

基于深度学习的乡村综合能源系统短期负荷预测

杨浚文¹, 徐志², 姜詠², 覃日升², 任敏², 赵男², 袁志伟³, 李海良⁴

(1. 云南电网有限责任公司, 昆明 650051; 2. 云南电网有限责任公司电力科学研究院, 昆明 650051; 3. 云南电网有限责任公司红河供电局, 云南 蒙自 661100; 4. 暨南大学国际能源学院/能源电力研究中心, 广东 珠海 519000)

摘要: 乡村综合能源系统短期负荷的准确预测对于乡村电网的安全稳定运行及新能源就地消纳具有重要意义。然而, 受农业生产节律和天气突变等多重因素耦合影响, 乡村综合能源系统负荷呈现出强随机性、非平稳性及多时间尺度特征, 导致传统方法预测精度受限。为此提出了一种基于VMD-Informer-BiGRU模型的短期负荷预测方法。该方法首先构建了包含农忙时段、环境温度、节假日等多维特征集, 以精准表征乡村特有的用能模式; 利用变分模态分解(variational mode decomposition, VMD)将原始负荷序列解耦为具有不同频率特性的模态分量, 有效消除随机波动干扰; 在此基础上引入Informer模型深度挖掘全局长期时间依赖, 并协同双向门控循环单元(BiGRU)捕捉局部突变特征, 实现多尺度时序特征的深度融合; 最后构建了组合预测模型。以中国某乡村地区实际用能场景为预测对象, 计算结果表明所提模型的预测精度高于其他模型。

关键词: 乡村综合能源系统; 负荷预测; Informer模型; 双向门控循环单元(BiGRU); 变分模态分解(VMD)

Short-Term Load Forecasting for Rural Integrated Energy Systems Based on Deep Learning

YANG Junwen¹, XU Zhi², JIANG He², QIN Risheng², REN Min², ZHAO Nan², YUAN Zhiwei³,
LI Hailiang⁴

(1. Yunnan Power Grid Co., Ltd., Kunming 650051, China; 2. Electric Power Research Institute of Yunnan Power Grid Co., Ltd., Kunming 650051, China; 3. Honghe Power Supply Bureau of Yunnan Power Grid Co., Ltd., Mengzi, Yunnan 661100, China; 4. Energy and Electricity Research Center, Jinan University, Zhuhai, Guangdong 519000, China)

Abstract: Accurate short-term load forecasting of rural integrated energy systems is of great significance for the safe and stable operation of rural power grids and local renewable energy consumption. However, due to the coupled effects of agricultural production rhythms, sudden weather changes, and other factors, the load of rural integrated energy systems exhibits strong randomness, non-stationarity, and multi-time-scale characteristics, which limit the forecasting accuracy of traditional methods. To address this issue, a short-term load forecasting method is proposed based on the VMD-Informer-BiGRU model. Firstly, a multidimensional feature set including farming periods, ambient temperature, and holidays is constructed to accurately characterize the unique energy consumption patterns in rural areas. Variational mode decomposition (VMD) is then used to decompose the original load sequence into modal components with different frequency characteristics, effectively reducing random fluctuation interference. On this basis, the Informer model is introduced to extract global long-term temporal dependencies, while the bidirectional gated recurrent unit (BiGRU) is employed to capture local mutation features, achieving deep fusion of multi-scale temporal features. Finally, a combined forecasting model is established. Using an actual rural energy consumption scenario in China as the case study, the results show that the proposed model achieves higher forecasting accuracy than other models.

Key words: rural integrated energy system; load forecasting; Informer model; bidirectional gated recurrent unit (BiGRU); variational mode decomposition (VMD)

0 引言

随着国家“双碳”战略与“乡村振兴”战略的深入实施,构建以高比例可再生能源为主体的新型农村电力系统已成为能源转型的关键阵地^[1-2]。在此背景下,农村地区凭借其广阔的地域空间和资源禀赋,正成为未来增量分布式光伏、生物质能等新能源规模化利用的主战场,催生了源网荷储高度协同的新型农村电网转型升级^[3-4]。然而,与负荷密度高、且相对规律的城市电网不同,农村电网广阔的地理分布使其负荷特性深受天气突变、农忙灌溉、农业加工等季节性生产节律及分布式电源出力波动的多重耦合影响,呈现出极强的随机性、非平稳性与“源荷双侧波动”特征^[5]。因此,开展高精度的乡村综合能源系统短期负荷预测,对于优化电力调度策略、保障农村电网安全稳定运行及促进新能源就地消纳具有重要的现实意义。

近年来,随着人工智能技术的快速迭代,深度学习算法因其强大的非线性映射与特征提取能力^[6-7],逐渐在负荷预测领域占据主导地位^[8-10]。相比于自回归综合移动平均(autoregressive integrated moving average model, ARIMA)^[11]、自回归移动平均(autoregressive moving average model, ARMA)^[12]等传统统计学模型难以准确描述非线性特性的局限,神经网络能更好地捕捉数据中的复杂动态关系^[13]。主流的单一深度学习模型主要包括长短期记忆网络(long short-term memory, LSTM)^[14]、门控循环单元(gated recurrent unit, GRU)及其双向变体(BiLSTM/BiGRU)等^[15-16]。Lakshmi等构建了基于长短期记忆网络(LSTM)的短期电力需求预测模型,利用门控机制挖掘长程依赖,有效提升了非线性负载预测精度^[17]。为了进一步挖掘序列的前后文双向依赖关系,Zhao等和Michael等构建了BiLSTM预测模型,有效提升了对波动性负荷的拟合能力^[18-19]。

此外,针对时间序列的局部特征捕捉,Wang等引入了时域卷积网络(temporal convolutional network, TCN),利用膨胀卷积结构扩大感受野,在捕捉长程依赖方面展现出优于传统循环神经网络(recurrent neural network, RNN)的潜力^[20-21]。与此同时,Zhan等应用Transformer模型,通过自注意力机制并行挖掘序列全局关联,显著增强了长程依

赖捕捉能力,但其二次方计算复杂度限制了长序列预测效率^[22-24]。尽管上述模型各具优势,但仍难以完全满足乡村综合能源系统短期负荷预测的高要求。具体而言,循环神经网络类模型在处理超长序列输入时仍面临梯度消失风险,且单一模型往往难以兼顾长周期趋势与短时突变特征的提取。

