

# Monitoring point optimization in lake waters

Gaoxuan Liu, Jiaoyan Ai, Jun Xu, Jianwu Zheng and Dongyi Yao

## ABSTRACT

In order to grasp the distribution of water quality index in lake water, taking Jinghu Lake of Guangxi University as the experimental object, an radial basis function (RBF) neural network was combined with a genetic algorithm on the basis of an unmanned ship to study the optimal selection of monitoring points. The single-objective and multi-objective optimization of water quality parameters were tested respectively and used to make the fitting distribution map. The results show that the genetic neural network has obvious advantages over the traditional isometric monitoring in the distribution error of water quality parameters, and the data reflected by the results are still accurate and effective at least six weeks after optimization. The results show that a genetic neural network can significantly improve the efficiency of water quality monitoring.

**Key words** | genetic algorithm, layout optimization, neural network, water quality monitoring

Gaoxuan Liu  
Jiaoyan Ai (corresponding author)  
Jun Xu  
Jianwu Zheng  
Dongyi Yao  
School of electrical engineering,  
Guangxi University,  
Nanning, Guangxi 530000,  
China  
E-mail: [shinin@vip.163.com](mailto:shinin@vip.163.com)

## INTRODUCTION

With the increase in outdoor recreational activities, the degree of human activity and changes to the natural environment, such as lakes, have become a problem (Li 2012). Therefore, water quality monitoring is particularly important for mastering the changing and the trend of water quality. In water quality monitoring, the layout of the monitoring points directly affects the efficiency and accuracy of the monitoring work. The optimal selection of monitoring points will improve the working efficiency of the measurement staff and save economic expenditure (Wang *et al.* 2013). Long-term effective data collection and analysis of lake waters can help meet laws outlined for lake environmental changes, the distribution of various parameters, and thus prevent and control water pollution in a timely and effective manner (Bai *et al.* 2012).

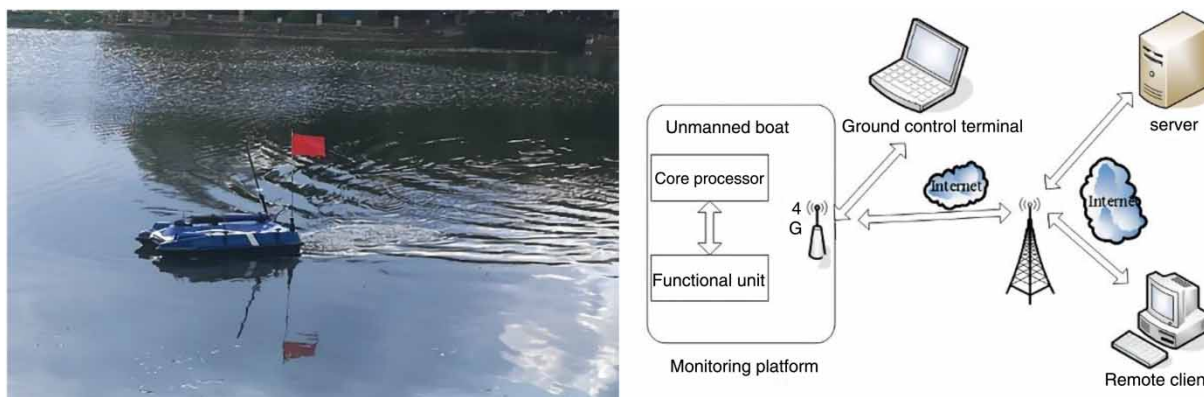
In recent years, as people pay more attention to water resources, better management of water resources is being explored. For lakes, monitoring stations to monitor water quality are used, but construction costs are high and maintenance is difficult (Liu *et al.* 2013). Mobile water quality

monitoring is a feasible method, but the existing mobile monitoring equipment has the disadvantages of large volume, inconvenient carrying, high energy consumption and secondary pollution. Therefore, we designed an autonomous mobile water quality monitoring system. It mainly consists of a monitoring platform (water quality monitoring unmanned ship), ground control terminal, remote client and hand-hold terminal (Figure 1). The system can realize autonomous movement through path editing and realize on-line monitoring of temperature, pH, dissolved oxygen, conductivity and chlorophyll a.

At present, for the optimization of lake water monitoring points, common methods include cluster analysis, dynamic closeness method, corresponding analysis, matter element analysis and other mathematical statistics methods. Among them, the cluster analysis method is simple in the case of small sample data, but ignores the interconnection of data space distribution (Mahbub *et al.* 2010). The dynamic closeness method can reflect the dynamic changes of water quality parameters at different times and perform cluster analysis but does not reflect the overall spatial distribution (Cui *et al.* 2015). The correspondence analysis method can intuitively put many sample variables on the same graph at

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

doi: 10.2166/ws.2020.147



**Figure 1** | Monitoring platform (water quality monitoring unmanned ship) and water quality control system.

the same time, but the results differ greatly for different evaluation environments (Zheng *et al.* 2007). The concept of matter-element analysis is clear and the calculation is simple, but the actual geographical location and environment of the monitoring point are not considered (Wang *et al.* 2015). Based on the above situation, this experiment uses a water quality monitoring unmanned ship, relying on its fast and efficient data monitoring characteristics, and adopts the method of combining a genetic algorithm with an radial basis function (RBF) neural network. The genetic algorithm utilizes the rule ‘survival of the fittest’ and has a good global search ability. RBF neural network has good generalization ability for spatial fitting of data, and a quick learning convergence speed, can complement each other, a combination which meets the experimental requirements of robustness and accuracy, and keeps the fast convergence certain, reflecting the parameter distribution of waters (Simon 1994).

