第 52 卷 第 2 期：663-673 高电压技术 Vol.52, No.2: 663-673
2026 年 2 月 28 日 High Voltage Engineering February 28, 2026
**DOI**: 10.13336/j.1003-6520.hve.20250447

# 基于双重注意力机制与时空门控融合的风电集群功率短期预测

孟高军[1]，吴佳凤[1]，张仰飞[1]，于琳琳[2]，刘海涛[1]
（1. 江苏省配电网智能技术与装备协同创新中心(南京工程学院)，南京 211167;
2. 国网河南省电力公司经济技术研究院，郑州 450052）

**摘　要**：随着全球能源结构向低碳化转型加速，风电并网规模的持续扩大使得其功率波动性与时空耦合特性成为制约电网安全经济运行的一大难题。针对风电集群功率预测中的时空依赖性和风电场间协同机制不明问题，该文提出了一种基于双重注意力机制与时空门控融合的预测方法。首先，利用最大互信息系数(maximal information coefficient, MIC)量化非线性气象特征的隐式影响，结合 Pearson 矩阵解析风电场间协同关系，有效解耦了风电集群的时空耦合特性。其次，基于注意力强化和抑制参数冗余改进卷积网络，建立时间和通道双重注意力机制的时空双驱动特征提取架构，自适应调整特征权重。最后，设计了时空门控机制模块，增强对时空特征的捕捉能力，有效解决时空特征融合不充分的问题。实验结果表明，所提模型在集群风电场数据集上的预测性能显著优于基线模型，具有较强的泛化能力与应用性。

**关键词**：风电集群；短期功率预测；卷积神经网络；双重注意力机制；时空特性

## Short-term Wind Power Prediction for Wind Farm Clusters Based on Dual Attention Mechanism and Spatio-temporal Gated Fusion

MENG Gaojun[1], WU Jiafeng[1], ZHANG Yangfei[1], YU Linlin[2], LIU Haitao[1]
(1. Jiangsu Distribution Network Intelligent Technology and Equipment Collaborative Innovation Center(Nanjing Institute of Technology), Nanjing 211167, China; 2. Research Institute of Economics, State Grid Henan Electric Power Co., Zhengzhou 450052, China)

**Abstract**：With the accelerated global energy structure transformation towards low-carbonization, the continuous expansion of grid-connected wind power capacity has made its power volatility and spatio-temporal coupling characteristics to be a critical challenge for power grid safety and economy. To address the issues of unclear spatio-temporal dependencies and inter-wind farm collaboration mechanisms in wind power cluster forecasting, this paper proposes a prediction method based on dual attention mechanisms and spatio-temporal gating fusion. Firstly, the maximal information coefficient is utilized to quantify the implicit influence of nonlinear meteorological features; combined with Pearson correlation matrix analysis of inter-wind farm cooperative relationships, the spatio-temporal coupling characteristics are decoupled. Secondly, an improved convolutional network with attention reinforcement and parameter redundancy suppression is developed to establish a spatio-temporal dual-driven feature extraction architecture with dual attention mechanisms in temporal and channel dimensions, and adaptively adjusts feature weights. Finally, a spatio-temporal gating mechanism module is designed to enhance spatio-temporal feature capturing capability, thereby effectively solving the problem of insufficient spatio-temporal feature fusion. Experimental results demonstrate that the proposed model significantly outperforms baseline models on wind farm cluster datasets, exhibiting strong generalization ability and practical applicability.

**Key words**：wind power cluster; short term power prediction; convolutional neural network; dual attention mechanism; spatio-temporal characteristics

## 0　引言

随着我国“双碳”政策的推行，截至 2024 年 8 月，全国累计风电装机容量约 4.7 亿 kW，同比增长 19.9%[[1]]。风能作为清洁能源，已经成为我国能源结构中的重要一环[2]。然而，风电出力受气象条件影响显著，具有非线性、时变性和不确定性，给电网调度以及及时消纳带来巨大挑战[3]。同时，在大型风电基地建设与海上风电规模化开发的双重驱动

**基金资助项目**：国家自然科学基金(52377053)；国网河南省电力公司科技项目(5217L0230007)。
Project supported by National Natural Science Foundation of China (52377053), Science and Technology Project of State Grid Henan Electric Power Company (5217L0230007).

下，未来风电开发将继续向集群化演变，这使得传统单机组的预测模式难以满足电网协同调度需求，对风电集群功率预测性能及预测精准度提出了更高要求[4]。

风电预测方法主要分为两类，一是基于数值天气预报(numerical weather prediction，NWP)，求解满足的高维非线性方程组，得到风机轮毂高度的风速、风向等信息，然后结合风电机组功率曲线计算得到风机的输出功率[5]。而 NWP 数据与实际风电场数据间存在误差且偏理想化，因此只适用于中长期预测。另一类方法一般根据历史运行数据和历史 NWP 数据的关联性进行统计分析为基础，建立 NWP 数据与风电功率之间的映射关系[6]。这类方法包括时序模型、机器学习与深度学习。时序模型未考虑气象因素的影响，预测精度不理想，需要结合人工神经网络进行组合预测；机器学习模型如人工神经网络、支持向量机，可较好得从数据中获取风电变化规律，但存在高数据要求以及参数选取问题[7-9]。深度学习模型尤其是神经网络，在大规模数据集上能提供更高的准确性和更好的性能[10]。因此，本文选用深度学习进行建模。

现有的深度学习研究主要分为两大类：一是对循环神经网络进行改良优化，一是卷积神经网络。而在风电预测方面，时序卷积神经网络因其可并行计算的优势，广受青睐。文献[11]借助神经网络的自学习能力，针对易过拟合现象，构建了基于鲸鱼算法和模拟退火组合优化(whale optimization algorithm - shuffle attention, WOA-SA)的短期预测模型，有效减小了预测误差。文献[12]针对风电场的迁移学习修正，对风速进行多重集成与修正，有效提升了数据不足时的预测性能。文献[13]针对数值天气预报与实际数据的误差，基于双向长短时记忆网络(bi-directional long short-term memory, BILSTM)预测风电，有效优化了 NWP 数据。文献[14]针对单一风电场的预测无法满足电网调度需求的问题，提出一种多重集成的集群短期风电功率预测方法，精度明显提升。文献[15]针对大型风电场的功率预测采用单点位风速外推预测代表性差的局限,提出基于高斯混合模型聚类的风电场短期功率预测方法，该方法在分组功率预测中表现出了显著优势。但以上研究都集中在单一发电场领域，未考虑到同一集群内的风电场间的协同特性。

此外，现今的风电预测大多在特征选取层面存在特征选取片面的问题。文献[16]应用平衡迭代规约和聚类,对集群内各场站按照二维坐标聚类分析，较常规方法在预测上有一定优越性，但没有考虑到集群内子场之间的出力相关性。文献[17]在数据中引入风切变物理量，由长短时记忆全卷积网络(long short-term memory - fully convolutional networks, LSTM-FCN)提取环境数据与海上风电出力间的深层非线性关联，有效提高了预测精度。文献[18]对原始风电数据进行缺失值插补，在时间卷积网络的基础上，引入注意力机制调整不同特征的权值，在高预测精度的同时兼顾了鲁棒性。文献[19]基于注意力机制改进 Transformer 耦合卷积神经网络,使用来自不同风电场和时间段的多个数据集进一步加强了所提算法的可靠性和泛化能力。

