

引用格式:王炳铮,岳云涛,李炳华,等.基于GRA/EEMD-Informer混合模型的光储直柔配电系统多数据预测方法[J].电力科学与技术学报,2024,39(3):86-95.
Citation: WANG Bingzheng, YUE Yuntao, LI Binghua, et al. Multi-data prediction method based on GRA/EEMD-Informer hybrid model for photovoltaic-storage-direct-flexible distribution system[J]. Journal of Electric Power Science and Technology, 2024, 39(3): 86-95.

基于 GRA/EEMD-Informer 混合模型的光储直柔 配电系统多数据预测方法

王炳铮¹, 岳云涛¹, 李炳华², 万珊珊¹

(1.北京建筑大学电气与信息工程学院,北京 100044;2.悉地(北京)国际建筑设计顾问有限公司,北京 100013)

摘要:针对现有时间序列模型预测光储直柔配电系统短期发电用电数据精度不高的问题,提出一种基于灰色关联度分析/集合经验模态分解(grey relation analysis, ensemble empirical mode decomposition, GRA/EEMD-Informer)的光储直柔配电系统多数据预测模型,通过灰色关联度分析、模态分解,结合自注意力蒸馏机制,有效捕捉输出和输入之间较精确的长程相关性耦合,降低了时空复杂度,极大缓解了传统编解码的局限性。将已建成并投入使用的光伏发电站某月数据、典型办公建筑某月电力数据及电动汽车充电站运行数据作为原始数据,以均方误差、平均绝对误差、均方根误差作为评价指标对模型进行检验,并进行消融实验与分析,最后与长短期记忆网络(long short-term memory, LSTM)、基于粒子群优化(particle swarm optimization, PSO)算法的长短期记忆网络(PSO-LSTM)、Transformer时间序列预测方法对比,结果表明该方法的拟合程度明显高于其他预测方法,验证了GRA/EEMD-Informer算法对提高预测能力的有效性和实用性。

关键词:光储直柔;GRA/EEMD-Informer;光伏发电输出功率;建筑电力负荷;电动汽车负荷

DOI: 10.19781/j.issn.1673-9140.2024.03.010 **中图分类号:** TM71 **文章编号:** 1673-9140(2024)03-0086-10

Multi-data prediction method based on GRA/EEMD-Informer hybrid model for photovoltaic-storage-direct-flexible distribution system

WANG Bingzheng¹, YUE Yuntao¹, LI Binghua², WAN Shanshan¹

(1.School of Electricity and Information Engineering, Beijing University of Civil Engineering and Architecture, Beijing 100044, China;

2.CADG Group, Beijing 100013, China)

Abstract: Addressing the issue of low accuracy in predicting short-term power generation and consumption data for the photovoltaic-storage-direct-flexible distribution system using existing time series models, a multi-data prediction model based on grey relation analysis (GRA), ensemble empirical mode decomposition (EEMD), and Informer, namely the GRA/EEMD-Informer is proposed. This model effectively captures the precise long-range correlation coupling between outputs and inputs through grey relation analysis, modal decomposition, combined with a self-attention distillation mechanism. It reduces spatiotemporal complexity and significantly alleviates the limitations of traditional encoding and decoding methods. Using data from a photovoltaic power station, the electricity consumption of typical office building, and an electric vehicle charging station for a certain month as the original data, the model is tested using evaluation metrics such as mean squared error (MSE), mean absolute error (MAE), and root mean squared error (RMSE). Ablation experiments and analyses are conducted, and the results are compared with those of long short-term memory (LSTM), particle swarm optimization-based LSTM (PSO-LSTM), and the transformer time series prediction method. The results indicate that the proposed method exhibits significantly higher fitting accuracy than other prediction

收稿日期:2022-11-06;修回日期:2024-04-07

基金项目:国家自然科学基金(61902016);北京建筑大学研究生创新科研项目(07081022003)

通信作者:岳云涛(1971—),男,博士,副教授,主要从事控制理论与控制工程等研究;E-mail:yueyuntao@bucea.edu.cn

methods, verifying the effectiveness and practicality of the GRA/EEMD-Informer algorithm in enhancing prediction capabilities.

Key words: PEDF; GRA/EEMD-Informer; photovoltaic power output; building power load; electric vehicle load

太阳能光伏系统因其具有用地成本低、传输损耗少等优点成为被重点推广的建筑可再生能源发电方式^[1-2]。从产能总量上看,光伏系统可能会在一定程度上缓解用电难的问题,但产能与用能时序上的不匹配仍会给电网带来极大负担,影响电网正常运行^[3-4]。针对该问题清华大学江亿院士提出了光储直柔(photovoltaic, energy storage, DC, flexible, PEDF)新型配电系统发展思路,通过需求侧响应的用电模式提高能源利用率,其中源—荷预测数据是后期储能容量配置及柔性调度策略的基础。

因此,出于PEDF配电系统自身及运营者的需求,需对光伏输出功率、建筑电力负荷、电动汽车负荷进行短期预测,通过短期电力规划在较大范围调节从电网的取电量^[5]。目前,学者对于预测问题有较多实验和理论研究,可分为持续、统计和学习模型。持续性模型假定在相似日输出功率不发生变化。常见的统计学方法有聚类分析法、主成分分析法等,适合中长期预测,对已有数据的平稳性要求很高,较难反映非线性影响^[6-9]。机器学习和深度学习是近几年常用方法。机器学习模型包含人工神经网络,但负荷序列具有随机波动性等特点,通过经验获得的特征会导致数据缺失或者冗余,影响预测精度^[10-11]。深度学习是预测领域中重要的研究方向^[12]。典型方法有卷积神经网络(convolutional neural network, CNN)、循环神经网络(recurrent neural network, RNN)等。长短期记忆网络(long short-term memory, LSTM)是一种特殊的RNN,相比普通的RNN, LSTM能够在长序列中有更好表现,解决了长序列训练过程中的梯度消失和梯度爆炸问题,因此很多学者提出基于LSTM的混合模型进行负荷预测^[13-18]。但RNN属于序列模型,需要以排队序列化方式处理信息,注意力权重需要等序列全部输入模型之后才能确定,这会造成大量时间开销,难以实现并行处理^[19-21]。而Transformer模型是一种基于注意力(attention)机制来加速深度学习算法的模型,可以进行并行化计算,但其也受到二次时间复杂度、高内存占用率以及编码器—解码器(encoder-decoder)模型固有体系的限制^[22-24]。

