第 52 卷 第 2 期：639-654 高电压技术 Vol.52, No.2: 639-654
2026 年 2 月 28 日 High Voltage Engineering February 28, 2026
DOI: 10.13336/j.1003-6520.hve.20250444

基于分组交叉门控机制图卷积神经网络的源荷超短期功率联合预测方法

赵婉冰，杨　强，陈源奕
（浙江大学电气工程学院，杭州 310027）

摘　要：多场站多类型源荷功率联合预测场景中，传统的时间序列预测模型难以高效提取多类型源荷的时空相关特征和内在隐含联系，导致功率预测精度不足。针对该问题，该文提出一种融合改进图卷积神经网络和分组交叉门控机制的源荷超短期功率联合预测方法。首先，构建了基于改进型时空图卷积神经网络的源荷超短期功率联合预测模型，可充分提取源侧(风电、光伏)与荷侧(电负荷、热负荷)功率的时空关联特征。进而，设计了一种分组交叉门控机制并集成到改进的图卷积网络中，可使不同类型源荷信息得以交叉调制，有效利用了源荷之间的内在隐含联系，从而显著提升了多场站多类型源荷超短期功率联合预测的精度。最后，基于我国北部某地区的风光电源和电热负荷数据进行了对比实验，结果验证了所提方法的有效性和优越性。
关键词：超短期功率预测；功率联合预测；图卷积网络；分组交叉门控机制；时空相关性

Ultra-short-term Source-load Joint Forecasting Method Based on Group Cross-gating Mechanism and Graph Convolutional Neural Network

ZHAO Wanbing, YANG Qiang, CHEN Yuanyi
(College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China)

Abstract: In the scenario of joint power forecasting for multiple sites and multiple types of sources and loads, traditional time-series forecasting models are often difficult to efficiently extract spatiotemporal correlations and inherent relationships among different types of sources and loads, resulting in insufficient forecasting accuracy. To address this issue, this paper proposes an ultra-short-term source-load joint forecasting method that integrates an improved graph convolutional neural network (GCN) with a group cross-gating mechanism. First, an ultra-short-term joint power forecasting model based on the improved GCN is constructed to fully extract the spatiotemporal correlation features of source (wind, solar) and load (electricity, heat) power. Then, a group cross-gating mechanism is designed and incorporated into the improved GCN, thus the cross-modulation of information between different types of sources and loads is realized, effectively utilizing their inherent relationships. This significantly improves the accuracy of ultra-short-term joint power forecasting for multiple sites and multiple types of sources and loads. Finally, based on historical data of wind power, solar power, electric load, and heat load from a region in northern China, comparative experiments are conducted. The results verify the effectiveness and superiority of the proposed method in terms of prediction accuracy.
Key words: ultra-short-term power prediction; joint power prediction; graph convolutional network; group cross-gating mechanism; spatio-temporal correlation

0　引言

随着全球能源系统低碳转型发展逐步深入，风电和光伏等可再生能源已成为电力系统的重要组成部分[1]。然而，风电和光伏受自然条件(如风速、光照强度等)影响显著，其出力呈现不确定性和间歇性，对电力系统安全稳定运行构成了巨大挑战[2]。随着城市化进程加快和居民生活水平提高，电力负荷与热负荷的需求也日益复杂多变，对综合能源系统供需平衡和高效调度提出了更高要求[3]。因此，精确的源荷(风光电源和电热负荷)功率预测，对于提升系统运行可靠性、经济性和灵活性具有重要意义[4-5]。

在综合能源系统中，光伏和风电受自然条件影

基金资助项目：国家自然科学基金(52177119)。
Project supported by National Natural Science Foundation of China (52177119).

响具有波动性，而电、热负荷需求受社会经济活动、气候条件及用户行为模式等多重因素影响，具有高度动态性和不确定性[6-7]。源荷之间由于共享驱动因素和系统运行的动态平衡特性，存在复杂的相互作用和依赖关系。若能够利用这些变量间潜在的耦合关系，构建包含复杂相互作用的预测模型以联合预测多种源荷功率，则有助于提升每种变量预测的准确性，并为优化调度提供更加精确、全面的信息。

近年来，机器学习算法在可再生能源发电和负荷预测领域受到广泛关注。卷积神经网络 (convolutional neural network, CNN)[8-9]、 长短期记忆神经网络(long short-term memory, LSTM)[10-11]、门控循环单元网络(gated recurrent unit, GRU)[12] 、Transformer 网络[13]等神经网络模型的引入，提升了源/荷功率的预测性能。文献[14]采用改进的 LSTM 网络进行超短期风电功率预测，具有良好的时序学习能力。文献[15]构建了基于 CNN-LSTM 的混合模型，提高了负荷预测精度。文献[16-17]分别采用改进的 GRU 和 Transformer 网络建立了短期光伏功率预测模型。然而，上述研究工作主要集中于单一类型电源或负荷的预测，忽视了不同类型源荷之间的相互联系，不具备多类型源荷组合下的功率预测能力。为解决该问题，部分文献针对多种不同变量联合预测开展了研究。文献[18]利用光伏和风电之间的潜在相关性，通过引入混合框架来预测多类型能源发电功率。文献[19]构建了涵盖电、热的综合负荷预测模型。然而，上述模型只考虑了源与源(风电与光伏)之间、荷与荷(电负荷与热负荷)之间的相关性，并未深入挖掘源与荷之间的隐性耦合关系。

上述部分工作所提出的预测模型(如 LSTM[10-11]、GRU[12])使用了门控机制以提升模型对时间序列的处理能力，但仍存在一定局限：模型仅能通过门控单元对单一类型的输入变量在时间维度上进行信息过滤，但在处理多源异构数据时难以充分利用源侧(如光伏、风电)、荷侧(电负荷、热负荷)多种变量的丰富特征，导致不同变量之间缺乏信息交互。特别是风/光多场站功率与电/热多区域负荷联合预测场景下，传统门控机制无法实现对多维度时空信息的精细化提取，限制了功率联合预测精度。

图卷积神经网络(graph convolutional network, GCN)因其具备对图结构数据出色的表征学习能力，在可再生能源分布式发电或负荷预测中展现出优势[20-22]。在图结构中，节点代表实体，边代表实体之间的连接关系。各节点和边都可以有自己的属性或特征。例如，多场站光伏功率预测场景中，各光伏电站可视为图结构中的节点，功率序列和天气序列可以表示为节点的属性。文献[22]提出了一种基于 GCN 的居民用户负荷预测模型，提高了集群负荷预测精度。文献[23]引入了时空图神经网络预测多站点的光伏发电功率。然而，上述研究大多仅关注单一类型的功率预测，难以满足综合能源系统多区域、多类型能源和多时空维度的功率预测需求。

鉴于综合能源系统中源荷功率时空关联和多元变量交互的复杂特性，本文考虑综合能源系统中多种源荷变量间的时空耦合特性，基于多维融合的时空图卷积网络，构建了一种多场站多变量源荷功率联合预测算法模型。通过挖掘多种新能源出力和多类型负荷之间的潜在关联，设计了一种以分组交叉为核心操作的新型门控机制，通过源-源、荷-荷、源-荷间的信息交互，不同变量信息之间得以高效地互相补充、交叉控制。将分组交叉门控机制与时空图卷积网络结合，可有效表达时空结构依赖与变量耦合特性，从而提升多源-多荷-多区域下的功率预测精度。本文结合真实数据，对所提出方法进行了实验验证，并和现有典型深度学习时间序列预测模型进行对比实验，结果验证了所提方法的有效性和优越性。

1 问题分析与模型框架

1.1 问题分析

实际场景中，多场站和区域的源侧(如风电、光伏等可再生能源发电)与负荷侧(包括电负荷与热负荷等)的功率变化并非孤立现象，而是存在多方面的隐性耦合关系。首先，综合能源系统中的各场站和各片区均位于同一地区和时区，它们的发电功率或负荷需求具有很强的时空相关性。该相关性既体现在时间上的同步波动，如在冬季供暖季，各区域的电热负荷会同步攀升；也体现在空间上的联动效应，如区域性强风天气会导致多个风电场出力同时激增，而大范围雾霾则可能使相邻片区光伏电站的发电效率同步下降。对于气象敏感型电源，时空相关性更为突出，例如：阴雨天气下不仅所有光伏场站的辐照度会同步降低，伴随的气压变化还可能引发周边风电场的出力波动。这种时空耦合特性使得综合能源系统的运行需要建立多维协同调控机制。

其次，不同类型的源荷功率之间也存在一定的

相关性，源侧可再生能源出力的波动可以直接或间接地反映负荷侧的用电模式与需求水平[24]。具体而言，可再生能源发电(如风电、光伏)受气象条件(如风速、光照强度、温度等)的显著影响，而这些气象因素同时也是影响负荷需求(尤其是制冷、供暖等负荷)的关键因素。例如，高温天气不仅会增加光伏发电功率，也会提升空调等制冷设备的用电负荷。因此，通过充分挖掘和利用隐性耦合关系，构建可反映各变量间交互机制的联合预测模型，对于提高系统功率预测精度具有重要意义。

1.2　模型框架

根据上述分析，本文提出了一种新型的基于分组交叉门控机制的图卷积神经网络模型(group cross-gating graph convolutional neural network, GCGGN)，用于多个场站或片区的风力发电功率、光伏发电功率、电负荷和热负荷这 4 种变量的联合预测。本节介绍本文提出的 GCGGN 网络的整体网络结构以及输入数据组织形式，网络中各模块的具体结构在第 2.1~2.3 节展开说明，使用 GCGGN 进行预测的整体流程在第 2.4 节阐述。

GCGGN 网络的输入为 4 种源荷变量(风电、光伏、电负荷、热负荷)的功率历史值及相关影响因素的历史值，输出为这 4 种源荷功率在各场站/片区的预测值，网络整体结构如图 1 所示。相比于传统模型，GCGGN 通过分组交叉门控机制，挖掘不同源荷变量间的潜在关联，并结合时空图卷积网络，实现多类型、多站点的联合预测，为综合能源系统调度提供数据基础。此外，GCGGN 具备处理复杂实际场景的能力。例如，当区域综合能源系统包含不同数量的光伏/风电场站或电/热负荷片区、且各待预测量受异质因素影响时，该模型可通过灵活建模异构输入变量，有效解决不同类型源荷空间分布数量差异及影响因素多样性的问题。

