DOI: 10.3969/j.issn.2097-583X.2026.04.003　　文章编号：2097-583X(2026)04-0533-09

面向风电功率预测的多尺度自适应图卷积方法

吕晓飞[1]，黄全全[1]，韩华春[2*]，王　智[1]，曾雨竹[1]

(1. 国网电力科学研究院有限公司 电网运行风险防御技术与装备全国重点实验室，江苏 南京 211106；
2. 国网江苏省电力有限公司电力科学研究院，江苏 南京 211103)

摘　要：风电功率预测面临风况不可预测性与随机性的双重挑战。现有基于图神经网络的预测模型通常仅从单一角度构建图结构，难以全面准确地刻画风电场间复杂的空间相关性。针对该问题，提出Dual-MSGNN模型，通过融合频域分析与自适应图卷积，捕获多时间尺度和空间尺度下风电序列的动态相关性。该方法从时间和空间两个维度构建图结构：一方面将时间序列分解为不同尺度，并引入自注意力机制提取序列内的时间依赖关系；另一方面采用自适应图卷积层，自动学习各时间尺度下序列间的空间相关模式。在真实风电数据集上的实验结果表明，Dual-MSGNN模型在预测精度和鲁棒性方面均具有显著优势，验证了其有效性。

关键词：风电功率预测；多尺度图卷积；自适应图卷积

中图分类号：TM614　　**文献标志码：A**　　**开放科学(资源服务)标识码(OSID)：**

Multi-scale adaptive graph convolutional method for wind power forecasting

Lü Xiaofei[1], Huang Quanquan[1], Han Huachun[2*], Wang Zhi[1], Zeng Yuzhu[1]

(1. State Key Laboratory of Technology and Equipment for Defense Against Power System Operational Risks, State Grid Electric Power Research Institute, Nanjing 211106, China;
2. State Grid Jiangsu Electric Power Co., Ltd. Research Institute, Nanjing 211103, China)

Abstract: Wind power forecasting faced the dual challenges of unpredictability and randomness in wind conditions. Existing forecasting models based on graph neural networks were usually constructed from only a single perspective to build graph structures, which made it difficult to comprehensively and accurately characterize the complex spatial correlations among wind farms. To address this issue, a Dual-MSGNN model was proposed, which integrated frequency domain analysis and adaptive graph convolution to capture the dynamic correlations of wind power series across multiple temporal and spatial scales. In this method, graph structures were constructed from both temporal and spatial dimensions: on the one hand, the time series were decomposed into different scales, and a self-attention mechanism was introduced to extract intra-sequence temporal dependencies; on the other hand, an adaptive graph convolution layer was adopted to automatically learn inter-sequence spatial correlation patterns under each time scale. Experimental results on real-world wind power datasets demonstrated that the Dual-MSGNN model achieved significant advantages in prediction accuracy and robustness, thus validating its effectiveness.

Key words: wind power forecasting; multi-scale graph convolution; adaptive graph convolution

0　引言

近年来，我国风力发电装机规模快速增长，风电已成为重要的电力能源。受风能固有的随机性与波动性影响，准确的风电功率预测对保障电力系统安全稳定运行至关重要。预测误差不仅会影

收稿日期：2026-06-29；**修回日期：**2026-07-28；*.**通信联系人，**E-mail：598543427@qq.com

基金项目：国家重点研发计划项目(2022YFB2402700)

作者简介：吕晓飞(1986－)，男，山东济南人，高级工程师，硕士，主要从事新能源并网控制技术研究；韩华春(1988－)，女，山东济宁人，正高级工程师，博士，主要从事新能源发电功率预测技术研究。

引用格式：吕晓飞，黄全全，韩华春，等. 面向风电功率预测的多尺度自适应图卷积方法[J]. 信阳师范大学学报(自然科学版)，2026，39(4)：533-541.

Lü Xiaofei, Huang Quanquan, Han Huachun, et al. Multi-scale adaptive graph convolutional method for wind power forecasting[J]. Journal of Xinyang Normal University (Natural Science Edition), 2026, 39(4): 533-541.

响电网调度计划的制定,还可能引发电量失衡、线路过载等问题,从而增加电网运行风险[1]。因此,实现高精度风电功率预测(Wind Power Forecasting,WPF)是提升电网安全运行水平的重要基础。

目前,国内外学者围绕风电功率预测已开展了大量研究,但现有方法仍存在一定局限[2]。文献[3]采用线性回归方法建立输入与输出之间的关系,但难以有效刻画风电功率的复杂非线性特征。文献[4]采用支持向量回归(Support Vector Regression,SVR)等机器学习方法,在一定程度上提高了预测性能,但其应用主要局限于超短期预测场景。随着深度学习技术的快速发展,卷积神经网络(Convolutional Neural Network,CNN)被用于提取空间特征,长短时记忆网络(Long Short-Term Memory,LSTM)擅长时序建模[5],Transformer则能够捕获长距离时序依赖,进一步提高预测精度[6-8],然而该类模型对时间噪声较为敏感[9]。近年来,风电场空间相关性与风电不确定性问题日益受到关注。文献[10]融合空间相关性与Stacking集成学习以提升预测性能;文献[11]结合深度误差反馈学习与注意力机制,提高了短期预测精度;文献[12-13]将风电时空相关性和不确定性引入暂态稳定约束最优潮流分析,为时空建模提供了新的思路。然而,现有图神经网络方法大多仅基于单一先验知识构建图结构,难以充分表征风电场之间复杂的时空关联特征,预测精度仍有进一步提升空间。已有研究表明,多模型融合策略可有效提高预测性能[14]。

