

DOI: 10.19783/j.cnki.pspc.241617

基于时序卷积神经网络和纵横交叉算法的低压台区负荷预测

丁伟锋¹, 周震震¹, 谢振华², 肖耀辉¹, 黄和燕¹, 何森¹

(1. 中国南方电网有限责任公司超高压输电公司电力研究院, 广东 广州 510620;

2. 广东工业大学自动化学院, 广东 广州 510006)

摘要: 精准的电力负荷预测对低压台区运维至关重要。为提升台区电力负荷预测精度, 提出一种纵横交叉算法(crisscross optimization algorithm, CSO)优化卷积注意力机制(convolutional block attention module, CBAM)和时序卷积神经网络(temporal convolutional network, TCN)的低压台区电力负荷预测模型。首先, 建立以时序卷积神经网络为基础的预测模型, 提取电力负荷输入序列隐含的时间规律。其次, 在模型输入侧引入CBAM模块, 通过在通道和空间上与模型输入进行加权, 提高模型对关键特征的敏感性。最后, 为解决模型参数易陷入局部最优、模型泛化性不高的问题, 提出使用CSO对CBAM-TCN的全连接层进行二次优化。以广东省某地两个典型低压台区实测电力负荷数据集进行仿真建模, 结果表明所提组合预测方法性能优于其他对比模型, 并对其有效性进行了验证。

关键词: 低压台区; 负荷预测; 纵横交叉算法; 时序卷积神经网络; 卷积注意力机制

Low-voltage distribution area load forecasting based on temporal convolutional network and crisscross optimization algorithm

DING Weifeng¹, ZHOU Zhenzhen¹, XIE Zhenhua², XIAO Yaohui¹, HUANG Heyan¹, HE Sen¹

(1. CSG EHV Electric Power Research Institute, Guangzhou 510620, China; 2. School of Automation,

Guangdong University of Technology, Guangzhou 510006, China)

Abstract: Accurate power load forecasting is crucial for the operation and maintenance of low-voltage distribution areas. To improve the accuracy of power load forecasting, this paper proposes a low-voltage load forecasting model that integrates a crisscross optimization algorithm (CSO) with a convolutional block attention module (CBAM) and a temporal convolutional network (TCN). First, a forecasting model is established based on TCN to extract the implicit temporal patterns of the input sequence of power loads. Second, a CBAM module is introduced at the model input side to apply channel-wise and spatial-wise weighting, thereby enhancing the model's sensitivity to key features. Finally, to address issues such as local optima and limited generalization, the CSO algorithm is proposed to perform secondary optimization on the fully connected layer of the CBAM-TCN model. Using real power load datasets from two typical low-voltage substations in Guangdong province for simulation and modelling, the results show that the proposed hybrid forecasting method outperforms other comparative models and effectively validates its superiority.

This work is supported by the Young Scientists Fund of National Natural Science Foundation of China (No. 52207166).

Key words: low-voltage distribution area; load forecasting; crisscross optimization algorithm; temporal convolutional network; convolutional block attention module

0 引言

配变台区处于电网末端, 其故障大多是因为负

荷过大, 引起局部过热导致保险片熔断、接线柱头烧毁等, 台区故障将直接影响该区域范围内的居民用电, 造成停工停产, 对经济造成一定损失, 供电单位会受到用电客户的投诉^[1]。当前对于台区故障的处理往往处于被动状态, 无法及时预知台区面临的重过载风险。因此通过台区负荷进行提前预测, 有利于供电单位及时掌握台区运行态势, 对后续指

基金项目: 国家自然科学基金青年科学基金项目资助(52207166); 中国南方电网有限责任公司科技项目资助(012000KK52200014)

导改善台区的供电稳定性、提高供电质量显得尤为重要。

电力负荷预测依赖于数据驱动, 通过海量的电力负荷数据与相关的特征因素进行建模, 负荷预测技术大致分为传统方法与机器学习方法^[2-5]。传统方法基本建立在统计学的基础上, 存在预测精度受数据质量影响大、非线性拟合能力差的弊端。早期机器学习模型多为浅层结构, 如文献[6]提出一种考虑实时电价为特征因素的支持向量机(support vector machine, SVM)电力负荷预测模型; 文献[7]提出基于改进决策树的短期电力负荷预测模型, 但模型结构相对简单, 预测效果受超参数影响大, 容易发生过拟合现象。相对于浅层模型, 深度机器学习模型拥有更强大的特征表达能力、泛化能力和大数据处理能力, 如文献[8]使用卷积神经网络(convolutional neural network, CNN)挖掘学习多尺度分量特征, 实现用户级超短期负荷预测; 文献[9]基于多任务学习实现综合能源系统负荷短期预测, 但尚未充分利用特征之间的时序关系, 模型对特征因素的敏感度一致。

电力负荷数据属于时间序列数据, 预测模型若能学习到时序上的隐藏关系, 将会提升负荷的预测精度。且神经网络模型常采用梯度下降法更新网络权值^[10-11], 同时模型输入的特征因素过多, 网络对于特征因素的关注程度是一样的, 如何减轻冗余特征对于电力负荷预测结果的影响, 需要进一步研究^[12-13]。

针对上述问题, 本文构建了一种纵横交叉算法(crisscross optimization algorithm, CSO)优化卷积注意力模块(convolutional block attention module, CBAM)和时序卷积神经网络(temporal convolutional network, TCN)的台区电力负荷预测模型。为解决常规预测模型对电力负荷和特征因素时序规律敏感性不高的问题, 采用 TCN 获取时间序列深层隐含关系; 赋予特征因素不同的关注度, 降低冗余因素影响, 设计了独特的动态特征加权 CBAM-TCN 电力负荷预测模型; 针对神经网络模型采用梯度下降法训练后, 模型参数易陷入局部最优、模型泛化性不高的问题, 提出使用 CSO 算法对 CBAM-TCN 的全连接层进行二次优化, 以此获得较高的电力负荷预测精度。实验结果表明, 本文所提 CSO-CBAM-TCN 的预测精度高于其他模型。

1 基于 CBAM 的 TCN 预测模型

1.1 TCN

对于电力负荷预测任务, 研究者常使用循环神经网络^[14-15]、长短时记忆网络^[16-18]等捕捉负荷、温

度等数据的时间规律特性, 但是这些网络结构的运算机制较为复杂, 网络训练时间长, 若输入样本的维度较大, 网络的预测效果并不佳, 甚至在处理长时间序列时会出现梯度爆炸或消失的风险。TCN 融入了一维卷积、因果卷积和扩张卷积, 并在网络结构里加入了残差连接, 是一种善于处理时间序列的网络, 因此本文将该网络应用于台区电力负荷预测领域。