如图 1 所示，GCGGN 的输入可分别组织为 4 个三维张量，各张量为整合了时间、空间和通道维度的图结构数据。不同时刻数据沿时间维度排列，不同场站/片区的数据沿空间维度排列，不同影响因素(如风速、光照等)的数据沿通道维度排列。4 类输入变量基于源-荷双侧属性划分为 2 组：源侧包含风力发电与光伏发电信息张量，负荷侧涵盖电负荷与热负荷信息张量，如此分组旨在强化组内变量的关联耦合特性。模型输出同样为 4 个张量，分别对应各风电场、光伏站点发电功率预测值以及各片区电

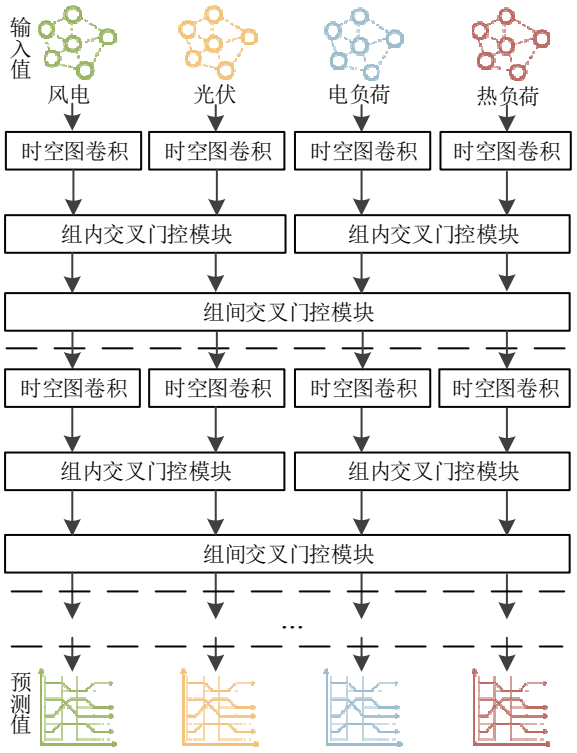


图 1　网络整体结构

Fig.1　Overall network architecture

负荷与热负荷预测结果。网络结构按层级划分(在图 1 中用虚线分隔)，各层级在时间维度做下采样，逐步将数据压缩至单时间点输出。每个层级均由时空图卷积模块、组内交叉门控模块和组间交叉门控模块组成，时空图卷积用于提取时空相关性，2 类交叉门控模块通过变量间交叉调制充分利用其内在联系。各模块具体设计见第 2 章。

2　基于分组交叉门控机制图卷积神经网络的源荷功率联合预测模型

2.1　时空图卷积模块

在 GCGGN 的每个层级中，为了充分提取每种变量自身在各场站/片区和各时间点之间的时空相关性，4 个输入张量(或中间结果特征图)首先被各自的时空图卷积模块(如图 2 所示)处理。

具体而言，每个时空图卷积模块中，输入张量 $\boldsymbol{X}_{\text{in}} \in \mathbb{R}^{v_{\text{in}} \times c_{\text{in}} \times t_{\text{in}}}$ 首先通过神经网络的时间线性投影层(简称时间投影层)进行处理(其中 v_{in}、c_{in} 和 t_{in} 分别表示输入图的节点数、通道数和时间点数)。考虑到不同通道可能具有不同的潜在特性，输入张量的每个通道 $\boldsymbol{X}_{\text{in},i}$ 分别使用不同投影矩阵 $\boldsymbol{W}_{\text{T},i} \in \mathbb{R}^{t_{\text{in}} \times t_{\text{out}}}$ 进行投影(通道索引 $i \in 1,2,\cdots,c_{\text{in}}$)，可表示为：

$$\boldsymbol{X}_{\mathrm{T},i}=\boldsymbol{X}_{\mathrm{in},i}\boldsymbol{W}_{\mathrm{T},i} \tag{1}$$

式中：$\boldsymbol{W}_{\mathrm{T},i}\in\mathbb{R}^{t_{\mathrm{in}}\times t_{\mathrm{out}}}$ 表示第 i 个通道的时间投影矩阵，该矩阵是可学习的神经网络参数，负责将形状为 $v_{\mathrm{in}}\times 1\times t_{\mathrm{in}}$ 的张量 $\boldsymbol{X}_{\mathrm{in},i}$ 映射为一个形状为 $v_{\mathrm{in}}\times 1\times t_{\mathrm{out}}$ 的张量 $\boldsymbol{X}_{\mathrm{T},i}$。在本文实现的网络中，$t_{\mathrm{out}}$ 往往小于 t_{in}，以达到在时间维度上逐步下采样的目的。将每个通道的结果沿着通道维度进行拼接，组成时间投影层的输出张量 $\boldsymbol{X}_{\mathrm{T}}\in\mathbb{R}^{v_{\mathrm{in}}\times c_{\mathrm{in}}\times t_{\mathrm{out}}}$。这一时间投影层用于提取输入数据在时间维度上的相关性。在时间投影层之后，张量被包含空间线性投影层(简称空间投影层)和通道线性投影层(简称通道投影层)的图卷积模块进行处理。其中，空间投影层和上文所述的时间投影层类似，但其作用的维度是空间维度，可表示为：

$$\boldsymbol{X}_{\mathrm{S},i}=\boldsymbol{W}_{\mathrm{S},i}\boldsymbol{X}_{\mathrm{T},i} \tag{2}$$

式中：$\boldsymbol{W}_{\mathrm{S},i}\in\mathbb{R}^{v_{\mathrm{out}}\times v_{\mathrm{in}}}$ 表示第 i 个通道的空间投影矩阵；$\boldsymbol{X}_{\mathrm{T},i}$ 表示时间投影层输出的第 i 个通道。随后，沿着通道维度将输出进行拼接，得到 $\boldsymbol{X}_{\mathrm{S}}\in\mathbb{R}^{v_{\mathrm{out}}\times c_{\mathrm{in}}\times t_{\mathrm{out}}}$。在本文实现的网络中，$v_{\mathrm{in}}=v_{\mathrm{out}}$。

最后，使用通道投影层进行通道维度上的信息交互，可表示为：

$$\boldsymbol{X}_{\mathrm{F}}=\boldsymbol{X}_{\mathrm{S}}\boldsymbol{W}_{\mathrm{F}}+\boldsymbol{B} \tag{3}$$

式中：通道投影矩阵 $\boldsymbol{W}_{\mathrm{F}}$ 为可学习的参数，该层将 $\boldsymbol{X}_{\mathrm{S}}\in\mathbb{R}^{v_{\mathrm{out}}\times c_{\mathrm{in}}\times t_{\mathrm{out}}}$ 映射为 $\boldsymbol{X}_{\mathrm{F}}\in\mathbb{R}^{v_{\mathrm{out}}\times c_{\mathrm{out}}\times t_{\mathrm{out}}}$。通道投影用于在给定时间点整合特定场站/片区不同属性的信息，其参数在不同节点和时间上共享。此外，该层的参数还包括一个可训练的偏置项 $\boldsymbol{B}$。$\boldsymbol{X}_{\mathrm{F}}$ 经过非线性激活之后最终输出本模块。如果将不同通道的时间投影矩阵和空间投影矩阵各自视作一个整体，则上述的几个投影层可总体表示为：

$$\boldsymbol{X}_{\mathrm{out}}=\sigma\left(\boldsymbol{W}_{\mathrm{S}}(\boldsymbol{X}_{\mathrm{in}}\boldsymbol{W}_{\mathrm{T}})\boldsymbol{W}_{\mathrm{F}}+\boldsymbol{B}\right) \tag{4}$$

式中：$\sigma(\cdot)$ 表示非线性激活函数(在具体实现中采用了 ReLU 函数)。该公式具有与标准的图卷积块 $\boldsymbol{X}_{\mathrm{out}}=\sigma\left(\boldsymbol{A}_{\mathrm{gcn}}\boldsymbol{X}_{\mathrm{in}}\boldsymbol{W}\right)$ 相似的形式，其中 $\boldsymbol{A}_{\mathrm{gcn}}=\tilde{\boldsymbol{D}}^{-1/2}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{D}}^{-1/2}$ 是输入图的归一化邻接矩阵，$\tilde{\boldsymbol{A}}$ 表示带自环的邻接矩阵，$\tilde{\boldsymbol{D}}$ 是 $\tilde{\boldsymbol{A}}$ 的度矩阵，$\boldsymbol{W}$ 是权重矩阵。在本文设计的神经网络中，$\boldsymbol{W}_{\mathrm{S},i}$ 可以视为第 i 个通道的 $\boldsymbol{A}_{\mathrm{gcn}}$，其初始化数值为基于气象因素等历史时空特征而构建的归一化邻接矩阵，继而在模型训练阶段通过学习更新。本文邻接矩阵的具体计算方法参见附录 A。

2.2 组内交叉门控模块

在每种类型的变量被各自的时空图卷积模块处理之后，每组 2 种变量的特征图被输入给组内交叉门控模块，如图 2 所示。如前所述，风电功率和光伏功率之间以及电负荷和热负荷之间存在较强的相关性，例如风力和光伏发电功率这组变量高度依赖于局部天气状况，而电负荷和热负荷这组变量都取决于局地的用户行为。而本文提出的组内交叉门控模块(具体结构见图 3)可以视作一种特殊的交叉注意力机制，使得每组 2 个变量的中间结果能够相互交互，控制彼此的信息在神经网络中的流动，进而得到更准确的联合预测结果。

本节以光伏发电信息和风力发电信息的特征图为例来说明组内交叉门控模块的计算步骤(电负荷和热负荷信息的交叉门控在公式形式上与之相同，不再重复说明)。记交叉门控模块的输入为 $\boldsymbol{X}_{\mathrm{solar}}$ 和 $\boldsymbol{X}_{\mathrm{wind}}$，分别为光伏发电信息和风力发电信息对应的 2 个时空图卷积模块输出的特征图。2 者空间维度和通道数有可能不同，如一个分支的输入数据来自 6 个光伏站点，另一个分支输入数据来自 5 个风

