International Journal of Industrial Engineering, 30(6), 1629-1642, 2023

SPORTS ECOTOURISM DEMAND PREDICTION USING IMPROVED FRUIT FLY OPTIMIZATION ALGORITHM

Yu Xie[1], Hui Jiang[2], Lei Wang[3], and Chunlin Wang[4*]

[1]College of Physical Education
Chuzhou University
Chuzhou, China

[2]Institute of Physical Education
Dezhou University
Dezhou, China

[3]Department of Physical Education
Tangshan Normal University
Tangshan, China

[4]College of Physical Education
Sichuan Normal University
Chengdu, China
[*]Corresponding Author email: wangchunlin@sicnu.edu.cn

With improvements in the national consumption level, the tourism industry is playing an increasingly important role in the national economy, and the proportion of tourism revenue in GDP is constantly increasing. In this paper, we first improve the standard Drosophila algorithm by adaptively adjusting the fly population number and search step size while optimizing the initial iteration position and improving the local search ability and search efficiency. Then, the improved algorithm (FOA) is combined with the echo state network to establish a two-stage combined prediction model called the Adaptive Fruit Fly Optimization Algorithm-Echo State Network (AFOA-ESN). The experimental results show that the AFOA-ESN model has higher prediction accuracy compared to other prediction models, and the convergence rate and prediction accuracy of the AFOA-ESN are better than the standard ESN and FOA-ESN, proving the effectiveness of model improvement.

Keywords: Sports; Eco-Tourism; Fruit Fly Optimization Algorithm (FOA); Adaptive Fruit Fly Optimization Algorithm-Echo State Network (AFOA-ESN).

(Received on January 5th, 2023; Accepted on November 14, 2023)

1. INTRODUCTION

Since the beginning of the 21st century, tourism has ushered in a golden age of rapid development. In 2017, China's total tourism revenue reached 5.4 trillion yuan, an increase of 15.14% from 4.69 trillion yuan in 2016, which is much higher than the GDP growth rate. In the past decade, tourism revenue in GDP rose from 4.05% in 2007 to 6.53% in 2017. With the growth of tourism income, the sports ecological tourism industry has also undergone qualitative changes. According to relevant data, China's tourism industry is gradually shifting from the second stage to the third stage, and the future market prospects are very broad. National policy support and the endogenous growth of tourism supply and demand have jointly created the prosperity and development of tourism for more than ten years. The World Tourism Council's research report predicts that by 2027, the number of Chinese "new tourist households" (those with new incomes of more than USD 35,000) will exceed 64 million, which is far above second-ranked India (9.4 million new tourist households predicted by 2028). The direct contribution of China's tourism industry to GDP will jump to about USD 1.3 trillion (based on 2017 fixed price estimates). Beyond the current number-one country, i.e., the US (USD 509.4 billion), China will have a 10-year compound growth rate of 6.6% and a global ranking of fifth place. In the next 10 years, China's tourism industry will create about 34 million jobs directly, which is far higher than second-ranked India (9.38 million), and this will help to ensure domestic employment. Information on the current and future levels of tourism demand and its contribution to the economy is very important due to the business institutions and decision-making departments of the government. For countries or regions in the Caribbean Sea,

DOI: 10.23055/ijietap.2023.30.6.8841 ISSN 1943-670X

tourism market prosperity is the most critical measure by which to predict the macro economy of the country or region (Han *et al*., 2013). Tourism products are perishable, and accuracy is critical to the problem of tourism demand prediction.

To date, many studies have been conducted on measurement models, decision vector machine models, gray theoretical models, and artificial neural network models, but in the face of nonlinear prediction problems, the prediction effect is different. Zhang and Chen (2013) proposed echo state networks (ESNs); compared with traditional algorithms, the model construction mode for these networks is different. The unique structure of this method opens up new research ideas and directions. In the echo state neural network reserve pool and traditional neural network algorithm, the connection weight in the network initialization reserve pool is composed of the input weight and output weight. In the process of training, the input weight and reserve pool weight values do not change. Only the size of the reserve pool output weight changes, and this effectively avoids complex, time-consuming training of the artificial neural network. Echo-state neural networks can deal with nonlinear problems well, and the training process is simple and efficient. In nonlinear prediction problems, an echo state neural network has many advantages compared with traditional artificial neural networks and other types of algorithms and has great application prospects. However, in the face of different types of problems and data, the traditional ESN model may have strong random and non-optimal initial parameter settings, and the output results are uncertain. Therefore, this method can further enhance the prediction effect of ESN through some optimization methods. The selection of parameters is a major consideration. Current research mainly combines intelligent algorithms to obtain optimized parameter inputs so as to help build suitable prediction models. However, it is difficult to find the appropriate objective function for some parameters, such as internal weights, so applying combinatorial optimization techniques is one way to alleviate the disadvantages brought about by parameter randomization.

The fruit fly optimization algorithm (FOA) is a new heuristic global search optimization algorithm that was proposed in recent years. It has the advantages of a simple calculation process, strong global convergence, short execution time, self-organization and adaptability, and strong portability. Increasing numbers of scholars have applied the intelligent algorithm to the traditional deterministic optimization problem. Most of the current research results were analyzed using an ESN or the FOA algorithm alone, but in this paper, we study the combined prediction model combining an ESN and the FOA to find a model that can improve the accuracy and reliability of prediction. In addition, with the rise in tourism demand, the academic community's demand for regional tourism is also gradually becoming high, but for the study of sports, ecological tourism demand is still relatively low. Combined with the above content, we select sports ecological tourism demand as a research direction based on the construction of the AFOA-ESN model to predict the direction of tourism demand. The proposed AFOA algorithm effectively improves the convergence speed, the prediction accuracy, and the initial position optimization ability. We also apply the combined FOA-ESN prediction model and an FOA-ESN-based model to provide auxiliary support for tourism industry demand prediction and operational-related decisions. The study refers to the development of sports ecotourism.

2. RELATED WORK

2.1 Research on the Demand Prediction of Sports Ecotourism

Zhang *et al*. (2015) took the number of visitors from the top ten source countries to Hong Kong as the dataset. According to four prediction models, ARIMA (Autoregressive Integrated Moving Average Model), ALDM (Alternating Direction Method of Multipliers), ECM (Engine Control Module), and VAR (Variation), several different predictive models were constructed, and after testing the prediction effect of the above methods, it was found that the combination of models can improve the effect of tourism prediction. Wang *et al*. (2016) combined linear and nonlinear statistical models to make predictions for time series with possible nonlinear features, testing the prediction accuracy of the combined model by using time series datasets of the outbound travel demand in Taiwan and comparing a single prediction model with a model combining the two methods. It was found that several combined predictive models had more accuracy than a single prediction model and could identify the turning point of the change in the tourism environment. Zhang *et al.* (2014) studied the prediction effect of four methods: the time series model, the metrology model, the gravitational model, and the expert decision system. They found that the simplest and lowest-cost time series models are suitable for practitioners. The gravitation model is suitable for solving international tourism problems. The expert decision system is suitable for scenarios where some data are not available. Wu and Zhou (2017) used seven quantitative prediction methods, sampling predictions of the tourist flow at 24 starting sites; 1 year and 2 years were selected for prediction. The results showed that the l-year forecast period was better than the 2-year period. In the process of data collection and the research itself, Song *et al*. (2020) studied the relationship between prediction accuracy, data characteristics, and research characteristics, and Wang *et al*. (2021) made a comprehensive analysis of nearly 2000 published tourism prediction studies. This indicated that data characteristics and prediction accuracy vary in different types of prediction methods, so selecting the appropriate prediction model according to the data characteristics can effectively reduce the prediction cost.

