

DOI: 10.19783/j.cnki.pspc.231357

基于改进 Q 学习算法和组合模型的超短期电力负荷预测

张丽^{1,2}, 李世情¹, 艾恒涛¹, 张涛¹, 张宏伟³

(1. 河南理工大学电气工程与自动化学院, 河南 焦作 454003; 2. 河南省煤矿装备智能检测与控制重点实验室, 河南 焦作 454003; 3. 国网山西省电力公司临汾供电公司, 山西 临汾 041000)

摘要: 单一模型在进行超短期负荷预测时会因负荷波动而导致预测精度变差, 针对此问题, 提出一种基于深度学习算法的组合预测模型。首先, 采用变分模态分解对原始负荷序列进行分解, 得到一系列的子序列。其次, 分别采用双向长短期记忆网络和优化后的深度极限学习机对每个子序列进行预测。然后, 利用改进 Q 学习算法对双向长短期记忆网络的预测结果和深度极限学习机的预测结果进行加权组合, 得到每个子序列的预测结果。最后, 将各个子序列的预测结果进行求和, 得到最终的负荷预测结果。以某地真实负荷数据进行预测实验, 结果表明所提预测模型较其他模型在超短期负荷预测中表现更佳, 预测精度达到 98% 以上。

关键词: Q 学习算法; 负荷预测; 双向长短期记忆; 深度极限学习机; 灰狼算法

Ultra-short-term power load forecasting based on an improved Q-learning algorithm and combination model

ZHANG Li^{1,2}, LI Shiqing¹, AI Hengtao¹, ZHANG Tao¹, ZHANG Hongwei³

(1. School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454003, China;
2. Henan Key Laboratory of Intelligent Detection and Control of Coal Mine Equipment, Jiaozuo 454003, China;
3. Linfen Power Supply Company, State Grid Shanxi Electric Power Company, Linfen 041000, China)

Abstract: The prediction accuracy of a single model will deteriorate because of load fluctuations when making ultra-short-term load forecasting. To solve this problem, this paper proposes a combinatorial prediction model based on a deep learning algorithm. First, variational mode decomposition is used to decompose the original load sequence to obtain a series of sub-sequences. Secondly, a bidirectional long short-term memory network and an optimized deep extreme learning machine are used to predict each sub-sequence. Thirdly, the improved Q-learning algorithm is used to weight the prediction results of the bidirectional long short-term memory network and of the deep extreme learning machine to obtain those of each sub-sequence. Finally, the prediction results of each subseries are summed to obtain the final load prediction results. The results show that the prediction model proposed in this paper performs better than other models in ultra-short-term load forecasting, with a prediction accuracy of more than 98%.

This work is supported by the National Natural Science Foundation of China (No. 52177039).

Key words: Q-learning algorithm; load forecasting; bi-directional long short-term memory; deep extreme learning machine; grey wolf optimization algorithm

0 引言

随着我国能源结构的进一步转型, 电动汽车等新型电力负荷的可移动性使负荷的随机性与不确定性更加复杂^[1], 微电网在孤岛运行状态下对于负荷侧

的波动更为敏感, 因此精准的负荷预测对于微电网的安全运行具有重大意义^[2]。

负荷预测按照预测期限的长短可以分为超短期、短期、中期以及长期预测^[3], 超短期负荷预测是预测未来 15 min 到几小时的负荷值, 用于在线监控电力设备的运行状况。负荷预测一般有回归分析法和时间序列法^[4]。中长期负荷预测受外界因素的影响较大, 因此一般采用回归分析法。超短期负荷预测与历史负荷关联度较高, 一般采用时间序列法。

基金项目: 国家自然科学基金项目资助(52177039); 河南省高等学校重点科研项目资助(24A470006); 河南省科技攻关项目(242102241027)

近年来深度学习方法被广泛应用于负荷预测^[5], 文献[6]采用最小二乘支持向量机对模态分解后的子序列进行预测, 文献[7]采用改进时域卷积网络对电力负荷进行预测, 文献[8-9]采用双向长短期记忆(Bi-directional long short-term memory, BiLSTM)网络进行超短期电力负荷预测。

由于电力负荷本质上存在着波动性, 通过模态分解可进一步提取序列的特征, 文献[10]采用变分模态分解(variational model decomposition, VMD)对电力负荷数据进行分解, 通过模态分解降低了数据的复杂度, 解决了原始数据存在的随机性问题, 随后采用时域卷积网络提取数据之间的关联性, 使预测精度得到了提高。然而单一模型对不同频域的子序列预测效果不同, 限制了预测精度的进一步提高, 因此组合模型越来越多地应用到负荷预测中。文献[11]采用误差倒数法将长短期记忆(LSTM)网络与极致梯度提升两个基础模型组合起来, 但误差倒数法对异常值敏感度较高, 会影响权重分配的准确性, 因此需要一种鲁棒性更好且更加合理的方法对模型的预测结果进行权重分配。文献[12]采用 Q 学习(Q-Learning, Q-L)算法对模型的光伏功率预测结果进行权重分配, 提高了预测精度, 但 Q-L 算法的动作选择采用贪婪机制, 因此其智能化不够。

针对上述问题, 本文提出了一种改进 Q 学习(improved Q-L, IQ-L)算法为组合模型中基础模型的预测结果进行权重分配, 构建了变分模态分解、双向长短期记忆网络、改进灰狼算法(improved grey wolf optimizer, IGWO)、优化深度极限学习机(IGWO optimizes deep extreme learning machine, IDELM)的超短期负荷预测模型(VMD-BiLSTM-IQ-IDELM)。VMD将原始负荷序列分解成不同频域的分量, 提取负荷序列在各个频域的特征, 每个分量同时采用 BiLSTM 和 IDELM 进行预测, 然后采用 IQ-L 算法对预测结果进行权重分配, 在面对不同频域的子序列时能够调整权重大小, 该权重分配方法综合利用了两种基础模型的优点。以某地真实负荷数据进行实验, 验证了本文所提方法的有效性。

1 超短期负荷预测思路

本文提出的预测模型框架如图 1 所示, 负荷本质上存在着波动性, 经过对负荷数据进行模态分解后, 可以得到一组规律性较强的分量, 但分解后的子序列中存在着模态混叠现象, 单一模型预测具有很大的局限性, 因此对每个模态分量采用组合模型预测的方式。

1) 对负荷数据进行预处理, 对缺失负荷数据采

用历史负荷数据中相同时刻的数据补齐, 随后将一

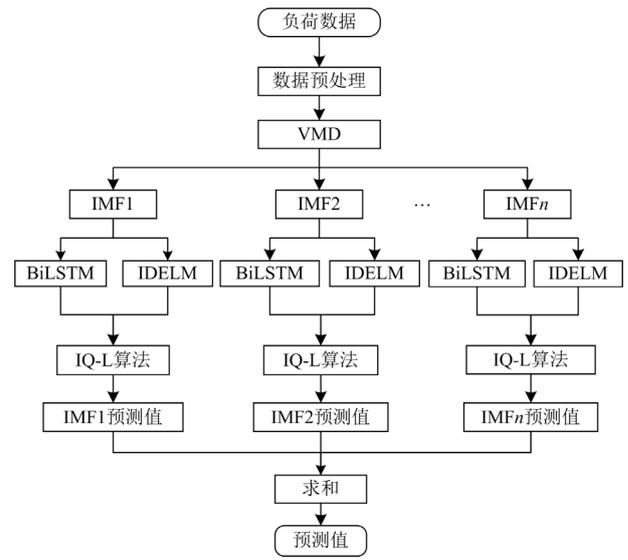


图 1 预测模型框架

Fig. 1 Framework of forecasting model

维负荷数据进行重构。

2) 采用分解效果较好的 VMD 将负荷数据分解成一系列不同频域特征的分量(IMF1—IMF_n), 每个子序列同时采用 BiLSTM 与 IDELM 进行预测。