针对现有模型在时空建模维度单一、特征挖掘不充分等方面的问题,提出一种双分支并行的Dual-MSGNN模型,该模型由多关系图构建模块与多关系图计算模块组成。为充分捕捉风电场间复杂的时空关联关系,多关系图构建模块分别从空间和时间两个维度构建空间感知图、语义感知图及时间子序列全连接感知图。在多关系图计算模块中,首先对原始时间序列进行快速傅里叶变换,并将其映射至多个显著的时间尺度相关空间,从而提取不同尺度下的时序相关性;同时,构建全连接图卷积器,沿时间子序列部署图神经网络以捕获时空信息;此外,引入多头自注意力机制,进一步强化对内部时序模式的建模能力。本文的主要贡献可概括如下:

1)通过深入分析风电时序数据特性,揭示了风电场机组间存在的复杂时空耦合规律,并据此设计Dual-MSGNN模型,有效挖掘和表征风电数据中的时空信息。

2)为同时解决风电场数据中的空间相关性和时间依赖性问题,分别构建了空间感知图、语义感知图及全连接感知图,全面反映风电场的时空特征。

3)引入多头自注意力机制与自适应图卷积网络的组合结构,通过这两种网络结构的协同作用,实现多个时间尺度及多个时间序列上并行搭建图卷积网络。

4)在真实数据集上进行的大量实验结果表明,Dual-MSGNN模型在风电功率预测任务中持续优于现有深度学习模型,验证了其有效性和优越性。

1　方法介绍

1.1　总体介绍

本文提出的风电功率预测整体框架Dual-MSGNN如图1所示,主要由两个核心模块构成:多关系图构建模块与多关系图计算模块。

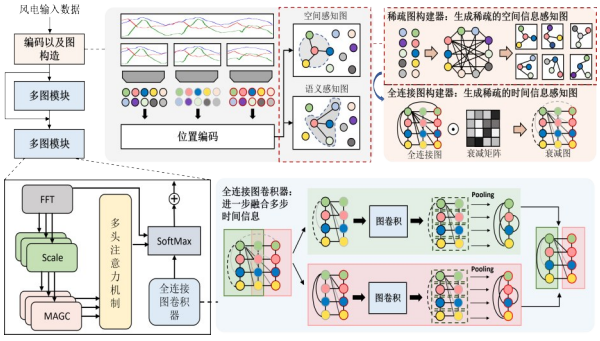


图1　所提方法框架

Fig. 1　Framework of the proposed method

具体而言,多关系图构建模块接收原始输入数据,分别生成稀疏的空间信息感知图与稀疏的时间信息感知图。这两类图结构能够有效捕捉长序列时间信息及多变量间复杂的空间耦合关系。在多关系图计算模块中,数据通过两条并行的计算路径进行特征提取。其中,空间信息感知图数据首先输入多尺度感知模块,经多尺度自适应图卷积层处理,以揭示不同序列间的相关性;同时,引入多头注意力机制对多尺度特征进行加权融合。另一条路径中,时间信息感知图数据则通过

全连接图卷积器进行计算，实现不同时间子序列上图神经网络的部署与时空信息提取。最终，两条路径的输出经SoftMax函数进行融合，生成预测结果。

1.2 多关系图构建模块

风电场间的空间相关性是影响风电功率预测模型性能的关键因素之一。为充分挖掘风电场站点之间的潜在关联，现有基于图神经网络的预测方法通常借助图结构建模空间依赖关系。然而，并非所有风电场节点之间均存在显著的空间相关性，若将所有节点两两连接构造完全图，不仅会引入大量冗余信息，还可能增加模型过拟合的风险。因此，如何构建稀疏且高效的空间感知图结构，成为提升模型泛化能力与预测精度的关键问题。

1.2.1 稀疏图构建

风速是影响风电功率的关键气象因素，因此在图结构的构建过程中，不仅需考虑风电场间的欧氏距离，还应通过波动互相关系数量化节点间的关联强度。从概率统计角度来看，相关性较强的风电场在风电功率概率分布上应具有较高的相似性；同时，地理位置等物理因素也使得相邻风电场间的相互影响更为显著，故地理距离和位置信息亦不可忽视。

为此，首先计算风电场之间的欧氏距离（实际地理距离），并为每个节点选取Top-K近邻构建空间距离图，记为$G^D=(\nu,\varepsilon,A^D)$；其次，基于风速变化序列的相似性，度量节点间的语义关系，构建风速语义感知图，并提取Top-K最相关节点作为语义邻居，记为$G^N=(\nu,\varepsilon,A^N)$。通过融合上述两类图结构，模型可同步感知物理位置邻近性与风电场动态变化特征，从而提升风电功率预测的准确性与鲁棒性。以下分别给出两种图矩阵的具体构建方式。