To date, many scholars have studied the driving factors of the change in sports ecotourism demand. Huang *et al*. (2005) used a measurement model to predict Hong Kong tourism data from 2001 to 2008. The factors that had the largest impact on the number of tourists in Hong Kong were tourism costs, economic conditions (measured using the income of tourist origin), the price of competing products, and the reputation of tourists. These factors can be used as input variables to provide important information for the prediction of tourists in Hong Kong. Chen *et al*. (2019) conducted a systematic analysis of tourism needs and studied a relatively general algorithm to explore the influence of qualitative non-economic factors on tourism, including leisure time factors and climate factors. The results showed that leisure time and climate factors have more influence on tourists than economic factors. This proves the importance of qualitative non-economic factors in tourism motivation theory and demand analysis. Niu *et al*. (2013) incorporated business sentiment indicators into their model. The experimental results showed that business sentiment indicators improved the accuracy of fit and prediction. Expanding business sentiment surveys to tourism could create substantial benefits. Wang and Zou (2013) introduced consumer expectations in time series models. Four models were tested using autoregression (AR), autoregressive integrated moving average (ARIMA), self-excitation threshold autoregression, and the Markov transfer matrix.

To sum up, sports ecotourism products are perishable, and their accuracy is critical to tourism demand prediction. The mismatch between tourism resources and demand will cause huge losses in social resources and capital and affect social and economic development. The prediction model construction method, data selection, and algorithm selection will have a great impact on tourism demand prediction. Therefore, prediction research in respect of sports ecotourism is of some significance.

2.2 Research on Fruit Fly Optimization Algorithm (FOA)

Wang and Zou (2013) presented a new population intelligence algorithm with global search capability and fast convergence properties. This algorithm is relatively simple, has a fast calculation speed, fast solution speed, and high accuracy, and is currently used in many combinatorial optimization and continuous optimization fields. The most widely used application of the fruit fly optimization algorithm (FOA) in continuous optimization is parameter selection, such as the structural parameters of ANNs (Wang and Zou, 2012), penalty parameter of SVR, and conversion coefficient (Li *et al*., 2022; Xi and Han, 2014). The FOA shows strong search power in the above studies, so we consider the selection of the predicted input variables using the FOA.

The fruit fly optimization algorithm (FOA) itself has some shortcomings; for example, it is not suitable for solving problems with negative arguments, and it is not sufficient for solving complex stability. Additionally, the convergence precision of the algorithm is not precise, and it is especially prone to causing local extrema. Therefore, improvement of the FOA has high application value and is a research hotspot. Zhang *et al*. (2015) proposed the first-in-first-out (FIFO) improved algorithm and selected a test function to test the FIFO improvement effect. The results show that the optimization algorithm for the forecasting of FIFO is superior to the traditional flies. Nazir *et al*. (2021) introduced a cloud model in the FOA. The constructed algorithm, CM-FOA, improved the flies. The cloud model can track the fly's optimization iteration process and random movement. Dehghani *et al*. (2020) proposed a new FOA in order to improve the algorithm’s global optimization ability. He *et al*. (2019) proposed an adaptive fly optimization algorithm based on the population and the current optimum using probability replication. To continue the optimization, Xie *et al*. (2020), by increasing the inertia, changed the diminishing nonlinear characteristics and the relationship between the individual and the group to build an improved fly optimization algorithm. He *et al*. (2018) proposed a fly optimization algorithm using multiple groups, a new method of community monitoring. The algorithm only has a few parameters, and the calculation process is simple. Zhang *et al*. (2018) balanced the population diversity and stability of the parameters, creating a level probability strategy and new mutations. The LP-flies optimization algorithm was proposed. The algorithm achieved the ideal result in the optimization of the continuous function and processing joint supply problems. Luo *et al*. (2019) used two fly populations to put forward the DDSC-flies optimization algorithm. By updating the overall optimum and restructuring subgroups to exchange information, two subgroups improved the population dynamic search and avoided slow convergence due to random positions. The search accuracy and speed are superior to the traditional FOA.

3. RESEARCH ON SPORTS ECO-TOURISM DEMAND FORECASTING BASED ON IMPROVED FOA ALGORITHM

3.1 Standard FOA algorithm

The fruit fly optimization algorithm was developed by Pan (2011). The movement of the bionic simulation fly model was added to the FOA algorithm in order to find the optimal solution; the goal was a relatively new heuristic algorithm. The FOA has strong global search ability and fast convergence performance. In addition, because the algorithm is simple in structure, it involves fewer variable parameters, so the operation quantity is lower than most other heuristic algorithms. The experimental results show that the FOA is successful in the reserve pool echo state network parameter optimization of operable functions and optimization ability.

The FOA mainly consists of five steps: population initialization, random flight, flavor concentration determination, flavor concentration, and position labeling. The logical structure of the FOA is shown in Figure 1.

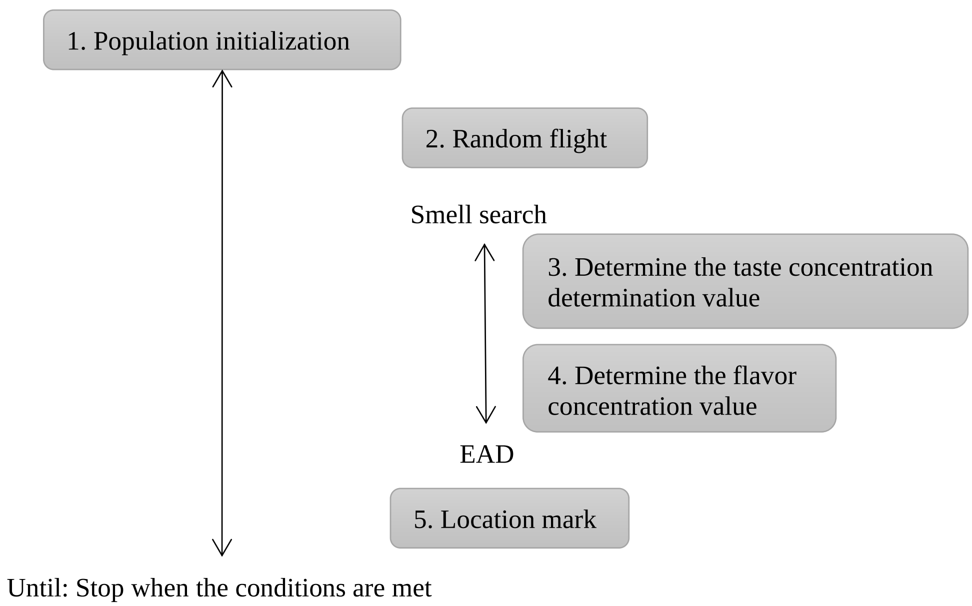

Figure 1. Logical structure of the FOA algorithm.

