DOI: 10.19666/j.rlfd.202602018

基于相似日聚类-TCN-Informer 组合模型的短期电力负荷预测研究

吴　涛[1]，洪　滂[2]，刘端阳[3]，曹　强[2]，王　亚[2]，胡　胤[2]，王　昊[2]，
余　威[2]，杨永军[1]，罗　睿[1]，陈　武[4]
（1. 西安热工研究院有限公司，陕西　西安　710054；
2. 华能湖北能源销售有限责任公司，湖北　武汉　430077；
3. 华北电力大学能源动力与机械工程学院，北京　102206；
4. 华能江西能源销售有限责任公司，江西　南昌　330000）

［摘　　要］【目的】针对现有短期电力负荷预测模型难以同时学习用电模式差异与长程依赖、致使高波动场景预测精度受限的问题，本文提出一种“相似日聚类-TCN-Informer”组合模型。【方法】首先，利用模糊 *C* 均值算法依据历史负荷与日历特征将样本划分为工作日、周末和节假日 3 类相似日，以削弱跨模式噪声；其次，对每一簇数据先由时序卷积神经网络（temporal convolutional network，TCN）提取局部特征，再引入概率稀疏注意力 Informer 捕获长程依赖并生成簇级预测；随后融合气象与电价等外生变量，并采用贝叶斯优化进行全局超参数搜索。【结果】基于澳大利亚能源市场 AEMO-NSW 30 min 分辨率数据的 24 h 预测实验表明，该模型的均方根误差、平均绝对误差分别较 LSTM 降低 16.59%和 14.66%，较 TimesNet 降低 12.33%和 11.61%，平均绝对百分比误差仅为 3.02%。【结论】结果验证了所提方法在精度与鲁棒性上的显著优势，可为电网调度与现货交易提供可靠决策支持。

［关　键　词］短期电力负荷预测；相似日聚类；时序卷积神经网络；Informer；长程依赖

［引用本文格式］吴涛，洪滂，刘端阳，等. 基于相似日聚类-TCN-Informer 组合模型的短期电力负荷预测研究[J]. 热力发电，2026, 55(8): 45.　Wu T, Hong M, Liu D Y, et al. Short-term power load forecasting via a similar-day clustering-TCN-Informer hybrid model[J]. Thermal Power Generation, 2026, 55(8): 45.

Short-term power load forecasting via a similar-day clustering-TCN-Informer hybrid model

Wu Tao[1], Hong Mang[2], Liu Duanyang[3], Cao Qiang[2], Wang Ya[2], Hu Yin[2], Wang Hao[2], Yu Wei[2], Yang Yongjun[1], Luo Rui[1], Chen Wu[4]
(1. Xi'an Thermal Power Research Institute Co., Ltd., Xi'an 710054, China;
2. Huaneng Hubei Energy Sales Co., Ltd., Wuhan 430077, China;
3. School of Energy Power and Mechanical Engineering, North China Electric Power University, Beijing 102206, China;
4. Huaneng Jiangxi Energy Sales Co., Ltd., Nanchang 330000, China)

Abstract: [Objective] Existing short-term power load forecasting models struggle to simultaneously learn the inherent differences in electricity consumption patterns across diverse scenarios and effectively capture long-range temporal dependencies, so their forecasting accuracy is significantly restricted in high-volatility situations such as extreme weather events, peak load periods, or sudden changes in user behavior. To address this limitation, this paper proposes a novel hybrid forecasting framework named “similar-day clustering-temporal convolutional network (TCN)-Informer”. [Methods] First, the fuzzy *C*-means (FCM) clustering algorithm is employed to categorize historical load samples into three distinct similar-day clusters (working days, weekends, and holidays) based on

收 稿 日 期：2026-02-08　修 回 日 期：2026-04-03　接 受 日 期：2026-04-19
基 金 项 目：中国华能集团有限公司总部科技项目（HNKJ25-H15）
Supported by: Science and Technology Project of China Huaneng Group Co., Ltd. (HNKJ25-H15)
第一作者简介：吴涛（1993），男，硕士，工程师，主要研究方向为电力生产信息化智能化，wutaox@tpri.com.cn。

historical load data and calendar features (e.g., weekday/weekend flags, holiday labels). This preprocessing step effectively mitigates cross-pattern noise interference caused by heterogeneous consumption behaviors, laying a solid foundation for subsequent feature extraction. Second, for each clustered dataset, a causal dilated TCN is utilized to extract local temporal features and capture short- to medium-term load fluctuations, as its dilated convolution structure enables efficient expansion of the receptive field without increasing computational complexity while avoiding future information leakage. Subsequently, the Informer model equipped with a probabilistic sparse attention mechanism is introduced to capture long-range dependencies in load time series, which significantly reduces the computational burden compared to traditional attention mechanisms and enhances the model's ability to learn long-term temporal correlations. Moreover, exogenous variables closely related to power load, including meteorological factors (temperature, humidity, solar irradiation) and real-time electricity prices, are integrated into the forecasting framework to improve the comprehensiveness of input information. To optimize the model's hyperparameters, Bayesian optimization is adopted to efficiently search the global hyperparameter space, outperforming traditional grid search in both search efficiency and optimization performance. [Results] Experimental validations for 24-hour load forecasting are conducted using 30-minute resolution data from the AEMO-NSW power market, covering a one-year period with over 17 000 samples. The results demonstrate that the proposed model achieves remarkable performance: its root mean square error (RMSE) and mean absolute error (MAE) reduce by 16.59% and 14.65% compared with those of the classic LSTM model, and by 12.33% and 11.60% in contrast to the advanced TimesNet model, with the mean absolute percentage error (MAPE) as low as 3.02%. Particularly in high-volatility scenarios, the model maintains superior stability and accuracy. [Conclusion] These findings fully verify the prominent advantages of the proposed method in forecasting precision and robustness, which can provide reliable decision-making support for power grid dispatching, electricity spot trading, and energy resource optimization in practical power systems.

