

基于改进麻雀算法的短期负荷预测研究

孙夏丽, 李士心*, 王 坤, 刘清清

天津职业技术师范大学电子工程学院, 天津
Email: 2713229103@qq.com, *lishixin@tute.edu.cn

收稿日期: 2021年8月12日; 录用日期: 2021年9月8日; 发布日期: 2021年9月15日

摘 要

针对麻雀算法(SSA)早熟、易陷入局部最优解的问题,采用混沌策略反向学习初始化种群代替随机初始化种群,增强种群多样性;同时引入黄金正弦因子更新发现者位置以平衡局部和全局搜索能力;引入自适应t分布改进跟随者位置。与粒子群算法、灰狼算法作对比,在测试函数上验证改进麻雀算法的优越性。采用改进后的麻雀算法优化BP神经网络模型(GSASSA-BP)对短期负荷问题进行预测,结果表明,提出的GSASSA-BP模型在短期负荷问题上预测效果更好。

关键词

麻雀算法, 混沌, 黄金正弦, t分布, 预测, 短期负荷

Short Term Load Forecasting Based on Improved Sparrow Algorithm

Xiali Sun, Shixin Li*, Kun Wang, Qingqing Liu

School of Electronic Engineering, Tianjin University of Technology and Education, Tianjin
Email: 2713229103@qq.com, *lishixin@tute.edu.cn

Received: Aug. 12th, 2021; accepted: Sep. 8th, 2021; published: Sep. 15th, 2021

Abstract

In order to solve the problem of being premature and easy to fall into the local optimal solution of sparrow algorithm, chaos strategy is used to replace random initialization population to enhance the diversity of population. At the same time, the golden sine factor is introduced to update the location of the discoverer to balance the local and global search ability. Adaptive t-distribution is introduced to improve the follower position. Compared with particle swarm optimization algo-

*通讯作者。

rithm and gray wolf algorithm, the superiority of the improved sparrow algorithm is verified in the test function. The improved sparrow algorithm is used to optimize the BP neural network model to predict the short-term load problem. The results show that the proposed GSASSA-BP model has better prediction effect on the short-term load problem.

Keywords

Sparrow Algorithm, Chaos, Golden Sine, t Distribution, Forecast, Short Term Load

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

短期负荷预测是编排发电计划、调度计划、确保系统安全稳定运行的基础[1]，但易受气象条件、节假日等因素的影响，对其精确的预测显得尤为重要。

近年来，短期负荷已由传统的时间序列法[2]、灰色预测[3]、趋势外推法[4]等发展到今天的随机森林法[5]、模糊理论[6]、人工神经网络[7]、支持向量机[8]等方法。例如文献[9]采用自适应柯西变异粒子群对 LSTM 模型参数进行寻优，提高短期电力负荷预测的精度和稳定性；文献[10]提出多阶段优化的变分模态分解和粒子群算法优化支持向量回归的短期电力负荷预测模型，精度得到了提高；文献[11]利用花朵授粉全局搜索能力强特点对 BP 神经网络的权值阈值进行优化，提高了短期负荷预测的精度。

针对 BP 神经网络收敛速度较慢且易陷入局部极值缺点，采用混沌反向策略初始化种群；引入黄金正弦改进发现者的位置，并在跟随者位置更新中采用 t 分布策略的方法改进麻雀算法，利用改进后的麻雀算法优化 BP 神经网络，应用于短期电力负荷预测中，仿真结果表明，改进后的方法预测精度提高。

2. 麻雀算法

SSA 算法[12]与其他算法相比，具有收敛速度快，稳定性好等优点，但易陷入局部最优解。麻雀算法由发现者、跟随者和预警者组成的。设 n 只麻雀组成的群体及对应适应度函数分别可表示为：

$X = [x_1, x_2, \dots, x_n]^T$ 和 $F = [f(x_1), f(x_2), \dots, f(x_n)]^T$ 。发现者位置更新如下：

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \cdot \exp\left(-\frac{i}{\alpha \cdot T_{\max}}\right), & R_2 < ST \\ X_{i,j}^t + Q \cdot L, & R_2 \geq ST \end{cases} \quad (1)$$

式中， i 为迭代次数； $j = 1, 2, \dots, d$ 。 T_{\max} 为最大迭代次数； X_{ij} 为第 i 只麻雀在第 j 维中的位置； $\alpha \in (0, 1]$ 为随机数； R_2 ($R_2 \in [0, 1]$) 和 ST ($ST \in [0.5, 1]$) 分别为预警值和安全值； Q 是服从正态分布的随机数； L 为 $1 \times d$ 的全 1 矩阵。

