Elman neural network for predicting aero optical imaging deviation based on improved slime mould algorithm*

XU Liang[1], WANG Luyang[1], XUE Wei[2], ZHAO Shiwei[1], and ZHOU Liye[1]**

1. Tianjin Key Laboratory of Complex Control Theory and Application, School of Electronic Engineering and Automation, Tianjin University of Technology, Tianjin 300384, China

2. China Academy of Aerospace Science and Innovation, Beijing 100048, China

(Received 4 August 2022; Revised 21 September 2022)

This research suggests a methodology to optimize Elman neural network based on improved slime mould algorithm (ISMA) to anticipate the aero optical imaging deviation. The improved Tent chaotic sequence is added to the SMA to initialize the population to accelerate the algorithm's speed of convergence. Additionally, an improved random opposition-based learning was added to further enhance the algorithm's performance in addressing problems that the SMA has such as weak convergence ability in the late iteration and an easy tendency to fall into local optimization in the optimization process when solving the optimization problem. Finally, the algorithm model is compared to the Elman neural network and the SMA optimization Elman neural network model. The three models are assessed using four evaluation indicators, and the findings demonstrate that the ISMA optimization model can anticipate the aero optical imaging deviation in an accurate way.

Document code: A **Article ID:** 1673-1905(2023)05-0290-6

DOI https://doi.org/10.1007/s11801-023-2137-7

One of the key elements influencing the development of precision-guided weapons and equipment is aero optical imaging deviation. The current value of the aero optical imaging deviation may be immediately obtained by the airborne computer, which has direct application value for raising the hit rate. In order to quickly estimate the aero optical imaging deviation in actual engineering, this paper uses the data collected for typical operating conditions and an improved optimization algorithm in order to address the issues of high calculation cost, lengthy implementation period, and high cost of wind tunnel experiments and numerical simulations. Height, Mach, and angle of attack are a few variables that impact aero optical imaging deviation. Building a suitable network model to forecast output results for various inputs is the core of utilizing a neural network to anticipate imaging migration. The running environment of high-speed aircraft may therefore be simulated using the trained network[1-3].

Many academics have proposed various methods for anticipating aero optical imaging deviation. The least squares support vector machine for chaotic particle swarms was optimized by XUE et al, using an aero optical flow imaging deviation algorithm model, and the addition of chaotic algorithms increased the variety of the initial sequences of particle swarms[4]. WU et al[5] proposed an improved particle swarm optimization (PSO) algorithm to optimize the extreme learning machine's aero optical imaging deviation model, which increased prediction accuracy by dynamically transforming the inertia weights. YAO[6] proposed an improved aero optical imaging deviation algorithm model to optimize the back propagation (BP) neural network, which increased population diversity by introducing the concept of aggregation. CHEN[7] proposed to improve the atomic search algorithm to optimize the aero optical imaging deviation model of the extreme learning machine, and the Levy flight and gold sine algorithm with adaptive step length were added to improve the convergence speed and exploration accuracy of the algorithm. ZHANG et al[8] proposed aero optical imaging deviation prediction based on improved sparrow search algorithm to optimize BP neural network and added the idea of bird swarm algorithm flight behavior, which ensured global convergence and population diversity.

In this study, the neural network model used to forecast the aero optical imaging deviation is optimized using an improved slime mould algorithm (ISMA). Tent chaotic mapping and a random opposition-based learning are added to the ISMA, which increases exploration accuracy and speeds up convergence. An enhanced approach is then

* This work has been supported by the National Natural Science Foundation of China (Nos.61975151 and 61308120).

** E-mail: babudou_xuewei@163.com

used to refine the weights and thresholds of Elman to produce predictions for aero optical image deviation that are more precise.