LSTM、门控循环单元(gate recurrent unit, GRU)、Transformer等基于深度学习和神经网络的

预测模型在某些时间序列预测任务上已经取得了不错的成果。然而,时间序列预测仍然是一个具有挑战性的问题,因为模型的准确度取决于多种因素,如数据质量、模型选择、模型参数设置以及数据量等。因此,为了获得更准确的预测结果,需要对PEDF配电系统的具体数据进行综合考虑和评估。在实践中,时间序列预测模型在某些数据集上表现出色,但在其他数据集上则可能表现不佳,因此不能对时间序列预测模型的准确度一概而论^[24-26]。

针对以上问题,为提高预测精度及效率,提出基于灰色关联度分析/集合经验模态分解(grey relation analysis, ensemble empirical mode decomposition, informer, GRA/EEMD)-Informer混合模型的时间序列预测方法,对PEDF配电系统多数据进行预测,以均方误差、平均绝对误差和均方根误差作为模型性能评价指标,并与LSTM、基于粒子群优化(particle swarm optimization, PSO)算法的长短期记忆网络(PSO-LSTM)、Transformer、Informer模型进行对比,验证GRA/EEMD-Informer模型的有效性。

1 GRA/EMD-Informer预测方法研究

1.1 原始数据及处理

本研究采用3类数据集进行实验,分别为2018年6月整月光伏电站数据、典型办公建筑电力数据,以及2022年8月整月电动汽车充电站数据。光伏电站数据采用辐射量、组件温度、环境温度、相对湿度等因素作为输入变量,办公建筑电力数据采用独立台区采样区间的有功功率、无功功率和温度作为负荷特征,电动汽车充电站运行数据包括三相功率、每个阶段和总计的无功功率、视在功率及功率均值等。每组数据集共2 880组测量值,其中光伏电站、办公建筑运行数据为某单位非公开数据集,电动汽车充电站运行数据为自测数据集。

1.1.1 异常数据检测及替换

原始数据的随机噪声和异常数据等因素会导致预测结果产生误差,本研究对数据进行包括异常数据检测与替换、标准化及数据集划分处理。异常数据的检测采用 3σ 统计判别法,设定置信概率极限值,识别采集的原始数据中某个时间点的数据与前后时间点差值是否超限^[27]。首先选取 N 天原始数

据,一天中所有序列数据为 $\{x_1, x_2, x_3, \dots, x_i\}$,对应某天的第 i 时刻的测量值,分别计算数学期望 $E(i)$ 和标准差 $\sigma(i)$,则第 N 天 i 时刻数据的偏移率为

$$\xi(N, i) = \frac{|E(i) - X(N, i)|}{\sigma(i)} \quad (1)$$

通常一天内的时间点功率值是在 $(E(i) - 3\sigma(i), E(i) + 3\sigma(i))$ 范围内,如不在此范围则判定为异常点,将采取平均值法将异常点替换 x ,即

$$\bar{X}(N, i) = \frac{X(N-1, i) + X(N+1, i)}{2} \quad (2)$$

1.1.2 标准化处理

本研究所采用的原始数据存在极端最大、最小情况,因此选用标准差标准化 StandardScaler 类,对序列 $x_1, x_2, x_3, \dots, x_n$ 进行变换,将数据以列为组,按其属性删除平均值和缩放到单位方差来标准化特征。通过计算训练集的平均值和标准差,将其作为总体的均值和方差,转化函数为

$$x^* = \frac{x - \mu}{\sigma} \quad (3)$$

式中, μ 为样本数据均值; σ 为标准差。最终每个属性的所有数据都聚集在0附近,使新数据集变换为均值为0、标准差为1分布。

1.1.3 数据集划分

对新数据集划分为训练集、测试集及验证集,设定70%数据为训练集,分别设定1、3、7 d即96、288、672个数据点为测试集,剩余部分作为各类数据集的验证集。

1.2 GRA/EMD-Informer 时间序列预测模型

为提高PEDF配电系统多数据预测的精度,提出了GRA/经验模态分解(empirical mode decomposition, EMD)-Informer混合预测模型,削弱了光伏气象因素变量的空间复杂程度,剔除掉不必要的冗余信息,同时减弱了建筑电力负荷和电动汽车充电负荷序列的非线性。其中,应用于PEDF配电系统的GRA/EMD-Informer多数据预测模型流程如图1所示。首先采用GRA关联度模型根据自变量及因变量之间的相似程度筛选最优自变量^[27]。2个因素间关联度越接近1,表明影响程度越大。具体计算步骤包括:确定反映系统行为特征的参考数列和影响系统行为的比较数列;对光伏电站原始数据进行无量纲化处理,将各列气象数据分别与实时光伏功率值求差并取绝对值,计算参考数列与比较数列的关联系数。根据差值分别计算各气象因素和光伏功率的关联系数 ξ_i ,并计算关联度 γ ,最后将各个时刻的关联系数集中起来求其平均值,再进行排序和筛选。

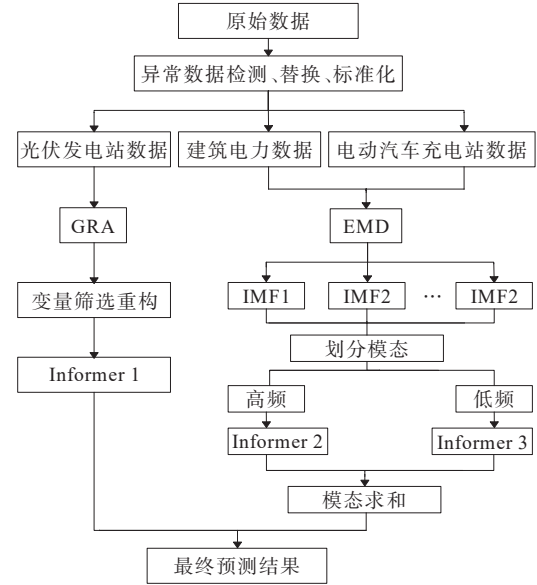


图1 GRA/EMD-Informer预测模型流程

Figure 1 Flowchart of predictive model based on GRA/EMD-Informer

计算过程为

$$x_k = |x_0(t_k) - x_i(t_k)| \quad (4)$$

$$\zeta_i = \frac{D_{\min_2} + \rho \cdot D_{\max_2}}{x_k + \rho \cdot D_{\max_2}} \quad (5)$$

