

DOI: 10.19783/j.cnki.pspc.240538

基于级联 MCNN-MMLP 双残差网络的短期负荷预测

余凯峰¹, 吐松江·卡日¹, 张紫薇², 马小晶¹, 王志刚¹

(1. 新疆大学电气工程学院, 新疆 乌鲁木齐 830049; 2. 清华四川能源互联网研究院, 四川 成都 610213)

摘要: 为了解决负荷特性复杂导致负荷预测精度低的问题, 提出了一种 GWO-VMD 和级联 MCNN-MMLP 双残差网络的短期负荷预测模型。首先, 利用由灰狼算法(grey wolf optimize, GWO)优化的变分模态分解(variational mode decomposition, VMD)对原始负荷数据进行处理, 降低原始负荷数据的复杂程度。其次, 使用多尺度卷积神经网络(multiscale convolutional neural networks, MCNN)和多层感知机(multi-layer perceptron, MLP)结合的双残差神经网络对各个模态进行迁移学习训练和预测, 并在 MLP 网络中引入多头注意力机制弥补网络信息瓶颈问题。最后, 再次使用 MCNN-MMLP 双残差模型对初步预测的误差进行预测并校正初值, 从而进一步提升预测精确度。通过对实际负荷数据进行分析, 本模型的均方误差为 5.024 (MW)^2 、均方根误差为 2.241 MW 、平均绝对百分比误差为 0.160% , 决定系数为 0.996 , 各性能指标均优于其他传统及智能负荷预测方法。

关键词: 负荷预测; 多尺度卷积神经网络; 双残差神经网络; 多头注意力机制; 迁移学习

Short-term load forecasting based on a cascade MCNN-MMLP double residual network

YU Kaifeng¹, TUSONGJIANG·Kari¹, ZHANG Ziwei², MA Xiaojing¹, WANG Zhigang¹

(1. School of Electrical Engineering, Xinjiang University, Urumqi 830049, China;

2. Sichuan Energy Internet Research Institute, Tsinghua University, Chengdu 610213, China)

Abstract: There is an issue of decreased accuracy in load forecasting due to the complexity of load characteristics, so a short-term load forecasting model based on GWO-VMD and cascaded MCNN-MMLP residual networks is proposed. First, the original load data is processed using variational mode decomposition (VMD) optimized by grey wolf optimization (GWO) to reduce the complexity of the original load data. Secondly, a dual residual neural network combining multiscale convolutional neural networks (MCNN) and a multi-layer perceptron (MLP) is employed for transfer learning training and prediction of each mode. A multi-head attention mechanism is introduced into the MLP network to address the information bottleneck issue. Lastly, the MCNN-MMLP double residual model is used to predict and correct the errors in the preliminary prediction, thereby further improving the accuracy of the forecast. The proposed model outperforms the traditional and intelligent load forecasting models. It achieves a mean square error of 5.024 MW^2 , root mean square error of 2.241 MW , mean absolute percentage error of 0.160% , and coefficient of determination of 0.996 based on actual load data.

This work is supported by the National Natural Science Foundation of China (No. 52067021 and No. 52207165).

Key words: load forecasting; multiscale convolutional neural network; double residual neural network; multi-head attention mechanism; transfer learning

0 引言

作为电力行业数字化转型的关键环节, 电力负

荷预测扮演重要角色。在电力市场中, 电力负荷预测为电力市场竞价和现货价格计算提供依据^[1-2]。精准的电力负荷预测有利于平衡电力市场供需关系, 从而保障电力系统和电力市场稳定高效运行^[3-4]。同时精确的负荷预测可以有效降低发电成本和碳排放量, 成为实现“双碳”目标的重要手段之一^[5-6]。目前, 国内外学者开展了大量电力负荷预测方面的研究工作。根据时间跨度, 可将电力负荷预测的研究

基金项目: 国家自然科学基金项目资助(52067021, 52207165); 新疆维吾尔自治区自然科学基金面上项目资助(2022D01C35); 新疆维吾尔自治区优秀青年科技人才培养项目资助(2019Q012)

分为长期负荷预测、中期负荷预测^[7]和短期负荷预测^[8-9]。而短期负荷预测对电网经济调度、用电规划和有效支持电网实时运行有指导作用^[10]。根据研究的预测模型,短期电力负荷预测方法分为传统预测方法^[11]、现代预测方法和组合预测方法^[12]三大类。

传统预测方法基于数理统计模型,如峰值负荷模型法、回归分析法^[13]、最小二乘法^[14]、灰色模型法等。文献[15]采用模糊多元线性回归模型进行假日负荷短期预测,但缺乏自主学习能力,对复杂非线性系统的预测精度较低。

在基于人工智能的现代预测方法中,许多研究者利用传统机器学习中的支持向量机^[16]、极限学习机^[17-18]、BP神经网络(back propagation neural network, BPNN)^[19]来进行预测。有研究人员利用优化算法对支持向量机(support vector machine, SVM)进行优化。文献[20]采用粒子群算法对SVM参数进行优化,提高了短期预测精度。随着深度学习的发展,卷积神经网络^[21](convolutional neural network, CNN)、长短期记忆神经网络^[22](long short-term memory, LSTM)、门控循环单元神经网络等深度学习模型都运用到负荷预测领域。文献[23]采用自适应局部迭代滤波和LSTM神经网络在深度学习架构中结合的方法实现短期预测。但为提高预测精度,自适应局部迭代滤波和LSTM神经网络都需要调整众多参数。

组合预测模型指通过不同模型和算法进行加权与融合形成的预测模型。针对复杂的负荷数据,在模型预测之前将负荷分解为多个子序列,然后将每个序列的预测结果叠加,得到最终的预测值^[24]。文献[25]首先将数据分解算法与神经网络相结合,然后通过经验模态分解将原始功率负荷分解为单个子组件,最后利用堆栈式LSTM神经网络分别对不同组件进行预测。通过实例验证其组合模型的预测精度较高,但是堆栈式LSTM网络的复杂性让参数训练变得更加困难且计算成本高,模型的预测精度难以进一步提升。文献[26]提出了基于特征筛选的CNN和双向长短期记忆网络(bidirectional long short-term memory, BiLSTM)相结合的模型,该模型首先对多维输入数据选取高相关性特征参数,然后使用CNN网络有效地提取特征信息,最后利用BiLSTM模型进行训练,以挖掘电力负荷数据中的时序性特点。通过实验验证,CNN-BiLSTM混合模型比单一模型具有更高的预测精度。然而在文献[27]中,多元模型的准确性一般会优于单变量模型,但在训练和测试中不能保证外生数据的可靠性。

综上所述,本文提出了一种灰狼算法优化的变

分模态分解(grey wolf optimize and variational mode decomposition, GWO-VMD)与级联多尺度卷积神经网络和多头感知机(multiscale convolutional neural networks and multi-layer perception with multi-head attention, MCNN-MMLP)双残差网络的短期负荷预测模型。首先,利用GWO对变分模态分解中模态分量个数 K 和二次惩罚因子 α 进行优化,降低原始负荷数据的复杂性。其次,利用MCNN提取分解时序数据特征,随后将提取到的特征输入MLP神经网络进一步提取特征。将MLP输出的一部分与上一部分输出进行分解重构作为下一个MCNN-MLP模块的输入,另外一部分直接预测输出。然后,为了解决MLP网络中的信息瓶颈问题,引入了多头注意力机制到网络中。最后,使用MCNN-MMLP双残差网络模型对初步预测误差进行预测并校正初值,以提高预测精确度。实验结果表明,本文方法能够有效地降低负荷数据复杂性,并提供更准确的负荷预测结果。同时为降低计算成本和节约资源,本文采用迁移学习对各模态进行训练,即将前一个模态训练好的预测模型(源模型)知识迁移到下一个模态预测模型(目标模型)。

1 理论基础

1.1 变分模态分解理论

变分模态分解是一种非递归的信号分解方法,用于将复杂信号分解为具有不同频率和能量的多个模态。将原始负荷序列 f 分解为一系列有限带宽模态函数 $\{u_k(t)\}, k=1,2,\dots,K$,其主要公式如下所述。