为进一步提升模型在复杂场景下的预测性能,研究者提出了多种混合预测模型,旨在融合不同算法的优势。目前常见的混合策略主要包括“特征提取+时序预测”和“信号分解+预测集成”两类。在特征融合方面,Li等和Chung等构建了基于注意力机制的CNN-LSTM(convolutional neural networks-long short-term memory)混合模型,利用CNN挖掘多维数据的空间相关性,再通过LSTM捕捉时间依赖性^[25-26];Niu等提出了CNN-BiGRU组合架构,实现了对光伏与负荷数据的时空特征深度挖掘^[27]。在信号分解方面,为应对负荷数据的非平稳性,张未等和王俊等采用变分模态分解(variational mode decomposition, VMD)结合LSTM模型,有效降低了序列的随机波动干扰^[28-29]。此外,针对长序列建模难题,Gong等和周璇等将Informer模型应用于负荷预测,通过概率稀疏注意力机制降低了Transformer的计算复杂度^[30-32]。然而,现有的混合模型大多侧重于单一维度的优化,缺乏对“全局长程依赖”与“局部突变细节”的协同建模,特别是针对受农忙及气候突变影响强烈的乡村负荷,现有方法在捕捉其多时间尺度耦合特性方面仍存在局限性。

鉴于乡村负荷数据兼具长周期季节性趋势与短时剧烈农事波动的复杂特性,本文提出一种基于VMD-Informer-BiGRU的组合预测模型。该模型采用“全局-局部协同”的建模思路:首先,利用变分模态分解(VMD)将非平稳的原始负荷序列解耦为多个频率稳定的模态分量;其次,引入Informer模型,利用其概率稀疏自注意力机制高效捕捉负荷序列的全局长程依赖特征;同时,针对乡村特有的由农忙或天气突变引起的局部负荷骤变,串联双向门控循环单元(BiGRU)进行二次特征提炼,以弥补Informer在局部细节刻画上的不足。最后,结合构建的农忙时段、环境温度等多维特征集,通过我国某乡村地区的实际用能场景进行算例分析,验证所提方法在多时间尺度下的预测有效性与鲁棒性。

1 VMD分解

VMD的核心思想是通过构建并求解一个约束变分问题来实现信号的自适应分解。该问题旨在寻找一组模态分量 $\{u_k\}$ 及其对应的中心频率 $\{\omega_k\}$,使得所有模态的估计带宽之和最小,同时约束所有模态之和能够精确重构原始信号 $f(t)$ 。该约束变分问题可表述为:

$$\min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \quad (1)$$

式中: ∂_t 为对时间 t 的一阶偏导数; $\delta(t)$ 为狄拉克冲激函数; $\|\cdot\|_2^2$ 为 L^2 范数; j 为虚数单位; $*$ 为卷积符号。

同时还需要满足式(2):

$$\text{s.t.} \sum_{k=1}^K u_k = f(t) \quad (2)$$

式中 K 为模态数量。

通过采用二次惩罚因子 α 和拉格朗日乘子 $\lambda(t)$ 这两种数学工具, VMD 方法将最初的约束问题进行转化, 从而构建出一个等价的、但无约束的增广拉格朗日表达式 L :

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle \quad (3)$$

式中 $\langle \cdot, \cdot \rangle$ 为内积运算。

该变分问题采用交替方向乘子法 (ADMM) 求解。通过迭代交替更新模态分量 u_k 、中心频率 ω_k 及拉格朗日乘子 λ , 直至满足收敛判据。最终对频域解进行逆傅里叶变换, 实现对原始信号的有效分解。

2 Informer-BiGRU 模型

2.1 Informer模型

Informer是一种基于Transformer架构改进的高效时间序列预测模型,旨在解决传统Transformer在处理长序列时计算复杂度高、内存占用大以及推理速度慢的问题。在本文构建的组合预测框架中,Informer模块主要承担提取负荷序列全局长程依赖(global temporal dependency)的任务。该模型保留

了经典的编码器-解码器(encoder-decoder)架构。其结构如图1所示。

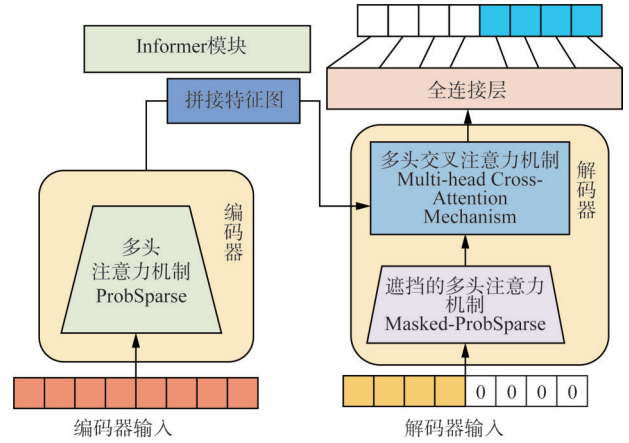


图1 Informer模型结构图

Fig. 1 Structure of Informer model

2.1.1 概率稀疏自注意力机制

在处理乡村综合能源系统负荷数据时,输入序列往往包含明显的周期性和趋势性。传统Transformer的自注意力机制需要计算所有查询(Query)与键(Key)的点积,其时间复杂度和空间复杂度均为 $O(L^2)$ (其中 L 为序列长度),这极大地限制了模型的运算效率。Informer指出,在自注意力矩阵中,只有少量的查询向量对长程依赖的捕捉起关键作用,呈现出显著的稀疏性分布。为此,本文采用概率稀疏自注意力机制(probSparse self-attention)替代全量注意力。该机制基于Kullback-Leibler(KL)散度,筛选出对注意力分布贡献最大的“活跃”查询向量,将计算复杂度降低至 $O(L \log L)$ 。对于输入张量,首先通过线性投影生成查询矩阵 Q 、键矩阵 K 和值矩阵 V 。ProbSparse机制仅选取稀疏度量值最高的Top- u 个查询向量(记为 $\bar{Q}$)参与点积运算,其余查询向量的注意力输出则由均值替代。其计算公式如下:

$$A(Q, K, V) = \text{Softmax} \left(\frac{\bar{Q} K^T}{\sqrt{d}} \right) V \quad (4)$$

式中: d 为输入维度; $\bar{Q}$ 为筛选后的稀疏查询矩阵; Softmax 为归一化指数函数。

2.1.2 生成式解码器

为了避免传统递归预测导致的误差累积,Informer采用生成式解码器。其输入由一段紧邻预测时刻的历史序列作为上下文引导,与待预测时刻

的占位符拼接而成。基于交叉注意力机制，解码器仅需一次前向传播即可聚合全局特征，直接并行生成未来所有时间步的预测结果。该机制不仅大幅提升了推理效率，更保证了长步长预测的稳定性，最终经全连接层映射得到初步预测特征。

2.2 BiGRU 模型

尽管 Informer 擅长捕捉长程依赖，但对局部剧烈波动的敏感度不足。为此，本文引入 BiGRU 以深度挖掘微观时序规律。作为 LSTM 的轻量化变体，GRU 通过更新门和重置门高效控制信息流，在保留非线性拟合能力的同时显著降低了计算复杂度，有效满足实时性要求。为充分利用序列的前后文双向特征，本文构建 BiGRU 网络，其结构如图 2 所示。