（1）空间距离图矩阵

在空间距离图矩阵的构建中，主要依据风电场之间的物理距离。鉴于距离是影响风电出力相关性的重要因素，基于地理位置构建距离图矩阵

$$A^D=\begin{pmatrix}1&\omega_{12}^D&\cdots&\omega_{1N}^D\\ \omega_{21}^D&1&\cdots&\omega_{2N}^D\\ \vdots&\vdots&\ddots&\vdots\\ \omega_{N1}^D&\omega_{N2}^D&\cdots&1\end{pmatrix}\in\mathbf{R}^{N\times N},\tag{1}$$

式中：$\omega_{ij}^D(i,j=1,2,\cdots,N)$表示节点$i$与节点$j$之间的距离权重，

$$\omega_{ij}^D=\begin{cases}\exp\left(-\dfrac{D_{ij}^2}{\sigma_D^2}\right),\ \exp\left(-\dfrac{D_{ij}^2}{\sigma_D^2}\right)\geqslant\xi,\\ 0,\qquad\text{其他},\end{cases}\tag{2}$$

D_{ij}为风电场i与j之间的地理距离；σ_D为稀疏权重因子，用于控制矩阵整体的稀疏程度；ξ为稀疏超参数，用于决定矩阵中零元素的数量。

（2）语义感知图矩阵

在语义感知图矩阵的构建中，选取与目标节点最相关的邻居节点作为语义邻居，以增强关键节点间的关联性并降低过拟合风险。采用风电功率概率分布的相似性来度量节点间的语义关联，并以Jensen-Shannon（JS）散度作为相似性度量指标。由于风电功率数据为连续随机变量，在计算概率分布前需进行离散化处理。具体地，将风电功率值按升序排列，均匀划分为K个子区间，统计各子区间内样本数占总样本数的比例，从而构建风电功率的概率分布：

$$P_i=\{p_i(1),\cdots,p_i(s),\cdots,p_i(N_s)\},\tag{3}$$

式中：$p_i(s)$表示第i个风电场数据落入第s个子区间的概率。

基于该分布，KL散度的计算公式为：

$$r_{\mathrm{KL}}(P_i|P_j)=\sum_{s=1}^{N_s}p_i(s)\ln\frac{p_i(s)}{p_j(s)}。\tag{4}$$

由于KL散度具有非对称性，无法准确反映两个分布之间的真实差异。

JS散度通过构造P_i与P_j的平均概率分布来解决该问题，其计算公式为：

$$r_{\mathrm{JS}}(P_i|P_j)=\frac{1}{2}\left(\sum_{s=1}^{N_s}p_i(s)\ln\frac{p_i(s)}{p_c(s)}+\sum_{s=1}^{N_s}p_j(s)\ln\frac{p_j(s)}{p_c(s)}\right),\tag{5}$$

式中：$p_c(s)=(p_i(s)+p_j(s))/2$。

据此，可构建语义感知图矩阵如下：

$$A^N=\begin{pmatrix}1&\omega_{12}^N&\cdots&\omega_{1N}^N\\ \omega_{21}^N&1&\cdots&\omega_{2N}^N\\ \vdots&\vdots&\ddots&\vdots\\ \omega_{N1}^N&\omega_{N2}^N&\cdots&1\end{pmatrix}\in\mathbf{R}^{N\times N},\tag{6}$$

$$\omega_{ij}^N=\begin{cases}\exp\left(-\dfrac{r_{\mathrm{JS}}}{\sigma_N^2}\right),\ \exp\left(-\dfrac{r_{\mathrm{JS}}}{\sigma_N^2}\right)\geqslant\eta,\\ 0,\qquad\text{其他},\end{cases}\tag{7}$$

式中：σ_N为稀疏权重因子，用于控制矩阵的稀疏性；η为稀疏超参数；$i,j=1,2,\cdots,N$。

1.2.2　全连接图构建器

全连接(Full Connection,FC)图构建器旨在捕获风电数据的时序信息。鉴于多变量时间序列通常具有较长的时间跨度,首先将每个传感器的时间序列划分为若干固定长度的子序列(补丁),以提取局部时序模式。设补丁尺寸为f,由输入序列X生成子序列集合$\{X_t\}_{t=1}^{\hat{L}}$,其中t为子序列索引,$X_t\in\mathbf{R}^{N\times f}$,$N$为传感器数量,子序列总数$\hat{L}=\lfloor L/f\rfloor$,其中$\lfloor\cdot\rfloor$表示截断操作。每个$X_t$包含所有传感器在该分段内的信号,即$X_t=\{x_{t,i}\}_{i=1}^{N}$,其中$x_{t,i}\in\mathbf{R}^f$。随后,采用编码器$f_c(\cdot|W_c)$对每个时间窗内的分段信号进行处理,该编码器在传感器层级上操作,以学习传感器级特征,即

$$x'_{t,i}=f_c(x_{t,i}|W_c)。$$

为保持子序列的位置信息,参照文献[15]引入位置编码。对于第i个传感器在第t个子序列中的特征,位置编码定义如下:

$$\vec{p}_t^{(m)}=f_p(t)^{(m)}=\begin{cases}\sin(\omega_k t), & m=2k,\\ \cos(\omega_k t), & m=2k+1,\end{cases}\tag{8}$$

式中:m为特征维度索引,ω_k为与维度相关的频率。获得各子序列的传感器特征后,进一步构建FC图。与传统方法仅考虑同一时刻或同一传感器内部关联不同,FC图将不同时间戳、不同传感器之间的节点全部连接,从而同时建模时空维度的关联关系。假设运行状态相似的传感器在特征空间中距离较近,故采用点积计算任意两个传感器特征之间的相似性。