综合上述问题，风电功率预测需要综合考虑风电场间的空间依赖性与时间序列特征，深入把握其出力耦合关系，并融合多模态特征数据。多种时间尺度预测中，超短期预测因时间跨度过短，难以支撑发电计划；中长期预测偏重宏观规划，无法满足电网实时调度需求。短期预测尽管对电网调度至关重要，却因气象与出力数据波动大、不确定性高、时空相关性复杂而面临诸多挑战。此外，大气动力学等多因素耦合使得影响权重难以界定，预测误差易引发发电计划偏离，干扰电力市场的价格信号。

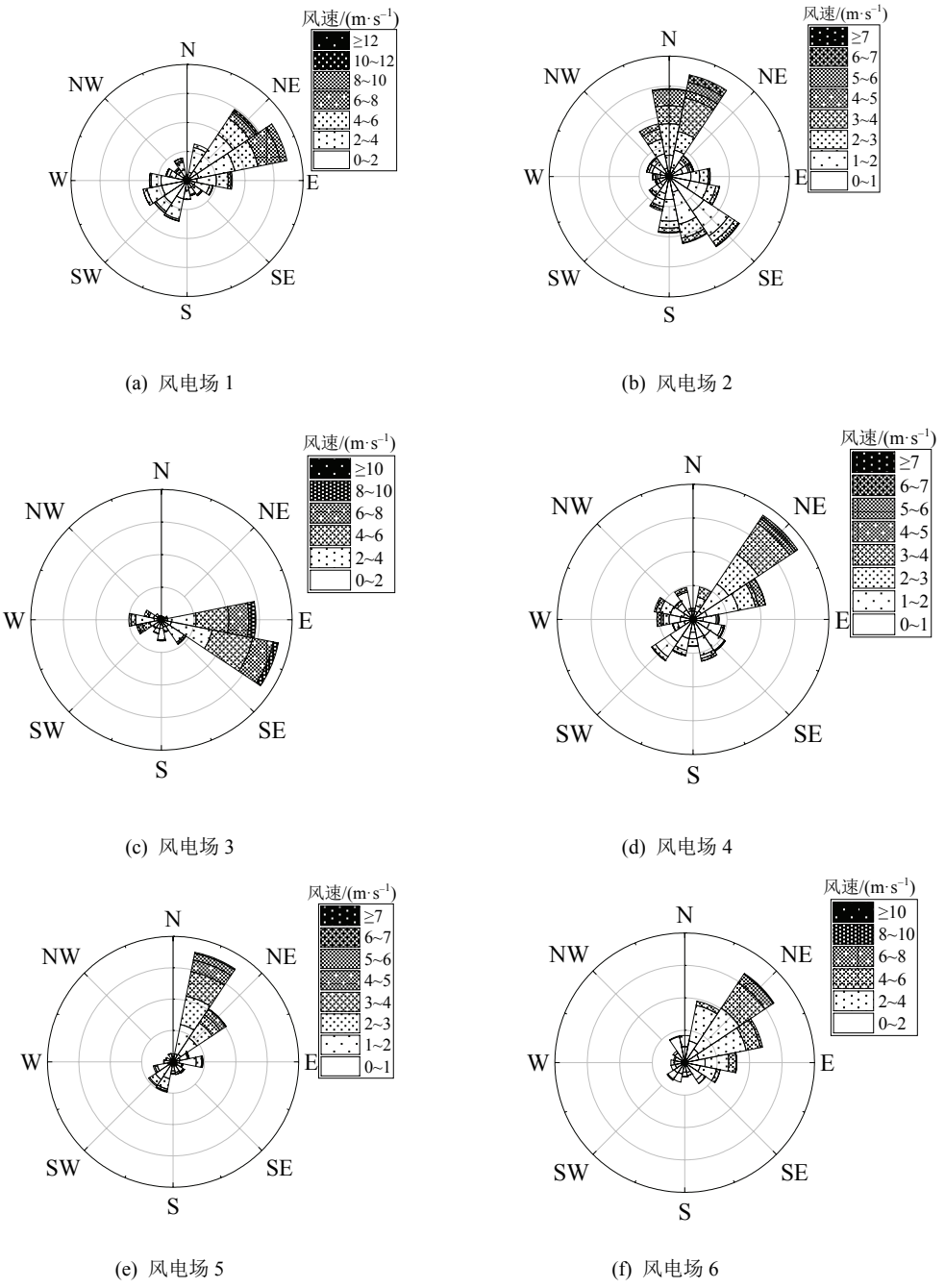
为此，本文提出一种基于双重注意力机制与时空门控融合的风电功率短期预测模型。该模型通过数据预处理模块，利用最大互信息系数(maximal information coefficient, MIC)和 Pearson 相关矩阵深入挖掘风电场间的潜在时空关联关系，构建多维节点属性特征矩阵与节点结构特征矩阵。在模型架构中，引入时间和通道双重注意力机制，对输入特征权重及风电场集群间的结构化关联进行自适应调整，以捕捉复杂的时空依赖关系；同时，通过时空门控融合机制有效表征时间序列的动态变化，进一步提升特征建模能力。最终，通过融合特征的深度处理与加权平均策略生成整体预测结果。该模型能够更加精确地对风电集群间的复杂时空关系进行建模，有效提升风电功率短期预测的预测性能。

## 1 风电场时空功率特性分析

风电场的地理分布和风资源的空间连续性使得相邻风电场之间呈现显著的时空关联性[20]。以北方某风电场集群为例，如图 1 所示，其 6 个风电场

的主导风向和风速表现出明显的方向性。由于地理邻近，风速和风向趋于一致，但局部地形差异仍会导致个别风电场的功率输出出现一定差异。此外，相邻风电场之间还存在显著的尾流效应。当气流经过上游风电场风力机时，因叶片的空气动力学作用对气流造成扰动，在风力机后方形成尾流区域。尾流区域内风速因能量损耗显著降低，同时风向因紊流特性发生偏移。这种变化直接影响下游风电场风力机的空气动力学性能，导致发电效率降低和功率输出波动加剧。

基于此，风电场时空关联性体现在：功率波动由相似风速、风向驱动，呈同步性；功率波动时间差异与风速相关，高风速时相邻风电场间时滞小，低风速时大；尾流效应使下游风电场功率输出低于上游，且波动时间和幅度有独特规律。




(a) 风电场 1　(b) 风电场 2

(c) 风电场 3　(d) 风电场 4

(e) 风电场 5　(f) 风电场 6

图 1　集群内 6 处风电场风资源分布图

Fig.1　Wind resource distribution map of the six wind farms in the cluster

## 2 基于注意力机制的集群功率短期模型

本文提出的基于双重注意力机制与时空门控融合的风电集群功率短期风电预测模型，其总体框架如图 2 所示。

该模型主要包括 3 个部分：

1）数据预处理模块：对多模态数据集(包括出力、风速、风向、温湿度等)进行预处理，利用孤立森林算法去除异常数据，并将离散的多模态数据转换为节点属性特征矩阵和节点结构特征矩阵。

2）改进卷积模块：通过多层卷积堆叠，逐步提取高层次的时空特征，引入双重注意力机制(时间和通道注意力)对各输入特征和集群各风电场间协同关系进行加权，帮助模型自适应调整各特征的重要性，增强对关键特征的聚焦能力。同时，加入时空门控机制，使得模型能够有效地捕捉风电功率波动中的长短期依赖。

3）输出模块：包括全连接层和回归层，用于生成最终短期功率预测结果。

### 2.1 数据预处理

#### 2.1.1 异常值处理

高维和大规模数据的复杂性给模型学习效率和泛化能力带来挑战。为简化数据并提升预测性能，常采用聚类分析和异常检测方法。孤立森林作为一种高效的无监督算法，其核心假设是异常样本更易被孤立。