$$\gamma_i = \frac{1}{n} \sum_{k=1}^n \xi_i \quad (6)$$

式中, D_{\min_2}, D_{\max_2} 为两级最小差和两级最大差, $D_{\min_2} = \min_i \min_k |x_0(t_k) - x_i(t_k)|$, $D_{\max_2} = \max_i \max_k |x_0(t_k) - x_i(t_k)|$; ρ 为可调节系数。削减变量后采用EMD模型将目标序列进行分解重构^[16-18]。根据原始数据序列 $x(t)$,寻找到所有极大值、极小值点,并采用三次样条插值函数拟合出原始数据的上、下包络线 $U_s(t), V_s(t)$ 并相减求出均值包络线 $m(t)$,再使用原始目标序列减去均值包络线得到中间信号 $c(t)$,判断 $c(t)$ 是否满足在整段数据内极值点个数和过零点个数的绝对差值小于1及任意时刻由极大值点形成的上包络线和由极小值点形成的下包络线的平均值为零。若仍存在负局部极大值和正局部极小值,说明这还不是一个本征模函数(intrinsic mode function, IMF),需要继续重复以上步骤进行筛选。当 $c(t)$ 满足以上条件即为一个IMF分量,令 $c(t) = I_{MF1}$,再用原始信号与 I_{MF1} 相减作为新的原始信号。重复以上步骤,直至到第 n 阶IMF分量或残差 $r_n(t)$ 是单调函数或者常量时完成分解,最终得到 n 个IMF分量和一个残差 $r(t)$ 。计算过程如下:

$$m(t) = \frac{U_s(t) + V_s(t)}{2} \quad (7)$$

$$c(t) = x(t) - m(t) \quad (8)$$

$$r(t) = x(t) - c(t) \quad (9)$$

$$x(t) = \sum_{i=1}^n c(t) + r_n(t) \quad (10)$$

根据模态分解结果将高、低频信号分量组合重构,通过Informer模型进行稀疏自注意力、自蒸馏机制、一步生成式解码得到预测结果。首先将自注意力机制写成概率形式^[28-29],即

$$\text{Attention}(q_i, K, V) = \sum_j \frac{k(q_i, k_j)}{\sum_l k(q_i, k_l)} V_j = E_{p(k|q)}[v_j] \quad (11)$$

式中, q_i, k_i, v_i 分别为矩阵 Q, K, V 的第 i 行, $p(k_i|q_i) = \frac{k(q_i, k_i)}{\sum_l k(q_i, k_l)}$, $k(q_i, k_i) = \exp\left(\frac{q_i k_i^T}{\sqrt{d}}\right)$, d 为输入维度。

再通过KL散度建立衡量标准,对query稀疏性进行度量与评价,挑选出得分最高的精英项,计算其query和key点积结果,其余lazy queries将自注意力层的输入直接取均值填充,保证每个概率稀疏自注意力(ProSparse self-attention)层输入和输出序列长度一致^[22]。具体计算过程为

$$K_L(q||p) = \ln \sum_{i=1}^{L_K} e^{\frac{q \cdot k_i^T}{\sqrt{d}}} - \frac{1}{L_K} \sum_{j=1}^{L_K} \frac{q \cdot k_j^T}{\sqrt{d}} - \ln L_K \quad (12)$$

$$M(q_i, K) = \ln \sum_{j=1}^{L_K} e^{\frac{q \cdot k_j^T}{\sqrt{d}}} - \frac{1}{L_K} \sum_{j=1}^{L_K} \frac{q \cdot k_j^T}{\sqrt{d}} \quad (13)$$

$$\overline{M}(q_i, K) = \max_j \left\{ \frac{q \cdot k_j^T}{\sqrt{d}} \right\} - \frac{1}{L_K} \sum_{j=1}^{L_K} \frac{q \cdot k_j^T}{\sqrt{d}} \quad (14)$$

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{\overline{Q} K^T}{\sqrt{d_{\text{key}}}}\right) V \quad (15)$$

式(12)~(15)中, K_L 为衡量相似性的散度; M 为稀疏度量; \overline{M} 为 M 的均值; L_K 为常数; \overline{Q}, Q 为同样大小的稀疏矩阵。通过编码器运用蒸馏机制抽取鲁棒的长序列输入的长期依赖,对有支配特征的优势特征进行特权化,在相邻的注意块之间加入卷积池化操作,对特征进行降采样。从 j 层到 $(j+1)$ 层的蒸馏过程为^[28-29]

$$X'_{j+1} = \text{MaxPool}\left(\text{ELU}\left(\text{convld}\left([X'_j]_{\text{AB}}\right)\right)\right) \quad (16)$$

式中, $[]_{\text{AB}}$ 表示多头自注意机制; convld 表示卷积滤波器在时间维度上的ELU激活函数。解码器由

2个相同的多注意层组成,输入中包括含有时间戳信息的开始序列及占位符,将其送入解码器之后,占位符中得到目标序列^[29]。其中,输入解码器中的向量可表示为

$$X(\text{de}) = \text{Concat}(X(\text{token}), X(0)) \in \mathbb{R}^{(L_{\text{token}} + L_y) \times d_{\text{model}}} \quad (17)$$

式中, $X(\text{token}) \in \mathbb{R}^{L_{\text{token}} \times d_{\text{model}}}$ 为输出序列的开始标记; $X(0) \in \mathbb{R}^{L_y \times d_{\text{model}}}$ 为目标序列的占位符,置为0。

解码器的隐藏注意力机制不会选择特定标记作为开始序列,利用生成式推理一步生成目标维度序列,一个全连接层获得最终输出。

1.3 预测误差评价指标

采用均方误差 e_{MSE} 、平均绝对误差 e_{MAE} 及均方根误差 e_{RMSE} 作为 GRA/EMD-Informer 预测结果误差评估指标,具体计算公式为

$$e_{\text{MSE}} = \frac{1}{n} \left(\sum_{i=1}^n y_i - \overline{y_i} \right)^2 \quad (18)$$

$$e_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^n |y_i - \overline{y_i}| \quad (19)$$

$$e_{\text{RMSE}} = \sqrt{\frac{1}{n} \left(\sum_{i=1}^n y_i - \overline{y_i} \right)^2} \quad (20)$$

式(18)~(20)中, n 为预测结果的数量; $y_i, \overline{y_i}$ 分别是第 i 个采样点的实际数据和预测数据。

2 PEDF 多数据预测结果与分析

2.1 GRA/EMD-Informer 模型预测结果

按电力需求预测的时间分类,短期预测的预测期一般为日预测和周预测。本研究通过 GRA/EMD-Informer 模型对光伏输出功率、建筑电力负荷、电动汽车负荷3个数据集进行训练,通过调整输入批尺寸(batch size)、迭代次数(epoch)、学习率(learning rate)确定最佳模型的参数,并逐步延长预测时间范围,分别在96、288、672个点(1、3、7 d)中选择最佳预测结果,如表1所示。

通过表1的误差评价指标结果可以看出, GRA/EMD-Informer 模型在电力负荷这种具有较稳定周期性的数据上会表现得更好,而对于光伏输出功率及电动汽车负荷,可能是由于特征维数预测各向异性及波动过大而造成误差指标偏高。不过整体来看,模型可以获得较好的预测误差指标结果,并且随着预测时间延长,误差可以控制为在一定范围内波动。为进一步分析与验证模型的预测结果,接下来结合各数据拟合曲线对仿真结果进行消融实验。

