

基于动态空间关联聚合的多区域负荷预测

刘璟璐, 曹文涛*, 董生爱, 张鹏飞, 张子昌

(沈阳工业大学 电气工程学院, 辽宁 沈阳 110870)

Multi-Regional Load Forecasting Based on Dynamic Spatial Correlation Aggregation

LIU Jinglu, CAO Wentao*, DONG Shengai, ZHANG Pengfei, ZHANG Zichang

(School of Electrical Engineering, Shenyang University of Technology, 110870, China)

Abstract: [**Objective**] Accurate multi-regional load forecasting (MRLF) is a critical foundation for ensuring grid stability and optimizing the economic dispatch of modern power systems. Traditional load forecasting methods face challenges in capturing the dynamic spatial dependencies among multi-regional loads. To address the problem that traditional methods are unable to extract dynamic spatial features among loads, a MRLF model based on decoupling and adversarial graph attention network (DAGAT) is proposed. [**Methods**] The model effectively extracted the dynamic spatial features between multi-regional loads through graph attention networks. Firstly, the regional load series was decomposed into trend and fluctuation components using discrete wavelet transform (DWT). Secondly, a two-channel feature extraction module was designed to extract the spatio-temporal features of the multi-regional load series based on the different characteristics of trend and fluctuation components. In addition, a generative adversarial network architecture based on zero-sum game was introduced for adversarial training of spatio-temporal prediction models, the resulting adversarial loss was weighted and combined with a traditional forecasting loss function as the loss function of the model. Finally, simulation validation was performed based on real multi-regional load data from the New York independent system operator. [**Results**] Comparison and analysis with mainstream models showed that the DAGAT model proposed in this paper was more superior than traditional machine learning models, prediction models that only capture temporal dependencies, and prediction models that express spatial correlations using static weight matrices. The superiority of the model stemmed from the fact that the graph attention network captured the spatial correlation between multi-regional loads by learning dynamically changing weights, and the DWT clearly represented the characteristics of the multi-regional load series, and the joint loss further optimized

the model parameters. [**Conclusion**] DAGAT effectively improves the accuracy of MRLF and solves the problem that dynamic spatial features between multi-regional loads cannot be captured, which is effective and superior in engineering practice.

Key words: decoupling; generative adversarial network; graph attention network; multi-regional load forecasting

摘 要: [**目的**] 准确的多区域负荷预测 (MRLF) 是确保电网稳定和优化现代电力系统经济调度的关键基础。传统负荷预测方法在捕获多区域负荷之间的动态空间依赖性时面临着挑战。为了解决传统方法无法提取负荷间动态空间特征的问题,本文提出了一种基于解耦和对抗图注意力网络 (DAGAT) 的 MRLF 模型。[**方法**] 该模型有效地通过图注意力网络提取了多区域负荷之间的动态空间特征。首先,采用离散小波变换 (DWT) 将区域负荷序列分解成趋势分量和波动分量。其次,根据趋势分量和波动分量的不同特征设计了一个双通道特征提取模块,用于提取多区域负荷序列的时空特征。此外,引入了一种基于零和博弈的生成对抗网络架构,用于时空预测模型的对抗训练,并将得到的对抗损失与传统的预测损失函数加权结合作为模型的损失函数。最后,基于纽约独立系统运营商的真实多区域负荷数据进行仿真验证。[**结果**] 与主流模型比较分析表明,相较于传统机器学习模型、仅能捕获时间依赖关系的预测模型以及采用静态权重矩阵表达空间相关性的预测模型,本文所提 DAGAT 模型更具优越性。该模型的优越性源于图注意力网络通过学习动态变化的权重捕获了多区域负荷之间的空间相关性,DWT 清晰表示了多区域负荷序列的特征,联合损失进一步优化了模型参数。[**结论**] DAGAT 有效提高了 MRLF 精度,解决了多区域负荷间动态空间特征无法捕获的问题,在工程实践中具有有效性和优越性。

关键词: 解耦;生成对抗网络;图注意力网络;多区域负荷预测

0 引言

新型电力系统中可再生能源占比不断上升,使得源荷时空分布不均衡特征日益显著^[1]。协同运行的多区域电力系统,能够有效促进新能源的跨区域消纳,实现多区域新能源资源的均匀分配^[2]。通过准确预测多区域负荷,能够计算各区域负荷可调容量,从而实现可调容量的动态优化配置^[3]。因此,仅靠单区域负荷预测无法满足跨区域协调调度需求,多区域经济调度亟需准确的多区域负荷预测(Multi-Regional Load Forecasting, MRLF)提供数据支撑,从而最大限度提高多区域电力系统的经济性^[4]。

深度学习作为一种人工智能方法,不仅可以多层神经网络学习输入和输出之间的复杂映射关系,还可以有效解决浅层模型中的梯度消失问题^[5]。目前,许多学者已经将深度学习模型应用于短期负荷预测^[6-9]。然而,与传统负荷预测不同的是,MRLF除了需要关注负荷的时间依赖性外,还需要关注多区域负荷间的空间依赖性^[10]。文献[11]通过分析不同区域负荷的相关性确定了多区域负荷间的空间关联关系,并利用门控多头网络捕获区域负荷间的空间相关性。文献[12]针对元胞空间关系构建和区域负荷空间关系捕获困难的问题,提出了一种基于时空图注意力的超短期区域负荷预测方法,并通过真实负荷数据验证了所提方法的有效性。文献[13]提出了一种用于区域智能能源系统的两阶段特征提取和混合倒置 Transformer 的短期区域负荷预测模型,提高了预测精度,确保了区域智能能源系统的稳定运行。文献[14]针对不同地理区域电力用户之间存在的空间关联性,提出了一种自适应地理图神经网络用于短期区域负荷预测。文献[15]充分考虑 I 类元胞之间的空间影响性,提出了一种基于空间影响性的时空图卷积网络短期空间负荷预测模型。文献[16]聚焦于多区域电力负荷的时空相关性,提出了一种基于图卷积网络(Graph Convolutional Network, GCN)的 MRLF 模型。

虽然现有 MRLF 方法对多区域负荷之间存在的空间依赖关系有所发掘,但是忽略了空间依赖关系的动态变化,严重影响了预测精度^[17]。因此,有效捕获区域负荷间的动态空间依赖性对提高

MRLF 精度至关重要。同时,多区域负荷数据作为一种时间序列数据,可以被分解为各种成分,包括趋势、季节性和残差等,每种成分都表现出不同的特征^[18-19]。然而,现有大多数方法并没有细化多区域负荷的特征。因此,分解多区域负荷序列来细化特征,对提高 MRLF 精度至关重要。此外,现有的 MRLF 模型仅以 L1 损失或 L2 损失作为模型的损失函数,而过度简化的损失函数可能不足以改善一些模型参数,从而影响模型预测性能^[20-21]。

因此,为解决现有 MRLF 模型忽视的问题,本文提出一种基于解耦和对抗图注意力网络(Decoupling and Adversarial Graph Attention network, DAGAT)的混合模型,其整体结构如图 1 所示。将多区域欧式负荷数据重构为非欧式结构数据,经离散小波变换(Discrete Wavelet Transform, DWT)解耦模块细化特征,设计时空网络捕获动态空间依赖性,并引入对抗性损失优化模型,显著提升了负荷预测精度。