对于任意两个传感器节点i、j,相似度为:

$$e_{tr,ij}=g_s(z_{l,i})\cdot\left(g_s(z_{r,j})\right)^{\mathrm{T}},$$

式中:$t,r\in[1,\hat{L}]$。函数$g_s(z)=z\boldsymbol{W}_s$用于提升表达能力,$\boldsymbol{W}_s$为可学习的权重矩阵。

随后,通过softmax函数将相关性值归一化至$[0,1]$区间,最终得到FC图$G^F=(\boldsymbol{Z},\boldsymbol{A}^F)$。邻接矩阵$\boldsymbol{A}^F=((e_{tr,ij})_{i,j=1}^{N})_{t,r=1}^{L}$表示所有时间子序列与传感器之间的相关性,$\boldsymbol{Z}=(z_{t,i})_{i=1,t=1}^{N,L}$。图$G^F$不仅包含同一时间步长内的空间关联,还刻画了跨时间步长的时空依赖关系,使模型能够对风电数据中全面的时间序列依赖进行建模。

然而,仅依据特征相似性构建FC图,忽略了时间距离对相关性的影响。实际中,相邻时间步长的数据通常具有更强的相关性,而时间间隔越大,相关性越弱。为此,设计时间距离衰减矩阵以增强FC图的表达能力。为直观说明,假设当前有3个子序列的图邻接矩阵,每个矩阵包含4个传感器信息,则邻接矩阵的每一行表示某传感器与所有子序列上所有传感器的相似关系。其中,前4列(传感器数量)表示与同一子序列内各传感器的相似性,因数据同时产生,应具有最强相关性;随后4列表示与下一子序列中各传感器的相似性,因位于不同子序列,其相关性应减弱并需通过衰减率衡量;最后4列相关性进一步下降,可采用另一个衰减系数表示。

基于上述分析,定义衰减矩阵$\boldsymbol{C}=(c_{tr,ij})_{i,j=1,i,r=1}^{N,\hat{L}}$,式中:$c_{tr,ij}=\delta^{t-r}$。该矩阵可用于调整不同子序列传感器相关性的强度。实际计算时,只需将原始邻接矩阵$\boldsymbol{A}^F$与衰减矩阵$\boldsymbol{C}$逐元素相乘,即可得到时间距离加权的邻接矩阵。该方法确保时间上邻近的传感器特征具有更强的相关性,而时间上远离的传感器相关性则相应减弱,更符合实际物理规律。

1.3　多关系图计算模块

多关系图计算模块的核心目的在于通过多尺度图神经网络提升整体模型的预测性能,以有效捕获风电数据中不同时间和空间尺度上的潜在特征。在多关系图构建模块完成风电传感器数据的精准提取,并分别形成空间与时间维度的图结构后,进一步对多尺度时空信息进行建模,主要包括以下4个步骤:1)识别输入时间序列的尺度;2)采用自适应图卷积提取尺度间序列相关性;3)利用多头注意力机制捕获序列内的时序依赖关系;4)通过SoftMax函数自适应融合不同尺度的表示。

1.3.1　多尺度识别器

多尺度识别器作用于编码后的原始数据,其本质是在不同时间尺度上构建对应的图结构。尺度的选取是影响模型性能的关键因素之一。考虑到风电序列数据固有的周期性特征,选取周期作为尺度划分的依据。

受到TimesNet[16]启发,采用快速傅里叶变换(Fast Fourier Transform, FFT)来检测数据的周期性,并将其作为时间尺度的来源,具体如下:

$$\begin{aligned}&\boldsymbol{F}=\mathrm{Avg}\left(\mathrm{Amp}\left(\mathrm{FFT}\left(\boldsymbol{X}_{\mathrm{emb}}\right)\right)\right),\\&f_1,\cdots,f_k=\underset{f\in\{1,\cdots,L/2\}}{\mathrm{argTopk}}\left(\boldsymbol{F}\right),s_i=\frac{L}{f_i}。\end{aligned}\tag{9}$$

式中:FFT(·)和Amp(·)分别表示FFT与幅度值计算,$\boldsymbol{X}_{\mathrm{emb}}$表示时序嵌入特征向量,向量$\boldsymbol{F}\in\mathbf{R}^L$表

示各频率对应的振幅。该结果通过在模型维度上使用Avg(·)进行平均得到。需要指出的是，随时间变化的输入数据可能呈现出不同的周期性，从而使模型能够自适应地检测不断变化的时间尺度。这种随时间变化的周期尺度具有稳定的内在相关性，从而使模型能够有效学习序列间与序列内相关性的动态演化特性。

根据所选定的时间尺度，可利用式(10)将输入数据沿时间维度进行调整，得到对应不同时间尺度的若干三维张量表示：

$$\boldsymbol{X}^i=\mathrm{Reshape}_{s_i,f_i}\left(\mathrm{Padding}\left(\boldsymbol{X}_{\mathrm{in}}\right)\right), \tag{10}$$

式中：Reshape(·)用于数据尺度的变换，Padding(·)用于沿时间维度对输入时间序列进行填充扩展，$X^i\in\mathbf{R}^{d_{\mathrm{model}}\times s_i\times f_i}(i\in\{1,\cdots,k\})$表示基于第$i$个时间尺度经过整形后得到的三维特征张量。

