

# 基于自适应参数聚合的联邦双分支图神经网络及其在用电量需求预测中的应用

钟磊<sup>1</sup>, 姜雪娇<sup>1</sup>, 郑楷洪<sup>2\*</sup>, 徐佳隆<sup>1</sup>, 符峥<sup>1</sup>, 高磊<sup>1</sup>, 李胜<sup>2</sup>

<sup>1</sup>中国南方电网海南电网有限责任公司, 海南 海口

<sup>2</sup>南方电网数字电网集团有限公司, 广东 广州

收稿日期: 2025年8月16日; 录用日期: 2025年9月15日; 发布日期: 2025年9月23日

## 摘要

用电量需求预测是智能电网领域的重要研究方向, 需要对用电量数据的时序特征和峰值变化进行精确建模。然而, 不同地区的用电量数据特性及其需求存在差异, 且部分地区的用电情况属于敏感行业数据, 需要进行隐私保护。联邦学习为用电量需求预测任务提供了一种兼顾隐私保护的可行方案, 但传统的联邦学习方法在该任务中面临两大挑战: 一是当本地客户端的用电量数据不足时, 其本地模型容易过拟合到含噪声的峰值变化; 二是不同客户端的数据特征可能存在显著差异, 部分客户端的用电量模式以周期性为主, 另一部分则具有明显的峰值变化, 导致不同客户端的本地模型参数不一致。相应地, 在服务器端聚合这些差异较大的模型会导致全局模型在部分区域的用电数据上出现过平滑现象, 降低其对峰值变化的拟合能力。为克服这些挑战, 本文提出了一种基于自适应参数聚合的联邦双分支图神经网络 (Federated Dual-branch Graph Neural Network with adaptive parameter aggregation, FDGN) 用于用电量需求预测。对于局部模型的构建, FDGN首先采用一种混合特征表示, 捕捉用电量数据的时序特性和数值波动特性。鉴于时序模式和峰值变化具有不同的数据特征且对预测均至关重要, FDGN采用双分支图对其分别建模: 首先设计了用电量时序图及其对应的多尺度图注意力网络以提取用电量数据的时序模式特征, 然后设计数值相关图及其对应的高斯自编码器, 将峰值变化建模为参数化高斯分布中的动态协方差特征, 从而缓解过拟合问题。最后, 融合这两部分特征以生成最终预测结果。对于全局模型的构建, FDGN通过在服务器端引入基于相似度的自适应动态参数聚合方法, 缓解全局模型的过平滑问题。实验结果表明, 在用电量需求预测中, FDGN优于常用的联邦学习方法。

## 关键词

联邦学习, 用电量需求预测, 图神经网络

\*通讯作者。

# Federated 2-Branch Graph Neural Network Based on Adaptive Parameter Aggregation and Its Application in Electricity Demand Forecasting

Lei Zhong<sup>1</sup>, Xuejiao Jiang<sup>1</sup>, Kaihong Zheng<sup>2\*</sup>, Jialong Xu<sup>1</sup>, Zheng Fu<sup>1</sup>, Lei Gao<sup>1</sup>, Sheng Li<sup>2</sup>

<sup>1</sup>Hainan Power Grid Co., Ltd., China Southern Power Grid, Haikou Hainan

<sup>2</sup>Digital Grid Research Institute, China Southern Power Grid, Guangzhou Guangdong

Received: Aug. 16<sup>th</sup>, 2025; accepted: Sep. 15<sup>th</sup>, 2025; published: Sep. 23<sup>rd</sup>, 2025

## Abstract

Electricity demand forecasting stands as a critical research area within smart grids, requiring accurate modeling of the temporal characteristics and peak variations inherent in consumption patterns. However, significant disparities exist in electricity data characteristics and demand patterns across different regions, with consumption data from certain areas constituting sensitive industrial data necessitating robust privacy protection. Federated learning offers a viable privacy-preserving solution for electricity demand forecasting tasks. Nevertheless, traditional federated learning approaches encounter two primary challenges in this context: 1) local models at clients with insufficient electricity data are prone to overfitting on noisy peak variations; and 2) substantial heterogeneity in data features exists among clients, leading to parameter inconsistency across local models. Consequently, aggregating these divergent models at the server results in an over-smoothed global model that compromises its ability to capture peak fluctuations effectively. To overcome these challenges, this paper proposes a Federated Dual-branch Graph Neural Network with adaptive parameter aggregation (FDGN) tailored for electricity demand forecasting. Within the local model framework, FDGN first constructs a hybrid feature representation capturing both the temporal evolution and numerical volatility of electricity consumption data. Recognizing that temporal patterns and peak variations possess distinct characteristics yet are equally crucial for prediction, FDGN employs a dual-branch graph for parallel modeling: it designs an electricity time-series graph coupled with a corresponding multi-scale attention network to extract temporal patterns, and constructs a numerical structure graph paired with a Gaussian autoencoder to represent peak variations as dynamic covariance features within a parameterized Gaussian distribution, thereby mitigating overfitting. These temporal features and covariance features are ultimately fused to generate the final prediction. For global model construction, FDGN addresses the over-smoothing issue during aggregation by introducing a similarity-based adaptive dynamic parameter aggregation mechanism at the server. Experimental results verify that FDGN outperforms conventional federated learning methods in electricity demand forecasting.

## Keywords

Federated Learning, Electricity Consumption Demand Forecasting, Graph Neural Network

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着智能电网技术演进,用电量需求预测已发展为支撑电网智能化运营的一项关键技术[1]。尽管现有研究已提出多种基于时间序列分析的预测模型[2],但该领域仍存在两个关键问题:其一,用电峰值波动特征的精确建模与其噪声鲁棒性之间的平衡;特别值得注意的是“数据孤岛”现象的存在,使得局部区域因训练样本不足而引发的模型过拟合问题尤为突出,制约了传统预测模型的实际应用性能[3][4]。其二,用电量数据固有的空间离散化分布特性,即数据分散存储于不同地理区域及产业部门[3],这种分布模式不仅因涉及敏感工业信息而需遵循严格的隐私保护规范,更导致数据难以进行集中式整合分析。在此背景下,联邦学习(Federated Learning, FL)作为一种新兴的分布式机器学习范式,为同时解决数据隐私保护与跨区域协同建模这一对矛盾需求提供了新的技术路径。

尽管联邦学习已被证实为用电量预测领域有效的分布式机器学习方法[5],当前大部分基于联邦学习的预测方法仍受制于两个难点:其一,“数据孤岛”效应导致客户端(如区域电力公司)本地数据稀缺时,模型易陷入过拟合困境,稀疏的历史用电记录难以全面建模复杂的用电行为模式,特别是对生产事件敏感且受噪声干扰的用电峰值波动特征[6],这种数据匮乏性会迫使局部模型过拟合到特定噪声模式,削弱其在新场景下的泛化性能[7];其二,传统的联邦聚合机制存在特征稀释风险,当整合具有显著参数差异的本地模型时(源于训练数据分布异质性[8]),全局模型容易过度平滑化,无法有效保留客户端特有的用电特征等关键判别信息[9]。上述挑战共同揭示了开发适用于用电量预测任务的 FL 框架的必要性,需在用电量需求预测任务中实现模型泛化能力与客户端特定适应性的平衡。