表1 GRA/EMD-Informer多数据集预测结果

Table 1 Multi-data prediction results based on GRA/EMD-Informer

| 数据集 | 误差评价 指标 | 数据点/个 | | |
|----------|------------|---------|----------|----------|
| | | 96(1 d) | 288(2 d) | 672(3 d) |
| 光伏发电功率预测 | e_{MSE} | 0.413 | 0.300 | 0.991 |
| | e_{MAE} | 0.426 | 0.412 | 0.682 |
| | e_{RMSE} | 0.643 | 0.548 | 0.995 |
| 建筑电力负荷预测 | e_{MSE} | 0.112 | 0.142 | 0.136 |
| | e_{MAE} | 0.252 | 0.298 | 0.283 |
| | e_{RMSE} | 0.336 | 0.378 | 0.369 |
| 电动汽车负荷预测 | e_{MSE} | 0.533 | 0.659 | 0.589 |
| | e_{MAE} | 0.575 | 0.644 | 0.615 |
| | e_{RMSE} | 0.730 | 0.812 | 0.767 |

2.2 消融实验结果与分析

2.2.1 GRA消融实验

为验证 GRA/EMD-Informer 的可行性,采用 Informer 模型对光伏电站的一周输出功率进行预测,在同等条件下,通过 GRA-Informer 模型对一个月的光伏发电站数据进行关联度计算并进行变量筛选。关联系数代表该子序列辐射、组件温度、相对湿度、环境温度、气压对与母序列实际输出功率对应维度上的关联程度值,数字越大代表关联性越强。实验结果表明,辐射、组件温度、相对湿度、环境温度、气压与实际输出功率的关联度分别为 0.520、0.770、0.769、0.760、0.096,其中与实际输出功率关联度最大的是组件温度,与实际输出功率关联度最小的是气压。因此将辐射、组件温度、相对湿度、环境温度作为最终的输入特征变量再次输入模型进行预测,Informer 及 GRA-Informer 预测结果如图 2、3 所示。

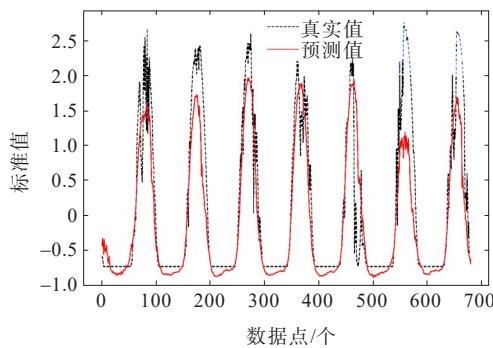


图2 Informer 光伏输出功率预测结果

Figure 2 PV output power prediction results based on informer

通过对比可以看出,Informer 模型预测结果的大体趋势较为准确,但峰值预测不尽如人意,其

672个点的预测结果的 e_{MSE} 、 e_{MAE} 、 e_{RMSE} 分别为 1.020、0.651、1.010。GRA-Informer 模型在维持初始学习率 0.000 1, epoch 为 5 等设置条件下,预测结果曲线拟合相对较好,其 e_{MSE} 、 e_{MAE} 、 e_{RMSE} 分别为 0.991、0.682、0.995,其 e_{MSE} 相对于 Informer 模型下降约 2.84%,模型精度有所提高,且从预测曲线来看,无论在预测趋势或峰谷值识别情况都有较好的表现。

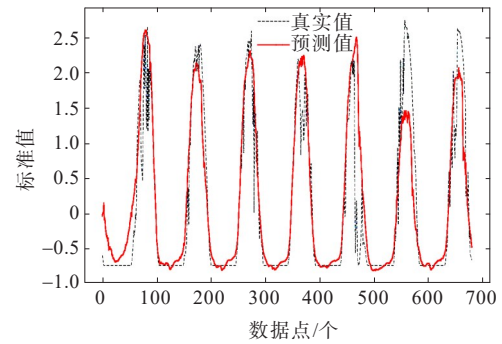


图3 GRA-Informer 光伏输出功率预测结果

Figure 3 PV output power prediction results based on GRA-Informer

2.2.2 负荷重构消融实验

通过 EMD-Informer 模型将办公建筑电力数据与电动汽车充电站原始信号模态分解为含有高、低频的正弦间歇性信号,两类数据都被分解为 9 个 IMF 和 1 个残差,9 个 IMF 反映了数据在不同影响因素和不同尺度的分布,逐渐由非平稳到平稳,残差反映序列长期的变化规律^[11]。同时计算 9 个模态子序列在月信号长度下的过零率,得到了电力负荷数据的高频分量索引包括 $I_{MF1} \sim I_{MF5}$ 及残差,低频 IMF 分量索引包括 $I_{MF6} \sim I_{MF9}$;电动汽车充电站功率数据高频分量索引包括 $I_{MF1} \sim I_{MF6}$ 及残差,低频 IMF 分量索引包括 $I_{MF7} \sim I_{MF9}$ 。分别将两类原始数据高频分量与残差叠加,低频分量叠加,结合原数据特征变量重构数据集,2 类数据集原始数据、高、低频及整体趋势分别如图 4(a)、(b) 所示。

通过图 4 可以看出,2 类数据集的高频部分包含更多的周期性信息,具有较强规律性,而低频部分包含的非周期性信息更多,通过 EMD-Informer 对 2 类高、低频重构数据集分别预测发现,2 类数据集的低频部分拟合精度比较好,与实际变化曲线较为接近。2 类数据集的误差指标 e_{MSE} 、 e_{MAE} 分别为 0.230、0.061, 0.100、0.363。对于高频部分,与实际趋势也较为吻合,其中 e_{MSE} 、 e_{MAE} 分别为 0.430、0.410, 0.770、0.604。对高、低频预测结果叠加后,计算 2 类数据集的 e_{MSE} 、 e_{MAE} 、 e_{RMSE} 分别为 0.326、

0.451、0.571、0.758、0.698、0.870。从整体来看,电力负荷数据集EMD-Informer模型比Informer模型的 e_{MSE} 和 e_{MAE} 更低,在672个点的预测长度下分别降低了11.17%、5.45%,证明了EMD-Informer在预测精度方面的优越性,但该模型对于电动汽车负荷预测的准确性并没有明显变化,这可能是方法或数据集的原因,需要进一步实验。