1) 因果卷积与扩张卷积

因果卷积由一系列的一维卷积堆叠而成, 其结构如图 1 所示, 可以明显地看出, 当前时刻的输出和历史输入密切相关^[19], 例如 t 时刻的输出 y_t 仅和当前时刻以及上一时刻的两个输入做卷积运算, y_t 和历史信息 $\{x_{t-3}, x_{t-2}, x_{t-1}, x_t\}$ 相关。

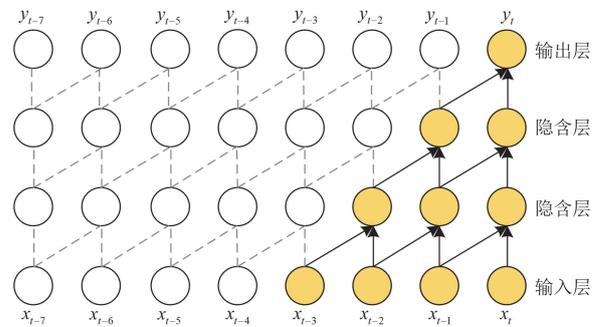


图 1 因果卷积结构图

Fig. 1 Structure of causal convolution

然而因果卷积在处理长时间序列时, 不得不增加隐含层的数量或者增大卷积核的大小, 这将使得网络结构变得复杂, 因此因果卷积存在感受野过小的缺陷。

TCN 在网络结构中加入了扩张卷积, 其优势在于: 在隐含层数、卷积核大小相同的时候, 扩张卷积可以获取更大的感受野, 有利于解决电力负荷长时间序列预测问题^[20]。扩张卷积的结构如图 2 所示。

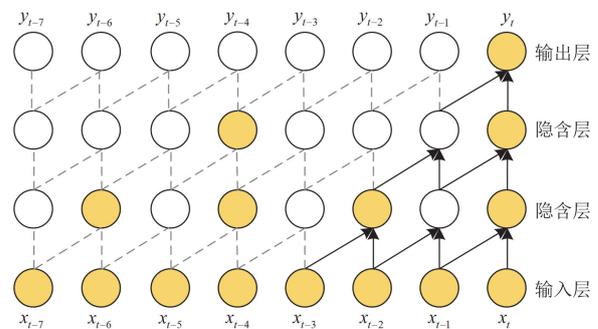


图 2 扩张卷积结构图

Fig. 2 Structure of extended convolution

假设图 2 结构的扩张系数 $d=[1,2^1,2^2]$ ，卷积核的大小为 2，对比图 1 的因果卷积，扩张卷积的 y_i 和历史信息 $\{x_{i-7}, x_{i-6}, \dots, x_{i-1}, x_i\}$ 紧密联系，感受野要更大。

假设网络的输入为 x ， f 为卷积核，则扩张卷积的计算公式为

$$F(s) = \sum_{i=0}^{k-1} f(i)x_{s-d \cdot i} \quad (1)$$

式中： $F(s)$ 表示扩张卷积第 s 个神经元的输出； k 为卷积核的大小；扩张系数 $d=2^i$ ， i 为网络层数。

从式(1)可以看出， $s-d \cdot i$ 反映了神经元输出只对历史输入做卷积运算。通过调节 k 和 d 的大小实现对感受野的控制，在增大这两个参数的情况下，网络输出受历史信息的影响更为长久。

2) 残差结构

由扩张卷积理论，增大网络层数(扩张系数)可以增大网络的感受野，但网络层数过多容易发生梯度爆炸或者梯度消失的问题，造成网络崩溃，网络学习能力反而下降。残差块很好地解决了因网络层数增多引起网络退化的问题^[21]，残差块在输出和输入之间引入了捷径，这使得低层能够较为简单地接收到高层传递过来的梯度，从而避免出现梯度爆炸或梯度消失的问题。

一维残差块如图 3 所示。假设残差块的输入为 a ，经过卷积后的输出为 $A(a)$ ，残差块内部加入了权值归一化、ReLU 激活函数和 Dropout 随机失活率，其中权值归一化用于限制权值范围，加速网络

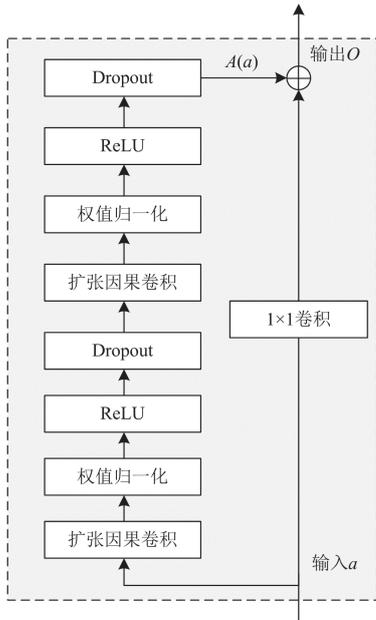


图 3 一维残差块

Fig. 3 One-dimensional residual block

收敛速度；ReLU 促使网络高效的梯度下降和反向传播；Dropout 减少了神经元之间复杂的共适应关系，使网络结构变简单，防止训练过程中出现过拟合现象。残差块输出 O 的计算公式为

$$O = \delta(a + A(a)) \quad (2)$$

式中： δ 为激活函数，一般选择 ReLU 函数。

1.2 CBAM

注意力机制的权重系数通常由神经网络生成，CBAM 则是依赖于卷积神经网络，是一种两维度注意力机制结构，分别为通道注意力和空间注意力^[22-23]。CBAM 是一种轻量化的注意力机制模块，模块参数少，加入 CBAM 并不会给整个网络带来过大负担，却可以提升整个模型的预测精度。CBAM 的结构如图 4 所示，可以看出：相比空间注意力，通道注意力的顺序通常在前，原始特征图经过通道注意力形成权重系数矩阵，再与原始特征图做乘积，空间注意力模块类似。