1) 利用各个模态函数 $u_k(t)$ 的单侧频谱与包含中心频率 ω_k 的指数项 $e^{-j\omega_k t}$ 混叠,从而将频谱调制到对应基频带,如式(1)所示。

$$[(\delta(t) + \frac{j}{\pi t}) \cdot u_k(t)] e^{-j\omega_k t} \quad (1)$$

式中, $\delta(t)$ 为狄拉克函数。

2) 利用高斯平滑法估计出各模态信号带宽,求解带约束条件的变分问题,如式(2)所示。

$$\begin{cases} \min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t [(\delta(t) + \frac{j}{\pi t}) \cdot u_k(t)] e^{-j\omega_k t} \right\|^2 \right\} \\ \text{s.t. } \sum_k u_k = f \end{cases} \quad (2)$$

式中, ω_k 表示各个模态对应的中心频率, $\{\omega_k\} = \{\omega_1, \omega_2, \dots, \omega_K\}$ 。

1.2 灰狼优化算法

灰狼算法是受灰狼群体的协作和竞争行为启发,其核心思想是将待优化问题映射为一个多维空

间中的搜索空间, 其中每个灰狼个体表示一个潜在的解, 并通过模拟灰狼的行为来搜索最优解。

灰狼算法数学模型如式(3)、式(4)所示。

$$\begin{cases} \mathbf{D}_\alpha = |C_1 \cdot \mathbf{Z}_\alpha - Z| \\ \mathbf{D}_\beta = |C_2 \cdot \mathbf{Z}_\beta - Z| \\ \mathbf{D}_\delta = |C_3 \cdot \mathbf{Z}_\delta - Z| \end{cases} \begin{cases} Z_1 = \mathbf{Z}_\alpha - A_1 \cdot \mathbf{D}_\alpha \\ Z_2 = \mathbf{Z}_\beta - A_2 \cdot \mathbf{D}_\beta \\ Z_3 = \mathbf{Z}_\delta - A_3 \cdot \mathbf{D}_\delta \end{cases} \quad (3)$$

$$Z(t+1) = \frac{Z_1 + Z_2 + Z_3}{3} \quad (4)$$

式中: \mathbf{D}_α 、 \mathbf{D}_β 、 \mathbf{D}_δ 分别表示三只头狼 α 、 β 、 δ 与其他种群个体之间的方向向量; Z 表示猎物与狼群之间的距离; Z_1 、 Z_2 、 Z_3 分别表示三只头狼与猎物的距离; \mathbf{Z}_α 、 \mathbf{Z}_β 、 \mathbf{Z}_δ 分别表示三只头狼 α 、 β 、 δ 决定种群个体下一步的移动向量; 参数 C_1 — C_3 、参数 A_1 — A_3 分别表示猎物和灰狼之间的距离和搜索的范围; $Z(t+1)$ 表示 t 次迭代后的猎物与狼群之间的距离。

1.3 深度学习框架

一维 CNN 最大的特点是一维矢量输入, 一维卷积核越大, 网络参数越少。因此一维 CNN 可以使用更大的卷积核获得更大的感受野^[28]。一维卷积原理如图 1 所示, 其中 R_1 — R_3 是历史时间序列。

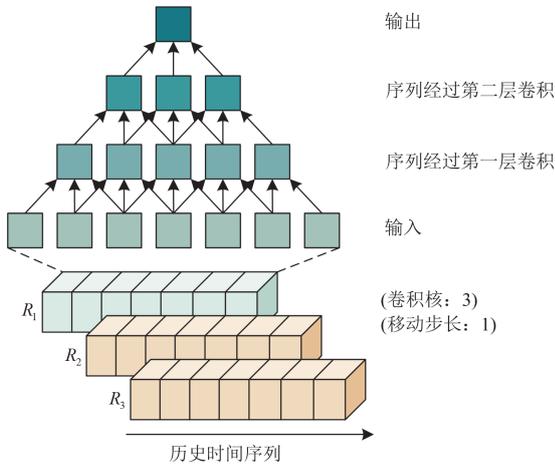


图 1 一维 CNN 原理图

Fig. 1 One-dimensional CNN schematic diagram

以 R_1 为例, 当卷积核在原始序列上水平滑动时, 每 3 个相邻元素加权生成输出序列的新元素。生成的输出序列作为下一层的输入, 直到输出层。一维 CNN 的卷积计算公式如式(5)所示。

$$y_i^{\text{out}} = f\left(\sum_{m=1}^K \omega_m u_{i+m}^{\text{in}} + b\right) \quad (5)$$

式中: y_i^{out} 为输出序列中第 i 个元素的值; $f(\cdot)$ 为激

活函数; K 为卷积核的大小; ω_m 为卷积核中第 m 个元素的权值; u_{i+m}^{in} 为输入序列中第 i 个元素的值; b 为卷积核的偏置项。

MLP 网络通常用于解决分类和回归问题。MLP 网络具有较强的逼近能力, 可以逼近任何复杂的非线性函数关系。前向传播公式如式(6)所示。

$$o_j^m = Y\left(\sum_{i \in (i-1)} \omega_j^m p_i^{m-1} + b_j^m\right) \quad (6)$$

式中: o_j^m 为第 m 层中第 j 个神经元的值; $Y(\cdot)$ 为激励函数, 该函数与输出值、上一层的神经元 p_i^{m-1} 有关, 其作用主要是使网络能够更好地处理非线性问题; ω_j^m 为连接到神经元的参数; b_j^m 为相应的偏移量。

在双残差网络中每一个 MCNN-MMLP 模块具有两个分支。 C^l 分支为正向预测分支, 与其他模块中的正向预测分支共同构成最终的预测输出; N^l 分支为分解重构分支, 与上一模块的输入分支形成残差并输入至下个模块, 如图 2 所示。

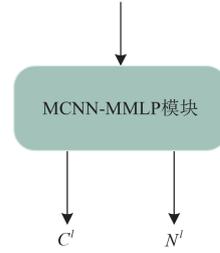


图 2 双残差结构图

Fig. 2 Dual residual architecture diagram

正向预测分支和分解重构分支生成的过程如式(7)一式(9)所示。

$$h_l^C = \text{ReLU}(\omega_l^C n_l + b_l^C) \quad (7)$$

$$h_l^N = \text{ReLU}(\omega_l^N n_l + b_l^N) \quad (8)$$

$$C^l = L_{\text{INEAR}}(h_l^C), N^l = L_{\text{INEAR}}(h_l^N) \quad (9)$$

式中: h_l^C 和 h_l^N 分别为第 l 块中正向预测和分解重构的映射; ReLU 为激活函数; ω_l^C 和 ω_l^N 为连接到神经元的权重; n_l 为第 l 块经过 MLP 网络映射后的输出; b_l^C 和 b_l^N 为相应的偏移量; L_{INEAR} 为线性映射输出。

1.4 多头注意力机制

多头注意力机制是一种注意力模型的扩展形式, 用于提高模型的表达能力和泛化能力。多头注意力模型把输入向量输入多组权重矩阵然后映射到输出, 其中查询(query)、键(key)、值(value)和输出都是 d 维的向量。多头注意力机制的优势在于每个子空间可以学习不同的特征表示, 从而提高模型的

表达能力。

多头注意力机制计算如式(10)一式(11)所示。

$$P_i = S_{\text{softmax}}(\alpha_i Q_i K_i^T) V_i \quad (i = 1, 2, \dots, h) \quad (10)$$

$$Y_{\text{output}} = c_{\text{concat}}(P_1, P_2, \dots, P_h) W_o \quad (11)$$

式中： P_i 为第*i*个头的计算结果； S_{softmax} 为多分类的激活函数； Q_i 、 K_i 、 V_i 分别为第*i*个子空间中进行线性变换的结果； c_{concat} 为拼接操作； α_i 为缩放因子； W_o 为输出的线性变换； Y_{output} 为多头注意力机制计算结果。