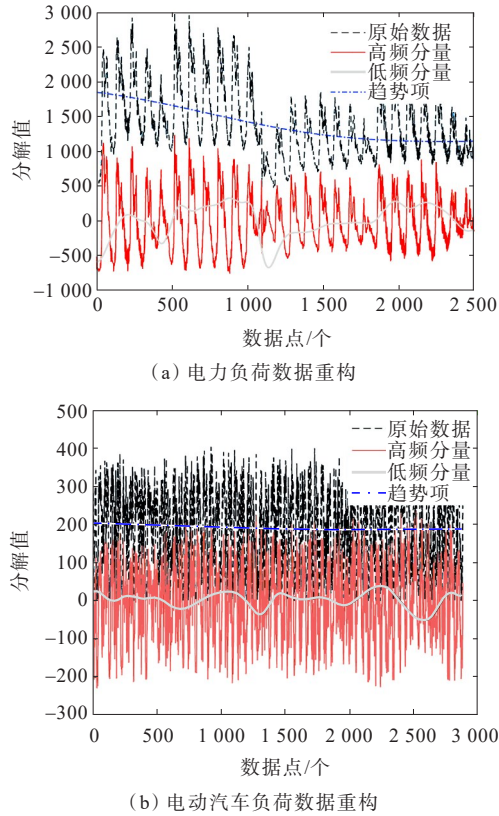


图4 重构数据的高、低频及趋势

Figure 4 High and low frequency and trend of reconstructed data

2.3 改进与误差分析

2.3.1 基于辅助噪声的模型改进

对EMD-Informer模型进行测试和验证,发现通过EMD的频率数据存在高、低频组合分量模式混叠问题,导致后半部分分量没有实际物理意义。针对以上问题,采用一种辅助噪声方法对EMD进行改进:首先定义原始信号 $x(t)$ 和噪声标准差 σ ;对原始信号 $x(t)$ 加入一个正态分布随机白噪声 $r(t)$,得到带噪声的信号 $s(t)=x(t)+r(t)$;利用式(4)~(10)对 $s(t)$ 进行EMD,同样得到一组固有模式函数和一个残差。重复以上2个步骤,每次加入新的正态分布白噪声序列,直到得到50组分解结果;针对每组分解结果,将对应的IMF组合起来,得到一组IMF。最后将每次得到的IMF进行集成平均处理得到最终结果。改进两类数据分解效果如图5、6所示。

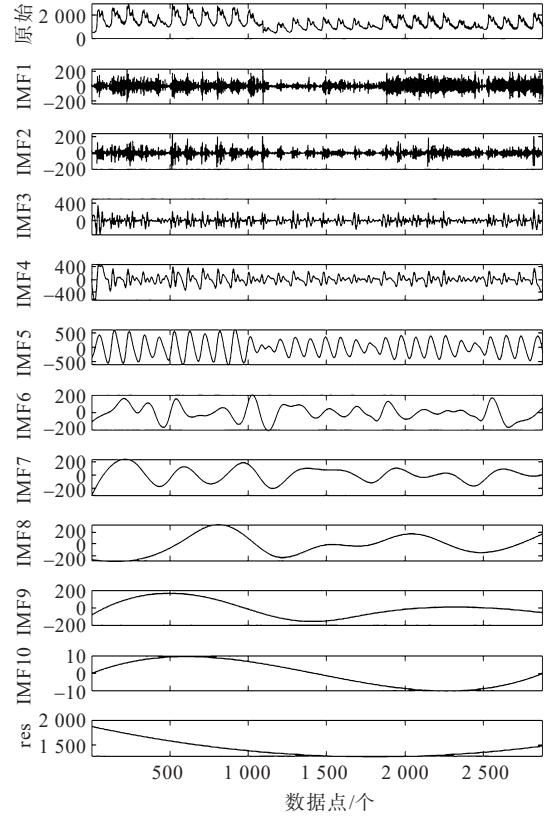


图5 办公建筑电力负荷EEMD结果

Figure 5 Decomposition results of office building power load based on EEMD

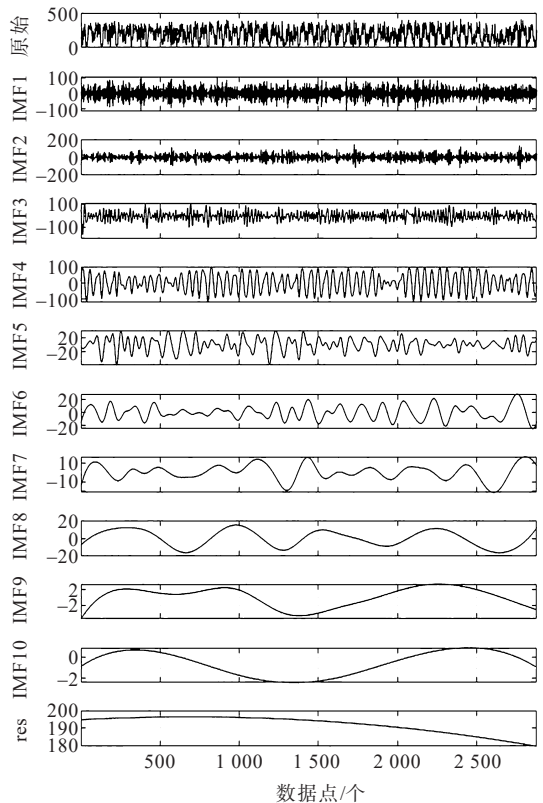


图6 电动汽车负荷EEMD结果

Figure 6 Decomposition results of electric vehicle load based on EEMD

筛选高、低频与重构数据集的流程与 EMD-Informer 相同,不再赘述。通过改进的 EMD-Informer 即 EEMD-Informer 再次对两类即 4 个数据集分别进行预测,结果表明无论是高频还是低频部分,两类数据集的拟合性能更强。对高、低频预测结果叠加后计算两类数据集的 e_{MSE} 、 e_{MAE} 、 e_{RMSE} 分别为 0.136、0.283、0.369,0.589、0.615、0.167。对于建

筑电力数据集,该模型 e_{MSE} 和 e_{MAE} 进一步降低,最优结果在 672 个点的预测长度下分别降低了 58.28%、40.67%。同时,该模型对于电动汽车负荷预测的精度较 Informer 也有了显著提高, e_{MSE} 和 e_{MAE} 最优结果分别降低了 20.60%、10.74%。