3) 采用 BiLSTM 对输入网络的负荷子序列进行时序特征的提取^[13], 然后利用其对时间序列变化趋势的敏感性对负荷子序列进行预测。

4) 采用 IDELM 对输入网络的负荷子序列进行预训练和降维操作, 实现对负荷数据特征的提取, 利用其良好的预测精度及泛化能力对负荷子序列进行预测。

5) 采用 IQ-L 算法对 BiLSTM 和 IDELM 的预测结果进行权重分配, 形成优势互补, 提高预测精度。

6) 将各个子序列的预测结果进行相加, 得到负荷的预测值。

2 预测模型的构建

2.1 数据预处理

本文采用时间序列滚动预测的方式, 利用前 24 个时刻的负荷值预测下一个时刻的负荷值。首先将一维负荷序列进行重新构造, 获得输入矩阵和输出矩阵, 分别如式(1)和式(2)所示。

$$\mathbf{X}_{in} = \begin{bmatrix} n_1 & n_2 & n_3 & \cdots & n_{1416} \\ n_2 & n_3 & n_4 & \cdots & n_{1417} \\ \vdots & \vdots & \vdots & & \vdots \\ n_{24} & n_{25} & n_{26} & \cdots & n_{1439} \end{bmatrix} \quad (1)$$

$$\mathbf{Y}_{out} = [n_{25} \quad n_{26} \quad n_{27} \quad \cdots \quad n_{1440}] \quad (2)$$

为了减少数据自身对预测结果的影响, 对数据进行归一化处理, 使数据处于[0,1]之间。

2.2 变分模态分解(VMD)

VMD 相对于其他模态分解方法有更坚实的数学理论基础且能够有效地减少模态混叠现象的发生, 但分解效果受分解模态数 K 的影响较大, K 值的确定一般采用中心频率法^[14], 当两个相邻模态分量的中心频率相距较近时, 会发生模态混叠现象。本文 K 值从 3 开始选取, 可以得到不同 K 值下各个模态的中心频率, 一旦某个 K 值下两个模态分量的中心频率相距较近, 便选取前一个 K 值作为 VMD 的分解层数。

2.3 双向长短期记忆(BiLSTM)神经网络的构建

BiLSTM 由前向 LSTM 和后向 LSTM 构成, 前向 LSTM 正常提取负荷序列的历史信息, 后向 LSTM 反向追溯负荷序列的历史信息, 相较于单一的 LSTM, 这种网络结构能够对负荷序列的时序特征进行深度挖掘, 有效地提高了模型的非线性学习能力以及预测的准确度。BiLSTM 的神经网络结构如图 2 所示。

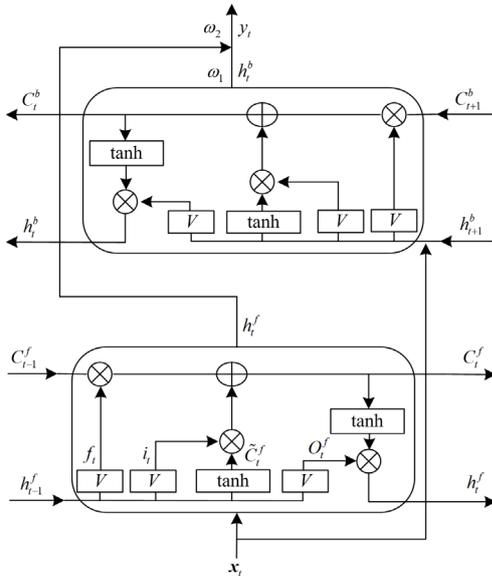


图 2 BiLSTM 神经网络结构

Fig. 2 Structure of BiLSTM neural network

图 2 中: x_t 为 t 时刻隐藏层的输入矩阵; h_t^f 、 h_t^b 分别为 t 时刻 f 长短期记忆网络和 b 长短期记忆网络隐藏层的输出; C_t^f 、 C_t^b 分别为 t 时刻 f 长短期记忆网络和 b 长短期记忆网络隐藏层的细胞记忆; V 为激活函数; “ \oplus ” 为元素相加; “ \otimes ” 为元素相乘。本文中: x_t 为 t 时刻输入网络的 24 个时刻的负荷值; h_t^f 、 h_t^b 为网络输出对应的预测负荷值。单个 LSTM 神经网络的计算过程为

$$\begin{aligned} f_t &= V(\omega_f \cdot [h_{t-1}^f, x_t] + b_f) \\ i_t &= V(\omega_i \cdot [h_{t-1}^f, x_t] + b_i) \\ \tilde{C}_t^f &= \tanh(\omega_c \cdot [h_{t-1}^f, x_t] + b_c) \\ C_t^f &= f_t \otimes C_{t-1}^f \oplus i_t \otimes \tilde{C}_t^f \\ O_t^f &= V(\omega_o \cdot [h_{t-1}^f, x_t] + b_o) \\ h_t^f &= O_t^f \otimes \tanh(C_t^f) \end{aligned} \quad (3)$$

式中: ω_f 、 ω_i 、 ω_c 、 ω_o 为权重矩阵; b_f 、 b_i 、 b_c 、 b_o 为偏置量矩阵; \tanh 为双曲正切激活函数; f_t 、 i_t 、 O_t^f 、 \tilde{C}_t^f 分别表示遗忘门、输入门、输出门和候选记忆细胞单元的输出。