跟随者的位置更新为：

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{\text{worst}} - X_{i,j}^t}{i^2}\right), & i > \frac{n}{2} \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| \cdot A^+ \cdot L, & \text{otherwise} \end{cases} \quad (2)$$

式中, X_p, X_{worst} 分别为当前全局最优和最差的位置。 A 为 $1 \times d$ 的全 1 或 -1 矩阵, 且 $A^+ = A^T (AA^T)^{-1}$ 。预警麻雀占种群的 10%~20%, 位置更新为:

$$X_{i,j}^{t+1} = \begin{cases} X_{best}^t + \beta \cdot |X_{i,j}^t - X_{best}^t|, & f_i > f_g \\ X_{i,j}^t + K \cdot \left(\frac{X_{i,j}^t - X_{worst}^t}{(f_i - f_w) + \varepsilon} \right), & f_i = f_g \end{cases} \quad (3)$$

其中, X_{best} 为当前全局最优位置。 β 为服从(0,1)正态分布的随机数。 $K \in [-1,1]$ 的随机数, f_i, f_g 和 f_w 分别为当前麻雀个体、全局最佳和最差的适应度值。 ε 为最小的常数。

3. 改进麻雀算法

3.1. 混沌策略反向学习初始化种群

随机生成初始种群虽能保证初始位置均匀分布, 但不能保证解的质量, 影响算法的收敛速度及精度。本文采用反向学习初始化的方法引入 Logistic 映射策略[13]。首先利用 Logistic 映射生成 N 个向量, 再使用反向学习生成相对应的 N 个反向初始解, 然后从初始解和反向解中选择前 N 个适应度较好的个体作为初始化种群。

3.2. 黄金正弦因子

2017 年, Tanyildizi 等人提出黄金正弦算法(Golden-SA) [14], 具有寻优能力强、设置参数少、原理简单等特点。利用上述方式生成的 N 个个体, 在发现者的位置更新上采用黄金正弦因子, 改进后的更新方式如下:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \times |\sin(R_1)| + R_2 \times \sin(R_1) \times |x_1 \times P_i^t - x_2 \times X_{i,j}^t|, & R_2 < ST \\ X_{i,j}^t + Q \cdot L, & R_2 \geq ST \end{cases} \quad (4)$$

式中, $R_1 \in [0, 2\pi], R_2 \in [0, \pi]$ 为随机数, 黄金分割系数 x_1, x_2 缩小搜索空间, 平衡搜索和开发, 使个体趋向最优值, $x_1 = -\pi + (1 - \tau) \times 2\pi, x_2 = -\pi + \tau \times 2\pi$, 黄金分割数 $\tau = (\sqrt{5} - 1) / 2$ 。

3.3. 自适应 t 分布

在 SSA 跟随者位置中引入自适应 t 分布[15]变异策略, 可以增加种群的多样性, 避免陷入局部最优。更新后的位置更新方式如下:

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{worst}^t - X_{i,j}^t}{i^2}\right), & i > \frac{n}{2} \\ X_p^{t+1} + X_p^{t+1} * t(iter), & \text{otherwise} \end{cases} \quad (5)$$

式中, X_p^{t+1} 为当前发现者的最佳位置。

3.4. 改进算法性能测试

3.4.1. 测试函数及参数设置

为验证改进麻雀算法的性能, 与粒子群算法(PSO) [16]、灰狼算法(GWO) [17]进行比较, 参数设置如表 1 所示。选取 6 个标准测试函数进行测试, 详见表 2。本文测试实验环境为: Inter(R) Core(TM) i7-4700HQ CPU @ 2.40GHz, 8GB 内存, Window7 系统和 matlab R2019b。种群规模为 30, 最大迭代次数为 1000, 维数为 30。

Table 1. Parameter setting
表 1. 参数设置

算法	参数设置
PSO	$c_1 = c_2 = 1, \omega = 0.3, V_{\max} = 2, V_{\min} = -2$
GWO	α 从 2 递减至 0
SSA	PD = 20%, SD = 10%
GSASSA	PD = 20%, SD = 10%