Step 1: Population initialization. The population initialization stage needs to define the basic characteristics of the flies. First, the size of the group (Sizepop) determines the number of random positions at each iteration, the maximum number of iterations (Maxgen), and the total number of flies through olfactory foraging. X_{axis} and Y_{axis} represent the initial position (the current position represents the moment during iteration); the search step ($Slegth$) represents the maximum distance each fly can fly during each iteration.

Step 2: Random flight. Each fly within the population is given a random flight direction, and the location information variables X_i and Y_i are recorded. The fly population will search for an odor within the search step range. The location update formula is as follows:

$$X_i = X_{axis} + S_{length} \tag{1}$$
$$Y_i = Y_{axis} + S_{length} \tag{2}$$

Step 3: Calculate the flavor concentration determination value (S_i). The taste concentration determination value of the Drosophila population is used because the abscissa and ordinate of the Drosophila position in some functions cannot express the Drosophila position, so the reciprocal value of the distance function is used as the obtained concentration determination value S_i. S_i is not the final flavor concentration, and the parameter needs to be substituted into the function to be optimized. The flavor concentration determination value is not counted in this algorithm.

$$Dist_i = \sqrt{Xi^2 + Yi^2} \tag{3}$$
$$S_i = 1/Dist_i \tag{4}$$

Step 4: Determine the flavor concentration value ($Smell_i$). The flavor concentration values of the Drosophila population behave mathematically as the dependent variable size of the function to be optimized and represent the prediction error size of the neural network optimization in this study. According to the flavor concentration determination value of the fruit fly population, the food flavor concentration is obtained using formulas 3-5, usually assuming that the higher the flavor concentration, the closer the fruit fly is to the food.

$$Smell_i = f(S_i) \tag{5}$$

Step 5: Position tag. In a fruit fly population, before flight to the local optimal solution, every individual fly in a population is noted according to its position to calculate the corresponding density. The taste level at the location of the maximum will be selected as the group of fruit flies together fly to that position; namely, the starting position of the next iteration. Further, if the optimal value of the current iteration ($bestSmell$) is better than before the global optimal ($bestSmell\_global$), the current iteration is used to update the $bestSmell$ value of $bestSmell\_glibal$ for the global optimal.

$$[bestSmell \quad bestindex] = max(Smell_i) \tag{6}$$
$$bestSmell\_\_global = bestSmell \tag{7}$$
$$X_{axis} = X(bestindex) \tag{8}$$
$$Y_{axis} = Y(bestindex) \tag{9}$$

The iterative process of the FOA includes steps 2 through 5 of the above steps if the stop condition is implemented (i.e., the number of iterations accumulates to the maximum iteration number), and the current $bestSmell\_glibal$ is extracted as the optimal solution found in the whole Drosophila search process.

3.2 ESN Echo State Network Prediction Model

Echo state networks (ESNs) have a reservoir pool that differs from traditional neural network algorithms. When the network is initialized, the connection weights are randomly generated. At the beginning of an ESN startup, a reservoir is generated to serve as a base, and then a hidden space with dynamic and complex properties is formed. The connection weights between neurons within the reservoir are not adjusted with the training process. The reserve pool is composed of input weight and output weight. In the process of training, we keep the status of the input weight and reserved pool weight value and change the size of the reserved pool output weight, usually according to the ESN actual output value and target value error. The least-squares method is used to adjust the weight, effectively avoiding the artificial neural network's complex, lengthy training.

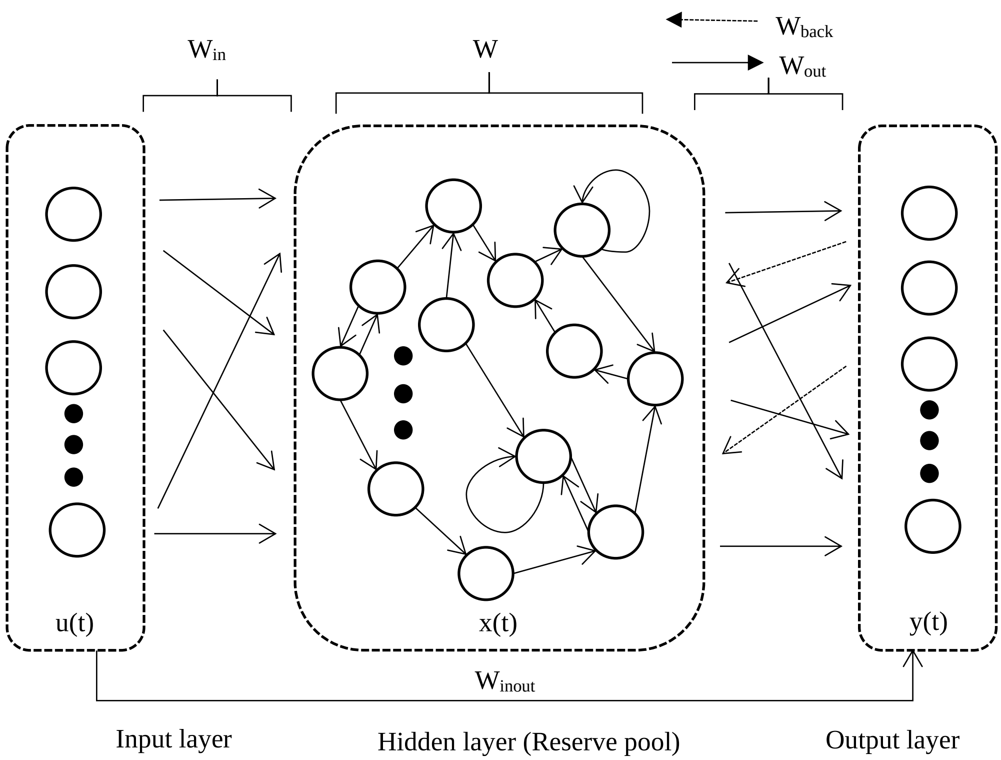

Figure 2. ESN network structure diagram.

A typical ESN network contains three levels where the input layer is used to convert the input information into the initial activation signal, the reserve pool further converts the input layer activation information into the output layer information because of its complex internal connection structure, and the reserve pool has the non-linear dynamics of the recurrent neural network. The network structure is shown in Figure 2, where a solid line is a necessary connection, and a dotted line is an optional connection.

The input layer of the ESN has K nodes, the value of K depends on the dimension of the input information, the univariate prediction model K is equal to l, and the multivariate prediction model K is greater than l. There are N nodes in the hidden layer (i.e., the reserve pool), representing the number of neurons in the reserve pool. The number of nodes in the output layer is L, depending on the dimension of the output information.

The connection weights of the input layer, reserve pool, and output layer are represented by W_{in}, W, and W_{out}, respectively. If there is no feedback structure, the network weights do not contain W_{back}. Several weight matrices usually have different scales, depending on the input, output information format and the internal connection mode. The weight to be adjusted during training is the weight W_{out} between the reserve pool and the output layer nodes.