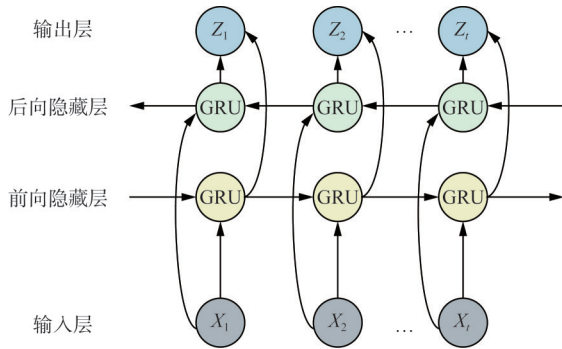


图 2 BiGRU 模型结构图

Fig. 2 Structure of BiGRU model

如图 2 所示，BiGRU 由前向与后向 GRU 叠加构成。前向层沿时间正序提取历史影响，后向层沿逆序捕捉未来反馈，从而实现对序列双向时序特征的完整提取。在每个时间步 t ，BiGRU 的最终输出由两个方向的隐藏状态拼接而成，即 $h_t = [\vec{h}_t \oplus \overleftarrow{h}_t]$ ，其中 $\vec{h}_t$ 和 $\overleftarrow{h}_t$ 分别为前向和后向隐藏层状态，表示向量拼接操作。在组合模型中，BiGRU 紧接 Informer 解码器之后，利用双向递归特性对输出特征进行二次提炼，重点增强模型对乡村负荷中由极端天气或突发性行为引发的短时、高频波动的感知能力。

2.3 Informer-BiGRU 模型构建

为同时利用 Informer 在长序列全局时序依赖建模方面的优势，以及 BiGRU 在捕捉局部短时剧烈波动特征方面的强项，本文构建了如图 3 所示的串联混合预测模型。

2.4 负荷预测结果评价指标

为了评估预测模型的效果和精准度，选择拟合

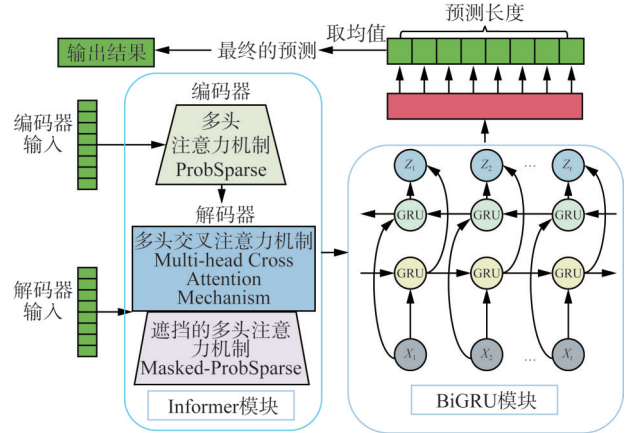


图 3 Informer-BiGRU 网络的结构图

Fig. 3 Architecture of the Informer-BiGRU network

优度检定 R^2 、均方误差 E_{MSE} 、平均绝对误差 E_{MAE} 和均方根误差 E_{RMSE} 作为预测评估指标，这些指标能够真实具体地反映模型对乡村综合能源系统短期负荷的预测误差情况和效果。评价指标公式可表示为：

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (5)$$

$$E_{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6)$$

$$E_{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (7)$$

$$E_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (8)$$

式中： $\hat{y}_i$ 为第 i 时刻乡村负荷的预测值； y_i 为第 i 时刻负荷的真实值； $\bar{y}$ 为乡村负荷真实值的平均值； n 为测试样本数量。

2.5 VMD-Informer-BiGRU 模型的构建

为系统性解决乡村综合能源系统短期负荷预测中存在的非线性强、季节性明显及局部波动剧烈等难题，本文构建了 VMD-Informer-BiGRU 组合预测模型，模型结构如图 4 所示。

该流程具体执行步骤如下。

1) 数据导入与特征工程：导入并且可视化原始数据，构建包含温湿度等常规天气特征，以及农忙时段、季度等乡村特有特征的多维特征集。

2) 时序划分与归一化：将按时间顺序划分训练集与测试集，防止信息泄露。采用 Min-Max 方法将特征缩放至 $[0, 1]$ 区间，测试集归一化参数严格复

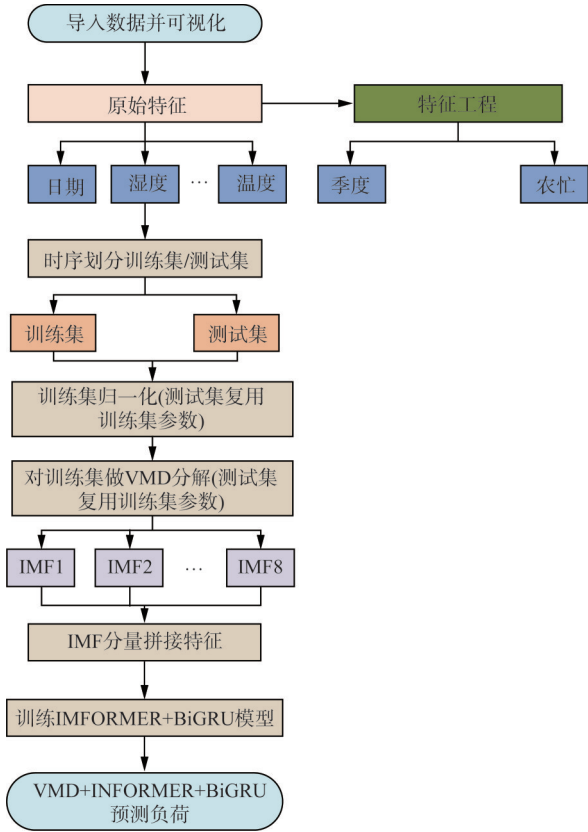


图 4 VMD-Informer-BiGRU 流程图

Fig. 4 Flow chart of VMD-Informer-BiGRU

用训练集统计量。

3)VMD 特征增强: 对训练集负荷进行 VMD 分解, 提取 K 个本征模态分量(IMFs)作为增广特征拼接到原始矩阵。测试集复用训练集分解参数, 确保处理一致性。

4)Informer-BiGRU 混合建模: 将多维张量输入组合模型, 先由 Informer 捕捉全局长程依赖, 再经 BiGRU 提取局部突变特征。

5)预测输出: 经全连接层映射输出预测值, 并通过反归一化还原量纲。该流程系统性应对了数据复杂性, 实现了高精度预测。

通过这一系列步骤, VMD-Informer-BiGRU 模型能够系统性地应对原始数据中的多重复杂性, 从而实现比单一模型更精准、更鲁棒的预测性能。

3 数据准备与设备条件

3.1 数据准备

本文的数据集源自国内南部某乡村地区 2018 年的逐时负荷与天气记录, 该时间分辨率能够精确

刻画负荷的日内波动细节, 满足短期负荷预测对于高频颗粒度的要求。其中, 前 6 982 个点用于模型训练, 其余 1 778 个点则用于性能测试。

3.2 设备条件

本研究的算例基于 Python3.12 实现, 实验硬件平台为: Intel Core Ultra 7 处理器、NVIDIA GeForce RTX 4070 Ti Super (12GB GDDR6X 显存) 显卡、32 GB 内存及 2 TB 固态硬盘。