1 问题描述

准确的区域负荷预测可以平衡发电与需求的关系,提高电力调度的准确性和效率。MRLF 问题旨在通过分析历史区域负荷数据与地理位置、气象条件及经济等相关影响因素之间的潜在联系,从而预测未来一段时间的区域负荷。定义第 i 个负荷节点在 t 时刻的区域负荷为 $x_t^i \in \mathbf{R}$;所有区域负荷节点在 t 时刻的区域负荷为 $\mathbf{X}_t = [x_t^1, \dots, x_t^i, \dots, x_t^N]^T \in \mathbf{R}^N$;所有区域负荷节点在 T_1 时间片的区域负荷为 $\mathbf{X}_{:T_1} = \{\mathbf{X}_1, \dots, \mathbf{X}_{T_1}\} \in \mathbf{R}^{T_1 \times N}$ 。本文的目的是基于过去 T_1 步的历史区域负荷数据,学习一个映射函数 F_θ ,通过 F_θ 映射出未来 T_2 步的区域负荷。MRLF 问题可以表示为

$$\{\mathbf{X}_{t+1}, \mathbf{X}_{t+2}, \dots, \mathbf{X}_{t+T_2}\} = F_\theta\{\mathbf{X}_{t-T_1+1}, \mathbf{X}_{t-T_1+2}, \dots, \mathbf{X}_t\} \quad (1)$$

式中: θ 为可学习参数。

然而,随着某一区域负荷的持续增加,其周边区域的负荷也会受到影响,从而产生负荷之间相互影响的空间关联特性。因此,需要构建区域负荷关系图 $G=(V,E,A)$ 来描述负荷之间的空间依赖性。其中, V 为区域负荷节点的集合; E 为两个区域负荷节点之间联系的集合; A 为图邻接矩阵,

用来描述任意两个区域负荷节点之间的空间依赖性。因此, MRLF 问题将被重新描述为

$$\{X_{t+1}, X_{t+2}, \dots, X_{t+T_2}\} = F_{\theta} \{ \{X_{t-T_1+1}, X_{t-T_1+2}, \dots, X_t\}, G \} \quad (2)$$

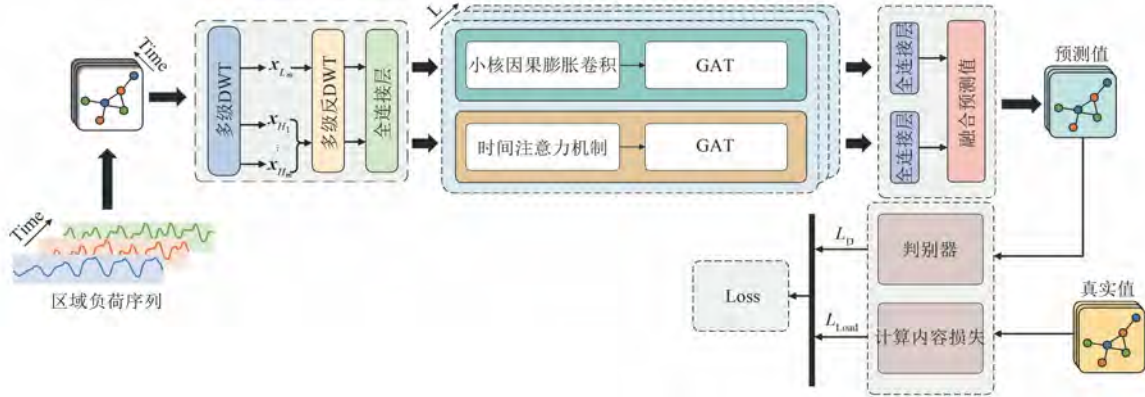


图1 DAGAT 模型整体结构

Fig. 1 The overall structure of DAGAT model

2 DAGAT 模型

2.1 DWT 解耦

时间序列数据通常由多个成分组成^[22]。为了清晰地表示负荷数据以捕捉负荷序列的特征, 将区域负荷序列分解成趋势分量和波动分量, 且趋势分量和波动分量之间互不影响。

在 MRLF 框架下应用 DWT 对负荷序列进行分解。选择 DWT 的原因是负荷序列与风电序列相似, 两者都是离散序列^[23]。DWT 是专门为处理离散序列而设计的。DWT^[21] 过程如图 2 所示。

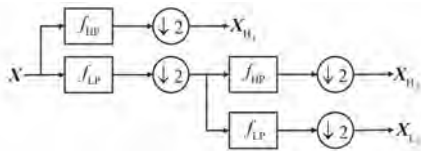


图2 DWT 过程

Fig. 2 DWT process

图 2 中, 第一级 DWT 通过一个低通滤波器 f_{LP} 和一个高通滤波器 f_{HP} 将输入序列 $X_{:T_1} \in \mathbf{R}^{T_1 \times N}$ 分解为一个低频分量 $X_{L_1} \in \mathbf{R}^{\frac{T_1}{2} \times N}$ 和一个高频分量 $X_{H_1} \in \mathbf{R}^{\frac{T_1}{2} \times N}$ 。同理, 第二级 DWT 将 $X_{L_1} \in \mathbf{R}^{\frac{T_1}{2} \times N}$ 分解为一个子低频分量 $X_{L_2} \in \mathbf{R}^{\frac{T_1}{4} \times N}$ 和一个子高频分量 $X_{H_2} \in \mathbf{R}^{\frac{T_1}{4} \times N}$ 。DWT 过程可以表示为

$$\begin{cases} X_{H_1} = (f_{HP} \otimes X_{:T_1})_{(\downarrow)} \\ X_{L_2} = [f_{LP} \otimes (f_{LP} \otimes X_{:T_1})_{(\downarrow)}]_{(\downarrow)} \\ X_{H_2} = [f_{HP} \otimes (f_{LP} \otimes X_{:T_1})_{(\downarrow)}]_{(\downarrow)} \end{cases} \quad (3)$$

式中: $\otimes$ 为卷积运算; $\downarrow$ 为降低采样频率。

DWT 会导致低频分量和高频分量时间序列长度变短。为了保证后续不同频率的分量返回到时域时长度一致, 需要对 X_{L_2} , X_{H_2} 和 X_{H_1} 进行上采样操作。利用反低通滤波器 f_{LP}^T 和反高通滤波器 f_{HP}^T 来实现反 DWT。再通过全连接层得到趋势分量 $X_{trend} \in \mathbf{R}^{T_1 \times N \times d}$ 和波动分量 $X_{fluc} \in \mathbf{R}^{T_1 \times N \times d}$ 。反 DWT 过程可以表示为

$$\begin{cases} X_{trend} = W^{LP} f_{LP}^T \otimes [f_{LP}^T \otimes (X_{L_2})_{\uparrow}]_{\uparrow} + b^{LP} \\ X_{fluc} = W^{HP} \{ f_{LP}^T \otimes [f_{HP}^T \otimes (X_{H_2})_{\uparrow}]_{\uparrow} + \\ f_{HP}^T \otimes (X_{H_1})_{\uparrow} \} + b^{HP} \end{cases} \quad (4)$$

式中: $\uparrow$ 为提高采样频率; W^{LP} , W^{HP} , b^{LP} 和 b^{HP} 为模型参数矩阵, W^{LP} , W^{HP} , b^{LP} , $b^{HP} \in \mathbf{R}^d$ 。