针对联邦学习框架下用电量需求预测存在的局部数据过拟合与全局模型过平滑挑战,本文提出了一种基于自适应参数聚合的联邦双分支图神经网络(Federated Dual-branch Graph Neural Network with adaptive parameter aggregation, FDGN)。所提出的 FDGN 在联邦学习架构中同时建模用电数据的时序特征和峰值变化特征,其主要工作集中在局部模型和全局模型构建两方面,具体而言,在局部建模阶段,FDGN 首先整合时间序列分析技术,构建用电量数据的联合特征表示,充分表达用电数据的时序特征和峰值变化特征。然后基于这种联合特征表示,FDGN 设计了双分支图结构来实现时序特征和峰值变化特征的建模和处理:双分支图结构的第一个分支是设计的用电量时序图及其对应的多尺度图注意力网络以提取用电量数据的时序模式特征,捕获用电数据中跨时间尺度的周期模式与趋势;另一个分支是设计的数值相关图及其对应的高斯自编码器,对于易导致过拟合的峰值变化特征,该分支将峰值变化特征建模为参数化高斯分布中的期望和动态协方差,通过在预测均值上引入动态协方差,从而在预测峰值变化数据时产生可控的预测波动,在保障预测精确性的同时通过这种波动性缓解过拟合问题。最后,FDGN 融合两个分支的特征以生成最终预测结果。实验表明,FDGN 在 2 个真实电网数据集上的测试结果较基线有所提升,验证了 FDGN 的有效性。

本文的主要贡献如下:

- 1) 提出了基于自适应参数聚合的联邦双分支图神经网络 FDGN,分别建模用电量数据的时序特征和峰值变化模式。通过整合序列分析和双分支图,FDGN 减轻了客户端数据有限导致的过拟合,捕捉用电峰值变化。

- 2) 针对全局模型提出了动态融合聚合方法,在全局模型构建过程中引入相似度参数,通过动态聚合与参数化高斯分布形式的建模来缓解过平滑问题。旨在提高不同用电模式下的预测精度和泛化能力。

本文的结构安排如下:第 2 节回顾了用电量预测的相关工作。第 3 节介绍所提出的 FDGN 框架。第 4 节通过实验验证所提出的 FDGN 的有效性,最后一节总结全文。

## 2. 相关工作

### 2.1. 传统用电量预测方法

传统的用电量预测方法主要包括基于时间序列的方法、基于回归的方法和灰色预测方法[2]。时间序列方法分析用电量的历史时间序列数据以预测未来值。然而, 这些方法假设数据具有一定的统计规律性, 且对数据异常值敏感[3]。基于回归的方法, 如多元线性回归, 建立用电量与影响因素之间的函数关系[2][10]。但它们在处理复杂的非线性关系时往往面临困难。灰色预测方法适用于数据有限的预测场景, 但长期预测的精度相对较低[11]。

### 2.2. 基于机器学习的预测方法

随着机器学习技术的发展, 神经网络、支持向量机和随机森林等方法已广泛应用于用电量预测[12]。神经网络, 特别是深度神经网络, 能够自动从数据中提取特征, 并具有强大的非线性拟合能力[13]。例如, 长短期记忆(Long Short-Term Memory, LSTM)网络能有效处理时间序列数据, 并在用电量预测中取得了良好效果[14]。然而, 这些基于机器学习的方法通常需要大量集中数据进行训练, 在面临数据孤岛问题时难以实现[3]。深度学习的出现标志着该领域的变革, 随后出现了用于分层特征提取的多尺度 LSTM 架构和双向 LSTM (Bidirectional LSTM, BiLSTM)模型[15], 它们结合双向上下文以提高短期预测精度。同时, 频域方法, 如经验模态分解(Empirical Mode Decomposition, EMD) [16]和傅里叶分析[17]的信号分解, 也得到了广泛应用。尽管基于傅里叶的神经架构能有效捕捉周期性模式, 但它们往往在时间分辨率和频率定位之间难以平衡, 可能导致瞬态特征丢失。基于小波的多分辨率分析[18]通过使用适合异质时间序列数据的自适应基函数缓解了这一问题, 但却引入了更高的计算复杂度。

### 2.3. 联邦学习应用

联邦学习是一种保护隐私的机器学习范式, 支持在去中心化数据源上协同训练共享模型, 同时保持数据机密性[19]。与传统的集中式方法不同, FL 采用客户端 - 服务器架构, 客户端(如移动设备、医院或工业传感器)使用自己的数据训练本地模型, 仅将模型更新发送给中央服务器进行聚合[20]。其主要挑战包括处理非独立同分布(non-IID)数据分布、优化通信效率和减轻安全漏洞[21]。最新进展提出了联邦平均(Federated Averaging, FedAvg) [22]、自适应聚合算法和差分隐私机制[23]等解决方案, 展示了 FL 在医疗、金融和物联网领域的适用性。智能电网因其分布式基础设施和严格的数据隐私要求, 成为 FL 的一个有吸引力的应用场景[24]。FL 使电力公司和用户能够协同训练电力需求预测和负载均衡模型[25], 而无需集中收集数据。

## 3. 提出的 FDGN 模型

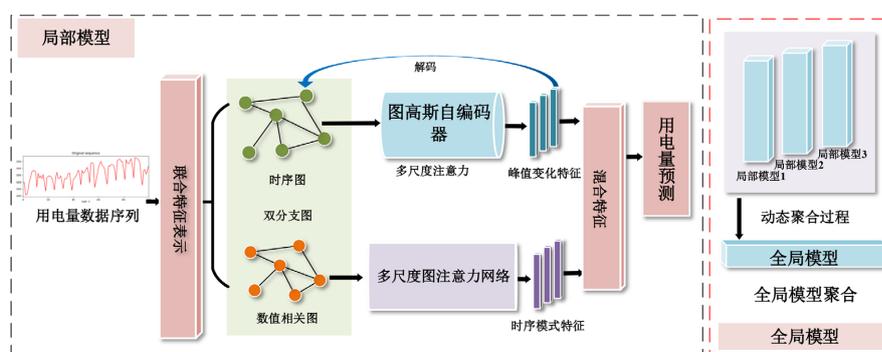


Figure 1. The model graph of FDGN  
图 1. FDGN 模型图

本文提出了具有自适应参数聚合的联邦双分支图神经网络用于用电量需求预测。下面详细介绍该模型。FDGN 的模型结构如图 1 所示。

本文提出了一种基于具有自适应参数聚合的联邦双分支图神经网络的用电量需求预测方法。具体过程如下。

### 3.1. 用电量需求预测局部模型构建

本文认为用电量数据本质上具有双分支属性：时间维度反映用电量随时间的变化，用电量沿时间序列变化；空间维度反映用电量数据峰值波动之间的数值相关性。当前用电量数据的预测依赖于时序变化趋势和用电量峰值变化的数值相关性。因此，为了建模用电量数据的这些分支属性，本文设计了双分支图。

