

王 俊, 王继焯, 程 坤, 等. 基于双层优化 VMD-LSTM 的农村超短期电力负荷预测[J]. 沈阳农业大学学报, 2024, 55(1):92-102.

WANG Jun, WANG Jiye, CHENG Kun, et al. Ultra-short-term power load forecasting based on two-layer optimization VMD-LSTM[J]. Journal of Shenyang Agricultural University, 2024, 55(1):92-102.

## 基于双层优化 VMD-LSTM 的农村超短期 电力负荷预测

王 俊<sup>1</sup>, 王继焯<sup>1</sup>, 程 坤<sup>1</sup>, 方 均<sup>1</sup>, 鞠丹阳<sup>2</sup>

(1. 沈阳农业大学 信息与电气工程学院, 沈阳 110161;

2. 国网内蒙古东部电力有限公司 呼伦贝尔供电公司, 内蒙古 呼伦贝尔 162650)

**摘 要:** 稳定的供电是农村发展建设的有力保障, 而电力负荷水平是建设效果的重要衡量标准, 因此建立精确的负荷预测模型可以更准确直观显现电力负荷情况, 为供电公司制定决策提供有力支撑。由于 LSTM 负荷预测模型在数据预测方面存在收敛性差、预测精度不高等问题, 为提高模型的预测精度, 提出一种基于双层优化 VMD-LSTM 的超短期电力负荷预测方法。首先提出麻雀算法优化变分模态分解 (sparrow variational mode decomposition, SVMD), 通过 SVMD 将原始数据转化为模态分量 (intrinsic mode functions, IMF); 其次采用改进樽海鞘群算法 (association salp swarm algorithm, ASSSA) 优化 LSTM 模型。通过引入 4 种策略增强标准樽海鞘算法优化能力; 最后将各模态分量分别代入到新模型并进行叠加预测。选取辽宁省某市某乡村 10kV 变压器真实历史负荷数据, 以均方根误差 (RMSE)、平均绝对误差 (MAE)、平均绝对百分比误差 (MAPE)、拟合度 ( $R^2$ ) 作为评价指标, 并与其他基础预测模型进行对比, 结果表明, 改进后的算法在计算精度、稳定性方面均优于其他基础预测模型。

**关键词:** 长短期预测; 双层优化; 樽海鞘群算法; 变分模态分解; 叠加预测

中图分类号: TM714

文章编号: 1000-1700(2024)01-0092-11

文献标识码: A

开放科学 (资源服务) 标识码 (OSID):



## Ultra-short-term Power Load Forecasting Based on Two-layer Optimization VMD-LSTM

WANG Jun<sup>1</sup>, WANG Jiye<sup>1</sup>, CHENG Kun<sup>1</sup>, FANG Jun<sup>1</sup>, JU Danyang<sup>2</sup>

(1. College of Information and Electrical Engineering, Shenyang Agricultural University, Shenyang 110161, China;

2. State Grid Inner Mongolia Eastern Electric Power Co., Ltd, Hulunbuir Power Supply Company, Hulunbuir Inner Mongolia 162650, China)

**Abstract:** Stable power supply is a guarantee for rural development and construction, and the level of power load is an important measure of the construction effect. Therefore, establishing a precise load prediction model can more accurately and intuitively show the power load, and provide a strong support for the formulation of decision-making for power supply companies. Since the LSTM load forecasting model has problems such as poor convergence and low forecasting accuracy in data forecasting, in order to improve the forecasting accuracy of the model, an ultra-short-term power load forecasting method based on two-layer optimization VMD-LSTM is proposed. First, the sparrow algorithm is proposed to optimize the variational mode decomposition (sparrow variational mode decomposition, SVMD), and the original data is converted into modal components (intrinsic mode functions, IMF) through SVMD; secondly, the improved salp swarm algorithm (association

收稿日期: 2023-10-10

基金项目: 国家电网公司科技项目 (SGTYHT/23-JS-001); 国家自然科学基金项目 (61903264)

第一作者: 王 俊 (1979-), 男, 博士, 教授, 博士生导师, 从事智能配电网、可再生能源电力并网运行与控制研究, E-mail: wddream2016@syau.edu.cn

salp swarm algorithm, ASSSA) to optimize the LSTM model. The optimization ability of the standard salp algorithm is enhanced by introducing four strategies; finally, each modal component is substituted into the new model and superimposed prediction is performed. The real historical load data of 10 kV transformer in a certain city and village in Liaoning Province is chosen, the root mean square error (RMSE), average absolute error (MAE), average absolute percentage error (MAPE), and fitting degree ( $R^2$ ) are taken as evaluation indicators, and compared with other basic prediction models. The results show that the improved algorithm is superior to other basic prediction models in terms of calculation accuracy and stability.

**Key words:** long-term prediction; double-layer optimization; salp swarm algorithm; variational mode decomposition; superposition forecast

农村电网建设是国家农村发展建设的重要一环,对乡村振兴起着重要的作用,电力负荷水平是其中一项重要指标,预测的精确度直接影响到对乡村地区电力负荷水平的判断,进而对之后的供电计划产生深远影响。超短期负荷预测一般用于在线监测、事故预防及紧急事故处理等,精准的负荷预测对电网的稳定运行及安全经济起着巨大的作用,电力负荷预测的准确性受到多个影响因素制约,得到精确预测结果相对困难,因此提升模型的预测准确性显得尤为重要。负荷预测主要包括传统预测和机器学习。传统预测利用智能电表实时测量数据<sup>[1]</sup>进行建模,机器学习包括人工神经网络、支持向量机等,其中长短期记忆神经网络LSTM为代表的深度学习被广泛运用。由于单一模型已经难以满足精度需要,大量学术文献把研究方向集中在组合模型。LightGBM<sup>[2]</sup>、随机森林<sup>[3]</sup>、误差修正<sup>[4]</sup>、XGBoost<sup>[5]</sup>、word2vec<sup>[6]</sup>进行结合提高模型预测精度,为进一步完善预测模型,引入了卷积和注意力机制的概念<sup>[7-9]</sup>。随着智能算法的成熟,越来越多学者通过模拟各生物的生活习性而创造出大量智能优化算法并加入到负荷预测领域,如粒子群算法<sup>[10]</sup>、鲸鱼算法<sup>[11]</sup>,通过计算机大量训练数据代替人工,避免了因经验造成的模型参数非最优从而导致预测结果不够精确。