孤立森林通过随机选择特征及其切分值构建多棵孤立树。持续进行递归划分，直到子集仅包含一个样本或达到预设深度。样本的异常分数 $s(l,n)$ 基于其在所有树中的平均路径长度 $E(h(l))$获得[21]。其计算式如下：

$$s(l,n)=2^{-\frac{E(h(l))}{c(n)}} \tag{1}$$

$$E(h(l))=\frac{1}{K}\sum_{k=1}^{K}h_k(l) \tag{2}$$

式中：$c(n)$为归一化常数；$n$ 为原始样本数量；$K$ 为孤立树总数；$h_k(l)$ 为样本 $l$ 在第 $k$ 颗孤立树中的路径长度；$E(h(l))$为平均路径长度，路径越短，样本越容易被孤立，异常分数越高。

#### 2.1.2 特征提取

为精准解析气象-风电功率耦合机理，本研究构建了 MIC 与 Pearson 相关系数的双模态分析框架，从非线性与线性双重视角解构其作用机制。

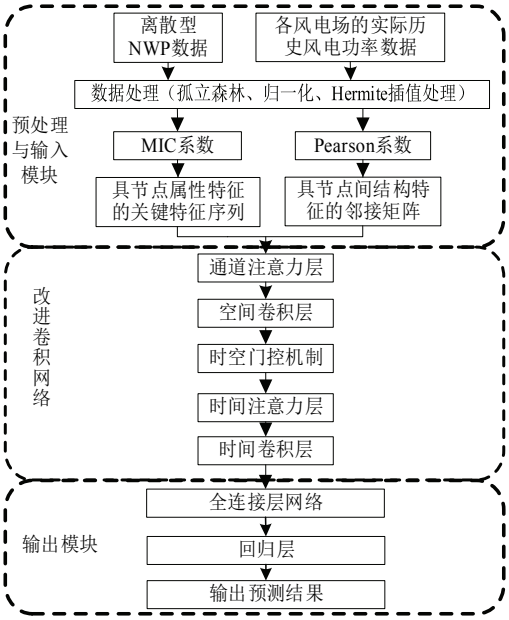



图 2 风电集群功率预测流程框架

Fig.2 Wind cluster power prediction process framework

MIC 基于信息论，无需对数据分布作假设，可在捕捉量化复杂的非线性关系的同时检测多变量间的高阶作用，挖掘关键信息，为分析提供深层支持；Pearson 方法依托协方差架构，可高效解析风电场群的关联特性，其显著性检验机制确保了推断的可靠性。该融合方法通过双维度特征（线性、非线性）识别能力互补，可以量化表征风电场群时空耦合关系，减少冗余信息，提升模型计算效率。

其中，互信息定义为：

$$I(X;Y)=H(X)-H(X\mid Y)=\sum_{x\in X}\sum_{y\in Y}p(x,y)\log_2\frac{p(x,y)}{p(x)p(y)} \tag{3}$$

式中：$I(X;Y)$为变量 $X$ 和 $Y$ 之间的互信息；$H(X)$为 $X$ 的熵；$H(X|Y)$为条件熵；$p(x,y)$为联合概率密度函数；$p(x)$、$p(y)$均为边缘密度函数。

MIC 系数表示为：

$$M_{\mathrm{IC}}(X;Y)=\max_{|G|_X\cdot|G|_Y<B}\frac{I(X;Y)\big|_G}{\log_2(\min(|G|_X,|G|_Y))} \tag{4}$$

式中：$G$ 表示在变量 $X$ 和 $Y$ 值域上划分的一个网格；$|G|_X$ 和 $|G|_Y$ 分别为该网格在 $X$ 和 $Y$ 上的划分格数；$|G|_X\cdot|G|_Y<B$ 为网格数量的约束条件；$B$ 为取样本总数的 0.6 次方。

通过 Pearson 系数量化集群风电场间功率波动的线性协同性，其计算式为：

$$\rho_{xy}=\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2\sum(y_i-\bar{y})^2}} \tag{5}$$

式中：$x_i$、$y_i$为风电场功率序列；$\bar{x}$、$\bar{y}$为其均值；$\rho_{xy}$为这两个风电场之间的相关系数。

由预处理得到的节点属性特征矩阵作为卷积网络的输入部分，为模型提供风电场出力预测的基础环境信息。不同的气象因素如风速、风向、温度等，通过卷积层的卷积核在属性特征矩阵上滑动进行卷积操作，提取不同尺度和层次的气象特征。节点结构特征矩阵将风电场在空间上进行建模，与由 Pearson 得到的用以调整权重的邻接矩阵，挖掘风电场间空间上的出力相互影响和协同变化关系。

### 2.2 改进卷积网络

风电集群中的机组参数种类繁多，且数据呈现出显著的时间异质性和多样性。在仅依靠相关性分析进行特征筛选时，往往难以全面捕捉到对功率预测至关重要的关键特征。这种方法可能忽视了非线性和高阶关联，导致信息丢失并降低模型的预测精度。与此同时，若直接使用所有参数进行建模，高维数据会导致模型复杂度大幅增加，进而引发“维度灾难”，限制模型的性能表现。

尽管传统卷积神经网络(convolutional neural network, CNN)在特征提取方面具有优势，但随着网络深度的增加，模型的参数量显著增长，容易导致过拟合问题。针对以上问题，本文对 CNN 架构进行的改进主要从以下两方面着手：一是引入注意力机制，通过动态调整特征权重，突出关键因素的影响，提高模型捕获关键信息的能力；二是利用权重衰减等方法抑制参数冗余，降低模型复杂度并提升泛化性能。

#### 2.2.1 注意力学习强化

注意力机制通过为输入特征分配不同的重要性权重，优化特征选择和加权处理。本文引入双重注意力机制(dual attention mechanism, DAM)，结合时间和通道维度的注意力权重，提升模型对时空序列数据的关键信息捕捉能力。

首先，输入数据通过乱序操作进行重排，打破特征之间的固有关联性。随后，通过注意力机制来分配权重。对于输入特征序列 $\boldsymbol{S}=[\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_N]$，注意力机制通过计算特征的重要性得分 $e_j$，进而得到权重系数 $\alpha_j$，其计算式为：

$$\alpha_j=\frac{\exp(e_j)}{\sum_{j\le N}\exp(e_j)} \tag{6}$$

式中：$N$ 为特征序列长度；$e_j$ 通过前馈神经网络获得，加权后的特征向量表示为：

$$\hat{\boldsymbol{s}}_j=\alpha_j\boldsymbol{s}_j \tag{7}$$

这种加权方式使得模型聚焦于与功率预测密切相关的特征，减少冗余信息，增强模型预测能力。

而 DAM 通过同时关注时间和特征(通道)两个维度，能够更全面地捕捉时间序列数据中的关键信息。具体来说，时间注意力机制(temporal attention)旨在识别和加强不同风电场之间的时序关联，进而优化时序信息的建模；而通道注意力机制(channel attention)则关注对风电预测最为关键的输入特征，以优化特征选择过程[22]。这种双重聚焦的策略使得模型能够动态调整对不同时间步和特征的关注程度，显著提升预测的准确性和鲁棒性，其架构图如图 3 所示。