#### 3.1.1. 数据预处理与双分支图构建

在预处理中，进行特征融合增强，使用滑动窗口计算统计特征(均值/方差)。使用基于局部加权回归的季节-趋势分解(Seasonal-Trend Decomposition Using LOESS, STL)将数据分解为三个分量：[趋势分量  $T$ 、季节分量  $S$ 、残差分量  $R$ ]。然后，对用电量数据应用集成经验模态分解(Ensemble Empirical Mode Decomposition, EEMD)以获得模态函数(IMF)分量。因此，对于每个时间片，数据向量  $f$  由  $[data, T, S, R, Mean, Variance, IMF_1, \dots, IMF_k]$  组成，从而构建用电量的联合数据表示。

接下来，构建用电量数据的双分支图。我们在公式(1)中将用电量数据表示为双分支图。

$$DualG = (V, E_{temp}, E_{sim}) \quad (1)$$

设时间序列为  $X = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^{n \times d}$ ，其中每个时间点  $x_t$  对应一个图节点  $v_t$ ，节点特征向量为  $f_t = [x_t, T_t, S_t, R_t, mean_t, variance_t, IMF_{1t}, \dots, IMF_{kt}]$ 。DualG 包含两种类型的边，一种边用于建模时序关系，表示为  $E_{temp}$ ，对应的邻接矩阵为  $A_{temp}$ 。另一种边建模用电量数据中的数值相关性，表示为  $E_{sim}$ ，对应的邻接矩阵为  $A_{sim}$ 。具体而言，边的计算过程如下。

为了建模用电量数据的时间特性，时序边定义如公式(2)

$$E_{temp}(t) = \left\{ (f_t, f_{t+k}) = 1 - \frac{k}{K} \mid k \in [1, \dots, K-1] \right\} \quad (2)$$

其中  $K$  是时间窗口， $f_t = [x_t, T_t, S_t, R_t, mean_t, variance_t, IMF_{1t}, \dots, IMF_{kt}]$ ， $E_{temp}$  是边。在这种情况下，时间序列中与当前时间片  $t$  较近的节点将具有较高的权重，连接权重随时间衰减。这与用电量随时间变化的相关模式一致。

为了在建模用电量时序模式的同时确保预测值的准确性，我们需要建模用电量的数值相关性。基于这些相关性，指导当前时间片的用电需求预测。因此，我们定义特征相似性边  $E_{sim}$  如公式(3)

$$E_{sim}(t, s) = \left\{ (f_t, f_s) \mid \text{Cosine}(f_t, f_s) > \tau \right\} \quad (3)$$

其中  $\text{Cosine}$  表示余弦相似度。为了防止过平滑、易于预测的用电量数据的干扰，我们设置阈值  $\tau$  来过滤、留下具有明显波动性的数据。然后，基于这些数据的相关性，建模用电量过程中的数值关系。

由此，我们构建了双分支图。接下来，利用这些双分支图设计基于图神经网络的预测模型。

#### 3.1.2. 用于用电量需求预测的图神经网络

FDGN 设计了一种多尺度的图注意力机制来建模双分支图。

##### 1) 对于时序图：

FDGN 设计了多尺度图自注意力神经网络，提取不同尺度的注意力特征，建模不同时间序列长度的时序特征。具体而言，FDGN 利用多尺度图注意力网络提取多尺度注意力特征，它包含三个核心部分：(1) 量化图中的节点之间关联的相似度量；(2) 建模输入关系差异影响的注意力系数；(3) 构建多尺度聚合的注意力特征。

两个输入向量之间的相似度计算如公式(4)

$$Sim_{ij}^h = \text{Cosine}\left(A_{temp}^h \times P^h \times f_i, A_{temp}^h \times P^h \times f_j\right), \quad (4)$$

其中  $A_{temp}^h$  是图邻接矩阵的  $h$  尺度邻域， $Sim_{ij}^h$  是  $i$  和  $j$  之间的  $h$  尺度相似度特征， $P^h$  是可学习权重矩阵，基于此  $h$  尺度相似度函数， $h$  尺度注意力系数定义如公式(5)

$$\alpha_{ij}^h = \text{softmax}\left(Sim_{ij}^h\right), \quad (5)$$

其中  $\alpha_{ij}^h$  是注意力系数。FDGN 计算  $H$  个尺度的注意力系数，然后基于  $H$  个尺度的注意力系数计算  $i$  的注意力特征如公式(6)

$$feature_i = \sum_h \sum_j \alpha_{ij}^h \times Q^h \times f_j, \quad (6)$$

其中  $Q^h$  是可训练权重。

## 2) 对于数值相关图：

FDGN 设计了基于多跳注意力的图高斯自编码器，用于建模用电量的数值特征。本文设计的基于多跳注意力的编码器分为 3 跳，过滤出与当前用电量值最相关的特征。

### ➤ 第一跳：数值相关信息过滤

以节点  $i$  的注意力特征为例，首先全局扫描输入的数值相关图的图节点，计算初步注意力权重，然后选择重要区域或节点。

$$\alpha_{ij}^{(1)} = \text{softmax}\left(\frac{(A_{sim} f_i) * (A_{sim} f_j^T)}{\sqrt{d_k}}\right) \quad (7)$$

其中， $A_{sim}$  是邻接矩阵，超参数  $d_k$  设置为 1， $f$  是特征， $T$  是转置操作。

### ➤ 第二跳：依赖关系建模

基于第一跳的结果，进一步分析所选区域内的子结构。

$$\alpha_{ij}^{(2)} = \alpha_{ij}^{(1)} + \text{softmax}\left((W \circ A_{sim}) f_i * (W \circ A_{sim}) f_j^T\right) \quad (8)$$

这里， $W$  是随机初始化的可学习掩码矩阵。

### ➤ 第三跳：通过高阶推理整合为参数化高斯分布

聚合多跳结果，形成最终的注意力分布特征。

$$Z_i = N\left(\sum_{k=1}^2 W_1 \left(\sum_j \alpha_{ij}^{(k)} f_j\right), \theta \times \sum_{k=1}^2 W_2 \left(\sum_j \alpha_{ij}^{(k)} f_j\right)\right) \quad (9)$$