除去预测模型本身优化,数据集处理方面也很重要,常用方法有经验模态分解、变分模态分解、小波变换等。经验模态分解通过与贝叶斯<sup>[12]</sup>、PCA主成分分析<sup>[13]</sup>、TCN时域网络<sup>[14]</sup>、ARIMA<sup>[15]</sup>进行组合预测,将数据集分解为模态分量而不需要提前分析研究,在此基础上发展出的聚类经验模态分解<sup>[16]</sup>、CEEMDAN<sup>[17]</sup>、EEMD<sup>[18]</sup>可以在一定程度上进行优化,但由于经验模态分解本身存在模态混叠和端点效应,这种改进所取得效果比较有限。而变分模态分解可以有效避免这两种问题,经验模态分解逐渐被取代为变分模态分解<sup>[19-24]</sup>,变分模态将数据分解为模态分量通过VMD将数据分解后进行重组,进一步提高了模型的精确度,但这两者都没有考虑到算法自身参数的最优解问题,因此利用智能算法对VMD关键参数各模态初始中心约束强度( $\alpha$ )和模态个数( $k$ )进行智能寻优<sup>[25-28]</sup>,避免人为设置参数导致预测结果精度不够的问题。将处理数据集、智能算法与预测模型进行结合<sup>[29-31]</sup>,充分利用各部分特点最大限度提高模型的训练效果与精确度。上述文献充分说明了智能算法优化VMD及组合算法在电力负荷预测中的可行性。

在分析研究以上成果后,提出一种双层优化组合模型,一层麻雀算法优化VMD,二层改进樽海鞘群算法优化LSTM,数据集经第一层分解为模态分量后分别带入到二层模型训练,再进行叠加预测。通过实验仿真,将最后的预测结果与基础模型进行比较,并利用4种评价指标验证其性能以及在农村电力负荷预测方面的可行性。

## 1 麻雀算法优化变分模态分解

### 1.1 标准变分模态分解(VMD)

2014年,DRAGOMIRETSKIY提出了变分模态分解(variational mode decomposition, VMD),该方法通过分量窄带建立约束优化问题估计各个信号分量,VMD是一种比较完善、技术比较成熟的信号处理技术,能够有效地克服传统EMD、CEEMD等算法在处理时存在的模式混淆、终端影响等问题,能够得到更加稳定的负荷数据。利用此技术,能够尽量降低大多数噪声,提高信号的平滑性,通过分解得到若干组模态分量。设定待处理信号原始负荷时序 $f(t)$ ,运用VMD进行信号分解的约束表达式为:

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

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

式中分析了各模态成分的分布及中心频率。其中,  $k$  为模态分解的次数,  $*$  为一个卷积运算符采用拉格朗日乘积算子, 求解带约束的变分问题。相反的情况下, 包络熵变得很低, 公式变为:

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

式中:  $\lambda$  为 Lagrange 算子;  $\alpha$  为二次惩罚因子。运用交替方向乘子 ADMM 寻优迭代后得到模态分量和中心频率。交替寻优后的  $u_k, \omega_k, \lambda$  表达式为:

$$u_k^{n+1}(\omega) = \frac{f(\omega) - \sum_{i \neq k} u_i(\omega) + \lambda(\omega) / 2}{1 + 2\alpha(\omega - \omega_k)^2} \quad (4)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |u_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |u_k^{n+1}(\omega)|^2 d\omega} \quad (5)$$

$$\lambda^{n+1}(\omega) = \lambda^n(\omega) + \gamma \left( f(\omega) - \sum_k u_k^{n+1}(\omega) \right) \quad (6)$$

式中:  $\gamma$  为噪声容忍度,  $u_k^{n+1}(\omega), u_i(\omega), f(\omega)$  和  $\lambda(\omega)$  进行傅里叶变换后可得到  $u_k^{n+1}(t), u_i(t), f(t), \lambda(t)$ 。

### 1.2 麻雀算法优化变分模态原理

VMD 和 EMD 由于理论框架不同, 二者存在诸多差异, VMD 有 5 个参数, alpha 所代表的惩罚系数和 k 所代表的分解个数对最终结果影响最大, 采用麻雀算法同时搜寻 alpha 和 k。避免主观因素对参数选取的影响, 以包络熵极小值作为适应度函数, 当 IMF 中噪声较多, 特征信息较少时, 包络熵值较大, 反之, 则包络熵值较小。

信号  $x(i)=i=1, 2, \dots, N$  的包络熵依据公式为:

$$\left. \begin{aligned} E_p &= -\sum_{i=1}^N p_i \lg p_i \\ p_i &= a(i) / \sum_{i=1}^N a(i) \end{aligned} \right\} \quad (7)$$

式中:  $a(i)$  为 VMD 分解后的 k 个模态成分经过 Hilbert 解调所得的包络熵, 其中, 对  $a(i)$  进行了标准化处理,  $N(i)$  是对  $a(i)$  进行了统计处理后获得的一个概率分布,  $N(i)$  是抽样点数目, 对这些概率分布的一个序列进行了统计处理, 将其作为一个包络熵来求取。

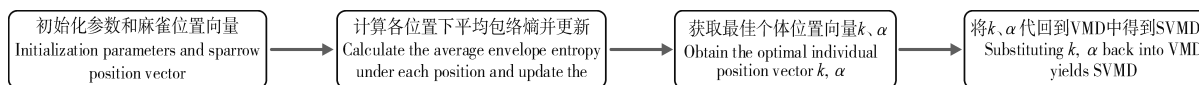


图1 麻雀算法优化VMD流程

Figure 1 Sparrow algorithm to optimize the VMD process

## 2 改进樽海鞘群算法 (ASSSA) 优化 LSTM 模型

### 2.1 基本算法原理

樽海鞘群算法 (salp swarm algorithm, SSA) 是 SEYEDALI 提出的一种智能优化算法。该算法模拟了樽海鞘链的群体行为, 初始化樽海鞘的值  $x_j^i$ 。

$$x_j^i = rand(N, D) \times (ub_j - lb_j) + lb_j \quad (8)$$

式中:  $N$  为樽海鞘种群的大小;  $D$  为空间维度;  $lb_j$  为每一维搜索空间下限;  $ub_j$  为每一维搜索空间上限;  $i \in [1, 2, \dots, N]$ ;  $j \in [1, 2, \dots, N]$ 。