Recursive neural networks are characterized by short-term memory and function to process nonlinear systems, and the activation function of neurons can be selected according to the actual requirements. The information at the current moment is input from the input layer to the reserve pool. Combined with the feedback information of the state at the last moment and the output layer in the reserve pool, the input signal of the neuron is formed together through a certain weight, and then the input signal is converted into a new round of state vector $x(t+1)$ through the activation function. The updated formula is as follows:

$$x(t+1) = f\left(W_{in} \cdot u(t+1) + W \cdot x(t) + W_{back} \cdot y(t)\right) \tag{10}$$

f represents the activation function that constitutes the state vector $x(t+1)$. $x(t)$ represents the state of the reserve pool in the previous step, and the state vector $x(0)$ at time 0 can be randomly generated. $u(t)$ represents the input vector at step t, the length of which depends on the dimension of the input information. $y(t)$ represents the output vector in step t, and $y(t+1)$ represents the next step, as determined by the next input vector $u(t+1)$, the next state vector $x(t+1)$, and the output vector $y(t)$ in the current step. The specific formulas are as follows:

$$u(t) = [u_1(t), u_2(t), \dots u_k(t)]^T \tag{11}$$
$$x(t) = [x_1(t), x_2(t), \dots x_k(t)]^T \tag{12}$$
$$y(t) = [y_1(t), y_2(t), \dots y_k(t)]^T \tag{13}$$
$$y(t+1) = f_{out}\{W_{out} \cdot (u(t+1), x(t+1), y(t))\} \tag{14}$$

3.3 Improved Fruit Fly Optimization Algorithm to Optimize the Combinatorial Prediction Model of ESN

3.3.1 Improvement Ideas

Due to the key factors and structure setting, the standard FOA has four characteristics: the number of individuals in the fly population is determined during the initial stage; the range of activity of each fly (maximum random step) is not fixed; the initial position of the fly population is randomly determined and unique, and the initial position of the next iteration is determined by the location of the current optimal solution. These four features lead to the following problems in the implementation of the standard FOA: (l) The individual number of the Drosophila population is fixed. In the process of optimization, the Drosophila population may still consume more computing power despite the poor overall performance of local solutions, thus affecting the operation efficiency of the algorithm. Alternatively, where the local solution performs better, the Drosophila population may miss the optimal solution due to an insufficient number of optimal individuals. (2) The maximum random step size of Drosophila may cause an excessive number of iterations in the locations where the local solution performance is poor, and the local solution cannot be fully mined in the local optima due to the large step size range of the local solution, which affects the efficiency and optimization ability of the algorithm. (3) The initial position of the Drosophila population is random and unique, resulting in a large mass fluctuation of the initial solution before the iteration starts, which will affect the effect of the subsequent iteration calculation.

In order to avoid several problems encountered by the standard FOA, the AFOA algorithm is proposed. The specific improvement ideas are as follows:

(1) Initial positions are optimized to introduce informative flies. Similar to the fruit fly population optimization process, a random position search before the start of the iteration will return the information from the random position solution to the Drosophila population, which selects the optimal location as the initial position for entering the iteration. To expand the initial

position search ability of the information flies, the AFOA algorithm gives them a larger random search step size (which can be set to a global search), after which the fruit fly population conducts a local search.

(2) Introducing a variable maximum search step size strategy. The AFOA algorithm introduces the scaling factor, which is the ratio of the current value to the target value of the solution (the scaling factor is the ratio of the $mape$ to the target $mape$ of the current iteration), and the value of the scaling factor multiplied by the standard search step is the maximum search step of the next iteration. $StepLength$ is the standard search step; $StepLength(t+1)$ is the maximum search step for the next iteration. The advantage of a random step size is that the current local optimal solution can increase when the solution is poor and quickly jump out of this range. When the current local optimal solution is good, the step size can be reduced to increase the local search ability at this position.

$$StepLength(t+1) = \frac{mape(t)}{mape\_t\,arg\,e\,t} * StepLength\_\_S\,tan\,d\,ard \tag{15}$$

(3) Introduction of a variable Drosophila population size strategy. This strategy can be simply understood as follows: when the population of the current position is stimulated by a higher odor concentration, the number of flies searching for food increases; the number of flies in the population decreases due to a low odor concentration. This strategy can improve the search power of the Drosophila population at the superior local locations of the current solution while reducing the time-consuming excess at locations near the non-optimal solution. Similar to the variable search step strategy, the variable population size strategy introduces a scaling factor, which is the ratio of the target value of the solution to the current value (the scaling factor is the ratio of the target $mape$ to the current iteration), and the scaling factor multiplied by the standard population size is the size of the Drosophila population in the next iteration.

$$Sizepop(t+1) = \frac{mape\_t\,arg\,e\,t}{mape(t)} * Sizepop\_\_S\,tan\,d\,ard \tag{16}$$

Termination condition judgment. When the maximum number of iterations of the initial setting is not reached ($Maxgen$), the iteration count t is increased by 1. When the iteration stops, the global optimal solution ($Smell\ best$) and the optimal position ($X\_best\ ,\ Y\_best$) are output.

3.3.2 The AFOA-ESN Combined Prediction Model

In this study, the improved fruit fly optimization algorithm was combined with the ESN network to construct a new combinatorial prediction model, i.e., AFOA-ESN (Figure 3). Compared with traditional ESN, the AFOA-ESN model uses the global search capability of the FOA for parameter optimization, so it does not need to manually debug ESN parameters to quickly construct the prediction model applicable to the current problem. Compared with the FOA-ESN model, the AFOA-ESN model can automatically adjust the population size and the maximum random step size according to the current optimization state, and optimize the local position search and local position jumping out ability. The computational efficiency and optimization effect are also improved. Further, the AFOA-ESN algorithm introduces informative flies to optimize the initial position and improve the convergence ability of the algorithm.

The AFOA-ESN is a two-stage combinatorial predictive model, where the first stage optimizes the ESN parameters with the AFOA algorithm, and the ESN network is trained on the test set to determine its key parameter values. The second stage is the prediction stage, i.e., the offline ESN with optimized parameters:

(1) Data preprocessing. Data are input into the combined prediction system. In order to prevent the size of the input data from affecting the network, data processing is usually performed, and the preprocessed data are used as the input network of the model for prediction. X' represents the normalized data, X denotes the raw data; X_{min} is the minimum of the raw data and X_{man} is the maximum of the raw data. The common normalized data processing methods have the following formula:

$$X' = \frac{X - X_{min}}{X_{man} - X_{\min}} \tag{17}$$

(2) Parameter initialization: the parameters of the AFOA algorithm and ESN network need to be initialized, and the parameters of the AFOA algorithm include standard population size, standard maximum random step size, and maximum iteration number; its data information is determined by flight. The parameters include spectral radius size, input node, output node, sparsity, etc.

(3) ESN parameter optimization: There are two layers of the AFOA cycle structure. The first cycle structure is the random position of the Drosophila population, and the second is the cycle structure of the iterative count. When the fly's

position is random, the smell in that direction is strong, and the stronger the smell, the more likely the next flight target is in that direction. After a series of screening flights, the final parameters and processes are determined.

(4) ESN network prediction: This link needs to refer to the predicted error value, judge the optimization effect of AFOA, and finally find the best network parameters and predicted values.