2.2 时间注意力机制

通过解耦得到的趋势分量和波动分量具有完全相反的时间特征。趋势分量的变化持续时间长、幅度小; 波动分量的变化通常是突然的, 持续时间短、幅度大。区域负荷序列的不同成分如图 3 所示。

趋势分量的任何两个时间片之间都存在强相关性; 而波动分量的相邻时间片存在强相关性, 时间间隔较远的时间片的相关性相对较弱。基于趋

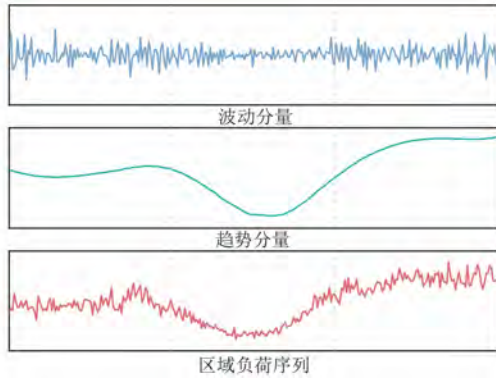


图 3 区域负荷序列的不同成分

Fig. 3 Different components of regional load series

势分量的特征,选择时间注意力机制来捕获趋势分量的时间特征。

时间注意力机制是一种动态的时间选择机制,结构如图 4 所示。时间注意力机制可以为输入序列的某个时间步长分配更大的权重,使模型能够专注于更关键的信息,以提高模型的性能和效率^[25]。其不仅可以实现所有历史信息交互,还可以选择更重要的信息进行关注。获取时间步 t 上的趋势分量的过程可以表示为

$$\begin{cases} x_{itrend}^{natt} = \sum_{i=1}^t \alpha_{i,i}^n (W_{VT} x_{itrend}^n) \\ \alpha_{i,i}^n = \frac{\exp[(W_{KT} x_{itrend}^n)^T (W_{QT} x_{itrend}^n)]}{\sum_{j=1}^t \exp[(W_{KT} x_{itrend}^n)^T (W_{QT} x_{jtrend}^n)]} \end{cases} \quad (5)$$

式中: x_{itrend}^n 为负荷节点 n 在时间步 t 上的趋势分量; x_{itrend}^{natt} 为 x_{itrend}^n 的加权聚合结果; $W_{VT} \in \mathbf{R}^d$ 为线性变换参数; $x_{itrend}^n, x_{jtrend}^n$ 分别为负荷节点 n 在 i 时刻 j 时刻的趋势分量; W_{QT}, W_{KT} 为线性变换矩阵, $W_{QT}, W_{KT} \in \mathbf{R}^{d \times d}$; $\alpha_{i,i}^n$ 为负荷节点 n 在时间步 t 的时间注意力权重。

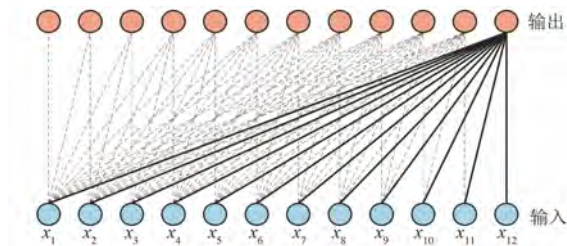


图 4 时间注意力机制

Fig. 4 Time attention mechanism

2.3 小核因果膨胀卷积

由于波动分量的变化持续时间短、幅度大。

因此,采用因果卷积来捕获波动分量的时间特征。因果卷积是一种专门用于处理时间序列数据的卷积神经网络,其当前时间步的输出仅与当前时间步和一些先前时间步的输入相关。针对高频分量的特点,本文采用较小的卷积核进行卷积。此外,为了保持因果卷积具有一定的感受野,将膨胀卷积的思想融入因果卷积,形成小核因果膨胀卷积,其结构如图 5 所示。

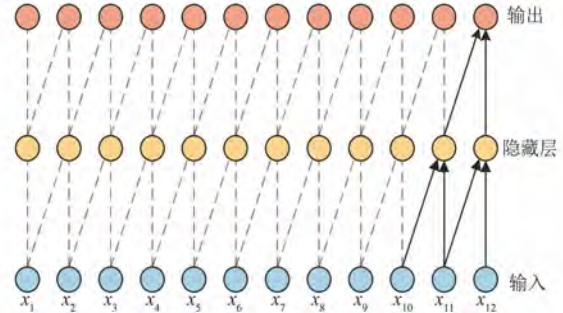


图 5 小核因果膨胀卷积

Fig. 5 Small kernel causal dilation convolution

小核因果膨胀卷积过程可以表示为

$$X \otimes f_c(t) = \sum_{\tau=0}^K f_c(\tau) X(t - \tau) \quad (6)$$

式中: $X \in \mathbf{R}^{T_1}$ 为时间序列; $f_c \in \mathbf{R}^K$ 为滤波器, K 为因果卷积核大小; τ 为第几个卷积核。

经过时间注意力机制和小核因果膨胀卷积捕获时间依赖关系后,得到趋势分量的时间注意力表示 $X_{itrend}^{att} \in \mathbf{R}^{T_1 \times N \times d}$ 和波动分量的因果卷积表示 $X_{fluc}^{conv} \in \mathbf{R}^{T_1 \times N \times d}$ 。

2.4 图注意力网络 (Graph Attention network, GAT)

实际上,随着某一区域负荷的持续增加,其不仅对周边区域的负荷造成影响,也会受到周边区域负荷对其的影响^[15],从而呈现出空间依赖性。因此,本文使用基于密度的噪声应用空间聚类 (Density-Based Spatial Clustering of Applications with Noise, DBSCAN) 算法对多区域电力负荷数据进行聚类操作,以呈现出负荷间的空间依赖性。DBSCAN 算法的聚类结果如图 6 所示,可见区域负荷被聚类为 3 种类型。

图 7 展示了第一类和第二类区域负荷全年的负荷曲线。由图 7 可知,同类区域负荷具有相同的变化趋势。

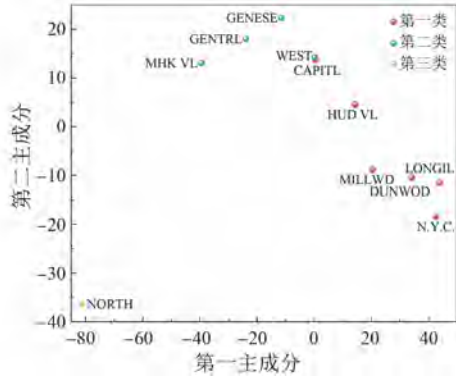


图6 DBSCAN 算法的聚类结果

Fig. 6 Clustering results of the DBSCAN algorithm

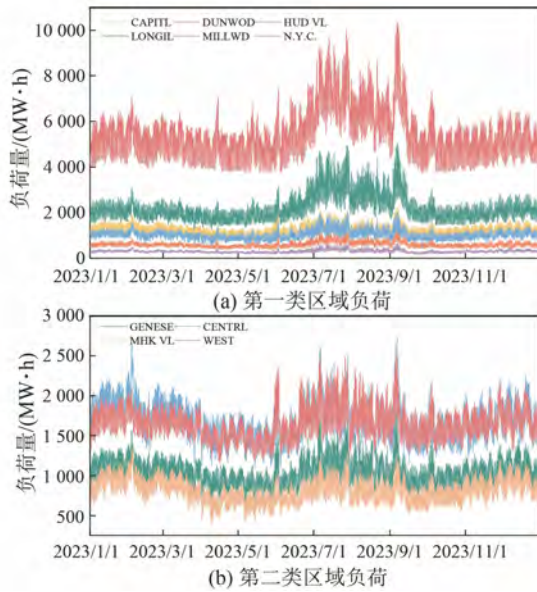


图7 各类区域负荷曲线对比

Fig. 7 Comparison of various load curves

区域负荷分布图如图8所示。结合聚类结果和区域负荷分布图,可以直观地观察到同一类的区域负荷在地理位置上是相邻的。因此,区域负荷之间的空间依赖性也应被视为区域负荷序列中固有的特征。但是,如果需要考虑区域负荷之间的空间依赖性,区域负荷数据的结构将发生变化,从以往整齐排列的欧式数据转化为非欧式数据^[26]。传统网络只能处理欧式数据,提取区域负荷序列内的时间依赖关系。图神经网络正是为处理非欧式数据而设计的,可以捕获区域负荷间隐藏的空间依赖关系。动态空间依赖关系示意图如图9所示。

