Water Quality Parameters Identification Model Based on Artificial Fish Swarm Algorithm with Adaptive Parameter Optimization

Xiaohui Chen, Wenzhou Yan, Kang Liu

Xi'an University of Architecture and Technology, Xi'an 710055, Shaanxi, China

Abstract
In view of the bad convergence performance and low precision of standard artificial fish swarm algorithm in the water quality properties identification, this paper put forward an improved identification model based on adaptive parameters optimization. Firstly, it optimized the immune cloning and selection algorithm (ICSA) in periodic mutation operator and selection operator. Then it introduced the diversity and immune memory property of the optimized immune cloning and selection algorithm into the artificial fish swarm algorithm and optimized the parameters adaptively. The diversity maintaining strategy based on concentration mechanism was adopted to keep a certain of artificial fish concentration of different fitness values in the new generation of artificial fish swarm. Finally, the improved artificial fish swarm algorithm was used in the identification of water quality parameters. The simulation experiments show that the improved artificial fish swarm algorithm in this paper has better convergence performance than traditional ones and has lower error in the water quality property identification.
Key words: ARTIFICIAL FISH SWARM ALGORITHM, ICSA, OPERATOR OPTIMIZATION, IMMUNE MEMORY, CONCENTRATION MECHANISM, DIVERSITY MAINTAINING STRATEGY

1. Introduction

Water is an important resource for the survival of mankind, raising billions of people on the earth and promoting the development of industry, agriculture and even the society, so that it becomes an irreplaceable material basis for human survival. However, with the expansion of human activities, the development of the world economy, the rapid increase of population and the acceleration of urbanization, the pollution sources and the amount of sewage are increasing rapidly. The water body has gradually been polluted and influenced the water use [1]. The worse water environment quality makes water pollution one of the global issues of concern. Therefore, in the study of the use of water, not only the amount of water should be considered, but also the water quality in natural water should be analyzed and predicted.

Due to the limitation of hydrology and water quality monitoring, most of channel segments lack relatively accurate and reliable water quality model. Liu combined with the actual data change trend to get the general reasons for the impact on the prediction accuracy. The error detection analysis proved that there was significantly difference between the error correction model and modified residual error model in the vicinity of the data changing point. The latter one can improve the accuracy of some non-stationary random number sequences, whose data was closer to the actual data details. Pan Jun pointed out that the traditional solution of the grey dynamic model group method was to improve the prediction accuracy of the model group, but it cannot solve the error problems [2]. Song Yuejun processed the original monitoring data sequence by exponential smoothing process into a sequence of exponential law, and then GM (1, 1) model was constructed. However, when the random fluctuations of the forecasting data sequence, a single GM (1, 1) model fitting effect was poor, which limi-

ted the prediction accuracy. Gray metabolism GM (1, 1) model removed the old data in time and constantly added new information, and has achieved good results in the practical application [3]. A logarithmic transformation of ammonia nitrogen original data sequence was conducted by Li Ruzhong in the main stream of the Huaihe River in dry season to construct a grey dynamic model group consisting of six GM (1, 1) models. The results showed that this model can effectively improve the fitting effect of random fluctuation data sequence compared with single GM (1, 1) model [4]. Lapedes et al first applied neural network to the learning and prediction of the time series simulation data generated by computer. Then the intelligent technology represented by neural network has been widely used in the field of forecasting [5]. A. Lapedes and R. Farber have published the first article about the application of the neural network into forecast, and it is gradually applied to solve the problem of the hydrology [6]. In the process of water quality simulation, Maier et al. used ANN model to predict the parameters of water quality model [7]. Other researchers have applied ANN technology to the water environment. According to the distribution of automatic water quality monitoring systems of Dongjiang River, Li Ying proposed two kinds of water quality prediction models based on adaptive neural network: prediction of downstream water quality from the upstream water quality; prediction of future water quality based on the current water quality. A static and dynamic learning algorithm based on orthogonal polynomial basis was also given out [8]. According to the mapping relationship between organic matter in river surface water and its main influencing factors, Yang Yuchuan established the BP network model with higher accuracy of prediction [9]. Qiu Jing built a forecast model based on the three layer BP network, improved it, and simulated the water quality of the Yangtze River basin. The results are objective and have strong generalization ability [10].

