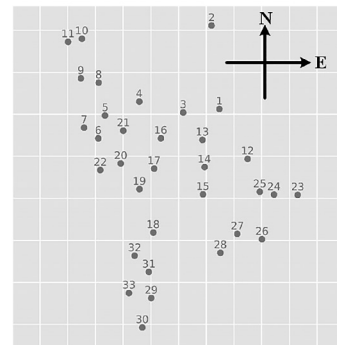


图4 承德市某风电机组集群

Fig. 4 Comparison of wind turbine cluster in Chengde

如上图所示,33 台风电机组呈不规则分布。

选取 2023 年 4 月 27 日至 2023 年 7 月 23 日的数据作为数据集,时间分辨率为 5 min。数据包括各台风电机组的风速和功率。该数据集共计 25 167 帧,按照 8 : 1 : 1 的比例分为训练集、验证集和测试集。

### 3.2 对比模型

为了衡量本文所提模型的性能,选取包括本文所提模型在内的5种算法进行对比,其中包括GRU<sup>[25]</sup>,LSTM<sup>[26]</sup>,Transformer<sup>[18]</sup>,TPTransformer,STTransformer。

GRU是一种先进的循环神经网络(RNN)架构,主要用于处理序列数据。与传统的RNN相比,GRU通过引入更新门和重置门来解决长期依赖问题,从而有效地捕捉序列中的时间动态。更新门帮助模型决定在新信息到达时保留多少过去的信息,而重置门则决定忽略多少过去的信息。这使得GRU在处理时间序列预测任务时表现出色。

LSTM是另一种广泛应用的RNN架构,专门设计用于避免标准循环神经网络中的长期依赖问题。LSTM通过引入3个门(输入门、遗忘门和输出门)和一个细胞状态,能够在长序列中有效地保持和传递信息。这些门的结构允许网络有选择地保留或忽略信息,使其在许多序列建模任务中非常有效。

Transformer是一种基于自注意力机制(self-attention)的深度学习模型,于2017年由Vaswani等提出。它摒弃了传统的循环和卷积网络结构,通过并行处理序列的每个元素,显著提高了训练效率。自注意力机制允许模型直接关注输入序列的任何部分,从而更有效地捕捉远距离依赖。Transformer已成为自然语言处理领域的核心,广泛应用于时间序列任务。

时态感知Transformer(temporal perceptive transformer, TPTransformer)对传统Transformer网络的多头自注意力部分进行改进,使得模型能够捕捉时间序列趋势变化中的信息,从而增强了对时间信息的利用能力。

时空感知Transformer(spatial and temporal perceptive transformer, STTransformer)在TPTransformer的基础上加入了动态时空图卷积网络。这一改进有助于提取不同风电机组间的空间相关性,并增强模型对空间信息的利用能力。

对比模型包括横向对比实验和纵向消融实验。横向对比实验通过与GRU、LSTM模型比较,证明了本文提出模型的优越性。纵向消融实验则通过逐步添加各改进模块,展现了每个模块的贡献。

### 3.3 任务说明及评价指标

本研究的目标是进行风电集群功率预测,模型输入包括33台风电机组过去24个时间步的风速和功率数据,预测目标是这33台风电机组未来连续6个时间步的功率值。

本文选择平均绝对误差(mean absolute error, MAE)、均方根误差(root mean square error, NRMSE)和决定系数 $R^2$ 作为单步预测误差评价指标。这3项指标在评估风电功率预测模型的性能时具有较好的表现,计算公式为

$$MAE = \frac{1}{Q} \sum_{q=Q} |x_q(i) - \hat{x}_q(i)| \quad (16)$$

$$RMSE = \sqrt{\frac{1}{Q} \sum_{q=Q} (x_q(i) - \hat{x}_q(i))^2} \quad (17)$$

$$R^2 = 1 - \frac{\sum_{q=Q} (x_q(i) - \hat{x}_q(i))^2}{\sum_{q=Q} (x_q(i) - \bar{x})^2} \quad (18)$$