其中， $W_1$  和  $W_2$  是可学习参数，特征被参数化为高斯分布的均值和协方差。解码器随后通过合成这些加权关联来增强注意力的表示。

### 3.1.3. 基于双分支特征的用电量预测

经过上述步骤，合成双分支特征计算用电量数据的最终预测，如公式(10)所示。

$$y = \text{SLP-layer}(Z \odot \text{feature}). \quad (10)$$

这里, SLP-layer 是输出预测结果的全连接输出层,  $Z \odot feature$  表示  $Z$  与  $feature$  的元素积。任务目标是最小化如公式(11)所示的损失。

$$Loss = \frac{1}{N} \sum_n \left\{ \lambda \sum_{i,j} \left[ \text{Cosine}(Z_i^n, Z_j^n) - A_{sim\_i,j} \right]_n^2 + \text{CrossEntropy}(y^n, label^n) \right\}. \quad (11)$$

其中,  $CrossEntropy$  表示交叉熵,  $\lambda$  表示相似性度量系数, 用于平衡重构损失和交叉熵, 这里我们取 0.005。

### 3.2. 基于模型参数相似度的全局自适应动态参数聚合

为了解决客户端模型在全局模型聚合中的过平滑问题, 我们提出了一种基于模型参数相似度的自适应参数聚合方法。对于全局模型, 通过参数差异检测和权重调整解决过平滑问题, 同时在局部模型中引入全局正则化项, 防止本地模型对不一致的模型参数过拟合。

在服务器端:

- 差异度量: 计算第  $i$  个本地局部模型参数  $w_i$  与当前全局参数  $w_g$  之间的  $L2$  距离  $D_i = \|w_i - w_g\|^2$ 。
- 权重生成: 使用 softmax 函数计算聚合权重  $\alpha_i$ 。
- 聚合执行如公式(12)所示

$$w_{g\_new} = \sum_i \alpha_i w_i + \varepsilon w_{g\_old} \quad (12)$$

其中  $\varepsilon$  是历史参数保留系数。

在客户端, 使用接收的全局模型更新本地模型时, 引入全局模型正则化项, 损失函数如公式(13)所示

$$Loss = \frac{1}{N} \sum_n \left\{ \lambda \sum_{i,j} \left[ \text{cosine}(Z_i^n, Z_j^n) - A_{sim\_i,j} \right]_n^2 + \text{CrossEntropy}(y^n, label^n) \right\} + \beta \|Parameter_{local}, Parameter_{global}\|^2 \quad (13)$$

这里, parameter 指用于预测的全连接层 SLP-layer 的模型参数。

## 4. 实验分析

我们的实验通过比较分析评估三个关键结论: (1) 所提出的 FDGN 在预测方面优于基准方法; (2) 提出的双分支图相比传统的序列建模方法具有优势; (3) 基于相似度的自适应参数聚合能有效缓解参数不一致问题。

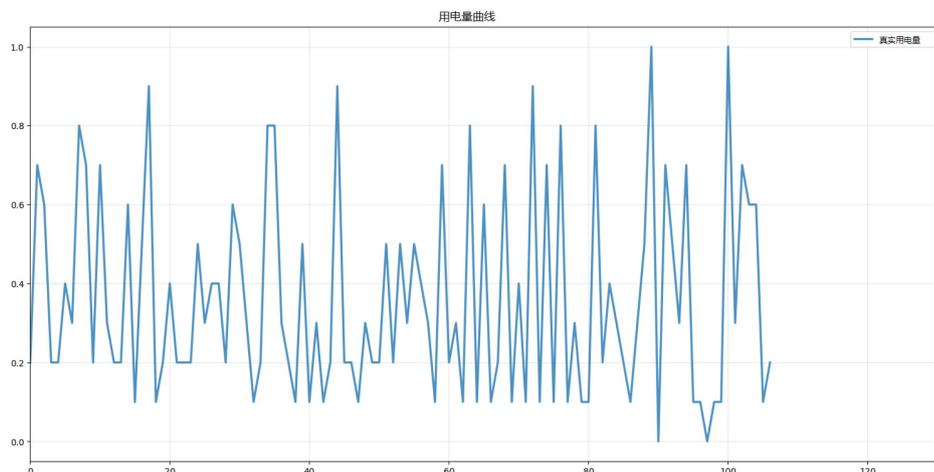
### 4.1. 数据集

本文使用 2 个数据集进行实验。

1) 源于 Kaggle 的摩洛哥得土安的用电量数据集[26] (Data\_kaggle)

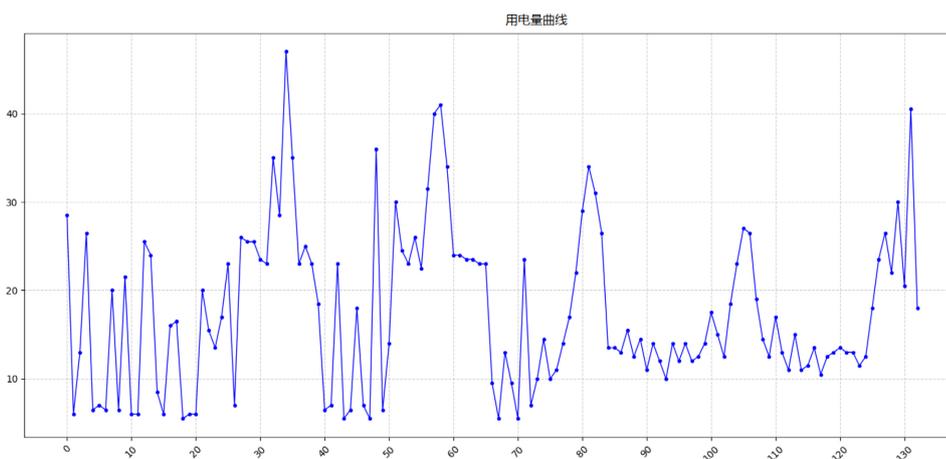
(<https://www.kaggle.com/datasets/fedesoriano/electric-power-consumption>)

该数据集包含 52,416 条观测记录, 间隔为 10 分钟, 每条记录具有 9 个特征: 时间戳(10 分钟窗口)、环境温度、相对湿度、风速和一般扩散流——该术语描述低温流体排放(<0.2°C 至 100°C), 通常与硫化物丘和细菌垫等地质构造相关。该数据集还包括三个指定区域(区域 1~3)的用电量指标, 而环境变量(温度、湿度、风速、扩散流)作为特征向量  $f$  中的辅助属性。其中区域 2 的用电量趋势如图 2 所示。其中, X 轴表示收集的数据点数量(按时间顺序, 该数据集采用 10 分钟窗口), Y 轴表示归一化的用电量。