对于时间注意力，给定输入特征矩阵 $\boldsymbol{A}\in\mathbf{R}^{T\times F}$($T$ 为特征序列长度，$F$ 为特征维度)，通过时间注意力网络生成时间权重 $\boldsymbol{a}_\mathrm{t}$，计算式为：

$$\boldsymbol{a}_\mathrm{t}=f(\boldsymbol{W}_\mathrm{t}\cdot\boldsymbol{A}+\boldsymbol{b}_\mathrm{t}) \tag{8}$$

式中：$f(\cdot)$表示 softmax 函数；$\boldsymbol{W}_\mathrm{t}$、$\boldsymbol{b}_\mathrm{t}$为可学习的参数。类似地，对于通道注意力，生成通道权重 $\boldsymbol{a}_\mathrm{c}$ 的计算式为：

$$\boldsymbol{a}_\mathrm{c}=f(\boldsymbol{W}_\mathrm{c}\cdot\boldsymbol{A}^\mathrm{T}+\boldsymbol{b}_\mathrm{c}) \tag{9}$$

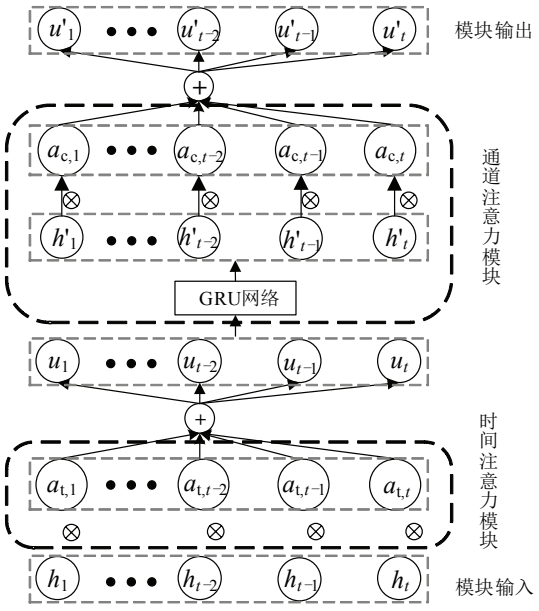



图 3 双重注意力机制架构图

Fig.3 Architecture diagram of dual-attention mechanism

式中：$\boldsymbol{W}_{\mathrm{c}}$、$\boldsymbol{b}_{\mathrm{c}}$为可学习的参数。最终的注意力网络输出是原始特征矩阵与 $\boldsymbol{a}_{\mathrm{t}}$、$\boldsymbol{a}_{\mathrm{c}}$的加权组合。

通过在学习强化方面引入双重注意力机制，结合时间和通道维度的注意力权重，可有效提升模型对时空序列数据的关键信息捕捉能力。

#### 2.2.2 权重衰减

权重衰减(Weight Decay)是常见的正则化技术，旨在通过对模型权重施加惩罚，避免模型过度拟合训练数据，尤其在处理高维数据时，效果更显著。具体而言，权重衰减通过在损失函数中引入与权重大小相关的惩罚项(其 L2 范数)，有效限制了权重参数的增大，从而防止模型在复杂特征空间中过拟合[23]。其基本优化目标为：

$$L_{\mathrm{reg}} = L + \lambda \sum_{d=1}^{m} \left\| \theta_d \right\|^2 \tag{10}$$

式中：$L$ 为原始损失函数；$\lambda$ 为正则化系数；$\theta_d$ 为第 $d$ 个模型参数，$m$ 为参数总数。通过控制 $\lambda$，可在防止模型过拟合的同时减小维度带来的计算负担。

### 2.3 时空门控机制模块

风电出力的时序数据通常具有周期性和长期趋势，传统的时序模型难以捕捉这种复杂的时间依赖性。为增强对风电场间空间依赖性的建模，在两层注意力之间形成信息流的有效传递与整合方式，优化特征和时间维度的双重信息处理能力，本文在网络中引入门控融合机制模块。该模块基于门控循环单元(gate recurrent unit, GRU)，能够动态选择风电场间信息传递强度，从而更好地处理风电出力数据中的长短期时间依赖性[24]。

GRU 网络通过重置门(reset gate)和更新门(update gate)控制信息的流动，以灵活地捕捉长期和短期依赖关系，如图 4 所示。

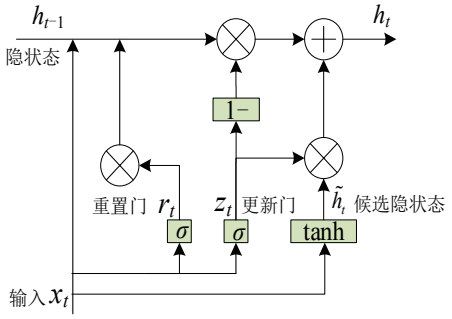



图 4 典型 GRU 网络架构

Fig.4 Typical GRU network architecture

GRU 的计算式如下：

$$r_t = \sigma(x_t \boldsymbol{W}_{\mathrm{xr}} + h_{t-1} \boldsymbol{W}_{\mathrm{hr}} + \boldsymbol{b}_{\mathrm{r}}) \tag{11}$$

$$z_t = \sigma(x_t \boldsymbol{W}_{\mathrm{xz}} + h_{t-1} \boldsymbol{W}_{\mathrm{hz}} + \boldsymbol{b}_{\mathrm{z}}) \tag{12}$$

$$\tilde{h}_t = \tanh(x_t \boldsymbol{W}_{\mathrm{xh}} + (r_t \odot h_{t-1}) \boldsymbol{W}_{\mathrm{hh}} + \boldsymbol{b}_{\mathrm{h}}) \tag{13}$$

$$h_t = z_t \odot \tilde{h}_t + (1 - z_t) \odot h_{t-1} \tag{14}$$

式中：$\boldsymbol{W}_{\mathrm{xr}}$、$\boldsymbol{W}_{\mathrm{xz}}$、$\boldsymbol{W}_{\mathrm{xh}}$分别是与当前输入 $x_t$ 相乘的，用于重置门、更新门和候选状态的权重矩阵；$\boldsymbol{W}_{\mathrm{hr}}$、$\boldsymbol{W}_{\mathrm{hz}}$、$\boldsymbol{W}_{\mathrm{hh}}$分别是与上一时刻隐藏状态 $h_{t-1}$ 相乘的，用于重置门、更新门和候选状态的权重矩阵；$\boldsymbol{b}_{\mathrm{r}}$、$\boldsymbol{b}_{\mathrm{z}}$、$\boldsymbol{b}_{\mathrm{h}}$ 为用于重置门、更新门和候选状态的偏置参数矩阵；$\sigma$ 为 sigmoid 激活函数；$r_t$ 为重置门状态；$z_t$ 为更新门状态；$\tilde{h}_t$ 为候选隐藏状态；$h_t$ 为隐藏层的最终输出；tanh 为非线性激活函数，确保候选隐状态中的值保持在区间(−1,1)中；⊙为 Hadamard 乘积。

空间门控机制进一步增强了风电场之间的信息传递能力，通过门控系数 $g_{xy}$ 动态调整不同风电场之间的信息传递强度。其更新式为：

$$g_{xy} = \sigma(W_{\mathrm{g}} \left[ h_x \left\| h_y \right. \right]) \tag{15}$$

式中：$W_{\mathrm{g}}$为其可学习参数；$[\cdot \| \cdot]$表示向量拼接操作；$g_{xy}$为风电场 $x$ 和 $y$ 之间的门控系数。