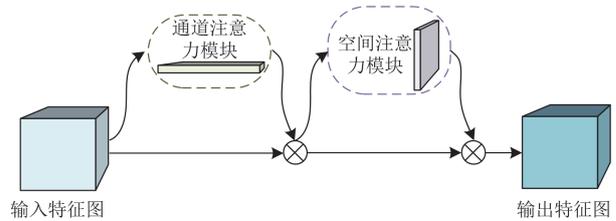


图 4 CBAM 结构图

Fig. 4 Structure of CBAM

假设某一层特征图 $N \in \mathbf{R}^{C \times H \times W}$ ，则通道注意力 $M_u \in \mathbf{R}^{C \times 1 \times 1}$ ，空间注意力 $M_v \in \mathbf{R}^{1 \times H \times W}$ ， C 、 H 和 W 分别表示特征图的通道数、高度和宽度。

$$N' = M_u(N) \otimes N \quad (3)$$

$$N'' = M_v(N') \otimes N' \quad (4)$$

式中： N' 为通道注意力模块和 N 的元素的乘积结果； \otimes 为矩阵元素张量积； N'' 为空间注意力模块和 N' 的乘积结果。

通道注意力模块具体结构如图 5 所示，为了有效提取通道上的有效信息，首先采用最大值池化和平均值池化并行的方式压缩得到两个局部特征图；然后在其后接入一个共享权重的多层感知机，进行非线性映射；最后需要对非线性映射的特征图信息做加法，再经 Sigmoid 函数映射到 $[-1,1]$ ，就可以得到通道注意力权重矩阵 $M_u(F)$ ，过程可以表达为

$$M_u(N) = \sigma(\text{MLP}(\text{Maxpool}(N)) + \text{MLP}(\text{Avgpool}(N))) \quad (5)$$

式中： σ 为 Sigmoid 函数； $\text{MLP}(\cdot)$ 表示多层感知机映射； $\text{Maxpool}(\cdot)$ 和 $\text{Avgpool}(\cdot)$ 分别表示最大值池化和平均值池化操作。

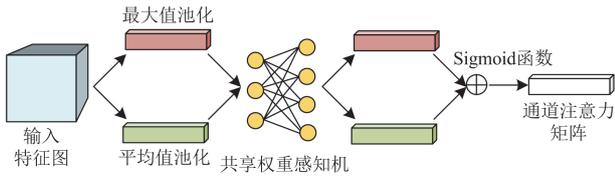


图5 通道注意力模块结构图

Fig. 5 Structure diagram of channel attention module

空间注意力模块的具体结构如图6所示,空间注意力模块可获取输入特征图不同位置信息,首先对输入特征图在通道上进行最大值池化和平均值池化操作,得到两个维度一致的特征图;然后将这两个特征图在通道上进行合并得到特征图 F ;最后利用 1×1 的卷积核做卷积操作,再经过 Sigmoid 函数输出后,即得到空间注意力权重矩阵。过程可以表达为

$$M_v(F) = \sigma(\text{Conv}(\text{Maxpool}(N); \text{Avgpool}(N))) \quad (6)$$

式中: $M_v(F)$ 表示空间注意力权重矩阵; $\text{Conv}(\cdot)$ 表示卷积操作。

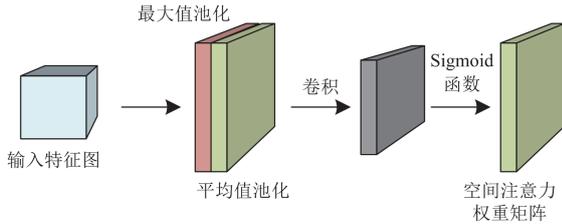


图6 空间注意力模块结构图

Fig. 6 Structure diagram of spatial attention module

2 基于模型参数优化的预测模型

2.1 CSO

CSO 启发于“中庸思想”和遗传算法的基因交叉机制,是一种性能优异的群智能元启发式算法^[24-26]。CSO 的显著特点就是横向交叉和纵向交叉相结合的双交叉机制,在寻优过程中,两种交叉算子交替进行,不断迭代,直至迭代次数达到上限或者损失满足要求,纵横交叉算法流程图如图7所示。

横向交叉是对种群个体之间进行交叉互换信息。以 S 代表种群,种群中的个体数为 M , 每个个体的维度为 D , 用 $S(p)$ 、 $S(q)$ 分别代表种群中两个不同个体, $S(p, w)$ 、 $S(q, w)$ 分别代表个体 p 、 q 的第 w 维度信息,则这两个个体横向交叉交换信息的公式为

$$S_{hc}(p, w) = r_1 \cdot S(p, w) + (1 - r_1) \cdot S(q, w) + c_1 \cdot (S(p, w) - S(q, w)) \quad (7)$$

$$S_{hc}(q, w) = r_2 \cdot S(q, w) + (1 - r_2) \cdot S(p, w) + c_2 \cdot (S(q, w) - S(p, w)) \quad (8)$$

式中: $S_{hc}(p, w)$ 、 $S_{hc}(q, w)$ 分别为父代 $S(p)$ 和 $S(q)$ 做横向交叉后产生的新个体; r_1 、 r_2 代表两组父代在相同维度上交换信息的概率, r_1 、 $r_2 \in \text{Rand}(0, 1)$; c_1 、 $c_2 \in \text{Rand}(0, 1)$ 。

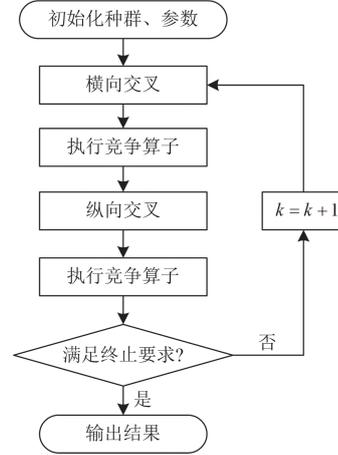


图7 纵横交叉算法流程图

Fig. 7 Flow chart of CSO algorithm

纵向交叉的对象是横向交叉后的子代,是对种群中的单个个体不同维度上的信息进行交叉互换,这样的交叉方式可以增加种群多样性,同时能够使个体在一定的条件下摆脱局部最优,增强寻优能力。纵向交叉的数学描述为

$$S_{vc}(p, w_1) = r \cdot S(p, w_1) + (1 - r) \cdot S(p, w_2) \quad (9)$$

式中: $S_{vc}(p, w_1)$ 为纵向交叉产生的子代; r 用于衡量个体不同维度信息交换的程度, $r \in \text{Rand}(0, 1)$; $S(p, w_1)$ 、 $S(p, w_2)$ 分别表示父代个体 S 的第 w_1 和 w_2 维度信息。

2.2 CSO 优化 CBAM-TCN

本文将 CSO 应用于本文提出的 CBAM-TCN 台区负荷预测深度学习模型参数优化,以确保该模型在初步使用梯度下降法进行训练后,模型参数中的阈值与偏置没有陷入局部最优。优化流程如图8所示,图中: j 表示迭代次数; N_{\max} 表示最大迭代次数。

在优化的过程中,需要确定适应度函数,本文选取均方误差和正则项作为适应度函数,均方误差保证预测结果贴近实际值,正则项的引入让模型不发生拟合现象,提升模型的泛化能力。则适应度函数可以用式(10)表示。

$$l = \frac{1}{G} \sum_{b=1}^G (g_b - \bar{g}_b)^2 + \lambda \sum_{c=1}^Q \theta_c^2 \quad (10)$$