2.3.2 实际误差分析

消融实验所得到的误差指标结果如表 2~4 所示,

表 2 消融实验 1—光伏输出功率预测结果对比
Table 2 Ablation experiment 1—comparison of predicted PV output power

| 方法 | 批尺寸 | 数据点/个 | | | | | | | | |
|------------------|-----|-----------|-----------|------------|-----------|-----------|------------|-----------|-----------|------------|
| | | 96(1 d) | | | 288(3 d) | | | 672(7 d) | | |
| | | e_{MSE} | e_{MAE} | e_{RMSE} | e_{MSE} | e_{MAE} | e_{RMSE} | e_{MSE} | e_{MAE} | e_{RMSE} |
| Informer | 24 | 0.607 | 0.695 | 0.779 | 0.506 | 0.628 | 0.712 | 1.020 | 0.680 | 1.010 |
| | 48 | 0.768 | 0.762 | 0.876 | 0.579 | 0.643 | 0.761 | 1.013 | 0.652 | 1.006 |
| | 96 | 0.428 | 0.480 | 0.654 | 0.553 | 0.599 | 0.744 | 1.068 | 0.654 | 1.033 |
| GRA/EMD-Informer | 24 | 0.477 | 0.586 | 0.690 | 0.300 | 0.412 | 0.548 | 0.991 | 0.682 | 0.995 |
| | 48 | 0.560 | 0.642 | 0.748 | 0.318 | 0.433 | 0.564 | 1.001 | 0.703 | 1.000 |
| | 96 | 0.413 | 0.426 | 0.643 | 0.436 | 0.393 | 0.660 | 1.052 | 0.607 | 1.025 |

表 3 消融实验 2—电力负荷预测结果对比
Table 3 Ablation experiment 2—comparison of power load prediction

| 方法 | 批尺寸 | 数据点/个 | | | | | | | | |
|-------------------|-----|-----------|-----------|------------|-----------|-----------|------------|-----------|-----------|------------|
| | | 96(1 d) | | | 288(3 d) | | | 672(7 d) | | |
| | | e_{MSE} | e_{MAE} | e_{RMSE} | e_{MSE} | e_{MAE} | e_{RMSE} | e_{MSE} | e_{MAE} | e_{RMSE} |
| Informer | 24 | 0.363 | 0.423 | 0.602 | 0.359 | 0.481 | 0.599 | 0.367 | 0.477 | 0.606 |
| | 48 | 0.399 | 0.487 | 0.632 | 0.390 | 0.511 | 0.625 | 0.423 | 0.506 | 0.650 |
| | 96 | 0.478 | 0.535 | 0.692 | 0.455 | 0.562 | 0.674 | 0.564 | 0.592 | 0.751 |
| GRA/EMD-Informer | 24 | 0.256 | 0.415 | 0.506 | 0.234 | 0.389 | 0.484 | 0.480 | 0.590 | 0.693 |
| | 48 | 0.305 | 0.471 | 0.552 | 0.274 | 0.436 | 0.524 | 0.601 | 0.667 | 0.775 |
| | 96 | 0.303 | 0.461 | 0.550 | 0.262 | 0.428 | 0.512 | 0.326 | 0.451 | 0.571 |
| GRA/EEMD-Informer | 24 | 0.112 | 0.252 | 0.336 | 0.142 | 0.298 | 0.378 | 0.136 | 0.283 | 0.369 |
| | 48 | 0.194 | 0.327 | 0.441 | 0.205 | 0.346 | 0.453 | 0.233 | 0.367 | 0.483 |
| | 96 | 0.212 | 0.345 | 0.460 | 0.209 | 0.345 | 0.458 | 0.281 | 0.399 | 0.530 |

表 4 消融实验 2—电动汽车负荷预测结果对比
Table 4 Ablation experiment 2—comparison of predicted electric vehicle load

| 方法 | 批尺寸 | 数据点/个 | | | | | | | | |
|-------------------|-----|-----------|-----------|------------|-----------|-----------|------------|-----------|-----------|------------|
| | | 96(1 d) | | | 288(3 d) | | | 672(7 d) | | |
| | | e_{MSE} | e_{MAE} | e_{RMSE} | e_{MSE} | e_{MAE} | e_{RMSE} | e_{MSE} | e_{MAE} | e_{RMSE} |
| Informer | 24 | 0.699 | 0.667 | 0.836 | 0.676 | 0.658 | 0.822 | 0.742 | 0.689 | 0.861 |
| | 48 | 0.847 | 0.744 | 0.920 | 0.698 | 0.660 | 0.835 | 0.858 | 0.759 | 0.926 |
| | 96 | 0.878 | 0.761 | 0.937 | 0.795 | 0.733 | 0.891 | 0.862 | 0.754 | 0.928 |
| GRA/EMD-Informer | 24 | 0.807 | 0.739 | 0.898 | 0.701 | 0.664 | 0.837 | 0.758 | 0.698 | 0.870 |
| | 48 | 0.831 | 0.725 | 0.911 | 0.716 | 0.665 | 0.846 | 0.872 | 0.757 | 0.933 |
| | 96 | 0.877 | 0.758 | 0.936 | 0.767 | 0.691 | 0.876 | 0.840 | 0.741 | 0.917 |
| GRA/EEMD-Informer | 24 | 0.533 | 0.575 | 0.730 | 0.659 | 0.644 | 0.812 | 0.589 | 0.615 | 0.767 |
| | 48 | 0.656 | 0.637 | 0.810 | 0.673 | 0.648 | 0.820 | 0.598 | 0.611 | 0.773 |
| | 96 | 0.817 | 0.705 | 0.903 | 0.706 | 0.658 | 0.840 | 0.638 | 0.632 | 0.799 |

可以看出在预测偏移量增大、预测总长度增加的情况下, GRA/EEMD-Informer模型预测的精度更高,泛化能力更强。但对于电力负荷和电动汽车负荷预测结果,确实会由于一些不可抗拒因素,如天气因素、经济因素、节假日因素、大用户设备因素、数据传输因素等造成数据预测结果与实际结果存在一定误差。

2.4 多种模型预测性能分析

所提模型与 LSTM、PSO-LSTM、Transformer 时间序列方法进行性能对比实验,划分相同训练集,设置多种预测时长,对各网络模型参数调优,以同样误差指标结果进行评价,得到结果如表5所示。除本研究提出的模型外,较为优秀的预测结果加粗标记。