## 3 算例分析

### 3.1 仿真设置

为验证所提模型在风电出力预测中的有效性，本文选择北方某集群中的 6 处风电场作为实验对象，采用 2023 年 7—12 月期间的风电出力数据与气象数据进行模型训练与验证。实验中，数据集按照 7:3 的比例划分为训练集和测试集，最大训练轮次设定为 100，初始学习率为 0.01，每 400 轮衰减为 0.1 倍，参数设置如表 1 所示。其中包含对比模型：鲸鱼优化算法-模拟退火-极限学习机(whale optimization algorithm- simulated annealing- extreme learning machine, WOA-SA-ELM)，时间卷积-双向长短期记忆网络(temporal convolutional network-bi-directional long short-term memory, TCN-BiLSTM)与 LSTM-FCN。

为了全面评估模型的预测性能，本文采用了均方根误差(root mean square error, RMSE)$R_{\mathrm{RMAE}}$和平均绝对误差(mean absolute error, MAE)$R_{\mathrm{MAE}}$等评价指标，旨在量化模型在实际应用中的预测精度和误差表现。RMSE 反映预测值与真实值的偏差，数值

越小表示预测越精确，模型稳定性越高；MAE 衡量预测值偏离真实值的平均程度，数值越小预测效果越好；决定系数(coefficient of determination, $R^2$)评估模型拟合程度，越接近 1 表明拟合效果越好，相关性越强；相对预测误差(relative prediction error, RPD)$R_{\mathrm{RPD}}$ 则用于衡量模型准确性，RPD 大于 2 通常表示良好的预测能力；平均绝对百分比误差(mean absolute percentage error, MAPE)$R_{\mathrm{MAPE}}$ 值越小表示预测结果的相对偏差越小，能够直观反映预测结果的相对准确性。其表达式分别为：

$$R_{\mathrm{RMSE}}=\sqrt{\frac{1}{M}\sum_{q=1}^{M}(\overline{u}_q-u_q)^2} \tag{16}$$

$$R_{\mathrm{MAE}}=\frac{1}{M}\sum_{q=1}^{M}\left|\overline{u}_q-u_q\right| \tag{17}$$

$$R^2=1-\frac{\sum_{q=1}^{M}(u_q-\overline{u})^2}{\sum_{q=1}^{M}(\overline{u}_q-\overline{u})^2} \tag{18}$$

$$R_{\mathrm{RPD}}=\frac{\sigma_{\mathrm{obs}}}{R_{\mathrm{RMSE}}} \tag{19}$$

$$R_{\mathrm{MAPE}}=\frac{1}{M}\sum_{q=1}^{M}\frac{\left|u_q-\overline{u}_q\right|}{\left|\overline{u}_q\right|}\times 100\% \tag{20}$$

式中：$M$ 为用于评估的预测点总数；$\overline{u}_q$ 为 $q$ 时刻的真实值；$u_q$ 为预测值；$\overline{u}$ 是真实值的均值；$\sigma_{\mathrm{obs}}$ 为真实值的标准差。

### 3.2 数据预处理分析

为提升模型的预测精度，本文重点研究风电场 1，并通过计算 MIC 系数、构建邻接矩阵精确表征风电的时空特性，以构建包含气象特征与空间结构的复合输入体系。具体的 MIC 系数与 Pearson 矩阵如图 5、图 6 所示。

基于 MIC 系数分析筛选，模型输入特征选择为 10 m、30 m、50 m、轮毂高度的风速与风角度以及温湿度等气象数据，旨在捕捉风电场不同层次的风资源及其对出力预测的影响。同时该数值可作为时间权重 $\boldsymbol{a}_{\mathrm{t}}$ 的初始值依据，强化其关键信息聚焦效能。

如图 6 直观反映了各风电场之间的功率相关性。风电场 1 与风电场 4 的相关性较高(0.67)，说明两者可能受到相似气象条件的影响；与风电场 2 的相关性较低(0.20)，表明两者存在较大差异。将此 Pearson 系数归一化后作为通道权重 $\boldsymbol{a}_{\mathrm{c}}$ 的初始值依据，使高相关通道的初始权重更高，这使模型在训练初期即对协同性强的风电场产生注意力偏置，抑制低相关风电场的冗余信息干扰。同时，该矩阵作为空间结构特征输入至后续时空门控机制，通过邻接矩阵的边权显式定义风电场间的信息传递强度，实现对空间关系的结构化建模。

表 1 仿真参数设置表

Tabel 1 Simulation parameter settings

| 模型 | 参数名称 | 参数值/设置 |
|---|---|---|
| WOA-SA-ELM | ELM 隐藏层神经元数 | 100 |
| | 模拟退火初始温度 | 100 |
| | 鲸鱼优化最大迭代次数 | 100 |
| TCN-BILSTM | TCN 卷积核大小 | 3 |
| | TCN 滤波器数量 | 16 |
| | BILSTM 单元数 | 6 |
| LSTM-FCN | 卷积层 1 | ([1,1], 32) |
| | 卷积层 2 | ([1,1], 64) |
| | LSTM 单元数 | 50 |
| 本文模型 | 卷积层 1 | ([1,1], 64) |
| | 卷积层 2 | ([1,1], 128) |
| | GRU 单元数 | 100 |

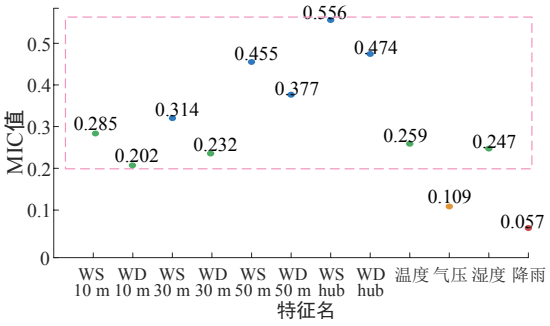



图 5 各特征与风电出力间的 MIC 系数点图

Fig.5 Dot plot of MIC coefficients between various features and wind power output

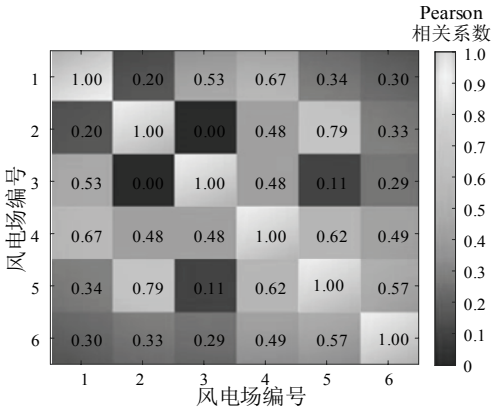



图 6 Pearson 相关系数矩阵

Fig.6 Pearson correlation coefficient matrix

### 3.3 模型对比与结果分析

为了对比验证本文所提方法的有效性，仿真实验选取预测模型 WOA-SA-ELM、TCN-BILSTM、LSTM-FCN 作为性能测试对比。具体如表 2 所示。

通过表 2 数据及图 7 所示预测对比图可见，本文模型在多个指标上占优，RMSE 较 3 个对比模型分别优化了约 1.75、1.68、0.25，MAE 分别优化了约 1.51、1.43、0.26，MAPE 较 3 者低 0.66、1.87、0.21 个百分点，这表明预测值与真实值的偏差和误差幅度较小；在拟合指标 $R^2$ 上，分别优化了约 0.05、0.04、0.006，RPD 则分别优化了 1.19、1.06、0.28，这意味着对集群短期功率变化的解释能力更强，预测精度具明显优势。