对每个樽海鞘个体按照目标函数计算适应度值并排序,取最小值为全局最优位置,即为食物源位置。

领导者的位置更新为:

$$X_j^i(k) = \begin{cases} F_j + c_1 [c_2(ub_j - lb_j) + lb_j] c_3 \geq 0.5 \\ F_j + c_1 [c_2(ub_j - lb_j) + lb_j] c_3 < 0.5 \end{cases} \quad (9)$$

式中:  $k$  为当前迭代次数;  $X_j^i$  为第 1 个樽海鞘(领导者)在  $j$  维空间的位置。

$F_j$  为食物源在第  $j$  维空间的位置,它是整个樽海鞘群搜索的目标,  $c_2$  和  $c_3$  为  $[0, 1]$  之间的随机数,  $c_2$  决定在第  $j$  维空间移动的步长,  $c_3$  确定了移动的正负方向,  $c_1$  随迭代次数的增加而非线性减小,  $c_1$  是最重要的参数,定义如下:

$$c_1 = 2e^{-\left(\frac{4t}{T_{\max}}\right)^2} \quad (10)$$

式中:  $t$  为当前迭代次数;  $T_{\max}$  为种群最大迭代次数。

(1) 樽海鞘追随者的位置更新为:

$$x_j^i(k) = \frac{1}{2} [x_j^i(k-1) + x_j^{i-1}(k-1)] \quad (11)$$

式中:  $i \geq 2$ ,  $X_j^i$  为第  $i$  只樽海鞘在第  $j$  维的位置。

(2) 计算更新后每只樽海鞘每一维适应度值,若适应度值小于食物源位置的适应度值,则将此位置替换食物源位置。

(3) 判断是否满足最大迭代数,若是满足,则输出结果,否则转为步骤 3 继续迭代,直到满足条件为止。

## 2.2 改进的樽海鞘群算法

2.2.1 自适应螺旋搜索策略 在 SSA 的领导者种群更新中, SSA 在最优个体周围进行随机位置更新,受金枪鱼群算法螺旋觅食策略的启发,将该螺旋觅食策略引入到 SSA 领导者位置更新中。SSA 从原策略和螺旋觅食策略中随机选择一种策略对个体位置进行更新,避免其他个体陷入局部最优。螺旋觅食策略具体为:

$$\alpha_1 = a + (1 - a) \cdot \frac{t}{t_{\max}} \quad (12)$$

$$\beta = e^{bl} \cdot \cos(2\pi b) \quad (13)$$

$$l = e^{3\cos\left[\left(\frac{t_{\max} + 1}{t} - 1\right)\pi\right]} \quad (14)$$

式中:  $a$  为一个常数,取值为 0.7;  $b$  为一个 0 到 1 均匀分布的随机数;  $X_{br}^i$  为最优个体或者在搜索空间中随机生成的一个个体。全局探索阶段,螺旋搜索策略需要搜索更广泛的空间,因此  $X_{br}^i$  为搜索空间中随机生成的一个个体;在后期的局部开发阶段,更多围绕最优个体进行螺旋搜索,因此  $X_{br}^i$  取最优个体的位置信息。个体在迭代时,会依照概率随机选择使用原始搜索策略或者螺旋搜索策略。

2.2.2 自适应种群分布策略 自适应种群分布策略公式为:

$$Pop_{leader} = c \times \left[ \sin\left(\frac{-\pi t}{2t_{\max}} + 0.5\pi\right) - 0.2 * rand \right] + d \quad (15)$$

$$Pop_{follower} = N - Pop_{leader} \quad (16)$$

式中:  $c$  为控制领导者和跟随者数量的比例系数。分析公式可知,  $Pop_{leader}$  的值随迭代数的增大,领先者数目不断减小,从随者数目不断增多。更多的“樽海鞘”跟随者则深入到了最好的位置,同时为了保证算法后期仍有部分个体进行全局探索,因此,引入参数  $d$ ,  $d$  为确保  $Pop_{leader}$  不为 0 的随机数。

2.2.3 正弦惯性权重递减策略 追随者位置更新与当前个体和其前一个体位置有重大关系。但是,由于盲目跟随的行为,只能单方面接受信息,因此,搜索效果会受到一定程度限制。

在此基础上增加了正弦惯性权重递减策略,具体公式为:

$$\omega = \omega_{\max} - \frac{(\omega_{\max} - \omega_{\min})}{1 + 0.5e^{8 - 0.06t}} \tag{17}$$

式中: $\omega_{\max}$  为最大惯性权重; $\omega_{\min}$  为最小惯性权重。

2.2.4 双重反向学习策略 该方法可有效地避免陷入局部极值问题。具体公式为:

$$X_i^0 = lb - ub - X_i^l \tag{18}$$

式中: $X_i^0$  是  $X_i^l$  对应的反向解。

为了进一步增强种群多样性,解决传统反向学习方法产生的解不一定是最优解问题,利用Tent混沌映射的随机性、遍历性,辅助产生新的解集。将Tent的混沌映射与反向学习相结合,得到Tent反向学习机制。具体数学模型描述为:

$$X_i^{T_0} = lb + ub - \lambda_i \cdot X_i^l \tag{19}$$

$$\lambda_{i+1} = \begin{cases} 2\lambda_i, 0 \leq \lambda_i \leq 0.5 \\ 2(1 - \lambda_i), 0.5 < \lambda_i \leq 1 \end{cases} \tag{20}$$