Key words: short-term power load forecasting; similar-day clustering; temporal convolutional network; Informer; long-range dependency

近年来，“双碳”战略目标与可再生能源大规模并网正深刻改变电力系统的运行形态。新能源出力的不确定性放大了负荷波动幅度，对电网调度安全与经济性提出更高要求[1]。短期电力负荷预测（short-term load forecasting，STLF）作为电力市场交易、机组组合和分时电价制定的基础，其精度直接关系到源−网−荷各环节的协同优化[2]。因此，构建能够兼顾负荷多尺度时空特征、适应复杂用电场景的高精度 STLF 模型，已成为当前电力系统与人工智能交叉研究的热点之一。

近年来，深度学习因其对非线性时序特征的提取能力，成为短期负荷预测研究的主要方法。早期工作多聚焦于经典长短时记忆（long short-term memory，LSTM）神经网络结构中融入注意力机制，以强化关键特征的动态权重分配。Lin 等人提出的双阶段特征−时间注意力−LSTM 即显著提高了区间概率预测精度[3]。与此同时，分解−重构思想被证明可有效削弱原始负荷序列的混叠特性，Yang 等人通过双变量经验模态分解结合多变量多尺度排列熵（bivariate empirical mode decomposition-multivariate multiscale permutation entropy，BEMD-MMPE）组合 LSTM，并借助贝叶斯优化实现参数自适应，获得了更稳健的区间预测[4]。针对深度模型训练难、解释性弱的问题，Eren 等人综述了 LSTM 神经网络及其衍生架构在在线、自适应及不确定性度量方面的进展，指出集成与迁移学习将成为下一阶段的主要方向[5]。为降低模型对超长序列的依赖，Chen 等人利用时间序列聚类与早期分类策略，将 LightGBM 与局部模式识别结合，兼顾了实时性与精度[6]。进一步地，Fazlipour 等人设计了堆叠式 LSTM-自编码器并在预训练阶段引入多级注意力，实现了对超长依赖的捕获及随机权值初始化难题的缓解[7]。这些早期探索为后续多模态、多任务负荷预测模型的演进奠定了方法学基础。

尽管基于循环神经网络（recurrent neural network，RNN）的 LSTM 神经网络能够捕捉负荷序列的时序依赖，但其串行迭代特性在并行计算及长距离梯度传播方面仍面临瓶颈。受一维卷积与膨胀因子相结合的启发，时序卷积神经网络（temporal convolutional network，TCN）近年被引入负荷预测领域。TCN 依赖因果卷积确保输出不泄露未来信息，通过残差块与膨胀卷积在保持感受野指数级扩张的同时有效缓解梯度消散问题，兼具并行推理与长序列建模优势。Zulfiqar 等人将 TCN 应用于住宅级 STLF，证明其在提取多尺度时序特征方面优于传统 LSTM，实现了精度提升[8]。赵洋等进一步在 TCN 中植入自注意力以动态调节不同时间尺度的卷积核权重，显著降低了负荷峰值误差[9]。Türkoğlu

等人的综述指出，TCN 在在线增量学习与异常检测场景具有天然适配性，已被视作继 RNN 之后的另一条主流技术路线[10]。结合特征工程与优化算法亦可放大 TCN 优势，Zuo 等人通过 TCN 作为局部模式提取器，实现了精度、稳定性与收敛速度的多指标最优[11]。因此，引入 TCN 为核心的卷积式并行结构，可有效提高负荷预测的精度。

虽然基于循环神经网络与卷积神经网络的架构可提取时间序列多尺度依赖，然而这些网络往往侧重单向或局部，难以在更长的历史窗口内保持信息流动。针对长序列预测瓶颈，基于注意力机制的 Transformer 及其后续 Informer 以 Self-Attention 与长依赖建模能力被引入电力时序预测，能够在保持线性复杂度的同时实现对超长序列的高效编码[12]。任建吉等在 CNN-BiLSTM 框架中嵌入注意力机制，显著强化了负荷与气象等多源特征的耦合关系，从而提升了超短期电力负荷预测精度[13]。Qu 等人提出的 Forwardformer 将多尺度前向自注意力融入 Transformer 架构，实现了对日前负荷的高效长序列建模，为短期负荷预测提供了新的 Transformer 范式[14]。范杏蕊利用改进的 Autoformer 在网络内部引入渐进式时序分解，并以 Nystrom 自注意力高效逼近标准注意力，使模型在保持较低时间复杂度的同时精准捕捉负荷的复杂趋势与不确定性，提升短期电力负荷预测精度[15]。周师琦等基于 Autoformer 通过多尺度解耦和概率化贝叶斯优化，能够在捕获高波动负荷子序列的同时动态量化预测不确定性，提升短期电力负荷预测的精度与鲁棒性[16]。

同时，最近的研究着重于通过不同模块的融合，以兼顾局部特征提取与长程依赖建模的双重需求，提高负荷预测的准确性。部分工作探索了 TCN 与 Transformer 的串行或并行叠加结构。为提升校园热电联产系统的蒸汽负荷预测精度，有学者构建了混合 Transformer-TCN-GRU 模型，利用 TCN 提取多尺度局部模式，借助 Transformer 捕获全局依赖，在长达 8 年的实测数据上获得了 0.96 的确定系数[17]。针对短期电力负荷的强波动性，另有研究设计了一种双模块框架，先通过数据增强模块引入峰值与趋势信息，再以多层 TCN 与 Transformer 相结合的深度学习模块实现高精度预测，在澳大利亚多个公开数据集上验证了其有效性[18]。进一步地，面向含可再生能源的综合能源系统多元负荷预测，有研究提出了 TCN-FECAM-Informer 模型，利用频域增强通道注意力机制改进 TCN 的特征提取能力，并将处理后的特征序列输入 Informer 网络，实验结果表明该结构在多元负荷预测中显著优于传统循环神经网络[19]。上述研究表明，TCN 与 Transformer 类模型的有机融合能够在局部时序特征与全局依赖关系之间形成有效互补。然而，现有叠加模型大多将卷积与注意力模块简单串行拼接，未充分考虑负荷数据中工作日、周末与节假日等不同用电模式的异质性，且普遍缺乏针对高维特征的自适应筛选机制，难以在复杂场景下实现精度与鲁棒性的同步提升。