相比之下，WOA-SA-ELM 因 ELM 的随机初始化权重机制，易陷入局部最优解，且 WOA 与 SA 的搜索策略和参数调整机制难以适应风电数据的动态变化。TCN-BILSTM 在处理长序列时，TCN 的卷积结构导致信息传递损失，BILSTM 虽能双向学习序列信息，但无法有效应对风电数据的非平稳性和强噪声特性。LSTM-FCN 虽然 LSTM 在处理时间序列数据时有一定能力，但面对复杂多变的风电数据，捕捉长期依赖关系的能力受限，FCN 在特征提取与整合上也难以满足风电数据复杂模式的需求。

如图 8 所示误差对比图进一步展示了各模型的误差表现，本文模型低误差区间内占比明显高于其他模型，误差控制能力显著占优，尤其对较大误差的抑制效果显著。这表明本文模型不仅能够精确预测风电出力，还能保持预测结果的稳定性和一致性。

为研究本文模型在不同规模训练数据下的泛化能力，设计验证环节为：首先根据 3.1 节仿真设置中所述划分训练集、测试集后，对训练集采用随机分层抽样方法提取其 30%、50%、70%的子集，确保各规模子集保留原始数据的时间分布特征。所有模型在不同规模训练集上训练时，保持参数设置不变，最终均在测试集上评估性能，结果见表 3。

据此可直观看出，随着数据集规模增大，WOA-SA-ELM 模型的 $R^2$ 值先升后降，TCN-BILSTM 模型提升幅度有限，而本文模型则从 0.916 9 稳步攀升至 0.950 2，拟合效果持续提升，表明本文模型对不同规模数据的适应能力更强，泛化性优势显著；同时，WOA-SA-ELM 和 TCN-BILSTM 等模型的 RMSE 值波动明显，本文模型的 RMSE 值在 30%时为 3.391 8，50%时上升至 5.784 7，70%时降至

表 2 预测模型结果对比表

Tabel 2 Comparison table of predictive modeling results

| 模型 | RMSE | MAE | $R^2$ | MAPE/% | RPD |
|---|---|---|---|---|---|
| WOA-SA-ELM | 6.746 1 | 4.940 2 | 0.897 8 | 7.46 | 3.131 3 |
| TCN-BILSTM | 6.681 2 | 4.851 4 | 0.905 3 | 8.67 | 3.254 4 |
| LSTM-FCN | 5.245 8 | 3.685 5 | 0.938 2 | 7.01 | 4.038 8 |
| 本文模型 | 4.995 9 | 3.424 4 | 0.943 9 | 6.80 | 4.319 4 |

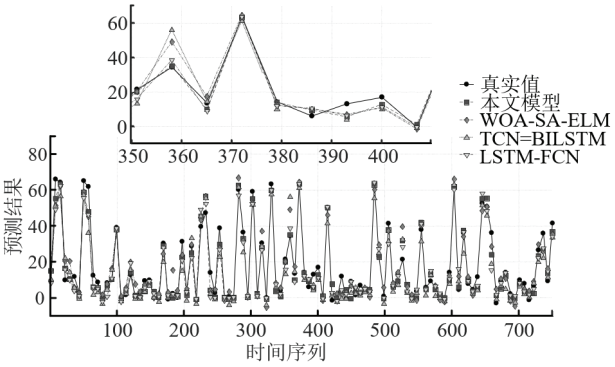



图 7 预测结果对比图

Fig.7 Comparison chart of predicted results

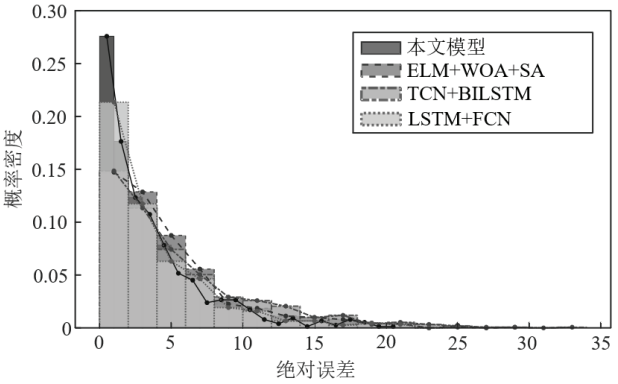



图 8 预测结果绝对误差对比图

Fig.8 Comparison of absolute errors in prediction results

表 3 预测模型在 30%、50%、70%训练集下的指标对比表

Table 3 Comparison table of indicators for prediction models under 30%, 50%, and 70% training sets

| 模型 | 数据集 | RMSE | MAE | $R^2$ | MAPE/% | RPD |
|---|---|---|---|---|---|---|
| WOA-SA-ELM | 30% | 3.905 1 | 2.828 1 | 0.889 9 | 10.41 | 3.013 7 |
| | 50% | 6.017 1 | 4.265 4 | 0.913 7 | 7.14 | 3.411 2 |
| | 70% | 6.369 8 | 4.754 4 | 0.906 1 | 7.67 | 3.277 6 |
| TCN-BILSTM | 30% | 4.728 6 | 2.992 3 | 0.888 1 | 9.03 | 3.006 5 |
| | 50% | 6.835 2 | 4.896 7 | 0.889 6 | 7.17 | 3.163 6 |
| | 70% | 6.993 4 | 4.866 5 | 0.896 7 | 7.95 | 3.193 4 |
| LSTM-FCN | 30% | 3.555 5 | 2.404 3 | 0.908 7 | 10.07 | 3.312 8 |
| | 50% | 5.721 2 | 4.077 3 | 0.922 0 | 7.37 | 3.741 3 |
| | 70% | 5.577 5 | 3.850 3 | 0.928 0 | 7.33 | 4.009 9 |
| 本文模型 | 30% | 3.391 8 | 2.148 3 | 0.916 9 | 9.51 | 3.710 1 |
| | 50% | 5.784 7 | 4.080 9 | 0.920 3 | 7.08 | 3.683 1 |
| | 70% | 4.639 3 | 3.118 8 | 0.950 2 | 7.32 | 4.523 1 |

4.639 3，整体波动幅度远小于其他模型；在 MAPE 指标上，本文模型在 30%、50%、70%数据集下的 MAPE 值分别为 9.51%、7.08%、7.32%，相较于 WOA-SA-ELM 和 TCN-BILSTM 等模型，其波动幅度更小且数值整体更低。这说明在面对数据集规模这一扰动时，本文模型的预测误差率的稳定性更优，鲁棒性更高。

### 3.4 消融实验

为进一步验证本文模型各关键组件的有效性，通过消融实验验证时间注意力机制、通道注意力机制和 GRU 模块在整体模型中的作用。实验结果如表 4、图 9 所示。

基础模型 CNN-GRU 虽然能够捕捉一定的时空特征，但因对特征间复杂关系建模能力的限制，其 RMSE 和 MAE 分别为 7.018 4 和 5.069，$R^2$ 仅为 0.838 9，预测精度存在提升空间。在加入注意力机制(attention, ATT)后，模型对时空特征的捕捉能力显著提高，RMSE 降低约 0.14，MAE 减少约 0.1，表明注意力机制在提升预测准确性方面发挥了积极作用。但此时对 MAPE 的改进不足，且 $R^2$ 和 RPD 提升有限，说明单一的注意力机制难以充分挖掘多维特征间的内在关联。相比之下，本文模型引入双重注意力，结合时间和通道两个维度的特征权重优化，各项指标得到有效优化，尤其 RMSE 降幅达到约 1.88，MAE 降幅近 1.55，MAPE 降低近 0.64%，充分验证了双重注意力机制在深度建模特征交互关系中的重要作用。