Elman is a dynamic recurrent neural network built on the foundation of the BP neural network. To achieve the goal of memory, a successor layer is added to the hidden layer. As a result, the network has the capacity to adjust to the specifics of the event, improving the network's overall stability. The Elman network structure is shown in Fig.1, which is divided into four layers, namely, an input layer, an output layer, a hidden layer, and a recurrent layer. The input, output, and hidden layers share a structure with a feedforward network. The input layer is a signal transmission function, the recurrent layer is used to remember the output value of the hidden layer a moment before, and the activation function of the hidden layer generally selects the Signmoid function. The output of the hidden layer enables dynamic modeling by taking over the layer's auto-join to the hidden layer's input, making it sensitive to historical data. Based on this structure, Elman networks have more computational power and network stability than BP networks and are capable of internal feedback, storing, and using output data from earlier times[9,10].

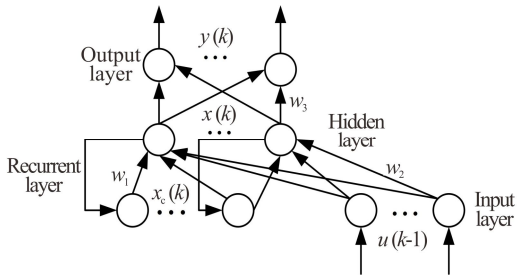

Fig.1 Elman neural network structure diagram

The performance of the network is greatly influenced by the number of layers and nodes in the hidden layer. Too many layers and nodes will make the network more complex, slow down computation, and potentially result in overfitting. According to empirical formula Eq.(1) and the trial approach, the general setting network implicit layer description is determined

$$L = \sqrt{m+n} + a, \tag{1}$$

where L is the number of hidden layer nodes, m is the number of output layer nodes, n is the number of input layer nodes, a is an adjustment parameter, and it takes the range from 1 to 10 for training to find the optimal value.

LI et al[11] proposed the SMA, a new meta-heuristic algorithm, in 2020. The algorithm primarily models the behavior and morphological alterations of slime polyphenols during the foraging process, models the positive and negative feedback processes produced by the mucosal body during foraging, and displays the three correlations between the shape change and contraction mode of the mucosal vein tube, resulting in three distinct foraging forms. Although the SMA outperforms other conventional algorithms in terms of convergence speed and searches precision, it still has the drawback of being unable to escape local extremes later in the iteration.

As the slime approaches the food source, the oscillator produces a propagation wave that increases the cytoplasmic flow through the veins. The faster the flow of cytoplasm, the thicker the veins. This allows when slime bacteria come into contact with multiple food sources, different networks of veins form between multiple food sources, and the venous network is related to the quality of food sources. Model the behavior of slime close to food as a mathematical equation to simulate how mucous bacteria contract.

The simulated shrinkage mode of the slime bacteria is shown as

$$\overrightarrow{X(t+1)} = \begin{cases} \overrightarrow{X_b(t)} + \overrightarrow{v_b} \cdot (\overrightarrow{W} \cdot \overrightarrow{X_A(t)} - \overrightarrow{X_B(t)}), r < p \\ \overrightarrow{v_c} \cdot \overrightarrow{X(t)}, r \geq p \end{cases}, \tag{2}$$

where v_b is a parameter with a range of $[-a, a]$, v_c from 1 linearly decreasing to 0, t is the current number of iterations, X_b is the currently found odor concentration of the highest individual location, is the location of X slime bacteria, and is a random extraction of two X_A individuals X_B from the group, is the weight of W slime bacteria.

The parameter p is calculated according to the most suitable value and the optimal value of the current individual, and the definition of p is defined as

$$p = \tanh |S(i) - DF|, \tag{3}$$

where $i \in 1, 2, ..., n$, $S(i)$ represents the fitness of X, and DF is the best fitness for all iterations.

The parameter a is shown as

$$v_b = [-a, a], \tag{4}$$

$$a = \arctan h\left(-\left(\frac{t}{\max_t}\right)+1\right), \tag{5}$$

where t is the current number of iterations, and $\max_t$ is the maximum number of iterations.