## METHODS

### Data sources

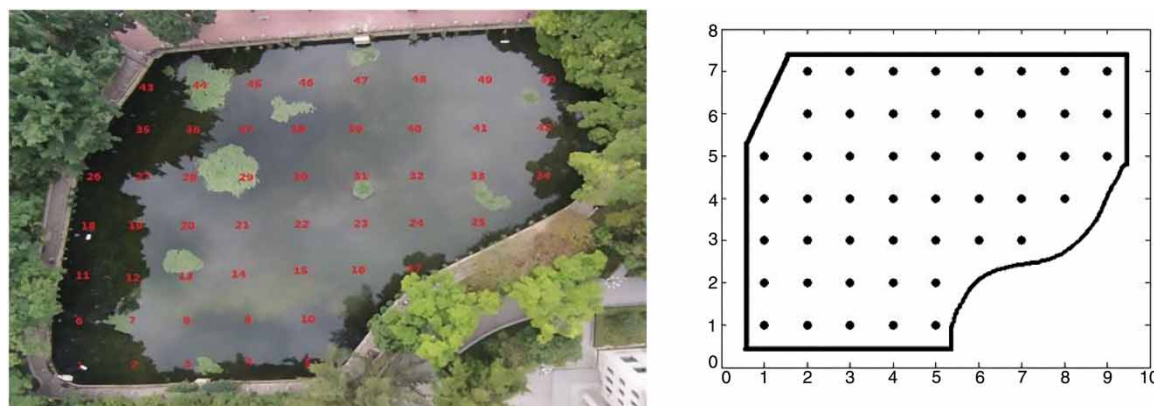
Jinghu Lake is located at Guangxi University, with an area of about 3,000 m<sup>2</sup>. It is a typical small landscape lake. The previous management and maintenance of Jinghu Lake was generally determined by random sampling analysis or by visual experience. This method cannot fully grasp the water quality of Jinghu Lake, and it is difficult to judge and predict its change rule. A comprehensive understanding of the water quality requires adequate water quality testing,

but a large amount of water quality testing requires a lot of manpower and financial resources. Therefore, we use the developed water quality monitoring system to obtain water quality information more efficiently and economically.

The extraction of water quality parameters will be extracted by an unmanned ship. It can detect temperature (T), pH, dissolved oxygen (DO), conductivity (COND) and chlorophyll a (Chl-a) by carrying a water quality monitoring sensor. Total phosphorus (TP) and total nitrogen (TN) are monitored in the laboratory through water collected in sampling bottles in unmanned ships. From October 2018, we conducted an eight-week water quality test on Jinghu Lake, choosing Tuesday mornings each week to test the water quality. The Jinghu lakes were roughly divided into 50 grid areas according to their size and numbered, with the center of each grid area selected as the monitoring point (Figure 2), in order to simplify subsequent operations in the algorithmic model, the coordinates in Figure 2 have been designed to match the individual monitoring points). When using the developed water quality monitoring system for water quality monitoring, the longitude and latitude of each monitoring point are calibrated through the electronic map in the ground control terminal, and then the unmanned ship is navigated to each monitoring point through GPS technology for water quality detection and water sampling. The water quality parameters for the first week are shown in Table 1.

### RBF neural network

In this study, the geographical coordinates of the monitoring points are used as the input of the RBF neural network



**Figure 2** | Distribution of monitoring points in Jinghu Lake.

**Table 1** | Water quality parameters of the monitoring site in the Jinghu Lake

**Monitoring point water quality parameter table**

Monitor point number	T (°C)	pH	DO (mg/L)	COND (μS/cm)	Chl-a (μg/L)	TP (mg/L)	TN (mg/L)
1	30.49	6.84	4.53	106	6.1128	0.073	1.276
2	30.4	6.78	5.60	110	5.8300	0.042	1.395
3	30.29	6.96	5.27	98	6.9616	0.048	0.893
...	...	...	...	...	...	...	...
50	34.91	7.10	9.78	92	9.2842	0.085	0.853

fitting, and the water quality parameter values are used as the output. The RBF neural network was used to establish the relationship between the geographic coordinates of the monitoring points and the water quality parameter value (Chen *et al.* 2013). The function expression is as follows:

$$Z = f(x, y) \quad (1)$$

where (x, y) is the geographic coordinate of the sampling point, and Z is the water quality parameter value. According to the information of the existing samples, use the neural network to train the training samples, and converge the relationship between the coordinates of the monitoring points and the water quality parameter values into the network. Then random geographic coordinates are entered and simulations are performed using the network to obtain a more optimal neural network parameter setting (Broomhead & David 1988; Hanbay *et al.* 2007).