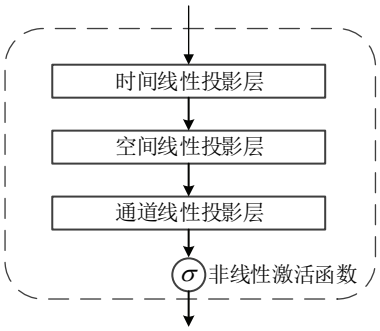


图 2 时空图卷积模块

Fig.2 Spatio-temporal graph convolutional block

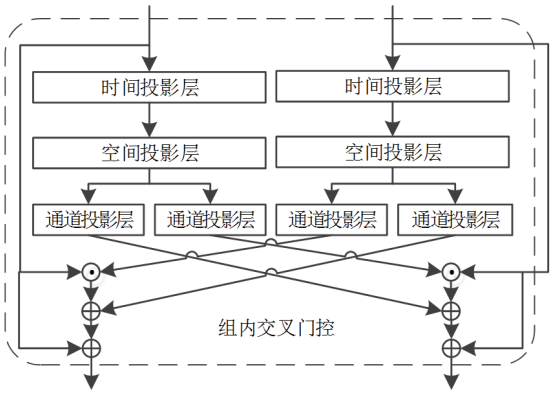


图 3 组内交叉门控模块

Fig.3 Intra-group cross-gating block

电站点，且二者包含的特征数(发电功率和气象变量)也可能存在差异。因此，神经网络先对输入的张量在时间维度进行投影，然后分别对 2 个张量在空间和通道维度上进行投影，表示为：

$$G_{\text{solar1}}\left(\boldsymbol{X}_{\text{solar}}\right)=\boldsymbol{W}_{\text{S,solar}}\left(\boldsymbol{X}_{\text{solar}}\boldsymbol{W}_{\text{T,solar}}\right)\boldsymbol{W}_{\text{F,solar1}} \quad (5)$$

$$G_{\text{solar2}}\left(\boldsymbol{X}_{\text{solar}}\right)=\boldsymbol{W}_{\text{S,solar}}\left(\boldsymbol{X}_{\text{solar}}\boldsymbol{W}_{\text{T,solar}}\right)\boldsymbol{W}_{\text{F,solar2}} \quad (6)$$

$$G_{\text{wind1}}\left(\boldsymbol{X}_{\text{wind}}\right)=\boldsymbol{W}_{\text{S,wind}}\left(\boldsymbol{X}_{\text{wind}}\boldsymbol{W}_{\text{T,wind}}\right)\boldsymbol{W}_{\text{F,wind1}} \quad (7)$$

$$G_{\text{wind2}}\left(\boldsymbol{X}_{\text{wind}}\right)=\boldsymbol{W}_{\text{S,wind}}\left(\boldsymbol{X}_{\text{wind}}\boldsymbol{W}_{\text{T,wind}}\right)\boldsymbol{W}_{\text{F,wind2}} \quad (8)$$

式中：$\boldsymbol{W}$ 为各个线性投影层的参数，下标 S、T 和 F 用以区分作用的维度，下标 solar 和 wind 用以区分作用的张量；$G(\cdot)$ 表示生成门控权重或门控偏置的操作。式(5)—式(8)中多个维度的线性投影层不仅提取了输入张量的特征，还在空间和通道维度上跨分支对齐了 2 个张量的形状，从而支持后续的交叉门控操作。此外，式(5)—式(8)在通道维度上对每个张量分别应用 2 次线性投影，以分别生成门控权重和门控偏置，在式中分别以下标“1”和“2”标记。

随后，交叉门控操作按以下方式进行：

$$\hat{\boldsymbol{X}}_{\text{solar}}=\boldsymbol{X}_{\text{solar}}\odot G_{\text{wind1}}(\boldsymbol{X}_{\text{wind}})+G_{\text{wind2}}(\boldsymbol{X}_{\text{wind}}) \quad (9)$$

$$\hat{\boldsymbol{X}}_{\text{wind}}=\boldsymbol{X}_{\text{wind}}\odot G_{\text{solar1}}(\boldsymbol{X}_{\text{solar}})+G_{\text{solar2}}(\boldsymbol{X}_{\text{solar}}) \quad (10)$$

式中：$\odot$ 表示逐元素乘法；$\hat{\boldsymbol{X}}_{\text{solar}}$ 和 $\hat{\boldsymbol{X}}_{\text{wind}}$ 分别表示经过交叉门控调制后的光伏分支与风电分支特征图。这一步交叉门控操作使 2 个分支的数据相互作用，控制彼此的信息流。最后，将来自本模块输入的残差连接添加到输出中并进行非线性激活，得到本模块的输出：

$$\boldsymbol{Y}_{\text{solar}}=\text{ReLU}(\hat{\boldsymbol{X}}_{\text{solar}}+\boldsymbol{X}_{\text{solar}}) \quad (11)$$

$$\boldsymbol{Y}_{\text{wind}}=\text{ReLU}(\hat{\boldsymbol{X}}_{\text{wind}}+\boldsymbol{X}_{\text{wind}}) \quad (12)$$

式中：ReLU(·) 为非线性激活函数。需要注意的是，本文虽以光伏-风力发电功率的联合预测为例阐述组内交叉门控模块的运算机制，但其架构设计具有普适性，该模块同样适用于电负荷和热负荷这组变量之间的交叉调制。

2.3　组间交叉门控模块

如第 1.1 节所述，不仅每组内的 2 种变量具有较强的关联，而且 2 组变量之间也蕴含跨组的潜在耦合关系。为了捕捉这种联系并将其用于提升预测精度，本文设计了组间交叉门控模块，让 2 组变量互相对对方的特征图进行调制。组间交叉门控模块的输入为风力发电信息、光伏发电信息、电负荷信息和热负荷信息这 4 类变量的特征图(分别记作 $\boldsymbol{Y}_{\text{solar}}$、$\boldsymbol{Y}_{\text{wind}}$、$\boldsymbol{Y}_{\text{elect}}$ 和 $\boldsymbol{Y}_{\text{heat}}$)，输出为经过交叉调制后的这 4 种变量的特征图。具体计算步骤如图 4 所示(简明起见，残差连接在图中省略)，详细解释如下文所述。

首先，利用神经网络的线性层分别对风电信息和光伏信息的特征图在时间和空间维度上进行投影，得到中间结果：

$$\boldsymbol{Y}'_{\text{solar}}=\boldsymbol{W}'_{\text{S,solar}}(\boldsymbol{Y}_{\text{solar}}\boldsymbol{W}'_{\text{T,solar}}) \quad (13)$$

$$\boldsymbol{Y}'_{\text{wind}}=\boldsymbol{W}'_{\text{S,wind}}(\boldsymbol{Y}_{\text{wind}}\boldsymbol{W}'_{\text{T,wind}}) \quad (14)$$

式中：各参数下标的含义与式(5)—式(8)中类似。此处的空间维度上的投影操作使得风电和光伏特征图的空间维度相同，以便下一步的拼接操作。之后拼接操作可表示为：

$$\boldsymbol{Y}_{\text{sw}}=\text{Cat}(\boldsymbol{Y}'_{\text{solar}},\boldsymbol{Y}'_{\text{wind}}) \quad (15)$$

式中：Cat(·) 表示沿通道维度将 2 个张量拼接(concatenate)。易知，张量 $\boldsymbol{Y}_{\text{sw}}$ 同时包含了光伏和风力发电 2 种变量的信息。然后，基于 2 个线性层对该张量进行通道维度上的投影，分别生成用于门控调制电负荷和热负荷特征图的增益值：

$$\boldsymbol{G}_{\text{elect}}=\boldsymbol{Y}_{\text{sw}}\boldsymbol{W}_{\text{G,elect}} \quad (16)$$

$$\boldsymbol{G}_{\text{heat}}=\boldsymbol{Y}_{\text{sw}}\boldsymbol{W}_{\text{G,heat}} \quad (17)$$

式中：$\boldsymbol{W}_{\text{G,elect}}$ 和 $\boldsymbol{W}_{\text{G,heat}}$ 为通道维度上的线性层参数。以上这 2 步通道维度上的投影使得 $\boldsymbol{G}_{\text{elect}}$ 和 $\boldsymbol{G}_{\text{heat}}$ 的通道数分别与电负荷和热负荷的特征图 $\boldsymbol{Y}_{\text{elect}}$ 和 $\boldsymbol{Y}_{\text{heat}}$ 的通道数相同。

而后，$\boldsymbol{Y}_{\text{elect}}$ 和 $\boldsymbol{Y}_{\text{heat}}$ 分别被 $\boldsymbol{G}_{\text{elect}}$ 和 $\boldsymbol{G}_{\text{heat}}$ 调制，并经过残差连接和非线性激活后得到本模块的 2 个输出：

$$\boldsymbol{Z}_{\text{elect}}=\text{ReLU}(\boldsymbol{G}_{\text{elect}}\odot\boldsymbol{Y}_{\text{elect}}+\boldsymbol{Y}_{\text{elect}}) \quad (18)$$

$$\boldsymbol{Z}_{\text{heat}}=\text{ReLU}(\boldsymbol{G}_{\text{heat}}\odot\boldsymbol{Y}_{\text{heat}}+\boldsymbol{Y}_{\text{heat}}) \quad (19)$$

式中：ReLU 函数的数学特性可强制过滤负值输出，从模型架构层面满足了功率非负的物理规律。通过

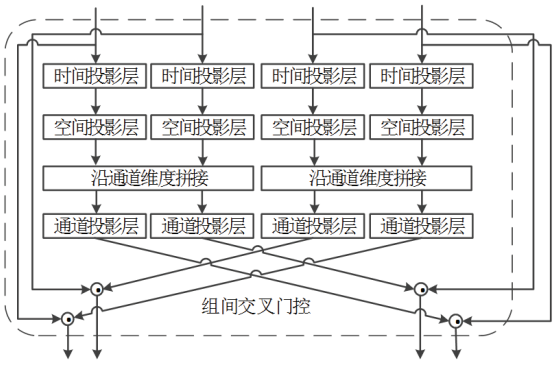


图 4　组间交叉门控模块

Fig.4　Inter-group cross-gating block

式(18)、(19)所表达的组间交叉门控机制，风电和光伏的信息可以调制电负荷和热负荷的预测结果。

上文描述了基于光伏和风电信息变量来调制电负荷和热负荷信息的过程。以电负荷和热负荷信息调制风电和光伏信息的步骤与之类似，数学表达式的形式完全相同，这里不再赘述。

总体而言，相较于的图神经网络或传统的时序预测模型，本文提出的 GCGGN 模型在结构上引入了一种新型的组内(源-源、荷-荷)和组间(源-荷)双层分组交叉门控机制，使得模型能够更充分地利用不同变量之间的隐含关联以进行动态权重调节，增强了多源变量间的信息互补与交叉控制能力，进而提升了多种不同变量联合预测的精度；同时结合图卷积模块，有效捕捉了区域间能源站点的空间相关性，提升了多场站集群预测的精度。该融合设计增强了模型对不同变量、地区、气象条件等的适应能力和泛化性能，在后续消融实验中也验证了该设计在提升预测精度方面的有效性。