GAT的注意力机制能够学习到不同时间步变化的邻居节点权重,实现对邻居节点信息的加



图8 区域负荷分布图

Fig. 8 Regional load distribution map

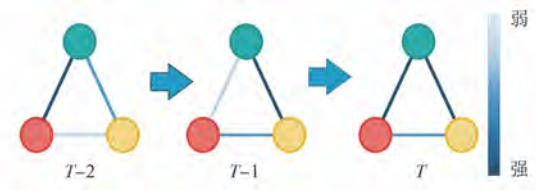


图9 动态空间依赖关系

Fig. 9 Dynamic spatial dependency relationship

权聚合以及动态捕获节点间的空间依赖性;而GCN由于使用的是相同的平均权重,所以在聚合邻居节点信息时是静态的。因此,本文采用GAT来捕获区域负荷间的动态空间依赖性。

由于区域负荷数据为非欧式结构。因此需要通过图邻接矩阵描述节点间的连接关系。本文采用动态时间规整(Dynamic Time Warping, DTW)算法确定区域负荷间的连接关系,建立图邻接矩阵。图邻接矩阵 A 的元素 A_{ij} 为

$$A_{ij} = \begin{cases} 1, d(i, j) \geq k \\ 0, d(i, j) < k \end{cases} \quad (7)$$

式中: k 为阈值; $d(i, j)$ 为负荷节点 i 和 j 中心之间的欧氏距离。

在时间步 t 上,负荷节点 n 和 i 间的空间注意力权重 $\beta_{t,i}^n$ 为

$$\beta_{t,i}^n = \frac{\exp[(\mathbf{W}_s^k \mathbf{x}_t^n)^T (\mathbf{W}_s^Q \mathbf{x}_t^i)]}{\sum_{j=1}^N \exp[(\mathbf{W}_s^k \mathbf{x}_t^n)^T (\mathbf{W}_s^Q \mathbf{x}_t^j)]} \quad (8)$$

式中: $\mathbf{x}_t^n$ 、 $\mathbf{x}_t^i$ 和 $\mathbf{x}_t^j$ 分别为节点 n 、 i 和 j 在时间步 t 上的区域负荷; $\mathbf{W}_s^Q$ 、 $\mathbf{W}_s^k$ 为线性变换矩阵, $\mathbf{W}_s^Q, \mathbf{W}_s^k \in \mathbf{R}^{d \times d}$ 。

负荷节点 n 在时间步 t 上的节点信息聚合结果 $\mathbf{x}_t^{ngat}$ 可以表示为

$$x_t^{ngat} = \sum_{i=1}^N \beta_t^{n,i} (W_S^V x_t^n) \quad (9)$$

式中: $W_S^V \in \mathbf{R}^d$ 为线性变换参数。

通过叠加精心设计的双通道时空模块, 可以获得趋势分量和波动分量的不同时空特征。然后, 通过融合模块可推断出未来的区域负荷。

2.5 融合和对抗训练模块

由两个全连接层组成的预测器将趋势分量、波动分量的时空特征表示 $\mathbf{X}_{trend}^{gat}$ 、 $\mathbf{X}_{fluc}^{gat}$ 转换为趋势分量、波动分量的预测表示 $\hat{\mathbf{Y}}_{trend}^f$ 、 $\hat{\mathbf{Y}}_{fluc}^f$ 。为了得到未来区域负荷的预测表示 $\hat{\mathbf{Y}}^f \in \mathbf{R}^{T_2 \times N \times d}$, 需要将 $\hat{\mathbf{Y}}_{trend}^f$ 和 $\hat{\mathbf{Y}}_{fluc}^f$ 进行融合。区域负荷瞬息万变从而导致区域负荷的波动分量会存在分布偏移的问题。为了解决分布偏移, 需要对时间步 t 内的波动分量进行加权求和, 从而得到更精准的未来区域负荷。整个过程可以表示为

$$\begin{cases} \hat{y}_t^{nf} = \hat{y}_{trend}^f + \sum_{i=T_1+1}^{T_1+t} \eta_{t,i}^n (W_F^V \hat{y}_{fluc}^{nf}) \\ \eta_{t,i}^n = \frac{\exp[(W_F^K \hat{y}_{trend}^{nf})^T (W_F^K \hat{y}_{fluc}^{nf})]}{\sum_{j=T_1+1}^{T_1+t} \exp[(W_F^K \hat{y}_{trend}^{nf})^T (W_F^K \hat{y}_{jfluc}^{nf})]} \end{cases} \quad (10)$$

式中: $\hat{y}_{fluc}^{nf}$ 、 $\hat{y}_{jfluc}^{nf}$ 分别为负荷节点 n 在 i 时刻 j 时刻的波动预测值; $\hat{y}_{trend}^{nf}$ 为负荷节点 n 在时间步 t 上的趋势预测值; $\eta_{t,i}^n$ 为注意力权重; W_F^K 、 W_F^K 为线性变换矩阵; W_F^V 为线性变换参数。

在得到 $\hat{\mathbf{Y}}^f$ 后, 通过一个全连接层便可得到未来区域负荷 $\hat{\mathbf{Y}} \in \mathbf{R}^{T_2 \times N}$, 然后引入生成对抗网络进行对抗性训练, 进一步优化模型参数。

在模型训练过程中, 通过最大化判别器的损失, 能够有效提高预测结果的真实性^[27]。首先, 判别器对模型推断出来的未来区域负荷进行判别, 并返回一个对抗损失; 然后, 将对抗损失和内容损失相结合进一步优化生成器的模型参数, 使生成器能够生成更精确的未来区域负荷, 而更精确的未来区域负荷也会使判别器更严格。因此, 在生成器和判别器互相博弈的过程中, 模型性能更加优越, 从而预测出比较精确的未来区域负荷。整个过程可以表示为

$$\min_c \max_D V(D, G) = E_{Y \sim P_{data}(Y)} [\log D(Y)] + E_{X \sim P_X(X)} \{\log \{1 - D[G(X)]\}\} \quad (11)$$

式中: $V(D, G)$ 为生成对抗网络的博弈目标函数; $D(\cdot)$ 为判别器; $G(\cdot)$ 为生成器; $\mathbf{Y}$ 为实际区域负荷数据; $E_{Y \sim P_{data}(Y)} [\log D(Y)]$ 为 $\log D(Y)$ 的期望; $\mathbf{X}$ 为基于历史区域负荷数据推断的未来区域负荷数据; $E_{X \sim P_X(X)} \{\log \{1 - D[G(X)]\}\}$ 为 $\log \{1 - D[G(X)]\}$ 的期望。