BiLSTM 通过综合前向 LSTM 和后向 LSTM 的预测值得到最终的预测值 y_t , 其中 ω_1 、 ω_2 分别为对应的隐藏层输出权重。

$$y_t = \omega_2 \cdot h_t^f + \omega_1 \cdot h_t^b \quad (4)$$

2.4 深度极限学习机(DELM)模型的构建

根据文献[15], 本文构建的 DELM 是由多个 ELM-AE 组成的, ELM-AE 能够提取输入数据的特征, ELM-AE 隐藏层的输出矩阵 H 为

$$H = g(\omega X + B) \quad (5)$$

式中: $g(\cdot)$ 是激活函数; ω 为 ELM-AE 输入层到隐藏层的权重矩阵; X 为 ELM-AE 的输入数据矩阵; B 为偏置量矩阵。

对于降维和高维特征表达, ELM-AE 的隐藏层输出权重矩阵 β 为

$$\beta = \left(\frac{I}{C} + H^T H \right)^{-1} H^T X \quad (6)$$

式中: C 为正则化参数; I 为单位矩阵。

对于等维度表达, 输出权重矩阵 β 为

$$\beta = H^{-1} X \quad (7)$$

DELM 的结构如图 3 所示, 输入数据被逐层地提取特征信息, 使模型可以学习到比原始数据更有用的特征, 有效地提高了模型的预测精度。

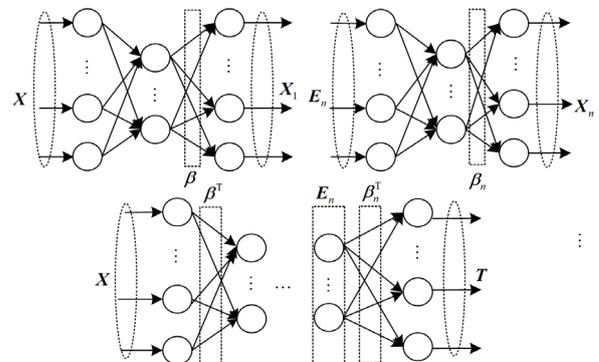


图 3 深度极限学习机的网络结构

Fig. 3 Network structure of the DELM

DELM 在 ELM-AE 的训练过程中, ELM-AE 输入层权重采用随机初始化的方式, 因此预测结果波动性较大且稳定性较差。为了解决此问题, 采用优化算法来优化其输入层权重, 以达到理想的预测结果。因此, 本文采用改进灰狼算法对其进行优化。

2.5 改进灰狼算法优化深度极限学习机

2.5.1 传统灰狼算法存在的问题

传统 GWO 具有强大的全局寻优能力、收敛迅速及参数较少的特点^[16], 因此受到广泛关注。但该算法也同时存在着以下问题。

1) GWO 的收敛因子呈线性递减, 然而实际种群的搜索是一个非线性的过程, 线性的收敛因子不能很好地平衡算法前期的探索能力和后期的开发能力。

2) GWO 的位置更新方程中, 以头狼引导位置的更新且三只头狼对位置更新的影响力相同^[17], 因此算法不仅易陷入局部最优, 而且收敛速度较慢。

针对以上问题, 本文从收敛因子和位置更新方程两个方面对算法进行改进。

2.5.2 收敛因子的改进

GWO 的收敛因子影响灰狼种群优化迭代过程的探索能力和开发能力^[18], 为了合理表征实际的优化过程, 改进的收敛因子 a 如式(8)所示。

$$a = \cos\left(\frac{t}{t_{\max}} \cdot \pi\right) + 1 + k \cdot R \quad (8)$$

式中: 第一部分保证收敛因子呈余弦曲线形式下降; t 是当前迭代次数; t_{\max} 是最大迭代次数; $k \cdot R$ 为扰动, 使算法在迭代过程中跳出局部最优, 其中 k 是控制扰动大小的参数, 取 0.05; R 为 [0,1] 内的随机数。

改进前后的收敛因子对比如图 4 所示。由图 4 可以看到, 在狼群包围的早期阶段, 收敛因子 a 缓慢下降, 便于狼群进行大范围的探索, 达到了灰狼种群多样化的目的, 提高了算法前期的探索能力。在狼群包围的后期阶段, 收敛因子 a 缓慢下降, 有利于加强算法后期的开发能力。

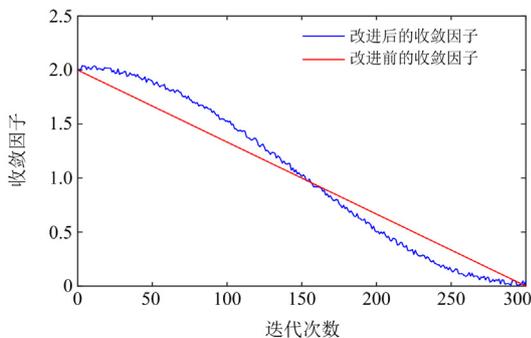


图 4 收敛因子对比

Fig. 4 Comparison of convergence factors

2.5.3 位置更新方程的改进

为了进一步提高 GWO 的收敛速度并使算法跳出局部最优, 本文提出了一种新的位置更新方程, 如式(9)所示。

$$\bar{X}(t+1) = \frac{A_1 \bar{X}_1 + A_2 \bar{X}_2 + A_3 \bar{X}_3}{3} + b_1 \cdot r_1 \cdot (\bar{X}_j(t) - \bar{X}(t)) \quad (9)$$

其中:

$$A_1 = \frac{\frac{1}{D_1}}{\frac{1}{D_1} + \frac{1}{D_2} + \frac{1}{D_3}} \quad (10)$$

$$A_2 = \frac{\frac{1}{D_2}}{\frac{1}{D_1} + \frac{1}{D_2} + \frac{1}{D_3}} \quad (11)$$

$$A_3 = \frac{\frac{1}{D_3}}{\frac{1}{D_1} + \frac{1}{D_2} + \frac{1}{D_3}} \quad (12)$$

式中: $\bar{X}(t+1)$ 为个体位置更新后得到的位置向量; b_1 取 0.05; $r_1 \in [0,1]$; \bar{X}_1 、 \bar{X}_2 、 \bar{X}_3 分别为 α 狼、 β 狼、 δ 狼的个体向量; $\bar{X}_j(t)$ 为群体在第 t 次迭代中随机选择的个体向量; A_1 、 A_2 、 A_3 分别为对应于 α 狼、 β 狼、 δ 狼的权重; D_1 、 D_2 、 D_3 分别为灰狼群体与 α 狼、 β 狼、 δ 狼间的欧氏距离。

当灰狼群体离 3 只头狼中任何一只距离较近时, 其权重就会增大, 随着迭代次数的增加, 所有的狼都会向 α 狼集中, 因此 α 狼的作用会越来越凸显, 从而加快收敛速度, 式(9)中的第二部分是群体中随机选取灰狼个体引导灰狼种群更新, 目的是防止算法陷入局部最优。

2.5.4 IGWO 性能测试

为了验证算法改进后的性能, 选用 GWO、粒子群优化(particle swarm optimization, PSO)算法、麻雀搜索算法(sparrow search algorithm, SSA)与 IGWO 进行对比验证, 测试函数的范围和最小值如表 1 所示, 表达式如式(13)所示。

表 1 测试函数

Table 1 Test function

函数	范围	最小值
F_1	[-100,100]	0
F_2	[-10,10]	0
F_3	[-5.12,5.12]	0
F_4	[-600,600]	0

$$\left\{ \begin{aligned} F_1 &= \sum_{i=1}^n x_i^2 \\ F_2 &= \sum_{i=1}^n |x_i| + \prod_{i=1}^n |x_i| \\ F_3 &= \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10] \\ F_4 &= \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos \frac{x_i}{\sqrt{i}} + 1 \end{aligned} \right. \quad (13)$$