2.4 基于 GCGGN 的预测模型

图 5 给出了所提 GCGGN 模型对源荷功率进行预测的流程图。首先，需要获取模型预测所需的前一段时间(5 h)的历史信息，主要包括各光伏场站和风电场站的发电功率、各片区的电负荷和热负荷功率以及影响风光电源出力的多种气象数据(如温度、湿度、风速、辐照度等)。其次，对采集到的历史数据进行预处理，剔除明显异常的数值，并做归一化。而后，将历史数据组织为 4 个具有图数据结构的张量(如第 1.2 节所述)，分别包含各时间和空间的风电、光伏、电负荷和热负荷信息。在此基础上，将这 4 个张量输入给已训练的 GCGGN 模型进行预测，模型输出也是 4 个具有图结构的张量，存放了各场站(或片区)源/荷功率的预测值。

3 算例分析

3.1 算法实现细节

本节使用我国北部某地区范围内的 5 个风电场、6 个光伏电站、3 个片区的电负荷和 3 个片区的热负荷开展实验。选择各场站和片区 2019 年 10 月至 2019 年 12 月共 3 个月的实况功率数据和气象数据进行建模分析，数据分辨率为 15 min。本文数据集信息在附录 B 中介绍，风电、光伏、电负荷、热负荷的实际功率值范围分别为 0~100、0~100、10~25、1~5 MW，部分源荷功率典型曲线如附录 B 图 B1 所示。对于风电场和光伏电站，使用功率历史值及相关气象影响因素的历史值作为网络的输入特征。对于电负荷和热负荷，使用二者的历史值本身作为网络的输入特征。采用滑动时序窗口，以 20 步历史观测序列作为输入，预测下一时刻目标值。数据集分割为训练集(70%)、验证集(10%)与测试集(20%)。实验基于 PyTorch 框架搭建神经网络模型，采用 Adam 优化器训练 200 个 epoch，使用的损失函数为 L1 损失，批大小(batch size)为 32，初始学习率为 0.000 5。本文基于上述数据集和超参数设置对 GCGGN 性能进行评估。

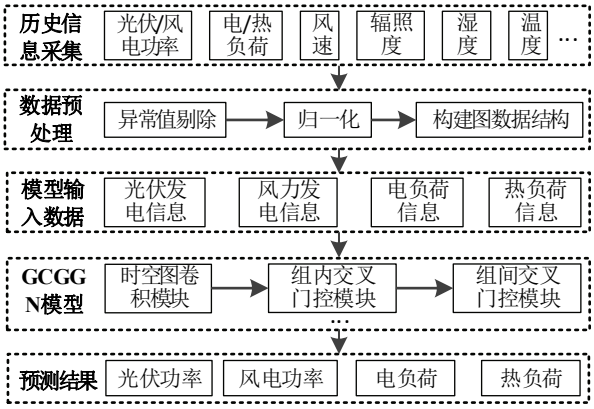


图 5 预测流程图

Fig.5 Flow chart of prediction

3.2 数据预处理

从现场采集得到的功率数据和气象数据通常存在一定量的异常值和缺失值，且存在量纲不统一的问题。因此，在将各类数据输入模型之前，本文先对数据进行以下 3 项预处理操作。

1）剔除显著的异常值。部分气象参数需保持非负特性(如风速、辐照度)，其中夜间时段辐照度应趋近于 0；温度、气压等参数的实测值需处于该地区历史观测的合理区间。针对明显偏离合理范围的异常值，采用相邻时刻均值进行修正。

2）补全缺失数据。由于设备故障等原因，数据集中的个别数据出现了缺失，因此本文用线性插值法补全了此类数据。

3）归一化。特征量纲与数量级的差异会干扰模型训练，通过归一化至[0,1]区间消除量纲差异，可稳定神经网络的训练过程并提升训练效率。本文归一化的方法为最小-最大归一化方法，可表示为：

$$\boldsymbol{X}_{\text{norm}}=\frac{\boldsymbol{X}-\boldsymbol{X}_{\min}}{\boldsymbol{X}_{\max}-\boldsymbol{X}_{\min}} \tag{20}$$

式中：$\boldsymbol{X}$ 和 $\boldsymbol{X}_{\text{norm}}$ 分别是某一特征向量归一化前和归一化后的数值；$\boldsymbol{X}_{\max}$ 和 $\boldsymbol{X}_{\min}$ 分别是该特征向量的最大值和最小值。

3.3　模型评价指标

本文采用的误差评价指标有下述 4 种：归一化的平均绝对误差(normalized mean absolute error, NMAE, δ_{NMAE})、归一化的均方根误差(normalized root mean square error, NRMSE, δ_{NRMSE})、决定系数(R^2)和归一化的平均偏差(normalized mean bias error, NMBE, δ_{NMBE})。NMAE 和 NRMSE 衡量预测误差，数值越小表示精度越高；R^2 反映预测值与真实值的相关性，越接近 1 表明模型的拟合效果越好；NMBE 评估误差的系统性偏差程度，但数值可能受正负抵消影响，需与 NMAE 等指标结合以更全面判断预测值的偏离趋势。

1）归一化的平均绝对误差(NMAE)

$$\delta_{\text{NMAE}}=\frac{1}{N}\sum_{i=1}^{N}\left(\left|y'(i)-y(i)\right|\right) \tag{21}$$

2）归一化的均方根误差(NRMSE)

$$\delta_{\text{NRMSE}}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y'(i)-y(i)\right)^2} \tag{22}$$

3）决定系数(R^2)

$$R^2=1-\frac{\sum_{i=1}^{N}\left(y'(i)-y(i)\right)^2}{\sum_{i=1}^{N}\left(y(i)-\bar{y}\right)^2} \tag{23}$$

4）归一化的平均偏差(NMBE)

$$\delta_{\text{NMBE}}=\frac{1}{N\bar{y}}\sum_{i=1}^{N}\left(y'(i)-y(i)\right) \tag{24}$$

式中：$y(i)$ 和 $y'(i)$ 分别为归一化后的真实值和预测值；N 为样本数；$\bar{y}=\frac{1}{N}\sum_{i=1}^{N}y(i)$ 为真实值的平均值。

3.4　实验结果与分析

表 1—表 4 中展示了本文所设计模型、消融实验以及和其他模型对比实验的结果，表 1—表 4 分别对应风电功率、光伏功率、电负荷、热负荷的预测结果，其最后一行展示了本文设计的 GCGGN 模型的性能，各源/荷变量的 NMAE 误差均在 1%~2%之间，且其 $R^2\approx1$、$\delta_{\text{NMBE}}\approx0$。实验结果表明 GCGGN 预测结果的精确度较高，且没有明显的系统性偏差，能够满足超短期源荷功率预测的精度需求。此外，GCGGN 具备接近实时的推理速度(毫秒级)，在 intel core i7(16G 内存)的计算机上对 4 种变量进行单次功率预测仅耗时 21.7 ms(不同模型的推理时间详见附录 B 表 B1)。该推理时间远低于超短期预测 15 min 的时效要求，充分满足实时性预测需求。

表 1　不同模型预测性能对比(风电)

Table 1　Comparison of prediction performance of different models(WTs)

模型	NMAE/%	NRMSE/%	R^2/%	NMBE/%
MLP	3.54	4.91	97.32	0.67
LSTM	2.61	4.32	97.93	−1.48
GRU	2.76	4.53	97.72	1.87
Transformer	2.54	3.93	98.28	−1.65
BasisFormer	2.55	4.61	97.64	−0.16
iTransformer	2.16	4.15	98.09	0.65
STGCN	2.31	4.19	98.05	0.59
GN	2.28	3.63	98.54	−0.73
CGGN	1.99	3.34	98.76	−0.64
GCGGN	1.85	3.22	98.85	−0.55

表 2　不同模型预测性能对比(光伏)

Table 2　Comparison of prediction performance of different models(PVs)

模型	NMAE/%	NRMSE/%	R^2/%	NMBE/%
MLP	2.76	6.56	94.45	2.72
LSTM	2.37	5.21	96.50	3.87
GRU	1.91	3.13	98.74	0.49
Transformer	2.15	3.09	98.77	0.58
BasisFormer	1.99	3.68	98.25	−3.16
iTransformer	1.80	3.54	98.38	0.52
STGCN	1.88	3.17	98.70	1.29
GN	1.82	4.16	97.77	0.66
CGGN	1.29	3.37	98.53	0.43
GCGGN	1.27	3.20	98.68	0.42

表 3　不同模型预测性能对比(电负荷)

Table 3　Comparison of prediction performance of different models(electrical load)

模型	NMAE/%	NRMSE/%	R^2/%	NMBE/%
MLP	1.85	2.42	87.77	3.17
LSTM	1.72	2.39	88.08	0.60
GRU	1.58	2.24	89.53	0.03
Transformer	1.64	2.23	89.62	3.43
BasisFormer	1.52	2.25	89.45	−0.11
iTransformer	1.30	2.02	91.51	−0.22
STGCN	1.58	2.26	89.34	−0.15
GN	0.94	1.54	95.05	0.29
CGGN	0.91	1.55	94.98	0.28
GCGGN	0.83	1.43	95.72	0.26

GCGGN 模型输出的部分场站预测结果如图 6 所示。图中的蓝色实线表示 GCGGN 的预测结果，红色点表示相应时刻处的真实值，采用归一化结果

表 4 不同模型预测性能对比(热负荷)

Table 4 Comparison of prediction performance of different models(heat load)

模型	NMAE/%	NRMSE/%	R^2/%	NMBE/%
MLP	2.73	3.89	89.38	−1.13
LSTM	2.59	3.30	92.36	−1.08
GRU	2.46	3.06	93.43	−2.95
Transformer	2.41	3.15	93.04	−2.84
BasisFormer	2.28	2.90	94.11	−1.32
iTransformer	2.22	2.83	94.40	−0.56
STGCN	2.35	2.99	93.72	−1.22
GN	2.13	2.80	94.50	−0.55
CGGN	2.08	2.74	94.73	−0.52
GCGGN	2.07	2.72	94.80	−0.51

绘制曲线图。从图 6 可以看出，对于光伏发电功率等局部波动较平缓的变量，GCGGN 的预测结果和真实值非常接近。对于热负荷这类局部高频抖动的变量，尽管在拐点处出现可见的误差，但误差的量级依然较小(2%~3%之间)，在可接受的范围内。总体上，本文的 GCGGN 可以精确预测多种类型源荷功率的变化。