2.6 损失函数

为有效提升模型预测性能, 本文采用结合内容损失和对抗损失的双重损失机制进行联合优化, 其数学表达式为

$$L = L_{Load} + \lambda L_D \quad (12)$$

$$L_{Load} = \sum_{t=T_1+1}^{T_1+T_2} \sum_{n=1}^N |x_t^n - \hat{y}_t^n| \quad (13)$$

$$L_D = \log \{1 - D[G(X)]\} \quad (14)$$

式中: L_{Load} 为内容损失; L_D 为对抗损失; λ 为损失权重。

通过最小化预测值与真实值之间的绝对误差, 能够有效衡量模型的预测性能。内容损失能够提供稳定的梯度, 避免梯度爆炸, 提高模型的鲁棒性。通过最大化对抗损失, 能够迫使模型学习真实的数据分布, 从而生成更贴合实际的预测值。

3 算例分析

3.1 数据集分析与预处理

本文利用纽约 11 个区域 2023 年全年的区域负荷数据来验证 DAGAT 模型的性能。区域负荷数据的采样间隔为 1 小时, 则每个区域负荷节点有 8 760 个数据点。采用 Z-score 方法对数据集进行归一化处理。同时, 数据集按 6:2:2 的比例分为训练集、验证集和测试集。Z-score 归一化过程可以描述为

$$x_{norm} = \frac{x - x_{mean}}{x_{std}} \quad (15)$$

式中: x_{mean} 为所有区域负荷的平均值; x_{std} 为所有区域负荷的标准差; x 为区域负荷真实值; x_{norm} 为归一化值。

本文根据前 $T_1 = 12$ 小时的区域负荷来预测未来 $T_2 = 12$ 小时的区域负荷。

3.2 模型参数与评价指标

DAGAT 各个模块的超参数设置如表 1 所示。

此外,本文采用平均绝对误差 (Mean Absolute Error, MAE)、平均绝对百分比误差 (Mean Absolute Percentage Error, MAPE) 和均方根误差 (Root Mean Square Error, RMSE) 三个评价指标来评价 DAGAT 的预测性能,其表达式分别为

$$MAE = \frac{1}{s} \sum_i |\hat{y}_i - y_i| \quad (16)$$

$$RMSE = \sqrt{\left(\frac{1}{s} \sum_i (\hat{y}_i - y_i)^2\right)} \quad (17)$$

$$MAPE = \frac{1}{s} \sum_i \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \quad (18)$$

式中: s 为时间步长; $\hat{y}_i$ 为区域负荷预测值, y_i 为区域负荷真实值。

表 1 模型超参数设置

Tab. 1 Model hyperparameter settings

参数名称	参数值	参数名称	参数值
Batch-size	64	DWT level	1
Input-dim	1	Output-dim	1
Features	128	Kernel-size	2
Num-layer	2	Enepoch	100
Wavelet	Coiflets	Learning rate	0.001

3.3 算例结果分析

为验证所提 DAGAT 模型的预测性能,将其与各种主流模型进行比较分析。包括:自回归差分移动平均 (Autoregressive Integrated Moving Average, ARIMA)、支持向量回归 (Support Vector Regression, SVR) 此类传统的机器学习模型; Dlinear、Pyraformer、Autoformer、PatchTST、长短期记忆 (Long Short-Term Memory, LSTM) 网络和 Crossformer 此类仅能捕获时间依赖关系的预测模型; 时空图卷积网络 (Spatio-Temporal Graph Convolution Network, STGCN)、自适应图卷积递归网络 (Adaptive Graph Convolutional Recursive Network, AGCRN)、动态自适应和对抗图卷积网络 (Dynamic Adaptive and Adversarial Graph Convolutional Network, DAAGCN)、多元时间序列图神经网络 (Multivariate Time series Graph Neural Network, MTGNN)、STNorm 以及 GWNet 此类采用静态权重矩阵表达空间相关性的预测模型。不同模型的对比结果如图 10 所示。所有评价指标经三次试验取平均值,模型评价指标结果如表 2

所示。

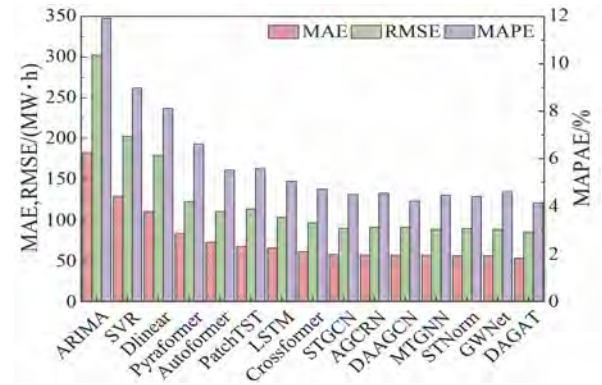


图 10 不同模型的性能对比结果

Fig. 10 Comparative results of the performance of different models

表 2 模型评价指标结果

Tab. 2 Results of model evaluation indicators

模型名称	MAE/(MW·h)	RMSE/(MW·h)	MAPE/%
ARIMA	183.054	302.839	11.92
SVR	128.414	202.464	8.98
Dlinear	110.304	179.174	8.12
Pyraformer	84.060	122.730	6.63
Autoformer	72.886	110.387	5.52
PatchTST	67.703	113.608	5.57
LSTM	65.730	103.416	5.04
Crossformer	61.196	96.670	4.72
STGCN	57.607	90.009	4.49
AGCRN	57.427	91.100	4.54
DAAGCN	56.690	91.350	4.24
MTGNN	56.439	88.957	4.46
STNorm	56.243	89.598	4.39
GWNet	55.949	88.515	4.60
DAGAT	52.954	85.056	4.16

由图 10 和表 2 可知,所有能捕获区域负荷空间依赖关系的模型优于只能捕获区域负荷序列内时间依赖关系的模型。这表明区域负荷不仅具有很强的时间依赖性,还具有空间依赖性。

此外,为了验证 DAGAT 在捕获区域负荷空间依赖性方面的优越性,将其与同样具备空间依赖性捕获能力的 STGCN、AGCRN 和 MTGNN 模型进行比较。与 STGCN 相比,DAGAT 的 MAE、RMSE 和 MAPE 分别降低了 4.653 MWh、4.953 MWh 和 0.33%;与 AGCRN 和 MTGNN 相