4 算例分析

4.1 基于乡村综合能源系统短期负荷的特征工程

乡村综合能源系统负荷受农业节律、天气及生活习惯等多重因素驱动, 具有强非线性特征。鉴于短期负荷预测主要依赖于近期微观天气条件与用户作息规律, 而非长期宏观经济指标, 为挖掘潜在规律, 本文构建了包含农忙时段、有效温度、季度及小时索引的关键特征集。在时间维度上, 提取“小时”与“季度”刻画日、季周期波动; 在生产与气候维度上, 引入“农忙”反映农业冲击, 构建“有效温度”量化热累积效应。多维特征的引入旨在还原短期负荷的动态机理, 为 VMD-Informer-BiGRU 模型奠定数据基础。

4.1.1 基于农业物候的农忙时段特征构建

与城市电网不同, 乡村综合能源系统负荷特性与农业生产周期呈现强耦合性, 其峰值波动往往由高功率农业设备的集中使用主导。为捕捉农事活动引起的负荷突变, 本文引入农业物候学概念, 构建“农忙时段”特征变量, 根据作物生长周期将全年划分为 3 个关键高负荷阶段: 3 月中旬至 5 月上旬的春灌返青期, 因需水量大集中开启大功率水泵灌溉导致负荷持续抬升; 6 月上旬至中旬的“三夏”抢收抢种期, 多台农机设备运行叠加夏季制冷负荷, 极易形成全年负荷最高峰; 9 月下旬至 10 月中旬的秋收加工期, 大规模加工设备投入使用, 再次引发负荷显著波动。

基于上述物候规律, 本文构建了离散化的农忙特征变量 $F_{\text{farm}}(t)$, 通过分段映射函数将连续的时间序列划分为不同的农业生产状态, 具体定义如下。

$$F_{\text{farm}}(t) = \begin{cases} 1, & \text{春灌 (3 月 15 日 - 5 月 10 日)} \\ 2, & \text{三夏 (6 月 1 日 - 6 月 20 日)} \\ 3, & \text{秋收 (9 月 20 日 - 10 月 15 日)} \\ 0, & \text{非农忙时期 (其他时段)} \end{cases} \quad (9)$$

4.1.2 有效温度特征构建

在乡村综合能源系统中,气温是影响负荷变化的最主要天气因素,但负荷响应存在显著的“热惯性”滞后。为准确量化热量积累效应,本文构建“有效温度 T_{eff} ”特征,通过对当前及历史气温进行加权组合,模拟热量在时间维度上的衰减过程。其计算公式如下。

$$T_{\text{eff}}(t) = 0.5T(t) + 0.3T(t-1) + 0.2T(t-2) \quad (10)$$

式中 $T(t)$ 、 $T(t-1)$ 和 $T(t-2)$ 分别为当前时刻、前1时刻和前2时刻的实际气温值。权重系数采用递减分布,反映了距离当前时刻越近的气温对负荷影响越显著的物理规律。

4.1.3 多尺度时间周期特征构建

为刻画乡村综合能源系统负荷显著的“日-季”双重周期性,本文构建小时索引与季度标识特征。其中,小时索引(0-23)用于精细刻画居民作息及设备日内运行规律;季度特征(1-4)则捕捉气候更替及农事轮作引起的长期季节性差异。引入双维度特征有助于模型快速锁定微观与宏观层面的负荷周期性模式。

4.2 基于皮尔逊相关系数的特征分析

为了消除冗余特征对模型训练的干扰并提升预

测效率,本文采用皮尔逊相关系数对各输入特征与负荷之间的线性相关性进行量化筛选。特征-负荷相关系数热力图如图5所示,各特征的关联强度与统计显著性检验结果如表1所示。

由图5可知,水汽压、湿度、季度及时刻索引与负荷呈显著正相关。其中,水汽压相关系数高达0.50,主导作用显著;时刻与季度特征则验证了周期性驱动机制。农忙类型与负荷呈负相关,这是由于农忙时段的负荷特性呈现两极分化,虽然6月“三夏”期间叠加高温制冷导致负荷激增,但3月和9—10月的农忙期处于春秋气候舒适区,基础负荷较低。而非农忙期涵盖了7—8月的盛夏高温及12—1月的深冬采暖期,整体平均负荷水平较高,从而导致统计意义上农忙特征与负荷呈现弱负相关。

在此基础上,风速、降水量和能见度相关系数均低于0.1,予以剔除。此外,针对温度类特征的强共线性,为避免冗余,本文优先保留物理意义明确的“有效温度”。该特征融合了历史气温的累积滞后效应,相比单纯体感温度,更能准确刻画乡村负荷的“热惯性”机理。最终,筛选出水汽压、有效温度、湿度、季度、压强、时刻索引及农忙类型7个关键特征用于建模。

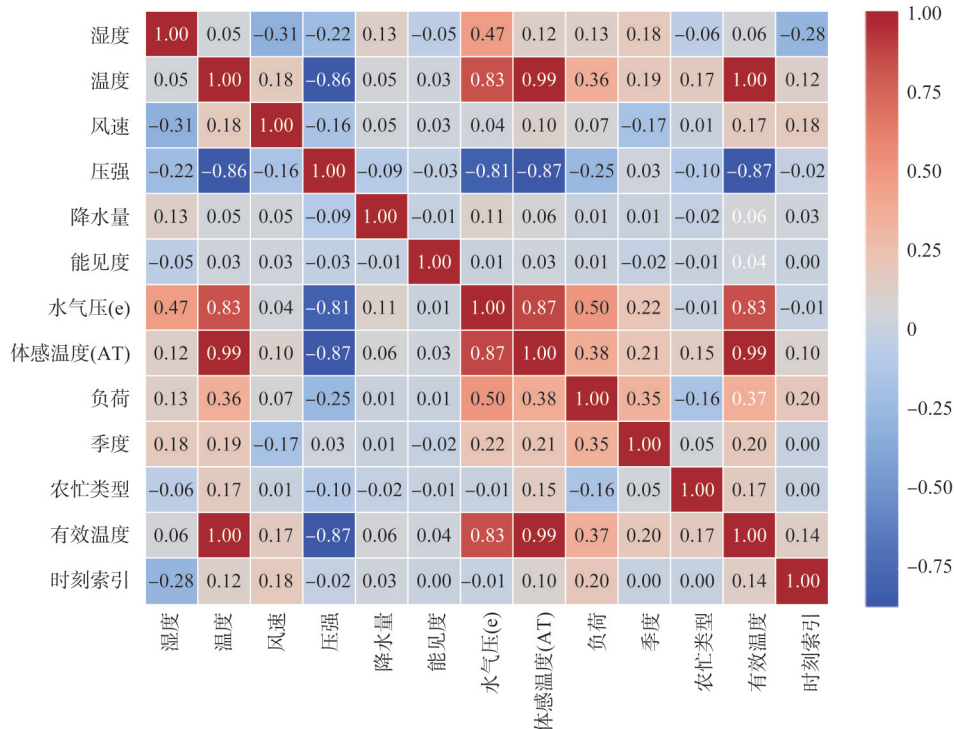


图5 乡村综合能源系统负荷数据的皮尔逊相关系数热力图

Fig. 5 Heatmap of Pearson correlation coefficient for load data in rural integrated energy systems