**Figure 2.** Electricity consumption data in Zone 2  
**图 2.** 区域 2 的用电量数据

2) 第二个用电量数据集(Data\_Southern)来自中国南方电网运营辖区的 5 个地区(2025 年 12 月 1 日~5 月 31 日)(<https://pan.baidu.com/s/1F7diPhFfdCIZG01tmRIOeg?pwd=r2vy>)。用电量数据来自南方电网的 5 个地区, 每个地区有小时级的用电量记录。不同地区的用户结构(包括工业、居民等)不同, 导致用电量分布特征存在显著差异。图 3 显示了中国南方电网某地区的用电量数据, 其中, X 轴表示收集的数据点数量, Y 轴表示归一化的用电量数据。



**Figure 3.** Consumption data in an area of the China Southern Power Grid  
**图 3.** 中国南方电网某地区的用电量数据

如图 3 所示, 用电量数据呈现近似周期性和峰值变化。这 2 个数据集均可用于用电量预测任务。在实验中, 第一个数据集包含 3 个不同地区的用电量数据, 第二个数据集包含 5 个不同地区的用电量数据。我们将这些用电量数据作为不同客户端的数据, 创建联邦学习环境, 以验证本文提出的 FDGN 算法的有效性。

## 4.2. 基准模型

本文采用联邦学习下的一组基准模型, 进行评估和比较。

1) 第一组将常见的用电量预测方法与典型的联邦学习框架(如平均联邦学习 AvgFL 和基于相关性的主动客户端选择策略 FedCor)相结合, 形成诸如 AvgFL-SVR、AvgFL-BiLSTM、AvgFL-多尺度注意力模型(AvgFL-MSattention)、FedCor-SVR、FedCor-BiLSTM 和 FedCor-多尺度注意力模型(FedCor-MSattention)等模型。其中, SVR 指支持向量回归模型, BiLSTM 指双向 LSTM 模型, 多尺度注意力模型指多尺度注意力模型。

2) CNN-LSTM FED [27]: 这是一种基于数据增强的预测模型。在该研究中, 使用生成对抗网络(GANs)生成用电量数据, 通过数据增强缓解客户端数据的过拟合问题。

3) 自适应堆叠 LSTM [28]: 该用电量预测框架融合了自适应学习、联邦学习和边缘计算概念。

4) SparseMoE [29]: 混合专家(MoE)架构的专家网络通过基于 Transformer 的深度学习模型 Metaformer 实现。它利用指数移动平均操作和池化算子进行预测。

### 4.3. FDGN 的实验结果分析

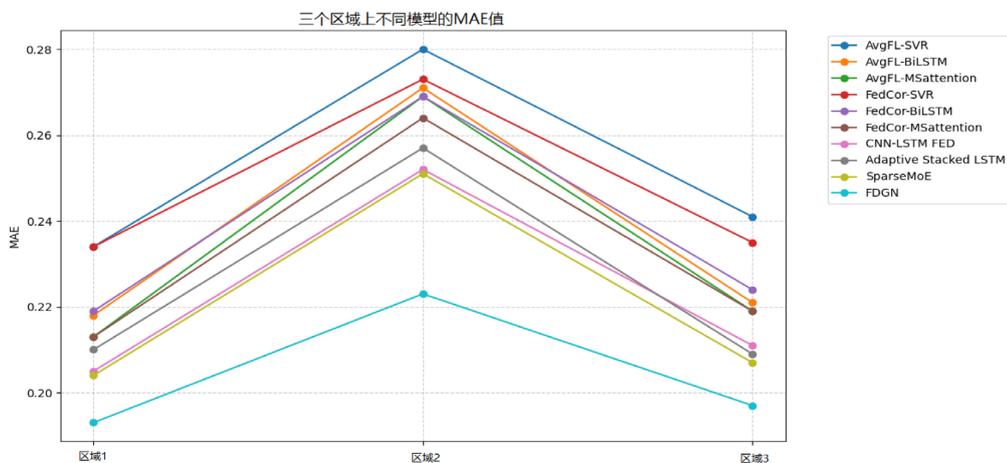
#### 4.3.1. 在 Data\_kaggle 上的验证

对于用电量预测, 我们首先使用得土安用电量数据(Data\_kaggle)验证 FDGN 的有效性。表 1 和图 4 记录了预测结果, 展示了用电量预测情况。

**Table 1.** Test MSE on Data\_kaggle dataset

**表 1.** Data\_kaggle 数据集上的测试 MSE

模型	区域 1	区域 2	区域 3
AvgFL-SVR	0.179	0.257	0.211
AvgFL-BiLSTM	0.169	0.238	0.197
AvgFL-MSattention	0.167	0.232	0.194
FedCor-SVR	0.164	0.244	0.207
FedCor-BiLSTM	0.152	0.222	0.192
FedCor-MSattention	0.153	0.225	0.189
CNN-LSTM FED	0.135	0.205	0.183
Adaptive Stacked LSTM	0.134	0.213	0.179
SparseMoE	0.127	0.206	0.181
<b>FDGN</b>	<b>0.115</b>	<b>0.193</b>	<b>0.173</b>



**Figure 4.** Test MAE on Data\_kaggle dataset

**图 4.** Data\_kaggle 数据集中的测试 MAE

表 1 和图 4 结果证实, 在 Data\_kaggle 用电量数据集上, 所提出的 FDGN 在所有区域均实现了最低的 MSE 和 MAE。FDGN 优于 FedCor 和 AvgFL 等传统联邦学习方法。此外, FDGN 的性能优于 CNN-LSTM FED 等基于数据增强的 FL 方法和自适应堆叠 LSTM 等自适应 FL 方法。值得注意的是, FDGN 在高波动期间保持了稳健的性能, 表明其在建模基准用电量和用电量峰值方面的能力有所增强。区域 2 的预测用电量和原始用电量如图 5 所示。