比,DAGAT 同样具有优越性。这是因为 STGCN、AGCRN 和 MTGNN 模型使用相同的权重聚合邻居节点信息,为一个静态过程;而 DAGAT 中的

GAT 通过注意力机制为不同的相邻居点分配不同的聚合权重。DAGAT 模型聚合邻居节点信息的权重动态变化过程如图 11 所示。

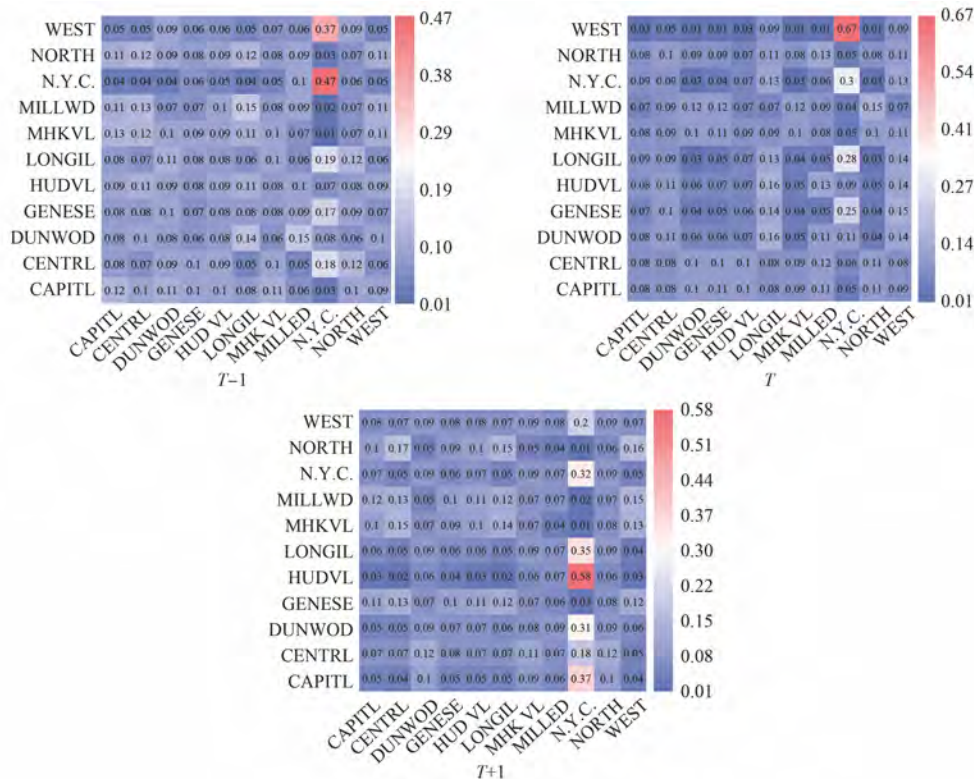


图 11 权重动态变化过程

Fig. 11 Dynamic weight change process

从图 11 可知,不同时刻的信息聚合权重明显不同,权重越大表示邻居节点信息对当前节点的影响越大;同时也能避免一些空间关联性较低的节点对当前节点造成干扰。动态变化的聚合权重能够捕获区域负荷间的动态空间依赖关系,从而提高模型的预测性能。

由 DAGAT 模型与 AGCRN、GWNNet^[10] 和 DAAGCN^[28] 此类具有动态图结构的模型的结果可知,相较于 AGCRN、DAAGCN 和 GWNNet, DAGAT 的 MAE 分别下降了 4.473 MWh、3.736 MWh 和 2.995 MWh, DAGAT 模型仍然具有优势。这是因为 AGCRN、GWNNet 和 DAAGCN 虽然具有动态图结构,能够动态描述节点之间的连接关系,但无法动态衡量影响大小,因为其节点信息聚合过程仍遵循 GCN 的静态权重分配。

为了更直观地分析模型之间的预测性能,图 12 展示了不同模型的预测误差分布结果。其中,横坐标为预测值与真实值的差值区间,纵坐标为

差值落在某一差值区间的次数。

由图 12 可知,ARIMA 模型等预测性能较差的模型其误差分布分散。DAGAT 模型的误差分布相对集中,主要在 0 附近。可见,与其他预测模型相比,DAGAT 模型在大多数情况下预测误差较小,在区域负荷预测中显示出巨大的潜力。

为了进一步分析 DAGAT 模型的预测性能,随机选择四个预测窗口,如图 13 所示。图中参考线的左侧为区域负荷的历史值,右侧为预测值和相应的实际值。可见,与其他模型相比,DAGAT 模型的预测值与实际值之间的偏差较小。

DAGAT 模型在 7 天时间内的预测结果如图 14 所示。可见,在大多数情况下,预测曲线总是与实际曲线重叠,随着实际值的变化而变化。

综上,模型比较结果验证了 DAGAT 模型在区域负荷预测上的优越性。

3.4 消融试验

为进一步验证 DWT、对抗训练和 GAT 对模

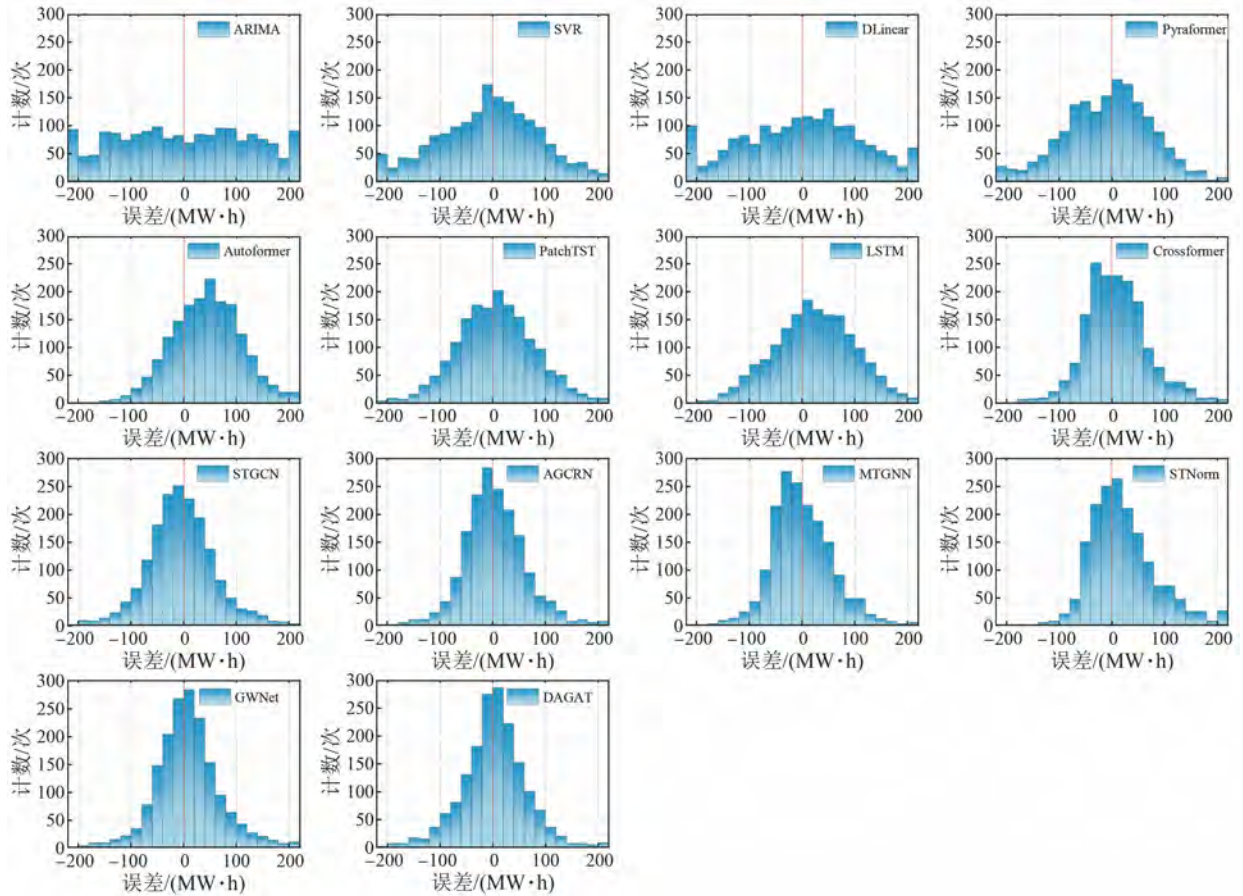


图 12 不同模型的预测误差分布

Fig. 12 Distribution of prediction errors for different models

型的影响。本文对 DWT、双重损失机制和 GAT 模块进行了消融试验,结果如表 3 所示。其中,无 DWT 的 DAGAT 表示模型中去除了 DWT 模块,直接对输入的原始序列进行特征提取;无 GAT 的 DAGAT 表示模型中去除 GAT 模块,仅通过捕获数据的时间依赖性进行预测;无 GAN 的 DAGAT 表示不引入双重损失模型,即模型的训练目标是使内容损失最小化。