为进一步对比 GCGGN 模型的预测值和真实值，以真实值作为横坐标，以每个真实值对应的模型预测值作为纵坐标，绘制出预测结果的散点图，如图 7 所示(包括了风电、光伏、电负荷、热负荷这 4 种变量的所有功率预测结果)。每类源/荷变量的功率散点图如附录 B 图 B2 所示。若散点的分布越接近 $y=x$ 函数曲线，则说明预测值和真实值越接近。观察图 7 和附录 B 图 B2 可以发现，预测结果基本能拟合出 $y=x$ 函数曲线(即对角线)，并且绝大多数点都分布在离对角线距离较小的范围内，说明预测值和真实值较为接近。综合以上分析可知，所提出的 GCGGN 预测方法在处理具有强非线性的超短期源荷功率预测方面具有优越的性能。

为彰显本文方法在多种场景下的有效性和鲁棒性，图 8 展示了 GCGGN 在 3 种典型光伏气象场景(晴天、阴雨天、多云)以及 3 种典型风电气象场景(强风、微风、风速突变)下的预测结果及误差。各子图左侧坐标轴为功率值，右侧坐标轴为 NMAE。在光伏发电功率预测方面，GCGGN 展现出优秀的非线性波动适应性。晴天场景中，模型的超短期预测误差稳定，仅日出日落时光照骤变导致 NMAE 略有上升，白天时段 NMAE 维持在 2%以内(见图 8(a))；阴雨天光伏出力低，虽在 52~54 时段功率骤降致 NMAE 略有上升，但绝大多数时段误差

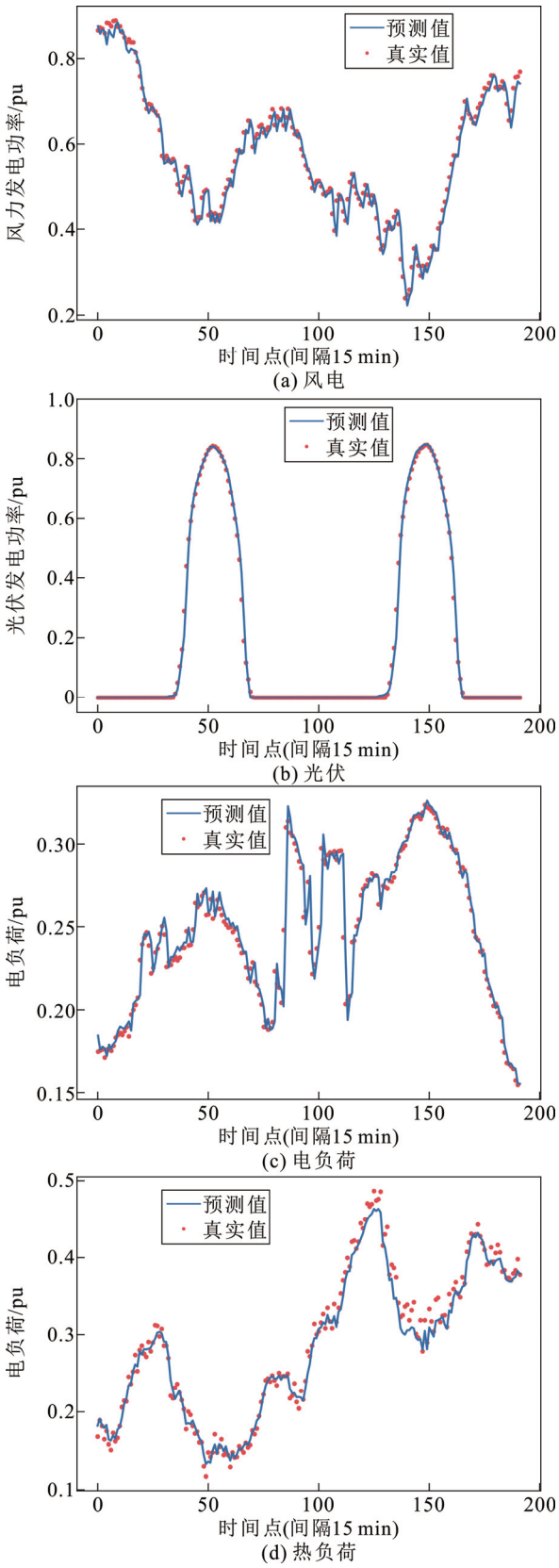


图 6 源荷功率预测结果

Fig.6 Result of source/load power prediction

可控(见图 8(b))；多云天气下，仅在第 46、52 时间点云层遮挡瞬间出现较大误差(见图 8(c))。此外在

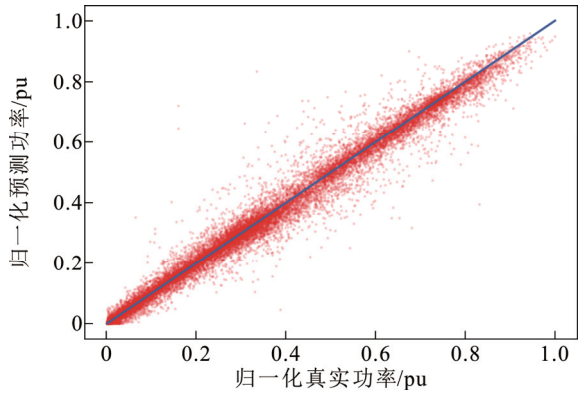


图 7　源荷功率预测结果散点图

Fig.7　Scatter plot of source/load power prediction results

各天气条件下，GCGGN 对夜间光伏发电功率的预测结果均几乎为 0，符合物理规律。以上各场景结果表明，模型能精准捕捉光伏发电的波动规律。在风电功率预测中，虽然风电时间序列随机性较强，但 GCGGN 仍能紧密跟随实际功率趋势。在强风(图 8(d))、微风(图 8(e))、风速突变(图 8(f))气象条件下，GCGGN 预测结果的 NMAE 均基本稳定在 2%以下，仅在风速突变瞬间误差小幅上升，体现出高预测精度与复杂场景下的稳定性。综上，GCGGN 通过时空图卷积模块与交叉门控机制，引入了时空关联信息和不同变量间的关联信息，精准刻画了功率变化趋势，有效降低了光照或风速突变场景下的预测偏差，在各类天气条件下具有较强的鲁棒性。

本文设计的 GCGGN 中包含了组间交叉门控模块和组内交叉门控模块这 2 个可以提升联合预测精度的新模块，其中利用组内交叉门控模块提取每组内不同种类的发电功率或负荷之间的时空关联信息，利用组间交叉模块采集源、荷 2 组变量之间具有的隐含联系。为了评估这 2 个模块对模型性能的贡献，本文进行了如下消融实验。首先，在完整的 GCGGN 中移除组间交叉门控模块，该模型记作 CGGN(cross-gating graph convolutional network)。在 CGGN 的基础上再去除组内交叉门控模块，得到一个完全不含交叉门控机制的图卷积神经网络，记作 GN(graph convolutional network)。上述 2 个简化模型均采用和 GCGGN 相同的超参数进行训练，与 GCGGN 的对比结果记录在表 1—表 4 的后 3 行中。由表 1—表 4 可知，删除组间/组内交叉门控模块均会导致模型预测性能下降，尤其对于风力发电功率/光伏发电功率这 2 个依赖局部天气条件且具有强关联性的被预测量，其预测精度降幅更为显著。这表明本文设计的组内/组间交叉门控模块可显著提升预测准确度，印证了该模块提取的各变量时空耦合信息对模型性能的关键作用。此外，与 LSTM、GRU 等利用传统门控机制实现单变量、单场站时序预测的模型相比，GCGGN 的性能提升显著，证明了图神经网络与文本所提出的分组交叉门控机制相结合的优势。总之，通过多变量组内和组间的交叉调制，同一区域内各预测量潜在的时空特征得到了有效提取，各预测变量之间隐性关联被预测模型充分利用，有效提高了模型预测的准确性。

3.5　不同模型预测性能对比

为验证 GCGGN 的方法先进性，本文将 GCGGN 和目前主流时间序列预测模型进行了对比，在同样环境下实现并训练了 7 种不同时序预测模型：(1)经典序列预测模型：LSTM[25]、GRU[26]；(2)自注意力架构：Transformer[27]及其改进型 BasisFormer[28]、iTransformer[29]；(3)全连接神经网络：移除 GCGGN 所有交叉门控模块，并将时空图卷积替换为多层感知机(multilayer perceptron, MLP)结构；(4)时空图卷积网络：STGCN。因传统时序预测模型缺乏多变量联合建模能力，对比模型采用分变量独立训练策略，针对光伏、风电、电负荷、热负荷 4 类变量分别构建预测模型并进行测试。

各模型的预测数据评价结果见表 1—表 4。结果表明 GCGGN 在所列方法中具有最高的预测精度。综合多种误差指标来看，GCGGN 的预测性能显著优于 LSTM、GRU、MLP 和多种基于 Transformer 的算法。具体而言，与对比模型中性能次优且先进的 iTransformer 算法相比，GCGGN 在发电功率(包括风电、光伏)预测方面的 NMAE 分别降低了 14.4%和 29.4%，NRMSE 分别降低了 22.4%和 9.60%；负荷(包括电负荷、热负荷)预测方面的 NMAE 分别降低了 36.2%和 6.8%，NRMSE 分别降低了 29.2%和 3.9%；R^2 和 NMBE 指标也均有改善，分别更接近 1 和 0。尽管部分数据集上本文模型的 NMBE 绝对值略大于对比方法，但结合 NMAE 指标来看，这是由于其他方法因不同时刻误差符号不一致而出现了抵消效应。热负荷预测中各模型的 NMBE 多为负，主要原因是测试集集中在冬季，取暖需求上升使实际负荷整体呈现上升趋势。可见，本文模型在多个预测对象、多种评价指标上相较对比模型表现出优越性能，验证了本文方法的先进性。