1.5 迁移学习

迁移学习是一种机器学习方法，通过将某个领域中学到的知识应用于相关的领域，以便在目标领域能够更好地学习和表现。迁移学习中的两个基本概念是领域和任务。其中领域 D 由特征空间 X 和边缘概率分布 $P(X)$ 组成，领域定义为 $D = \{X, P(X)\}$ 。任务由标签空间 Y 和回归模型 $f(x)$ 共同组成，任务定义为 $T = \{Y, f(x)\}$ 。迁移学习的策略是给定源域 D_s 和任务 T_s ，在新的目标域 D_t 和任务 T_t 中，通过学习 $D_s \neq D_t$ 或 $T_s \neq T_t$ 条件下源域与目标域知识的相似性，提高目标域回归模型的泛化性能。

2 基于级联 MCNN-MMLP 双残差网络的负荷功率预测模型

为提高负荷功率预测精度，本文提出了一种 GWO-VMD 和级联 MCNN-MMLP 双残差网络的预测模型。首先，对原始负荷数据进行数据预处理，并使用 GWO-VMD 对处理后的数据进行模态分解。

然后，利用未分解的时序数据训练源模型，并使用迁移学习的方法训练各个模态的目标模型。最后，利用初步预测数据和真实负荷数据得到误差数据集并再次使用 MCNN-MMLP 堆叠模型进行二次预测得到最终预测结果，预测流程如图 3 所示。

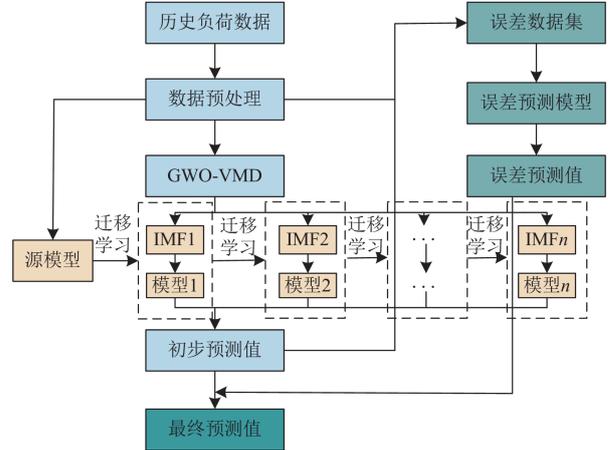


图 3 级联 MCNN-MMLP 双残差网络的负荷功率预测模型流程图

Fig. 3 Flowchart of load power prediction model of cascaded MCNN-MMLP dual residual network

MCNN-MMLP 双残差网络模型主要由多尺度卷积模块、MLP 模块和多头注意力机制组成，如图 4 所示。图 4 中，Conv1 表示卷积核为 1 的卷积层，Conv3 表示卷积核为 3 的卷积层。为增强模型非线性表达特征的能力、减小计算量，本文使用卷积核为 3 的卷积层进行级联代替卷积核为 5 的卷积层。在多尺度卷积模块中 BN 为批量归一化。不同卷积核可以提取出更全面、更丰富的特征信息，同时对

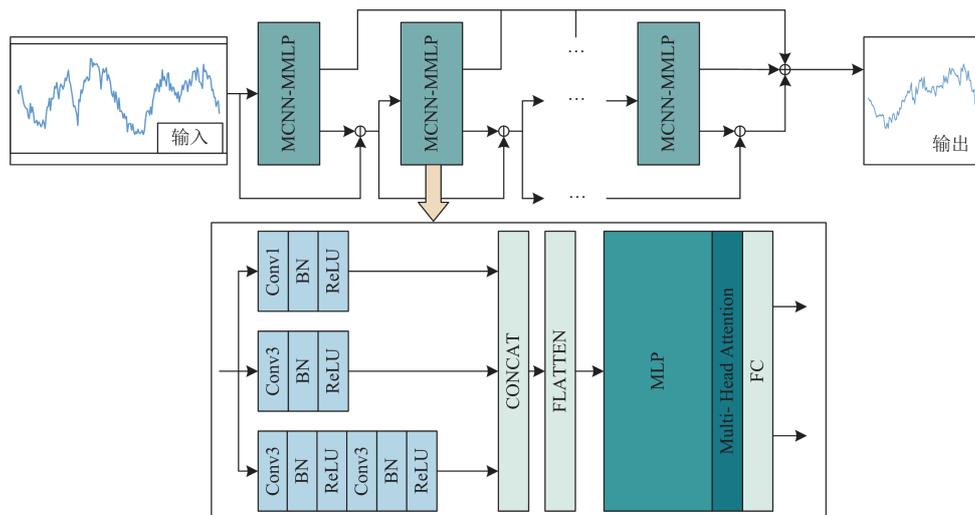


图 4 MCNN-MMLP 双残差网络的负荷功率预测模型结构图

Fig. 4 Structural diagram of load power prediction model of MCNN-MMLP dual residual network

局部变化更加敏锐, 能够更好地捕捉到时序数据中的局部模式和变化趋势, 从而提高对复杂数据下的预测能力。其次, 使用 CONCAT 函数和 FLATTEN 函数将获得的多尺度特征进行特征融合和铺展。然后, 使用嵌入多头注意力机制的 MLP 网络学习从输入特征与预测值之间的映射关系, 输出正向预测结果和分解重构分支。最后将 MCNN-MMLP 模块堆叠形成 MCNN-MMLP 双残差网络模型进行预测。

3 模型算例及实验结果分析

本文利用甘肃某地级市电网公司采集的真实负荷数据进行算例分析, 采样时间为 2022 年 1 月 1 日至 2023 年 8 月 27 日, 数据粒度为每 15 min 一个点, 共计 57 984 条数据。为提高模型的稳定性和准确性, 采用 3-sigma 准则对负荷数据进行预处理, 针对异常和缺失数据本文采用线性插值法进行填充。为降低特征之间的尺度差异对预测效果造成的影响, 本文对历史负荷数据进行归一化处理^[29]。将数据集中前 80% 作为训练集, 后 20% 作为测试集。其中, 采用 2023 年 8 月 27 日 00:00 至 2023 年 8 月 27 日 23:45 共 96 个点作为实验样本数据。

3.1 预测结果评估指标

为评估模型预测效果、比较其性能优劣, 本文采用 4 个电力负荷预测常用的评估指标: 均方误差 (mean squared error, MSE)、均方根误差 (root mean squared error, RMSE)、平均绝对百分比误差 (mean absolute percentage error, MAPE) 和决定系数 (R^2), 计算公式分别如式(12)~式(15)所示。

$$M_{SE} = \frac{1}{N} \sum_{t=1}^N (y_t - y'_t)^2 \quad (12)$$

$$R_{MSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - y'_t)^2} \quad (13)$$

$$M_{APE} = \frac{100\%}{N} \sum_{t=1}^N \left| \frac{y_t - y'_t}{y_t} \right| \quad (14)$$

$$R^2 = 1 - \frac{\sum_{t=1}^N (y'_t - y_t)^2}{\sum_{t=1}^N (\bar{y}_t - y_t)^2} \quad (15)$$

式中: y_t 和 y'_t 分别为 t 时刻负荷功率真实值和预测值; \bar{y}_t 为真实负荷功率的平均值; N 为用于测试的样本数目。

3.2 滑动窗口对模型性能的影响

本文使用基于历史负荷数据的预测模型。通过分析历史数据中负荷的变化趋势和规律, 来预测未来一段时间内的负荷情况。在模型建立过程中, 滑

动窗口的大小是一个重要的参数, 其决定了模型训练时训练样本的长度和数量。因此, 滑动窗口的大小对预测效果产生一定的影响。

为降低实验偶然性对结果的影响, 本文所有实验测试结果为多次实验的平均值。通过每增加 4 条样本数据的宽度来对窗口较优值进行寻优。表 1 为不同宽度的滑动窗口预测效果对比。从表 1 中可以看出, 当滑动窗口宽度为 4 时预测效果最为理想。与其他滑动窗口宽度相比, MSE 分别下降了 4.804 (MW)^2 、 11.104 (MW)^2 、 4.069 (MW)^2 、 5.231 (MW)^2 , RMSE 分别下降了 1.15%、2.62%、0.98%、1.26%, MAPE 分别下降了 2.68%、7.39%、3.79%、5.14%, R^2 分别上升了 0.34%、0.93%、0.34%、0.46%。其中滑动窗口宽度的改变对 MAPE 的影响最大, 选择宽度为 4 的滑动窗口使预测值更加接近真实值。相比宽度为 8 和 12 的滑动窗口, 较小的滑动窗口可以减少非平稳性的影响, 使模型更容易捕捉到数据的短期趋势和模式。当滑动窗口宽度继续增加, 预测效果也有所提升。使用较大的滑动窗口可以使模型更加稳定, 更多的历史数据可以提供更多的信息, 有助于减少数据的随机性和噪声。