Table 2. Benchmark function
表 2. 基准测试函数

函数名	函数	取值范围	最大值	函数名	函数	取值范围	最大值
Sphere	f_1	[-100, 100]	0	Schwefel2.1	f_4	[-100, 100]	0
Schwefel2.22	f_2	[-10, 10]	0	Rastrigin	f_5	[-5.12, 5.12]	0
Schwefel1.2	f_3	[-100, 100]	0	Griewank	f_6	[-600, 600]	0

3.4.2. 测试结果

针对同一函数，每种算法独立运行 20 次，并统计 4 种算法在 6 个基准函数上的平均值、标准差，结果如表 3 所示；寻优曲线趋势如图 1 所示。

由表 3 可知，对 $f_1 - f_4$ ，GSASSA 的寻优结果优于其他 3 种优化算法，GSASSA 能寻到最优值 0，而 PSO、GWO、SSA 中，GWO 在 f_1 、 f_4 上的均值和标准差优于其他 2 种智能算法，SSA 在 f_2 、 f_3 上的寻优效果强。对于函数 f_5 、 f_6 ，SSA 和 GSASSA 都能寻到最优解 0，GWO 的寻优效果相对于 PSO 较优。

由图 1(a)~(f)可知，GSASSA 算法的寻优速度及精度都优于其他几种算法。

Table 3. Comparison with the optimization results of standard optimization function
表 3. 与标准优化函数寻优结果对比

算法	函数	f_1	f_2	f_3	f_4	f_5	f_6
PSO	均值	1.43E+03	2.73E+01	1.04E+05	1.48E+01	1.87E+02	1.47E+01
	标准差	4.40E+02	3.94E+00	3.79E+04	2.53E+00	2.46E+01	3.83E+00
GWO	均值	2.34E-58	1.41E+00	7.35E+00	1.47E-14	2.84E-15	2.08E-03
	标准差	4.06E-58	6.16E+00	7.88E+00	2.33E-14	1.24E-14	4.28E-03
SSA	均值	1.50E-46	2.75E-50	1.85E-17	9.28E-10	0	0
	标准差	6.15E-46	1.20E-49	7.19E-17	3.79E-09	0	0
GSASSA	均值	0	0	0	0	0	0
	标准差	0	0	0	0	0	0