评价指标为 30 次独立重复实验的平均最优适应度值和标准偏差, 平均最优适应度值 (M) 的计算

公式如式(14)所示, 其主要用于评估算法的搜索精度, 标准偏差 (S_{td}) 的计算公式如式(15)所示, 其主要用于评估算法的稳定性^[19], 种群数设置为 30, 迭代次数为 300。

$$M = \frac{1}{P} \sum_{i=1}^P f_i \quad (14)$$

$$S_{td} = \sqrt{\frac{1}{P-1} \sum_{i=1}^P (f_i - M)^2} \quad (15)$$

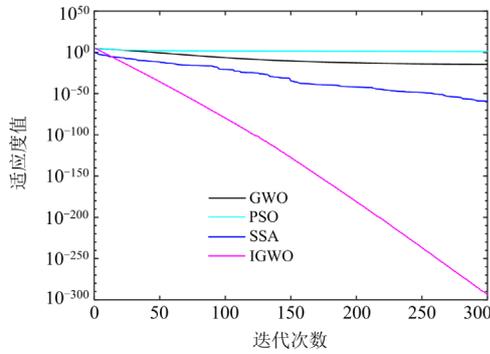
式中: P 是实验重复的次数; f_i 是每次独立运行的最优适应度值。评价指标的值如表 2 所示, 收敛曲线如图 5 所示。

表 2 评价指标对比

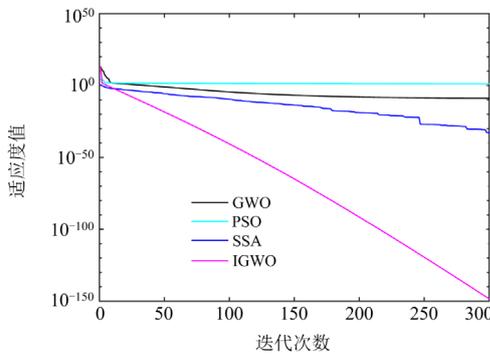
Table 2 Comparison of evaluation indexes

函数	GWO		PSO		SSA		IGWO	
	M	S_{td}	M	S_{td}	M	S_{td}	M	S_{td}
F_1	7.44×10^{-15}	1.48×10^{-14}	8.83	3.21	1.39×10^{-57}	7.66×10^{-57}	1.57×10^{-294}	0
F_2	1.45×10^{-9}	9.66×10^{-10}	15.26	4.52	1.31×10^{-33}	5.02×10^{-33}	2.11×10^{-149}	4.08×10^{-149}
F_3	7.81	6.69	108.27	40.66	0	0	0	0
F_4	0.0045	0.0093	0.6847	0.0956	0	0	0	0

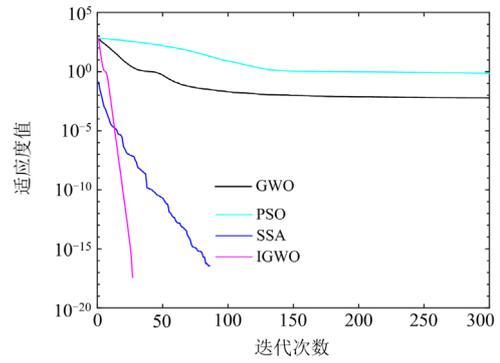
从图 5 的 F_1 、 F_2 收敛曲线可以看出, IGWO 相比于其他优化算法, 未陷入局部最优。从 F_3 、 F_4 收敛曲线可以看出, IGWO 在 50 次迭代内已达到全



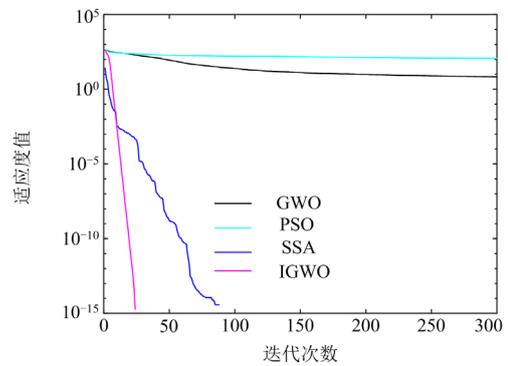
(a) F_1 收敛曲线



(b) F_2 收敛曲线



(c) F_3 收敛曲线



(d) F_4 收敛曲线

图 5 收敛曲线

Fig. 5 Convergence curve

局最优,收敛速度最快。从表2中可以看到,在单峰函数 F_1 、 F_2 的测试中,IGWO的平均最优适应度值以及标准偏差都是最小的,IGWO的性能超过了其他算法。在多峰函数 F_3 、 F_4 的测试中,IGWO的性能超过了GWO、PSO,与SSA的性能相同。

2.5.5 IGWO 优化 DELM 的过程

利用 IGWO 良好的寻优能力,可以在较小的训练误差下找到 ELM-AE 的最优输入层权重,提高预测精度与预测稳定性。IGWO 优化 DELM 的流程图如图 6 所示。

Step1: 设置 DELM 的隐藏层数和节点数以及 IGWO 的种群数和迭代次数。

Step2: 在 ELM-AE 输入层权重寻优范围内随机生成初始权重,IGWO 以均方误差为适应度函数,开始寻优。

Step3: 判断是否达到迭代上限,如果达到,停止寻优,以此时的寻优结果为最佳输入层权重,反之,继续寻优,直到达到迭代上限。

Step4: 将 IGWO 寻优得到的输入层权重代入到 ELM-AE 中。

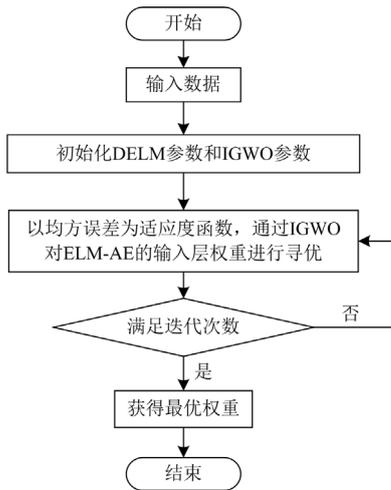


图 6 IGWO 优化 DELM

Fig. 6 DELM optimized by IGWO

2.6 改进 Q 学习算法分配权重

2.6.1 Q 学习(Q-L)算法

Q-L 算法是一种强化学习方法,该算法通过智能体连续试错来获取经验和知识,建立状态-行动关系的 Q 表,并不断优化迭代,获得最佳策略,使期望回报最大。 Q 表的迭代公式为

$$Q^{K+1}(S_K, a_K) = Q^K(S_K, a_K) + \alpha \left[\begin{matrix} R(S_K, S_{K+1}, a_K) + \\ \gamma \max_{a'} Q^K(S_{K+1}, a') - \\ Q^K(S_K, a_K) \end{matrix} \right] \quad (16)$$

式中: $Q^K(S_K, a_K)$ 为在 S_K 状态下执行动作 a_K 的 Q 值; $R(S_K, S_{K+1}, a_K)$ 为状态 S_K 到状态 S_{K+1} 的奖励值; a' 为 S_{K+1} 动作空间的任意一个动作值; α 为学习因子,取 0.01; γ 为折扣因子,取 0.9。

2.6.2 改进 Q 学习(IQ-L)算法

在原始的 Q-L 中,往往采用贪婪机制进行动作选取^[20],缺乏智能性且易陷入局部最优,因此提出一种动作选择机制,如式(17)所示。

$$a_m = \begin{cases} a_{Q_{\max}} & \delta \leq DG/K \\ a_{\text{random}} & \delta > DG/K \end{cases} \quad (17)$$