式中: G 表示训练集样本总数; θ_c 表示阈值或偏置; Q 为用于训练的阈值或偏置的总数量; g_b 和 \bar{g}_b 分别

表示台区负荷的真实值和预测值； λ 表示正则项的系数，取 0.005。

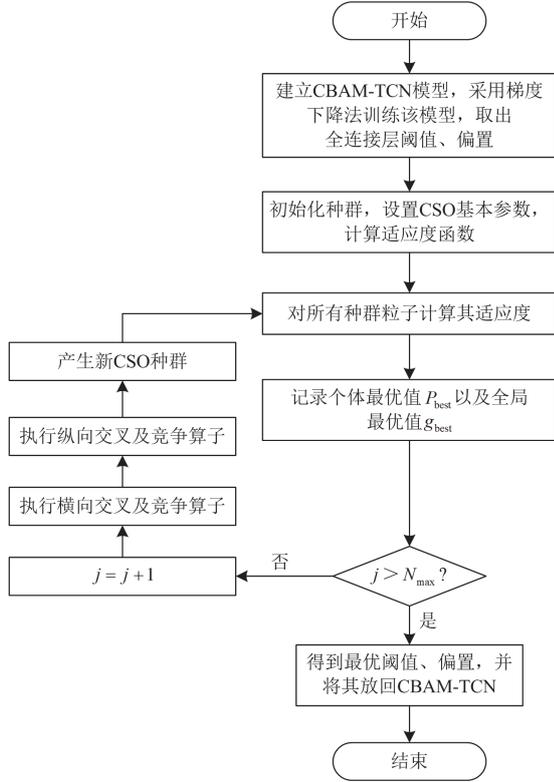


图 8 CSO 优化 CBAM-TCN 流程

Fig. 8 Flow chart of CBAM-TCN optimized by CSO

3 算例分析

本文所有建模仿真在 Windows11 系统和 PyCharm 软件中进行，由深度学习 TensorFlow 和 Keras 框架搭建神经网络模型。

3.1 数据准备

本文选取 2021 年广东省汕头市 2 个典型低压台区实测历史电力负荷数据集进行仿真建模，该数据集每隔 15 min 采样一次，每日采集 96 个点，共 2976 个点，采集的数据信息包含了温度、天气、台区台账信息和负荷大小、时间特征等。选取最后两天数据作为测试集，验证模型的预测性能，其余数据作为训练集。

为研究特征变量与电力负荷之间的相关性，采用皮尔逊相关系数法(Pearson correlation coefficient, PCC)定量分析关联程度，皮尔逊相关系数 $P_{CC} \in [-1,1]$ ， P_{CC} 的正和负分别代表特征与预测结果呈现正相关或负相关，绝对值越大，这种相关联系越强。表 1 是 2 个台区 P_{CC} 绝对值大小排序前 5 的特征，相关性分析间接说明负荷大小与温度、节假日、天气等因素密切关联。

表 1 相关性最强的 5 个特征变量

Table 1 Five characteristic variables with the strongest correlation

工业台区		居民台区	
特征	P_{CC} 参数	特征	P_{CC} 参数
最低温度	0.312	节假日	0.375
工作日	0.293	平均气温	0.314
平均气温	0.285	最高气温	0.287
最高温度	0.241	周末	-0.273
节假日	-0.211	小雨	0.192

神经网络模型的参数设置采用 Adam 优化器，迭代次数为 200，Batch_size 为 64，损失函数为均方误差函数，学习率为 0.01，激活函数为 ReLU 函数。其中 CBAM-TCN 网络的参数设置如下：CBAM 的输入和输出的维度相同；TCN 的卷积核为 2，残差块堆栈数为 2，扩张系数为 $[1, 2^1, 2^2, 2^3]$ ，Dropout 率为 0.05；TCN 后部接入两层全连接层，本文的电力负荷预测是单步预测，故输出层神经元个数为 1。

3.2 预测误差评价指标

负荷预测属于曲线回归问题，因此本文将选择平均绝对误差 M_{AE} 、均方根误差 R_{MSE} 、平均绝对百分比误差 M_{APE} 和决定系数 R^2 作为台区负荷预测结果的评价指标，各评价指标的计算公式分别为

$$M_{AE} = \frac{1}{E} \sum_{e=1}^E |z_e - \bar{z}_e| \quad (11)$$

$$R_{MSE} = \frac{1}{E} \sum_{e=1}^E (z_e - \bar{z}_e)^2 \quad (12)$$

$$M_{APE} = \frac{1}{E} \sum_{e=1}^E \left| \frac{z_e - \bar{z}_e}{z_e} \right| \times 100\% \quad (13)$$

$$R^2 = 1 - \frac{\sum_{e=1}^E (z_e - \bar{z}_e)^2}{\sum_{e=1}^E (z_e - \bar{z})^2} \quad (14)$$

式中： E 为样本总数； z_e 和 \bar{z}_e 分别表示 e 时刻下台区负荷的真实值和预测值； \bar{z} 表示台区负荷真实值的平均值。上述指标中， M_{AE} 、 R_{MSE} 、 M_{APE} 越小， R^2 越大，说明模型的预测误差越小，模型的预测性能越好。

3.3 样本输入步长确定

模型的输入与输出步长之间存在着某种时序上的依赖关系，输入步长可以理解为预测目标前对应的采集点个数，输出步长则对应每次预测的采集点个数。每次预测一个采集点负荷为单步预测，每次预测多个采样点为多步预测。通过大量研究者的时间步长实验总结，输出步长越长，预测精度反而会下降，因此本文选择单步预测。

合适的样本输入步长设置对模型的预测精度存在一定的影响, 下面将对输入步长进行实验, 选择经验测试的方法对最佳步长进行调节实验, 选择 CBAM-TCN 模型作为基准模型, 设置输入步长分别为 4、6、8、...、24 步, 实验结果对应的 M_{AE} 、 R_{MSE} 、 M_{APE} 和 R^2 评价指标如图 9 所示。

从不同输入步长的仿真结果可以看出, 当输入步长为 16 时, CBAM-TCN 模型的 M_{AE} 、 R_{MSE} 和 M_{APE} 评价指标最小, R^2 最大, 工业台区的 4 项指标分别为 8.618 kW、11.431 kW、4.037% 和 0.942, 居民台区的 4 项指标分别为 4.6 kW、6.194 kW、4.183% 和 0.957, 两个台区的 M_{APE} 和 R^2 综合评价指标相近。输入步长在 16—24 步变化时, M_{AE} 、 R_{MSE} 和 M_{APE} 误差指标在整体呈现增大趋势, R^2 整体呈现减小趋势, 由于输入步长变大、模型输入维度变大时模型