Weight W is shown as

$$W\left(SmellIndex(i)\right) = \begin{cases} 1 + r\log\left(\left(b_F - S(i)\right)/\left(b_F - w_F\right)+1\right), \text{condition} \\ 1 - r\log\left(\left(b_F - S(i)\right)/\left(b_F - w_F\right)+1\right), \text{others} \end{cases}, \tag{6}$$

where condition represents $S(i)$ in the first half of the population and r is a random value within the interval of [0,1], b_F is the optimal fit obtained during the current iteration, and w_F is the worst adaptation obtained during the current iteration.

It simulates the positive and negative feedback relationship between the width of the vibrio vein and the food concentration (adaptation value) in Eq.(6). Based on the above principle, the formula for the location update of the slime is shown as

$$X(t+1)=\begin{cases} rand(UB-LB)+LB\ ,rand \prec z \\ X_b(t)+v_b\left(WX_A(t)-X_B(t)\right), r \prec p \\ v_c X(t), r \geq p \end{cases} \quad (7)$$

where *LB* and *UB* represent the lower and upper limits of the search range, and *rand* are *r* random values within the interval of [0,1]. *z* is a parameter that weighs the search and development phases, and is set to 0.03.

Chaos, as a non-linear phenomenon common in nature, is applied by many scholars to optimize search problem because of its randomness, traversal and regularity[12]. The optimization search algorithm after joining chaos can not only maintain the diversity of population, but also help the algorithm to jump out of local optimization, which improves the global search ability of the algorithm. Ref.[13] adds Tent chaos to the sparrow search algorithm, Ref.[14] adds Tent chaos to the whale optimization algorithm, and Ref.[15] uses Tent chaos to improve the gray wolf optimization algorithm. The improved algorithm has fast convergence speed and strong global search capability. The initialization of the SMA plays an important role in the convergence speed and the accuracy of the search. When the mucous bacteria are initialized, most of the initial positions of the slime bacteria are generated immediately. If the initial population is evenly distributed in the search space, it is of great help to improve the algorithm.

Tent mapping has uniform probability density, power spectral density and ideal correlation characteristics. The mathematical expression is shown as

$$y_{i+1}=\alpha-1-\alpha|y_i|, \alpha \in (1,2]. \quad (8)$$

When $\alpha \leq 1$, the system is in a stable state, when $\alpha>1$, the system is in a state of chaos, and when $\alpha=2$, it is the center Tent map. The mathematical expression is shown as

$$y_{i+1}=\begin{cases} 2y_i, & 0<y_i \leq 0.5 \\ 2(1-y_i), & 0.5<y_i \leq 1 \end{cases}, \quad (9)$$

where *i* denotes the number of maps, and y_i represents the *i* number of function value.

Chaos sequences are generated in a variety of ways, and the uniformity of the sequence is better with Tent mapping. Tent mapping is simple in structure and has good traversal uniformity, but there are small periods in the Tent mapping iteration sequence. This is because the Tent chaos map belongs to the distribution of [0,1], whose function distribution is concentrated between 0.2 and 0.8. The distribution between 0—0.2 and 0.8—1 is poor[16]. Ref.[17] proposed an improved Tent chaos mapping, introducing random variables on the original Tent mapping expression, and improving the traversal of the Tent chaos mapping between 0 and 1 more evenly distributed than the Tent chaos mapping. Therefore, this paper uses the improved Tent mapping generated by the chaotic sequence initialization mucosal optimization algorithm population. The improved Tent chaos mapping function is shown as

$$y_{i+1}=\begin{cases} 2y_i+rand(0,1)\times\dfrac{1}{NT}, 0 \leq y_i \leq 0.5 \\ 2(1-y_i)+rand(0,1)\times\dfrac{1}{NT}, 0.5 \prec y_i \leq 1 \end{cases}, \quad (10)$$

where *i*=0, 1, 2…, *NT* is the number of particles in the chaos sequence, and *rand*(0,1) is the random number between 0 and 1.