综上所述，现有短期负荷预测研究虽在相似日划分、TCN 卷积与 Transformer 长程依赖建模方面各有进展，但仍存在三大问题：1）多数方法仅针对单一负荷模式训练，难以同时刻画工作日、周末与节假日的用电差异；2）卷积与注意力尚未形成统一训练框架，局部细节与全局趋势的交互效率不足；3）长序列模型普遍侧重复杂度压缩，缺乏特征筛选与超参数自适应。为弥补上述不足，本文提出“相似日聚类-TCN-Informer”组合模型：首先，采用模糊 C 均值将历史样本软聚类为工作日、周末与节假日 3 类相似日，引入差异化先验；其次，在每一簇内以因果膨胀 TCN 提取局部高频特征，再借助 Informer 的 ProbSparse-Attention 捕获长程依赖，形成“分群–卷积–稀疏注意”三级协同机制；最后，通过贝叶斯优化完成超参数搜索，构建出兼具多模式适配与局部–全局协同能力的高精度短期负荷预测模型。

1 研究方法

1.1 SD-TCN-Informer 整体架构

本文提出了一种融合相似日选取（similar days selection）、TCN 与 Informer 模型的混合深度学习预测框架（SD-TCN-Informer），旨在解决电力负荷预测中存在的噪声干扰严重、局部特征难以提取以及长程依赖丢失等关键问题。模型架构如图 1 所示，模型数据流依次经过样本筛选、局部特征提取与全局长序预测 3 个核心处理单元，最终实现对电力负荷的预测。

首先，针对电力负荷数据受气象条件（温度、湿度）及日历效应（工作日、节假日）影响的特点，模型引入了相似日选取模块，通过计算预测日与历史样本在气象特征及日期类型上的相似度，筛选出与预测目标最为接近的历史负荷序列，构建高相关

性的输入数据集。

随后，经过筛选的相似日多变量序列被输入 TCN 模块进行深层特征挖掘。TCN 采用因果膨胀卷积结构，能在保证不泄露未来信息前提下，通过指数级扩大的感受野捕捉序列中的高频波动与局部短时依赖关系，为后续的长序列建模奠定基础。

最后，由 TCN 模块输出的高维特征向量被送入 Informer 预测网络，利用生成式推理机制，接收编码器输出的全局上下文信息，一步生成电力负荷预测值。

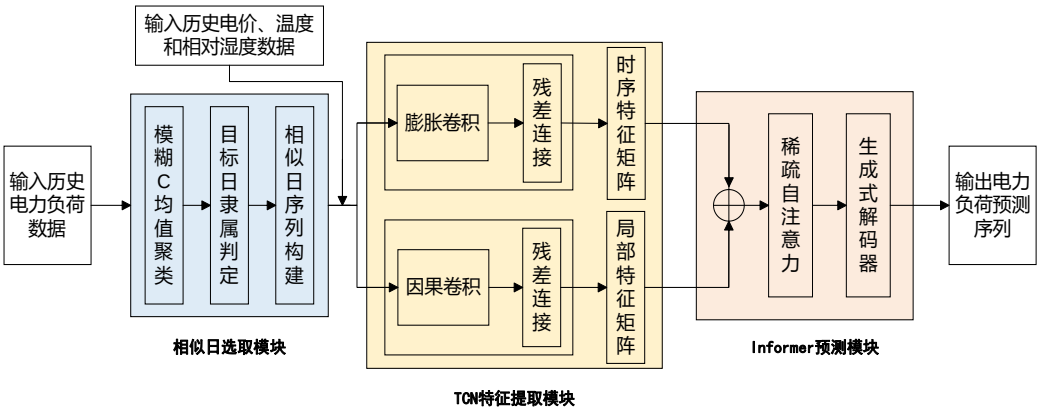


图 1 SD-TCN-Informer 模型架构

Fig.1 Architecture of the SD-TCN-Informer

1.2 基于模糊 *C* 均值的相似日聚类模块

电力负荷具有显著的周期性和日历效应，不同日期类型的负荷曲线形态差异巨大。直接使用全部历史数据进行训练往往会引入噪声，降低模型对特定日期的预测精度。为此，本文引入模糊 *C* 均值聚类算法（fuzzy C-means，FCM）对历史样本进行软聚类，构建相似日集合。FCM 允许一个数据点以不同程度的隶属度属于多个类别，更适合处理电力负荷数据中某些日期属性界限模糊的问题，其目标是通过迭代，最小化以下目标函数 J_m[20]：

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m} \mid \boldsymbol{x}_i - \boldsymbol{c}_j \mid^2 \tag{1}$$

式中：N 为样本总数；C 为聚类簇数，C=3，对应工作日、周末、节假日；m 为模糊加权指数，控制聚类的模糊程度；$\boldsymbol{x}_i$ 为第 i 个样本的特征向量，包含历史负荷、温度、湿度等特征；$\boldsymbol{c}_j$ 为第 j 个聚类中心的向量；u_{ij} 为第 i 个样本属于第 j 类的隶属度，满足 $\sum_{j=1}^{C} u_{ij} = 1$。

模型训练前，首先利用 FCM 算法计算历史样本的聚类中心 $\boldsymbol{c}_j$。对于待预测的目标日 $\boldsymbol{x}_{\text{target}}$，首先提取其特征向量，计算其与各聚类中心的欧氏距离，确定其所属的主导类别（即隶属度最大的类别）。随后，从历史数据库中筛选出所有属于该主导类别的样本，组成相似日集合。该集合将作为后续 TCN-Informer 模型的训练输入。

1.3 TCN 局部特征提取模块

在 SD-TCN-Informer 框架中，TCN 承担局部高频特征提取的任务，挖掘负荷序列中的短时突变和日内波动规律。TCN 采用了因果卷积（causal convolution），确保在 t 时刻的预测仅依赖于 t 时刻及以前的历史信息，严禁未来信息泄露。为了在不增加网络深度的情况下扩大感受野（receptive field），捕捉更长的历史信息，TCN 引入了膨胀卷积（dilated convolution）。