表 1 不同宽度的滑动窗口预测效果对比

Table 1 Prediction effect comparison of sliding windows with different widths

滑动窗口宽度	MSE/(MW) ²	RMSE/MW	MAPE/%	R ²
4	204.194	14.289	1.015	0.867
8	208.998	14.456	1.043	0.864
12	215.298	14.673	1.096	0.859
16	208.263	14.431	1.055	0.864
20	209.425	14.471	1.070	0.863

3.3 本文模型有效性分析

3.3.1 不同模型直接预测效果对比

为验证所提出的 MCNN-MMLP 双残差网络模型的性能, 搭建使用广泛的 CNN、GRU、LSTM、CNN-LSTM-Attention 预测模型进行对比。具体预测结果见表 2。

表 2 不同模型预测效果对比

Table 2 Prediction effect comparison of different models

模型	MSE/(MW) ²	RMSE/MW	MAPE/%	R ²
CNN	276.796	16.637	1.266	0.819
GRU	289.365	17.011	1.291	0.811
LSTM	454.425	21.317	1.667	0.704
CNN-LSTM-Attention	496.042	22.272	1.698	0.676
本文模型	204.194	14.289	1.015	0.867

由表 2 得知, 本文提出的模型在直接预测效果上优于对比模型。MSE 相比其他模型下降了 26.23%、29.43%、55.06%、58.84%, RMSE 分别下降了 14.11%、16.00%、32.97%、35.84%, MAPE 分别下降了 19.83%、21.77%、39.11%、40.22%, 而 R^2 相比其他模型分别提升了 5.86%、6.90%、23.15%、28.25%。对比模型中, CNN 模型可以捕捉局部特征, 而对于长期依赖关系的建模相对较弱, 并且 CNN 对输入数据的细节敏感, 导致模型效果欠佳。GRU 模型需要从输入序列数据中学习对于预测任务有用的特征。而输入序列中特征不足, 缺少重要的信息, 导致模型难以准确预测。LSTM 模型在处理不规则、波动大的时序数据时, 无法充分利用其循环结构, 从而导致预测效果不佳。CNN-LSTM-Attention 模型通常用于多变量的时序预测中, 而在单变量预测中 CNN-LSTM-Attention 模型效果不够理想, 其原因为数据特征经过 CNN 提取后, LSTM 难以捕捉数据间长期依赖关系, 导致注意力机制未能明确关注到重要特征信息。为进一步展示不同模型的预测效果之间的差异, 图 5 展示了不同模型在 8 月 27 日(96 个样本点)的预测结果同真实值的对比。观察图 5 中的预测曲线, 本文提出模型的预测结果更加贴合真实值, 预测值的波动情况没有明显的滞后性。

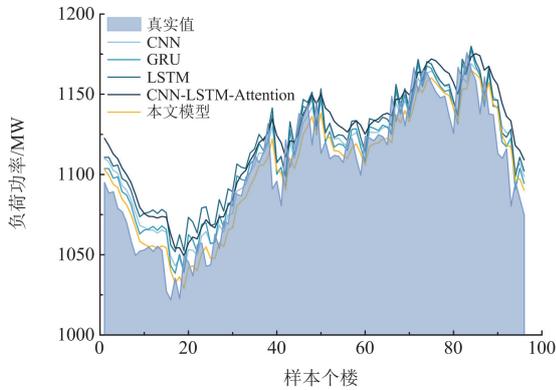


图 5 不同模型预测曲线对比图

Fig. 5 Comparison of prediction curves of different models

3.3.2 消融实验

为验证 MCNN-MMLP 双残差网络模型中各个模块的必要性, 本文在 MCNN-MMLP 双残差网络模型的基础上进行消融实验。为体现结构变化对预测结果的影响而非模型层数的影响, 使用卷积核为 1 的卷积网络代替多尺度卷积网络(D1), 使用卷积核为 3 的卷积网络代替多尺度卷积网络(D2), 使用卷积核为 1 的卷积网络代替 MLP 网络(D3)。同时在本文模型(D7)的基础上分别取消多头注意力机制(D4)

和取消卷积网络(D5), 最后利用卷积核为 1、3 的卷积层合成双通道卷积网络(D6)。消融实验结果如表 3 所示。

表 3 消融实验预测效果对比

Table 3 Prediction effect comparison in ablation experiments

模型	MSE/(MW) ²	RMSE/MW	MAPE/%	R ²
D1	218.399	14.778	1.109	0.857
D2	211.677	14.549	1.082	0.862
D3	237.573	15.413	1.179	0.845
D4	214.494	14.646	1.096	0.860
D5	209.944	14.489	1.088	0.863
D6	206.914	14.384	1.042	0.865
D7	204.194	14.289	1.015	0.867

从实验结果中可以看出, MLP 网络模块对 MCNN-MMLP 双残差网络模型的性能有较大影响。本文采用的数据集在小范围内波动大、非线性程度高, 而 MLP 网络可以通过多个隐藏层和非线性激活函数来表达复杂的非线性关系, 可以更好地捕捉非线性关系。将多尺度卷积模块换成单一的或是双通道的卷积模块, 模型的性能都在一定程度上降低, 证明多尺度卷积可以为模型提供更丰富的特征信息, 以更准确地预测未来的值。通过在不同尺度下进行卷积操作, 模型能够捕捉到不同层次上的相关特征, 从而提高预测的准确性。然后, 将多头注意力机制取消, 模型的性能也有所减弱, 对于一些重点信息没有进行关注从而导致信息的丢失, 根据实验结果表明, 多头注意力机制有助于整体预测效果的提升。

3.3.3 GWO-VMD 实验及分析

原始数据复杂程度高、小范围内波动大, 不利于数据特征的提取和后续的预测。GWO-VMD 在时序预测中通过对时序数据进行分解, 提取出数据的各个模态, 并使用这些模态来进行预测。在 VMD 算法中模态分量个数 K 和二次惩罚因子 α 是两个重要的参数, K 决定分解的本征模态函数(intrinsic mode function, IMF)个数, α 决定原始数据经 VMD 分解后各 IMF 的带宽。采用 GWO 算法对 VMD 参数进行优化时, 使用经验熵作为适应度函数。经验熵是一种常见的信息熵的估计方法, 对于 VMD 分解后的各个 IMF 分量进行经验概率密度函数计算并得到经验熵值, 将所有 IMF 分量的经验熵加权求和得到总的适应度值。通过计算适应度函数, 分析经验熵值从而了解数据的复杂程度。最终通过 GWO 算法的迭代过程来寻找最小的经验熵值, 以获得最佳的参数组合。具体优化 VMD 参数过程如

图 6 所示。首先, 定义适应度函数为各模态经验熵的加权和; 其次, 随机生成初始灰狼群体, 计算每个 IMF 分量的经验熵, 并综合得到总适应度。通过协作与竞争更新每只灰狼的位置, 优化参数以接近最优解。迭代调整直至满足条件, 输出最优参数及其对应的 VMD 分解。