Set the population number to *N*, assign *N* initial values of y_i to y_0, and generate chaos variables y_{bi} according to Eq.(11), $i=1,2,...n$. The chaotic variable y_{bi} is inversely mapped to the corresponding individual search space x_{bi} variable.

$$x_i=l_i+(u_i-l_i)y_i, i=0,1,2..., \quad (11)$$

where the search range of x_i is [l_i, u_i].

The opposition-based learning, which generates a reverse solution based on the current solution, is an improved strategy that was put forth by TIZHOOSH et al in 2005 for the field of group intelligence[18]. This strategy increases the search ability of the algorithm and lessens the likelihood that the algorithm will fall into local optimization. The calculation formula for the opposition-based learning is displayed as

$$x_d=UB+LB-x_t, \quad (12)$$

where *UB* and *LB* are the upper and lower bounds of the search space, respectively, and the current solution is x_t.

The inverse solution produced by the opposition-based learning lacks randomness and cannot successfully increase the population diversity in the search space because it is a specific value from the present solution. In order to further increase population diversity and improve populations' capacity to avoid local optimality, LONG et al[19,20] proposed an improved random opposition-based learning, demonstrated as

$$x_d=UB+LB-rand[0,1]\cdot x_t, \quad (13)$$

where x_d represents the random reverse solution, and *rand*[0,1] is a random number in [0,1].

The greedy principle is employed to determine whether to update the slime mold position because there is no assurance that the slime mould fitness value will be higher after a mutation. That is, the slime mould position is only changed when the mutant slime mould has a higher adaptability score. The update location equation is displayed as

$$X_i(t+1)=\begin{cases} Xn_i(t), & f(Xn_i(t)) \leq f(X_i(t)) \\ Xs_i(t), & f(Xn_i(t)) > f(X_i(t)) \end{cases}, \quad (14)$$

where $i \in [1,N]$, $X_i(t+1)$ is the optimal slime mould position after greedy selection, $f(Xn_i(t))$ is the current fitness value, and $f(X_i(t))$ is the old fitness value.

The Elman neural network is optimized by the ISMA in three steps: determining the topology of the network, ISMA optimization, and Elman neural network prediction. In order to determine the weights and thresholds that need to be optimized in order to improve the slime optimization algorithm, Elman structures the network according to the number of input and output parameters.

He then optimizes the network's initial weights and thresholds in order to increase its capacity for global optimization. After finding the global optimal value to use as the initial weights and thresholds of the network, ISMA chooses various equations in accordance with the changes in weights and iteratively updates them continuously. The network then generates sample prediction values after training. The algorithm flowchart of the ISMA optimization of the Elman neural network is shown in Fig.2.

As input variables for the imaging deviation prediction model, the altitude, Mach number, angle of attack, line of sight roll angle, and line of sight inclination of the aircraft are employed. The imaging deviation value is used as the model's output variable. In the experiment, 600 randomly chosen data sets were used as the model training sample set and the remaining data were used as the model test sample set to calculate the aircraft imaging deviation. The empirical formula states that the hidden layer nodes range from 4 to 14 and that experimentation determines the neural network layout. According to the experimental findings, there are 8 nodes in the Elman hidden layer.

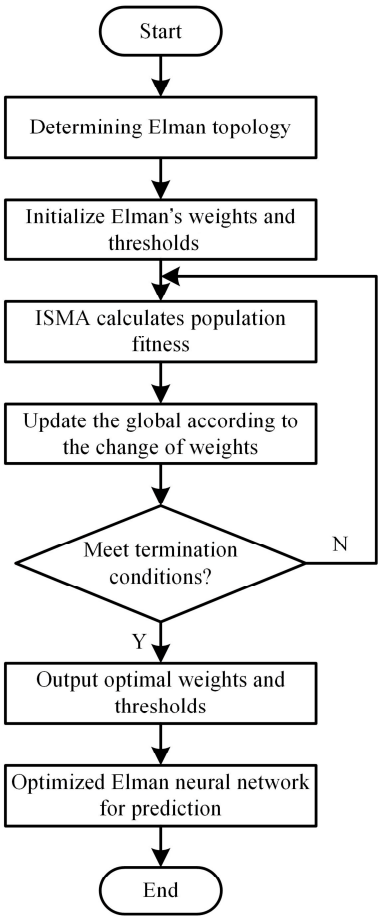

Fig.2 ISMA-Elman flowchart

Based on this, the ISMA is added to optimize the Elman network's initial weights and thresholds. This improves the network's capacity to do global optimization and prevents it from devolving into local optimization.