GCGGN 与对比模型的预测误差概率密度如图

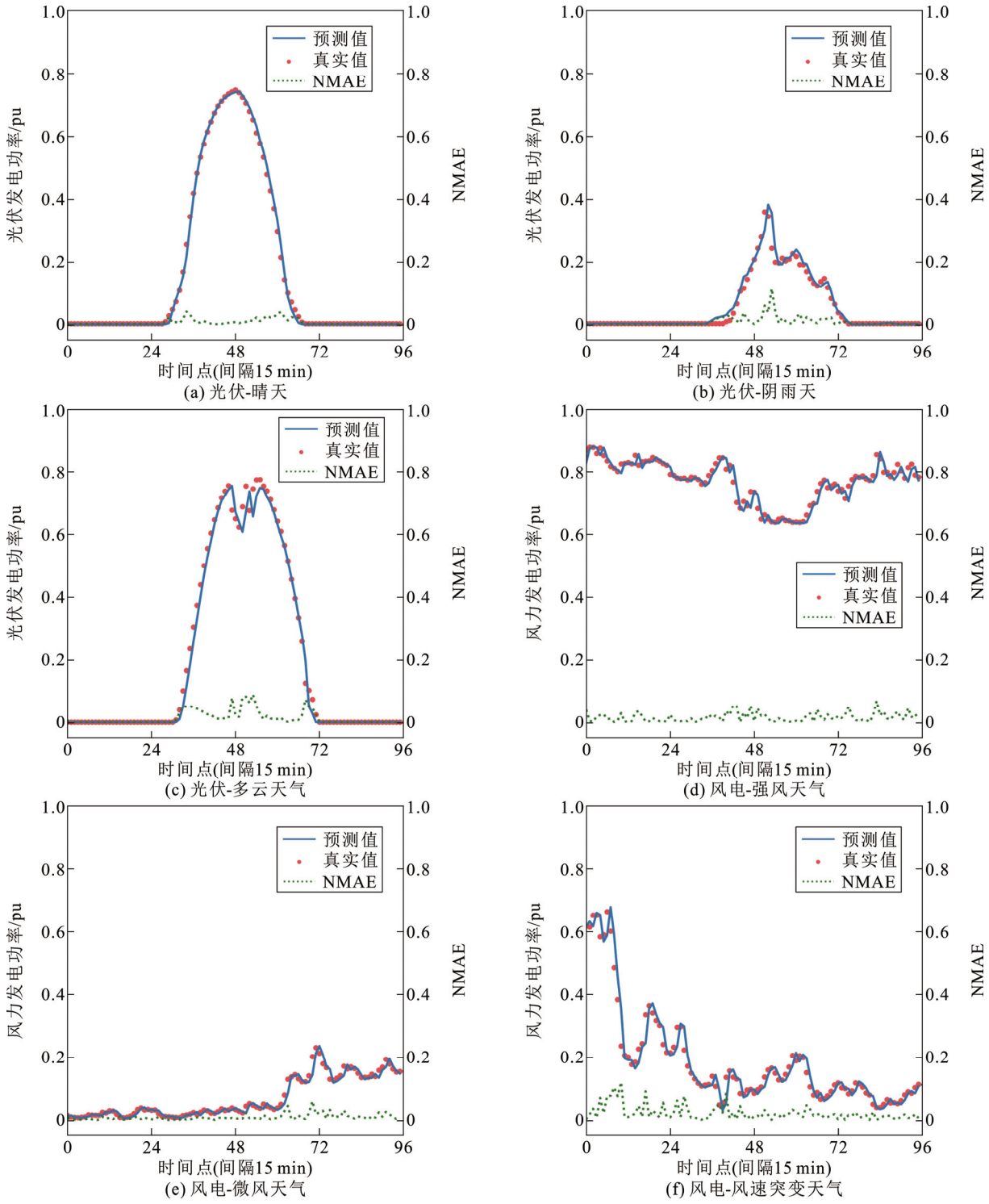


图 8 不同气象条件下典型场景预测结果

Fig.8 Prediction results of typical scenarios under different meteorological conditions

9 所示，描述了预测结果的总体误差分布情况。理论上，当概率密度曲线的尖峰对应的预测误差趋近于 0 值且分布呈现强收敛性时，表示模型具有更优的预测精度。观察图 9 可知，与其他对比方法相比，GCGGN 模型对风电、光伏、电负荷及热负荷的预测误差分布显著向 0 值集中，呈现出峰值更高、峰宽更窄的单峰形态(峰值更尖锐)。相比于本方法，MLP、iTransformer、BasisFormer 等模型的预测误差分布范围更宽，表明有较多的预测值与真实值相差较远，预测的准确性不足；而 LSTM、GRU、Transformer 模型的误差分布峰值点在某些数据集上会明显偏移零点，表明这些模型可能存在系统性

偏差，即整体趋向于高估或低估待预测的值。结果表明，相较于其他模型，本文所设计的 GCGGN 模型的误差分布更集中、偏差更小，从而展现出更高的预测精度和更好的预测稳定性。

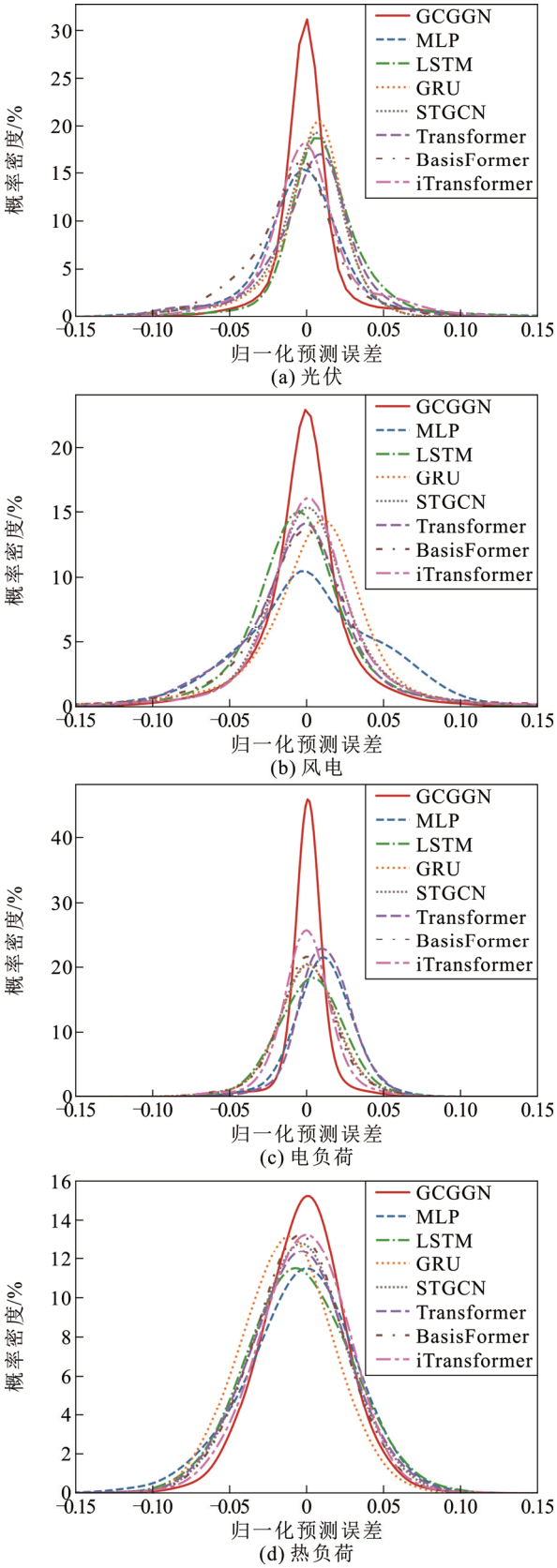


图 9　拟合出的各模型预测误差概率密度函数

Fig.9　Fitted probability density functions of prediction error for each model

3.6　多场站联合预测效果分析

为探究场群预测对单一场站预测精度的提升效应，本节通过调整风电/光伏站点及电/热负荷片区的参与数量，定量分析特定风电场的功率预测误差变化，实验结果整理于表 5。

如表 5 所示，当该区域内联合预测的源/荷规模增加时，单一场站的预测误差呈现下降趋势，从而验证了场群预测对提升预测精度的积极作用。该结果表明，本文设计的时空图卷积网络展现出强大的对空间相关性的提取能力，能够有效捕捉邻近场站/片区因相似气象条件或用户用能行为而产生的潜在关联，进而提升各源/荷功率的预测准确性，并增强它们参与联合预测的动力。这一优势表明 GCGGN 在提升单区域多源荷功率预测精度的同时，其分层图结构建模能力支持跨区域能源系统的时空特征融合，为综合能源系统的分布式源荷协同预测提供了可扩展解决方案。

综上所述，本研究提出的 GCGGN 模型在上述测试场景中均实现最优预测性能，所设计的组内/组间交叉门控模块和时空图卷积模块使得模型具备层次化特征交互机制与动态时空相关性提取能力，支持多站点风电和光伏功率、多区域电负荷和热负荷的联合预测，在多能源时空耦合预测任务中相较于其他方法展现出显著技术优势。

此外，针对源荷多变量联合预测场景，本文分析了各目标变量的数据特征，并以此切入剖析误差来源，具体分析见附录 C。

4　结论

1）针对多场站、多类型源荷功率时空特征耦合复杂、多源关联特征难以表征的问题，本文提出了一种基于改进型图卷积神经网络的超短期源荷功率联合预测模型，增强了风光出力和负荷的时空关联特征提取能力。

表 5　场站数量对同一风电场站预测性能的影响

Table 5　Influence of the number of stations on the prediction performance of a certain WT station

参与预测的场站数量	风电预测误差指标		
[风电，光伏，电负荷，热负荷]	NMAE/%	NRMSE/%	R^2/%
[1, 1, 1, 1]	1.84	3.47	97.81
[3, 3, 2, 2]	1.79	3.42	97.87
[5, 6, 3, 3]	1.70	3.39	97.91

2）本文通过设计组内和组间交叉门控机制，实现了多源信息的交叉调制与特征融合，增强了时序模型的时空特征提取能力，显著提升了多类型源荷功率联合预测精度。

3）基于我国北部某地区实例的实验结果表明，所提模型在预测精度上显著优于现有网络模型，并通过消融实验证实了所设计各模块对预测性能提升的作用。

未来研究将进一步拓展并验证预测模型在动态复杂工况和极端条件下的适应性与泛化能力。同时，基于超短期预测可开展系统运行控制与能量调控决策研究，以实现综合能源系统协同高效运行。

附录见本刊网络版(http://hve.epri.sgcc.com.cn)。

参考文献 References

[1] 任大伟，肖晋宇，侯金鸣，等. 双碳目标下我国新型电力系统的构建与演变研究[J]. 电网技术，2022，46(10)：3831-3839.
REN Dawei, XIAO Jinyu, HOU Jinming, et al. Construction and evolution of China's new power system under dual carbon goal[J]. Power System Technology, 2022, 46(10): 3831-3839.

[2] 黄炜栋，李 杨，李璟延，等. 考虑可再生能源不确定性的风-光-储-蓄多时间尺度联合优化调度[J]. 电力自动化设备，2023，43(4)：91-98.
HUANG Weidong, LI Yang, LI Jingyan, et al. Multi-time scale joint optimal scheduling for wind-photovoltaic-electrochemical energy storage-pumped storage considering renewable energy uncertainty[J]. Electric Power Automation Equipment, 2023, 43(4): 91-98.