式中: $X_i^{T_0}$  为种群中第  $i$  只个体对应的Tent反向学习机制生成的解,是对应的Tent混沌映射值。此外,根据光的凸透镜成像原理,提出透镜反向学习策略。该策略利用个体在焦点之外时会在凸透镜另一侧成倒立实像的特性建立数学模型,具体为:

$$X_i = \frac{lb + ub}{2} + \frac{lb + ub}{2} \cdot \frac{X_i}{k} \tag{21}$$

式中: $k$  为透镜成像缩放因子, $k$  的取值影响反向解的生成质量, $k$  值越小,生成的反向解范围越大; $k$  值越大,能提供较小范围的反向解,考虑到算法前期更多进行全局搜索,后期更多在局部位置。因此,提出一种动态调节的缩放因子公式:

$$k = \left[ 1 + \left( \frac{t}{t_{\max}} \right)^{1/3} \right]^c \tag{22}$$

双重反向学习策略,其一是Tent逆向学习;其二是透镜反向学习。该策略通过随机产生[0,1]的随机数与切换概率  $P$  比较大小,进而为算法选择不同的反向学习策略。在文中,切换概率设为  $P=0.5$ 。若  $\text{rand} < P$ ,选择Tent逆向学习,反之选择透镜反向学习。

### 2.3 长短期记忆神经网络LSTM模型

LSTM 是一种以 RNN 为基础进行改进的循环神经网络,它通过遗忘门、输入门和输出门来删除或添加“细胞状态信息”,使其只保留重要信息,因此避免了 RNN 中出现的梯度爆炸现象。相比于单一循环体结构,LSTM 有 3 个门:遗忘门、输入门、输出门。整体结构如图 2。

图 2 中: $x_t$  是  $t$  时刻的输入数据, $c_{t-1}$  表示在  $t-1$  时刻的网络细胞的记忆值, $h_{t-1}$  是  $t-1$  时刻 LSTM 模型的输出值。 $t$  时刻模型输出值分别为  $c_t$  和  $h_t$ 。

遗忘门(ForgetGate)的作用是决定上一时刻的单元状态  $c_{t-1}$  有多少保留到当前时刻的单元状态  $c_t$ 。输入门(InputGate)决定的是当前时刻网络的输入  $x_t$  有多少保存到单元状态  $c_t$ 。输出门(Output-Gate)的作用主要是为了控制长时记忆对当前输出的影响,输出门单元状态将共同决定网络模型最终

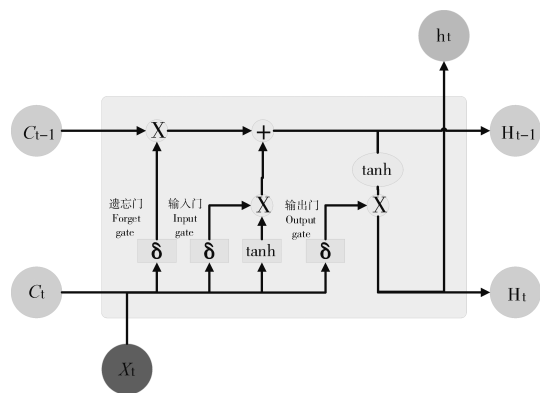


图2 LSTM结构图

Figure 2 LSTM structure diagram

的输出结果。

### 2.4 改进的樽海鞘群算法优化 LSTM 模型

由于 LSTM 模型里隐藏层神经元个数、最大周期及学习率等参数都会对模型预测结果产生较大影响。因此,利用 4 种改进策略优化原始算法的搜索方式、整体布局、跟进方式,通过反向解来解决最优解问题,改进后的算法对神经元个数、最大周期及学习率等参数进行智能迭代寻优,基本流程如图 3。

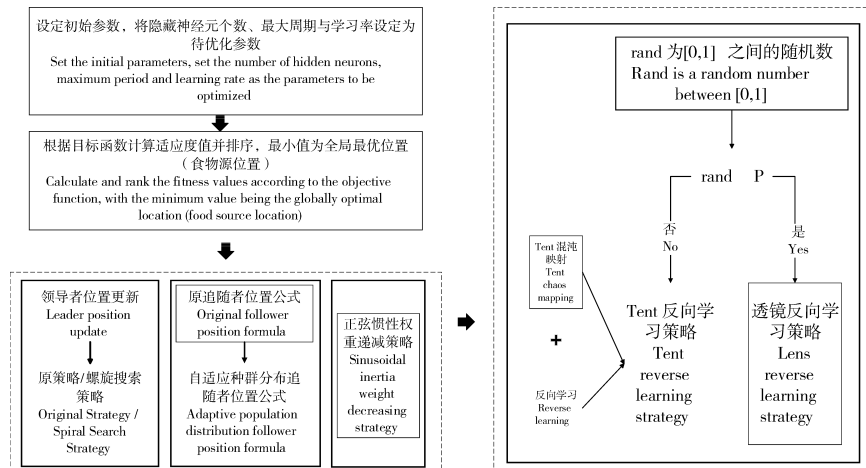


图 3 改进樽海鞘群算法优化 LSTM 参数

Figure 3 Optimization of LSTM parameters by improved salp swarm algorithm

## 3 双层优化 VMD-LSTM 负荷预测模型

在第一层优化模型里,SSA 可以模仿麻雀觅食和逃避捕猎者的行为,进而有效优化模型参数。假定  $N$  个种群中,SSA 算法可以在维度  $D$  上搜索出最佳解,并且可以根据发现者的位置更新,实现更加精确的搜索效果。发现者位置更新为:

$$x_{id}^{t+1} = \begin{cases} x_{id}^t \cdot \exp\left(\frac{-i}{\alpha T}\right), R_2 < ST \\ x_{id}^t + Q \cdot L, R_2 \geq ST \end{cases} \quad (23)$$

式中:  $x_{id}^{t+1}$  为  $t+1$  代第  $i$  只麻雀的位置;  $T$  为最大迭代次数;  $\alpha$  为均匀随机数,在  $(0, 1]$ ;  $Q$  为符合标准正态分布的随机数;  $L$  为矩阵;  $R_2$  为警戒值;  $ST$  为安全值。

加入者位置更新为:

$$x_{id}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{x\omega'_d - x_{id}^i}{i^2}\right), i > \frac{n}{2} \\ x_{id}^{t+1} + \frac{1}{D} \sum_{d=1}^D \left( rand\{-1, 1\} \cdot |x_{id}^i - x_{id}^{t+1}| \right), i \leq \frac{n}{2} \end{cases} \quad (24)$$