对于一维输入序列 $\boldsymbol{X}\in\mathrm{R}^{\mathrm{T}}$ 和卷积核 f:{0, …, k−1}→R，在时刻 t 的膨胀卷积操作 $F(t)$定义为[21]：

$$F(t) = (X *_d f)(t) = \sum_{i=0}^{k-1} f(i)\cdot x_{t-d\cdot i} \tag{2}$$

式中：d 为膨胀因子，通常随网络层数呈指数级增长；k 为卷积核大小；$*_d$ 表示膨胀卷积运算。这种设计使得 TCN 能够以对数级层数覆盖极长历史视窗，有效捕捉电力负荷数据中的多尺度时间特征。

为了解决深层网络的梯度消失问题，增强模型的稳定性，TCN 采用了残差块（residual block）结构。输入 x 与卷积层的输出 $F(x)$进行相加[22]：

$$o = \mathrm{Activation}(x + F(x)) \tag{3}$$

本文中的 TCN 层包含 2 个膨胀因子卷积层，每层后接权重归一化和 ReLU 激活函数，并引入 Dropout 层防止过拟合。经过 TCN 处理后，原始的一维负荷序列被映射为含局部上下文信息的高维特征向量，并作为后续 Informer 模块的输入。

1.4 Informer 长程依赖建模模块

本文在后端引入 Informer 模型，解决长序列时间预测（LSTF）难题。Informer 的核心创新包括概率稀疏自注意力机制以及生成式解码器（generative-style decoder）[23]。Zhou 等人在 Informer 中通过统计分析发现，自注意力分布通常呈长尾形态，仅有少数 Query-Key 对拥有显著权重[23]。Informer 引入 Kullback-Leibler（K-L）散度来度量注意力分布与均匀分布的差异，筛选出活跃查询（active queries）。概率稀疏自注意力定义为[23]：

$$\text{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\text{Softmax}\left(\frac{\bar{\boldsymbol{Q}}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V} \tag{4}$$

式中：$\bar{\boldsymbol{Q}}$ 为由稀疏度量筛选出的前 u 个主要查询向量（$u=c\ \ln L$）；d_k 为输入维度。该机制将时间复杂度从 $O(L^2)$显著降低至 $O(L\ln L)$，使得模型能够高效处理长达数周的电力负荷序列，精准捕捉周度、月度等长程依赖关系。

为了进一步压缩特征维度并突出关键信息，Informer 在编码器层之间引入了蒸馏操作。通过一维卷积（Conv1d）和最大池化（max pooling）层，逐步减半序列长度[23]：

$$X_{j+1}^{t}=\text{MaxPool}(\text{ELU}(\text{Conv1d}([X_j^t]_{\text{AB}}))) \tag{5}$$

该操作构建了金字塔式的特征提取结构，在保留重要长程特征的同时，大幅减少了内存占用。针对误差累积，Informer 采用生成式解码器，一次性输出所有预测时间步的结果。解码器的输入由 2 个部分组成：1）Start Token，从输入序列中截取的各种长度的历史片段；2）占位符，待预测部分的占位符。通过掩码多头注意力机制（masked multi-head attention），解码器利用编码器输出的特征图，直接通过全连接层一步生成未来 24 h（48 个点）的负荷预测值。

2 实验分析

2.1 数据与预处理

为验证所提模型在电力负荷预测任务中的有效性，本文选取澳大利亚能源市场运营商（Australian energy market operator，AEMO）新南威尔士州（NSW）地区的实际运行数据作为实验对象。数据集的时间跨度为 2018 年 1 月 1 日至 2020 年 12 月 31 日，采样间隔为 30 min。该数据集涵盖了长期的电力系统运行状况，包含约 52 608 个时间步的数据样本。

在变量选择上，考虑到电力系统受气象条件及市场经济因素的显著影响，本文选取了同一时间尺度下的区域温度、相对湿度以及电价作为多变量输入的外生协变量，以辅助模型捕捉负荷变化的复杂特征。各变量的时序分布概览如图 2 所示。

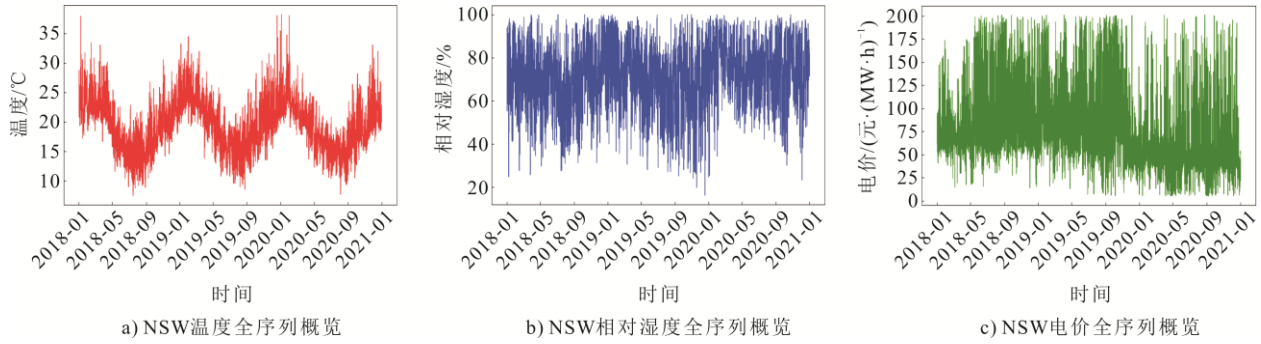


图 2　外生变量时序分布概览
Fig.2 Overview of the full exogenous variables series

由图 2 可见，作为外生变量的电价表现出极高的波动性和尖峰特征，这通常由电力市场的供需瞬时失衡引起；而温度和湿度则表现出较强的季节相关性。这些外生变量与负荷序列之间存在复杂的非线性耦合关系（如高温导致制冷负荷激增），这对预测模型的特征提取能力提出了更高的要求。因此，本文采用的 NSW 数据集是一个集多重季节性、非线性相关及高噪声于一体的典型时间序列，非常适于验证深度学习模型的预测性能。

2.1.1 *原始数据分析*

图 3 为 NSW 电力负荷全序列概览。由图 3 可见，电力负荷数据具有显著的周期性和非平稳性。序列呈现出明显的日度、周度和季节性波动模式：在日内尺度上，负荷通常在早晚出现双高峰，对应居民和工业用电的活跃时段；在季节尺度上，夏季和冬季由于空调及采暖需求，负荷水平往往高于春秋两季。