### Water temperature spatial distribution fitting

We take the temperature index of the first week as an example to carry out the experiment. The spatial distribution fitting of water temperature was completed by MATLAB. Firstly, the monitoring points are marked with coordinates, where x and y coordinates are defined as geographical location coordinates, and Z is the corresponding temperature value. Then, the real number is encoded according to the monitoring serial number, and the corresponding data table is made to facilitate the decoding operation of the information. The temperature data table is as shown in Table 2. Then, we use the 'meshgrid' function to perform an interpolation on the data. After processing, a total of 1,021 points of information were obtained, and we randomly selected 21 points as the test set and another 1,000 points as the training set used as RBF neural network training. The experiment used a trial and error method to adjust the parameters to achieve the ideal fitting effect. Finally, we

**Table 2** | Temperature data

Monitor point number	x	y	z (°C)
1	1	1	30.49
2	2	1	30.40
3	3	1	30.29
...	...	...	...
50	9	7	34.91

obtained the temperature spatial distribution data of Jinghu Lake in the first week (Figure 3).

### Genetic algorithm optimization

A genetic algorithm is a kind of evolutionary algorithm. It searches for the optimal solution based on the principle of 'survival of the fittest' to simulate the natural genetic mechanism (Kaya 2011). It has good global optimization and robustness (Chen 1995). This paper uses the data collected from the original water quality monitoring points as a reference standard, and then uses the spatial distribution of water quality indicators fitted by the RBF neural network as the fitness selection function, and uses genetic algorithms to optimize the number and spatial layout of the monitoring points. The following is an example of a single target optimization to describe the flow of a genetic algorithm; the principle flow chart is shown in Figure 4.

First, the sample points are numbered (from 1 to n) using real number coding, and randomly generate an initial set of individuals to form the initial population; each chromosome is an array of real numbers encoded and the initial population is a matrix array.

Then, the fitness value of each group of chromosomes was calculated. The fitness function is the error mean square error between the numerical distribution and the standard distribution of each set of chromosomes. The expression is:

$$MSE = \frac{1}{N} \sum_{i=1}^N (\hat{Z}_i - Z_i)^2 \quad (2)$$

where  $MSE$  is the error-average variance,  $N$  is the number of contrast points,  $\hat{Z}_i$  is the standard value of the water quality parameter, and  $Z_i$  is the corresponding fitting value. The smaller the error mean square error, the greater the fitness and the greater the probability of being selected.

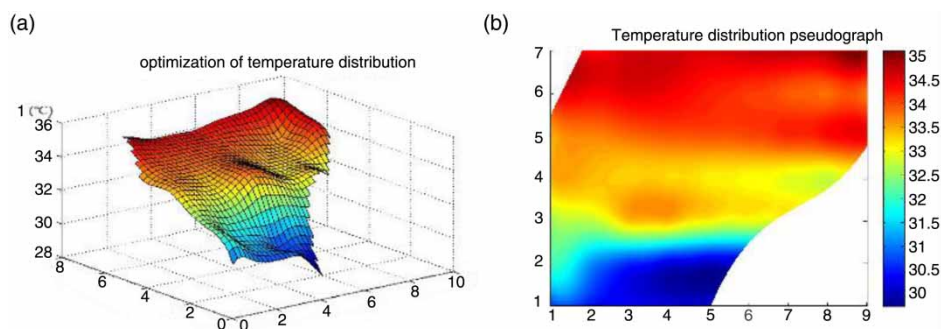
After setting a certain probability, the population is crossed, mutated, and selected to obtain the next generation group. The better group is selected by comparing it to the previous generation. The selection strategy adopted is to keep the best individuals from the parents directly involved in the selection competition of the offspring, thus avoiding the loss of good individuals from the parents and improving the overall level of the population (Manojkumar et al. 2015).

The objective function is set to the final selection criterion, and the objective function is set to the average absolute error. Its expression is:

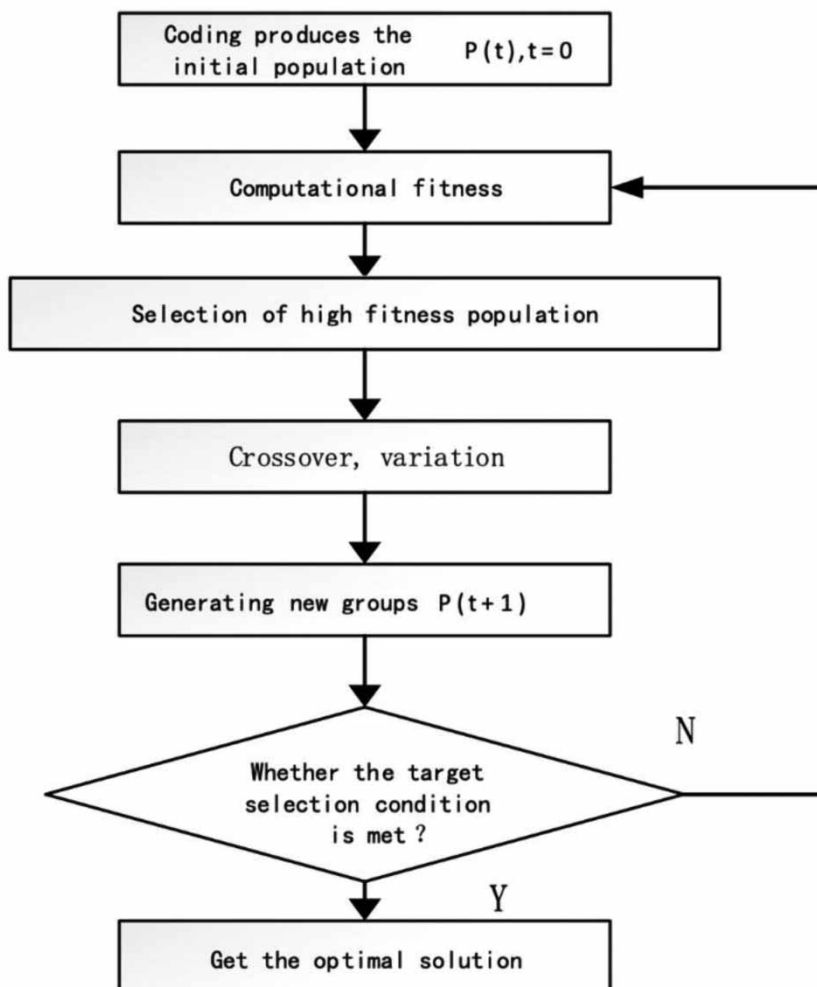
$$R = \frac{1}{N} (|\hat{Z}_1 - Z_1| + |\hat{Z}_2 - Z_2| + \dots + |\hat{Z}_n - Z_n|) \quad (3)$$

where  $R$  is the mean absolute error of the objective function.

Finally, determine whether the last selected individual meets the conditions of the target function, and if less than



**Figure 3** | Spatial distribution data of Jinghu Lake: (a) temperature 3D temperature distribution map; (b) pseudo-color image of temperature distribution.



**Figure 4** | Flow chart of genetic algorithm.

the setting error, directly output the selected chromosome; if not, return to step 2 and recalculate until the result meets the requirements of the target function. One iteration is the process by which an individual moves from calculating fitness to detecting whether the target selection condition is met.

### Multi-objective optimization

The individual fitness of the single-objective optimization algorithm is measured by the objective function, while the multi-objective optimization algorithm may have multiple conflicting optimization objectives at the same time. It is necessary to take the appropriate selection mechanism and fitness evaluation to quantify the objective function

(Ducheyne *et al.* 2004; Madeira *et al.* 2005; Yamachi *et al.* 2006).

Since various evaluation indicators have different dimensions, and the values are quite different, it is difficult to directly compare them. Therefore, it is necessary to normalize them. There are many standardized methods, and different normalized formulas may lead to different evaluation results (Liu 2010). This paper uses the min-max normalization for data processing, as shown in Equation (4).

$$\hat{z} = \frac{z - z_{\min}}{z_{\max} - z_{\min}} \quad (4)$$

where  $\hat{z}$  is the standard value of the water quality parameter,  $Z$  is the water quality parameter value,  $z_{\min}$  is the minimum value, and  $z_{\max}$  is the maximum value.



According to the seven water quality indexes that need to be considered, the seven sub-objective functions are set up as  $F_T(z)$ ,  $F_{pH}(z)$ ,  $F_{chl-a}(z)$ ,  $F_{DO}(z)$ ,  $F_{COND}(z)$ ,  $F_{TN}(z)$  and  $F_{TP}(z)$ . The main function is set to  $F(z)$ , and then the weight is set according to the importance of the data. The main function of the experiment is as follows in Equation (5).

$$F(z) = aF_T(z) + bF_{pH}(z) + cF_{chl-a}(z) + dF_{DO}(z) + eF_{COND}(z) + fF_{TN}(z) + gF_{TP}(z) \quad (5)$$

(a, b, c, d, e, f, g  $\geq 0$  and a + b + c + d + e + f + g = 1)

Then, the single objective optimization method is used for multi-objective optimization. The difference is that the selection function becomes the error of the seven sub-functions, but not the single index optimization.

## RESULTS AND DISCUSSION

### Single target monitoring point optimization

When studying the single-objective optimization problem, the layout optimization of water quality monitoring points was performed using the water body temperature indicator as an example. When the genetic algorithm is used to optimize the selection of monitoring points in the experiment, the initial population is first established. In this paper, the initial population size is set to 20 and 40, respectively, and the chromlength is 20 to constitute the initial population. The number of iterations is set to 30, 50 and 100 for the comparative optimization experiments, and the average error of 0.255 °C is selected as the threshold value of the target selection function. The number of the initial population size and number of iterations is chosen by analyzing the number of monitoring points, as well as the results obtained through the final optimization after extensive testing. The results obtained by the experiment are shown in Table 3.