表 1 各特征与负荷的相关性及其显著性检验结果

Tab. 1 Correlation analysis and significance test results of features and load

特征	相关系数	p 值	显著性
水气压(e)	0.498 1	0.001	极显著
体感温度(AT)	0.384 4	0.001	极显著
有效温度	0.368 7	0.001	极显著
温度	0.360 0	0.001	极显著
季度	0.348 4	0.001	极显著
压强	-0.254 5	0.001	极显著
每天第几小时	0.203 8	0.001	极显著
农忙类型	-0.156 5	0.001	极显著
湿度	0.126 6	0.001	极显著
风速	0.072 1	0.001	极显著
降水量	0.010 0	0.348	不显著
能见度	0.007 4	0.486	不显著

4.3 VMD 分解

乡村综合能源系统的负荷数据因其源荷多样性, 呈现出强烈的随机波动、非平稳性及多尺度特性。为有效处理此复杂序列, 本文采用 VMD 对原始负荷数据进行分解。其中, 分解层数 K 的选择对 VMD 的分解效果至关重要, 本文对 VMD 的关键参数进行了设置: 模态数 K 由中心频率法选定为 8, 惩罚因子设为 4 380, 收敛判据为 10^{-7} 。最终的分解效果参见图 6。

如图 6 所示, 负荷序列被成功分解为 8 个不同频率尺度的模态函数(IMF)。其中, IMF1 和 IMF2 为低频分量, 其曲线平滑, 反映了乡村负荷的长期基础趋势以及由季节交替导致的宏观周期性变化规律; IMF3—IMF5 为中频分量, 反映了受居民生活作息及常规农业生产活动影响的日、周周期性波动特征; IMF6—IMF8 为高频分量, 敏锐捕捉了负荷数据中由微天气突变、大功率农机随机启停等不确定因素诱发的非平稳随机扰动及高频噪声; 随着分解层级的递进, 各模态分量的中心频率逐渐升高, 波动随机性显著增强, 而能量幅值则相应衰减, 验证了该方法在多尺度特征提取方面的有效性。

4.4 模型预测

本实验是在 Python3.12 环境下搭建 PyTorch 深度学习框架, 构建 VMD-Informer-BiGRU 组合预测网络。模型采用滑动窗口技术进行预测, 并在网络层间引入 Dropout 机制以防止过拟合。输入数据采用前文所述的多维张量, 该张量融合了经 VMD 分

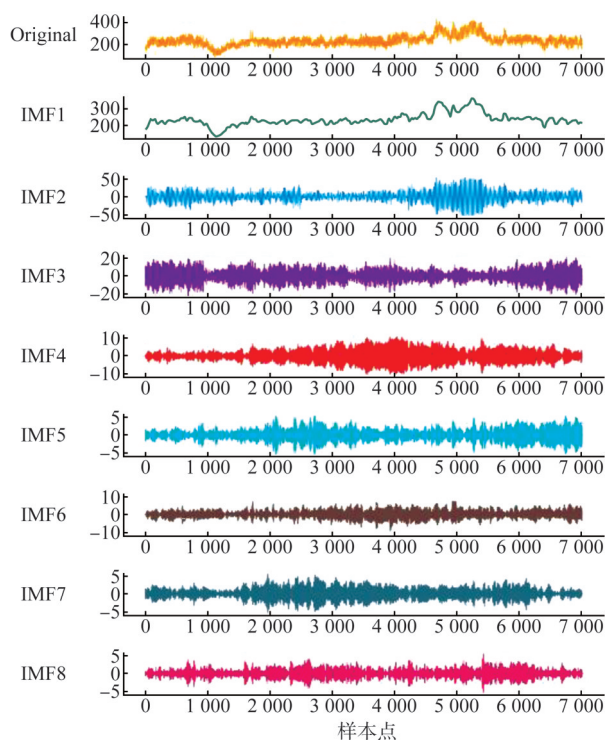


图 6 乡村综合能源系统负荷数据的 VMD 模态分量图

Fig. 6 Plot of VMD modal components for load data in rural integrated energy systems

解得到的 IMFs 分量、筛选后的天气特征及农忙时段等多维信息。模型的核心超参数设置如下: Informer 模块的编码器与解码器层数均设为 1, 多头注意力机制的头数设为 4。BiGRU 模块的隐藏层维度设为 64, 并在层间引入 Dropout 防止过拟合。模型训练采用均方误差作为损失函数, 优化器选用 Adam 算法。

4.4.1 预测结果可视化分析

为了评估模型的预测性能, 选取测试集中的前 1 000 个样本点进行可视化对比, 结果如图 7 所示。

如图 7 所示, 红色实线代表真实负荷, 蓝色虚线代表模型预测负荷。可以看出, VMD-Informer-

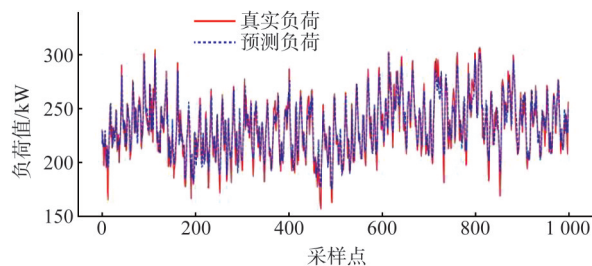


图 7 负荷预测值与真实值对比图

Fig. 7 Comparison of forecasted and actual load profiles

BiGRU 模型的预测曲线与真实负荷曲线高度贴合。无论是在负荷平稳变化的低谷期,还是在波动剧烈的高峰期,模型均能准确跟踪负荷的变化趋势。特别是针对乡村综合能源系统负荷特有的尖峰突变,得益于 BiGRU 对局部特征的二次提炼,模型展现出了优异的捕捉能力,有效克服了传统模型在极值点处的“滞后”或“削峰”现象。

4.5 预测模型对比分析

为了全面评估 VMD-Informer-BiGRU 组合模型在乡村综合能源系统负荷预测中的性能,本文设计了包含单一模型、基础组合模型以及不同深度学习架构混合模型在内的 8 种对比方案。