[3] 鲍 谚，石锦凯，陈世豪. 考虑负荷预测不确定性的快充站储能鲁棒实时控制策略[J]. 电力系统自动化，2023，47(10)：107-116.
BAO Yan, SHI Jinkai, CHEN Shihao. Robust real-time control strategy for energy storage in fast charging station considering load forecasting uncertainty[J]. Automation of Electric Power Systems, 2023, 47(10): 107-116.

[4] 董 存，王 铮，白捷予，等. 光伏发电功率超短期预测方法综述[J]. 高电压技术，2023，49(7)：2938-2951.
DONG Cun, WANG Zheng, BAI Jieyu, et al. Review of ultra-short-term forecasting methods for photovoltaic power generation[J]. High Voltage Engineering, 2023, 49(7): 2938-2951.

[5] 刘俊磊，刘新苗，卢 洵，等. 高比例新能源系统供需平衡分析方法和对策[J]. 高电压技术，2023，49(7)：2711-2724.
LIU Junlei, LIU Xinmiao, LU Xun, et al. Analysis methods and countermeasures of supply and demand balance of high proportion of new energy system[J]. High Voltage Engineering, 2023, 49(7): 2711-2724.

[6] HAN S, QIAO Y H, YAN J, et al. Mid-to-long term wind and photovoltaic power generation prediction based on copula function and long short term memory network[J]. Applied Energy, 2019, 239: 181-191.

[7] MASSIDDA L, MARROCU M. Total and thermal load forecasting in residential communities through probabilistic methods and causal machine learning[J]. Applied Energy, 2023, 351: 121783.

[8] 赵 洋，王瀚墨，康 丽，等. 基于时间卷积网络的短期电力负荷预测[J]. 电工技术学报，2022，37(5)：1242-1251.
ZHAO Yang, WANG Hanmo, KANG Li, et al. Temporal convolution network-based short-term electrical load forecasting[J]. Transactions of China Electrotechnical Society, 2022, 37(5): 1242-1251.

[9] 张昆明，蔡珊珊，章天晗，等. 考虑多维时域特征的行业中长期负荷预测方法[J]. 电力系统自动化，2023，47(20)：104-114.
ZHANG Kunming, CAI Shanshan, ZHANG Tianhan, et al. Medium-and long-term industry load forecasting method considering multi-dimensional temporal features[J]. Automation of Electric Power Systems, 2023, 47(20): 104-114.

[10] 王 琛，王 颖，郑 涛，等. 基于 ResNet-LSTM 网络和注意力机制的综合能源系统多元负荷预测[J]. 电工技术学报，2022，37(7)：1789-1799.
WANG Chen, WANG Ying, ZHENG Tao, et al. Multi-energy load forecasting in integrated energy system based on ResNet-LSTM network and attention mechanism[J]. Transactions of China Electrotechnical Society, 2022, 37(7): 1789-1799.

[11] HUANG X Q, LI Q, TAI Y H, et al. Time series forecasting for hourly photovoltaic power using conditional generative adversarial network and Bi-LSTM[J]. Energy, 2022, 246: 123403.

[12] 邹 智，吴铁洲，张晓星，等. 基于贝叶斯优化 CNN-BiGRU 混合神经网络的短期负荷预测[J]. 高电压技术，2022，48(10)：3935-3945.
ZOU Zhi, WU Tiezhou, ZHANG Xiaoxing, et al. Short-term load forecast based on bayesian optimized CNN-BiGRU hybrid neural networks[J]. High Voltage Engineering, 2022, 48(10): 3935-3945.

[13] 骆 钊，吴谕侯，朱家祥，等. 基于多尺度时间序列块自编码 Transformer 神经网络模型的风电超短期功率预测[J]. 电网技术，2023，47(9)：3527-3536.
LUO Zhao, WU Yuhou, ZHU Jiaxiang, et al. Wind power forecasting based on multi-scale time series block auto-encoder transformer neural network model[J]. Power System Technology, 2023, 47(9): 3527-3536.

[14] 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.

[15] 申洪涛，李 飞，史 轮，等. 基于气象数据降维与混合深度学习的短期电力负荷预测[J]. 电力建设，2024，45(1)：13-21.
SHEN Hongtao, LI Fei, SHI Lun, et al. Short-term power load forecasting based on reduction of meteorological data dimensionality and hybrid deep learning[J]. Electric Power Construction, 2024, 45(1): 13-21.

[16] 吉兴全，赵国航，叶平峰，等. 基于 QMD-HBi GRU 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.

[17] 董 俊，刘 瑞，束洪春，等. 基于 BIRCH 聚类的 L-Transformer 分布式光伏短期发电功率预测[J]. 高电压技术，2024，50(9)：3883-3893.
DONG Jun, LIU Rui, SHU Hongchun, et al. Short-term distributed photovoltaic power generation prediction based on BIRCH clustering and L-Transformer[J]. High Voltage Engineering, 2024, 50(9): 3883-3893.

[18] ZHENG J Q, DU J, WANG B H, et al. A hybrid framework for forecasting power generation of multiple renewable energy sources[J]. Renewable and Sustainable Energy Reviews, 2023, 172: 113046.

[19] 余 强，韩静娴，杨子梁，等. 基于改进双重压缩和激励与多头特征注意力机制的电-热负荷协同预测[J]. 电力自动化设备，2025，45(3)：201-208.
YU Qiang, HAN Jingxian, YANG Ziliang, et al. Collaborative fore-

casting of electricity-thermal load based on improved dual squeeze and excitation and multi-head feature attention mechanism[J]. Electric Power Automation Equipment, 2025, 45(3): 201-208.

[20] 戴浩男，张辰灏，甄　钊，等. 基于时空特征聚类和双层动态图卷积网络建模的短期净负荷预测[J]. 高电压技术，2024，50(9)：3914-3923.
DAI Haonan, ZHANG Chenhao, ZHEN Zhao, et al. Short-term net load forecasting based on temporal-spatial feature clustering and two-layer dynamic graph convolutional network modeling[J]. High Voltage Engineering, 2024, 50(9): 3914-3923.

[21] 张鹏飞，胡　博，何金松，等. 基于时空图卷积网络的短期空间负荷预测方法[J]. 电力系统自动化，2023，47(13)：78-85.
ZHANG Pengfei, HU Bo, HE Jinsong, et al. Short-term spatial load forecasting method based on spatio-temporal graph convolutional network[J]. Automation of Electric Power Systems, 2023, 47(13): 78-85.

[22] 董　雷，陈振平，韩富佳，等. 基于图卷积神经网络与 K-means 聚类的居民用户集群短期负荷预测[J]. 电网技术，2023，47(10)：4291-4301.
DONG Lei, CHEN Zhenping, HAN Fujia, et al. Short-term load forecasting of residential user groups based on graph convolutional neural network and K-means clustering[J]. Power System Technology, 2023, 47(10): 4291-4301.

[23] SIMEUNOVIĆ J, SCHUBNEL B, ALET P J, et al. Spatio-temporal graph neural networks for multi-site PV power forecasting[J]. IEEE Transactions on Sustainable Energy, 2022, 13(2): 1210-1220.

[24] MA X, PENG B, MA X X, et al. Multi-timescale optimization scheduling of regional integrated energy system based on source-load joint forecasting[J]. Energy, 2023, 283: 129186.

[25] YU Y, SI X S, HU C H, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. Neural Computation, 2019, 31(7): 1235-1270.

[26] CHO K, VAN MERRIËNBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. Doha, Qatar: Association for Computational Linguistics, 2014: 1724-1734.

[27] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: Curran Associates Inc., 2017: 6000-6010.

[28] NI Z L, YU H, LIU S Z, et al. BasisFormer: attention-based time series forecasting with learnable and interpretable basis[C]//Proceedings of the 37th International Conference on Neural Information Processing Systems. New Orleans, USA: Curran Associates Incorporated, 2023: 3118.

[29] LIU Y, HU T G, ZHANG H R, et al. iTransformer: inverted transformers are effective for time series forecasting[C]//Proceedings of the Twelfth International Conference on Learning Representations. Vienna, Austria: ICLR, 2024.



赵婉冰
2000—，女，博士生
研究方向为新能源功率预测与负荷预测、综合能源系统优化运行
E-mail: wb_zhao@zju.edu.cn

ZHAO Wanbing
Ph.D. candidate



杨　强(通信作者)
1979—，男，博士，教授，博导
研究方向为电力人工智能理论及应用
E-mail: qyang@zju.edu.cn

YANG Qiang
Ph.D., Professor
Corresponding author

收稿日期　2025-04-02　在线出版日期　2025-07-18　　编辑　程子丰

附录 A

具体而言，归一化邻接矩阵初始值的构建方法为：各节点的历史特征之间计算皮尔逊相关系数 r_{mn}，并放缩到[0, 1]之间得到 w_{mn}，再用度矩阵做归一化得到 $\boldsymbol{A}_{\text{gcn}}$。节点 m 和节点 n 的时间序列之间的相关系数 r_{mn} 计算公式如下：

$$r_{mn}=\frac{\sum_{t=1}^{T}\left(x_{m,t}-\bar{x}_m\right)\left(x_{n,t}-\bar{x}_n\right)}{\sqrt{\sum_{t=1}^{T}\left(x_{m,t}-\bar{x}_m\right)^2}\sqrt{\sum_{t=1}^{T}\left(x_{n,t}-\bar{x}_n\right)^2}} \tag{A1}$$

式中，$x_{m,t}$ 和 $x_{n,t}$ 分别为节点 m 和 n 在 t 时刻的特征；$\bar{x}_m$ 和 $\bar{x}_n$ 分别为节点 m 和 n 历史特征的平均值；T 为时间长度。线性放缩的公式为：

$$w_{mn}=\frac{1+r_{mn}}{2} \tag{A2}$$

式中，w_{mn} 即为带自环的邻接矩阵 $\tilde{\boldsymbol{A}}$ 中的元素。而后计算 $\tilde{\boldsymbol{A}}$ 的度矩阵 $\tilde{\boldsymbol{D}}$，并对 $\tilde{\boldsymbol{A}}$ 做归一化，作为 $\boldsymbol{W}_{\text{S},i}$ 的初始值。$\boldsymbol{W}_{\text{S},i}$ 的参数后续在网络训练阶段通过学习更新调整，保持为对称矩阵且各元素限幅在[0, 1]之间，以更好地反映各站点或地区之间的关联关系。