From the analysis of the experimental results, it can be concluded that for the same initial population size, the higher the number of iterations corresponds to the better solution obtained, and secondly, for the same number of iterations, the larger the initial population, the better the result obtained. Of these, at an initial population of 20,

**Table 3** | Temperature iteration optimization

Initial population size	Number of iterations	Average error	Optimal solution
20	30	0.2449	27
	50	0.2403	27
	100	0.2412	26
40	30	0.2438	27
	50	0.2468	26
	100	0.2402	26

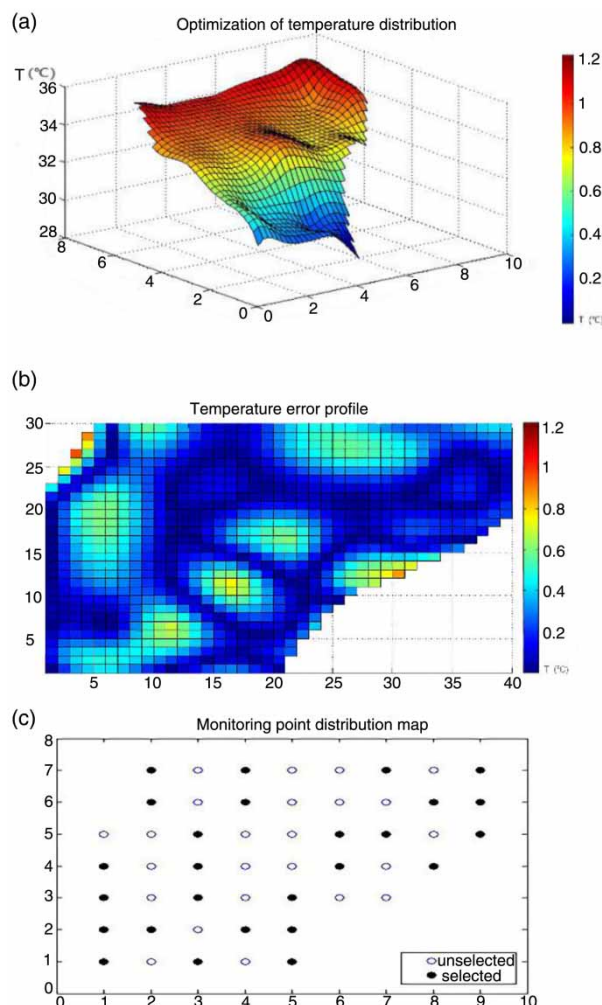
100 iterations is one less in the selection of monitoring sites than 30 and 50 iterations. And with a population size of 40, the selection of one fewer monitoring points for iterations 50 and 100 than for the optimal solution obtained by iterations 30, and the average error for iterations 100 versus 50 was somewhat reduced.

Since the initial population is larger, the more iterations, the smaller the average error and the smaller the number of monitoring points. Therefore, the initial population size is chosen to be 40, and the iteration is 100 times the experimental optimal solution. The results of optimization of single target monitoring points on temperature are shown in Figure 5.

It can be seen from the figure that 92.8% of the fitting errors of the optimized temperature distribution are below 0.2°C, while only 0.29% are above 0.6°C (Figure 5(b)). It is ideal to fit the temperature distribution of Jinghu Lake at that time. From the comparison of the figures, we can see that the fitting error is relatively large in some places where the temperature fluctuates greatly and in the edge zone, but the whole can well reflect the real situation (Figure 5(a)). It is believed that the experiment has successfully completed the selection of temperature monitoring points in Jinghu Lake. Aiming at the initial 50 monitoring points, 26 optimal monitoring points are obtained through genetic neural network optimization (Figure 5(c)). Similarly, the optimization of single water quality parameter of pH, Chl-a, DO, COND, TN and TP is similar to that of temperature optimization.

### Multi-objective optimization

In the actual water quality monitoring process, it is often necessary to monitor a variety of water quality indicators, so the optimization of monitoring points cannot be



**Figure 5** | The results of optimization of single target monitoring points on temperature: (a) temperature distribution optimization diagram; (b) temperature distribution error diagram; (c) the selection of monitoring points.

single-objective optimization, but multi-objective optimization is required. If single-objective optimization is used, the results are different water quality parameters need to select different monitoring points for water quality testing, which is not feasible in real monitoring operations. Therefore, this paper introduces complex multi-objective optimization through single-objective optimization, which can complete the optimization of monitoring points suitable for multiple water quality parameters and monitoring.

For the multi-objective optimization experiment, the initial population size setting is set to 40 and the number of iterations in sequence is set to 50 for the optimization