式中: $a_{Q_{\max}}$ 是当前状态下最大 Q 值对应的动作; a_{random} 是随机选择动作; G 可以控制随机选择动作的次数,在本文中取 6; D 是当前迭代次数; K 取 10 000; δ 为 [0,1] 内的随机数。在前期 $\delta > DG/K$, 因此动作选择是随机的,这是经验的积累过程,在后期 $\delta \leq DG/K$, 动作选择为最大 Q 值所对应的动作,这是利用经验的过程。改进的动作选择机制可以避免通过贪婪机制选择动作而陷入局部最优。

2.6.3 求解组合模型权重

本文采用 IQ-L 算法将 BiLSTM 与 IDELM 的预测结果进行赋权并求和,获得每个子序列的预测结果,其中 $c_1 + c_2 = 1$, c_1 为 BiLSTM 预测结果的权重, c_2 为 IDELM 预测结果的权重,流程如下。

Step1: 建立状态集 S 和动作集 Z 。将权重 c_1 作为环境状态,以 0.0001 为间隔将 [0,1] 其划分为 $\{\Delta f_1, \Delta f_2, \dots, \Delta f_{10001}\}$,共 10 001 个状态作为状态集 S ,以动作集 $Z = [-0.0001, 0.0001]$ 建立行表示状态和列表示动作的 Q 表。

Step2: 建立损失函数 L 与奖惩机制 R 。

$$L = \sqrt{\sum_{t=1}^N [A(t) - \hat{A}(t)]^2} / (N-1) \quad (18)$$

$$R = \begin{cases} +1 + L_m - L_{m+1} & L_{m+1} < L_m \\ -1 + L_m - L_{m+1} & L_{m+1} > L_m \end{cases} \quad (19)$$

式中: $A(t)$ 为负荷真实值; $\hat{A}(t)$ 为预测值; N 为样本点数目; L_m 为第 m 次选择的动作所得到的损失函数值。

Step3: 动作选择机制,如式(17)所示。

Step4: 更新 Q 表,如式(16)所示。

Step5: 判断 Q 表是否收敛,如收敛并且在以后的 1000 次学习中都收敛,则输出权重,否则转到 Step3。

采用 IQ-L 算法对组合预测模型进行权重分配,其流程图如图 7 所示。

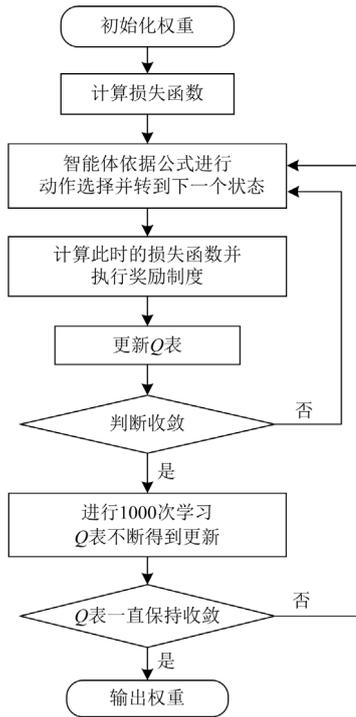


图7 IQ-L 算法分配权重

Fig. 7 Weights assigned by using IQ-L algorithm

3 仿真实例

3.1 数据来源

本文所采用的数据来源于河南某县的电力负荷数据, 3 个数据集时间跨度分别为 2019 年 1 月 1 日—1 月 15 日、5 月 1 日—5 月 15 日和 9 月 1 日—9 月 15 日, 分辨率为 15 min, 每天 96 个时刻的负荷值, 以每个数据集前 1344 个时刻的负荷数据为训练集, 后 96 个时刻的负荷数据为测试集。1 月份的负荷序列及其分解结果如图 8 所示, 3.3 节的前 4 个实验对比都以 1 月份的数据作为数据集。

由图 8 可得, 负荷序列经过 VMD 后, 各个子序列相对于原始负荷序列的复杂度降低, 其中 IMF1 反映了负荷波动的整体趋势, IMF2、IMF3 反映了负荷序列的局部特征。

3.2 模型参数选取

为了保证实验的公平性和合理性, VMD 的 K 取不同值时, 各个模态的中心频率如表 3 所示。双向长短期记忆网络隐藏层的神经元数量为 50, 训练次数为 200, IDELM 隐藏层的层数以及神经元节点数对实验结果影响很大, 过多会导致预测时间变长, 过少则不能充分提取输入数据的特征, 影响预测精度, 因此隐藏层的层数以及每层神经元节点数的设置要在一定的范围内, IDELM 的层数与节点数实验如表 4 所示。

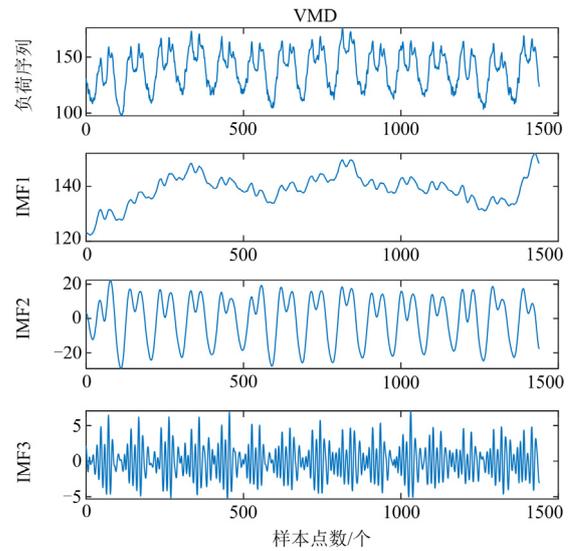


图8 负荷数据分解

Fig. 8 Decomposition of load data

表3 不同 K 值下各模态的中心频率

Table 3 Central frequencies of each mode under different K values

模态	$K=3$	$K=4$	$K=5$	$K=6$
1	2.5342×10^{-6}	1.07×10^{-6}	1.06×10^{-6}	1.05×10^{-6}
2	0.0140	0.0105	0.0105	0.0106
3	0.0772	0.0224	0.0224	0.0223
4	—	0.0824	0.0803	0.0802
5	—	—	0.1505	0.1500
6	—	—	—	0.4553

表4 深度极限学习机参数设置实验

Table 4 Experiment on parameter setting of DELM

隐藏层数	各层神经元节点数			平均绝对百分比误差/%	运行时间/s
	1	2	3		
2	10	20	—	1.8634	0.4728
	10	30	—	1.7466	0.6166
	20	30	—	1.6250	0.5468
	30	20	—	1.6109	0.4273
3	20	30	40	2.6834	0.7057

由表 3 可得, 当 $K=4$ 时, 模态 2 和模态 3 的中心频率接近, 此时出现了模态混叠现象, 因此, VMD 的 $K=3$ 较为合适, 惩罚系数取 1000。由表 4 可得, 增加隐藏层数并没有提高预测的准确性, 因此深度极限学习机的层数设置为 2, 第一层的节点数设置为 30, 第二层的节点数设置为 20, 采用 sigmoid 激活函数。