消融实验结果表明，本文模型的优越性来源于时间注意力和通道注意力机制的协同作用，以及 GRU 模块对时序特征的有效建模。各模块的有机融合显著增强了模型对复杂风电场环境下风电出力的预测能力，为时空多维特征的深度挖掘提供了有效解决方案。

### 3.5 超参数变化对预测结果影响

利用测试集数据对预测过程中影响模型的两个超参数——学习率和正则化参数 $\lambda$ 进行研究。每个实验重复 10 次，在测试集上统计平均 RMSE、MAE、$R^2$、RPD。图 10、11 为不同超参数选择下得到的误差箱图。

从图 10 可以观察到，随着学习率的逐渐减小，模型的预测误差呈现出先减小后增大的趋势。当学习率为 0.01 时，模型的预测性能达到最佳。相反，当学习率设置为 0.1 时，模型的更新步伐过快，导致权重在优化过程中跳过最优解区域，进而导致预测性能下降；而当学习率设置为 0.000 1 时，优化过程中的每次更新幅度过小，未能有效调整权重，从而影响模型的优化效果。图 11 的结果表明，随着正则化参数 $\lambda$ 的变化，模型的预测误差变化并不明显，但在 $\lambda$ 值为 0.001 附近预测结果更为稳定。

表 4 消融实验预测模型结果对比表

Tabel 4 Comparison table of predictive modeling results for ablation experiments

| 模型 | RMSE | MAE | $R^2$ | MAPE/% | RPD |
|---|---|---|---|---|---|
| CNN-GRU | 7.018 4 | 5.069 | 0.838 9 | 7.44 | 2.920 5 |
| CNN-GRU-ATT | 6.881 4 | 4.970 5 | 0.900 7 | 7.45 | 3.181 7 |
| 本文模型 | 4.995 9 | 3.424 4 | 0.943 9 | 6.81 | 4.319 4 |

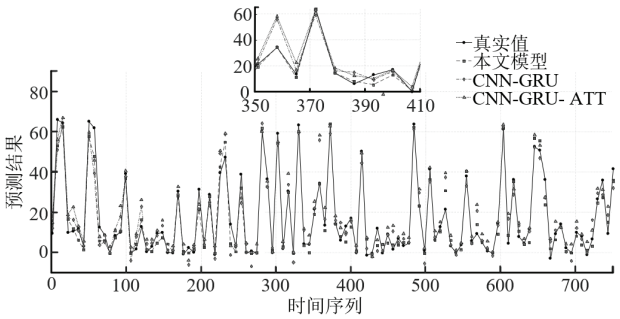
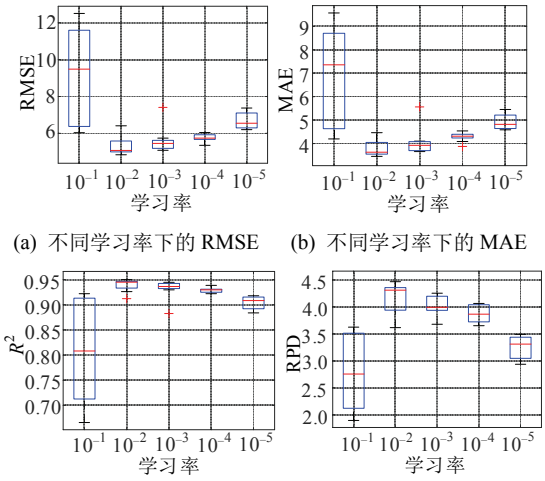



图 9 消融实验预测对比图

Fig.9 Comparison of ablation experiment predictions




(a) 不同学习率下的 RMSE (b) 不同学习率下的 MAE

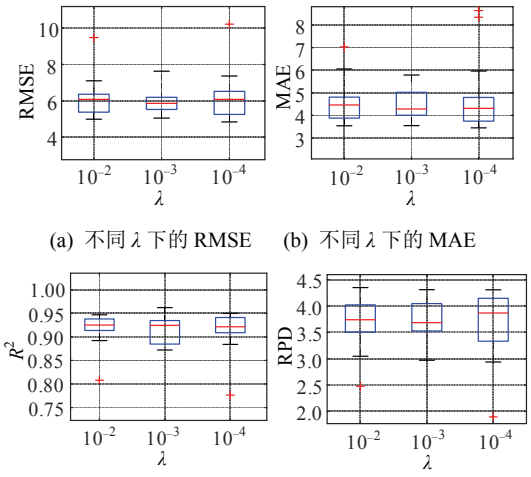
(c) 不同学习率下的 $R^2$ (d) 不同学习率下的 RPD

图 10 不同学习率下的评价指标箱形图

Fig.10 Box plots of evaluation indicators at different learning rates

## 4 结论

本文针对风电集群功率短期预测中面临的时空依赖性和风电场间协同机制不明问题，提出了一




(a) 不同 $\lambda$ 下的 RMSE (b) 不同 $\lambda$ 下的 MAE

(c) 不同 $\lambda$ 下的 $R^2$ (d) 不同 $\lambda$ 下的 RPD

图 11 不同 $\lambda$ 下的评价指标箱形图

Fig.11 Box plots of evaluation indicators at different $\lambda$

种基于双重注意力机制与时空门控融合的预测模型。主要结论如下：

1）模型突破传统特征筛选方法的局限，利用 MIC 与 Pearson 相关系数的线性-非线性双维度特征分析方法，实现了非线性气象特征与风电场间动态关联强度的量化解析，使得输入特征的物理可解释性增强，为后续建模奠定了坚实基础。

2）所提出双重注意力机制结合时空门控融合的模型架构，通过注意力强化-参数抑制的优化策略，经过时间-通道双注意力模块捕获跨时间步的特征权重分布，并结合时空门控机制模块实现时空特征动态融合，形成了具有自适应特征选择能力的预测机制。

3）实验验证表明，本文模型在 RMSE、MAE、$R^2$、MAPE、RPD 5 项评估指标上均优于一般预测模型 (WOA-SA-ELM、TCN-BILSTM 和 LSTM-FCN)，尤其在控制预测误差方面表现良好；同时对比之下，本文模型在泛化性和鲁棒性方面展现出更为突出的特性，在不同规模数据集下，均能保持良好的性能表现。消融实验的结果也证明了各个模块对提升预测精度和模型鲁棒性的关键作用。

## 参考文献 References

[1] 国家能源局. 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. The National Energy Administration released the statistical data of the national power industry from January to August 2024[J]. Electric Power Environmental Protection, 2024(9): 39.

[2] 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)：2806-2818.
ZHANG Zhigang, KANG Chongqing. Challenges and prospects for constructing the new-type power system towards a carbon neutrality future[J]. Proceedings of the CSEE, 2022, 42(8): 2806-2818.

[3] 韩自奋，景乾明，张彦凯，等. 风电预测方法与新趋势综述[J]. 电力系统保护与控制，2019，47(24)：178-187.
HAN Zifen, JING Qianming, ZHANG Yankai, et al. Review of wind power forecasting methods and new trends[J]. Power System Protection and Control, 2019, 47(24): 178-187.

[4] 文云峰，杨游航，邢鹏翔，等. 多维因素制约下新能源消纳能力评估方法研究综述[J]. 中国电机工程学报，2024，44(1)：127-147.
WEN Yunfeng, YANG Youhang, XING Pengxiang, et al. Review on the new energy accommodation capability evaluation methods considering multi-dimensional factors[J]. Proceedings of the CSEE, 2024, 44(1): 127-147.