图 4 展示了 NSW 电力负荷全序列数据分布直方图。从统计特性来看，负荷分布近似于正态分布，但存在一定的偏度。这种分布特性反映了电力消费行为的统计规律，即大部分时间的负荷集中在均值附近（6 000~8 000 MW 区间），但也存在由于极端天气或特殊事件导致的极高或极低负荷情况。

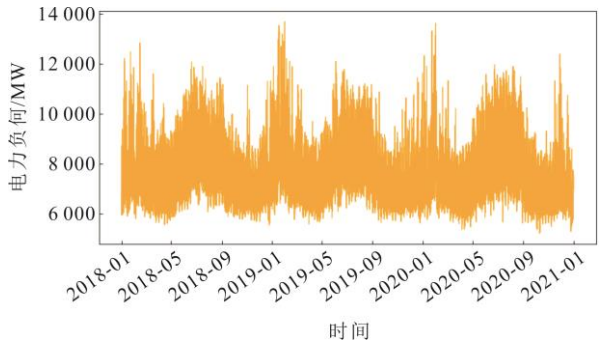


图 3 NSW 电力负荷全序列概览
Fig.3 Overview of the full NSW electrical load series

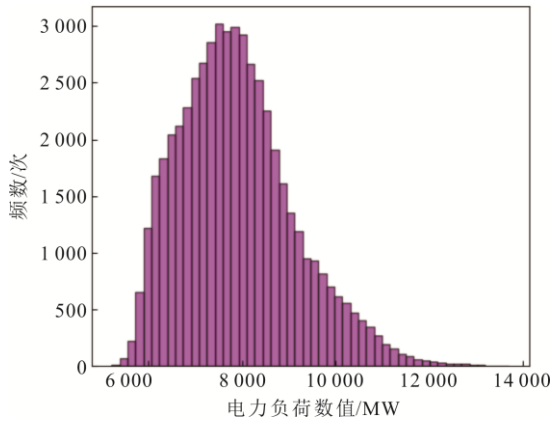


图 4 NSW 电力负荷全序列数据分布直方图
Fig.4 Histogram of the full series data distribution of NSW electrical load

2.1.2 数据预处理及划分

本文以历史电力负荷序列为输入，未来电力负荷为预测目标，输入序列步长为 96 步，输出序列步长为 48 步，即预测未来 24 h 的电力负荷值。为保证模型评估的客观性，本文将数据集严格按照时间顺序以 7:1:2 的比例划分为训练集、验证集和测试集。模型的训练与评估均基于训练集和验证集完成，而本文最终的性能评估及图表中展示的预测结果均是在测试集上得出的。

为了消除电力负荷与温度、湿度、电价等外生变量之间显著的量纲差异，并加速深度学习模型的训练收敛过程，本文引入了全局标准化处理模块。该模块采用标准差标准化方法，将所有输入特征映射至统一的数值区间。为了严格遵循时间序列分析中的因果约束并防止数据泄露，标准化所需统计量（均值与标准差）并非基于全量数据计算，而是严格限制在当前折（fold）的训练集范围内动态生成。

在验证策略上，本文采用时间序列交叉验证（TS-CV）框架。该方法通过扩展窗口机制，保持时间序列的自然顺序不被打乱：训练集的结束点随时间推移逐步向后滑动，而验证集始终取自训练窗口紧邻的未来时间段。在每一轮交叉验证中，标准化参数均重新计算并仅应用于对应的验证集和测试集。该方法不仅保证了评估过程的客观性，还有效模拟了实际电力系统中随着时间推移不断获取新数据进行滚动预测的真实场景。

2.2 评价指标

为了实现全面评估模型的预测性能，本文评估指标采用平均绝对误差（δ_{MAE}）、相关系数（C_{ORR}）、均方根误差（δ_{RMSE}）和平均绝对百分比误差（δ_{MAPE}），其计算公式为：

$$\delta_{\mathrm{MAE}}=\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i| \tag{6}$$

式中：n 为样本总数；y_i 为第 i 个样本的真实值；$\hat{y}_i$ 为 i 个样本的预测值。

$$C_{\mathrm{ORR}}=\frac{\sum_{i=1}^{n}(y_i-\bar{y})(\hat{y}_i-\bar{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2\sum_{i=1}^{n}(\hat{y}_i-\bar{\hat{y}})^2}} \tag{7}$$

式中：$\bar{y}$ 为真实值的平均数；$\bar{\hat{y}}$ 为预测值的平均值。

$$\delta_{\mathrm{RMSE}}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{y}_i)^2} \tag{8}$$

$$\delta_{\mathrm{MAPE}}=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right| \tag{9}$$

2.3 模型训练

本文提出的 SD-TCN-Informer 组合模型基于 PyTorch 深度学习框架搭建。实验硬件环境采用搭载第 13 代 Intel(R) Core(TM) i7-13620H 处理器和 NVIDIA GeForce RTX 4060 Laptop GPU 的计算平台。软件环境为 Python 3.8、CUDA 12.8 及 PyTorch 1.13。模型训练采用 Adam 优化器，初始学习率设为 10^{-4}，并引入早停机制以防止过拟合，最大迭代轮数设为 100，批次大小为 32。训练过程中采用标准化后的平均绝对误差 δ_{MAE} 作为损失函数。

图 5 展示了模型在训练集与验证集上的损失收敛曲线。由图 5 可见，模型在前 20 轮经历了快速

下降与震荡调整，随后损失值逐渐趋于平稳，验证集损失在第 53 轮达到最低点，随后呈现轻微上升趋势，整个训练过程未出现梯度消失或发散现象，证明了模型结构的稳定性。在计算效率方面，得益于 TCN 的并行卷积结构与 Informer 的概率稀疏注意力机制，模型有效降低了处理长序列的时间复杂度。在上述硬件配置下，单次迭代（epoch）的平均耗时约为 28 s，完整训练流程总耗时约为 32 min。