式中:  $x\omega'_d$  为最劣位置;  $x_{id}^i$  为最优位置;  $x_{id}^{t+1}$  为种群  $t+1$  次迭代时麻雀在  $d$  维最优位置。

侦察者位置更新为:

$$x_{id}^{t+1} = \begin{cases} x_{id}^t + \beta (x_{id}^t - x_{id}^t), f_i \neq f_g \\ x_{id}^t + K \left( \frac{x_{id}^t - x\omega'_d}{|f_i - f_{\omega}| + e} \right), f_i = f_g \end{cases} \quad (25)$$

式中:  $\beta$  为步长控制参数;  $K$  为  $[-1, 1]$  的随机数,表示麻雀移动方向;  $f_i, f_{\omega}, f_g$  分别表示当前麻雀适应度、最劣适应度和最佳适应度。

根据麻雀算法在寻优上所具有的优势,将VMD中两个重要参数  $\alpha$  和  $k$  作为变量,其余参数由于对实验结果影响不大设为固定参数,在寻优完成后,得出最优麻雀位置即为最佳参数,再将此时的参数代回到变量中成为已知参数,新模型通过寻优后的参数来分解数据集。再结合上文所提到的第二层模型来进行训练并做预测。

其主要流程为:(1)数据预处理,通过 Person 系数进行特征筛选,保留  $p$  绝对值小于 0.05 的强相关变量;(2)搭建第一层模型,对分解求出 VMD 的两个参量,并选择出最优的组合,代回到原模型;(3) SVMd 分解数据集,并将分解出来的 imf 分量采用 ASSSA-LSTM 神经网络分别进行训练;(4)将训练好的 imf 分量分别预测,得到预测值  $P_1$ 、 $P_2$ 、 $\dots$ 、 $P_n$  并将各模态分量叠加形成最终的预测结果;(5)评估预测结果。与真实电力负荷进行对比,通过相应的评价指标评估模型实际预测效果,整体预测框架如图 4。

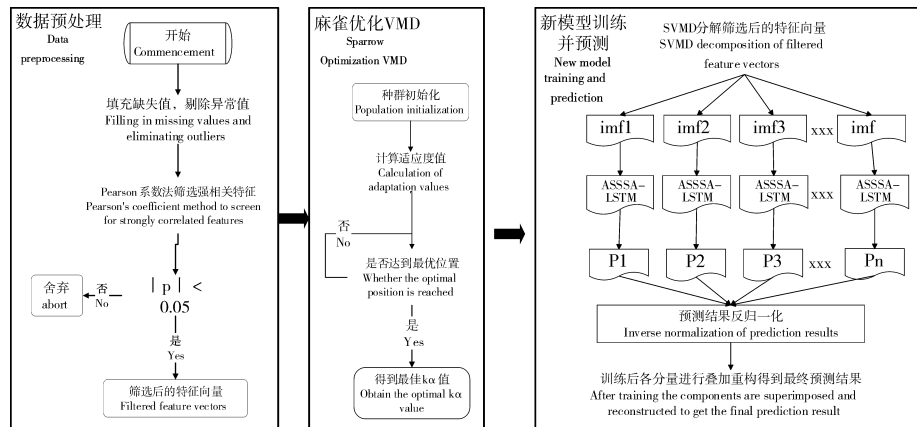


图4 整体流程图

Figure 4 Overall flowchart

### 4 算例分析

本次试验对辽宁省某市某乡村 10 kV 变压器 2020 年 7 月 1 日到 7 月 10 日的变压器历史负荷和有关特性进行了分析。采用以 15 min 为一个采样点的 21 维、3 456 条数据的电力变压器的历史负荷数据。影响负荷预测准确性的有时间、气象等因素。时间方面,白天期间早晨和中午的时间段,夜间负荷会有差别。天气变化也会影响负荷,比如雨雪天气和晴天数据相差会比较大。由于各特征与负荷间存在相关性差异,为便于模型训练,对数据集进行特征筛选,采用皮尔森相关系数检验法,以热力图的形式表现出来,显著性小于 0.05 为强相关特征(图 5),相关性系数绝对值越接近 1 则为强相关特征(图 6)。

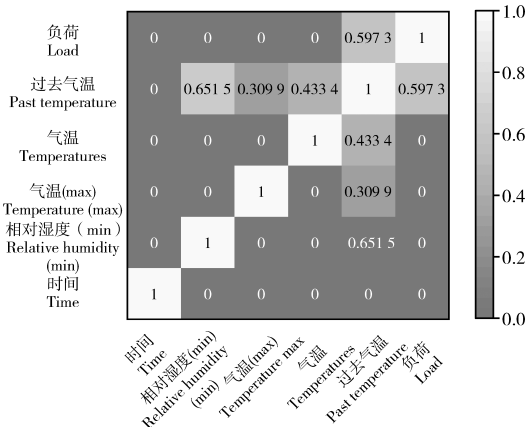


图5 皮尔逊相关性系数检验矩阵

Figure 5 Pilson correlation coefficient test matrix

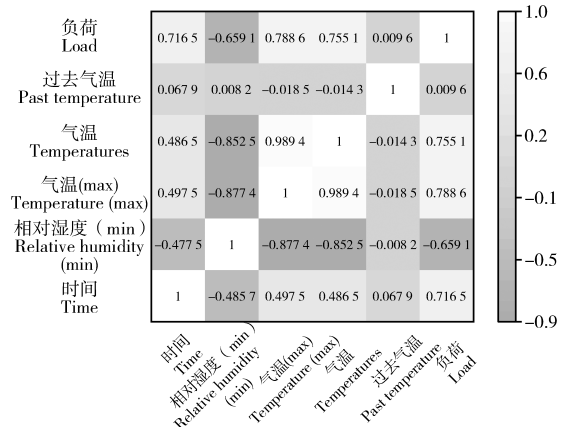


图6 皮尔逊相关性分析系数矩阵

Figure 6 Pilson correlation analysis coefficient matrix

图中 max 和 min 分别代表该特征最大和最小值,可以很明显看出时间、气温、相对湿度是对负荷影响较大的强相关特征,选取前 10 d 的负荷数据,以前 7 d 的数据作为训练后 3 d 的数据作为验证,设定步长为 10。