其中, $x_q$ 为真实值, $\hat{x}_q$ 为预测值, $i$ 为预测步长, $Q$ 为测试集样本数量。

### 3.4 预测结果分析

为了直观展示不同预测模型的性能,对第15号机组的单步预测结果进行了分析,并绘制了相应的曲线如图5所示。

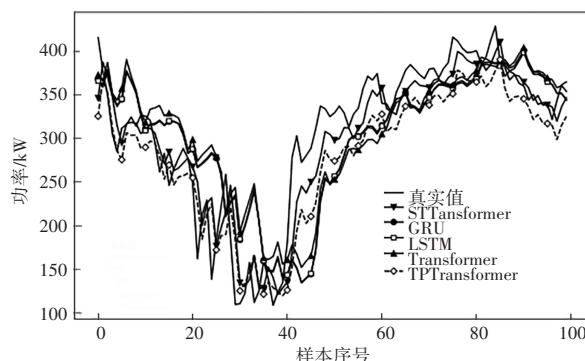


图5 15号机组各模型预测结果

Fig. 5 Prediction results of different models for turbine No. 15

图5中展示了多个预测模型在风电机组集群功率数据集上的单步预测性能,其中横轴表示样本序列,纵轴表示功率。黑色曲线代表真实值,其他曲线则分别对应STFormer、GRU、LSTM、TF和TPTF等模型的预测结果。

可以看出,各模型整体上均能较好地捕捉到真实值的变化趋势,但在局部细节的拟合能力上存在差异。本文提出的STFormer模型在大幅波动区间表现出更强的趋势跟随能力,与真实值更加接近。相比之下,LSTM、GRU、TF和TPTF模型在部分点上略有偏差,但仍能较好地反映数据的波动特征。

综合来看,STFormer模型的预测曲线与真实值的吻合度最高,其在捕捉数据的空间相关性和时序依赖性方面展现出卓越的能力。相较之下,其他模型在应对数据剧烈波动时的表现略显不足。

上述内容直观呈现了单台风电机组的单步预测结果。进一步而言,本文的研究对象为风电机组集群,因此为全面评估风电场的预测性能,后文将对其整体预测精度进行详细分析。

由于本文是时空预测任务,本节将从时、空两个角度对比各模型预测性能。首先分析时间维度,随着预测间隔增加,各模型误差曲线如图6所示,未经特殊设计的GRU、LSTM、Transformer模型在进行复杂的风电集群功率任务时预测误差始终较大,TPTransformer在Transformer模型的基础上对多头自注意力机制进行改进,提升了模型利用时序信息的能力,使模型在前4步预测的精度大大提高。本文所提出的STTransformer模型在TPTransformer模型的基础上将前馈神经网络替换为空间动态图卷积网络,

增强了模型利用空间信息的能力,进一步提升了预测精度。预测间隔的增加使任务更加困难,导致所有模型的预测性能下降。在大多数情况下,本文提出的模型表现最佳。相比其他模型,本文所提模型的 MAE 降低了 6.5%~

16.4%、RMSE 降低了 6.4%~15.3%。

表 1 给出了各预测间隔各模型的  $R^2$  数值,越接近 1 说明模型拟合效果越好。

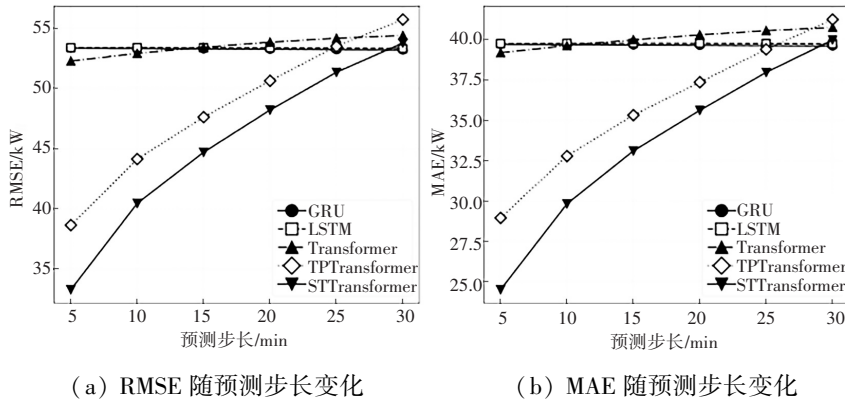


图 6 各模型预测误差随时间变化

Fig. 6 Temporal variation of prediction errors for different models

表 1 不同预测模型各预测间隔  $R^2$

Tab. 1  $R^2$  values for different prediction models at various prediction intervals

间隔/min	$R^2$				
	GRU	LSTM	Transformer	TpTransformer	STTransformer
5	0.703	0.700	0.704	0.851	<b>0.893</b>
10	0.703	0.700	0.696	0.807	<b>0.842</b>
15	0.703	0.700	0.690	0.777	<b>0.807</b>
20	0.703	0.699	0.684	0.749	<b>0.776</b>
25	0.703	0.699	0.678	0.722	<b>0.746</b>
30	0.704	0.699	0.674	0.691	<b>0.704</b>
平均	0.703	0.700	0.688	0.766	<b>0.795</b>