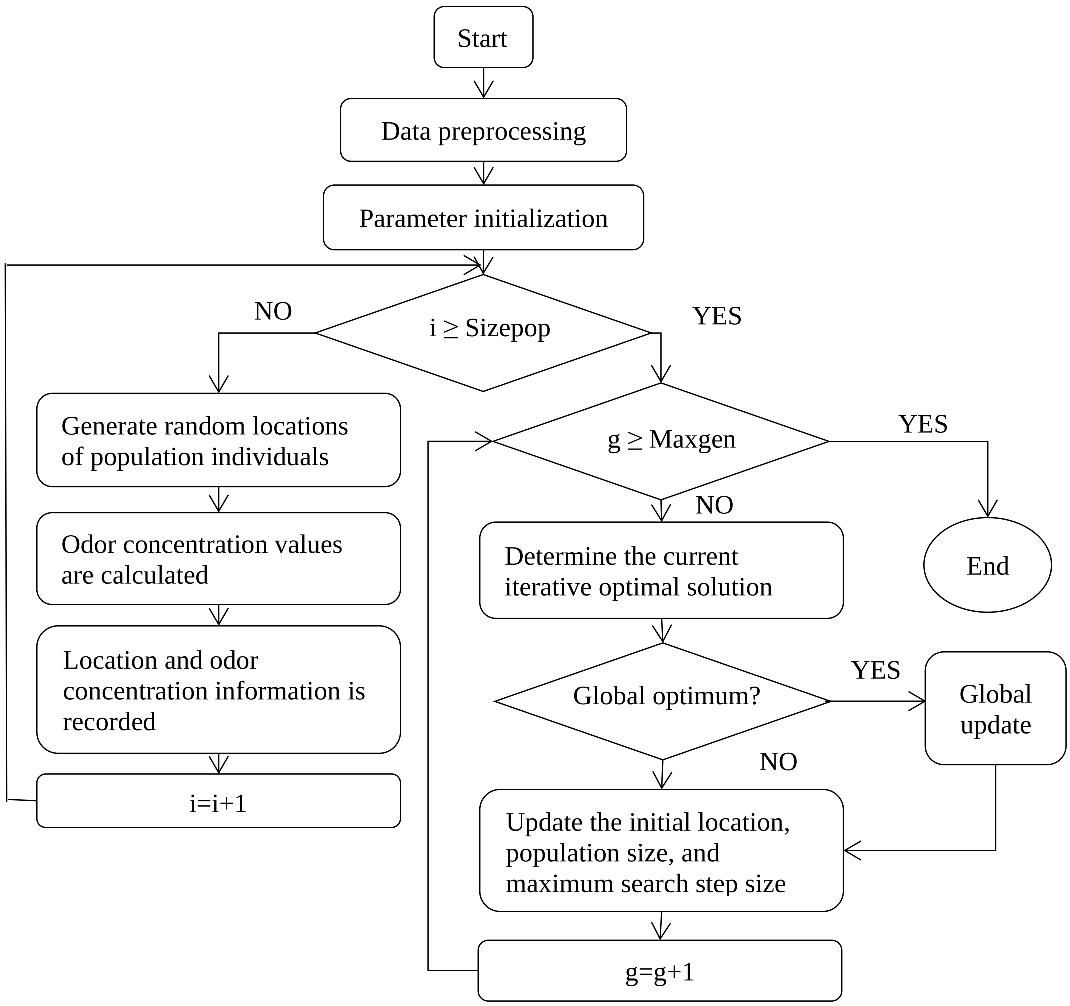

Figure 3. AFOA-ESN combined prediction model execution process.

4. EXPERIMENTAL ANALYSIS

4.1 Data Analysis

In this paper, the features and principles of the FOA and ESN neural network are reviewed, we put forward the AFOA-ESN combination prediction model, and we observe the prediction effect of the FOA-ESN and AFOA-ESN through experiments.

The experimental data are the monthly data in respect of the number of tourists in a certain place, which is the sum of the number of domestic tourists and the number of foreign inbound tourists; the number of tourists in a certain area is shown in Figure 4. In order to ensure consistency with the original test conditions, all data items use the Log10 (N) function, and the results are shown in Figure 5. The data collection period is from 2011 to April 2022, showing strong cyclical characteristics in form. These data are used to verify the ability of the FOA-ESN model to improve the ESN network and the prediction efficiency and prediction accuracy of the AFOA-ESN model compared with the FOA-ESN model. We also compare the model with the results of Sun, ANN, SVR, LSSVR, KELM-lin, KELM-poly, KELM-rbf, KELM-wav, and various algorithmic models combined with the search index (Fan *et al*., 2020). To agree with the original test conditions, all data were preprocessed using the Log10 (N) function.

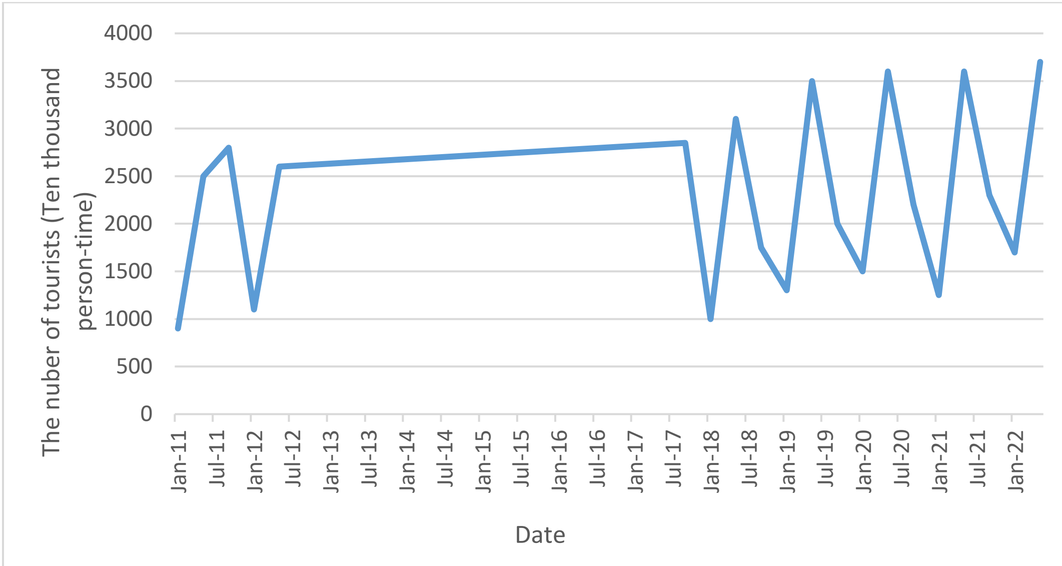

Figure 4. Number of tourists in a certain place (10,000 people).