[5] LIU T H, WEI H K, ZHANG K J. Wind power prediction with missing data using Gaussian process regression and multiple imputation[J]. Applied Soft Computing, 2018, 71: 905-916.

[6] 牛东晓，纪会争. 风电功率物理预测模型引入误差量化分析方法[J]. 电力系统自动化，2020，44(8)：57-65.
NIU Dongxiao, JI Huizheng. Quantitative analysis method for errors introduced by physical prediction model of wind power[J]. Automation of Electric Power Systems, 2020, 44(8): 57-65.

[7] WANG C, ZHANG H L, MA P. Wind power forecasting based on singular spectrum analysis and a new hybrid Laguerre neural network[J]. Applied Energy, 2020, 259: 114139.

[8] 陈 峰，余 轶，徐敬友，等. 基于 Bayes-LSTM 网络的风电出力预测方法[J]. 电力系统保护与控制，2023，51(6)：170-178.
CHEN Feng, YU Yi, XU Jingyou, et al. Prediction method of wind power output based on a Bayes-LSTM network[J]. Power System Protection and Control, 2023, 51(6): 170-178.

[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 Tao, GAO Jing, WANG Youyin, 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] 程 凯，彭小圣，徐其友，等. 基于特征选择与多层级深度迁移学习的风电场短期功率预测[J]. 高电压技术，2022，48(2)：497-503.
CHENG Kai, PENG Xiaosheng, XU Qiyou, et al. Short-term wind power prediction based on feature selection and multi-level deep transfer learning[J]. High Voltage Engineering, 2022, 48(2): 497-503.

[11] 马志远，王 勃，杨 茂，等. 基于二维风速修正和多重集成的两阶段迁移学习短期风电功率预测[J]. 高电压技术，2024，50(9)：3934-3943.
MA Zhiyuan, WANG Bo, YANG Mao, et al. Two-stage transfer learning short-term wind power prediction based on two-dimensional wind speed correction and multiple integration[J]. High Voltage Engineering, 2024, 50(9): 3934-3943.

[12] 刘吉成，朱玺瑞，于 晶. 基于 IWOA-SA-Elman 神经网络的短期风电功率预测[J]. 太阳能学报，2024，45(1)：143-150.
LIU Jicheng, ZHU Xirui, YU Jing. Short-term wind power prediction based on IWOA-SA-Elman neural network[J]. Acta Energiae Solaris Sinica, 2024, 45(1): 143-150.

[13] 苗长新，王 霞，李 昊，等. 基于数值天气预报风速误差修正的风电功率日前预测[J]. 电网技术，2022，46(9)：3455-3462.
MIAO Changxin, WANG Xia, LI Hao, et al. Day-ahead prediction of wind power based on NWP wind speed error correction[J]. Power System Technology, 2022, 46(9): 3455-3462.

[14] 李　聪，彭小圣，王皓怀，等. 基于 SDAE 深度学习与多重集成的风电集群短期功率预测[J]. 高电压技术，2022，48(2)：504-512.
LI Cong, PENG Xiaosheng, WANG Haohuai, et al. Short-term power prediction of wind power cluster based on SDAE deep learning and multiple integration[J]. High Voltage Engineering, 2022, 48(2): 504-512.

[15] 王一妹，刘　辉，宋　鹏，等. 基于高斯混合模型聚类的风电场短期功率预测方法[J]. 电力系统自动化，2021，45(7)：37-43.
WANG Yimei, LIU Hui, SONG Peng, et al. Short-term power forecasting method of wind farm based on Gaussian mixture model clustering[J]. Automation of Electric Power Systems, 2021, 45(7): 37-43.

[16] 杨　茂，彭　天，苏　欣. 基于预测信息二维坐标动态划分的风电集群功率超短期预测[J]. 中国电机工程学报，2022，42(24)：8854-8863.
YANG Mao, PENG Tian, SU Xin. Ultra-short term wind power prediction based on two-dimensional coordinate dynamic division of prediction information[J]. Proceedings of the CSEE, 2022, 42(24): 8854-8863.

[17] 李　卓，叶　林，戴斌华，等. 基于 IDSCNN-AM-LSTM 组合神经网络超短期风电功率预测方法[J]. 高电压技术，2022，48(6)：2117-2127.
LI Zhuo, YE Lin, DAI Binhua, et al. Ultra-short-term wind power prediction method based on IDSCNN-AM-LSTM combination neural network[J]. High Voltage Engineering, 2022, 48(6): 2117-2127.

[18] 张昊立，张　菁. 倪建辉，等. 引入注意力机制的 LSTM-FCN 海上风电功率预测[J]. 太阳能学报，2024，45(6)：444-450.
ZHANG Haoli, ZHHAG Jing, NI Jianhui, et al. LSTM-FCN offshore wind power forecasting with introduction of attention mechanism[J]. Acta Energiae Solaris Sinica, 2024, 45(6): 444-450.

[19] 张浩田，温　蜜，李晋国，等. 数据驱动的时间注意力卷积风电功率预测模型[J]. 太阳能学报，2022，43(10)：167-176.
ZHANG Haotian, WEN Mi, LI Jinguo, et al. Data driven time attention convolution wind power prediction model[J]. Acta Energiae Solaris Sinica, 2022, 43(10): 167-176.

[20] MULEWA S, PARMAR A M, DE A. Attention based Transformer coupled with convoluted neural network for ultra-short- and short-term forecasting of multiple wind farms[J]. International Journal of Green Energy, 2024, 21(6): 1238-1252.

[21] 杨　建，王　力，宋冬然，等. 基于孤立森林与稀疏高斯过程回归的风电机组偏航角零点漂移诊断方法[J]. 中国电机工程学报，2021，41(18)：6198-6211.
YANG Jian, WANG Li, SONG Dongran, et al. Diagnostic method of zero-point shifting of wind turbine yaw angle based on isolated forest and sparse Gaussian process regression[J]. Proceedings of the CSEE, 2021, 41(18): 6198-6211.

[22] 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[J]. 电力系统自动化，2022，46(7)：141-151.
SU Xiangjing, ZHOU Wenxin, LI Chaojie, et al. Interpretable offshore wind power output forecasting based on long short-term memory neural network with dual-stage attention[J]. Automation of Electric Power Systems, 2022, 46(7): 141-151.

[23] CAMAL S, GIRARD R, FORTIN M, et al. A conditional and regularized approach for large-scale spatiotemporal wind power forecasting[J]. Sustainable Energy Technologies and Assessments, 2024, 65: 103743.

[24] 吉兴全，赵国航，叶平峰，等. 基于 QMD-HBiGRU 的短期光伏功率预测方法[J]. 高电压技术，2024，50(9)：3850-3859.
JI Xingquan, ZHAO Guohang, YE Pingfeng, et al. Short-term PV forecasting method based on the QMD-HBiGRU model[J]. High Voltage Engineering, 2024, 50(9): 3850-3859.



**孟高军**
1987—，男，博士，教授，硕导
主要研究方向为新能源发电与并网、电能变换及存储
E-mail: gjun_m@126.com

MENG Gaojun
Ph.D., Professor



**吴佳凤**(通信作者)
2000—，女，硕士生
主要研究方向为新能源发电与并网控制技术、电力系统运行与控制
E-mail: jiaf_wu@163.com

WU Jiafeng
Corresponding author

收稿日期　2025-04-03　在线出版日期　2025-07-16　　编辑　陈　蔓