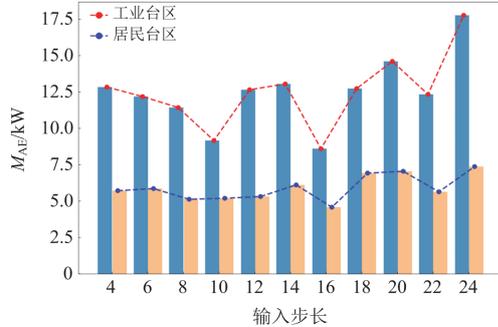
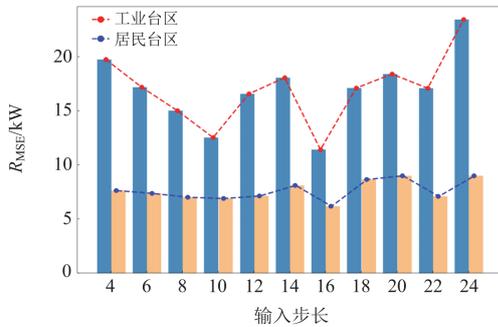
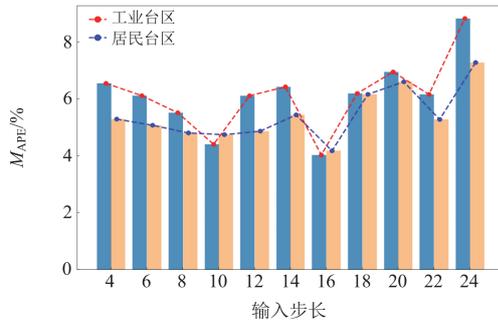
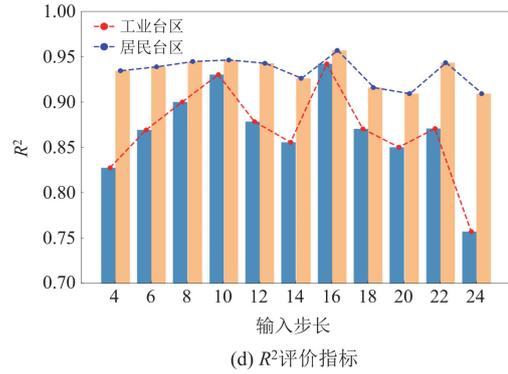
(a) M_{AE} 评价指标(b) R_{MSE} 评价指标(c) M_{APE} 评价指标(d) R^2 评价指标

图 9 不同输入步长下的评价指标

Fig. 9 Evaluation indicators under different input steps

不能从较长的时间数据序列中挖掘到与输出点之间的联系, 导致模型预测精度较低。因此接下来均采用输入步长为 16、输出步长为 1 进行实验。

3.4 CBAM-TCN 有效性验证

为了验证 CBAM-TCN 的有效性, 本节引入反向传播神经网络(back propagation neural network, BPNN)、CNN、TCN 模型作为比较, 对典型台区测试集的预测曲线如图 10 和图 11 所示, 对应的评价指标如表 2 所示。

由图 10、图 11 和表 2 可以得出:

对比两个典型台区 BPNN、CNN 和 TCN 模型的预测结果, CNN 和 TCN 模型预测结果均比 BPNN 要好, CNN 和 TCN 模型作为深度学习模型, 具有较为强大的数据处理能力和非线性拟合能力。TCN 的预测精度最高, 对于工业台区, TCN 相比于 BPNN, M_{AE} 、 R_{MSE} 、 M_{APE} 和 R^2 分别提升了 28.80%、19.88%、20.61% 和 5.54%; 对于居民台区, TCN 相比于 BPNN, M_{AE} 、 R_{MSE} 、 M_{APE} 和 R^2 分别提升了 15.87%、11.71%、17.27% 和 2.82%。由此可见, 相较于 BPNN、CNN 模型, TCN 能够挖掘各特征和电力

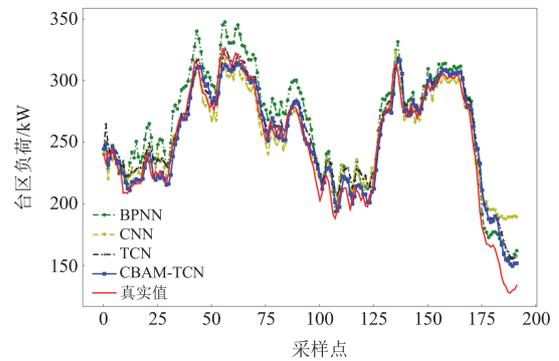


图 10 不同神经网络对工业台区测试集的预测曲线

Fig. 10 Prediction curves of different neural networks for industrial area test sets

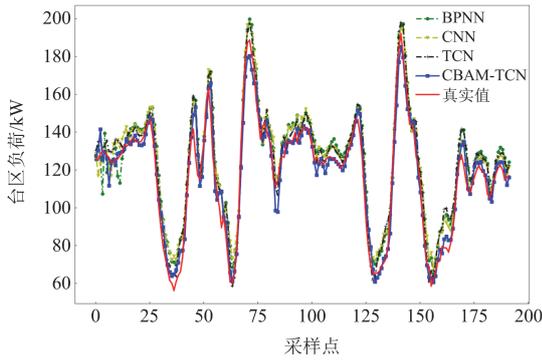


图 11 不同神经网络对居民台区测试集的预测曲线
Fig. 11 Prediction curves of different neural networks for residential area test sets

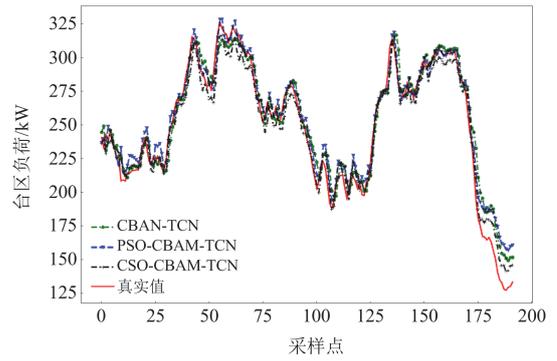


图 12 不同优化算法下对工业台区测试集的预测曲线
Fig. 12 Prediction curve of industrial area test set under different optimization algorithms

表 2 不同神经网络预测模型下的评价指标
Table 2 Evaluation indexes under different neural network prediction models