In view of the shortcomings of standard artificial fish swarm algorithm in the water quality identification, this paper put forward an improved model based on adaptive parameter optimization and conduct the simulation experiment to improve the effectiveness.

2. Artificial fish swarm algorithm

The vision of artificial fish swarm is similar to the region of optimization in each searching. Figure 1 constructs the visual model of artificial fish according to the activity of fish in reality.

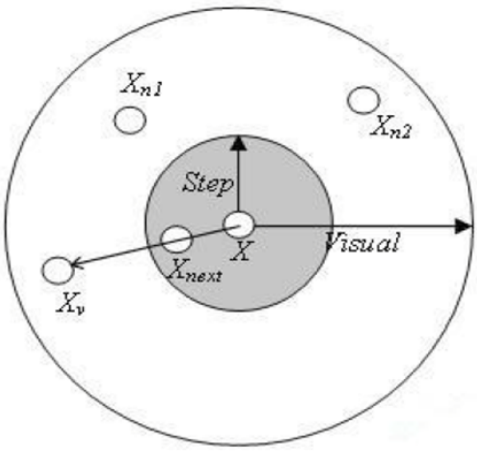

Figure 1. Visual model diagram of artificial fish swarm

As shown in Figure 1, X is the current state of artificial fish, $Visual$ is the field of view, and X_V is location of viewpoint of artificial fish. If the state of viewpoint X_V is better than current state, then one step length is added to the location of X_V to be X_{next}. If current state is better than the state of X_V, this artificial fish will continue searching in the field of view until it finds better state or follows other conditions to escape from the local optimal.

Here, current state is $X=(x_1,x_2,...,x_n)$ and state of view point is $X_V=(x_1^V,x_2^V,...,x_n^V)$. This process can be described as,

$$x_1^V = x_i + Visual \cdot Random() \quad (1)$$

$$i = Random() \quad (2)$$

$$X_{next} = \frac{X_V - X}{\| X_V - X \|} \cdot Step \cdot Random() \quad (3)$$

$Random$ function is generated randomly from 0 to 1. $Step$ is the step length.

The definitions of the parameters in artificial fish swarm algorithm are: $step$ is the step length of artificial fish; $visual$ is the field of view; try_number is the try number of artificial fish; δ is the index of

crowding degree; n is the total count. Fish swarm algorithm includes four standard behaviors similar to genetic algorithm (mutation, duplication, and crossover)

(1) Foraging behavior

$$X_j = X_i + visual \cdot rand \tag{4}$$

Where X_i is current state; X_j represents the random location in the field of view; Y_i and Y_j is the food concentration (fitting function value) in X_i and X_j, respectively. If $Y_i < Y_j$, the state X_i at this time t is updated to Equation (5)

$$X_i^{t+1} = X_i^t + \frac{X_j - X_i^t}{\left\| X_j - X_i^t \right\|} \cdot step \cdot rand \tag{5}$$

Otherwise,

$$X_i^{t+1} = X_i^t + visual \cdot rand \tag{6}$$

(2) Bunching behavior

X_i is current state of artificial fish; X_c is the central position; n_f is the number of artificial fish in the field of view; n is total number of artificial fish. If $Y_c > Y_i, \frac{n_f}{n} < \delta$, the position of artificial fish is updated.

$$X_i^{t+1} = X_i^t + \frac{X_c - X_i^t}{\left\| X_c - X_i^t \right\|} \cdot step \cdot rand \tag{7}$$

Otherwise foraging behavior is done.