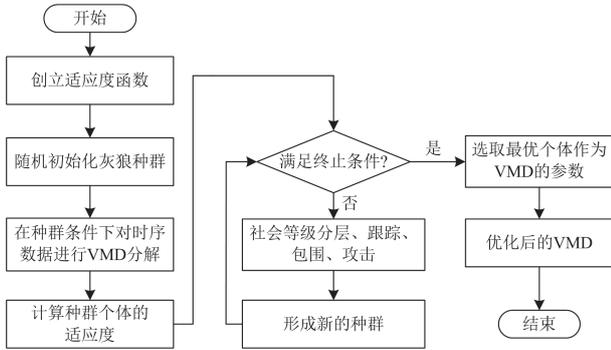


图 6 灰狼算法优化 VMD 参数流程图

Fig. 6 Flowchart of VMD parameters optimization based on grey wolf algorithm

本文将 K 的取值设置在 $[1, 10]$ 之间, α 的取值范围为 $[0, 2000]$, 种群数量设置为 50, 迭代 20 次, 迭代结果为 $K = 6$, $\alpha = 127.04$, 迭代过程中适应度值变化如图 7 所示。在迭代到第 5 次之后适应度值趋近于平稳, 将迭代后的最优个体代表的参数重新输入 VMD 中进行数据的分解。图 8 为选取部分采样点经过寻优后 VMD 分解的结果, 其中 IMF1 为负荷数据的趋势分量, 而 IMF2 和 IMF3 为负荷数据中周期性变化的分量, IMF4—IMF6 为高频分量, 其中包含随机性或者波动性较强的成分。

为验证 GWO-VMD 分解后的负荷数据对模型的预测效果有所提升, 本文将分解后的负荷数据分

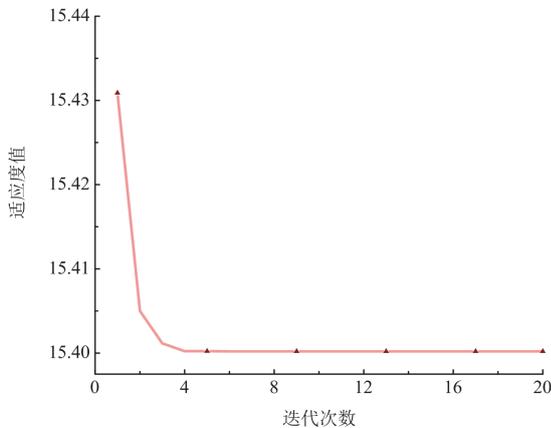


图 7 适应度值变化曲线

Fig. 7 Fitness value evolution curve

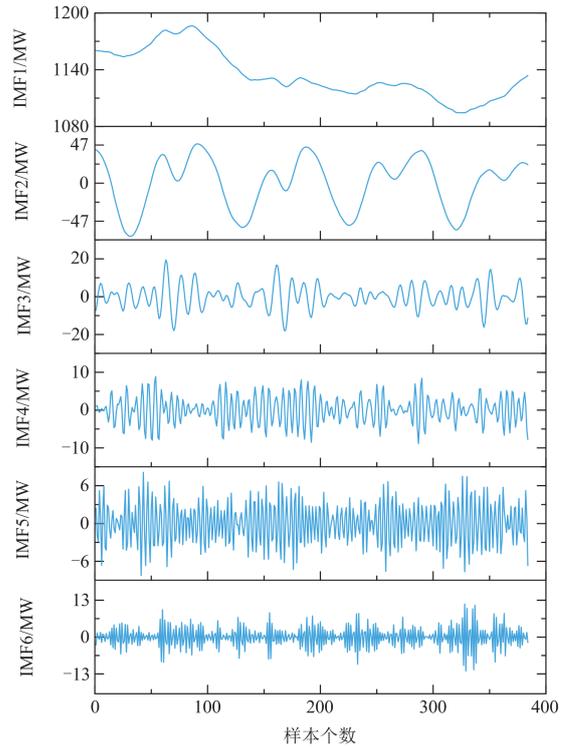


图 8 GWO-VMD 对负荷数据分解结果

Fig. 8 Decomposition results of load data by GWO-VMD

别输入到上文提到的所有模型中进行对比。对比结果如表 4 所示, 其中 T1 模型为 GWO-VMD-CNN, T2 模型为 GWO-VMD-GRU, T3 模型为 GWO-VMD-LSTM, T4 模型为 GWO-VMD-CNN-LSTM-Attention, T5 为 GWO-VMD 和本文提出的模型结合。

从表 4 中可以看出, 采用 GWO-VMD 分解后的负荷数据进行预测, 各个模型的预测效果都有所提升。本文模型在使用分解后的数据进行预测效果最优, 相比其他模型, MSE 分别降低了 78.29%、80.59%、88.51%、92.59%, RMSE 分别降低了 53.41%、56.77%、66.12%、72.78%, MAPE 分别下降了 57.66%、55.75%、65.38%、72.90%, R^2 分别提升了 4.55%、5.55%、10.14%、18.05%。为进一步说明 GWO-VMD 的有效性, 本文将直接预测的效果和使用分解数据进行预测的效果进行对比, 如图 9 所示。

表 4 使用分解数据的预测效果对比

Table 4 Prediction effect comparison with decomposed data

模型	MSE/(MW) ²	RMSE/MW	MAPE/%	R ²
T1	83.242	9.123	0.763	0.945
T2	96.696	9.833	0.730	0.936
T3	157.391	12.545	0.933	0.897
T4	243.867	15.616	1.192	0.841
T5	18.070	4.250	0.323	0.988

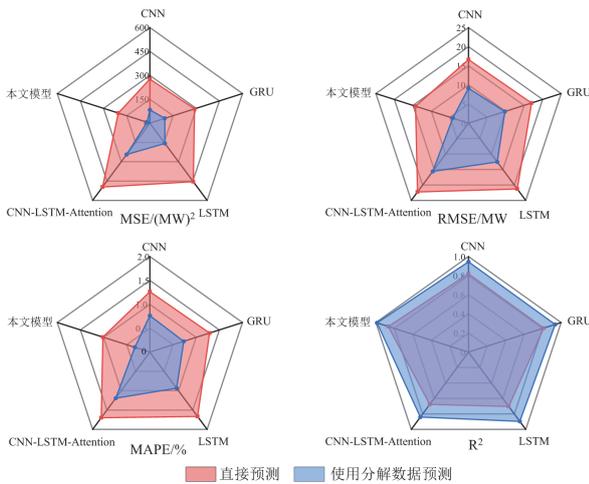


图 9 GWO-VMD 有效性论证对比图

Fig. 9 Comparison of validity demonstration with GWO-VMD

从图 9 可以可知，在使用分解后的负荷数据进行预测，预测效果有大幅度提升。其中本文提出的模型在直接预测和使用分解后的负荷数据进行预测效果都优于对比模型。在实验中本文使用了 CNN、GRU、LSTM、CNN-LSTM-Attention 4 个模型和本文模型进行对比实验，在 MSE 上使用分解数据的结果相比直接预测的结果分别降低了 69.92%、66.85%、65.36%、50.83%、91.15%；在 RMSE 上分别降低了 45.16%、42.19%、41.15%、29.88%、70.25%；在 MAPE 上分别下降了 39.68%、43.41%、44.03%、29.79%、68.17%；在 R^2 上分别上升了 15.38%、15.41%、27.59%、24.40%、14.08%。使用分解后的负荷数据，能够有效处理非线性和非平稳的数据。在负荷数据中包含着复杂的非线性关系和时间上的变化，通过对数据分解成多个模态函数，可以从原始数据中提取出隐含的特征信息，并对每个模态函数进行独立的分析和建模，从而更好地捕捉数据中的非线性和非平稳性特征。

图 10 为使用分解后的负荷数据各个模型的预测效果对比。从图 10 中可以看出，相比其他模型，本文提出的模型真实值和预测值的误差更小，模型预测值的曲线和真实值的曲线拟合程度高于其他模型。在一些波动较大的数据范围中，本文模型使用分解后的数据进行预测比直接预测的效果提升更为明显，同时其他模型的预测效果也都得到了不同程度的提升。