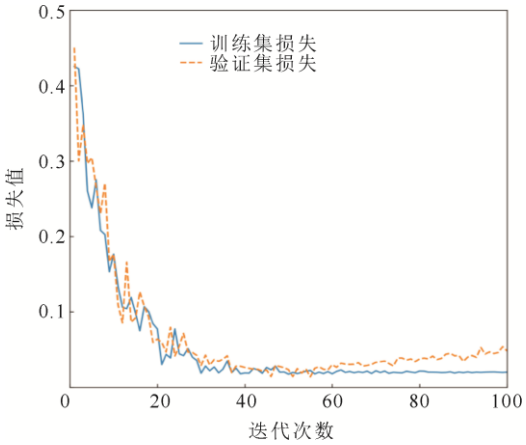


图 5　模型训练集与验证集损失收敛曲线
Fig.5 Loss convergence curves of the model's training set and validation set

3 实验结果与讨论

3.1 模型性能分析

将贝叶斯优化后的超参数组合代入本文提出的 SD-TCN-Informer 组合模型，在测试集上进行最终评估。在负荷基数较大情况下（7 000~9 000 MW），本文模型的平均绝对误差为 233.529，均方根误差仅为 296.71，平均绝对百分比误差控制在 3.02%，表现出极高预测精度。皮尔逊相关系数达 0.974 1，说明预测序列与真实负荷序列之间存在极强的正相关性，模型能够精准复现负荷的波动规律。图 6 为预测值与真实值的散点拟合情况。由图 6 可见，数据点紧密分布在 $y=x$ 理想对角线附近，未出现明显的离散现象。即便在 8 000 MW 以上的高负荷区间，模型依然保持了良好的拟合能力，证明 TCN 模块在提取局部高频特征方面的优势。

图 7 的直方图进一步展示了模型预测平均绝对误差的概率密度分布。误差分布呈现显著的右偏态，绝大多数样本的绝对误差集中在[0, 400]的低值区间内，长尾部分的极端误差极少。这表明模型具有极高的鲁棒性，在面对突发性负荷波动时仍能保持较低的偏差水平。

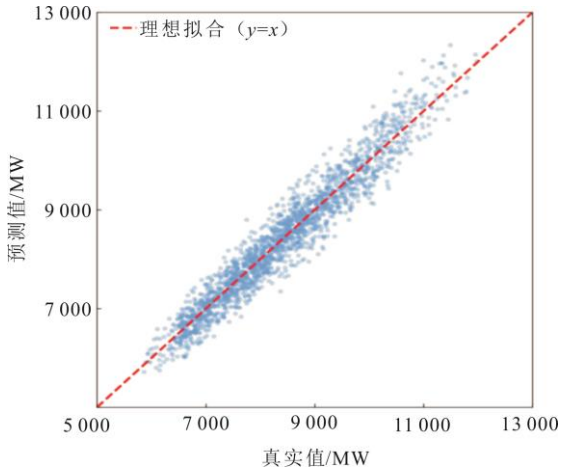


图 6　负荷预测值-真实值散点拟合图
Fig.6 Fitting scatter plot of predicted values vs. actual values of the load

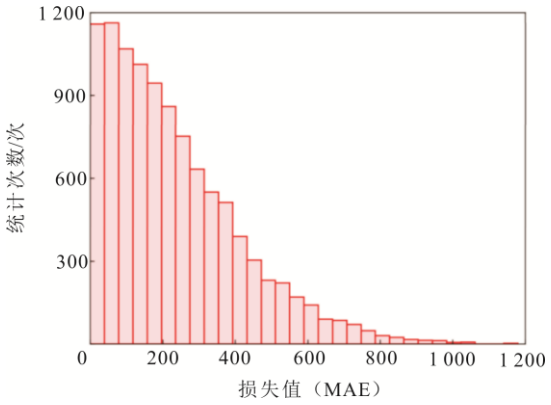


图 7 MAE 损失整体分布
Fig.7 Overall distribution of MAE loss

为了直观评估模型的时序跟踪能力，图 8 展示了测试集中不同日期 24 h 负荷预测走势。

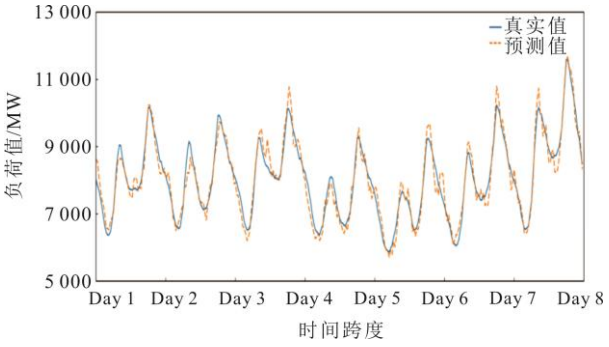


图 8　连续 1 周电力负荷预测结果对比
Fig.8 Comparison of continuous one-week power load prediction results

从图 8 可以看出，无论是早高峰的爬坡阶段，还是午间的平稳过渡期，预测曲线均能紧密跟随真实负荷曲线。特别是针对日内双峰结构（早晚高

峰），Informer 的 ProbSparse-Attention 机制发挥了关键作用，成功捕捉了负荷的长程依赖关系，实现了相位与幅值的双重精准对齐。

从图 8 也可以看出，真实负荷曲线在 7 天内呈现出明显的周期性波动，且工作日与周末的负荷模式存在显著差异。模型成功捕捉到这种周度依赖性，不仅在工作日的高峰期实现了紧密跟踪，在周末负荷整体下降时也能迅速调整预测基准。连续 1 周的预测曲线未出现相位偏移或显著的累积误差，证明模型在长周期序列预测中的鲁棒性。

此外，为了进一步探究模型在不同用电时段的误差分布特性，本文统计了测试集全时段的绝对误差，其箱线图如图 9 所示。由图 9 可见，各小时的误差中位数均保持在较低水平，且位置相对平稳，说明模型未出现特定的失效时段。尽管受随机用电行为影响，各时段均出现了一定数量的离群点，但这些点的数值量级并未失控，且箱体主体依然紧凑。这表明即便在面对突发扰动时，TCN 模块提取的局部特征仍能将预测误差限制在可接受范围内，模型整体具备良好的抗干扰能力。