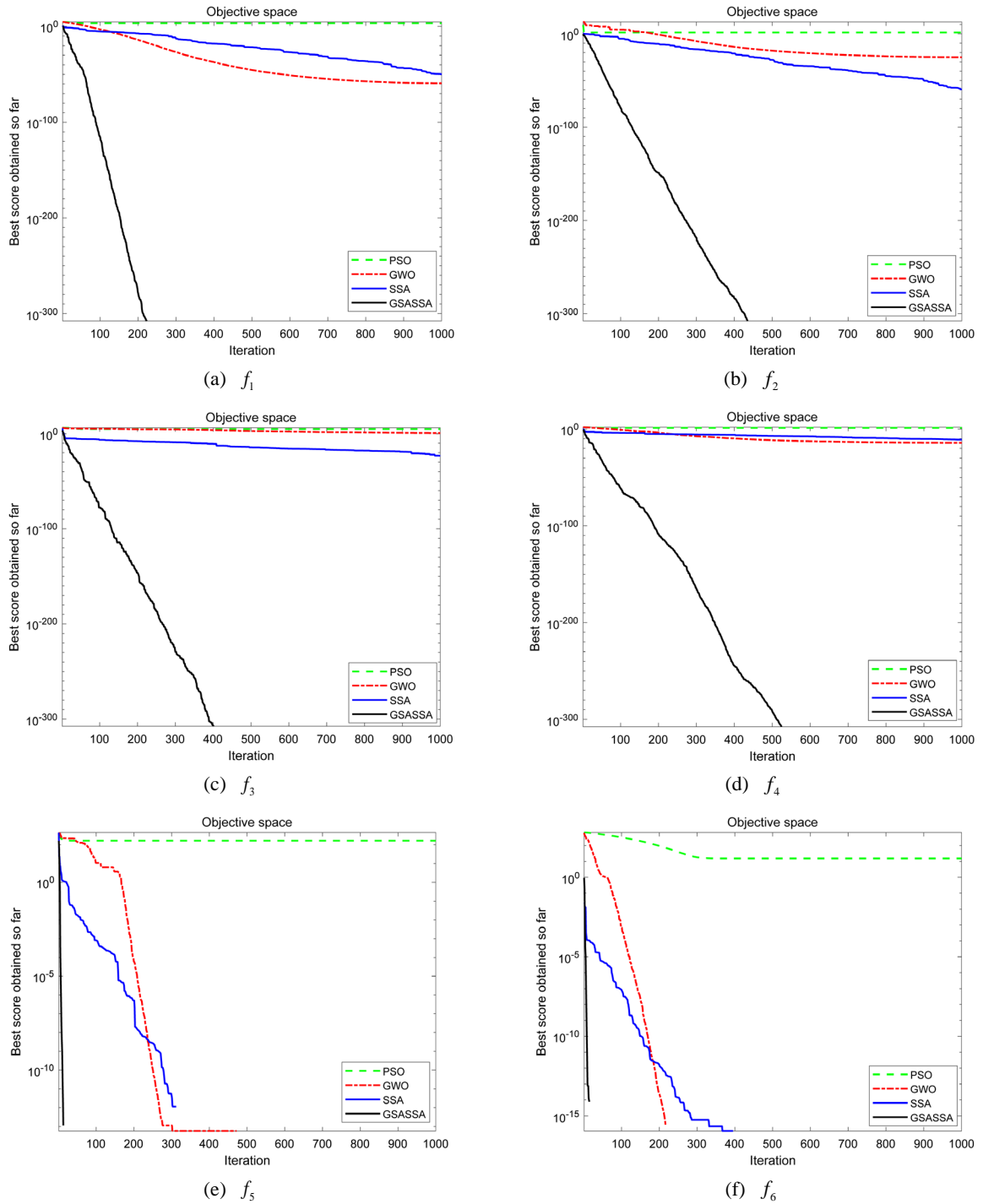


Figure 1. Optimization test results of different algorithms

图 1. 不同算法的寻优测试结果

4. GSASSA-BP 模型

4.1. BP 神经网络

BP 神经网络是采用梯度下降法，调整相应的权值和阈值，使网络的实际输出值和期望输出值的误差

均方差最小。BP神经网络一般由输入层、隐含层、输出层组成[18]。 $X = [x_1, x_2, \dots, x_n]$ 、 $Z = [z_1, z_2, \dots, z_k]$ 和 $Y = [y_1, y_2, \dots, y_m]$ 分别为输入层、隐含层和输出层，各层之间通过全连接进行连接。

4.2. GSASSA-BP 模型

GSASSA-BP 采用改进后的麻雀算法对神经网络的初始权值和阈值进行优化，使得优化后的 BP 神经网络收敛速度更快，全局搜索能力更强。GSASSA-BP 流程图[19]如图 2 所示。

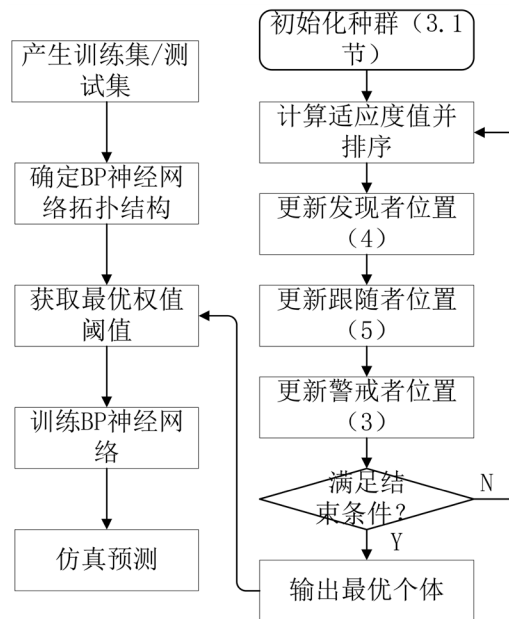


Figure 2. GSASSA-BP network model process
图 2. GSASSA-BP 网络模型流程

5. 算例分析

5.1. 数据选取及预处理

本文实验数据选取 2016 年电工数学建模竞赛负荷预测数据集，包括 2012.01 到 2015.01 的最高温度、最低温度、平均温度、相对湿度(平均)、降雨量及日需求量负荷(kWh)数据。其中，前 1099 天的日需求量负荷及气象条件作为训练数据，后 7 天的数据作为测试数据。