3.3 仿真及结果分析

3.3.1 预测模型对比

为了验证本文所提模型的有效性, 对 IDELM、

BiLSTM、BiLSTM-IQ-IDELM 与本文所提的 VMD-BiLSTM-IQ-IDELM 模型进行对比仿真实验，仿真实验预测曲线如图 9 所示，反映误差分布的箱线图如图 10 所示。负荷数据在采集过程中会因为采集设备故障以及一些其他因素导致异常，为了验证所提模型的鲁棒性，选取在一定比例归一化后的负荷数据中加入标准差为 0.2、噪声数据比例分别为 5%、10%、20% 的白噪声^[21]，预测结果如表 5 所示。实验指标采用均方根误差 X_{RMSE} 、平均绝对百分比误差 X_{MAPE} 以及平均绝对误差 X_{MAE} ^[22]对预测结果进行评价，未加入噪声情况下的负荷预测结果如表 6 所示。

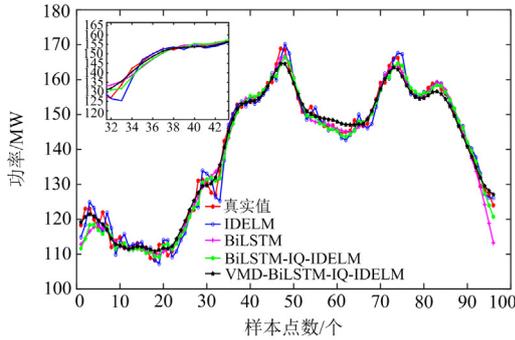


图 9 不同模型预测曲线对比图

Fig. 9 Comparison of prediction curves of different models

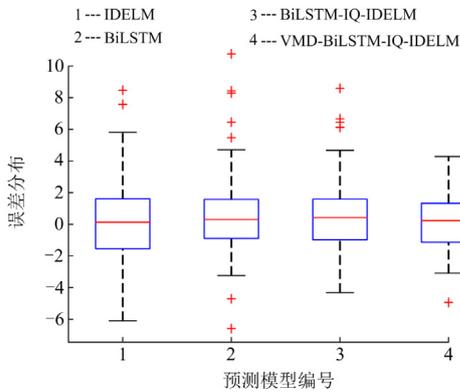


图 10 不同预测模型误差对比

Fig. 10 Errors comparison of different prediction models

表 5 不同噪声比例下的预测结果

Table 5 Forecasting results at different noise ratios

噪声比例/%	X_{RMSE}/MW	X_{MAE}/MW	$X_{MAPE}/\%$
0	1.7404	1.4115	1.0327
5	1.7453	1.4194	1.0386
10	1.7540	1.4284	1.0463
20	2.5057	1.9876	1.4063

由表 5 可以看出，在加入少量噪声后，模型的预测结果变化不大，在加入 20% 的噪声以后，模型

的预测准确性有所下降，然而还是超过了未加入噪声情况下的单一模型的预测结果，验证了所提模型具有较好的鲁棒性。

表 6 未加入噪声情况下的负荷预测结果指标

Table 6 Load forecasting result indicator without noise

模型	X_{RMSE}/MW	X_{MAE}/MW	$X_{MAPE}/\%$
IDELM	2.6832	2.0801	1.5430
BiLSTM	2.6066	1.7726	1.3305
BiLSTM-IQ-IDELM	2.3565	1.7686	1.3086
VMD-BiLSTM-IQ-IDELM	1.7404	1.4115	1.0327

由表 6 可以看出，单一模型预测效果较差，采用组合模型以后，BiLSTM-IQ-IDELM 模型相比于 IDELM、BiLSTM 单一预测模型， X_{RMSE} 分别下降了 12.17%、9.59%， X_{MAE} 分别下降 14.97%、0.22%， X_{MAPE} 分别下降 15.19%、1.64%。说明采用组合模型能够提高预测的精度，采用 VMD 以后再预测，相比于 BiLSTM-IQ-IDELM 模型， X_{RMSE} 下降 26.14%， X_{MAE} 下降 20.19%， X_{MAPE} 下降 21.08%。

3.3.2 权重分配方法对比

为了验证 Q-L 改进后的性能，将 Q-L、误差倒数法、等权重法、IGWO、IQ-L 几种权重分配方法对 BiLSTM 和 IDELM 的预测结果进行权重分配，得到最终预测结果，实验结果如图 11 所示，实验评价指标同上，结果如表 7 所示。

由表 7 可以看出，采用的权重分配方法不同，所得的预测结果也不相同，因为更好的权重分配方法能将两种基础模型的优点相结合，使预测结果更加精确。其中 IQ-L 权重分配方法表现最好，原因是避免了 Q-L 中采用贪婪机制选取动作而导致的局部最优，也避免了误差倒数法分配权重时对异常值敏感度过高。采用 IGWO 与 IQ-L 方法相比，IGWO 属于启发式算法，IQ-L 方法属于强化学习算法。IGWO 这种启发式算法，它的搜索过程是一个盲目探索的过程，

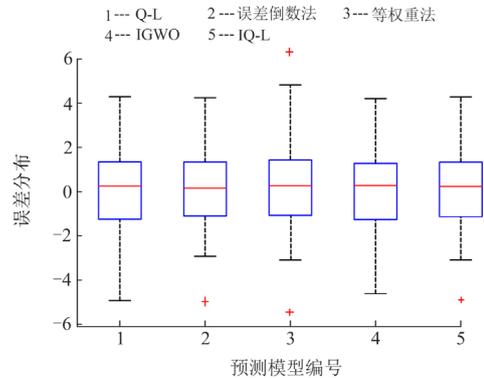


图 11 不同权重分配方法误差对比

Fig. 11 Errors comparison of different weight distribution methods

表 7 不同权重分配方法的负荷预测结果指标
Table 7 Load forecasting results metrics of different weight allocation methods

方法	X_{RMSE}/MW	X_{MAE}/MW	$X_{MAPE}/\%$
等权重	2.0155	1.5489	1.1469
误差倒数法	1.7606	1.4110	1.0328
Q-L	1.7506	1.4243	1.0429
IGWO	1.7411	1.4105	1.0331
IQ-L	1.7404	1.4115	1.0327

且适应度函数的设计影响寻优效果, 而 IQ-L 的搜索过程属于马尔可夫决策过程, 在两个基础模型预测结果权值分配的问题上, 采用马尔可夫决策过程的 IQ-L 方法显然比盲目探索的启发式算法更加高效。

IGWO 分配权重的实验中, 适应度函数为最小均方根误差, 种群数量为 20, 迭代次数为 50。

3.3.3 模型搭配合理性对比

为了验证本文所提模型中 BiLSTM 与 IDELM 是较为合适的搭配, 与 BiLSTM 与 IGWO 优化的支持向量回归(support vector regression, SVR)组合模型、BiLSTM 与 IGWO 优化的 BP 神经网络组合模型进行对比, IGWO 优化 SVR 的惩罚参数和核参数, 寻优范围分别为 [0.1,30] 和 [0.1,10], IGWO 优化 BP 神经网络的权值和偏置, 寻优范围均为 [-2,2], IGWO 优化 ELM-AE 的输入层权重, 寻优范围为 [-2,2], 种群数设置为 20, 迭代次数为 50。实验结果如图 12 所示, 实验评价指标同上, 结果如表 8 所示。