如表 1 所示,本文提出的模型在所有时间间隔预测中的决定系数值均最高,这突显了其优越性。观察表 1 可知,当预测间隔增加时,各模型均下降,当间隔增加到 30 min 时,各模型表现接近,为分析该情况,选取地理位置接近的两台风电机组,绘制其时滞相关系数如图 7 所示。

进一步降低。因此,对于超过 30 min 的功率预测问题,难度将会进一步增加。

绘制 33 台风电机组功率预测误差分布的箱线图如图 8 所示。由图 8 可知,本文提出的模型在风电场中 33 台风电机组的预测误差(包括最大值、最小值和中位数等指标)均达到最优,显示了该模型在控制单台机组预测误差方面的优越性能。

综合考虑全场风电机组的平均预测效果和每台机组的误差控制能力,本文提出的算法表现出更佳的预测效果。

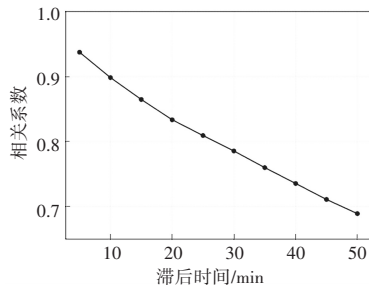


图 7 两台风电机组的时滞相关系数

Fig. 7 Correlation coefficients with time lag between two wind turbines

由图 7 可知,随着时间滞后的增加,两台风电机组的相关系数逐渐降低。在 30 min 后,相关性降至 0.8 以下,不再构成强相关性。此外,图中展示的是地理位置接近的两台风电机组。若风电机组之间的距离更远,其相关性将

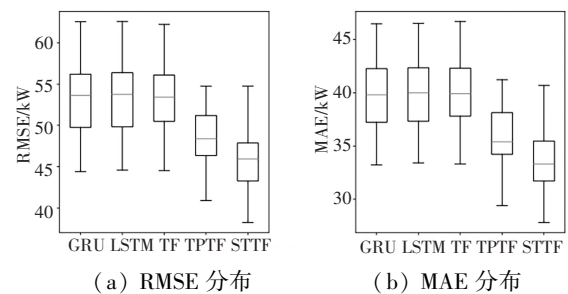


图 8 各风电机组预测误差分布

Fig. 8 Distribution of prediction errors for each wind turbine

## 4 结论

为应对风电机组集群功率预测中的挑战,本文提出了一种结合时空感知图卷积与注意力网络的创新预测模型。该模型的设计核心在于两个关键模块:空间相关性提取与时间相关性捕捉。

首先,在空间相关性提取方面,模型构建了一个空间动态图卷积网络,专门针对风电场的不规则布局进行优化。通过引入空间注意力机制,网络能够动态调整图邻接矩阵的权重,以适应不同工况条件下的变化。这种自适应机制增强了风电场内各风电机组间空间关系的表达能力,从而显著提升了预测精度。

其次,在时间相关性捕捉方面,本研究设计了一种具备全局与局部感知能力的时态感知注意力网络。与传统循环神经网络(RNN)架构不同,该网络通过时态感知机制有效缓解了信息衰减问题,并能够精准识别不同时间点上的功率变化趋势。这种方法显著提高了对时间序列信息的利用效率,增强了模型对风电功率动态变化的响应能力。

总体而言,本文提出的风电机组集群功率预测模型通过在空间和时间维度上的综合优化,实现了预测精度的显著提升。这一研究成果不仅有助于优化风电场的能源管理与调度,还能提升风电利用的稳定性和经济效益,为智能电网的发展提供了坚实支撑。