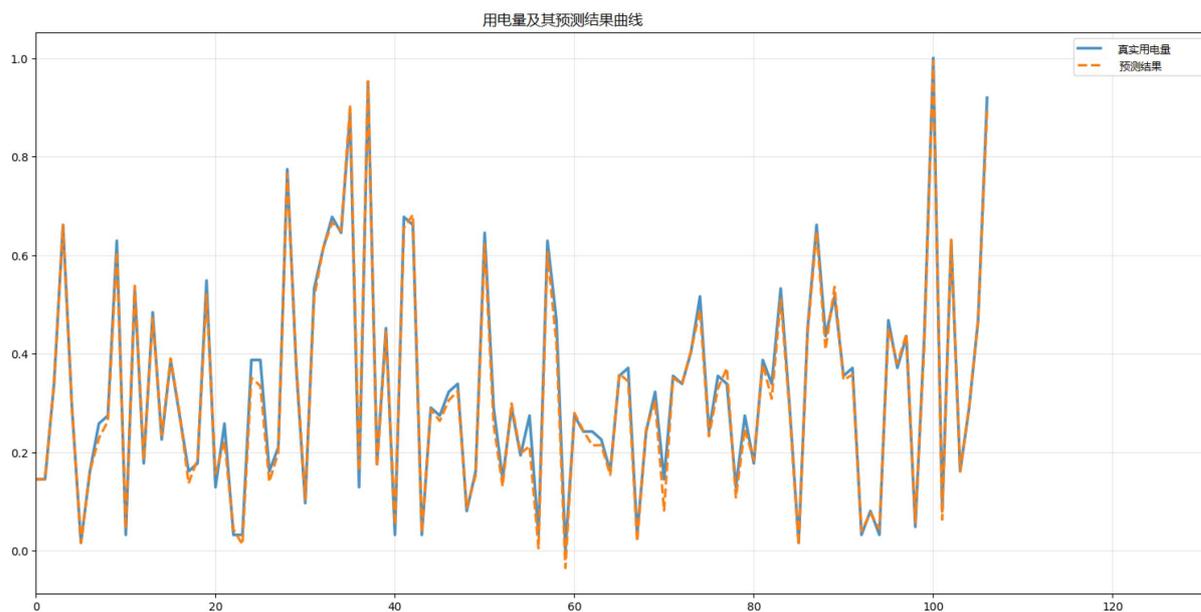


Figure 5. The forecasted electricity consumption curve in Zone 2

图 5. 区域 2 的预测用电量曲线

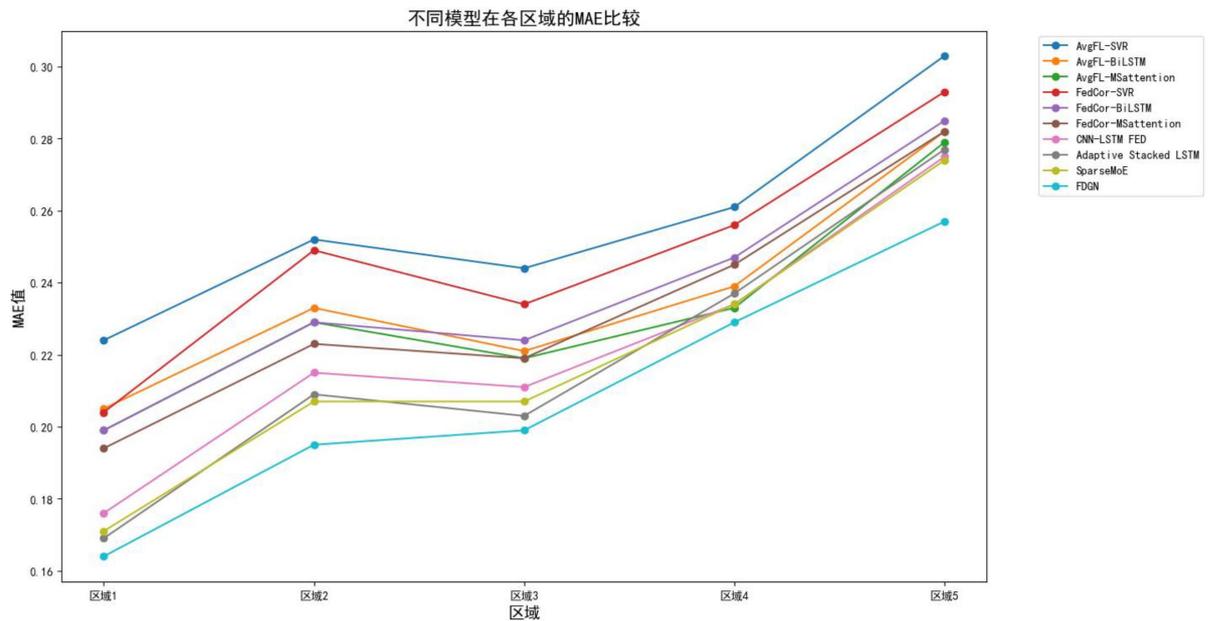
#### 4.3.2. 在 Data\_Southern 上的验证

在以下实验中, 我们展示了 FDGN 在中国南方电网数据集上的用电量预测性能。表 2 显示了 FDGN 使用中国南方电网数据集的用电量数据的预测性能。表 2 中以 MSE 表示的结果和测试 MAE 如图 6 所示。

Table 2. The global MSE of the FDGN model on the China Southern Power Grid dataset

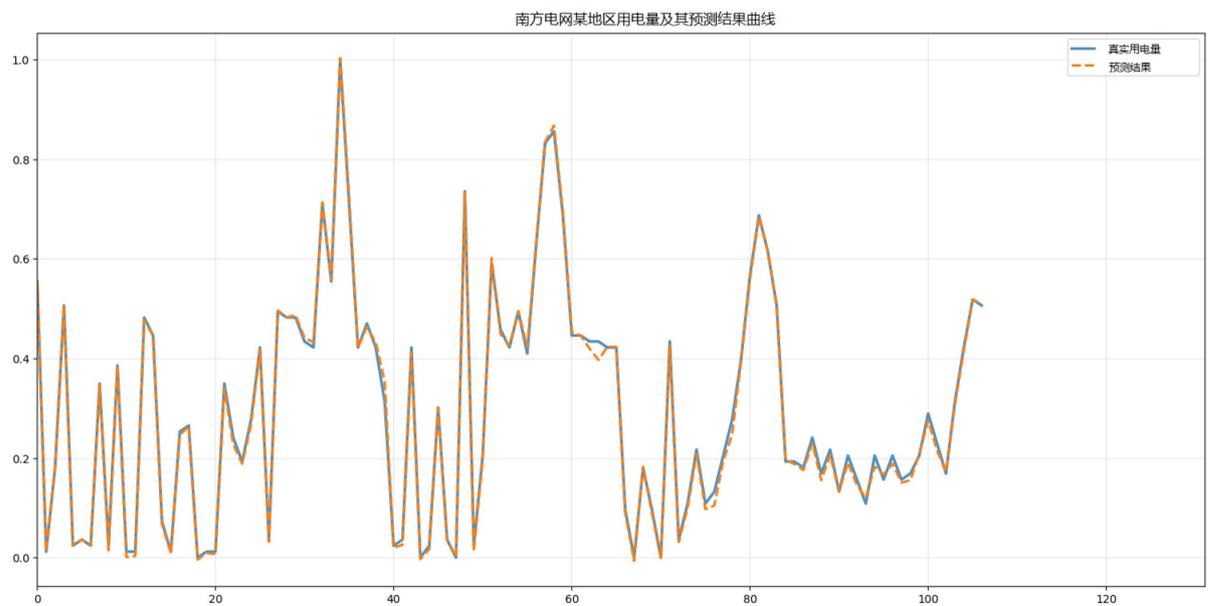
表 2. FDGN 模型在中国南方电网数据集上的全局 MSE

模型	区域 1	区域 2	区域 3	区域 4	区域 5
AvgFL-SVR	0.171	0.193	0.182	0.188	0.229
AvgFL-BiLSTM	0.159	0.182	0.171	0.183	0.217
AvgFL-MSattention	0.153	0.179	0.167	0.180	0.212
FedCor-SVR	0.151	0.188	0.175	0.189	0.223
FedCor-BiLSTM	0.143	0.179	0.169	0.181	0.211
FedCor-MSattention	0.141	0.173	0.165	0.179	0.207
CNN-LSTM FED	0.136	0.165	0.160	0.169	0.195
Adaptive Stacked LSTM	0.133	0.159	0.153	0.173	0.197
SparseMoE	0.129	0.156	0.156	0.172	0.194
<b>FDGN</b>	<b>0.122</b>	<b>0.155</b>	<b>0.144</b>	<b>0.164</b>	<b>0.191</b>