#### 4.1 麻雀算法的参数和优化 VMD 效果

以包络熵为适应度函数,模态分解数和惩罚因子作为优化参数,迭代次数初始设为 100 次,总共实验 50 次。由于每次运行都在迭代次数为 10 时趋于平稳,故将迭代次数改为 30,选取优化效果最好的曲线如图 7~图 9。

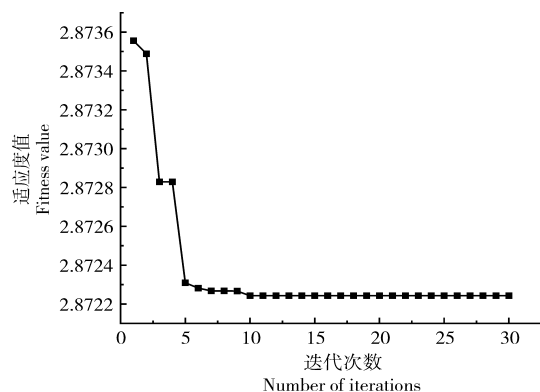


图7 适应度值曲线图

Figure 7 Adapt ability value curve

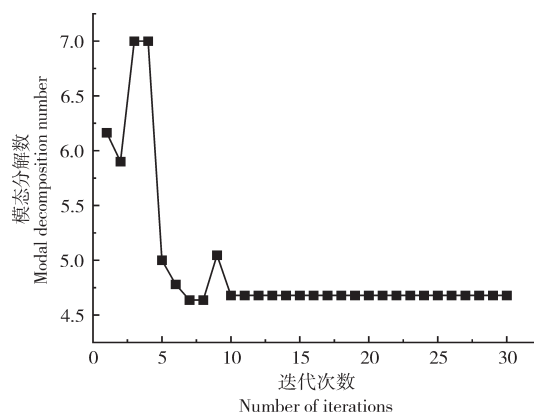


图8 模态分解数曲线图

Figure 8 Modal decomposition number curve

根据上文所述,麻雀最优位置即为  $k$ 、 $\alpha$  的值,试验结果得出 5 和 1 715 为最优解,此时将最优解代入到 SVM 中可以分解得到图 10 的模态分量 (intrinsic mode functions, IMF)。

#### 4.2 不同算法模型预测对比

为验证 SVM-ASSA-LSTM 模型的优越性和改进樽海鞘群算法的有效性和麻雀算法优化 VMD 参数的必要性,从而进一步说明所提出的方法对预测负荷有着明显提升和改善,分别选用 LSTM、PSO-LSTM、VMD-LSTM、SSA-LSTM 4 种基础模型去做对比,预测结果采用 RMSE、MAE、MAPE、 $R^2$  指标进行对比验证,函数表达式分别为:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (26)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (27)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (28)$$

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2} \quad (29)$$

式中:均方根误差 (RMSE)、平均绝对误差 (MAE)、平均绝对百分比误差 (MAPE) 随着数值减小,精确度增加。确定系数 ( $R^2$ ) 越大,精确度越高。 $y_i$  为第  $i$  点的实际负荷值, $\hat{y}_i$  为第  $i$  点的预测负荷值, $n$  为预测点个数。图 11 为不同模型预测效果对比,表 1 为试验结果数据对比。