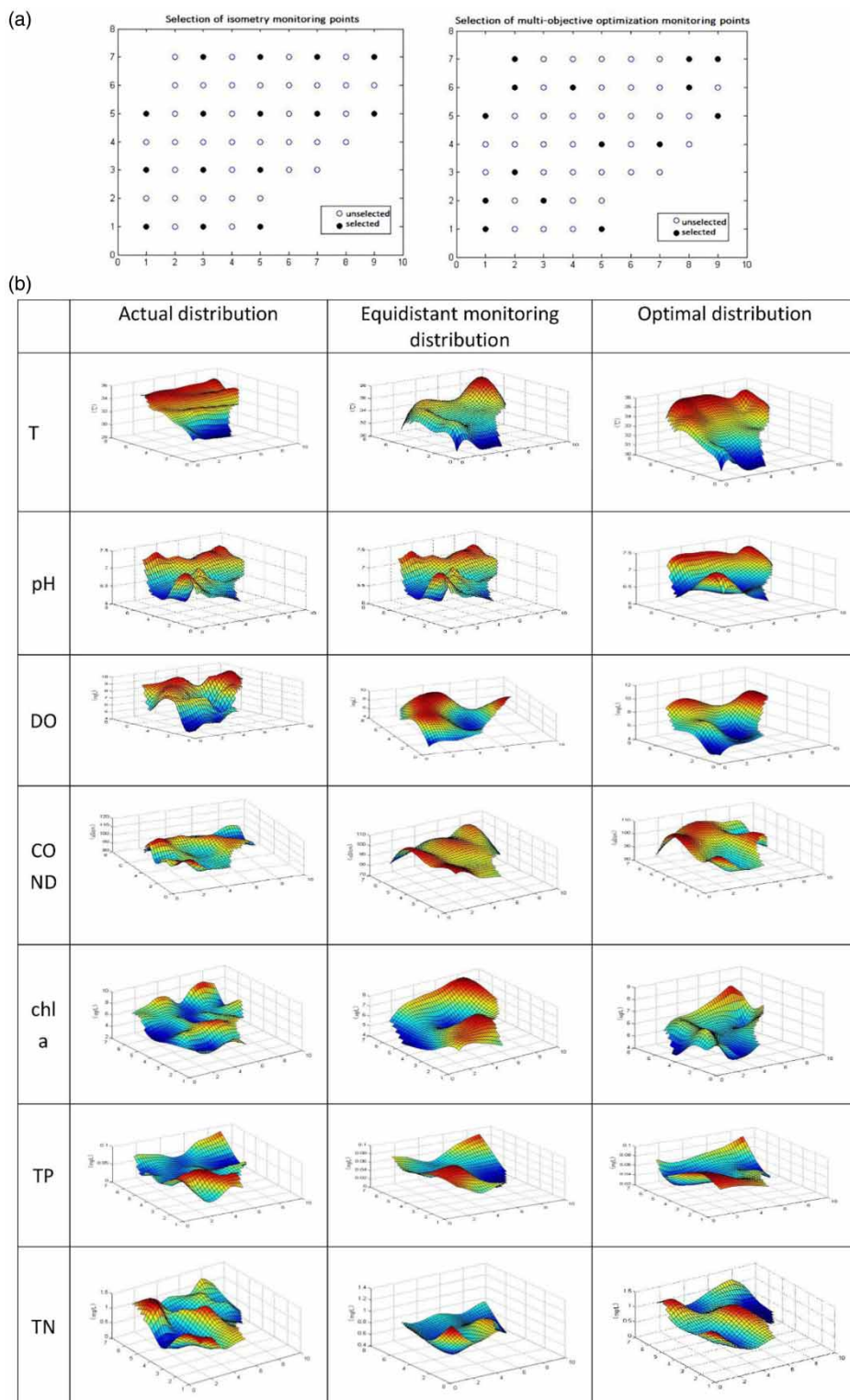
operation, which is compared with the traditional isometric sampling method (Figure 6(a)), and the experimental results are shown in Table 4.

From the experimental results, it can be seen that the fitting error of multi-objective genetic algorithm is better than that of traditional equidistant monitoring method on the whole when the same number of monitoring points are set, and the error of is reduced by 15.7%. The error of COND and TP is slightly higher than that of equidistant monitoring, which indicates that COND and TP are contradictory with other parameters in the optimization selection. The corresponding monitoring point selection after optimization is shown in Figure 6(a). The experiment shows that for the traditional equidistant sampling, multi-objective optimization can better represent the distribution of water quality change. Among them, the selection of monitoring points in the central area of lakes is less and more around, which indicates that the variation of water quality parameters in the central area is smaller and the variation of surrounding areas is larger. This may be due to the impact of lakeside trees and soil on water quality. The optimized monitoring selection avoids the shortcomings of excessive monitoring in the center and insufficient monitoring in other locations, and can reflect the water quality distribution more reasonably and accurately.

Spatial fitting of water quality parameters is shown in Figure 6(b). From the analysis of the optimization effect of single index, multi-objective optimization is not as effective as single-objective optimization. This is due to the contradiction of monitoring point selection in multi-objective optimization, and the optimal combination of monitoring points in single-objective optimization of each parameter is different, which makes the final multi-objective optimization results slightly different in the optimization of single index, but there are obvious advantages in the overall optimization of water quality parameters. Moreover, compared with the traditional equidistant monitoring, the monitoring fitting distribution optimized by multi-objective genetic neural network can reflect the actual situation more accurately in the overall water quality distribution and change trend.

### Optimization prediction and verification

Based on the optimization analysis of the first week of October, we get the monitoring points corresponding to single



**Figure 6** | The results of multi-target monitoring point optimization: (a) the selection of experimental monitoring point; (b) optimizing comparative maps.