5.1.1. 数据预处理

在数据收集时，由于各种原因会出现数据缺失的现象，对其处理方法如下：

$$I_d = (I_{d-1} + I_{d+1}) / 2 \tag{6}$$

其中， I_d, I_{d-1}, I_{d+1} 分别为第 d 天缺失点的数据，第 $d-1$ 天缺失的数据和第 $d+1$ 天缺失的数据。

5.1.2. 数据归一化处理

电力负荷数据中存在差异较大，为增加预测的准确性，需要对数据做归一化处理，使各种影响因素在同一数值范围内。

$$x_n = (x - x_{\min}) / (x_{\max} - x_{\min}) \tag{7}$$

其中， x 为即将归一化处理的数据， x_n, x_{\max}, x_{\min} 分别为处理后的数据、处理数据中的最大和最小值。

针对降雨量进行分段处理再进行量化取值, 具体如下表 4。

Table 4. Rainfall value table

表 4. 降雨量取值表

降雨量 R(mm)	0	0~10	10~25	25~50	50~100	100~250	>250
量化系数	0.8	0.7	0.6	0.5	0.4	0.3	0.2

5.1.3. 相关性分析

使用 Pearson 相关系数法对电力负荷数据和各气象条件进行相关性分析, Pearson 相关系数是用来描述两个变量 X 和 Y 之间的线性关系, 计算公式如下:

$$\rho_{x,y} = \frac{N \sum XY - \sum X \sum Y}{\sqrt{N \sum X^2 - (\sum X)^2} \sqrt{N \sum Y^2 - (\sum Y)^2}} \quad (8)$$

式中, X 为电力负荷数据, Y 为各气象条件, N 为数据总量个数, $\rho_{x,y} \in (-1,1)$, 结果越接近-1 或 1 表示两个变量之间的线性关系越强。经计算得, 日需求负荷分别与最高温度、最低温度、平均温度、相对湿度、降雨量的相关系数为 0.5947、0.6380、0.6391、0.1171、-0.1143, 存在相关性, 故把这些变量作为输入, 以提高电力负荷预测精度。

5.2. 实例预测及结果分析

本文采用数据对 BP、PSO-BP、SSA-BP 及 GSASSA-BP 神经网络的预测模型进行验证, 得到预测结果如下图 3 所示。

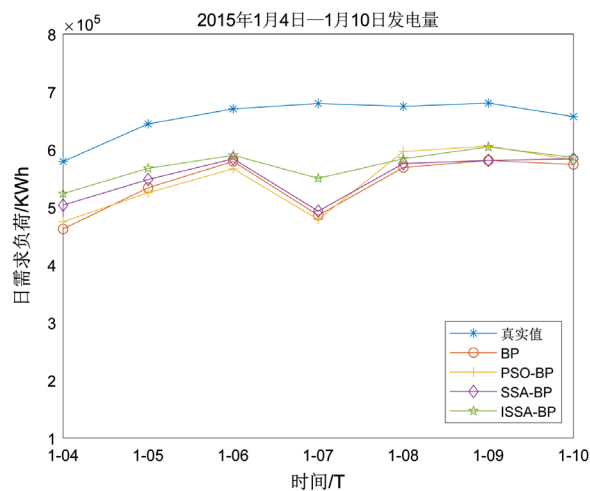


Figure 3. Simulation and prediction results of each prediction model

图 3. 各预测模型仿真预测结果图

由图可知, GSASSA-BP 的预测结果相较于 SSA-BP、PSO-BP、BP 更接近于实际负荷曲线。

为准确评估四种模型的预测精度, 采用平均绝对百分误差(MAPE)进行估计, 公式如下:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \quad (9)$$

各神经网络预测对比结果见下表 5:

Table 5. Comparison of prediction error results
表 5. 预测误差结果比较

预测模型	BP	PSO-BP	SSA-BP	GSASSA-BP
MAPE/%	17.47	16.52	15.54	12.57

由表 5 可得, GSASSA-BP 的 MAPE 比 BP 低 4.9%, 比 SSA-BP 低 2.97%, 说明优化后的麻雀算法预测效果能力更强。

6. 结束语

麻雀算法是一种新的智能算法。针对麻雀算法的局部搜索能力差的缺点, 使用 Logistic 反向学习初始化种群, 提高种群初始解的质量; 在发现者位置更新中引入黄金正弦因子协调搜索和开发的能力; 并在跟随者位置更新中加入 t 分布策略, 提高全局搜索能力。选用 BP 神经网络对短期电力负荷进行预测, 针对 BP 神经网络收敛速度慢、收敛精度不高的问题, 结合改进的麻雀算法提出 GSASSA-BP 模型, 使用该改进麻雀算法对 BP 神经网络的权值和阈值进行寻优, 实验结果表明, GSASSA-BP 模型提高了 BP 神经网络的收敛精度, 预测效果更好。