数据集	模型	评价指标			
		M_{AE}/kW	R_{MSE}/kW	$M_{APE}/\%$	R^2
工业台区	BPNN	15.345	17.431	6.588	0.866
	CNN	10.691	16.164	5.484	0.885
	TCN	10.926	13.966	5.230	0.914
	CBAM-TCN	8.618	11.431	4.037	0.942
居民台区	BPNN	8.662	10.180	8.34	0.885
	CNN	7.919	9.350	7.840	0.903
	TCN	7.287	8.988	6.900	0.910
	CBAM-TCN	4.600	6.194	4.183	0.957

负荷在时序上的依存关系，可以有效地对台区负荷时间序列进行预测。

对比 CBAM-TCN 和 TCN 的预测结果, CBAM-TCN 的预测曲线与台区负荷真实值最为贴近, 工业台区、居民台区的 M_{AE} 、 R_{MSE} 、 M_{APE} 和 R^2 评价指标分别提升了 21.12%、18.15%、22.81%、3.06% 和 36.87%、31.09%、39.38%、5.16%，提示效果显著，侧面反映出 CBAM 的有效性。CBAM 的加入没有带来过多的参数量，但是让网络对关键特征给予了更多的关注，可以有效地提升预测精度。

3.5 模型参数优化有效性验证

本文提出了一种基于 CSO 优化 CBAM-TCN 的预测模型(CSO-CBAM-TCN), 为验证其有效性, 引入 CBAM-TCN、PSO-CBAM-TCN 作为对比模型, 其中 CSO 的参数设置为: 迭代次数为 100, 横向、纵向交叉概率分别为 1、0.8, 种群大小为 30; PSO 的参数设置为: 惯性因子为 0.8, 粒子速度为 0.5, 种群大小为 30。两个典型台区的电力负荷预测曲线和对应的评价指标分别如图 12、图 13 和表 3 所示。

由图 12、图 13 和表 3 可以得出:

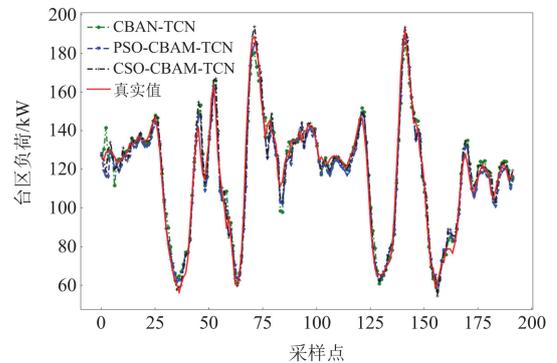


图 13 不同优化算法下对居民台区测试集的预测曲线
Fig. 13 Prediction curve of residential area test set under different optimization algorithms

表 3 不同优化算法情况下的评价指标
Table 3 Evaluation indicators under different optimization algorithms

数据集	模型	评价指标			
		M_{AE}/kW	R_{MSE}/kW	$M_{APE}/\%$	R^2
工业台区	CBAM-TCN	8.618	11.431	4.037	0.942
	PSO-CBAM-TCN	7.831	10.719	3.849	0.949
	CSO-CBAM-TCN	7.392	9.005	3.211	0.964
居民台区	CBAM-TCN	4.600	6.194	4.183	0.957
	PSO-CBAM-TCN	4.506	5.832	4.039	0.962
	CSO-CBAM-TCN	4.175	5.653	3.766	0.964

CBAM-TCN 和经过优化之后的模型相比较, 优化后的模型预测精度显著得到提升, CSO-CBAM-TCN 的提升效果最大。同一数据集, 不同模型在在负荷转折处的预测曲线差异最大, CSO-CBAM-TCN 的预测曲线最贴近真实值曲线。在工业台区数据集中, CSO-CBAM-TCN 比 CBAM-TCN, M_{AE} 、 R_{MSE} 、 M_{APE} 和 R^2 分别提升了 14.23%、21.22%、20.46% 和 2.34%; 在居民台区数据集中, CSO-

CBAM-TCN 比 CBAM-TCN, M_{AE} 、 R_{MSE} 、 M_{APE} 和 R^2 分别提升了 9.24%、8.73%、9.97%和 0.73%。可见, 选用群智能优化算法对神经网络的权值和阈值进行细微调整, 有助于避免这些参数陷入局部最优, 从而提升模型的拟合能力。

分析综合评价指标 R^2 , 对于工业和居民台区数据测试集, CSO-CBAM-TCN 比 PSO-CBAM-TCN 分别提升了 1.58%和 0.2%, 前者预测精度更高, 恰能说明 CSO 的优势所在。

图 14 为优化过程中的损失变化图, CSO 存在纵向交叉和横向交叉机制, 在解空间中具有更强的探索能力, 能够快速找到最优解, 完成收敛, 避免解过早地处于早熟状态。实验证实了 CSO 优化 CBAM-TCN 的可行性和有效性。

3.6 所提模型在不同季节下的表现

考虑到台区用户在不同季节的用电习惯差异, 配变台区的电力负荷存在一定的波动情况。为了保

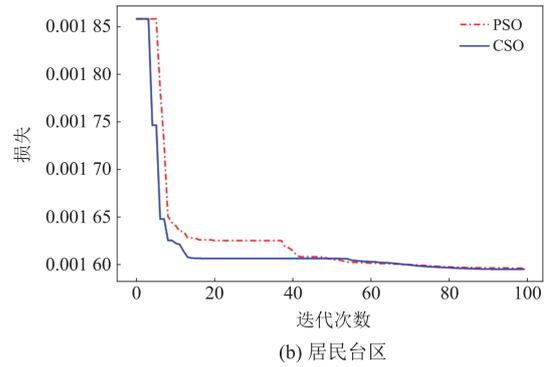
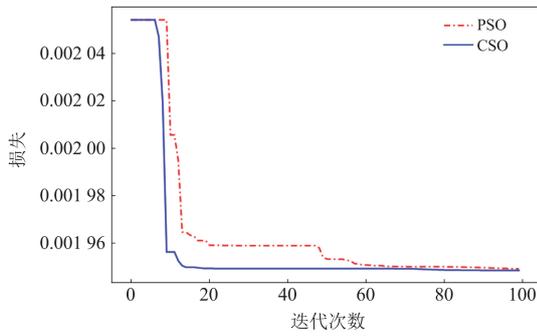


图 14 优化过程中的适应度函数变化图

Fig. 14 Fitness function change diagram in the optimization process

证试验结果的合理性和科学性, 验证本文提出的 CSO-CBAM-TCN 模型在不同季节下的台区负荷预测能够保持较高的精度, 本节分别选取了工业台区和居民台区在 2021 年 1 月(冬季)、4 月(春季)、7 月(夏季)和 10 月(秋季)4 个数据集进行建模, 选择每月最后两天作为测试集, 预测模型均采用本文提出的 CSO-CBAM-TCN 模型。仿真实验结果如图 15 和图 16 所示, 评价指标如表 4 所示。