The mean squared error (*MSE*), coefficient of determination (R^2), mean absolute error (*MAE*), and mean absolute percentage error (*MAPE*) were chosen as the detection criteria in order to assess the benefits and drawbacks of aero optical imaging deviation models. The *MSE* is the mean of the sum of squares of the corresponding point errors of the predicted data and the original data. The closer the mean square error is to 0, the higher the prediction accuracy of the model. The coefficient of determination (R^2) is taken at [0,1], and the closer to 1, the higher the degree of fit. The *MAE* represents the average of the absolute error between the predicted value and the true value, and the closer it is to 0, the more accurate the result. The *MAPE* is taken at $[0,+\infty)$, and the closer it is to 0, the higher the accuracy of the model.

$$MSE=\frac{1}{n}\sum_{i=1}^{n}\left(y_i-p_i\right)^2, \tag{15}$$

$$MAE=\frac{1}{n}\sum_{i=1}^{n}\left|y_i-p_i\right|, \tag{16}$$

$$MAPE=\sum_{i=1}^{n}\left|\frac{y_i-p_i}{p_i}\right|\times\frac{100\%}{n}, \tag{17}$$

$$R^2=\frac{\left(n\sum_{i=1}^{n}y_ip_i-\sum_{i=1}^{n}y_i\sum_{i=1}^{n}p_i\right)^2}{\left(n\sum_{i=1}^{n}y_i^2-\left(\sum_{i=1}^{n}y_i\right)^2\right)\left(n\sum_{i=1}^{n}p_i^2-\left(\sum_{i=1}^{n}p_i\right)^2\right)}, \tag{18}$$

where y_i is the predicted value of the ith sample, p_i is the target value of the ith sample, $i\in(1,2,...,n)$, and n is the number of samples.

The ISMA-Elman model is contrasted with the predictions from the SMA-Elman model and Elman neural network in order to assess the accuracy of the ISMA-Elman prediction model put forth in this study. The ISMA-Elman model and the SMA-Elman model have the same parameter values.

It is clear from the study and comparison of each model in Figs.3—5 and Tab.1 that the ISMA-Elman model's anticipated value and actual value differ by the least amount. The table shows that the ISMA-Elman model has a substantially greater prediction accuracy than the Elman model and the SMA-Elman model. The *MSE*, R^2, *MAE*, and *MAPE* have respective values of $1.322\,7\times10^{-11}$, 0.999 98, $2.564\,3\times10^{-6}$, and 0.711 3. R^2 is closer to 1 compared to the other two models, whereas *MSE*, *MAE*, and *MAPE* are closer to 0.

The fitness curves for SMA-Elman and ISMA-Elman are shown in Fig.6. The SMA-Elman model tends to stabilize after about 40 iterations, while the ISMA-Elman model tends to stabilize after around 15 iterations, as seen in the figure. The findings indicate that the model has faster convergence and a stronger capacity for global search than the ISMA-Elman model.