**Figure 6.** Test MAE of electricity consumption in the China Southern Power Grid dataset  
**图 6.** 中国南方电网数据集的测试 MAE

如表 2 和图 6 所示, FDGN 在所有区域均实现了最低的 MSE 和 MAE。这一性能突显了与传统用电量预测框架相比, FDGN 在数据拟合和预测精度方面的增强能力。Data\_Southern 区域 1 的工业预测用电量 and 原始用电量如图 7 所示。



**Figure 7.** The forecasted electricity consumption curve in Area 1 of Data\_Southern  
**图 7.** Data\_Southern 区域 1 的预测用电量曲线

接下来, 我们详细描述实验, 分析全局模型如何影响本地模型, 从而验证全局模型的聚合对本地数据的学习有积极促进作用。实验结果如下:

**Table 3.** The influence of the global model for clients (Test MSE) on Data\_Southern  
**表 3.** 全局模型对客户端的影响(Data\_Southern 上的测试 MSE)

模型	FDGN	FDGN-local
区域 1	0.122	0.147
区域 2	0.155	0.178
区域 3	0.144	0.171
区域 4	0.164	0.182
区域 5	0.191	0.221

如表 3 所示, FDGN-local 指不进行全局聚合的 FDGN 本地模型。FDGN 优于 FDGN-local, 这验证了全局模型的聚合对本地数据的学习有积极促进作用。

#### 4.4. FDGN 的消融实验

为验证所提出的 FDGN 框架的有效性, 我们针对其核心组件进行消融研究: 双分支图、用于建模峰值变化的高斯图注意力自编码器和动态融合聚合方法。评估了 4 个消融变体: 1) FDGN-temporal (移除双分支图中的数值结构图); 2) FDGN-AE (用 AE 替换高斯图注意力自编码器); 3) FDGN-Avg (用标准 AvgFL 替换动态融合聚合方法); 4) FDGN-AttAE (用传统注意力机制替代分层编码)。这些变体在两个用电量预测数据集上与完整的 FDGN 模型进行基准比较。实验结果如下。

**Table 4.** The ablation MSE of the FDGN on Tetouan-Electricity-Consumption dataset  
**表 4.** FDGN 在得土安用电量数据集上的消融 MSE

模型	区域 1	区域 2	区域 3
FDGN-AttAE	0.117	0.197	0.180
FDGN-Avg	0.123	0.201	0.179
FDGN-AE	0.121	0.197	0.179
FDGN-temporal	0.122	0.202	0.180
<b>FDGN</b>	<b>0.115</b>	<b>0.193</b>	<b>0.173</b>

**Table 5.** The ablation MSE of the FDGN on China Southern Power Grid dataset  
**表 5.** FDGN 在中国南方电网数据集上的消融 MSE

模型	区域 1	区域 2	区域 3	区域 4	区域 5
FDGN-AttAE	0.130	0.161	0.152	0.169	0.194
FDGN-Avg	0.129	0.161	0.153	0.167	0.192
FDGN-AE	0.127	0.159	0.151	0.166	0.191
FDGN-temporal	0.129	0.162	0.150	0.169	0.194
<b>FDGN</b>	<b>0.122</b>	<b>0.155</b>	<b>0.144</b>	<b>0.164</b>	<b>0.191</b>

如表 4 和表 5 所示, 实验量化了每个架构对整体性能的单贡献。所设计的组件提高了用电量预测的性能。

## 5. 结论

随着智能电网技术的快速发展, 用电量需求预测已成为支撑电网智能化运营的核心技术。尽管现有时间序列预测模型已取得一定进展, 但数据孤岛引发的局部过拟合问题与数据空间离散化带来的隐私保护挑战, 仍制约了传统方法的实际应用。联邦学习虽为协同建模与隐私保护提供了新思路, 但其在用电

量预测中仍面临局部数据稀缺导致的过拟合风险, 以及传统聚合机制造成的特征稀释问题。针对上述挑战, 本文提出的联邦双分支图神经网络(FDGN)通过自适应参数聚合机制, 创新性地设计了时序特征与峰值变化特征的双分支建模框架。实验结果表明, FDGN 在真实电网数据集上优于基线模型, 提升了预测精度与泛化能力。未来研究将进一步探索 FDGN 在跨区域电网协同调度中的动态优化潜力, 为智能电网的智能化升级提供更可靠的技术支撑。