4.5.1 多模型预测结果对比

8 种模型对乡村负荷变化的跟踪能力的对比图如图 8 所示。

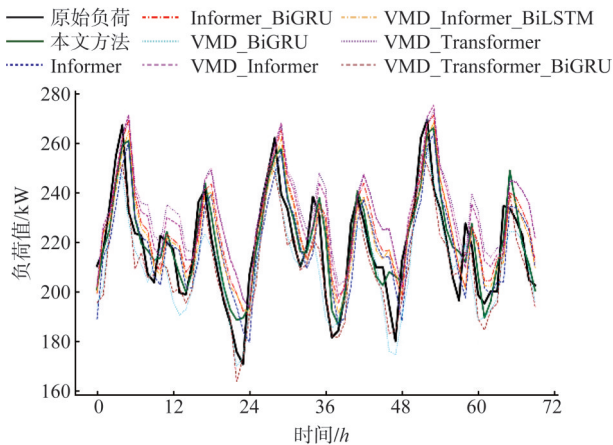


图 8 不同模型负荷预测结果对比图

Fig. 8 Comparison of load forecasting results among different models

从图 8 可以观察到,大多数模型均能捕捉到负荷的日周期波动趋势。然而,单一的 Informer 模型和 VMD-Informer 模型在负荷波峰处表现出明显的“削峰”现象,且相位存在滞后,难以准确响应乡村负荷的快速变化。引入 VMD 分解的模型在曲线平滑度上有所改善,但在极值点的拟合上仍有偏差。相比之下,本文提出的模型与真实值的贴合度最高,特别是在第 50–60 h 的高频波动区间,该模型能够敏锐捕捉到负荷的瞬时突变,展现出卓越的局部特征提取能力。

4.5.2 预测曲线拟合度分析图

基于前文定义的 4 项关键评价指标,本文设计

了包含单一模型、基础组合模型及深度学习混合模型在内的共计 8 组对比实验方案,并在统一的测试集上进行了严格的性能测试。各模型的具体预测性能指标如表 2 所示。此外,为消除不同量纲对误差评估的影响,表 2 中的误差指标(NMSE 值, NMAE 值, NRMSE 值)均基于归一化后的数据代入式(6)–(8)计算所得,并以百分比形式展示;而 R^2 则保持不变。表 2 不仅详细列出了各模型在上述 4 个维度上的量化评估结果,还通过直观的数据对比,清晰展示了从单一 Informer 模型到引入 VMD 分解、再到融合 BiGRU 模块的过程中,模型预测精度的逐步演进规律。

表 2 各模型性能指标对比

Tab. 2 Comparative analysis of performance metrics

across models					%
预测模型	NMSE 值	NMAE 值	NRMSE 值	R^2 值	
Informer	0.18	3.35	4.22	76.05	
Informer-BiGRU	0.17	3.30	4.15	76.83	
VMD-BiGRU	0.14	3.04	3.75	81.07	
VMD-Transformer	0.29	4.52	5.41	60.54	
VMD-Transformer-BiGRU	0.12	2.87	3.51	83.37	
VMD-Informer	0.28	4.31	5.24	62.97	
VMD-Informer-BiLSTM	0.16	3.19	4.00	78.40	
VMD-Informer-BiGRU	0.12	2.64	3.34	84.48	

如表 2 所示,本文提出的组合模型在所有评价指标上均表现最优。其 R^2 达到 84.48%, NRMSE 值降至 3.34%。与单一 Informer 模型相比, R^2 值提升了 8.43 个百分点, NRMSE 值降低了 0.88 个百分点,验证了“分解-重构-集成”策略的有效性。

在 Informer 基础上串联 BiGRU 后, R^2 从 76.05% 提升至 76.83%, NRMSE 值从 4.22% 降至 4.15%。这表明 BiGRU 有效弥补了 Informer 在提取局部短时依赖特征上的不足。

在同为 VMD-Informer 架构下, BiGRU 显著优于 BiLSTM, 且 NMAE 值降低了 0.55%。这说明针对乡村负荷这类非平稳序列,结构更精简的 BiGRU 网络在特征提取效率和泛化能力上更具优势。

本文通过在 VMD-Informer 后引入 BiGRU, 成功将 R^2 提升至全场最高的 84.48%。证明了 BiGRU 在强化局部细节捕捉、整合 VMD 多尺度特征方面的关键“桥梁”作用,而 VMD 的多尺度解耦则实现了高频噪声的平滑,最终达成各组件优势的互补。

相比于基准的 VMD-Transformer 模型, 本文提出的 VMD-Informer-BiGRU 模型的 NMAE 值从 4.52 % 显著降低至 2.64 %, 精度相对提升了 41.6 %。标志着本文所提出的模型在捕捉非线性负荷波动方面实现了显著的性能增益, 使残差分布更加集中于零误差基准线附近。考虑到 160~300 kW 的高负荷基数, 哪怕是 1% 的预测偏差也意味着数千瓦的电力缺口; 因此, 误差的进一步缩减能有效预防关键农事阶段的变压器过载跳闸, 为电网提供更高的安全决策裕度。同时, 高精度预测还能大幅降低系统的旋转备用冗余, 在优化运行经济性的同时, 最大程度释放分布式新能源的消纳潜力。

VMD-Transformer-BiGRU 和 VMD-Informer-BiGRU 具有同等组合架构, 使用 Informer 模块的 R^2 略高于使用 Transformer 模块, 且 NRMSE 值更低。这验证了 Informer 的概率稀疏注意力机制在处理长序列依赖时, 比传统 Transformer 具有更高的计算精度和特征捕捉能力。

综上所述, VMD-Informer-BiGRU 模型通过 VMD 的多尺度解耦、Informer 的全局长程建模以及 BiGRU 的局部细节优化, 实现了对乡村综合能源系统多元负荷的高精度预测。

4.6 模型计算复杂度与运行效率分析

为评估所提模型的复杂性与实际应用能力, 本文从参数规模 (Params)、计算复杂度 (FLOPs) 及时间开销等方面进行了对比实验与分析, 具体结果见表 3。

表 3 各模型复杂度与计算开销对比

Tab. 3 Comparison of model complexity and computational cost

预测模型	Params 值/M	FLOPs 值/M	训练时间 /(s·epoch ⁻¹)	推理时间 /ms
Informer	0.08	1.62	0.67	1.42
Informer-BiGRU	0.13	2.21	0.80	1.79
VMD-BiGRU	0.03	0.76	0.24	0.23
VMD-Transformer	0.09	1.63	0.78	1.34
VMD-Transformer-BiGRU	0.14	2.22	0.91	1.84
VMD-Informer	0.09	1.63	0.79	1.46
VMD-Informer-BiLSTM	0.15	2.42	0.97	1.72
VMD-Informer-BiGRU	0.14	2.13	0.86	1.63