基金项目

天津市科技特派员项目(19JCTPJC41500)。

参考文献

- [1] 夏博, 杨超, 李冲. 电力系统短期负荷预测方法研究综述[J]. 电力大数据, 2018, 21(7): 22-28.
- [2] 陈冬泮, 郑舟, 吴永峰. 基于时间序列的电力系统短期负荷预测问题分析[J]. 自动化应用, 2017(11): 99-101.
- [3] 许惠君, 王宗耀, 李中庆, 等. 基于改良灰色理论的电网短期负荷预测算法[J]. 计算机与现代化, 2021(3): 24-27+34.
- [4] 张彦宇, 肖茜. 国内外关于电力系统负荷预测的研究现状分析[J]. 山东工业技术, 2016(11): 215+208.
- [5] 李焱, 贾雅君, 李磊, 等. 基于随机森林算法的短期电力负荷预测[J]. 电力系统保护与控制, 2020, 48(21): 117-124.
- [6] 危志强. 基于改进模糊聚类算法的短期负荷预测研究[D]: [硕士学位论文]. 南昌: 南昌大学, 2020.
- [7] 范海虹. 基于萤火虫算法的短期电力负荷预测方法[J]. 中国电力, 2021, 54(3): 141-148.
- [8] 贾鹏, 杨炼鑫, 唐一鸣, 等. 基于 SVM 算法在电力负荷预测中的研究[J]. 科技视界, 2020(31): 14-16.
- [9] 魏腾飞, 潘庭龙. 基于改进 PSO 优化 LSTM 网络的短期电力负荷预测[J/OL]. 系统仿真学报, 2021, 33(8): 9. <http://kns.cnki.net/kcms/detail/11.3092.V.20200910.1734.013.html>, 2021-07-18.
- [10] 李文武, 石强, 李丹, 等. 基于 VMD 和 PSO-SVR 的短期电力负荷多阶段优化预测[J/OL]. 中国电力, 1-9. <http://kns.cnki.net/kcms/detail/11.3265.TM.20201221.1513.019.html>, 2021-07-18.
- [11] 牛庆, 曹爱民, 陈潇一, 等. 基于花朵授粉算法和 BP 神经网络的短期负荷预测[J]. 电网与清洁能源, 2020, 36(10): 28-32.
- [12] Xue, J. and Shen, B. (2020) A Novel Swarm Intelligence Optimization Approach: Sparrow Search Algorithm. *Systems Science & Control Engineering*, **8**, 22-34. <https://doi.org/10.1080/21642583.2019.1708830>
- [13] 王茜, 何庆, 林杰, 杨荣莹. 精英反向学习带扰动因子的混沌蚁狮算法[J]. 智能计算机与应用, 2020, 10(8): 51-57.
- [14] Tanyildizi, E. and Demi, R.G. (2017) Golden Sine Algorithm: A Novel Math Inspired Algorithm. *Advances in Electrical and Computer Engineering*, **17**, 71-79. <https://doi.org/10.4316/AECE.2017.02010>
- [15] 柳子来, 王健敏. 基于自适应 t 分布的改进粒子群实时任务调度算法[J]. 化工自动化及仪表, 2020, 47(5): 393-397+424.
- [16] Kennedy, J. and Eberhart, R.C. (1995) Particle Swarm Optimization. *IEEE International Conference on Neural Net-*

works, **4**, 1942-1948.

- [17] Mirjalili, S., Mirjalili, S.M. and Lewis, A. (2014) Grey Wolf Optimizer. *Advances in Engineering Software*, **69**, 46-61. <https://doi.org/10.1016/j.advengsoft.2013.12.007>
- [18] 谢宇希, 颜拥军, 李翔, 等. 基于 BP 神经网络的核探测器故障诊断方法研究[J/OL]. 原子能科学技术: 1-8. <http://kns.cnki.net/kcms/detail/11.2044.TL.20210401.1431.004.html>, 2021-07-18.
- [19] 刘伟军, 李念. 基于 MEA-BP 神经网络建筑工程造价指数预测[J]. 工程经济, 2021, 31(1): 23-27.