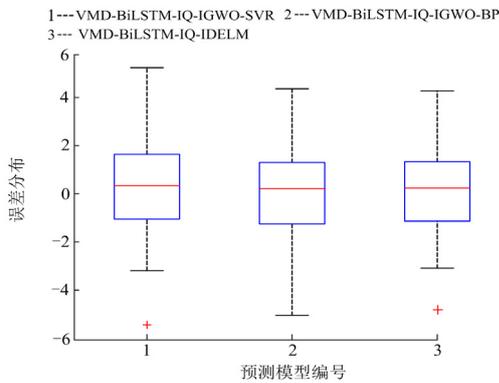


图 12 搭配合理性对比

Fig. 12 Comparison of collocation rationality

表 8 不同搭配模型的负荷预测结果指标
Table 8 Load forecasting results metrics of different collocation models

模型	X_{RMSE}/MW	X_{MAE}/MW	$X_{MAPE}/\%$
VMD-BiLSTM-IQ-IGWO-SVR	1.9885	1.6218	1.1758
VMD-BiLSTM-IQ-IGWO-BP	1.7577	1.4304	1.0463
VMD-BiLSTM-IQ-IDELM	1.7404	1.4115	1.0327

由表 8 可以看出, 模型搭配不同, 预测结果也不一样, 本文所提的 VMD-BiLSTM-IQ-IDELM 模型相比于 VMD-BiLSTM-IQ-IGWO-SVR 模型, X_{RMSE} 下降 12.47%, X_{MAE} 下降 12.96%, X_{MAPE} 下降 12.17%, 相比于 VMD-BiLSTM-IQ-IGWO-BP 模型, X_{RMSE} 下降 0.98%, X_{MAE} 下降 1.32%, X_{MAPE} 下降 1.29%, 因此 BiLSTM 与 IDELM 搭配较为合理。

3.3.4 优化算法对比

为了验证改进灰狼算法相比其他智能优化算法的寻优能力更强, 采用 VMD-BiLSTM-IQ-GWO-DELM、VMD-BiLSTM-IQ-PSO-DELM、VMD-BiLSTM-IQ-SSA-DELM 与 VMD-BiLSTM-IQ-IDELM 模型进行对比实验, 优化算法参数设置: 种群规模为 20, 迭代次数为 50。仿真实验预测曲线对比如图 13 所示, 箱线图如图 14 所示, 实验评价指标同上, 结果如表 9 所示。

由表 9 可以看出, 因为不同的优化算法对 ELM-AE 输入层权重寻优所得的值不同, 从而影响了模型的预测效果。相对于其他优化算法, IGWO 对 ELM-AE 输入层权重寻优效果更好。

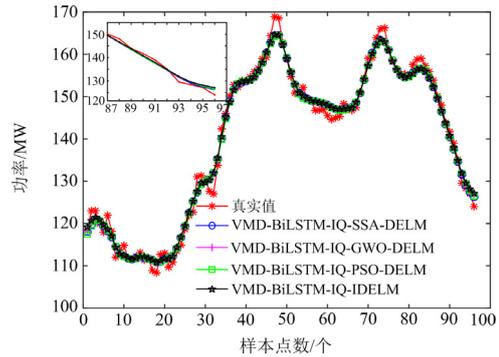


图 13 不同优化算法的预测曲线对比图

Fig. 13 Comparison of prediction curves of different optimization algorithms

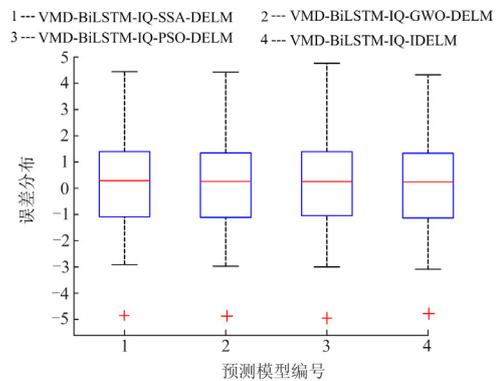


图 14 优化算法误差对比

Fig. 14 Errors comparison of optimization algorithms

表 9 不同优化算法的预测结果指标
Table 9 Prediction results metrics of different optimization algorithms

模型	X_{RMSE}/MW	X_{MAE}/MW	$X_{MAPE}/\%$
VMD-BiLSTM-IQ-PSO-DELM	1.8342	1.4736	1.0800
VMD-BiLSTM-IQ-GWO-DELM	1.7806	1.4287	1.0466
VMD-BiLSTM-IQ-SSA-DELM	1.7802	1.4241	1.0410
VMD-BiLSTM-IQ-IDELM	1.7404	1.4115	1.0327

3.3.5 不同月份对比

为了验证所提预测模型的有效性, 选取一年中的 3 个典型月份(1 月、5 月、9 月)进行预测, 各个月份负荷序列模态分解的 K 值采用中心频率法确定, 分别是 3、3、4。箱线图如图 15 所示, 实验评价指标同上, 结果如表 10 所示。

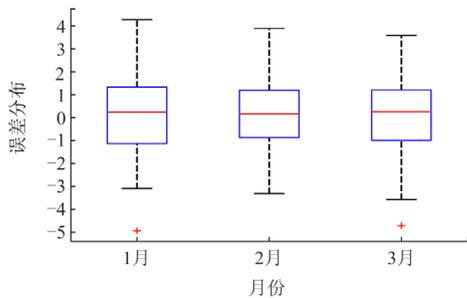


图 15 不同月份对比

Fig. 15 Comparison of different months

表 10 不同月份的负荷预测结果指标

Table 10 Load forecasting results metrics for different months

月份	X_{RMSE}/MW	X_{MAE}/MW	$X_{MAPE}/\%$
1	1.7404	1.4115	1.0327
5	1.4911	1.2069	0.9150
9	1.6085	1.2905	0.9220

由表 10 可得, 不同月份因为用电情况不同, 负荷的波动情况也不一样, 但各个月份预测结果的准确率都在 98%以上, 验证了所提模型的有效性。

4 结论

本文针对于超短期负荷预测难度大且准确度低的问题, 提出了基于 IQ-L 算法的组合预测模型, 结论如下:

- 1) 本文对 Q-L 算法的动作选择机制进行改进, 提高了算法的智能性, 保证了结果的预测精度;
- 2) 对 GWO 的收敛因子和位置更新方程进行了改进, 在对预测模型参数的优化中相比于其他优化算法, 寻优效果更好, 提高了模型预测精度与预测稳定性;

3) 对本文所提的模型与其他类似模型进行权重分配方法、搭配合理性、群智能优化算法以及不同月份对比实验, 证明了 IQ-L 算法分配权重的有效性、BiLSTM 与 IDELM 搭配的合理性、灰狼算法改进的有效性以及所提模型的普适性。