3.3.4 迁移学习实验及分析

将分解的数据直接输入模型进行训练，需要较长的训练周期。而使用迁移学习的方法，可以大大缩短训练时间，加快模型迭代收敛的速度。本文将

训练好的源模型进行迁移学习，在获得源模型的先验知识基础上，对其他模态进行模型训练。使用迁移学习进行模型训练和直接训练模型的损失函数曲线如图 11 所示。本文采取训练全部网络层的迁移策略，可以修改特征采集和预测部分的网络权重，使模型拟合程度更接近于真实数据，并且训练方法更加简单高效，无需额外设定超参数。

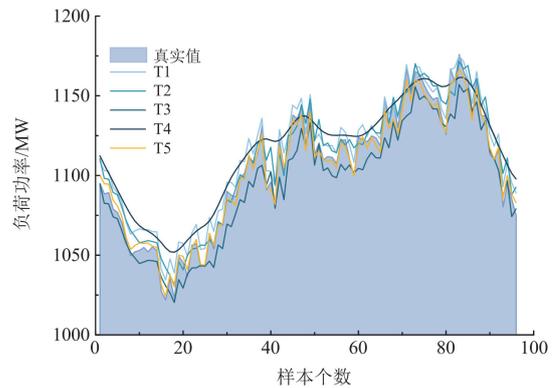
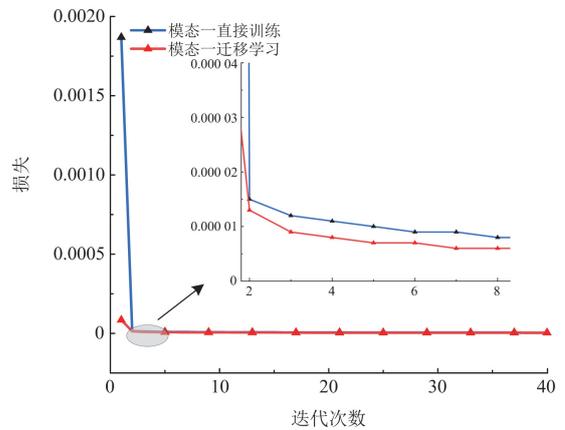
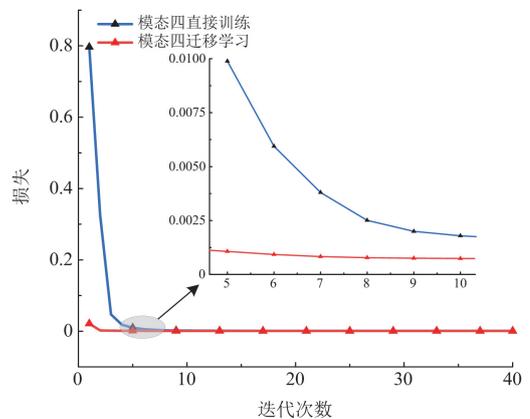


图 10 使用分解数据预测曲线对比图

Fig. 10 Prediction curves comparison with decomposed data



(a) 模态一损失函数曲线



(b) 模态四损失函数曲线

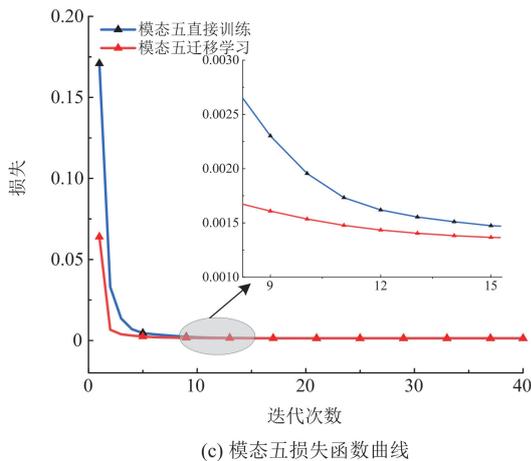


图 11 部分模态损失函数曲线对比图

Fig. 11 Comparison of partial mode loss function curves

从图 11 可以看出, 使用迁移学习后的模型收敛速度快于直接训练。使用迁移学习的训练模型在迭代 20 次左右收敛, 而直接训练的模型需要迭代 40 次左右才能收敛。因此使用迁移学习进行模型训练可以有效缩短训练时间、节约计算资源。

3.3.5 级联误差预测实验及分析

利用本文模型对分解后的负荷数据进行预测, 其预测效果提升显著。但由于分解数据中依旧存在模型无法直接表达的关系和复杂度, 最后预测结果和真实值存在一定的偏差。因此本文在误差的基础上采用级联预测的方式, 对误差进行二次预测。利用预测的误差值校正初值, 得到最终的预测结果。在利用本文模型的初步预测结果基础上, 使用本文模型和其他模型进行级联误差预测的效果对比, 如表 5 所示。

表 5 不同模型级联预测效果对比
Table 5 Cascaded prediction effect comparison of different models

级联误差预测模型	MSE/(MW) ²	RMSE/MW	MAPE/%	R ²
初步预测	18.070	4.250	0.323	0.988
CNN	7.867	2.804	0.217	0.994
GRU	15.452	3.930	0.301	0.990
LSTM	17.714	4.208	0.312	0.988
CNN-LSTM-Attention	9.899	3.146	0.237	0.993
本文模型	5.024	2.241	0.160	0.996

从表 5 中可知, 采用本文模型的级联预测方式最终预测效果最佳。与初步预测、其他模型进行级联预测相比, MSE 分别降低了 72.19%、36.13%、

67.48%、71.63%、49.24%; RMSE 分别降低了 47.27%、20.07%、42.97%、46.74%、28.76%; MAPE 分别降低了 50.46%、26.26%、46.84%、48.71%、32.48%; 而 R² 也有一定程度的提高。

本文模型级联的预测效果与其他模型级联的预测效果对比如图 12 所示, 可看出本文模型的级联预测效果最好。在一些细节拐点处, 利用本文模型进行级联预测的效果更加贴合真实值, 拟合程度更高。其他模型由于在初步预测的基础上进行级联预测, 总体预测趋势大致相符, 但在多处拐点都存在明显的预测偏差, 但使用本文模型进行级联预测在多处拐点上偏差更小。

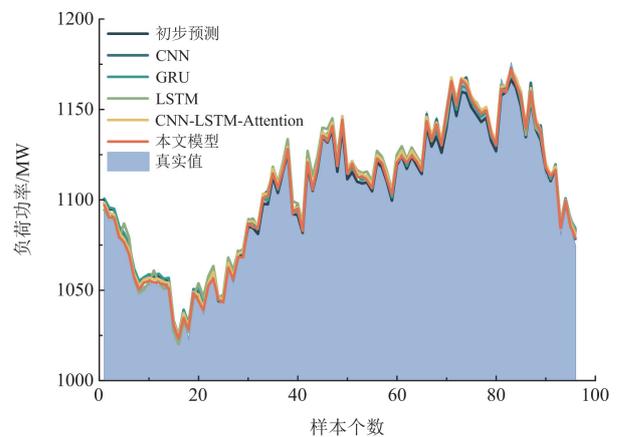


图 12 不同模型级联预测效果对比图

Fig. 12 Cascaded prediction effect comparison of different models

3.4 不同季节预测效果对比及分析

上述实验均以夏季某一典型日进行对比分析。而为验证本文提出的模型在不同季节的预测性能, 选取除夏季外剩余 3 个季节的典型日进行预测分析。在初步预测基础上, 使用本文模型和其他模型进行级联误差预测, 预测结果如表 6 所示。

通过表 5、表 6 综合对比分析预测结果可知, 本文模型面对不同季节时相比其他模型预测更加准确, 拟合能力更强。本文采用某地级市的真实负荷数据, 四季的负荷波动各不相同, 导致最终的预测效果有所差异。但使用本文模型进行初步预测, 各个季节的预测精确度都超过 98%。而经过本文模型的级联预测后, 各个季节的预测精确度超过 99%。从表 6 可知, 相比不同季节下其他模型的级联预测, 本文提出的模型在各个指标上都表现的更为优异。由于误差数据的波动性更强, 采用其他模型进行级联预测的效果可能会稍逊于初步预测。而在不同季节中, 使用本文模型进行级联预测的效果均优于初