(3) Rear-end behavior

X_i is current state of artificial fish, and X_j is the place where food concentration is highest in the field of view. If $Y_j > Y_i, \frac{n_f}{n} < \delta$, the position of artificial fish is updated.

$$X_i^{t+1} = X_i^t + \frac{X_j - X_i^t}{\left\| X_j - X_i^t \right\|} \cdot step \cdot rand \tag{8}$$

Otherwise foraging behavior is done.

(4) Random behavior

This behavior is the default behavior of artificial fish, meaning the fish can move randomly in the field of view, described as follows.

$$X_i^{t+1} = X_i^t \cdot visual \cdot rand \tag{9}$$

After executing this basis behavior, the fish swarm will be evaluated according to fitness function value and their next movement is coordinated until the global optimum is reached.

Through a lot of researches and simulation experiments about artificial fish swarm algorithm, it is found that this algorithm still has some defects. For example, there are many artificial fish singles in the foraging activity. All of these fishes will not be used, which causes the waste of resource. Because the more singles are, the more space the calculation requires and the lower the efficiency will be. In term of precision, standard artificial fish swarm algorithm also has big defects. When parameter setting is better, it can search the optimum solution. But when the parameter setting is inappropriate, only the scope or roughly interval of the solution oscillating around the optimum solution is obtained. When the artificial fish meets a much large searching space or the real problem is a large system, the efficiency of the entire algorithm will decrease significantly so as to influence the final optimum result. All of these defects will limit the application in practical problems.

3. Water quality parameter identification model based on artificial fish swarm algorithm with adaptive parameter optimization

3.1. Immune cloning algorithm based on operator optimization

In view of the defects of artificial fish swarm algorithm, this paper adopts the immune cloning selection algorithm (ICSA) to adaptively optimize parameters. After immune selection, global optimum of traditional algorithm appears around antibody with low affinity, which causes premature convergence of algorithm. Therefore, operator must be optimized.

(1)ICSA period mutation operator optimization

Cloning enlarges the population size. The mutation operation of the antibody after cloning will improve the diversity and extend the searching range to find better antibodies. The mutation probability is,

$$P_m(t) = \varepsilon \left| \cos \frac{2\pi}{T} t \right| + \gamma, t \geq 0 \tag{10}$$

Here, $P_m(t)$ is mutation probability; t is the generation time; T is mutation period; ε is mutation probability adjustment index; γ is the disturbance, $0 < \gamma << \varepsilon$ .When $t = kt/2$ and k is odd, the mutation probability will not be zero.

(2) ICSA selection operator optimization

$D(Ab_i)$ is defined as the concentration of the antibody Ab ; $S(Ab_i)$ is the number of antibodies similar to Ab_i, then

$$D(Ab_i) = S(Ab_i)/N, i = 1,2,...,N \tag{11}$$

$a(Ab_i)$ and $d(Ab_i)$ is defined as the affinity function and vector distance, respectively. Then,

$$d(Ab_i) = \sum_{j=1}^{N} | a(Ab_i) - a(Ab_j) | \tag{12}$$

$P_s(Ab_i)$ is defined as the selection probability of antibody Ab_i based on $D(Ab_i)$ and $d(Ab_i)$.

$$P_s(Ab_i)=\frac{d(Ab_i)\exp[-D(Ab_i)]}{\sum_{j=1}^{N}d(Ab_j)\exp[-D(Ab_j)]} \quad (13)$$

From Equation (13), when the

$$P_s(Ab_i)=\frac{d(Ab_i)\exp[-D(Ab_i)]}{\sum_{j=1}^{N}d(Ab_j)\exp[-D(Ab_j)]}$$

is constant, the more $d(Ab_i)$ is, the larger $P_s(Ab_i)$ will be; when $d(Ab_i)$ is constant, the more $D(Ab_i)$ is, the lower $P_s(Ab_i)$ will be, so as to ensure the diversity of antibody.