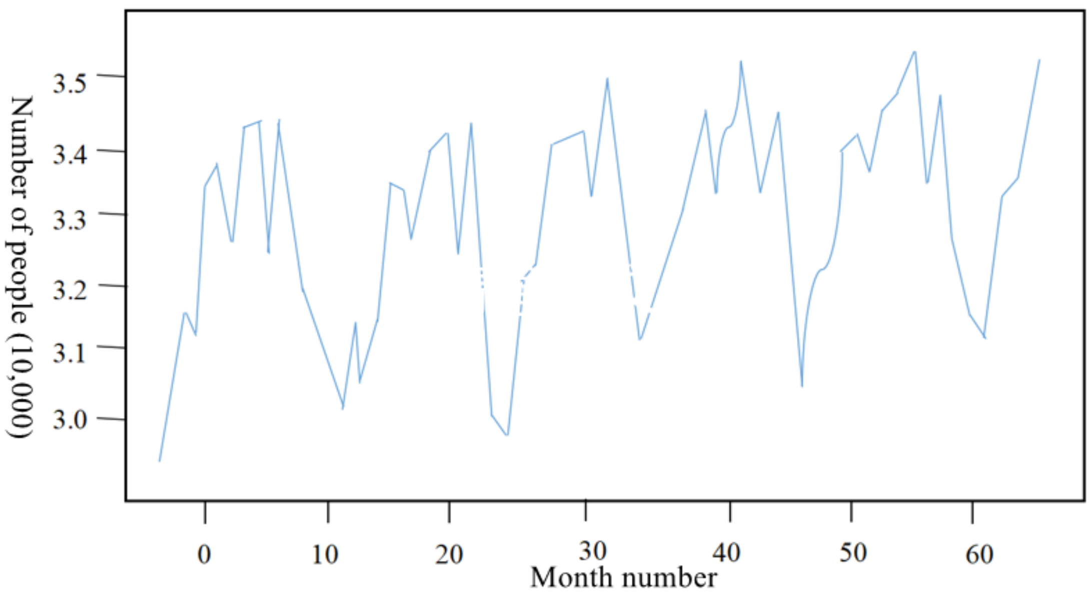

Figure 5. Number of visitors to a local area (Log10 (N)) (log of the raw data). The x-axis represents the month number; the y-axis represents the number of people (10,000).

In terms of parameters, the input data length of the ESN network is 36 months, so there are 36 nodes at the input layer. Each ESN network outputs only one data point, so the output node is set to 1. The FOA-ESN and AFOA-ESN combined prediction model is used to determine the size of the reserve pool and the size of the spectrum radius. In the construction stage of the neural network, only two parameters are assigned a certain range, and the specific size is generated by random numbers. Specifically, the size of the reserve pool is 50 to 500, and the spectral radius is 0.01 to 1.00. Other parameters of the ESN network are similar to those of experiment I. The activation function uses the tanh function, and the activation state memory parameter a is set to 0.20. During network training and testing, other parameters of the ESN remain unchanged except the reserve pool size and spectral radius. The maximum step length of the FOA algorithm is determined in the iteration process, where the single-step variation range of the reserve pool size is [-50,50], the spectrum radius variation range is [-0.25,0.25], the Drosophila population size is 10, and the maximum number of iterations is 100 (the population size and search step length of the AFOA algorithm are adaptive adjustments).

4.2 Error Evaluation Indexes

Three sets of experiments tested the application effects of the ESN, FOA-ESN, and AFOA-ESN models in tourism demand prediction from different dimensions. Typically, experiments require evaluating the prediction results combined with several different metrics of error in order to analyze and compare the effects of various prediction methods from multiple perspectives. The application of various error evaluation indexes can improve the comprehensiveness and objectivity of the prediction model effect measurement.

The mean absolute error (MAE) measures the average of all predicted values $\hat{y}_t$ versus the actual absolute error y_t, which is enough to reflect the difference between the predicted value and the real value. The mean squared error (MSE) is the sum of squares of the deviation between the predicted value and the true value divided using the number of prediction times, which will amplify the large error term and can be used to analyze the stability of the prediction time series error, with a wide range of applications. Root mean squared error ($RMSE$) is the square root on the basis of the mean squared error (MSE), which is also an error statistics method considering the absolute value scale. The mean absolute percentage error ($MAPE$) takes into account the proportion between errors and true values and can be used to compare errors between data of different sizes. Usually, ($MAPE$) values below l0% indicate excellent results.

$$MAE = \frac{\sum_{t=1}^{n} \left| \hat{y}_t - y_t \right|}{n} \tag{18}$$

$$MSE = \frac{\sum_{t=1}^{n} \left(\hat{y}_t - y_t \right) \wedge 2}{n} \tag{19}$$

$$RMSE = \left(\frac{\sum_{t=1}^{n} \left(\hat{y}_t - y_t \right)^2}{n} \right) \wedge 0.5 \tag{20}$$

$$MAPE = \frac{\sum_{t=1}^{n} \frac{\left| \hat{y}_t - y_t \right|}{y_t}}{n} \tag{21}$$

4.3 Predicting the Number of Visitors in an Area Based on the AFOA-ESN Model

The hardware and programming environment of this group of experiments were the same as those in Section 2. The method was the one-step prediction method, predicting the 37th data point with 1 to 36 data points. Finally, the 12 predicted values in the test set were compared with the real values and analyzed using two error indicators, MAPE and NMSE.

The following experimental results show that the FOA-ESN has good prediction accuracy and optimization effects, with a MAPE of 0.53% and NMSE of 0.67%, and both indexes are optimal. At the same time, the FOA-ESN model is very robust, and the prediction error of each data point is controlled within 1.59% (MAPE). The FOA-ESN can make the error gradient drop rapidly, which proves that this method can effectively solve the problem of sports ecotourism demand prediction.

As shown in Figure 6, the AFOA-ESN further improved the prediction accuracy compared to the FOA-ESN model, with a MAPE control within 0.41% and an NMSE of 0.66%. Further, due to the existence of information flies and adaptive mechanisms in the AFOA algorithm, the algorithm can obtain better initial positions than the FOA algorithm, with faster local search ability and iteration speed, and show faster convergence ability. The MAPE of the initial iteration of the AFOA-ESN model is around 0.5%, which is about 0.06% lower than that of the FOA-ESN model. In terms of convergence speed, at the tenth iteration of the AFOA-ESN model, the MAPE drops below 0.44% and below 0.42% after forty generations, and the error-index is lower than that of the FOA-ESN model. The improvement in the AFOA-ESN model shows the rationality and applicability of the optimization algorithm. The parameter settings are shown in Table 1.