## 基金项目

本工作得到了中国南方电网有限责任公司科技项目的支持(项目编号: 070000KC24110002)。

## 参考文献

- [1] González, A.M., Roque, A.S. and García-González, J. (2005) Modeling and Forecasting Electricity Prices with Input/output Hidden Markov Models. *IEEE Transactions on Power Systems*, **20**, 13-24. <https://doi.org/10.1109/tpwrs.2004.840412>
- [2] Yildiz, B., Bilbao, J.I. and Sproul, A.B. (2017) A Review and Analysis of Regression and Machine Learning Models on Commercial Building Electricity Load Forecasting. *Renewable and Sustainable Energy Reviews*, **73**, 1104-1122. <https://doi.org/10.1016/j.rser.2017.02.023>
- [3] Bhattarai, B.P., Paudyal, S., Luo, Y., Mohanpurkar, M., Cheung, K., Tonkoski, R., et al. (2019) Big Data Analytics in Smart Grids: State-of-the-Art, Challenges, Opportunities, and Future Directions. *IET Smart Grid*, **2**, 141-154. <https://doi.org/10.1049/iet-stg.2018.0261>
- [4] Hu, J., Harmsen, R., Crijns-Graus, W., Worrell, E. and van den Broek, M. (2018) Identifying Barriers to Large-Scale Integration of Variable Renewable Electricity into the Electricity Market: A Literature Review of Market Design. *Renewable and Sustainable Energy Reviews*, **81**, 2181-2195. <https://doi.org/10.1016/j.rser.2017.06.028>
- [5] Kairouz, P., McMahan, H.B., Avent, B., et al. (2021) Advances and Open Problems in Federated Learning. *Foundations and Trends® in Machine Learning*, **14**, 1-210.
- [6] Bousbiat, H., Bousseidj, R., Himeur, Y., et al. (2023) Crossing Roads of Federated Learning and Smart Grids: Overview, Challenges, and Perspectives. arXiv:2304.08602.
- [7] Silva, F.A.R., Orang, O., Javier Erazo-Costa, F., Silva, P.C.L., Barros, P.H., Ferreira, R.P.M., et al. (2025) Time Series Classification Using Federated Convolutional Neural Networks and Image-Based Representations. *IEEE Access*, **13**, 56180-56194. <https://doi.org/10.1109/access.2025.3554097>
- [8] Cheng, X., Li, C. and Liu, X. (2022) A Review of Federated Learning in Energy Systems. 2022 *IEEE/IAS Industrial and Commercial Power System System Asia (I&CPS Asia)*, Shanghai, 8-11 July 2022, 2089-2095. <https://doi.org/10.1109/icpsasia55496.2022.9949863>
- [9] Yang, Q., Liu, Y., Chen, T. and Tong, Y. (2019) Federated Machine Learning: Concept and Applications. *ACM Transactions on Intelligent Systems and Technology*, **10**, 1-19. <https://doi.org/10.1145/3298981>
- [10] Mohd Nizam, M.A., Sulaiman, S.A. and Ramli, N.A. (2024) Predictive Model for Electricity Consumption in Malaysia Using Support Vector Regression. In: Md. Zain, Z., Sulaiman, N., Mustafa, M., Shakib, M.N. and A. Jabbar, W., Eds., *Lecture Notes in Electrical Engineering*, Springer, 149-159. [https://doi.org/10.1007/978-981-97-3851-9\\_14](https://doi.org/10.1007/978-981-97-3851-9_14)
- [11] Chen, H. and Lee, C. (2019) Electricity Consumption Prediction for Buildings Using Multiple Adaptive Network-Based Fuzzy Inference System Models and Gray Relational Analysis. *Energy Reports*, **5**, 1509-1524. <https://doi.org/10.1016/j.egyr.2019.10.009>
- [12] Gomez, W., Wang, F. and Amogne, Z.E. (2023) Electricity Load and Price Forecasting Using a Hybrid Method Based Bidirectional Long Short-Term Memory with Attention Mechanism Model. *International Journal of Energy Research*, **2023**, Article ID: 3815063. <https://doi.org/10.1155/2023/3815063>
- [13] Wang, W., Shimakawa, H., Jie, B., Sato, M. and Kumada, A. (2025) BE-LSTM: An LSTM-Based Framework for Feature Selection and Building Electricity Consumption Prediction on Small Datasets. *Journal of Building Engineering*, **102**, Article 111910. <https://doi.org/10.1016/j.jobe.2025.111910>
- [14] Imani, M. (2021) Electrical Load-Temperature CNN for Residential Load Forecasting. *Energy*, **227**, Article 120480. <https://doi.org/10.1016/j.energy.2021.120480>
- [15] Wang, S., Wang, X., Wang, S. and Wang, D. (2019) Bi-Directional Long Short-Term Memory Method Based on Attention Mechanism and Rolling Update for Short-Term Load Forecasting. *International Journal of Electrical Power &*

- 
- Energy Systems*, **109**, 470-479. <https://doi.org/10.1016/j.ijepes.2019.02.022>
- [16] Lotfipoor, A., Patidar, S. and Jenkins, D.P. (2024) Deep Neural Network with Empirical Mode Decomposition and Bayesian Optimisation for Residential Load Forecasting. *Expert Systems with Applications*, **237**, Article 121355. <https://doi.org/10.1016/j.eswa.2023.121355>
- [17] Gashler, M.S. and Ashmore, S.C. (2016) Modeling Time Series Data with Deep Fourier Neural Networks. *Neurocomputing*, **188**, 3-11. <https://doi.org/10.1016/j.neucom.2015.01.108>
- [18] Liu, Y., Guan, L., Hou, C., Han, H., Liu, Z., Sun, Y., *et al.* (2019) Wind Power Short-Term Prediction Based on LSTM and Discrete Wavelet Transform. *Applied Sciences*, **9**, Article 1108. <https://doi.org/10.3390/app9061108>
- [19] Liang, F., Zhang, Z., Lu, H., *et al.* (2024) Communication-Efficient Large-Scale Distributed Deep Learning: A Comprehensive Survey. arXiv:2404.06114.
- [20] Huang, W., Wang, D., Ouyang, X., Wan, J., Liu, J. and Li, T. (2024) Multimodal Federated Learning: Concept, Methods, Applications and Future Directions. *Information Fusion*, **112**, Article 102576. <https://doi.org/10.1016/j.inffus.2024.102576>
- [21] Jia, N., Qu, Z., Ye, B., Wang, Y., Hu, S. and Guo, S. (2025) A Comprehensive Survey on Communication-Efficient Federated Learning in Mobile Edge Environments. *IEEE Communications Surveys & Tutorials*, **1**. <https://doi.org/10.1109/comst.2025.3535957>
- [22] Sun, T., Li, D. and Wang, B. (2022) Decentralized Federated Averaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **45**, 4289-4301. <https://doi.org/10.1109/tpami.2022.3196503>
- [23] Geyer, R.C., Klein, T. and Nabi, M. (2017) Differentially Private Federated Learning: A Client Level Perspective. arXiv:1712.07557.
- [24] Venkataramanan, V., Kaza, S. and Annaswamy, A.M. (2022) DER Forecast Using Privacy-Preserving Federated Learning. *IEEE Internet of Things Journal*, **10**, 2046-2055. <https://doi.org/10.1109/jiot.2022.3157299>
- [25] Ahmadi, A., Talaie, M., Sadipour, M., Amani, A.M. and Jalili, M. (2022) Deep Federated Learning-Based Privacy-Preserving Wind Power Forecasting. *IEEE Access*, **11**, 39521-39530. <https://doi.org/10.1109/access.2022.3232475>
- [26] <https://www.kaggle.com/datasets/fedesoriano/electric-power-consumption>
- [27] De Moraes Sarmiento, E.M., Ribeiro, I.F., Marciano, P.R.N., *et al.* (2024) Forecasting Energy Power Consumption Using Federated Learning in Edge Computing Devices. *Internet of Things*, **25**, Article 101050. <https://doi.org/10.1016/j.iot.2023.101050>
- [28] Abdulla, N., Demirci, M. and Ozdemir, S. (2024) Smart Meter-Based Energy Consumption Forecasting for Smart Cities Using Adaptive Federated Learning. *Sustainable Energy, Grids and Networks*, **38**, Article 101342. <https://doi.org/10.1016/j.segan.2024.101342>
- [29] Wang, R., Bai, L., Rayhana, R. and Liu, Z. (2024) Personalized Federated Learning for Buildings Energy Consumption Forecasting. *Energy and Buildings*, **323**, Article 114762. <https://doi.org/10.1016/j.enbuild.2024.114762>