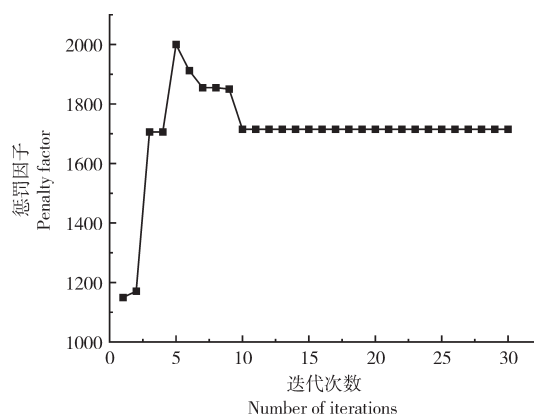


图9 惩罚因子曲线图

Figure 9 Penalty factor curve

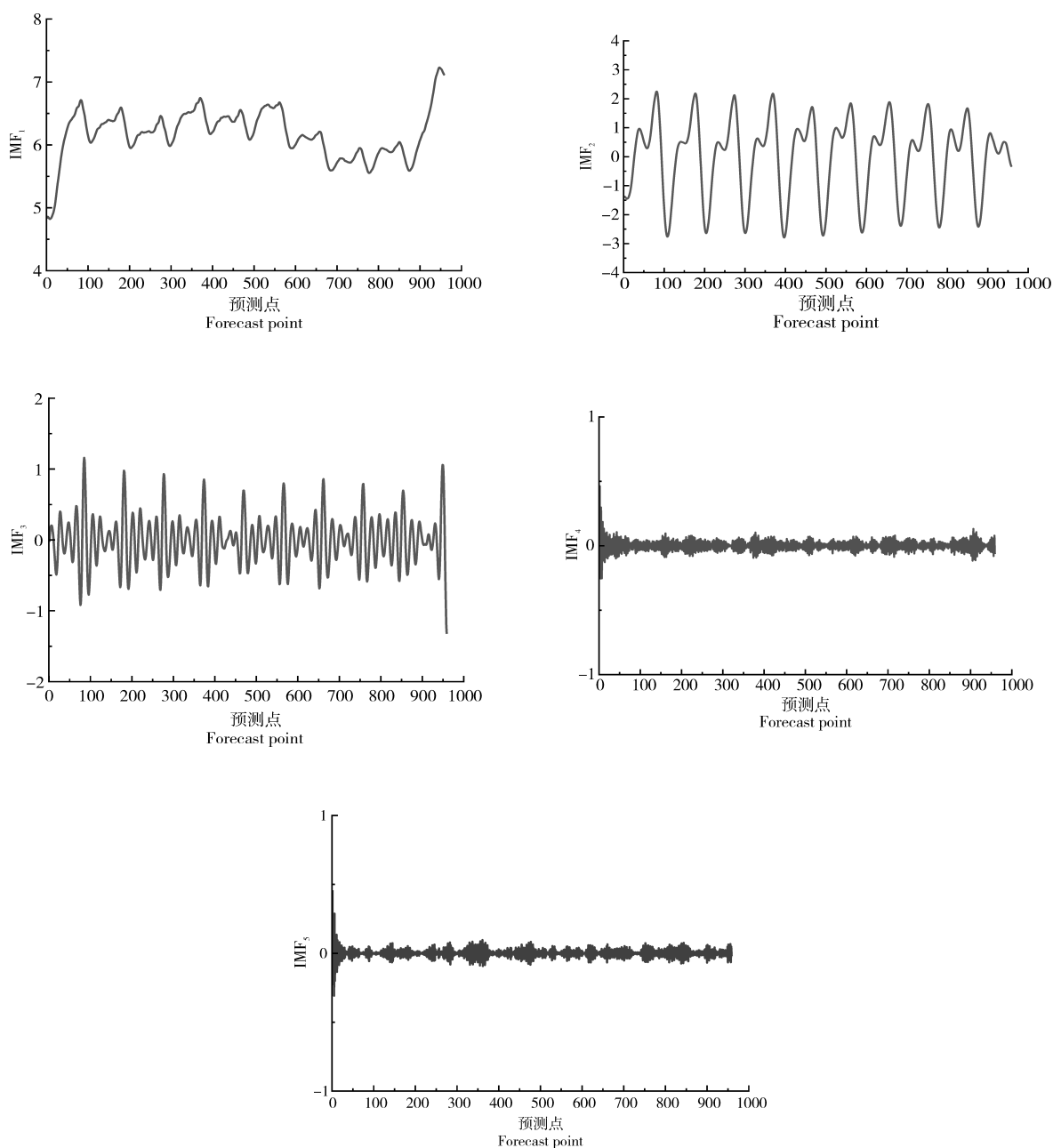


图10 各模态分量图

Figure 10 Diagram of each modal component

可以明显看出,  $RMSE$  指标较对比新模型分别降低 0.200 8, 0.115, 0.202 8, 0.070 3,  $MAE$  降低 0.133 1, 0.100 6, 0.147 2, 0.052 1,  $MAPE$  降低 2.656 1%, 1.867 9%, 2.512 5%, 0.631 5%,  $R^2$  增加 0.04 4, 0.02 1, 0.05 1, 0.01 2, 为验证新模型泛化性, 将双层优化应用于  $LSSVM$  模型和  $GRU$  模型, 结果如图 12。

试验结果再次证明(表2), 模型具有泛化性的特点。有效验证了改进樽海鞘群优化算法的优越性和  $VMD$  参数智能寻优的可靠性以及融合策略的有效性, 可针对农村地区负荷情况进行精确预测。

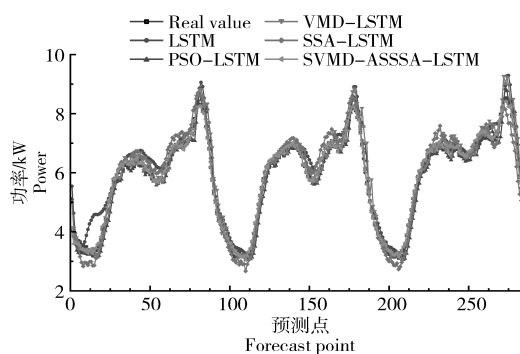


图11 4种指标对比

Figure 11 Four indicators comparison

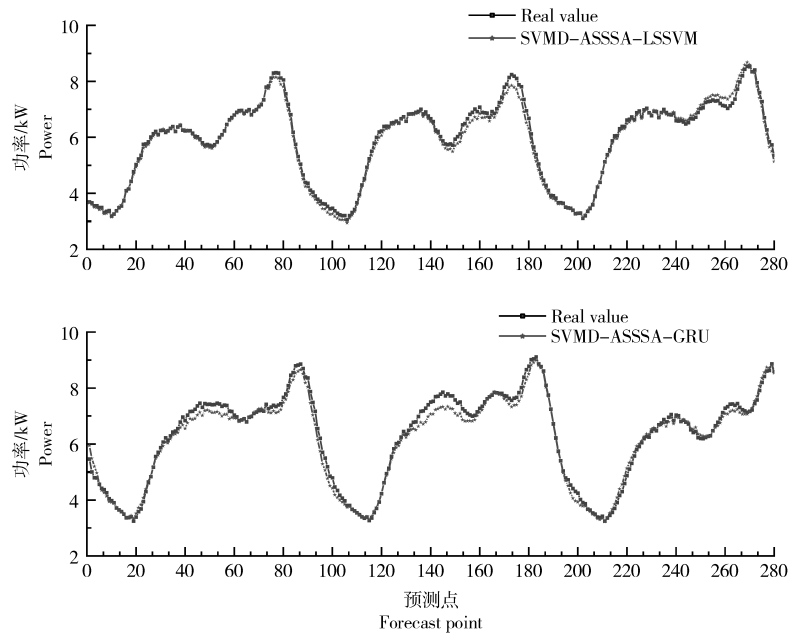


图12 模型验证结果

Figure 12 Model verification results

表1 试验结果对比

Table 1 Experimental results comparison

| 算法种类<br>Types of algorithms | RMSE   | MAE    | MAPE/% | R <sup>2</sup> |
|-----------------------------|--------|--------|--------|----------------|
| LSTM                        | 0.3760 | 0.2690 | 5.2107 | 0.941          |
| PSO-LSTM                    | 0.2902 | 0.2365 | 4.4225 | 0.964          |
| VMD-LSTM                    | 0.3780 | 0.2831 | 5.0671 | 0.934          |
| SSA-LSTM                    | 0.2489 | 0.1880 | 3.1861 | 0.973          |
| SVM-D-ASSSA-LSTM            | 0.1752 | 0.1359 | 2.5546 | 0.985          |

表2 两种模型验证结果

Table 2 Two model verification results

| 模型 Model          | RMSE   | MAE    | MAPE/% | R <sup>2</sup> |
|-------------------|--------|--------|--------|----------------|
| SVM-D-ASSSA-GRU   | 0.2135 | 0.1659 | 2.6131 | 0.9820         |
| SVM-D-ASSSA-LSSVM | 0.1657 | 0.1274 | 2.1427 | 0.9871         |