(3) Convergence analysis of Improved ICSA algorithm

The state transfer of improved ICSA algorithm can be described by following random process.

$$A_n \xrightarrow{cloning} B_n \xrightarrow{mutation} C_n \xrightarrow{selection} A_{n+1} \quad (14)$$

This state transfer from A_n to C_n construct the markoff chain. The random process $\{A_n \mid n=1,2,...\}$ is still a markoff process. If X and S is the searching space and state space, $s_i \in S$ is the i state of S, and a is the affinity function, then

$$s^* = \{x \in X \mid a(x) = \max_{x_i \in X}\{a(x_i)\}\} \quad (15)$$

It is defined, if for arbitrary original state distribution, there is

$$\lim_{n\to\infty}\sum_{s_i \cap s^*} p\{A_n^i\} = 1 \quad (16)$$

then the algorithm is convergent.

The convergence analysis of improved ICSA is described as follow.

$p_{ij}(n)$ is defined as the transfer probability of $\{A_n\}$, $p_{ij}(n)=p\{A_{n+1}^j / A_n^i\} \ge 0$, and $p\{A_n^i\}$ is $p_i(n)$ $p_n = \sum_{i\notin I} p_i(n)$.

According to the properties of markoff chain,

$$p_{n+1} = \sum_{s_i \in S}\sum_{j\notin I} p_i(n)p_{ij}(n) = \sum_{i\in I}\sum_{j\notin I} p_i(n)p_{ij}(n) + \sum_{i\notin I}\sum_{j\notin I} p_i(n)p_{ij}(n) \quad (17)$$

$$\sum_{i\notin I}\sum_{j\notin I} p_i(n)p_{ij}(n) + \sum_{i\notin I}\sum_{j\in I} p_i(n)p_{ij}(n) = \sum_{i\notin I} p_i(n) = p_n \quad (18)$$

$$\sum_{i\notin I}\sum_{j\notin I} p_i(n)p_{ij}(n) = p_n - \sum_{i\notin I}\sum_{j\in I} p_i(n)p_{ij}(n) \quad (19)$$

When $i \in I, j \notin I$,

$$p_{ij}(n) = 0 \quad (20)$$

when $i \notin I, j \in I$

$$p_{ij}(n) \ge 0 \quad (21)$$

Adding equations (19), (20) and (21) into equation (17), it is obtained,

$$p_{n+1} = p_n - \sum_{i\notin I}\sum_{j\in I} p_i(n)p_{ij}(n) \le p_n \quad (22)$$

If $\lim_{n\to\infty} p_n = 0$, then

$$1 \ge \lim_{n\to\infty}\sum_{s_i \cap s^*} p_i(n) \ge \lim_{n\to\infty}\sum_{i\in I} p_i(n) = 1 - \lim_{n\to\infty} p_n = 1 \quad (23)$$

Therefore, the improved ICSA algorithm is convergent to probability 1.

3.2. Artificial fish swarm algorithm based on improved ICSA algorithm

Improved ICSA algorithm is utilized for adaptive parameter optimization of the artificial fish swarm algorithm. The excess diversity and immune memory is introduced into the algorithm to improve the global searching ability and avoid the local solution. The lower the concentration of antibody is, namely artificial fish, the higher probability is. The higher the concentration of antibody is, the lower probability is. Under this condition, high fitness single can be stored to ensure the diversity of antibody and avoid the premature phenomenon.

This algorithm flow is written as follows.

(1) To determine the parameter value: define the step length $Step$, the visual range $Visual$, number of artificial fish in the swarm N, crowding degree index δ and cycling time t_N.

(2) To initialize fish swarm: randomly generate N fishes in the domain of definition $X_i, i=1,2,...,N$ to have the initial fish swarm P_0. Updating the bulletin board if necessary until the initialization of whole fish swarm.

(3) To generate the fish with immune memory: foraging, rear-end and bunching behavior; calculate the optimal value p_k in current fish swarm P_k, updating bulletin board according to corresponding content in the course of evolution and set p_k as immune to store the fish into the memory vault. If it meets the ending condition, the operation end, otherwise process continues.