通过对比各项评价指标, 不难发现本文所提 CSO-CBAM-TCN 模型在不同台区、不同季节数据集下的综合评价指标相近, 这一定程度上说明所提模型适用于不同场景下的台区电力负荷预测, 泛化性能较好。



(a) 工业台区

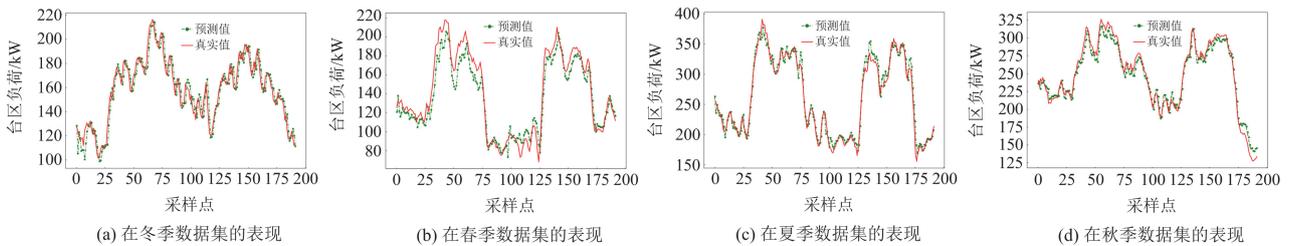


图 15 所提模型在工业台区不同季节下的表现

Fig. 15 Performance of the proposed model in industrial areas in different seasons

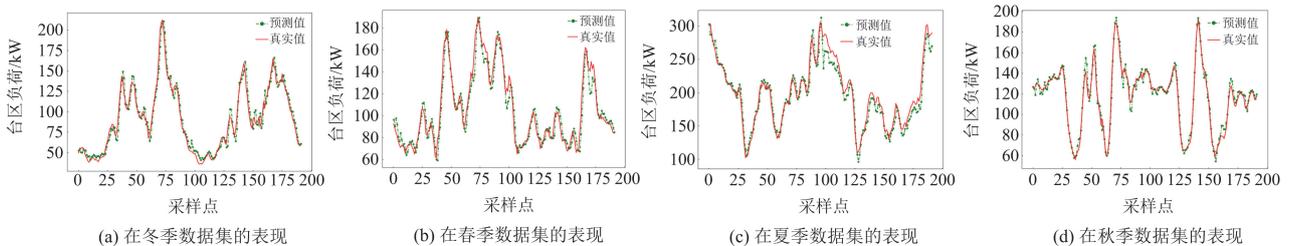


图 16 所提模型在居民台区不同季节下的表现

Fig. 16 Performance of the proposed model in residential areas in different seasons

表 4 所提模型在不同季节数据集下的评价指标

Table 4 Evaluation indicators of the proposed model under different seasonal datasets

台区 类型	季节	评价指标			
		MAE/kW	$RMSE/kW$	$MAPE/\%$	R^2
工业 台区	冬(1月)	5.678	7.107	3.701	0.927
	春(4月)	9.175	11.403	6.547	0.929
	夏(7月)	9.082	12.106	3.626	0.964
	秋(10月)	7.392	9.005	3.211	0.964
居民 台区	冬(1月)	5.777	7.382	6.935	0.964
	春(4月)	5.799	7.72	5.52	0.947
	夏(7月)	9.744	12.51	4.876	0.936
	秋(10月)	4.175	5.653	3.766	0.966

4 结论

针对低压配变台区预测精度低的问题, 本文基于 TCN 和 CSO 的低压台区电力负荷预测方法 CSO-CBAM-TCN, 利用实测数据进行仿真验证。实验证明 TCN 相对于 BPNN、CNN 等模型具有更高的预测精度, TCN 能够有效提取高维特征之间的紧密耦合关系和时序规律隐含关系。在 CBAM-TCN 预测模型中, 提出使用 CBAM 对模型特征输入在通道和空间两个维度赋予不同的关注度, 相对于 TCN 模型, CBAM 可为模型带来显著的精度提升。所提 CSO-CBAM-TCN 模型中, 实验证明采用 CSO 算法优化 CBAM-TCN 权值和阈值参数的可行性, 可进一步提升预测精度。而与其他对比模型相比, 所提 CSO-CBAM-TCN 具有更优异的预测性能, 模型泛化性能更好。

参考文献

- [1] 魏艳霞, 许锐琦, 殷豪. 基于深度信念网络的配变台区跳闸预测[J]. 智慧电力, 2020, 48(6): 119-124.
WEI Yanxia, XU Ruiqi, YIN Hao. Trip prediction of distribution transformer station based on deep belief network[J]. Smart Power, 2020, 48(6): 119-124.
- [2] ZHANG X, KONG X, YAN R. Data-driven cooling, heating and electrical load prediction for building integrated with electric vehicles considering occupant travel behavior[J]. Energy, 2023, 264.
- [3] 徐利美, 赵金, 李裕民, 等. 考虑多变量建模的中期负荷预测模型[J]. 南方电网技术, 2024, 18(11): 79-87.
XU Limei, ZHAO Jin, LI Yumin, et al. Medium-term load forecasting model considering multivariate modeling[J]. Southern Power System Technology, 2024, 18(11): 79-87.
- [4] SUN Fan, HUO Yaojia, FU Lei, et al. Load-forecasting method for IES based on LSTM and dynamic similar days with multi-features[J]. Global Energy Interconnection, 2023, 6(3): 285-296.
- [5] 高洪超, 陈启鑫, 雷振, 等. 考虑事件数据的产消者负荷预测方法及极端误差抑制策略[J]. 电力系统自动化, 2023, 47(13): 47-55.
GAO Hongchao, CHEN Qixin, LEI Zhen, et al. Load forecasting method and extreme error suppression strategy for prosumers considering event data[J]. Automation of Electric Power Systems, 2023, 47(13): 47-55.
- [6] 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.
- [7] 黄星知, 刘星, 张文娟, 等. 改进决策树在电网超短期负荷预测中的应用[J]. 能源与环境, 2022, 44(7): 265-271.
HUANG Xingzhi, LIU Xing, ZHANG Wenjuan, et al. Application of improved decision tree in power grid ultra-short-term load forecast[J]. China Energy and Environmental Protection, 2022, 44(7): 265-271.
- [8] 臧海祥, 陈玉伟, 程礼临, 等. 基于多尺度分量特征学习的用户级超短期负荷预测[J]. 电网技术, 2024, 48(6): 2584-2592.
ZANG Haixiang, CHEN Yuwei, CHENG Lilin, et al. User level ultra-short-term load forecasting based on multi-scale component feature learning[J]. Power System Technology, 2024, 48(6): 2584-2592.
- [9] 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.
- [10] 陈顺. 基于深度学习的输电线路短时负荷电流预测[D]. 广州: 广东工业大学, 2022.
CHEN Shun. Short term load current prediction of transmission lines based on deep learning[D]. Guangzhou: Guangdong University of Technology, 2022.
- [11] 孟安波, 陈顺, 王陈恩, 等. 基于混沌 CSO 优化时序注意力 GRU 模型的超短期风电功率预测[J]. 电网技术, 2021, 45(12): 4692-4700.
MENG Anbo, CHEN Shun, WANG Chenen, et al. Ultra-short-term wind power prediction based on chaotic CSO optimized temporal attention GRU model[J]. Power System Technology, 2021, 45(12): 4692-4700.
- [12] 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