表 3 消融试验结果

Tab. 3 Ablation experiment results

模型名称	MAE/(MW·h)	RMSE/(MW·h)	MAPE/%
无 DWT	62.1776	98.569 9	4.81
无 GAT	59.591 7	95.252 9	4.65
无 GAN	54.846 0	86.420 1	4.31
DAGAT	52.954 0	85.056 6	4.16

由表 3 可知,无 DWT 的 DAGAT 预测效果最差,这是因为直接将原始序列输入到模型中,会导

致针对趋势分量和波动分量特征不同而设计的时空特征提取模块失去其针对性。与无 GAT 的 DAGAT 相比,DAGAT 的 MAE、RMSE 和 MAPE 分别降低了 6.637 7 MWh、10.196 3 MWh 和 0.46%,这是因为 GAT 模块凭借有效捕获区域负荷间的动态空间依赖关系,能够提高模型的预测精度。与无 GAN 的 DAGAT 相比,DAGAT 的 MAE、RMSE 和 MAPE 分别降低了 1.862 MWh、1.363 5 MWh 和 0.15%,这是因为对抗训练过程能不断优化预测网络模型的参数,从而可以准确地学习区域负荷的潜在特征,以获得准确的未来区域负荷。

4 结语

本文考虑到区域负荷间存在动态空间依赖性以及负荷序列特征难以捕获,提出了一种基于 DAGAT 模型的短期 MRLF 方法。与主流模型的比较结果表明,本文所提 DAGAT 模型具有优越性。该模型的优越性源于 GAT 通过学习动态变

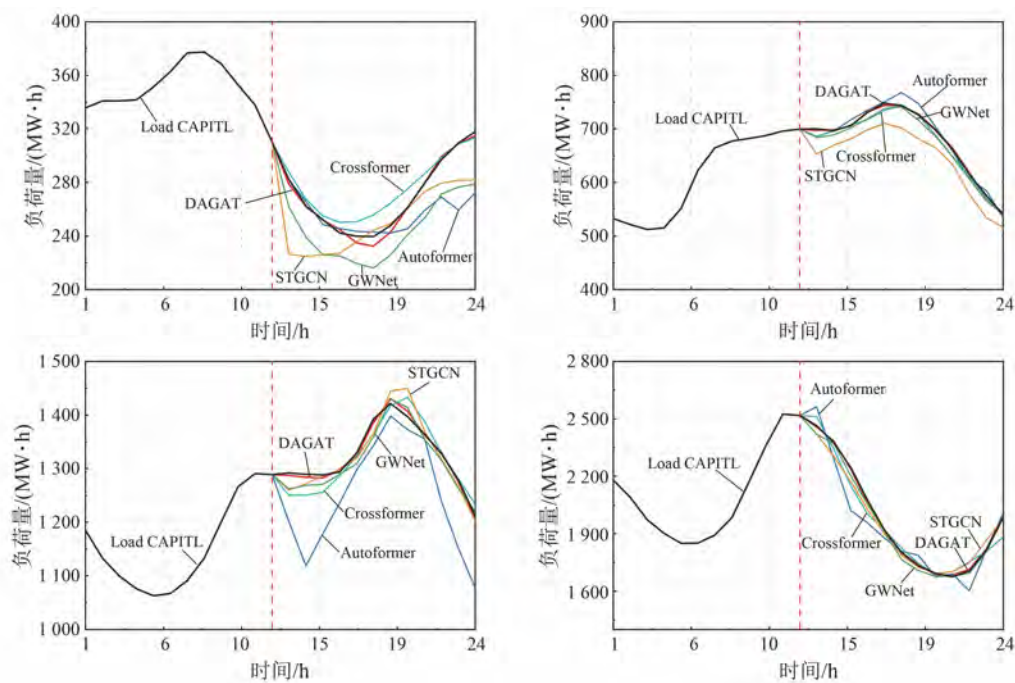


图 13 不同模型预测值曲线对比

Fig. 13 Comparison of predicted value curves of different models

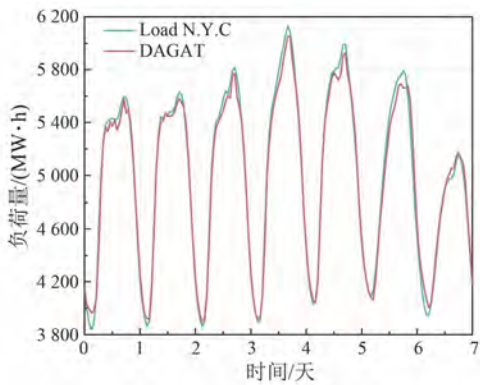


图 14 预测值曲线与真实值曲线比较

Fig. 14 Comparison between predicted value curve and actual value curve

化的权重捕获了多区域负荷间的空间相关性, DWT 清晰表示了多区域负荷序列的特征,联合损失进一步优化了模型参数。DAGAT 不仅有效提高了 MRLF 精度,还为捕获区域负荷间动态变化的空间依赖关系提供了基础,提高了其在工程实践中的有效性和可信度。

利益冲突声明

所有作者声明不存在利益冲突。

All authors disclose no relevant conflict of interests.

作者贡献

刘璟璐进行了方案设计、内容总结与论文撰写,曹文涛进行了试验研究,董生爱、张鹏飞和张子昌参与了论文的审核与修改。所有作者均阅读并同意了最终稿件的提交。

The scheme design, content summary and paper writing were carried out by Liu Jinglu. The experiment study was conducted by Cao Wentao. The manuscript was reviewed and revised by Dong Shengai, Zhang Pengfei and Zhang Zichang. All authors have read the last version of paper and consented for submission.

参考文献

[1] GONG D Y, CHEN N, JI Q F, et al. Multi-scale regional photovoltaic power generation forecasting method based on sequence coding reconstruction [J]. Energy Reports, 2023, 9(S10): 135-143.

[2] HUANG H H, HUANG Y H. A novel green learning artificial intelligence model for regional electrical load prediction [J]. Expert Systems with Applications, 2024, 256: 124907.