步预测，尤其在春季提升效果更为明显。本文采用的真实负荷数据来自某工业城市，春季作为工业生产旺季用电量较大，导致春季负荷波动规律不明显。因此，春季的预测结果较其他三季有所下降，但级联预测的优势更加突出。

表 6 不同季节级联预测效果对比
Table 6 Comparison of cascaded prediction effect in different seasons

季节	级联误差预测模型	MSE/(MW) ²	RMSE/MW	MAPE/%	R ²
春季	初步预测	500.306	22.367	1.473	0.549
	CNN	217.440	14.745	0.938	0.804
	GRU	409.762	20.242	1.439	0.631
	LSTM	420.111	20.496	1.301	0.621
	CNN-LSTM-Attention	242.205	15.562	0.942	0.782
	本文模型	127.647	11.298	0.726	0.885
秋季	初步预测	12.944	3.597	0.264	0.984
	CNN	11.189	3.345	0.250	0.986
	GRU	15.624	3.952	0.285	0.981
	LSTM	16.728	4.090	0.305	0.980
	CNN-LSTM-Attention	12.286	3.505	0.262	0.985
	本文模型	10.910	3.303	0.252	0.987
冬季	初步预测	18.268	4.274	0.316	0.986
	CNN	13.793	3.713	0.272	0.989
	GRU	17.958	4.237	0.303	0.986
	LSTM	24.891	4.989	0.372	0.981
	CNN-LSTM-Attention	15.168	3.894	0.280	0.988
	本文模型	13.717	3.703	0.278	0.990

3.5 不同实际数据集预测对比及分析

为进一步验证本文模型的泛化能力，本文使用江苏某地级市实际负荷数据进行预测，采样范围为2022年12月28日至2022年12月31日(时间间隔为1h)，预测结果对比如图13所示，其指标对比结果如表7所示。

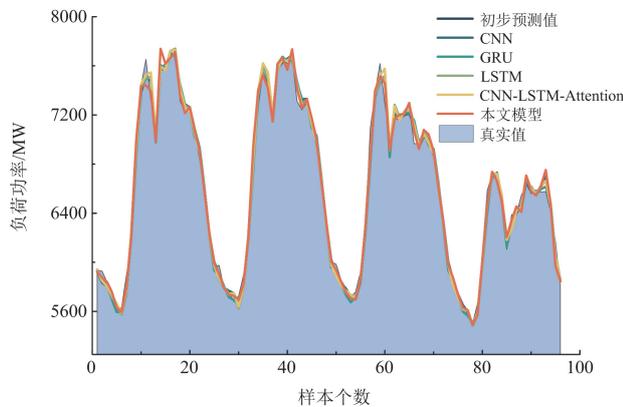


图 13 不同模型级联预测效果对比

Fig. 13 Cascaded prediction effect comparison of different models

表 7 不同模型级联预测指标对比
Table 7 Cascaded prediction index comparison of different models

级联误差预测模型	MSE/(MW) ²	RMSE/MW	MAPE/%	R ²
初步预测	6125.726	78.267	0.875	0.986
CNN	5006.572	70.757	0.835	0.988
GRU	7033.597	83.866	0.905	0.984
LSTM	5856.66	76.528	0.872	0.987
CNN-LSTM-Attention	6144.648	78.387	0.870	0.986
本文模型	4882.813	69.877	0.764	0.989

从表7得知，使用本文模型进行级联预测可有效降低误差，提高预测的精确度；而从图13分析可知，相较于其他预测模型，本文所建模型不仅在负荷平稳点能准确预测，而且在负荷拐点处(峰值及谷值附近)能更好地贴近变化趋势，因此总体预测效果优于其他模型。

4 结论

针对当前负荷特性复杂化导致预测精度低的问题，本文提出了GWO-VMD和级联MCNN-MMLP双残差网络的短期负荷预测模型，得出以下结论。

- 1) 针对目前预测模型在应对复杂负荷数据时预测精确度不高的问题，提出了利用GWO-VMD对负荷数据进行分解降低数据的复杂度。利用GWO对VMD中模态分量个数 K 和二次惩罚因子 α 进行寻优。
- 2) 为提升预测的精准度，提出了MCNN-MMLP双残差网络模型进行预测。每个模块都使用多尺度卷积神经网络提取数据特征，然后将特征输入到多层感知机进行输出。并且每个模块输出分为两部分：一部分用于分解重构作为下一模块的输入，另一部分直接进行预测输出。
- 3) 为进一步提高预测的精确度，本文利用MCNN-MMLP双残差网络模型进行级联预测，即对初步预测的误差进行二次预测，预测效果得到了进一步的提升。
- 4) 由于对各个模态都需要进行模型训练，本文采用迁移学习的方法加快模型训练，减少计算资源的浪费。

本文提出的基于GWO-VMD和级联MCNN-MMLP双残差网络的短期负荷预测模型为在复杂负荷数据的应用场景下提供了一种解决思路，与多种模型对比表现优异。本文模型的实验评估仅限于

短期预测, 尚未考虑现场实际部署时其他突发因素及中长期预测场景。因此, 未来有必要开展更深入的研究, 以满足实际负荷预测需求。

参考文献

- [1] GHADIMI N, AKBARIMAJD A, SHAYEGHI H, et al. Two stage forecast engine with feature selection technique and improved meta-heuristic algorithm for electricity load forecasting[J]. *Energy*, 2018, 161: 130-142.
- [2] 吴军英, 路欣, 刘宏, 等. 基于 Spearman-GCN-GRU 模型的超短期多区域电力负荷预测[J]. *中国电力*, 2024, 57(6): 131-140.
WU Junying, LU Xin, LIU Hong, et al. Ultra-short-term multi-region power load forecasting based on Spearman-GCN-GRU model[J]. *Electric Power*, 2024, 57(6): 131-140.
- [3] 封钰, 宋佑斌, 金晟, 等. 基于随机森林算法和粗糙集理论的改进型深度学习短期负荷预测模型[J]. *发电技术*, 2023, 44(6): 889-895.
FENG Yu, SONG Youbin, JIN Sheng, et al. Improved deep learning model for forecasting short-term load based on random forest algorithm and rough set theory[J]. *Power Generation Technology*, 2023, 44(6): 889-895.
- [4] MEMARZADEH G, KEYNIA F. Short-term electricity load and price forecasting by a new optimal LSTM-NN based prediction algorithm[J]. *Electric Power Systems Research*, 2021, 192: 106995.
- [5] 韩富佳, 王晓辉, 乔骥, 等. 基于人工智能技术的新型电力系统负荷预测研究综述[J]. *中国电机工程学报*, 2023, 43(22): 8569-8592.
HAN Fujia, WANG Xiaohui, QIAO Ji, et al. Review on artificial intelligence based load forecasting research for the new-type power system[J]. *Proceedings of the CSEE*, 2023, 43(22): 8569-8592.
- [6] 朱继忠, 苗雨旺, 董朝阳, 等. 基于 Attention-LSTM 与多模型集成的短期负荷预测方法[J]. *电力工程技术*, 2023, 42(5): 138-147.
ZHU Jizhong, MIAO Yuwang, DONG Zhaoyang, et al. Short-term load forecasting method based on Attention-LSTM and multi-model integration[J]. *Electric Power Engineering Technology*, 2023, 42(5): 138-147.
- [7] YAZICI I, BEYCA O F, DELEN D. Deep-learning-based short-term electricity load forecasting: a real case application[J]. *Engineering Applications of Artificial Intelligence*, 2022, 109: 104645.
- [8] WU Kailang, GU Jie, MENG Lu, et al. An explainable framework for load forecasting of a regional integrated energy system based on coupled features and multi-task learning[J]. *Protection and Control of Modern Power Systems*, 2022, 7(2): 349-362.
- [9] 邵必林, 纪丹阳. 基于 VMD-SE 的电力负荷分量的多特征短期预测[J]. *中国电力*, 2024, 57(4): 162-170.
SHAO Bilin, JI Danyang. Multi-feature short-term prediction of power load components based on VMD-SE[J]. *Electric Power*, 2024, 57(4): 162-170.
- [10] 王炜, 冯斌, 黄刚, 等. 基于自注意力编码器和深度神经网络的短期净负荷预测[J]. *中国电机工程学报*, 2023, 43(23): 9072-9084.
WANG Wei, FENG Bin, HUANG Gang, et al. Short-term net load forecasting based on self-attention encoder and deep neural network[J]. *Proceedings of the CSEE*, 2023, 43(23): 9072-9084.
- [11] CIULLA G, D'AMICO A. Building energy performance forecasting: a multiple linear regression approach[J]. *Applied Energy*, 2019, 253: 113500.
- [12] 董雷, 陈振平, 韩富佳, 等. 基于图卷积神经网络与 K-means 聚类的居民用户集群短期负荷预测[J]. *电网技术*, 2023, 47(10): 4291-4301.
DONG Lei, CHEN Zhenping, HAN Fujia, et al. Short-term load forecasting of residential user groups based on graph convolutional neural network and K-means clustering[J]. *Power System Technology*, 2023, 47(10): 4291-4301.
- [13] SONG K B, BAEK Y S, HONG D H, et al. Short-term load forecasting for the holidays using fuzzy linear regression method[J]. *IEEE Transactions on Power Systems*, 2005, 20(1): 96-101.
- [14] 毛李帆, 江岳春, 龙瑞华, 等. 基于偏最小二乘回归分析的中长期电力负荷预测[J]. *电网技术*, 2008, 32(19): 71-77.
MAO Lifan, JIANG Yuechun, LONG Ruihua, et al. Medium-and long-term load forecasting based on partial least squares regression analysis[J]. *Power System Technology*, 2008, 32(19): 71-77.
- [15] HAIDA T, MUTO S. Regression based peak load forecasting using a transformation technique[J]. *IEEE Transactions on Power Systems*, 1994, 9(4): 1788-1794.
- [16] 赵佩, 代业明. 基于实时电价和加权灰色关联投影的 SVM 电力负荷预测[J]. *电网技术*, 2020, 44(4): 1325-1332.
ZHAO Pei, DAI Yeming. Power load forecasting of SVM based on real-time price and weighted grey relational projection algorithm[J]. *Power System Technology*, 2020, 44(4): 1325-1332.
- [17] 张淑清, 要俊波, 张立国, 等. 基于改进深度稀疏自编码器及 FOA-ELM 的电力负荷预测[J]. *仪器仪表学报*, 2020, 41(4): 49-57.