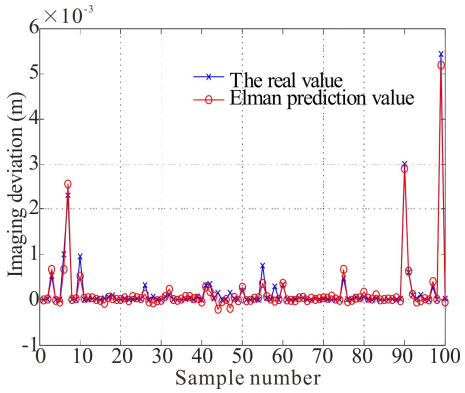

Fig.3 Elman predicted and real values (*MSE*=1.234 2×10^{-8}, R^2=0.973 45, *MAE*=6.343 7×10^{-5})

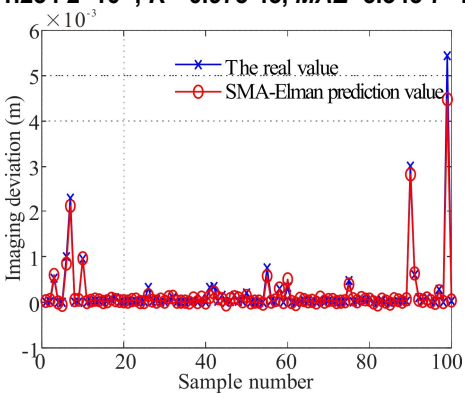

Fig.4 SMA-Elman predicted and real values (*MSE*=1.271 1×10^{-8}, R^2=0.989 59, *MAE*=4.899 4×10^{-5})

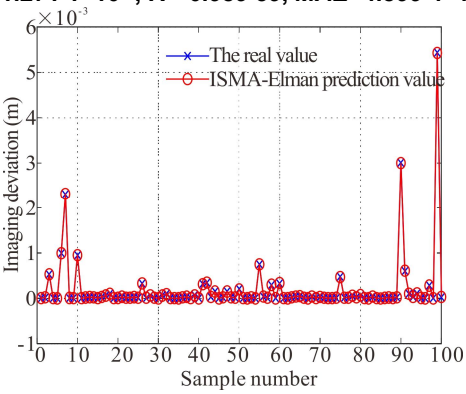

Fig.5 ISMA-Elman predicted and real values (*MSE*=1.322 7×10^{-11}, R^2=0.999 98, *MAE*=2.564 3×10^{-6})

Tab.1 Comparison of evaluation indicators of each model

Model	*MSE*	R^2	*MAE*	*MAPE* (%)
Elman	1.234 2×10^{-8}	0.973 45	6.343 7×10^{-5}	14.035 2
SMA-Elman	1.271 1×10^{-8}	0.989 59	4.899 4×10^{-5}	3.801 0
ISMA-Elman	1.322 7×10^{-11}	0.999 98	2.564 3×10^{-6}	0.711 3

In this paper, the ISMA-Elman model is used to predict the aero optical imaging deviation. The SMA has issues with the optimal problem, such as limited convergence ability in the late iteration and ease of falling into local optimum throughout the optimization process. The

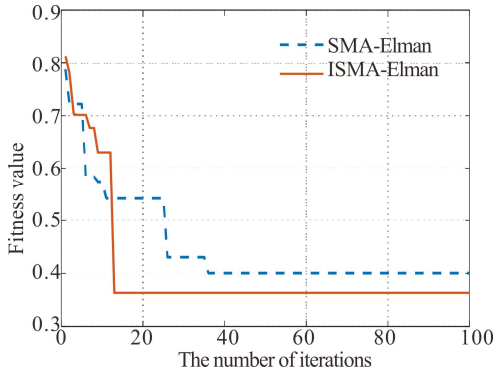

Fig.6 SMA-Elman and ISMA-Elman fitness curves

SMA is enhanced with the use of the modified Tent chaotic sequence to start the population and hasten algorithm convergence. Second, a random opposition-based learning is included to increase the population's variety and capacity to avoid local optima. ISMA is utilized to improve the weights and thresholds of the Elman neural network, which increases algorithm accuracy. The ISMA-Elman model assures population variety and global convergence while outperforming the Elman and SMA-Elman models in terms of prediction accuracy. It offers a reference for the quick engineering calculation of aero optical effects within a certain range and may quickly and correctly forecast the aero optical imaging deviation.

Statements and Declarations

The authors declare that there are no conflicts of interest related to this article.