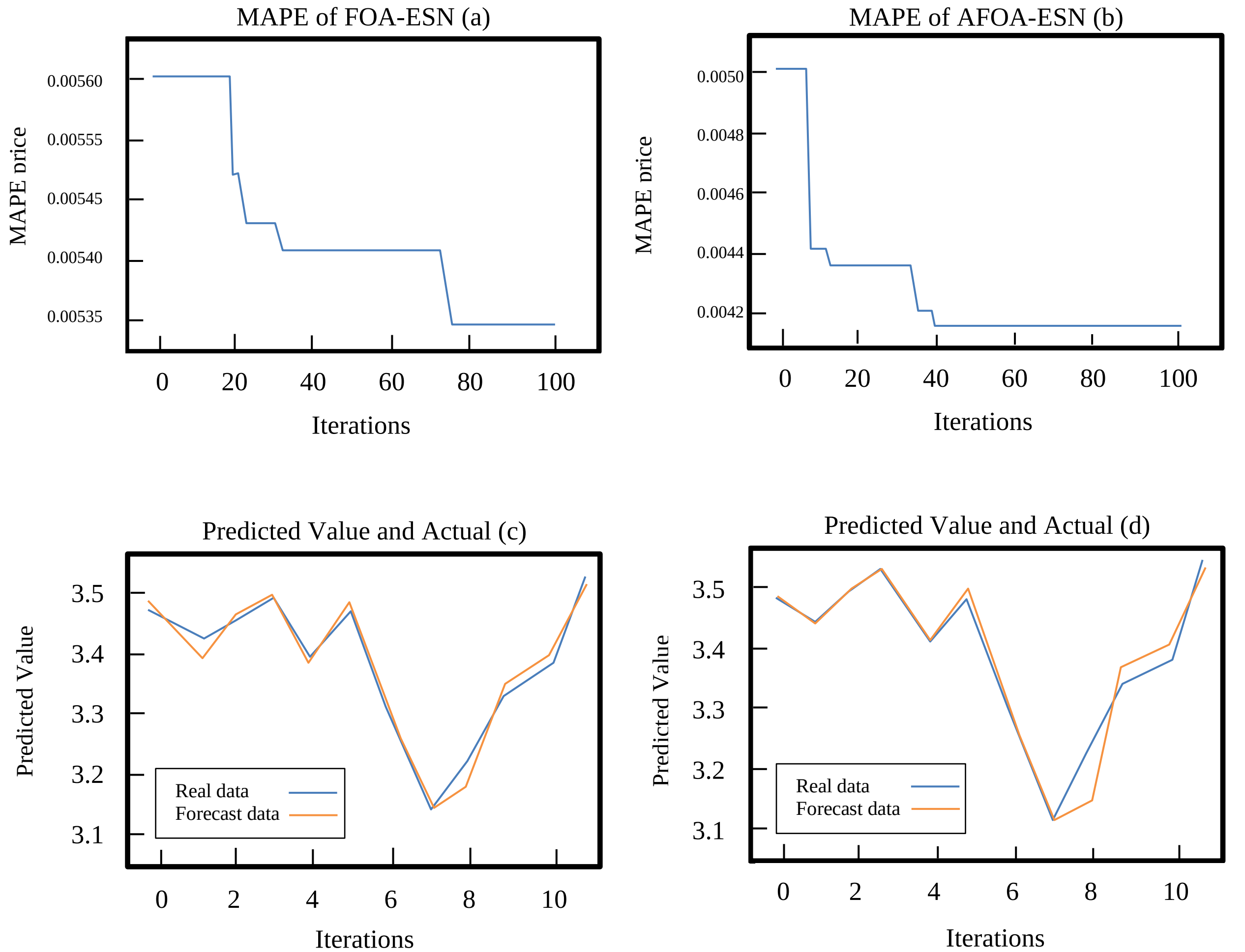

Figure 6. FOA-ESN error gradient chart (a); AFOA-ESN error gradient chart (b); FOA-ESN predicted number of tourists in Beijing (c); AFOA-ESN predicted number of tourists in Beijing (d).

Table 1 AFOA-ESN and other model parameter settings

Model	**Parameter setting**
ARIMA	The level is 0, the difference degree is 3, and the order is 3
BPNN	The maximum training times are 2000, the learning rate is 0.01, the number of nodes in the input layer is 6, the number of nodes in the output layer is 1, the number of hidden layers is 1, and the number of nodes in the hidden layer is 5. The activation functions of the hidden layer and the output layer are the tansig function and linear function purelin, respectively, and the training function is the trainlm function
PCA-BPNN PCA-ADE-BPNN	The maximum training times of BPNN are 2000, the learning rate is 0.01, the number of nodes in the input layer is 6, the number of nodes in the output layer is 1, the number of hidden layers is 1, and the number of nodes in the hidden layer is 5. The activation functions of the hidden layer and output layer are the tansig function and linear function purelin, respectively. The training function is the trainlm function. The population size of the ADE algorithm is 50, the maximum number of iterations is 30, the crossover probability is 0.6, and the scaling factor is 0.2 ~ 0.9
FOA-ESN	The number of nodes in the input layer and the number of nodes in the output layer of the ESN are 12, the number of nodes in the output layer is 1, the size of the reserve pool is 50 to 500, the spectral half-diameter is 1.0 to 2.0, the activation function uses the tanh function, and the activation state memory parameter a is set to 0.20. The population size of the FOA algorithm is 10, and the maximum number of iterations is 100

Model	**Parameter setting**
AFOA-ESN	The number of nodes in the input layer and the number of nodes in the output layer of the ESN are 12, the number of nodes in the output layer is 1, the size of the reserve pool is 50 to 500, the spectral half-diameter is 1.0 to 2.0, the activation function uses the tanh function, and the activation state memory parameter a is set to 0.20. The population size and search step of the AFOA algorithm are adaptively adjusted based on the objective function value, and the maximum number of iterations is 100

Table 2 Predictive value and error analysis of FOA-ESN, AFOA-ESN model, and contrast algorithm

Types of models	No index	Baidu	Google	Baidu + Google
ARIMAX	9.17	5.96	6.18	4.05
ANN	4.07	2.42	2.37	1.97
SVR	3.59	2.2	2.31	1.93
LSSVR	3.41	1.83	2.12	1.7
KELM-lin	3.06	1.31	1.55	0.9
KELM-poly	1.92	1.2	1.44	0.79
KELM-rbf	1.71	1.03	1.36	0.64
KELM-wav	1.85	1.1	1.35	0.73
FOA-ESN	0.53			
AFOA-ESN	0.41			

Based on the findings presented in Table 2, for the AFOA, the three prediction error indexes of the ESN (MAPE, MSE, and MAE) are better than those of the FOA-ESN model, showing that the prediction accuracy improvement effect is obvious. In terms of the validation set in the process of optimization, for the AFOA-ESN initial iteration, the MAPE is below 1.15%, which is about 0.15% below the previous FOA-ESN. The AFOA-ESN and FOA-ESN model values are lower than 0.85%. In conclusion, for the AFOA-ESN model, the initial position optimization, convergence speed, and accuracy are better than those of the FOA-ESN model. The results are shown in Table 3.

Table 3 Algorithm prediction error analysis

	MAPE (%)	**MSE**	**MAE**
ARIMA	14.66	4685.30	63.33
VAR	17.85	8560.00	79.52
PCA-VAR	14.34	5280.00	61.59
BPNN	6.97	2120.00	32.55
PCA-BPNN	6.98	1940.00	32.46
PCA-ADE-BPNN	6.42	1930.00	30.96
FOA-ESN	2.41	161.52	10.63
AFOA-ESN	1.56	71.71	6.86

5. CONCLUSIONS

In this paper, we proposed that the AFOA algorithm effectively improves the convergence speed, the prediction accuracy, and the initial position optimization ability. We combined the AFOA and ESN algorithm to build a new combined prediction model, i.e., the AFOA-ESN, and its operation principles and processes were discussed in detail. The combined prediction model effectively improves the problems of the standard ESN model, where parameters are random, unstable, and easily fall into local optima. We applied the AFOA-ESN model to sports ecotourism through different types of experiments. The application effect of the FOA-ESN for the tourism industry demand prediction problem and the improvement effect of the AFOA-ESN compared with the FOA-ESN model were verified in stages, and good experimental results were obtained. The MAPE for the initial AFOA-ESN iteration was below 1.15%, about 0.15% lower than the previous FOA-ESN. In addition, with the rise of the big data industry, increasing numbers of tourism-related databases have begun to be established, the most typical being the application of the network search index. In our subsequent research, we will use a variety of big data means to enrich the diversity of the input information and select prediction information suitable for the sports ecotourism industry.