1.3.2 全连接图卷积器

全连接图卷积器接收由全连接图构建器生成的包含多个子序列信息的FC图，旨在捕获风电数据中的时空依赖关系，实现表示学习。FC图应用图卷积的方法不仅难以有效捕获数据中的局部依赖性，还可能带来较高的计算成本。为此，本文提出一种基于移动池化的图神经网络(Moving-Pool GNN)，包括用于捕获局部时空依赖的移动窗口机制和用于提取高级特征的时间池化操作。

首先，采用大小为2的滑动窗口以步长1移动，可获取两个窗口，每个窗口包含2个子序列，每个子序列包含4个传感器。随后，在每个窗口内采用GNN进行计算。

采用消息传递神经网络(Message Passing Neural Network，MPNN)作为GNN的变体，以捕获窗口内图结构的时空依赖关系。MPNN的计算过程主要分为传播阶段与更新阶段。在传播阶段，相邻节点的信息被聚合至中心节点。给定第l层中第w个窗口的中心节点$z_{t,i}^l$，其相邻节点集合为$\{z_{r,j}^l\}_{j=1,r=w-M/2}^{N,w+M/2}$，这些节点跨越同一窗口内的$M$个子序列。中心节点与其各相邻节点的相关性为$\{e_{tr,ij}^l\}_{j=1,r=w-M/2}^{N,w+M/2}$。经传播后，中心节点聚合得到的特征为

$$h_{t,i}^l=\sum_{r=w-M/2}^{w+M/2}\sum_{j=1}^{N}z_{r,j}^l e_{tr,ij}^l。$$

在更新阶段，采用非线性函数对传播后的传感器特征进行变换，即$z_{t,i}^{l+1}=f_g(h_{t,i}^l|W_g)$。总体而言，MPNN基于传感器之间的相关性，在$M$个子序列上传播传感器信息，从而使模型能够在窗口内充分捕获时空依赖关系，并通过非线性更新增强特征的表示能力。在完成时空依赖捕获与传感器特征更新后，进一步采用时间池化操作提取每个窗口的高级特征。该操作借鉴卷积神经网络中的池化思想。

给定更新后的传感器特征$\{z_{t,i}^{l+1}\}_{t=w-M/2}^{w+M/2}$，对于第$i$个传感器，其在窗口$w$内对应$M$个子序列的特征，采用平均池化策略进行时间聚合，得到该窗口的传感器特征

$$z_{w,i}^{l+1}=\sum_{t=w-M/2}^{w+M/2}z_{t,i}^{l+1}/M。$$

通过在所有窗口上堆叠传感器特征，可构建新的FC图，作为后续层的输入。需要指出的是，本文仅采用单层结构，可直接利用各窗口获得的传感器特征作为输出。受多分支结构研究的启发，进一步集成了多个并行的图构建与卷积层。该设计允许各分支采用不同的权重初始化策略，从而使模型在训练过程中能够从多种综合信息视角更全面地捕获时空依赖关系，并获取更优的求解结果。将多个并行层的输出特征进行堆叠后，采用直接输出层将堆叠特征转换为最终的图表示，供后续预测模块使用。

1.3.3 多尺度自适应图卷积器

多尺度自适应图卷积器与全连接图卷积器并行处理，主要针对稀疏图构建器生成的空间距离图与语义感知图进行计算。首先，通过线性变换将对应于第i个尺度的张量投影回由N个变量表示的张量空间，定义如下：

$$\boldsymbol{H}^i=\boldsymbol{W}^i\boldsymbol{X}^i, \tag{11}$$

式中：$\boldsymbol{H}^i\in\mathbf{R}^{N\times s_i\times f_i}$为投影后的特征张量，$\boldsymbol{W}^i\in\mathbf{R}^{N\times d_{\mathrm{model}}}$为适应第$i$个尺度张量的可学习权重矩阵。在多尺度自适应图卷积器中，图学习过程引入两个可训练参数矩阵$\boldsymbol{E}_1^i\in\mathbf{R}^{N\times h}$和$\boldsymbol{E}_2^i\in\mathbf{R}^{N\times h}$，并通过二者相乘动态生成自适应邻接矩阵，公式如下：

$$\boldsymbol{A}^i=\mathrm{SoftMax}\left(\mathrm{ReLU}\left(\boldsymbol{E}_1^i\left(\boldsymbol{E}_2^i\right)^{\mathrm{T}}\right)\right), \tag{12}$$

式中：SoftMax函数用于对节点间的权重进行归一化，以确保序列间关系表达的平衡性与合理性。

在获得第i个尺度对应的自适应邻接矩阵$\boldsymbol{A}^i$后，采用Mixhop图卷积方法捕获序列间的空间相关性。该方法能够有效提取其他模型可能忽略的特征。图卷积过程定义为：

$$\boldsymbol{H}_{\text{out}}^{i}=\sigma\left(\bigcup_{j\in P}\left(\boldsymbol{A}^{i}\right)^{j}\boldsymbol{H}^{i}\right),\tag{13}$$