**Table 4** | Comparison of multi-objective optimization experiments

Experimental method	Monitoring points	Sum of errors	Average error						
			T	pH	Chl-a	DO	COND	TN	TP
Multi-objective optimization	15	1.0038	0.0304	0.039	0.1559	0.1935	0.1646	0.1944	0.2260
Isometric monitoring	15	1.1261	0.0944	0.041	0.1945	0.2038	0.1574	0.2214	0.2136

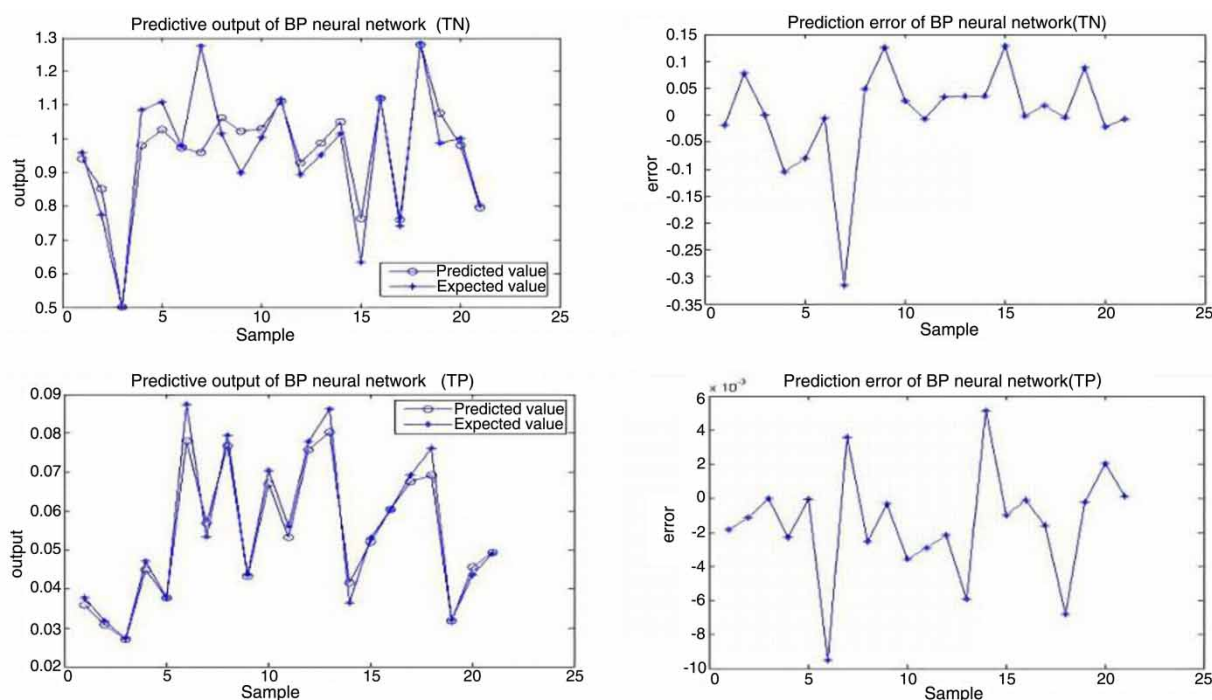
target and multi-objective optimization. In the following few weeks, we continue to monitor the water quality of Jinghu Lake to verify whether the first optimization of the monitoring site is effective for subsequent water quality monitoring.

Because TN and TP data cannot be detected directly by sensors, it is necessary to collect water samples for chemical detection, which will consume a lot of time and money. Therefore, we established a BP neural network for training multiple indicators, using temperature, pH, DO, COND and Chl-a as input variables, TN and TP as output variables. The training and prediction results of BP neural network are shown in Figure 7.

The experiment still uses the combination of sampling points obtained from the first week's multi-objective optimization for water quality monitoring. Prediction

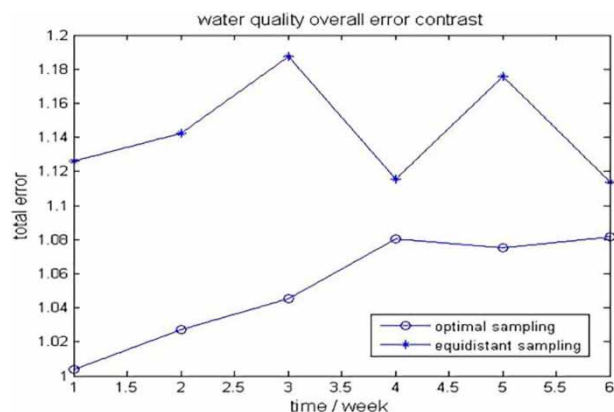
comparisons were made one week, three weeks and five weeks later, respectively. Their average error variations are shown in Table 5. The experimental results show that the error of TN and TP increases with time, ranging from 5.4% and 3.1% after one week to 15.3% and 18.1% after five weeks. On the one hand, the increase of errors is due to less training data; on the other hand, it is due to the influence of weather and human factors.

For the overall water quality of Jinghu Lake, the overall average error varies with the time of the week as shown in Figure 8. It can be seen from the figure that the overall error of multi-objective optimization monitoring shows an upward trend in the first four weeks, and then stabilizes to float at a certain value, while the traditional equidistant monitoring is random fluctuation in a relatively large error

**Figure 7** | BP neural network training and prediction.

**Table 5** | Average error of BP neural network prediction

Water quality parameters	Number of monitoring points	Average error (mg/L)		
		A week later	Three weeks later	Five weeks later
TN	15	0.1174	0.1534	0.1813
TP	15	0.0090	0.0103	0.0111

**Figure 8** | Comparison of the overall mean error of water quality.

range. Although the fitting error of multi-objective optimization is lower than that of equidistant sampling on the whole. In at least 6 weeks, the optimization effect of monitoring points is ideal.