(4) To generate new artificial fish in following two ways: 1) N new fishes from foraging behavior; 2) M new random fishes

(5) The selection of artificial fish based on concentration: calculating the selection probability of $N+M$ in step 4 and selecting N fishes to form the fish swarm Q_k.

(6) To update the fish swarm: using the immune memory fish in the memory vault to replace some

fishes q with bad fitness and generate new generation P_{k+1}, then turn to step 3

This paper takes the diversity maintaining strategy based on concentration mechanism to keep the fish of different fitness values at certain concentration. The concentration of fish i is defined as,

$$D(X_i) = \frac{1}{\sum_{j=1}^{N+M} \left| f(X_i) - f(X_j) \right|} \tag{24}$$

The selection probability based on fish concentration can be derived.

$$P(X_i) = \frac{\sum_{i=1}^{N+M} \left| f(X_i) - f(X_j) \right|}{\sum_{i=1}^{N+M} \sum_{j=1}^{N+M} \left| f(X_i) - f(X_j) \right|} \tag{25}$$

X_i represents the fish i and $f(X_i)$ is its fitness value. Form equation (25), the more antibodies similar to i, the lower selection probability will be. Therefore, the selection probability based on antibody concentration ensures the diversity in theory.

3.3. Water quality parameter identification model based on improved artificial fish swarm algorithm

In the optimization of water quality model parameters, a special measurement error between water quality variable calculation sequence and real observation sequence is taken as the objective function. The method of optimization is used to solve the appropriate parameters values which minimize the objective function. That is to say, starting from real water quality data, a group of parameters are searched out to make output data well fit the real observation data. This group of parameters is the optimal parameters, namely,

$$\min f(y, y^*, \theta) = f(y, y^*, \theta^*) \tag{26}$$

The advantage of water quality model parameter optimization is a global searching method which has greatly increased the reliability. In the parameter optimization, the error sum of squares between observation value and model output value is seen as the objective function.

$$\min f(y, y^*, \theta) = \min \sum_{i=1}^{n} \sum_{j=1}^{m} \lambda_i (y_{ij} - y_{ij}^*)^2 \tag{27}$$

From equation (27), the calculation patter of this proposed model can be summarized as an improved artificial swarm algorithm solving a constraint nonlinear programming problem.

4. Simulation experiment

To test the performance of the improved algorithm, this paper adopted the two dimensional G+ function to analyze this artificial fish swarm algorithm, then compared it with the traditional artificial fish swarm algorithm. The comparison results are shown as follows.

$$f(x) = \frac{1}{4000} \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right) \tag{28}$$