式中：$\boldsymbol{H}_{\text{out}}^{i}$ 为第 i 个尺度的融合输出特征，$\sigma(\cdot)$ 为激活函数，P 为整数邻接幂次集合，$\left(\boldsymbol{A}^{i}\right)^{j}$ 表示邻接矩阵 $\boldsymbol{A}^{i}$ 的 j 次幂。

最后，通过多层感知器(Multilayer Perceptron，MLP)将 H_{out}^{i} 投影回三维张量形式 $\hat{\boldsymbol{X}}^{i}\in\mathbf{R}^{d_{\text{model}}\times s_i\times f_i}$。

1.3.4　多头注意力机制与尺度融合器

经自适应图卷积器处理后，生成对应不同尺度的多个可学习图特征表示。多头注意力机制与尺度融合器负责捕捉不同尺度间的相关性，并对多尺度信息进行自适应聚合。对于每个时间尺度对应的张量 $\hat{\boldsymbol{X}}^{i}$，在尺度维度上应用多头自注意力机制

$$\hat{\boldsymbol{X}}_{\text{out}}^{i}=\text{MHA}_{s}\left(\hat{\boldsymbol{X}}^{i}\right),\tag{14}$$

式中：$\text{MHA}_{s}(\cdot)$ 表示在尺度维度上执行多头自注意力操作[8]。该操作将输入张量形状由 $B\times d_{\text{model}}\times s_i\times f_i$ 变换为 $Bf_i\times d_{\text{model}}\times s_i$，其中 B 为批量大小。尽管已有研究表明多头注意力机制在捕获时间序列长期依赖方面存在一定局限性，但本文通过尺度变换策略，将长时间跨度转换为周期长度，有效缓解了该问题。

完成多头注意力计算后，需要对 k 个不同尺度的张量 $\left\{\hat{\boldsymbol{X}}_{\text{out}}^{1},\cdots,\hat{\boldsymbol{X}}_{\text{out}}^{k}\right\}$ 进行聚合以进入下一层。首先，将每个尺度的输出张量重塑为二维矩阵 $\hat{\boldsymbol{X}}_{\text{out}}^{i}\in\mathbf{R}^{d_{model}\times L}$。随后，根据各尺度对应的振幅进行自适应融合，

$$\left[\hat{a}_1,\cdots,\hat{a}_k\right]=\text{SoftMax}\left(F_{f_1},\cdots,F_{f_k}\right),$$

$$\hat{\boldsymbol{X}}_{\text{out}}=\sum_{i=1}^{k}\hat{a}_i\,\hat{\boldsymbol{X}}_{\text{out}}^{i},\tag{15}$$

式中：F_{f_i} 为基于快速傅里叶变换(FFT)计算得到的第 i 个尺度对应的振幅值。本文采用SoftMax函数对各尺度振幅进行归一化，得到融合权重 $\hat{a}_i$。该策略借鉴了专家混合(Mixture of Experts，MoE)的思想，使模型能够依据各尺度在频域中的贡献度对其进行自适应加权，从而实现多尺度特征的有效融合。

2　实验分析

2.1　数据集

采用位于美国不同地区的两个风电场数据集(数据集1和数据集2)进行实验评估，二者均为美国国家可再生能源实验室公开收集的风电数据集的子集，最初应用于“东风集成与传输”研究项目。两个数据集的时间范围均覆盖2004年1月1日至2006年1月1日，共计包含105 264个数据采样点。每个数据点均包括经统一归一化处理后的风速与风电功率值。依据AIC、BIC和FPE标准，将时间滞后(Time Lag，TL)参数的最佳值设置为24。

2.2　参数设置

为统一优化策略，所有深度模型均采用Adam优化器，并选用额外损失作为损失函数。

为确保比较的公平性，数据集按时间顺序划分为3部分：前60%数据作为训练集，中间10%数据作为验证集用于超参数调优，剩余30%数据作为测试集用于样本外预测评估。同时，为降低模型初始化及训练过程中随机性对结果的影响，每个实验均独立重复运行10次，并取其性能指标的平均值作为最终结果。

2.3　对比方法

选取多种主流方法进行对比，涵盖传统机器学习模型SVR，深度学习领域的CNN、Transformer、LSTM和GNN等代表性模型。

SVR：通过设定容忍误差带，增强对异常值的鲁棒性，并提升预测泛化能力。

CNN：采用一维卷积提取时间序列中的局部模式，自动学习有效特征，减少对人工特征工程的依赖。

LSTM：通过门控机制实现信息的选择性保留与遗忘，有效捕获长期时序依赖。

STAGCM(Spatial-Temporal Adaptive Graph Convolutional Network with Attention)：利用GCN捕获节点间的拓扑结构关系，并通过时空注意力模块自适应学习不同时间步与空间邻居的重要性权重。

ISTAMGCN(Improved Spatial-Temporal Attention Multi-Graph Convolutional Network)：在STAGCM基础上引入改进的时空注意力机制与多图卷积结构，增强对动态空间依赖及长短期时序依赖的建模能力。