参考文献

- [1] 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.
- [2] 潘超, 鲁宝春, 孙丽颖, 等. 电力系统负荷预测方法综述[J]. 辽宁工业大学学报(自然科学版), 2016, 36(5): 296-299.
PAN Chao, LU Baochun, SUN Liying, et al. Review of load forecasting methods in power system[J]. Journal of Liaoning University of Technology (Natural Science Edition), 2016, 36(5): 296-299.
- [3] 朱继忠, 董瀚江, 李盛林, 等. 数据驱动的综合能源系统负荷预测综述[J]. 中国电机工程学报, 2021, 41(23): 7905-7924.
ZHU Jizhong, DONG Hanjiang, LI Shenglin, et al. Review of data-driven load forecasting for integrated energy system[J]. Proceedings of the CSEE, 2021, 41(23): 7905-7924.
- [4] 胡威, 张新燕, 李振恩, 等. 基于优化的 VMD-mRMR-LSTM 模型的短期负荷预测[J]. 电力系统保护与控制, 2022, 50(1): 88-97.
HU Wei, ZHANG Xinyan, LI Zhen'en, et al. Short-term load forecasting based on an optimized VMD-mRMR-LSTM model[J]. Power System Protection and Control, 2022, 50(1): 88-97.
- [5] 王海柱, 田馥铭, 张鹏, 等. 基于 CEEMD-FE 和 AOA-LSSVM 的短期电力负荷预测[J]. 电力系统保护与控制, 2022, 50(13): 126-133.
YANG Haizhu, TIAN Fuming, ZHANG Peng, et al. Short-term load forecasting based on CEEMD-FE-AOA-LSSVM[J]. Power System Protection and Control, 2022, 50(13): 126-133.
- [6] 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.
- [8] 任建吉, 位慧慧, 邹卓霖, 等. 基于 CNN-BiLSTM-Attention 的超短期电力负荷预测[J]. *电力系统保护与控制*, 2022, 50(8): 108-116.
REN Jianji, WEI Huihui, ZOU Zhuolin, et al. Ultra-short-term power load forecasting based on CNN-BiLSTM-Attention[J]. *Power System Protection and Control*, 2022, 50(8): 108-116.
- [9] YU M, NIU D, WANG K, et al. Short-term photovoltaic power point-interval forecasting based on double-layer decomposition and WOA-BiLSTM-Attention and considering weather classification[J]. *Energy*, 2023, 275: 127348.
- [10] 刘杰, 金勇杰, 田明. 基于 VMD 和 TCN 的多尺度短期电力负荷预测[J]. *电子科技大学学报*, 2022, 51(4): 550-557.
LIU Jie, JIN Yongjie, TIAN Ming. Multi-scale short-term load forecasting based on VMD and TCN[J]. *Journal of University of Electronic Science and Technology of China*, 2022, 51(4): 550-557.
- [11] 陈振宇, 刘金波, 李晨, 等. 基于 LSTM 与 XGBoost 组合模型的超短期电力负荷预测[J]. *电网技术*, 2020, 44(2): 614-620.
CHEN Zhenyu, LIU Jinbo, LI Chen, et al. Ultra short-term power load forecasting based on combined LSTM-XGBoost model[J]. *Power System Technology*, 2020, 44(2): 614-620.
- [12] 孟安波, 许炫淙, 陈嘉铭, 等. 基于强化学习和组合式深度学习模型的超短期光伏功率预测[J]. *电网技术*, 2021, 45(12): 4721-4728.
MENG Anbo, XU Xuancong, CHEN Jiaming, et al. Ultra short term photovoltaic power prediction based on reinforcement learning and combined deep learning model[J]. *Power System Technology*, 2021, 45(12): 4721-4728.
- [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] 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 Shuailong, ZHANG Li. Ultra short term power prediction of photovoltaic power generation based on VMD-LSTM and error compensation[J]. *Acta Energetica Sinica*, 2022, 43(8): 96-103.
- [15] 曾亮, 雷舒敏, 王珊珊, 等. 基于 OVMD-SSA-DELM-GM 模型的超短期风电功率预测方法[J]. *电网技术*, 2021, 45(12): 4701-4712.
ZENG Liang, LEI Shumin, WANG Shanshan, et al. Ultra-short-term wind power prediction based on OVMD-SSA-DELM-GM model[J]. *Power System Technology*, 2021, 45(12): 4701-4712.
- [16] ZHANG Q, LIU H, GUO J, et al. Improved GWO-MCSVM algorithm based on nonlinear convergence factor and tent chaotic mapping and its application in transformer condition assessment[J]. *Electric Power Systems Research*, 2023, 224: 109754.
- [17] 李云溟, 咸日常, 张海强, 等. 基于改进灰狼算法与最小二乘支持向量机耦合的电力变压器故障诊断方法[J]. *电网技术*, 2023, 47(4): 1470-1478.
LI Yunhao, XIAN Richang, ZHANG Haiqiang, et al. Fault diagnosis for power transformers based on improved grey wolf algorithm coupled with least squares support vector machine[J]. *Power System Technology*, 2023, 47(4): 1470-1478.
- [18] 龙文, 蔡绍洪, 焦建军, 等. 一种改进的灰狼优化算法[J]. *电子学报*, 2019, 47(1): 169-175.
LONG Wen, CAI Shaohong, JIAO Jianjun, et al. An improved grey wolf optimization algorithm[J]. *Acta Electronica Sinica*, 2019, 47(1): 169-175.
- [19] XUE J, SHEN B. Dung beetle optimizer: a new meta-heuristic algorithm for global optimization[J]. *The Journal of Supercomputing*, 2023, 79(7): 7305-7336.
- [20] 杨悦, 王丹, 胡博, 等. 基于改进多智能体 Q 学习的多源最优联合调频控制策略研究[J]. *电力系统保护与控制*, 2022, 50(7): 135-144.
YANG Yue, WANG Dan, HU Bo, et al. Multi-source optimal joint frequency modulation control strategy based on improved multi-agent Q-learning[J]. *Power System Protection and Control*, 2022, 50(7): 135-144.
- [21] 许言路, 武志凯, 朱赫炎, 等. 基于多尺度卷积神经网络的短期电力负荷预测[J]. *沈阳工业大学学报*, 2020, 42(6): 618-623.
XU Yanlu, WU Zhikai, ZHU Heyan, et al. Short-term load forecasting based on multi-scale convolutional neural network[J]. *Journal of Shenyang University of Technology*, 2020, 42(6): 618-623.
- [22] AN G, CHEN L, TAN J, et al. Ultra-short-term wind power prediction based on PVMD-ESMA-DELM[J]. *Energy Reports*, 2022, 8: 8574-8588.

收稿日期: 2023-10-20; 修回日期: 2024-02-25

作者简介:

张丽(1982—), 女, 博士, 副教授, 硕士生导师, 主要研究方向为电网需求侧响应、智能用电技术; Email: dqzhangli@hpu.edu.cn

李世情(1999—), 男, 通信作者, 硕士研究生, 主要研究方向为电力系统及其自动化。Email: 1014611023@qq.com

(编辑 姜新丽)