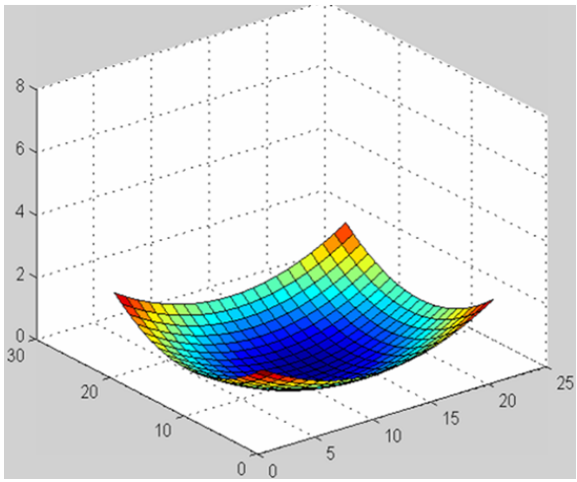

Figure 2. Two dimensional G + function image

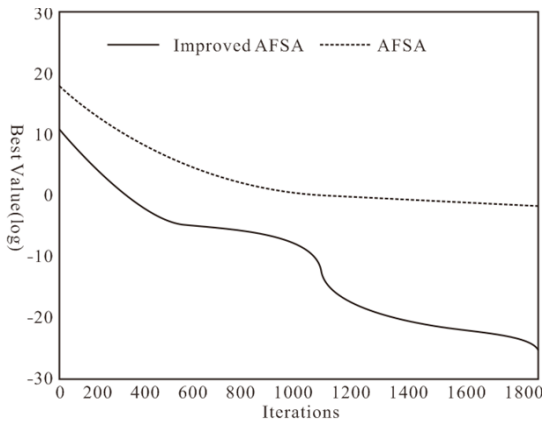

Figure 3. Convergence analysis of improved AFSA

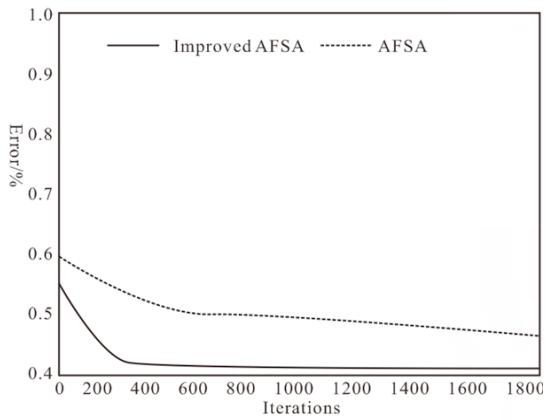

Figure 4. Error analysis of improved AFSA

Results show that the improved artificial fish swarm algorithm has better convergence performance

than standard artificial fish swarm algorithm and has smaller error in the application of water quality parameter identification.

5. Conclusions

Rapid economic development gives pressure on the water environment. Accurate prediction of water quality provides the basis for protection of water environment. Therefore, considering the defects in current standard artificial fish swarm algorithm in the identification of water quality properties, this paper puts forward a improved artificial fish swarm algorithm model based on adaptive parameter optimization. The simulation results show that this improved model has better convergence performance and has good effect in the application of water quality identification.

References

1. Fu C, Liu Y H. (2015) Simplex-differential evolution hybrid algorithm for parameter identification of river water quality model. *Journal of Hydroelectric Engineering*, 34(1), p.p.125-130.
2. Liu X D. (2015) General inversion optimization algorithm for multi-parameter identification of 1D river water quality model. *Journal of Hydroelectric Engineering*, 31(3), p.p.122-127.
3. Kang A Q. (2014) Parameter Identification of Water Quality Model Based on Framework of Water Resources Full Elements Allocation. *International Journal Hydroelectric Energy*, 29(8), p.p.14-17.
4. Hu F J. (2013) A rapid eye-to-hand coordination method of industrial robots. *Journal of Information and Computational Science*, 10(5), p.p.1489-1496.
5. Chen G Z, Xu X C. (2014) Application of a modified artificial fish swarm algorithm to identification of water quality parameters. *Journal of Hydroelectric Engineering*, No.2, p.p.108-113.
6. Luo G Y. (2014) Parameters identification of water quality model in branch backwater reach based on genetic algorithm. *China Environmental Science*, No.9, p.p.962-966.
7. Hu F J, Zhao Y W, Chen J. (2014) SIFT Feature Points Detection and Extraction of Three-Dimensional Point Cloud. *WIT Transactions on Information and Communication Technologies*, 60, p.p.603-611.
8. Deng Y X. (2014) Application of Bayes theorem in parameter identification for river water quality modeling. *Acta Scientiae Circumstantiae*, 28(3), p.p.568-573.
9. Zhu S, Mao G H. (2013) Parameters identification of river water quality model based on finite volume method-hybrid genetic algorithm. *Journal of Hydroelectric Engineering*, 26(6), p.p.91-95.
10. Wang. (2014) Parameter Optimization of Water Quality Model: Implementation of Genetic Algorithm and Its Control Parameters Analysis. *Chinese Journal of Environmental Science*, 26(3), p.p.61-65.