GNN-FFTransformer：利用GNN建模多变量间的空间关系，并通过特征融合实现时空信息的有效整合。

Informer：引入ProbSparse自注意力机制、自注意力蒸馏操作和生成式解码器，显著提升计算

效率与长程依赖捕获能力。

2.4 评价指标

采用归一化平均绝对误差(Normalized Mean Absolute Error, NMAE)和归一化均方根误差(Normalized Root Mean Square Error, NRMSE)作为预测性能的评价指标，其计算公式如下：

$$\mathrm{NMAE}=\frac{1}{N_p}\sum_{q=1}^{N_p}\left|y_q^i-\hat{y}_q^i\right|/M_N^i\times 100\%, \tag{16}$$

$$\mathrm{NRMSE}=\frac{1}{M_N^i}\sqrt{\frac{1}{N_p}\sum_{q=1}^{N_p}\left(y_q^i-\hat{y}_q^i\right)^2}\times 100\%, \tag{17}$$

式中：M_N^i表示第i个风电场的额定装机容量，N_p为测试集中的样本总数，y_q^i和$\hat{y}_q^i$分别表示第q条测试样本中第i个风电场风电功率的实际值与预测值。

2.5 结果分析

2.5.1 平均预测性能分析

依据前述参数设置，统计了区域内20个风电场的平均归一化平均绝对误差(NMAE)和平均归一化均方根误差(NRMSE)，用以评估不同方法在不同预测范围下的预测性能，结果如表1和表2所示。下面进一步分析各方法，并验证本文方法的优越性。

表1 不同基准方法和不同时间范围内所有风电场的平均NMAE

Tab. 1 Average NMAE of all wind farms under different benchmark methods and time ranges

模型	NMAE/%					
	10 min	20 min	1 h	3 h	4 h	5 h
SVR	2.577	3.832	7.674	13.877	15.824	19.541
CNN	2.456	3.624	7.352	13.671	15.735	19.053
LSTM	2.415	3.592	7.016	13.036	15.426	18.872
STAGCM	2.101	3.277	6.679	12.797	15.231	17.998
ISTAMGCN	2.043	3.228	6.534	12.698	14.996	17.954
GNN-FFTransformer	2.096	3.307	6.661	12.753	15.071	18.064
Informer	2.317	3.327	6.837	12.864	15.562	18.546
Dual-MSGNN	2.048	3.224	6.525	12.657	14.858	17.734

表2 不同基准方法和不同时间范围内所有风电场的平均NRMSE

Tab. 2 Average NRMSE of all wind farms under different benchmark methods and time ranges

模型	NRMSE/%					
	10 min	20 min	1 h	3 h	4 h	5 h
SVR	4.265	6.303	10.819	19.571	20.985	23.314
CNN	4.163	6.121	10.847	18.692	20.851	23.154
LSTM	4.077	5.921	10.443	18.293	20.486	22.685
STAGCM	3.843	5.783	10.169	17.377	19.606	22.005
ISTAMGCN	3.770	5.662	10.023	17.194	19.352	21.937
GNN-FFTransformer	3.830	5.785	10.175	17.233	19.564	22.064
Informer	4.021	5.831	10.221	17.462	19.769	22.446
Dual-MSGNN	3.224	3.781	10.018	17.089	19.235	21.782

从整体来看，混合模型的预测性能普遍优于传统单一模型。相比之下，单一深度学习模型在时空特征联合提取方面存在不足。例如，随着预测范围的增大，CNN和LSTM的误差增长更为显著，原因在于二者分别侧重于空间特征或时间特征提取，难以同时捕获时空耦合关系。因此，融合两类模型能够进一步提升预测性能。此外，在风电功率预测任务中，传统机器学习模型的准确性和稳定性均低于深度学习模型，主要由于SVR在处理具有潜在非线性关系的高维时间序列数据时能力有限，而深度学习模型具备更强的复杂特征提取能力。

在混合模型对比中，ISTAMGCN进一步与多种CNN及GCN方法进行比较。结果表明，基于GCN的方法在预测性能和稳定性上均优于CNN方法。Informer虽通过注意力机制降低了输入冗

余噪声的影响,但其误差仍高于多数GCN方法。这说明CNN在WPF任务中难以有效适应风电场在非欧几里得空间中的复杂拓扑结构,而GCN在空间关联建模方面具有天然优势。

在基于图模型的方法中,本文提出的Dual-MSGNN取得了最高的预测精度和最佳泛化能力,表明其能够有效捕获风电场数据中的时空依赖关系。通过构建多种时空图结构并设计相应的特征捕获策略,Dual-MSGNN实现了时间与空间信息的协同建模,同时结合动态组合方法与时空注意力机制,有效抑制了长时预测中的噪声积累。

此外,ISTAMGCN取得了与Dual-MSGNN较为接近的预测效果,表明多图机制在WPF任务中具有重要作用。在短期预测(如10 min)中,ISTAMGCN表现尤为突出,取得了最佳结果,说明其具有较强的短期时序建模能力。然而,随着预测范围的扩大,其对目标风电场先验空间信息的挖掘不足,且自注意力机制对远距离关联信息的捕获能力有限,导致信息利用不充分。相比之下,本文方法能够捕获更丰富的时空依赖关系,并有效抑制无关风电场信息的干扰,从而在更长预测范围内获得更优性能。

2.5.2　模块性能分析

为验证Dual-MSGNN中构建的多种图结构的有效性及多尺度建模策略的合理性,本文对整体框架进行了消融实验。具体地,在5 h预测范围内,分别去除空间图、时间图、邻居图以及多尺度建模模块,并测试各配置下的归一化平均绝对误差,以评估各模块对模型整体性能的贡献,实验结果如图2所示。