FUNDINGS

This study was supported by the Humanities and Social Science Research Projects in Anhui Universities in 2020: Research on the Integration and Development of Sports and Ecotourism in the Jianghuai Watershed (Project Number: SK2020A0510) and Philosophy and Social Science Project of Department of Education of Hubei province in 2021 Research on the construction path of football characteristic school in rural campus under the rural revitalization (Project number: 21G014).

REFERENCES

Chen, H.L., Li, S.M., Yang, Y.T., Li, J.W., Zeng, D., Ren, N., Wang, P.J., and Chen, Y.P. (2019). Method of Constructing Prediction Model Based on Sentinel Multiple Swarm Mechanism. CN109948675A. Retrieved on June 6, 2019, from https://static.tianyancha.com/patent/contentPDF/afa979ab860116a331f20ccbfc9b4567.pdf

Dehghani, M., Mardaneh, M., and Malik, O.P. (2020). FOA: 'Following' Optimization Algorithm for Solving Power Engineering Optimization Problems. *University of Mohaghegh Ardabili*, 46(1): 27.

Fan, Q., Meng, X.L., Xu, C.Q., and Yu, J.Y. (2019). Solution Method for Ill-conditioned Problems Based on a New Improved Fruit Fly Optimization Algorithm. *Journal of Applied Geodesy*, 14(1): 55-64.

Han, J.Y., Liu, C.Z., and Wang, L.G. (2013). Dynamic Double Subgroups Cooperative Fruit Fly Optimization Algorithm. *Pattern Recognition and Artificial Intelligence*, 26(11): 1057-1067.

He, C., Liu, C., Wu, T., Wu, Y., Xu, Y., and Chen, T. (2019). Rolling Bearings Fault Diagnosis Research Based on WKELM Optimized by Improved FOA Algorithm. *Computer Measurement & Control*, 75(86): 546-547.

He, D.W., Peng, J., Hu, J.H., Li, T., and Jia, W. (2018). Bearing Fault Diagnosis Based on a Modified CS-SVM Model Optimized by an Improved FOA Algorithm. *Journal of Vibration and Shock*, 37(18): 108-114.

Huang, Z.Q., Xu, Z.H., Sue, Boyd, David and Williams. (2005). A Fruit Fly Optimization Algorithm Based on Bacterial Chemotaxis. *Scientific Bulletin*, 50(21): 2365-2369.

Li, C., Zhao, G., Meng, J., Zheng, Z., and Yu, S. (2022). Multi-objective Optimization Strategy Based on Entropy Weight, Grey Correlation Theory, and Response Surface Method In Turning. *International Journal of Industrial Engineering: Theory, Applications and Practice*, 28(5): 7327.

Luo, S., Sarabandi, K., Tong, L., and Pierce, L. (2019). A SAR Image Classification Algorithm Based on Multi-feature Polarimetric Parameters Using FOA and LS-SVM. *IEEE Access*, 7: 175259-175276.

Nazir, H., Bajwa, I.S., Samiullah, M., Anwar, W., and Moosa, M. (2021). Robust Secure Color Image Watermarking Using 4D Hyperchaotic System, DWT, HbD, and SVD Based on Improved FOA Algorithm. *Security and Communication Networks*, 28(7): 6617944.

Niu, P.F., Ma, H.B., Li, G.Q., Ma, Y.F., Chen, G.L., and Zhang, X.C. (2013). Study on NO_x Emission Characteristics of Circulating Fluidized Bed Boilers Based on Support Vector Machine and Fruit Fly Optimization Algorithm. *Journal of Power Engineering*, 33(4): 267-271.

Pan, W.C. (2011). Fruit Fly Optimization Algorithm to Optimize Generalized Regression Neural Network to Conduct Business Performance Evaluation. *Journal of Taiyuan University of Technology: Social Science Edition*, 29(4): 1-5.

Song, J., Xu, B., and Yang, M.Z. (2020). Image Segmentation Based on the Adaptive Step-fly Optimization Algorithm. *Computer Engineering and Application*, 56(4): 184-190.

Wang, M.F., Gao, G.D., Shan, W.F., Gong, Q.Z., and Han, D.L. (2021). Deep Learning Method of Diabetic Retinopathy Classification Based on Fruit Fly Optimization Algorithm. CN112184697A. Retrieved on January 5, 2021, from https://wenku.baidu.com/view/06375fccad45b307e87101f69e3143323868f51d?fr=xueshu&_wkts_=1700017309543

Wang, T., Zhou, J.X., Yin, Y.J., Shen, X., and Zhou, Q. (2016). Exit Optimization Based on the Fruit Fly Optimization Algorithm. *Special Casting and Non-ferrous Alloy*, 36(3): 246-249.

Wang, X.G., and Zou, Z.J. (2013). Application of Support Vector Machine Parameter Optimization Based on Fruit Fly Optimization Algorithm in Ship Manipulation Prediction. *Journal of Shanghai Jiaotong University*, 47(6): 884-888.

Wang, X.G., and Zou, ZJ. (2012). Identification of a Ship Manipulation Response Model Based on a Fruit Fly Optimization Algorithm. *Journal of Dalian Maritime University*, 38(3): 1-8.

Wu, Y.X., and Zhou, Y.Q. (2017). A Fruit Fly Optimization Algorithm Based on Polar Coordinate Encoding. *Practice and Understanding of Mathematics*, 47(10): 161-170.

Xi, N., and Han, H. (2014). Multiple Swarm Fly Optimization Algorithms Solve Automated Warehouse Picking Job Scheduling Problems. *Journal of Wuhan University of Technology*, 36(3): 108-113.

Xie, X., Wang, W.R., Jiang, X.Y., Wan, T.C., and Hu, F.P. (2020). Research and Application of Image Edge Detection Algorithm Based on FOA. *DEStech Transactions on Environment, Energy and Earth Sciences*, DOI: https://doi.org/10.12783/dteees/peems2019/33988

Zhang, Q.T., Fang, L.Q., and Zhao, Y.L. (2015). Gemini Group Fruit Fly Optimization Algorithm with Levy Flight Characteristics. *Computer Applications*, (5): 1348-1352.

Zhang, X., and Chen, L. (2013). SVM Fault Diagnosis Based on the Fruit Fly Optimization Algorithm. *Electronic Design Engineering*, 21(16): 4.

Zhang, X.H., Hu, X., He, C., and Yu, C.M. (2018). A FOA-optimized RBF Algorithm-based Evaluation Research on E-commerce Websites. *Wireless Personal Communications*, 102: 2835–2851.

Zhang, Y., Xia, S.F., and Tang, D.S. (2014). Simulation Study of the Performance of Multimodal Function. *Journal of Jinan University (Natural Science and Medicine Edition)*, 35(1): 82-87.

Zhang, Y.J., Liu, W.Z., Fu, X.H., and Bi, W.H. (2015). Brillouin Scattering Spectrum Feature Extraction Based on Adaptive Variant Fruit Fly Optimization Algorithm and Generalized Regression Neural Network. *Spectroscopy and Spectral Analysis*, 35(22): 2916-2923.