References

[1] LI G C. Aero optics[M]. Beijing：National Defense Industry Press, 2006. (in Chinese)

[2] YIN X L. Aero optics principle[M]. Beijing：China Aerospace Publishing House, 2003. (in Chinese)

[3] XU L, CAI Y L. High altitude aero-optic imaging deviation prediction for a hypersonic flying vehicle[C]// 2011 IEEE International Conference on Imaging Systems and Techniques, May 17-18, 2011, Batu Ferringhi, Malaysia. New York：IEEE, 2011：210-214.

[4] WU Y, XUE W, XU L, et al. Optimized least-squares support vector machine for predicting aero-optic imaging deviation based on chaotic particle swarm optimization[J]. Optik, 2020, 206：163215.

[5] WU Y, XUE W, XU L, et al. Optimized ELM for predicting aero-optic imaging deviation based on improved PSO[J]. Journal of optoelectronics·laser, 2020, 31(01)：64-70. (in Chinese)

[6] YAO Y. Analysis and prediction of aero optical imaging deviation of typical aircraft[D]. Tianjin：Tianjin University of Technology, 2020. (in Chinese)

[7] CHEN X. Aero optical imaging deviation and prediction for different line of sight roll angles[D]. Tianjin：Tianjin University of Technology, 2021. (in Chinese)

[8] XU L, ZHANG Z Y, CHEN X, et al. Improved sparrow search algorithm based on BP neural networks for aero-optical imaging deviation prediction[J]. Journal of optoelectronics·laser, 2021, 32(06)：653-658. (in Chinese)

[9] ELMAN J L. Finding structure in time[J]. Cognitive science, 1990, 14(2)：179-211.

[10] ZHENG Y, ZHANG X, WANG X, et al. Predictive study of tuberculosis incidence by time series method and Elman neural network in Kashgar, China[J]. BMJ open, 2021, 11(1)：e041040.

[11] LI S, CHEN H, WANG M, et al. Slime mould algorithm：a new method for stochastic optimization[J]. Future generation computer systems, 2020, 111(12)：300-323.

[12] LIU L F, SONG Z D, YU H Y, et al. A modified fuzzy C-means (FCM) clustering algorithm and its application on carbonate fluid identification[J]. Journal of applied geophysics, 2016, 129：28-35.

[13] LV X, MU X D, ZHANG J, et al. Chaotic sparrow search optimization algorithm[J]. Journal of Beihang University, 2021, 47(08)：1712-1720. (in Chinese)

[14] LIN J, HE Q, WANG Q, et al. Optimization algorithm for sine and cosine whale based on chaos[J]. Intelligent computers and applications, 2020, 10(9)：43-48+52.

[15] MAO Q H, YANG L, WANG Y L. Fusion improves Tent chaos and simulated annealing gray wolf algorithm[J]. Practice and understanding of mathematics, 2021, 51(5)：147-161.

[16] YUE L F, YANG R N, ZHANG Y J, et al. Tent chaos and simulated annealing improvement of moth firefighting optimization algorithm[J]. Journal of Harbin Institute of Technology, 2019, 51(5)：146-154. (in Chinese)

[17] ZHANG N, ZHAO Z D, BAO X A, et al. Based on the improved tens chaotic gravitational search algorithm[J]. Control and decision, 2020, 35(4)：893-900. (in Chinese)

[18] TIZHOOSH H R. Opposition-based learning：a new scheme for machine intelligence[C]//Proceedings of International Conference on Computational Intelligence for Modelling, Control and Automation, and Intelligent Agent, Web Technologies and Internet Commerce, November 28-30, 2005, Vienna, Austria. New York：IEEE, 2005：695-701.

[19] LONG W, JIAO J J, LIANG X M, et al. A random opposition-based learning grey wolf optimizer[J]. IEEE access, 2019, 7：113810-113825.

[20] NAIK M K, PANDA R, ABRAHAM A. Adaptive opposition slime mould algorithm[J]. Soft computing, 2021, 25(22)：14297-14313.