从图2可以看出,去除空间图、时间图或邻居图均会导致模型性能明显下降,表明风电功率预测任务需要同时依赖空间关系与时间依赖信息的联合建模。此外,若不采用多尺度方式构建和运行图神经网络,模型性能亦出现显著衰减,说明在多尺度上进行分析是捕捉风电场数据复杂动态特性的关键。综合上述结果,通过构建多种图结构并引入多尺度建模策略,有效增强了对风电场数据中时空耦合依赖关系的捕获能力,从而显著提升了预测性能。

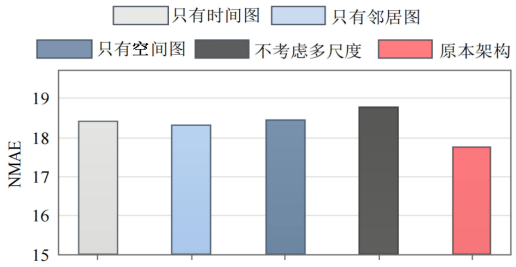


图2　消融实验结果

Fig. 2　Ablation experiment results

3　结论

本文提出了Dual-MSGNN风电功率预测模型,通过频域分析与自适应图卷积,实现对时间和空间尺度上序列间相关性的高效捕获,主要结论如下:

1) Dual-MSGNN模型通过引入多关系图构建模块,有效整合了风电场的空间位置与时间演化特征,提升了风电功率预测精度,同时增强了模型决策过程的可解释性。

2) 采用快速傅里叶变换对时间序列进行周期检测,增强了模型对不同时间尺度相关性的识别能力;自适应图卷积层与多头自注意力机制的协同引入,使模型能够自适应地捕捉动态时空依赖,进一步提升了对复杂气象条件下数据的适应性与泛化性能。

3) 在两个实际风电场数据集上的实验结果表明,Dual-MSGNN在绝大数预测范围内取得最佳预测性能,验证了其在风电功率预测任务中的有效性与优越性。

后续研究将聚焦于气候因素不确定性条件下的模型鲁棒性提升,探索更加多样化的图结构构建方式与高效的推理优化策略,以增强模型在不同气候场景下的适应能力与实际工程应用价值。

参考文献:

[1] Lakatos L, Hevessy G, Kovács J. Advantages and disadvantages of solar energy and wind-power utilization[J]. World Futures, 2011, 67(6): 395-408.

[2] 孙荣富, 张涛, 和青, 等. 风电功率预测关键技术及应用综述[J]. 高电压技术, 2021, 47(4): 1129-1143.

[3] Stathopoulos C, Kaperoni A, Galanis G, et al. Wind power prediction based on numerical and statistical models[J]. Journal of Wind Engineering and Industrial Aerodynamics, 2013, 112: 25-38.

[4] Awad M, Khanna R. Support vector regression[M]//Awad M, Khanna R. Efficient learning machines. Berkeley: Apress, 2015: 67-80.
[5] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems, 2017: 30.
[6] 余勇祥, 匡相宇, 韩建, 等. 融合物理约束与多尺度特征的Mamba-Transformer超短期风电功率预测[J]. 太阳能学报, 2026, 47(6): 296-305.
[7] Voulodimos A, Doulamis N, Doulamis A, et al. Deep learning for computer vision: a brief review[J]. Computational Intelligence and Neuroscience, 2018, 2018: 7068349.
[8] Lauriola I, Lavelli A, Aiolli F. An introduction to deep learning in natural language processing: models, techniques, and tools[J]. Neurocomputing, 2022, 470: 443-456.
[9] Wu N, Green B, Ben Xue, et al. Deep transformer models for time series forecasting: the influenza prevalence case [PP/OL]. arXiv(2020-01-23)[2026-06-20]. https://doi.org/10.48550/arXiv.2001.08317.
[10] 王小明,徐斌,尹元亚,等. 基于空间相关性与Stacking集成学习的风电功率预测方法[J]. 电力工程技术,2024,43(5): 224-232.
[11] 胡宇晗, 朱利鹏, 李佳勇, 等. 融合深度误差反馈学习和注意力机制的短期风电功率预测[J]. 电力系统保护与控制, 2024, 52(4): 100-108.
[12] 刘颂凯, 刘旺江, 胡畔, 等. 考虑风电出力时空相关性的暂态稳定约束最优潮流[J]. 电力工程技术, 2026, 45(6): 12-20.
[13] 刘颂凯,成思鋙,苏攀,等. 基于深度学习的含风电不确定性的暂态稳定约束最优潮流[J]. 电力工程技术,2026,45(5): 15-26.
[14] He Yaoyao, Wang Yun. Short-term wind power prediction based on EEMD-LASSO-QRNN model[J]. Applied Soft Computing, 2021, 105: 107288.
[15] Wang Yucheng, Xu Yuecong, Yang Jianfei, et al. Fully-connected spatial-temporal graph for multivariate time-series data[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2024, 38(14): 15715-15724.
[16] Wu Haixu, Hu Tengge, Liu Yong, et al. TimesNet: temporal 2D-variation modeling for general time series analysis [PP/OL]. V3. arXiv(2023-04-12)[2026-06-21]. https://doi.org/10.48550/arXiv.2210.02186.

责任编辑：陈松楠　申晓伟