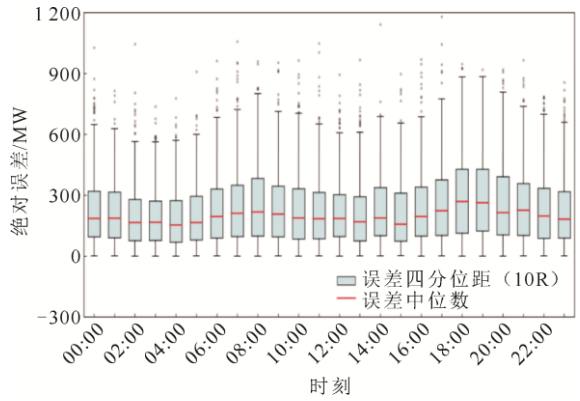


图 9 模型分时绝对误差分布特性
Fig.9 Distribution characteristics of hourly absolute error of the model

3.2 消融实验

为了定量评估“相似日聚类”“TCN 局部特征提取”及“Informer 长程建模”各组件的贡献，本文设计了消融实验。对比模型变体包括：1）本文提出的完整组合模型（SD-TCN-Informer）；2）移除 TCN 模块中残差连接结构仅保留卷积层，用于验证残差连接对深层特征提取的有效性（无 TCN 残差变体 SD-Informer）；3）将 Informer 中的 ProbSparse-Attention 替换为标准 Self-Attention，即退化为 Transformer 结构（SD-TCN-Transformer），验证稀疏注意力机制在捕捉长程依赖中的作用；4）去除相似日聚类模块的变体（TCN-Informer），用于验证相似日划分模块的作用。表 1 为消融实验结果对比。

表 1 消融实验结果对比
Tab.1 Comparison of ablation experiment results

模型变体	δ_{RMSE}	δ_{MAE}
SD-TCN-Informer	296.70	233.52
SD-TCN-Transformer	375.42	278.16
SD-Informer	452.76	313.76
TCN-Informer	348.55	265.38

由表 1 可见，完整模型的性能最优。当去除相似日聚类模块后，模型的均方根误差和平均绝对误差分别显著上升至 348.55 和 265.38。这证明若不显式区分工作日、周末与节假日等异质化用电模式，模型难以有效剥离跨模式噪声，导致整体预测精度明显下降。同时，当移除 TCN 的残差连接结构后，均方根误差大幅上升至 452.76，表明残差机制对于缓解梯度消失、提取深层时序特征至关重要。若将模型简化为单分支结构，平均绝对误差增加至 278.16，验证了多尺度特征融合在处理复杂电力负荷数据时的必要性。

3.3 对比实验

为了进一步验证所提方法的先进性，将其与当前主流的时间序列预测模型 LSTM 和 TimesNet 进行对比。所有模型均在相同的数据集和实验环境下运行，对比结果见表 2。

表 2 与主流模型预测性能对比
Tab.2 Comparison of prediction performance with mainstream models

对比模型	δ_{MAE}	δ_{RMSE}	C_{ORR}	δ_{MAE} 降低幅度/%	δ_{RMSE} 降低幅度/%
LSTM	273.62	355.71	0.945 8	14.66	16.59
TimesNet	264.18	338.45	0.961 2	11.61	12.33
本文模型	233.52	296.70	0.974 1		

由表 2 可知，传统的 LSTM 模型由于存在梯度消失问题，难以捕捉长序列依赖，预测精度最低，平均绝对误差为 273.62。TimesNet 模型凭借其先进的二维时序建模机制，能够有效提取复杂的时序特征，性能优于 LSTM 模型。然而，作为一种通用模型，TimesNet 模型未能针对电力负荷数据进行专门优化，在处理节假日等特殊模式时仍存在局限性，均方根误差为 338.45。相比之下，本文提出的组合模型在所有指标上均实现了最佳表现。具体而言，

与基准模型 LSTM 相比，本文模型的均方根误差和平均绝对误差分别降低了 16.59%和 14.66%；与 TimesNet 相比，本文模型的均方根误差和平均绝对误差降低了 12.33%和 11.61%。这充分证明了通过相似日聚类划分数据模式，并结合 TCN 与 Informer 优势的策略，能够有效解决短期电力负荷预测中“高波动”与“长依赖”难以兼顾的痛点。

4 结论与展望

本文针对短期电力负荷预测中存在的非线性强、波动性大及长程依赖难以捕捉的问题，提出了一种 SD-TCN-Informer 的组合预测方法。通过在澳大利亚 AEMO-NSW 30 min 公开数据集上的实验验证，得出以下结论。

1）在负荷预测准确率方面，利用模糊 C 均值算法将历史数据划分为不同相似日簇，有效降低了数据分布的复杂性，使模型能针对特定用电模式进行精准学习。TCN 模块有效提取了负荷序列的局部高频波动特征，而 Informer 网络通过稀疏注意力机制成功捕获了长序列的时间依赖性。两者结合使得模型在均方根误差和平均绝对误差指标上分别达到 296.71 和 233.53，预测精度显著优于单一的 LSTM 或 TimesNet 模型。表明模型在电力负荷预测中具有高准确性和强稳定性。

2）在工程应用方面，模型在 30 min 分辨率的 24 h 预测任务中表现出 3.02%的低平均绝对百分比误差和 0.974 1 的高相关性，且误差分布集中，具备良好的鲁棒性，可为电力系统的短期调度与现货交易提供可靠的数据支撑。

尽管本文提出的模型在电力负荷预测的准确性和稳定性方面整体表现更为突出，但是实验也具有一定的局限性，未来的研究工作将考虑引入更多因素作为协变量，进一步探索多源信息融合对极端天气下负荷预测精度的提升作用。

[参 考 文 献]

[1] 魏旭, 刘东, 高飞, 等. 双碳目标下考虑源网荷储协同优化运行的新型电力系统发电规划[J]. 电网技术, 2023, 47(9): 3648.
Wei X, Liu D, Gao F, et al. Generation planning of new power system considering source-grid-load-storage synergistic optimization under dual-carbon goals[J]. Power System Technology, 2023, 47(9): 3648.