## 参考文献

- [1] SUN S, LIU Y, LI Q, et al. Short-term multi-step wind power forecasting based on spatio-temporal correlations and transformer neural networks [J]. *Energy Conversion and Management*, 2023, 283: 116916.
- [2] 郑培, 于立军, 侯胜亚, 等. 基于卡尔曼滤波修正的多步风电功率预测[J]. *热能动力工程*, 2020, 35(4): 235-241.
- [3] TAWN R, BROWELL J. A review of very short-term wind and solar power forecasting [J]. *Renewable and Sustainable Energy Reviews*, 2022, 153: 111758.
- [4] 刘长良, 吴家佳. 基于核极端学习机的短期风电功率预测研究[J]. *热能动力工程*, 2017, 32(1): 95-100.
- [5] 殷豪, 董朕, 孟安波. 基于VMD-SE-IPSO-BNN的超短期风电功率预测[J]. *电测与仪表*, 2018, 55(2): 45-51.
- [6] GU B, HU H, ZHAO J, et al. Short-term wind power forecasting and uncertainty analysis based on FCM-WOA-ELM-GMM [J]. *Energy Reports*, 2023, 9: 807-819.
- [7] LIU Y, YANG T, TENG W, et al. Day-ahead energy power forecasting under carbon neutrality based on XGBoost [C]//2023 8th Asia Conference on Power and Electrical Engineering (ACPEE). Tianjin: IEEE, 2023: 2044-2053.
- [8] CHEN H, BIRKELUND Y, ZHANG Q. Data-augmented sequential deep learning for wind power forecasting [J]. *Energy Conversion and Management*, 2021, 248: 114790.
- [9] 侯倩, 郝晓光, 金飞, 等. 基于混合深度学习的风电功率预测及一次调频应用[J]. *热能动力工程*, 2023, 38(10): 167-175. DOI: 10.16146/j.cnki.rndlgc.2023.10.020.
- [10] 陈海鹏, 李赫, 阚天洋, 等. 考虑风电时序特性的深度小波-时序卷积网络的超短期风功率预测[J]. *电网技术*, 2023, 47(4): 1653-1665.
- [11] 向玲, 金子皓, 李林春. 基于注意力机制的IWOA-BiGRU超短期风电功率预测[J]. *华北电力大学学报(自然科学版)*, 2024, 51(4): 87-93, 102.
- [12] 易善军, 王汉军, 向勇, 等. 基于集成多尺度LSTM的短时风功率预测[J]. *重庆大学学报(自然科学版)*, 2021, 44(7): 75-81.
- [13] LI N, DONG J, LIU L, et al. A novel EMD and causal convolutional network integrated with transformer for ultra short-term wind power forecasting [J]. *International Journal of Electrical Power & Energy Systems*, 2023, 154: 109470.
- [14] TASICARA OGLU A. Evaluation of spatio-temporal forecasting methods in various smart city applications [J]. *Renewable and Sustainable Energy Reviews*, 2018, 82: 424-435.
- [15] 王晨, 寇鹏. 基于卷积神经网络和简单循环单元集成模型的风电场多风机风速预测[J]. *电工技术学报*, 2020, 35(13): 2723-2735.
- [16] ZHU Q, CHEN J, SHI D, et al. Learning temporal and spatial correlations jointly: a unified framework for wind speed prediction [J]. *IEEE Transactions on Sustainable Energy*, 2020, 11(1): 509-523.
- [17] ZHU X, LIU R, CHEN Y, et al. Wind speed behaviors feather analysis and its utilization on wind speed prediction using 3D-CNN [J]. *Energy*, 2021, 236: 121523.
- [18] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. *Advances in Neural Information Processing Systems*, 2017, 30: 5998-6008.
- [19] ZHU A X, LU G, LIU J, et al. Spatial prediction based on Third Law of Geography [J]. *Annals of GIS*, 2018, 24(4): 225-240.
- [20] ZHANG J, LIU D, LI Z, et al. Power prediction of a wind farm cluster based on spatiotemporal correlations [J]. *Applied Energy*, 2021, 302: 117568.
- [21] 刘长良, 赵陆阳, 王梓齐, 等. 基于时空注意力-Seq2Seq模型的多风电机组多步风速预测算法[J]. *太阳能学报*, 2023, 44(8): 420-429.
- [22] 王斌, 王栓虎, 张文慧. 基于多时间尺度的风电场无功补偿装置协调控制研究[J]. *自动化技术与应用*, 2025, 44(5): 52-56.
- [23] 廖卉莲, 李帅兵, 彭华坤, 等. 大惯量风电机组最优发电策略研究[J]. *自动化技术与应用*, 2025, 44(12): 10-13.
- [24] CHUNG J, GULCEHRE C, CHO K H, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling [DB/OL]. (2014-12-11). <https://arxiv.org/abs/1412.3555>.
- [25] GERS F A, SCHMIDHUBER J, CUMMINS F. Learning to forget: continual prediction with LSTM [J]. *Neural computation*, 2000, 12(10): 2451-2471.