## 5 结论

本研究结果表明,利用麻雀算法智能寻优的特点找出了VMD的两个最优参数,有效提升了算法性能和预测结果精度。在标准樽海鞘群算法的基础上进了4点改进,使基础模型有了更好的寻优和收敛效果,可以高效找到LSTM模型的最优参数,进行更精准的预测。两种优化模型结合后形成的SVM-D-ASSSA-LSTM模型相比于基础预测模型收敛速度更快,精度更高,预测效果更加稳定,对供电公司用电规划及未来农村电力发展建设也会起着更重要的作用。

### 参考文献:

- [1] 杨继革,严俊,陈丽春,等.基于智能电表的住宅短期电力负载预测[J].沈阳工业大学学报,2022,44(3):255-258.
- [2] 陈纬楠,胡志坚,岳菁鹏,等.基于长短期记忆网络和LightGBM组合模型的短期负荷预测[J].电力系统自动化,2021,45(4):91-97.
- [3] 伍乙杰,黄文灏,赖仕达,等.基于随机森林和双向长短期记忆网络的超短期负荷预测研究[J].电气自动化,2022,44(5):35-37,40.
- [4] 高明,郝妍.基于BiLSTM网络与误差修正的超短期负荷预测[J].综合智慧能源,2023,45(1):31-40.
- [5] 张照贝,顾春华,温蜜.基于XGBoost和QLRSTM的超短期负荷预测方法[J].计算机仿真,2022,39(1):90-95,110.

- [6] 梁敬东,崔丙剑,姜海燕,等.基于 word2vec 和 LSTM 的句子相似度计算及其在水稻 FAQ 问答系统中的应用[J].南京农业大学学报,2018,41(5):946-953.
- [7] 林 涵,郝正航,郭家鹏,等.基 TCA-CNN-LSTM 的短期负荷预测研究[J].电测与仪表,2023,60(8):73-80.
- [8] 邹 智,吴铁洲,张晓星,等.基于贝叶斯优化 CNN-BiGRU 混合神经网络的短期负荷预测[J].高电压技术,2022,48(10):3935-3945.
- [9] 刘立立,刘 洋,唐子卓,等.基于 Luong Attention 机制和特征优选策略的超短期负荷预测方法[J].电力系统及其自动化学报,2022,34(4):143-150.
- [10] 黄文静,戚建国.基于 LASSO 和 PSO-LSTM 的综合能源系统负荷预测[J].科技与创新,2023(23):13-15.
- [11] 路 航.基于 KPCA-WOA-BiLSTM 模型的短期电力负荷预测[D].哈尔滨:东北农业大学,2023.
- [12] 王雨前,王万雄.基于 EMD-Bayes-SVR 组合模型的短期电力负荷预测[J].电子科技,2023,36(12):64-71
- [13] 张雲钦,程起泽,蒋文杰,等.基于 EMD-PCA-LSTM 的光伏功率预测模型[J].太阳能学报,2021,42(9):62-69.
- [14] 闫文杰,徐志杰,薛桂香,等.基于 EMD-TCN 的多能供热系统太阳辐照度预测模型[J].太阳能学报,2023,44(11):182-188.
- [15] 梁浩彬.基于一种新型的 EMD-LSTM-ARIMA 算法的超短期电力负荷预测[J].现代工业经济和信息化,2023,13(10):194-198.
- [16] 刘亚琿,赵 倩.基于聚类经验模态分解的 CNN-LSTM 超短期电力负荷预测[J].电网技术,2021,45(11):4444-4451.
- [17] 黄 炜,陈 田.基于 SAE 与 CEEMDAN-BiLSTM 组合模型的短期电力负荷预测[J].计算机应用与软件,2022,39(7):52-58.
- [18] 齐少拴.基于 EEMD-SSA 组合模型的短期电力负荷预测研究[D].大庆:东北石油大学,2022.
- [19] 张淑清,宿新爽,陈荣飞,等.基于变分模态分解和 FABP 的短期电力负荷预测[J].仪器仪表学报,2018,39(4):67-73.
- [20] 汤义勤,邹宏亮,蒋 旭,等.基于 VMD 和贝叶斯优化 LSTM 的母线负荷预测方法[J].电网与清洁能源,2023,39(2):46-52,59.
- [21] 张 未,余成波,王士彬,等.基于 VMD-LSTM-LightGBM 的多特征短期电力负荷预测[J].南方电网技术,2023,17(2):74-81.
- [22] 方 娜,陈 浩,邓 心,等.基于 VMD-ARIMA-DBN 的短期电力负荷预测[J].电力系统及其自动化学报,2023,3(6):59-65.
- [23] 杨晶显,张 帅,刘继春,等.基于 VMD 和双重注意力机制 LSTM 的短期光伏功率预测[J].电力系统自动化,2021,45(3):174-182.
- [24] CAI C,LI Y,SU Z.Short-term electrical load forecasting based on VMD and gru-tcn hybrid network[J].Applied Sciences,2022,12(13):6647.
- [25] 王建元,张宇辉,刘 铖.基于参数优化 VMD 和改进 K 聚类判据融合的配电网故障选线方法[J].南方电网技术,2023,17(7):135-145.
- [26] 徐 岩,向益锋,马天祥.基于粒子群算法优化参数的 VMD-GRU 短期电力负荷预测模型[J].华北电力大学学报(自然科学版),2023,50(1):38-47.
- [27] GENG G,HE Y,ZHANG J.Short-term power load forecasting based on PSO-optimized vmd-tcn-attention mechanism[J].Energies,2023,16(12):4616.
- [28] HUANG W,SONG Q,HUANG Y.Two-stage short-term power load forecasting based on SSA - VMD and feature selection[J].Applied Sciences,2023,13(11):6845.
- [29] 肖烈禧,张 玉,周 辉,等.基于 IAOA-VMD-LSTM 的超短期风电功率预测[J].太阳能学报,2023,44(11):239-246.
- [30] 丁德邻,张 营,左洪福.基于 SSA-VMD-GRU 的锂电池剩余寿命预测方法研究[J].电子元件与材料,2023,42(9):1071-1078.
- [31] 严守平,萧展辉.基于自编码神经网络的电网故障预测算法[J].沈阳工业大学学报,2023,45(1):1-5.

[责任编辑 马迎杰]