由表 3 可知, 所提模型的参数量为 0.14 M, FLOPs 为 2.13 M。相较于基础的 VMD-Informer 模型, 虽因引入双向循环结构导致开销小幅增加, 但

与同等组合架构的 VMD-Informer-BiLSTM 相比, 得益于 BiGRU 更精简的门控机制, 本文模型在参数量和计算量上均更为轻量, 整体复杂度增长完全可控。

在运行效率方面, 所提模型单个 epoch 的训练时间为 0.86 s, 相较于同类多模块结构如 VMD-Transformer-BiGRU 的 0.91 s 具备更高的训练效率; 在推理阶段, 模型的单样本预测时间仅为 1.63 ms, 达到了毫秒级的响应速度。

综上所述, 所提模型在实现高精度预测的同时, 将计算开销控制在合理范围内, 能够切实满足乡村综合能源系统短期负荷预测的在线运行需求。

4.7 预测偏差分布与稳健性分析

为进一步验证 VMD-Informer-BiGRU 模型的稳健性并深度解析指标间的内在联系, 本节给出了模型的预测误差概率分布图。如图 9 所示, 预测残差呈现显著的正态分布特性, 概率密度峰值紧邻零误差基准线, 且平均误差仅为 -1.056 9 kW。这证明模型在整体预测上不存在明显的系统性偏差, 具有良好的数理统计稳健性。

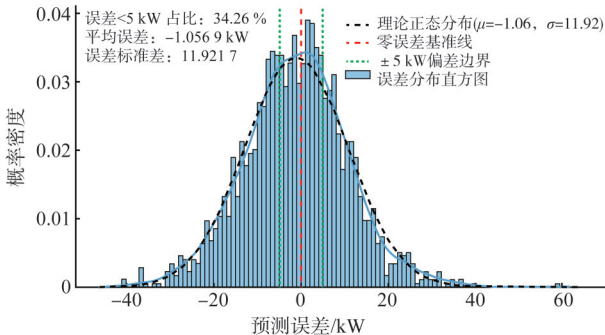


图 9 预测模型误差概率分布图

Fig. 9 Probability distribution plot of prediction model errors

针对乡村负荷在 160~300 kW 剧烈波动区间带来的预测挑战, 本模型表现出了优异的高基数偏差控制能力。统计结果表明, 约 34.26 % 的样本绝对误差严格控制在 5 kW 即对应约 1.6 %~3.1 % 的相对误差以内, 这一分布特征从微观层面支撑了表 2 中 NMAE 值低至 2.64 % 的评估结论。

此外, 通过观察图 9 两侧的“长尾”离群点可以发现, 部分偏差超过 20 kW 的样本点主要对应于图 8 中由大功率农机随机启停导致的极端脉冲波动。由于决定系数 R^2 值对残差平方项具有高度敏感性, 这些极少数反映乡村负荷强随机、非平稳特性的离

群点是限制 R^2 值进一步提升的主要物理根源。综上所述, 误差分布的集中性证明了模型在绝大多数运行工况下均能提供高精度的负荷预报。

5 结论

本文针对乡村综合能源系统负荷序列固有的非平稳性及多维耦合特征, 提出了一种基于 VMD-Informer-BiGRU 的短期负荷组合预测方法。通过算例分析, 得出以下结论。

1) 利用 VMD 技术将非平稳的原始负荷解耦为多个频率稳定的模态分量, 有效降低了序列的随机波动干扰, 将复杂的预测问题转化为多个相对平稳子序列的预测问题, 为后续深度学习模型的精准建模创造了有利条件。

2) 构建的 Informer-BiGRU 组合架构实现了“全局-局部”协同建模: Informer 利用概率稀疏自注意力机制高效捕捉序列的全局长程依赖, 而 BiGRU 则从双向维度精准提取局部突变特征, 两者结合有效解决了乡村负荷多时间尺度耦合的难题。

3) 基于中国某乡村实际数据的算例分析表明, 相较于 Informer、VMD-BiGRU 等对比模型, 本文所提模型在多项评价指标上均表现最优, 证实了其在解决乡村短期负荷预测问题上的有效性与鲁棒性。

参考文献

- [1] 陆佳琳, 方舒, 杜松怀, 等. 高比例可再生能源新型农村电网的技术挑战与展望[J]. 电力需求侧管理, 2022, 24(6): 38-43.
LU Jialin, FANG Shu, DU Songhuai, et al. Technical challenges and expectation of new rural power grid with high proportion of renewable energy [J]. Power Demand Side Management, 2022, 24(6): 38-43.
- [2] 王保民. 乡村振兴战略下能源电力面临的机遇与挑战[J]. 中国科技信息, 2022(7): 124-125.
WANG Baomin. Opportunities and challenges faced by energy and power under the rural revitalization strategy [J]. China Science and Technology Information, 2022(7): 124-125.
- [3] 董京, 何武, 刘雪飞, 等. 基于合作联盟的农村微电网光储协同配置与收益分配策略[J]. 电力自动化设备, 2025, 45(7): 10-18, 53.
DONG Jing, HE Wu, LIU Xuefei, et al. Cooperative allocation and profit distribution strategy of photovoltaic-storage systems in rural microgrids based on cooperative alliance [J]. Electric Power Automation Equipment, 2025, 45(7): 10-18, 53.
- [4] 韩璟琳, 胡平, 韩天华, 等. 考虑多能互动的农村区域电网负荷预测方法[J]. 电力系统及其自动化学报, 2024, 36(1): 152-158.
HAN Jinglin, HU Ping, HAN Tianhua, et al. Load forecasting method for rural regional power grids considering multi-energy interaction [J]. Proceedings of the CSU-EPSCA, 2024, 36(1): 152-158.
- [5] 梁睿, 金沫含, 朱慧君, 等. 基于多场景聚类 and 可解释时间融合 Transformer 网络的乡村地区净负荷区间预测[J]. 电网技术, 2025, 49(12): 5019-5027.
LIANG Rui, JIN Mohan, ZHU Huijun, et al. Net load interval forecasting for rural areas based on multi-scenario clustering and interpretable temporal fusion Transformer network [J]. Power System Technology, 2025, 49(12): 5019-5027.
- [6] 陶磊, 罗萍萍, 林济铿. 基于深度学习的直流微电网虚假数据注入攻击二阶段检测方法[J]. 中国电力, 2024, 57(9): 11-19.
TAO Lei, LUO Pingping, LIN Jikeng. Two-stage detection method for DC microgrid false data injection attack based on deep learning [J]. Electric Power, 2024, 57(9): 11-19.
- [7] 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.
- [8] 高岩, 吴汉斌, 张纪欣, 等. 基于组合深度学习的光伏功率日前概率预测模型[J]. 中国电力, 2024, 57(4): 100-110.
GAO Yan, WU Hanbin, ZHANG Jixin, et al. Day-ahead probabilistic prediction model for photovoltaic power based on combined deep learning [J]. Electric Power, 2024, 57(4): 100-110.
- [9] YANG W, SHI J, LI S, et al. A combined deep learning load forecasting model of single household resident user considering multi-time scale electricity consumption behavior [J]. Applied Energy, 2022(307): 118197.
- [10] SEKHAR C, DAHIYA R. Robust framework based on hybrid deep learning approach for short term load forecasting of building electricity demand [J]. Energy, 2023(268): 126660.
- [11] DUBEY A K, KUMAR A, GARCÍA-DÍAZ V, et al. Study and analysis of SARIMA and LSTM in forecasting time series data [J]. Sustainable Energy Technologies and Assessments, 2021(47): 101474.
- [12] BELMAHDI B, LOUZAZNI M, EL BOUARDI A. One month-ahead forecasting of mean daily global solar radiation using time series models [J]. Optik, 2020(219): 165207.
- [13] JIANG L, WANG X, LI W, et al. Hybrid multitask multi-information fusion deep learning for household short-term load forecasting [J]. IEEE Transactions on Smart Grid, 2021, 12(6): 5362-5372.
- [14] LI K, HUANG W, HU G, et al. Ultra-short term power load forecasting based on CEEMDAN-SE and LSTM neural network [J]. Energy and Buildings, 2023(279): 112666.
- [15] LIAO C, CHEN C, XIANG C, et al. Taxi-passenger's destination prediction via GPS embedding and attention-based BiLSTM model [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(5): 4460-4473.
- [16] 李延珍, 王海鑫, 杨子豪, 等. 基于非侵入式负荷分解的家庭负荷两阶段超短期负荷预测模型[J]. 电工技术学报, 2024, 39(11): 3379-3391.
LI Yanzhen, WANG Haixin, YANG Zihao, et al. A two-stage