- and Energy Systems, 2022, 137.
- [13] 欧阳福莲, 王俊, 周杭霞. 基于改进迁移学习和多尺度 CNN-BiLSTM-Attention 的短期电力负荷预测方法[J]. 电力系统保护与控制, 2023, 51(2): 132-140.
OUYANG Fulian, WANG Jun, ZHOU Hangxia. Short-term power load forecasting method based on improved hierarchical transfer learning and multi-scale CNN-BiLSTM-Attention[J]. Power System Protection and Control, 2023, 51(2): 132-140.
- [14] 袁郁, 杨超, 郑伟铭, 等. 基于 Bi-SRNN 的联邦学习区域电力短期负荷预测模型[J]. 电网与清洁能源, 2023, 39(10): 45-55.
YUAN Yu, YANG Chao, ZHENG Weiming, et al. The federated learning power short term load forecasting model based on Bi-SRNN[J]. Power System and Clean Energy, 2023, 39(10): 45-55.
- [15] 林帆, 张耀, 东琦, 等. 基于分位数插值和深度自回归网络的光伏出力概率预测[J]. 电力系统自动化, 2023, 47(9): 79-87.
LIN Fan, ZHANG Yao, DONG Qi, et al. Probability prediction of photovoltaic output based on quantile interpolation and deep autoregressive network[J]. Automation of Electric Power Systems, 2023, 47(9): 79-87.
- [16] 赵倩, 郑贵林. 基于 WD-LSSVM-LSTM 模型的短期电力负荷预测[J]. 电测与仪表, 2023, 60(1): 23-28.
ZHAO Qian, ZHENG Guilin. Short-term load forecasting based on WD-LSSVM-LSTM model[J]. Electrical Measurement & Instrumentation, 2023, 60(1): 23-28.
- [17] 张翔颖, 杨永标, 徐青山, 等. 基于多时段相似日理论的光伏功率组合预测方法[J]. 南方电网技术, 2023, 17(2): 57-65.
ZHANG Xiangying, YANG Yongbiao, XU Qingshan, et al. Photovoltaic power combination prediction method based on multi-temporal similarity day theory[J]. Southern Power System Technology, 2023, 17(2): 57-65.
- [18] 尹元亚, 潘文虎, 赵文广, 等. 基于 CEEMDAN 和 BiLSTM-AM 的超短期风速预测方法[J]. 电测与仪表, 2024, 61(9): 77-84.
YIN Yuanya, PAN Wenhui, ZHAO Wenguang, et al. Ultra-short-term wind speed prediction method based on CEEMDAN and BiLSTM-AM[J]. Electrical Measurement & Instrumentation, 2024, 61(9): 77-84.
- [19] TANG X, CHEN H, XIANG W, et al. Short-term load forecasting using channel and temporal attention based temporal convolutional network[J]. Electric Power Systems Research, 2023, 205.
- [20] LIU M, QIN H, CAO R, et al. Short-term load forecasting based on improved TCN and DenseNet[J]. IEEE Access, 2022, 10(1): 115945-115957.
- [21] 樊江川, 于昊正, 刘慧婷, 等. 基于多分支门控残差卷积神经网络的短期电力负荷预测[J]. 中国电力, 2022, 55(11): 155-162, 174.
FAN Jiangchuan, YU Haozheng, LIU Huiting, et al. Short-term load forecasting based on multi-branch residual gated convolution neural network[J]. Electric Power, 2022, 55(11): 155-162, 174.
- [22] 梁露, 张智晟. 基于多尺度特征增强 DHTCN 的电力系统短期负荷预测研究[J]. 电力系统保护与控制, 2023, 51(10): 172-179.
LIANG Lu, ZHANG Zhisheng. Short-term load forecasting of a power system based on multi-scale feature enhanced DHTCN[J]. Power System Protection and Control, 2023, 51(10): 172-179.
- [23] 王海鹏. 注意力机制下多尺度时间卷积的短期电力负荷预测[D]. 南昌: 南昌大学, 2022.
WANG Haipeng. Short-term load forecasting based on multi-scale temporal convolutional networks and attention mechanism[D]. Nanchang: Nanchang University, 2022.
- [24] 王雪, 刘林, 刘文迪, 等. 基于纵横交叉算法的新型电力系统惯量延迟优化控制策略[J]. 中国电力, 2024, 57(7): 12-20.
WANG Xue, LIU Lin, LIU Wendi, et al. A novel inertia delay optimization control strategy for new power systems based on crisscross optimization[J]. Electric Power, 2024, 57(7): 12-20.
- [25] 姚芳, 汤俊豪, 陈盛华, 等. 基于 ISSA-CNN-GRU 模型的电动汽车充电负荷预测方法[J]. 电力系统保护与控制, 2023, 51(16): 158-167.
YAO Fang, TANG Junhao, CHEN Shenghua, et al. Charging load prediction method for electric vehicles based on an ISSA-CNN-GRU model[J]. Power System Protection and Control, 2023, 51(16): 158-167.
- [26] TANG X, LI Z, XU X, et al. Multi-objective economic emission dispatch based on an extended crisscross search optimization algorithm[J]. Energy, 2022, 244.

收稿日期: 2024-12-04; 修回日期: 2025-04-21

作者简介:

丁伟锋(1995—), 男, 通信作者, 硕士, 研究方向为机器学习在电网自动化中的应用; E-mail: 179302400@qq.com

周震震(1986—), 男, 硕士, 高级工程师, 研究方向为数字电网。E-mail: Zhouzhengzhen@im.ehv.csg

(编辑 周金梅)