附录 B

本文数据集的部分源荷功率典型曲线如图 B1 所示。对于风电场，可用的特征包括 10 m、30 m、50 m 和轮毂高度处风速、气温和发电功率。对于光伏电站，可用的特征包括总辐照度、法向辐照度、水平辐照度、气温、气压和发电功率。

表 B1　不同模型推理时间对比

Table B1　Comparison of inference time cost of different models

模型	测试时间/s	模型	测试时间/s
MLP	0.001 2	BasisFormer	0.074 8
LSTM	0.028 8	iTransformer	0.045 2
GRU	0.010 4	STGCN	0.074 8
Transformer	0.055 2	GCGGN	0.021 7

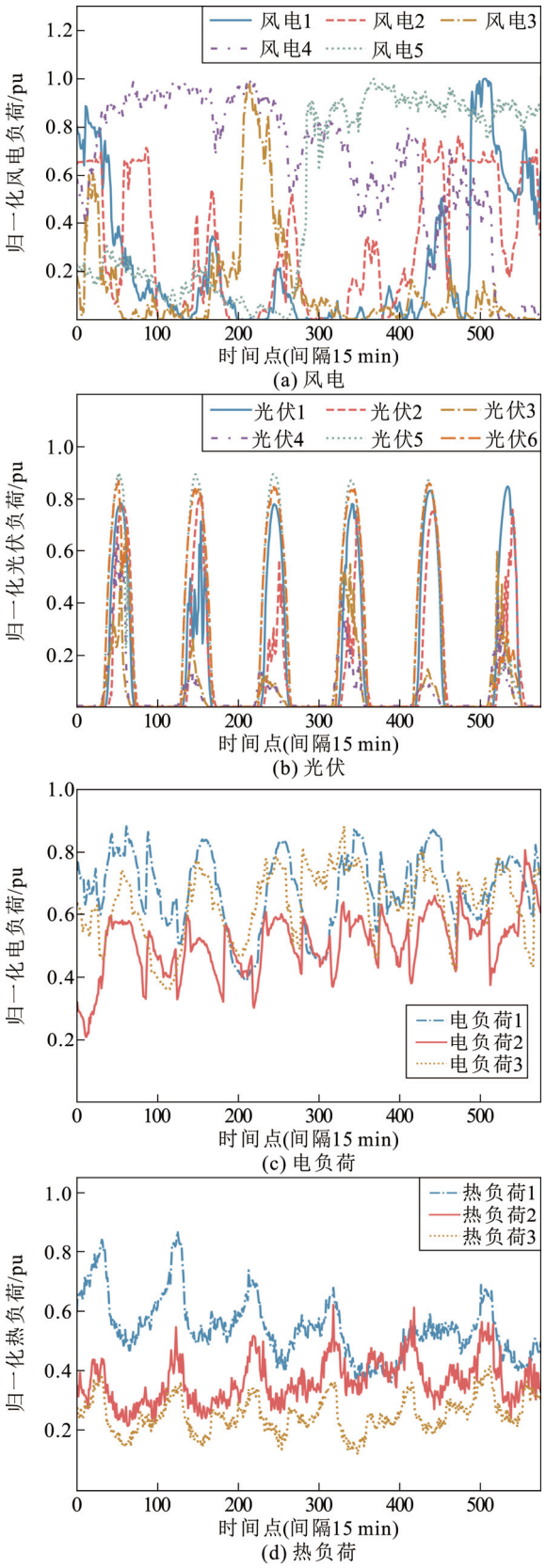


图 B1　数据集中部分源荷功率典型曲线

Fig.B1　Typical curves for the source/load powers in the dataset

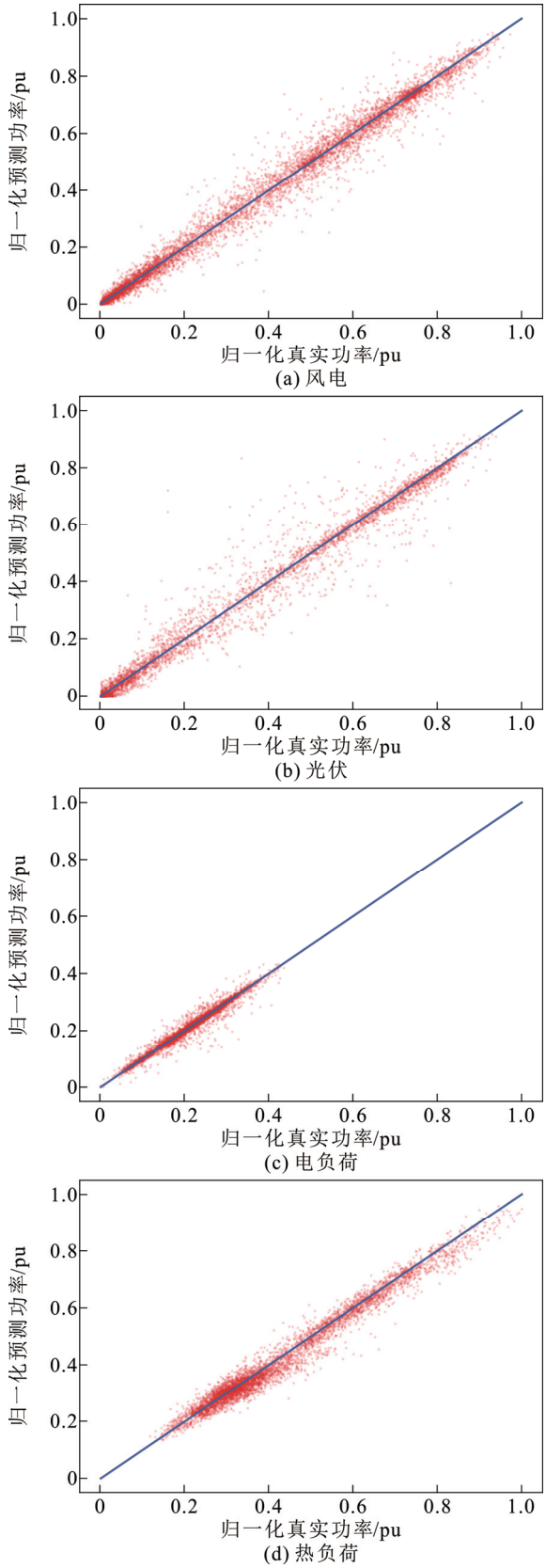


图 B2　各变量预测结果散点图

Fig.B2　Scatter plots of prediction results for each variable

附录 C

本文在相同模型架构下，针对风电、光伏、电负荷与热负荷功率的联合预测实验表明，热负荷预测误差 NMAE 为 2.07%，略高于其他预测对象，但其 RMSE 表现良好。为剖析热负荷预测 NMAE 较高的原因，本文先从数据特征入手，通过计算时间序列二阶差分绝对值的均值，量化序列的局部波动程度。二阶差分的计算公式为：

$$\Delta^2 x_t = x_t - 2x_{t-1} + x_{t-2} \tag{C1}$$

其绝对值可以很好地反映时间序列的局部波动性，即抖动、拐点出现的情况。当该指标较高时，表明序列局部波动性与非线性较强，曲线中存在较多拐点，曲线平滑性较差，模型拟合难度较大。

表 C1 显示，在本文所采用数据集中，对于光伏、风电、电负荷、热负荷的归一化功率时间序列，其二阶差分绝对值均值分别为 0.012 3、0.022 6、0.011 4 和 0.031 5，热负荷的该指标显著高于其他预测对象。这与图 B1 中热负荷功率曲线拐点和抖动较多的观察结果相符。热负荷曲线存在大量高频微小波动，功率变化率的数值与方向快速变化，导致对误差累积敏感的 NMAE 升高；但因变化率本身的幅度有限，因而未出现严重偏离真实值的预测偏差，NRMSE 仍处合理水平。

表 C1　各源荷功率数据的二阶差分绝对值均值

Table C1　Mean absolute value of second-order differences of power data for each source and load

统计指标	光伏	风电	电负荷	热负荷
$\Delta^2 x$ 绝对值均值	0.012 3	0.022 6	0.011 4	0.031 5

为了进一步探究数据特征对预测精度的影响，本文设计了对比实验，采用窗口长度为 5 和 3 的均值滤波来平滑热负荷序列，用滤波后的数据再次训练和测试模型，观察预测精度的变化。滤波前后曲线如图 C1 所示，实验结果如表 C2 所示。

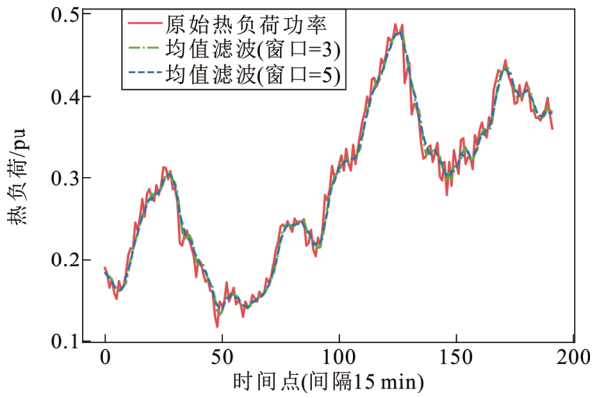


图 C1　热负荷原始数据与滤波后数据曲线

Fig.C1　Raw and filtered data curves for heat load

表 C2　滤波前后热负荷预测性能对比

Table C2　Comparison of heat load prediction performance before and after filtering

数据类型	热负荷			
	$\Delta^2 x$ 绝对值均值	NMAE/%	NRMSE/%	R^2/%
原始数据	0.031 5	2.07	2.72	94.80
滤波(窗口=3)	0.010 0	0.83	1.09	98.54
滤波(窗口=5)	0.005 5	0.52	0.68	99.67

由图 C1 和表 C2 可以看出，经不同程度的滤波处理后，热负荷时间序列的二阶差分绝对值均值显著下降，下降幅度分别为 68.3%和 82.5%，处理后的值接近电负荷的波动水平。预测结果的误差指标也相应降低，NMAE 分别降至 0.83%和 0.52%，NRMSE 降至 1.09%和 0.68%。这些统计特征和误差指标的变化验证了数据本身的局部波动性对预测精度的关键影响。

针对热负荷局部波动性的成因，已有研究指出冬季热负荷波动性以及不确定性较强[19]，该季度热负荷来源多元，供暖、生活热水等需求分散，不同建筑在热需求的时间分布与强度上存在显著差异。尽管热惯性在一定程度上抑制了负荷的大幅度波动(表现为 NRMSE 指标较低)，但多元需求叠加产生的高频微小变化会通过误差累积效应推高 NMAE 指标，凸显数据的局部波动、非线性等复杂特征对功率预测误差的重要影响。