- ultra-short-term household load forecasting model based on non-intrusive load decomposition [J]. Transactions of China Electrotechnical Society, 2024, 39(11): 3379 – 3391.
- [17] LAKSHMI D V, KUMAR K. Integrated LSTM and reinforcement learning framework for predictive power forecasting and energy optimization in battery management systems[J]. Journal of Energy Storage, 2025(138): 1186661.1 – 118666.13.
- [18] ZHAO Q, WANG S, CHEN Y, et al. A hybrid framework for short-term load forecasting based on optimized InMetra Boost and BiLSTM[J]. Energy, 2025(328): 136582.1 – 136582.10.
- [19] MICHAEL N E, HASAN S, AL-DURRA A, et al. Short-term solar irradiance forecasting based on a novel bayesian optimized deep long short-term memory neural network[J]. Applied Energy, 2022(324): 119727.1 – 119727.10.
- [20] WANG Y, CHEN J, CHEN X, et al. Short-term load forecasting for industrial customers based on TCN-LightGBM [J]. IEEE Transactions on Power Systems, 2021, 36(3): 1984 – 1997.
- [21] CHEN Q, LIU Y B, GE M F, et al. A novel Bayesian-optimization-based adversarial TCN for RUL prediction of bearings[J]. IEEE Sensors Journal, 2022, 22(21): 20968 – 20977.
- [22] ZHAN Y, YAN K, ZHENG X. Integrating transformers into physics-informed neural networks: An approach to lithium-ion battery state-of-health prognostics [J]. International Journal of Electrical Power & Energy Systems, 2025(172): 111173.1 – 111173.12.
- [23] XIA X, LUO Y, LI P, et al. Systematic evaluation of transformer-based time series forecasting models for post-processing WRF-simulated wind speed and predicting short-term power output[J]. Applied Energy, 2026(403): 127070.1 – 127070.10.
- [24] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. Advances in Neural Information Processing Systems, 2017, 30.
- [25] LI C, SHI J. A novel CNN-LSTM-based forecasting model for household electricity load by merging mode decomposition, self-attention and autoencoder [J]. Energy, 2025 (330): 136883.1 – 136883.11.
- [26] CHUNG W H, GU Y H, YOO S J. District heater load forecasting based on machine learning and parallel CNN-LSTM attention[J]. Energy, 2022(246): 123350.1 – 123350.13.
- [27] NIU D, YU M, SUN L, et al. Short-term multi-energy load forecasting for integrated energy systems based on CNN-BiGRU optimized by attention mechanism [J]. Applied Energy, 2022(313): 118801.1 – 118801.9.
- [28] 张未, 余成波, 王士彬, 等. 基于VMD-LSTM-LightGBM的多特征短期电力负荷预测[J]. 南方电网技术, 2023, 17(2): 74 – 81.
- ZHANG Wei, YU Chengbo, WANG Shibin, et al. Multi-feature short-term power load forecasting based on VMD-LSTM-LightGBM [J]. Southern Power System Technology, 2023, 17(2): 74 – 81.
- [29] 王俊, 王继烨, 程坤, 等. 基于双层优化VMD-LSTM的农村超短期电力负荷预测[J]. 沈阳农业大学学报, 2024, 55(1): 92 – 102.
- WANG Jun, WANG Jiye, CHENG Kun, et al. Ultra-short-term power load forecasting based on two-layer optimization VMD-LSTM [J]. Journal of Shenyang Agricultural University, 2024, 55(1): 92 – 102.
- [30] GONG M, ZHAO Y, SUN J, et al. Load forecasting of district heating system based on Informer [J]. Energy, 2022 (253): 124179.1 – 124179.11.
- [31] 周璇, 李可昕, 郭子轩, 等. 基于改进Informer的商业建筑短期用电负荷多步预测[J]. 华南理工大学学报(自然科学版), 2026, 54(1): 42 – 52.
- ZHOU Xuan, LI Kexin, GUO Zixuan, et al. Short-term power load multi-step forecasting for commercial building based on improved informer [J]. Journal of South China University of Technology (Natural Science Edition), 2026, 54(1): 42 – 52.
- [32] 俞胜, 孙可, 蔡华, 等. 结合极端梯度提升决策树与改进Informer的短期电力负荷预测方法[J]. 中国电力, 2025, 58(10): 195 – 205.
- YU Sheng, SUN Ke, CAI Hua, et al. A short-term power load forecasting method combining extreme gradient boosting decision tree with an improved informer [J]. Electric Power, 2025, 58(10): 195 – 205.

收稿日期: 2025-11-21; 网络首发日期: 2026-06-17

作者简介:

杨浚文(1985),男,高级工程师,硕士,从事电网规划、系统运行分析相关工作, junwenyang_1@163.com;

徐志(1984),男,通信作者,高级工程师(教授级),硕士,从事智能配电网与电能质量、直流输电与电力电子应用相关工作, xuzhi1123@163.com;

姜詠(1993),男,工程师,硕士,从事电能质量分析与控制相关工作, jh_yndky@outlook.com。