- ZHANG Shuqing, YAO Junbo, ZHANG Ligu, et al. Power load forecasting based on improved deep sparse auto-encoder and FOA-ELM[J]. Chinese Journal of Scientific Instrument, 2020, 41(4): 49-57.
- [18] 商立群, 李洪波, 侯亚东, 等. 基于特征选择和优化极限学习机的短期电力负荷预测[J]. 西安交通大学学报, 2022, 56(4): 165-175.
- SHANG Liqun, LI Hongbo, HOU Yadong, et al. Short-term power load forecasting based on feature selection and optimized extreme learning machine[J]. Journal of Xi'an Jiaotong University, 2022, 56(4): 165-175.
- [19] 冯伟, 蒋玮, 杨乐, 等. 基于 BP 神经网络方法的城市用电量预测模型研究[J]. 电网与清洁能源, 2018, 34(6): 19-25, 32.
- FENG Wei, JIANG Wei, YANG Le, et al. Research on forecasting model based on BP neural network for urban electricity consumption[J]. Power System and Clean Energy, 2018, 34(6): 19-25, 32.
- [20] 杨小明, 崔雪, 周斌, 等. 基于粒子群优化支持向量机的短期负荷预测[J]. 武汉大学学报(工学版), 2018, 51(8): 715-720.
- YANG Xiaoming, CUI Xue, ZHOU Bin, et al. Research on short term load forecasting based on particle swarm optimization-support vector machine[J]. Engineering Journal of Wuhan University, 2018, 51(8): 715-720.
- [21] 季玉琦, 严亚帮, 和萍, 等. 基于 K-Medoids 聚类与栅格法提取负荷曲线特征的 CNN-LSTM 短期负荷预测[J]. 电力系统保护与控制, 2023, 51(18): 81-93.
- JI Yuqi, YAN Yabang, HE Ping, et al. CNN-LSTM short-term load forecasting based on the K-Medoids clustering and grid method to extract load curve features[J]. Power System Protection and Control, 2023, 51(18): 81-93.
- [22] 姚浩然, 李成鑫, 郑秀娟, 等. 集成自适应啾啾模态分解和 BiLSTM 的短期负荷组合预测模型[J]. 电力系统保护与控制, 2022, 50(19): 58-66.
- YAO Haoran, LI Chengxin, ZHENG Xiujuan, et al. Short-term load combination forecasting model integrating ACMD and BiLSTM[J]. Power System Protection and Control, 2022, 50(19): 58-66.
- [23] 欧阳静, 杨吕, 尹康, 等. 基于 ALIF-LSTM 多任务学习的综合能源系统短期负荷预测[J]. 太阳能学报, 2022, 43(9): 499-507.
- OUYANG Jing, YANG Lü, YIN Kang, et al. Short-term load forecasting method for integrated energy system based on ALIF-LSTM and multi-task learning[J]. Acta Energiae Solaris Sinica, 2022, 43(9): 499-507.
- [24] 石卓见, 冉启武, 徐福聪. 基于聚合二次模态分解及 Informer 的短期负荷预测[J]. 电网技术, 2024, 48(6): 2574-2583.
- SHI Zhuojian, RAN Qiwu, XU Fucong. Short-term load forecasting based on aggregated secondary decomposition and Informer[J]. Power System Technology, 2024, 48(6): 2574-2583.
- [25] 刘建华, 李锦程, 杨龙月, 等. 基于 EMD-SLSTM 的家庭短期负荷预测[J]. 电力系统保护与控制, 2019, 47(6): 40-47.
- LIU Jianhua, LI Jincheng, YANG Longyue, et al. Short-term household load forecasting based on EMD-SLSTM[J]. Power System Protection and Control, 2019, 47(6): 40-47.
- [26] 朱凌建, 荀子涵, 王裕鑫, 等. 基于 CNN-BiLSTM 的短期电力负荷预测[J]. 电网技术, 2021, 45(11): 4532-4539.
- ZHU Lingjian, XUN Zihan, WANG Yuxin, et al. Short-term power load forecasting based on CNN-BiLSTM[J]. Power System Technology, 2021, 45(11): 4532-4539.
- [27] TAO Hong, SHU Fan. Probabilistic electric load forecasting: a tutorial review[J]. International Journal of Forecasting, 2016, 32(3): 914-938.
- [28] WANG Ke, MA Changxi, QIAO Yihuan, et al. A hybrid deep learning model with 1DCNN-LSTM-Attention networks for short-term traffic flow prediction[J]. Physica A: Statistical Mechanics and its Applications, 2021, 583: 126293.
- [29] 雷柯松, 吐松江·卡日, 伊力哈木·亚尔买买提, 等. 基于 WGAN-GP 和 CNN-LSTM-Attention 的短期光伏功率预测[J]. 电力系统保护与控制, 2023, 51(9): 108-118.
- LEI Kesong, KARI·Tusongjiang, YILIHAMU·Yaermaimaiti, et al. Prediction of short-term photovoltaic power based on WGAN-GP and CNN-LSTM-Attention[J]. Power System Protection and Control, 2023, 51(9): 108-118.

收稿日期: 2024-05-05; 修回日期: 2024-07-11

作者简介:

余凯峰(1999—), 男, 硕士研究生, 研究方向为负荷预测; E-mail: 107552204481@stu.xju.edu.cn

吐松江·卡日(1984—), 男, 通信作者, 博士, 副教授, 博士生导师, 研究方向为电力系统数字化、新能源发电与并网和人工智能与模式识别等。E-mail: minyun229@163.com
(编辑 张颖)