表5 各预测方法预测误差比较

Table 5 Comparison of prediction error among different prediction methods

| 方法 | 预测时间 范围/d | 光伏电站数据 | | | 建筑电力负荷数据 | | | 电动汽车负荷数据 | | |
|-------------------|--------------|--------------|--------------|------------|-----------|--------------|------------|--------------|--------------|------------|
| | | e_{MSE} | e_{MAE} | e_{RMSE} | e_{MSE} | e_{MAE} | e_{RMSE} | e_{MSE} | e_{MAE} | e_{RMSE} |
| GRA/EEMD-Informer | 1 | 0.413 | 0.300 | 0.991 | 0.112 | 0.142 | 0.136 | 0.533 | 0.659 | 0.589 |
| | 3 | 0.426 | 0.412 | 0.682 | 0.252 | 0.298 | 0.283 | 0.575 | 0.644 | 0.615 |
| | 7 | 0.643 | 0.548 | 0.995 | 0.336 | 0.378 | 0.369 | 0.73 | 0.812 | 0.767 |
| Transformer | 1 | 0.610 | 0.338 | 1.029 | 0.271 | 0.175 | 0.271 | 0.570 | 0.688 | 0.636 |
| | 3 | 0.498 | 0.464 | 0.680 | 0.391 | 0.325 | 0.388 | 0.591 | 0.650 | 0.641 |
| | 7 | 0.781 | 0.581 | 1.014 | 0.520 | 0.419 | 0.520 | 0.755 | 0.829 | 0.797 |
| PSO-LSTM | 1 | 0.429 | 0.518 | 1.729 | 0.241 | 0.193 | 1.729 | 0.752 | 0.856 | 0.819 |
| | 3 | 0.431 | 0.503 | 1.546 | 0.363 | 0.315 | 1.032 | 0.652 | 0.673 | 0.522 |
| | 7 | 0.654 | 0.719 | 1.314 | 0.490 | 0.439 | 1.314 | 0.867 | 0.925 | 0.905 |
| LSTM | 1 | 0.585 | 0.709 | 1.253 | 0.927 | 0.907 | 1.787 | 1.336 | 1.533 | 1.656 |
| | 3 | 0.578 | 0.626 | 0.795 | 0.613 | 0.664 | 0.895 | 1.092 | 0.794 | 0.872 |
| | 7 | 0.764 | 0.842 | 1.119 | 0.962 | 0.952 | 1.336 | 1.156 | 1.238 | 1.287 |

通过表5可以发现,所列时间序列方法对不同数据集间敏感程度不一,且不同预测数目和不同数据集间的预测结果波动较大,其中 PSO-LSTM 在 1 d 内预测相对稳定一些,Transformer 在 1、3 d 内的预测也有较好表现。不过将 GRA/EEMD-Informer 预测结果与其他算法在同条件下的结果进行对比,可以发现其误差及峰谷值偏移程度皆小于其他 3 种经典算法,因此从数据集的综合适用性、误差指标及峰谷值偏移程度来看,本研究提出的 GRA/EEMD-Informer 预测方法始终可以取得较高的预测精度,且在各种工况下对比其他模型预测精度至少可以提高 4%。总的来说,相比于其他 3 种预测方法,该方法在处理长序列和多尺度信息方面表现更好,具有更好的并行计算能力。

3 结语

本文结合 PEDF 的研究现状和混合模型思想提出一种 GRA/EEMD-Informer 时间序列预测模型,并与单一模型和其他方法所构建的模型进行对比,

得出如下结论。

1) 通过消融实验,对 GRA/EMD-Informer 添加辅助噪声改进,结果验证了 GRA/EEMD-Informer 预测模型可以有效抵抗样本中的噪声,具有较强的鲁棒性和可靠性。

2) GRA/EEMD-Informer 模型在同条件下与 LSTM、PSO-LSTM、Transformer 方法进行对比,具有更高的预测精度,更适合非线性复杂问题。

3) 本文所提的 GRA/EEMD-Informer 时间序列模型通过 3 个与 PEDF 配电系统的相关数据集进行 1、3、7 d 内的验证,可以将误差控制在较低的范围,预测精度较高,工程实用性强。

4) 在较长编码器输入时具有更丰富的信息及依赖性,使误差进一步降低。未来随着建筑数据量的增加,输入长序列更容易获得好的预测结果。

本文将计算机学科的深度学习方法引入 PEDF 配电系统中,基于历史数据得到高精度的短期预测结果,可以用于设计针对不同工况下的负荷自适应调节策略,对推进建筑配电系统电力规划有较大实际意义。

参考文献:

- [1] FONTENOT H, DONG B. Modeling and control of building-integrated microgrids for optimal energy management-A review[J]. *Applied Energy*, 2019, 254: 113689.
- [2] 郜捷,宋洁,王剑晓,等.支撑中国能源安全的电氢耦合系统形态与关键技术[J].*电力系统自动化*,2023,47(19): 1-15.
GAO Jie, SONG Jie, WANG Jianxiao, et al. Form and key technologies of integrated electricity-hydrogen system supporting energy security in China[J]. *Automation of Electric Power Systems*, 2023, 47(19): 1-15.
- [3] 郭怿,明波,黄强,等.考虑输电功率平稳性的水—风—光—储多能互补日前鲁棒优化调度[J].*电工技术学报*, 2023, 38(9): 2350-2363.
GUO Yi, MING Bo, HUANG Qiang, et al. Day-ahead robust optimal scheduling of hydro-wind-PV-storage complementary system considering the steadiness of power delivery[J]. *Transactions of China Electrotechnical Society*, 2023, 38(9): 2350-2363.
- [4] 吴晓刚,阎洁,葛畅,等.基于改进GRU-CNN的风光水一体化超短期功率预测方法[J].*中国电力*, 2023, 56(9): 178-186+205.
WU Xiaogang, YAN Jie, GE Chang, et al. Ultra-short-term power forecasting method for wind-solar-hydro integration based on improved GRU-CNN[J]. *Electric Power*, 2023, 56(9): 178-186+205.
- [5] INTEHA A, Nahid-Al-Masood, HUSSAIN F, et al. A data driven approach for day ahead short term load forecasting[J]. *IEEE Access*, 2022, 10: 84227-84243.
- [6] 董志强,郑凌蔚,苏然,等.一种基于IGWO-SNN的光伏出力短期预测方法[J].*电力系统保护与控制*, 2023, 51(1): 131-138.
DONG Zhiqiang, ZHENG Lingwei, SU Ran, et al. An IGWO-SNN-based method for short-term forecast of photovoltaic output[J]. *Power System Protection and Control*, 2023, 51(1): 131-138.
- [7] 方鹏,高亚栋,潘国兵,等.基于LSTM神经网络的中长期光伏电站发电量预测方法研究[J].*可再生能源*, 2022, 40(1): 48-54.
FANG Peng, GAO Yadong, PAN Guobing, et al. Research on forecasting method of mid-and long-term photovoltaic power generation based on LSTM neural Network[J]. *Renewable Energy Resources*, 2022, 40(1): 48-54.
- [8] 王海燕,刘佳康,邓亚平.基于预估—校正综合BP神经网络的短期光伏功率预测[J].*智慧电力*, 2023, 51(3): 46-52.
WANG Haiyan, LIU Jiakang, DENG Yaping. Short-term photovoltaic power forecasting based on predict-correct combination BP neural network[J]. *Smart Power*, 2023, 51(3): 46-52.
- [9] 邹文进,郝少飞,马刚,等.基于CEEMD-GA-BP神经网络的风光发电功率预测[J].*电网与清洁能源*, 2022, 38(3): 111-118.
ZOU Wenjin, HAO Shaofei, MA Gang, et al. Forecast of wind and solar power generation based on CEEMD-GA-BP neural network[J]. *Power System and Clean Energy*, 2022, 38(3): 111-118.
- [10] 王金玉,金宏哲,王海生,等. ISSA优化Attention双向LSTM的短期电力负荷预测[J].*电力系统及其自动化学报*, 2022, 34(5): 111-117.
WANG Jinyu, JIN Hongzhe, WANG Haisheng, et al. Short-term power load prediction of bidirectional LSTM with ISSA optimization attention mechanism[J]. *Proceedings of the CSU-EPSA*, 2022, 34(5): 111-117.
- [11] 王英立,陶帅,候晓晓,等.基于MIV分析的GA-BP神经网络光伏短期发电预测[J].*太阳能学报*, 2020, 41(8): 236-242.
WANG Yingli, TAO Shuai, HOU Xiaoxiao, et al. GA-BP neural network photovoltaic power generation short-term forecast based on miv analysis[J]. *Acta Energetica Solaris Sinica*, 2020, 41(8): 236-242.
- [12] 李毓洋,王媛媛,罗晓敏.基于卷积神经网络算法的高压发电机保护研究[J/OL].*长沙理工大学学报(自然科学版)*, 1-8[2024-03-12]. <https://doi.org/10.19951/j.cnki.1672-9331.20221116002>.
LI Yuyang, WANG Yuanyuan, LUO Xiaomin. Research on high-voltage generator protection based on CNN[J/OL]. *Journal of Changsha University of Science & Technology(Natural Science)*, 1-8[2024-03-12]. <https://doi.org/10.19951/j.cnki.1672-9331.20221116002>.
- [13] LI Q Q, XU Y, CHEW B S H, et al. An integrated missing-data tolerant model for probabilistic PV power generation forecasting[J]. *IEEE Transactions on Power Systems*, 2022, 37(6): 4447-4459.
- [14] 阳曾,丁施尹,叶萌,等.基于变分模态分解和深度学习的短期电力负荷预测模型[J].*电测与仪表*, 2023, 60(2): 126-131+146.
YANG Zeng, DING Shiyin, YE Meng, et al. Short-term load forecasting model based on VMD and LSTM[J]. *Electrical Measurement & Instrumentation*, 2023, 60(2): 126-131+146.
- [15] 周文,孟良,杨正富,等.基于时间序列的大型光伏电站发电短期预测[J].*电源技术*, 2021, 45(11): 1490-1494.
ZHOU Wen, MENG Liang, YANG Zhengfu, et al. Short-term forecasting method of large-scale photovoltaic power station based on time series[J]. *Chinese Journal of Power Sources*, 2021, 45(11): 1490-1494.