[3] 张浩, 陆海, 罗恩博, 等. 考虑资源灵活性的多区域综合能源系统协同调度[J]. 电网与清洁能源,

- 2024, 40(9): 148-154.
- ZHANG H, LU H, LUO E B, et al. Coordinated scheduling of multi-region integrated energy systems considering resource flexibility [J]. *Power System and Clean Energy*, 2024, 40 (9): 148-154.
- [4] LIU J C, LI T. Multi-step power forecasting for regional photovoltaic plants based on ITDE-GAT model [J]. *Energy*, 2024, 293: 130468.
- [5] EREN Y, KÜÇÜKDEMİRAL I. A comprehensive review on deep learning approaches for short-term load forecasting [J]. *Renewable and Sustainable Energy Reviews*, 2024, 189(B): 114031.
- [6] QU K, SI G Q, SHAN Z H, et al. Forwardformer: Efficient transformer with multi-scale forward self-attention for day-ahead load forecasting [J]. *IEEE transactions on power systems*, 2023, 39(1): 1421-1433.
- [7] 刘可真, 梁玉平, 代莹皓, 等. 基于 CNN-ISSA-GRU 模型的短期负荷预测方法[J]. *电机与控制应用*, 2022, 49 (8): 79-86.
- LIU K Z, LIANG Y P, DAI Y H, et al. Short-term load forecasting method based on CNN-ISSA-GRU model [J]. *Electric Machines & Control Application*, 2022, 49 (8): 79-86.
- [8] 李楠, 马宏忠. 基于综合预测模型和蒙特卡洛的电动汽车保有量及负荷预测方法研究[J]. *电机与控制应用*, 2022, 49(12): 74-80.
- LI N, MA H Z. Research on Electric vehicle ownership and load prediction method based on comprehensive prediction model and Monte Carlo [J]. *Electric Machines & Control Application*, 2022, 49 (12): 74-80.
- [9] 朱江, 汪帆, 曹春堂, 等. 基于 KLPP-K-means-BiLSTM 的台区短期电力负荷预测[J]. *电机与控制应用*, 2024, 51(3): 108-116.
- ZHU J, WANG F, CAO C T, et al. KLPP-K-means-BiLSTM based short-term power load forecasting for station areas [J]. *Electric Machines & Control Application*, 2024, 51 (3): 108-116.
- [10] LIU J L, DONG S A, ZHANG P F, et al. Load forecasting based on dynamic adaptive and adversarial graph convolutional networks [J]. *Energy and Buildings*, 2024, 312: 114206.
- [11] 耿贞伟, 李申章, 于凤荣. 考虑空间关联的超短期、短期多区域电力负荷单步和多步预测[J/OL]. *计算机工程*, 2024-07-16. <https://doi.org/10.19678/j.issn.1000-3428.0069490>.
- GENG Z W, LI S Z, YU F R. Ultra-short-term and short-term power load single-step and multi-step prediction considering spatial association [J/OL]. *Computer Engineering*, 2024-07-16. <https://doi.org/10.19678/j.issn.1000-3428.0069490>.
- [12] 赵紫昱, 陈渊睿, 陈霆威, 等. 基于时空图注意力网络的超短期区域负荷预测[J]. *电力系统自动化*, 2024, 48(12): 147-155.
- ZHAO Z Y, CHEN Y R, CHEN T W, et al. Ultra-short-term regional load forecasting based on spatio-temporal graph attention network [J]. *Automation of Electric Power Systems*, 2024, 48 (12): 147-155.
- [13] HUANG Z W, YI Y W. Short-term load forecasting for regional smart energy systems based on two-stage feature extraction and hybrid inverted Transformer [J]. *Sustainability*, 2024, 16 (17): 7613.
- [14] 朱涛, 金从友, YU Y M, 等. 自适应地理图神经网络短期电力负荷预测模型[J]. *中国测试*, 2024, 50(增刊2): 223-229.
- ZHU T, JIN C Y, YU Y M. Short-term power load forecasting model based on adaptive geographic map neural network [J]. *China Measurement & Test*, 2024, 50 (S2): 223-229.
- [15] 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 P F, HU B, HE J S, 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.
- [16] 吴军英, 路欣, 刘宏, 等. 基于 Spearman-GCN-GRU 模型的超短期多区域电力负荷预测[J]. *中国电力*, 2024, 57(6): 131-140.
- WU J Y, LU X, LIU H, et al. Ultra-short-term multi-region power load forecasting based on Spearman-GCN-GRU model [J]. *Electric Power*, 2024, 57 (6): 131-140.
- [17] WEI C Y, PI D C, PING M T, et al. Short-term load forecasting using spatial-temporal embedding graph neural network [J]. *Electric Power Systems Research*, 2023, 225: 109873.
- [18] XIAO W J, MO L, XU Z X, et al. A hybrid electric load forecasting model based on decomposition considering fisher information [J]. *Applied Energy*,

- 2024, 364: 123149.
- [19] ZHANG M Y, HAN Y, ZALHAF A S, et al. Accurate ultra-short-term load forecasting based on load characteristic decomposition and convolutional neural network with bidirectional long short-term memory model [J]. Sustainable Energy, Grids and Networks, 2023, 35: 101129.
- [20] SU Y X, HE Q Y, CHEN J, et al. A residential load forecasting method for multi-attribute adversarial learning considering multi-source uncertainties [J]. International Journal of Electrical Power & Energy Systems, 2023, 154: 109421.
- [21] OUYANG X C, YANG Y, ZHOU W, et al. Citytrans: Domain-adversarial training with knowledge transfer for spatio-temporal prediction across cities [J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 36(1): 62-76.
- [22] DUDWK G. STD: A seasonal-trend-dispersion decomposition of time series [J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(10): 10339-10350.
- [23] 李青, 张新燕, 马天娇, 等. 基于 SSA-CNN-BiGRU-Attention 的超短期风电功率预测模型 [J]. 电机与控制应用, 2023, 50(5): 61-71.
- LI Q, ZHANG X Y, MA T J, et al. Ultra-short term forecasting model of wind power based on SSA-CNN-BiGRU-Attention [J]. Electric Machines & Control Application, 2023, 50(5): 61-71.
- [24] FANG Y C, QIN Y J, LUO H Y, et al. STWave+: A multi-scale efficient spectral graph attention network with long-term trends for disentangled traffic flow forecasting [J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 36(6): 2671-2685.
- [25] 谢乐, 杨浙, 刘东. 基于多头注意力机制和门控循环单元神经网络的居民充电桩容量预测 [J]. 电机与控制应用, 2024, 51(3): 21-29.
- XIE L, YANG Z, LIU D. Residential charging station capacity prediction based on multi-head attention and gated recurrent unit neural network [J]. Electric Machines & Control Application, 2024, 51(3): 21-29.
- [26] WU D, LIN W X. Efficient residential electric load forecasting via transfer learning and graph neural networks [J]. IEEE Transactions on Smart Grid, 2022, 14(3): 2423-2431.
- [27] LI Y Y, SONG L D, HU Y, et al. Load profile inpainting for missing load data restoration and baseline estimation [J]. IEEE Transactions on Smart Grid, 2023, 15(2): 2251-2260.
- [28] 胡博, 张鹏飞, 黄恩泽, 等. 基于图 WaveNet 的电动汽车充电负荷预测 [J]. 电力系统自动化, 2022, 46(16): 207-213.
- HU B, ZHANG P F, HUANG E Z, et al. Graph WaveNet based charging load forecasting of electric vehicle [J]. Automation of Electric Power Systems, 2022, 46(16): 207-213.

收稿日期:2025-02-25

收到修改稿日期:2025-04-11

作者简介:

刘璟璐(1989-),男,博士,副教授,研究方向为分布式算法、人工智能模型,bradybest@163.com;

*通信作者:曹文涛(2000-),男,硕士研究生,研究方向为负荷预测,cwtaea@126.com。