## CONCLUSION

Experiments show that relative to the selection of traditional water quality monitoring points, genetic neural network in the accuracy of water quality parameters has been significantly improved, and the optimization effect over time is slightly reduced, but the overall error after its fitting is still less than the traditional isometric monitoring methods, greatly reducing the time and effort required to improve the efficiency and accuracy of water quality monitoring. The model used in this paper is not only applicable to the water quality parameters selected in this paper, for other different quantities and different kinds of water quality parameters monitoring is also applicable, only the application of genetic algorithms need to adjust the corresponding

parameters. In this paper, in order to reduce the significant interference of seasonal weather on the water quality parameter data when monitoring water quality data, each time the monitoring of water quality data to avoid high winds, rain and other influences on the climate, the subsequent study will add data under different climates to see if the algorithm can be adjusted according to the data, making the monitoring point optimization model more generalizable.

## ACKNOWLEDGEMENTS

This research was supported by Guangxi innovation driven development special fund project (AA17202032-2).

## DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

## REFERENCES

- Bai, J. H., Gao, H. F., Xiao, R., Wang, J. J. & Chen, H. 2012 [A review of soil nitrogen mineralization as affected by water and salt in coastal wetlands: issues and methods](#). *Clean – Soil Air Water* **40** (10), 1099–1105.
- Broomhead, D. S. & David, L. 1988 Multi-variable functional interpolation and adaptive networks. *Complex Systems* **2**, 321–355.
- Chen, L. 1995 *A Study of Optimizing the Rule Curve of Reservoir Using Object Oriented Genetic Algorithms*. PhD thesis, Department of Agricultural Engineering, National Taiwan University, Taipei.
- Chen, F. X., Cheng, J. C., Hu, Y. M., Zhou, Y. M., Zhao, Y. & Yi, J. C. 2013 Spatial prediction of soil chromium content based on RBF neural network. *Geographic Science* **33** (01), 69–74.
- Cui, H. B., Yin, Y. Q., Cui, Z. J., Wu, D. H., Jia, X. Q. & Liu, Y. 2015 Application of dynamic closeness method in the optimization of water quality monitoring points in the main stream of Tuhaihe River. *Environmental Monitoring and Warning* **7** (03), 17–21.
- Ducheyne, E. I., Wulf, R. R. D. & Baets, B. D. 2004 [Single versus multiple objective genetic algorithms for solving the even-flow forest management problem](#). *Forest Ecology and Management* **2** (201), 259–273.

- Hanbay, D., Turkoglu, I. & Demir, Y. 2007 Prediction of chemical oxygen demand (COD) based on wavelet decomposition and neural networks. *Clean – Soil Air Water* **35** (3), 250–254.
- Kaya, M. 2011 The effects of a new selection operator on the performance of a genetic algorithm. *Applied Mathematics and Computation* **217** (19), 7669–7678.
- Li, J. 2012 On the significance and content of water quality analysis. *Science and Technology Innovation Report* **21**, 233.
- Liu, X. T. 2010 Study on data normalization in BP neural network. *Mechanical Engineering and Automation* **6** (3), 122–126.
- Liu, S. M., Wu, X. & Ouyang, L. Y. 2013 Optimal location method of water quality monitoring points under uncertain nodal water volume. *Environmental Science* **34** (08), 3108–3112.
- Madeira, J. F. A., Rodrigues, H. & Pina, H. 2005 Multi-objective optimization of structures topology by genetic algorithms. *Advances in Engineering Software* **1** (36), 21–28.
- Mahbub, H., Syed, M. A. & Walid, A. 2010 Cluster analysis and quality assessment of logged water at an irrigation project. *Journal of Environmental Management, Eastern Saudi Arabia* **86** (01), 297–307.
- Manojkumar, R., Nitish, G. & Vibhu, T. 2015 Simulated binary jumping gene: a step towards enhancing the performance of real-coded genetic algorithm. *Information Sciences* **325** (20), 429–454.
- Simon, H. 1994 *Neural Networks: A Comprehensive Foundation*. McMillan Press, New York, NY, USA.
- Wang, Q. G., Li, S. B., Jia, P., Qi, C. G. & Ding, F. 2013 A review of surface water quality models. *The Scientific World Journal*. Article ID 231768, 7 pages. <https://doi.org/10.1155/2013/231768>.
- Wang, H., Liu, Z., Sun, L. N. & Luo, Q. 2015 Optimal design of river monitoring network in Taizihe River by matter element analysis. *PLoS One* **10** (5), 1–13.
- Yamachi, H., Tsujimura, Y., Kambayashi, Y. & Yamamoto, H. 2006 Multi-objective genetic algorithm for solving N-version program design problem. *Reliability Engineering & System Safety* **9** (91), 1083–1094.
- Zheng, L. C., Liu, Z. B., Zhou, Y. & Jiang, Y. P. 2007 Groundwater quality monitoring point optimization based on correspondence analysis. *Journal of Liaoning University of Engineering and Technology* **S2**, 260–262.

First received 9 March 2020; accepted in revised form 20 June 2020. Available online 3 July 2020