[2] 范茜茜, 王国强, 罗贺, 等. 基于 N-BEATS 与辅助编码器的短期电力负荷预测[J]. 电网技术, 2024, 48(4): 1612.
Fan X X, Wang G Q, Luo H, et al. Short-term power load forecasting based on N-BEATS and auxiliary encoder[J]. Power System Technology, 2024, 48(4): 1612.

[3] Lin J, Ma J, Zhu J, et al. Short-term load forecasting based on LSTM networks considering attention mechanism[J]. International Journal of Electrical Power & Energy Systems, 2022, 137: 107818.

[4] Yang D, Guo J, Sun S, et al. An interval decomposition-ensemble approach with data-characteristic-driven reconstruction for short-term load forecasting[J]. Applied Energy, 2022, 306: 117992.

[5] Eren Y, Küçükdemiral İ. A comprehensive review on deep learning approaches for short-term load forecasting[J]. Renewable and Sustainable Energy Reviews, 2024, 189: 114031.

[6] Chen Z, Chen Y, Xiao T, et al. A novel short-term load forecasting framework based on time-series clustering and early classification algorithm[J]. Energy & Buildings, 2021, 251: 111375.

[7] Fazlipour Z, Mashhour E, Joorabian M. A deep model for short-term load forecasting applying a stacked autoencoder based on LSTM supported by a multi-stage attention mechanism[J]. Applied Energy, 2022, 327: 120063.

[8] Zulfiqar M, Kamran M, Rasheed M B, et al. A hybrid framework for short term load forecasting with a novel feature engineering and adaptive grasshopper optimization in smart grid[J]. Applied Energy, 2023, 338: 120829.

[9] 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.
Zhao Y, Wang H M, Kang L, et al. Short-term power load forecasting based on temporal convolutional network[J]. Transactions of China Electrotechnical Society, 2022, 37(5): 1242.

[10] Türkoğlu A S, Erkmen B, Eren Y, et al. Integrated approaches in resilient hierarchical load forecasting via TCN and optimal valley filling based demand response application[J]. Applied Energy, 2024, 360: 122722.

[11] Zuo K. Integrated forecasting models based on LSTM and TCN for short-term electricity load forecasting[C]//2023 9th International Conference on Electrical Engineering, Control and Robotics (EECR). IEEE, 2023: 207.

[12] 李文超, 黄伟, 李鹏, 等. 多尺度特征融合增强的短期电力负荷预测深度学习网络[J]. 云南大学学报(自然科学版), 2025, 47(5): 839.
Li W C, Huang W, Li P, et al. Deep learning network for short-term power load forecasting enhanced by multi-scale feature fusion[J]. Journal of Yunnan University (Natural Sciences Edition), 2025, 47(5): 839.

[13] 任建吉, 位慧慧, 邹卓霖, 等. 基于 CNN-BiLSTM-Attention 的超短期电力负荷预测[J]. 电力系统保护与控制, 2022, 50(8): 108.
Ren J J, Wei H H, Zou Z L, et al. Ultra-short-term power load forecasting based on CNN-BiLSTM-attention[J]. Power System Protection and Control, 2022, 50(8): 108.

[14] Qu K, Si G, Shan Z, et al. Forward former: efficient transformer with multi-scale forward self-attention for day-ahead load forecasting[J]. IEEE Transactions on Power Systems, 2023, 39(1): 1421.

[15] 范杏蕊, 李元诚. 基于改进 Autoformer 模型的短期电力负荷预测[J]. 电力自动化设备, 2024, 44(4): 171.
Fan X R, Li Y C. Short-term power load forecasting based on improved autoformer model[J]. Electric Power Automation Equipment, 2024, 44(4): 171.

[16] 周师琦, 王俊帆, 赖俊升, 等. 结合贝叶斯 Autoformer 的多维自适应短期电力负荷概率预测方法[J]. 电子与信息学报, 2024, 46(12): 4432.
Zhou S Q, Wang J F, Lai J S, et al. Multi-dimensional adaptive short-term power load probability forecasting method combining Bayesian autoformer[J]. Journal of Electronics & Information Technology, 2024, 46(12): 4432.
[17] Alsamraee S A, Khanna S, Lund H, et al. A hybrid transformer-TCN-GRU based model for thermal load forecasting of a large university campus[J]. Energy, 2026, 334: 140114.
[18] Li Q, Qiu Y, Liu W, et al. Improving load forecasting accuracy through data augmentation and TCNs-transformer[C]//2024 4th Power System and Green Energy Conference (PSGEC). IEEE, 2024: 65.
[19] Li M, Zhang T, Yang H, et al. Multiple load forecasting of integrated renewable energy system based on TCN-FECAM-Informer[J]. Energies, 2024, 17(20): 5181.
[20] 徐衍会, 张蓝宇, 宋歌. 基于核的模糊 *C* 均值逐层聚类算法在负荷分类中的应用[J]. 电力建设, 2015, 36(4): 46.
Xu Y H, Zhang L Y, Song G. Application of kernel-based fuzzy C-means hierarchical clustering algorithm in load classification[J]. Electric Power Construction, 2015, 36(4): 46.
[21] 李辰浩, 郑红, 张佳鑫, 等. 融合时序卷积网络与傅里叶分解的时序预测模型[J]. 华东理工大学学报（自然科学版）, 2026, 52: 1.
Li C H, Zheng H, Zhang J X, et al. Time series prediction model integrating temporal convolutional network and Fourier decomposition[J]. Journal of East China University of Science and Technology (Natural Science Edition), 2026, 52: 1.
[22] 宋浏阳, 金烨, 郭旭东, 等. 基于自适应权重时间卷积网络的剩余使用寿命预测方法[J]. 北京化工大学学报(自然科学版), 2024, 51(3): 76.
Song L Y, Jin Y, Guo X D, et al. Remaining useful life prediction method based on adaptive weight temporal convolutional network[J]. Journal of Beijing University of Chemical Technology (Natural Science Edition), 2024, 51(3): 76.
[23] Zhou H, Zhang S, Peng J, et al. Informer: beyond efficient transformer for long sequence time-series forecasting[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(12): 11106.

（责任编辑　杜亚勤）