- [16] 徐岩,向益锋,马天祥.基于EMD-CNN-LSTM混合模型的短期电力负荷预测[J].华北电力大学学报(自然科学版),2022,49(2):81-89.
XU Yan, XIANG Yifeng, MA Tianxiang. Short-term power load forecasting method based on EMD-CNN-LSTM hybrid model[J]. Journal of North China Electric Power University (Natural Science Edition), 2022, 49(2): 81-89.
- [17] 刘文杰,刘禾,王英男,等.基于完整自适应噪声集成经验模态分解的LSTM-Attention网络短期电力负荷预测方法[J].电力建设,2022,43(2):98-108.
LIU Wenjie, LIU He, WANG Yingnan, et al. Short-term power load forecasting method based on CEEMDAN and LSTM-Attention network[J]. Electric Power Construction, 2022, 43(2): 98-108.
- [18] FAN C D, DING C K, ZHENG J H, et al. Empirical mode decomposition based multi-objective deep belief network for short-term power load forecasting[J]. Neurocomputing, 2020, 388: 110-123.
- [19] 王福忠,王帅峰,张丽.基于VMD-LSTM与误差补偿的光伏发电超短期功率预测[J].太阳能学报,2022,43(8):96-103.
WANG Fuzhong, WANG Shuaifeng, ZHANG Li. Ultra short term power prediction of photovoltaic power generation based on vmd-lstm and error compensation[J]. Acta Energiae Solaris Sinica, 2022, 43(8): 96-103.
- [20] 张震,李孟洲,李浩方,等.基于VMD-LSTM-MLR的短期电力负荷预测[J].水电能源科学,2021,39(10):208-212.
ZHANG Zhen, LI Mengzhou, LI Haofang, et al. Short-term power load forecasting based on VMD-LSTM-MLR[J]. Water Resources and Power, 2021, 39(10): 208-212.
- [21] 王俊杰,毕利,张凯,等.基于多特征融合和XGBoost-LightGBM-ConvLSTM的短期光伏发电量预测[J].太阳能学报,2023,44(7):168-174.
WANG Junjie, BI Li, ZHANG Kai, et al. Short-term photovoltaic power generation prediction based on multi-feature fusion and XGBoost-LightGBM-ConvLSTM[J]. Acta Energiae Solaris Sinica, 2023, 44(7): 168-174.
- [22] LIU Y J, MENG F D, ZHOU J, et al. Faster depth-adaptive transformers[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(15): 13424-13432.
- [23] LI Y Y, LIN Y, XIAO T, et al. An efficient transformer decoder with compressed sub-layers[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(15): 13315-13323.
- [24] ZHOU H Y, ZHANG S H, PENG J Q, et al. Informer: beyond efficient transformer for long sequence time-series forecasting[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(12): 11106-11115.
- [25] HAN Z Y, ZHAO J, LEUNG H, et al. A review of deep learning models for time series prediction[J]. IEEE Sensors Journal, 2021, 21(6): 7833-7848.
- [26] ARUNKUMAR K E, KALAGA D V, MOHAN SAI KUMAR C, et al. Comparative analysis of gated recurrent units (GRU), long short-term memory (LSTM) cells, autoregressive integrated moving average (ARIMA), seasonal autoregressive integrated moving average (SARIMA) for forecasting COVID-19 trends[J]. Alexandria Engineering Journal, 2022, 61(10): 7585-7603.
- [27] 班斓.基于长短记忆神经网络的短期光伏功率预测研究[D].南宁:广西大学,2021.
BAN Lan. Research on short-term photovoltaic power prediction based on long-term memory neural network [D]. Nanning: Guangxi University, 2021.
- [28] MA C B, ZHANG P, SONG F, et al. KD-informer: a cuff-less continuous blood pressure waveform estimation approach based on single photoplethysmography[J]. IEEE Journal of Biomedical and Health Informatics, 2023, 27(5): 2219-2230.
- [29] LU W, GAO L F, LI Z B, et al. Prediction of long-term elbow flexion force intervals based on the informer model and electromyography[J]. Electronics, 2021, 10(16): 1946.