Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy

Pol. J. Environ. Stud. Vol. 33, No. 5 (2024), 1-9
DOI: *10.15244/pjoes/186001* ONLINE PUBLICATION DATE:

Original Research

Short-Term Early Warning Method for Algal Bloom Risk Based on a Neural Network and a Two-Dimensional Hydrodynamic Model

Na Luo[1,†], Xiaochao Gu[2,†], Kai Gao[2,*], Zeli Li[2], Pengyu Mei[2], Zhen Zhang[2], Zilu Wang[2,3], Deyue Su[4]

[1] Tianjin Tianbin Tongsheng Environmental Technology Co., Ltd, Tianjin 300191, China
[2] Tianjin Eco-Environmental Monitoring Center, Tianjin 300191, China
[3] Tianjin Tianbin Ruicheng Environmental Technology Engineering Co., Ltd, Tianjin 300191, China
[4] Yuqiao Reservoir Administration of Diversion Project in Tianjin City, Tianjin 301900, China
† These authors contributed equally

Received: 23 January 2024
Accepted: 12 March 2024

Abstract

Eutrophication of lakes and reservoirs and the occurrence of cyanobacteria blooms are two of the major environmental problems facing the whole world. However, if cyanobacteria bloom outbreaks can be predicted in advance, there is enough time to implement various measures to reduce ecological harm and health risks and greatly reduce economic losses. Yuqiao Reservoir was a typical shallow-water lake reservoir that was also faced with a greater risk of cyanobacteria bloom outbreaks in summer and autumn. In this study, a BP neural network model and a two-dimensional hydrodynamic numerical model were constructed based on meteorological, hydrological, water quality, and terrain data for Yuqiao Reservoir. The neural network model was used to simulate the biogeneration and extinction processes of cyanobacteria bloom biomass, while the two-dimensional hydrodynamic numerical model was used to simulate the migration and accumulation processes of cyanobacteria. The two models were coupled by an algal attenuation coefficient for the short-term warning of algal bloom. The results of the Nash coefficient evaluation showed that the coupled model had good overall efficiency and could effectively simulate the algal density both in time and space. The results of scenario analysis showed that the change in flow field in a shallow water reservoir had a great influence on algae migration and accumulation, and water transfer, wind speed, and direction all affected the flow field distribution and then the formation of bloom. The model application results showed that the short-term bloom prediction method constructed in this study had advantages in forecasting accuracy and displayed an effect.

Keywords: Algae bloom, Early warning, Reservoir, Model

Introduction

The eutrophication of global freshwater ecosystems has accelerated, resulting in algae blooms becoming increasingly prominent. The rapid increase or aggregation of biomass can lead to the occurrence of algae blooms. Algae blooms in water can result in the death of fish due to hypoxia, hinder the growth of underwater plants, and damage the quality of the water source [1, 2]. In 1947, the first algae bloom occurred in Apopka Lake, Florida, United States. During the 1970s, a cyanobacteria outbreak occurred in Lake Kasumigaura, the second-largest lake in

*e-mail: kaigao123@sina.cn;
Tel.:+86-022-87671979

Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy

Japan. In China, algae blooms occurred successively in Tai Hu and Chao Lake in 2007. Once an algae bloom occurs, it can result in significant damage to the aquatic environment, and the repair process is highly complex and slow [3]. It was found that the meteorological factor was the main factor in the current condition of water quality in Lake Dianchi [4]. Several results explained the strong influence of monsoon rainfall on spatial and annual water chemistry variations, and a strong positive correlation (r=0.62) between water temperature and CHL-a was observed [5]. In addition, several studies have shown that the abnormal proliferation of phytoplankton in mesotrophic drinking water reservoirs was influenced by nutrient, hydrological, and meteorological factors [6]. Algae blooms have been predicted based on meteorology and nutrient input source type. The occurrence of algal blooms could be related to a number of factors. Early warning for algal blooms in lakes and reservoirs could provide additional time to take action to mitigate the losses associated with these blooms.

Using a water environmental model to study algae bloom early warning has become a recent trend in the field. The widely used water environmental models can be approximately divided into two categories: numerical models and statistical models. The former had been used to digitize lakes or reservoirs, and calculations based on such data can be used to reveal the physical, chemical, and biological evolution processes of a water environment in advance [7, 8]. Statistical models discovered the essence of the measured object and predicted its future behavior by collecting, processing, analyzing, and interpreting data [9-11].

Hydrodynamic simulation was the basis for all water quality simulations. With the continuous development of computer technology and the efforts of many scholars, two-dimensional hydrodynamic numerical simulation has undergone further development. The hydrodynamic simulation of the numerical model used was based on the Saint-Venant equation. Since the 1980s, various studies have verified and applied this approach. The development of hydrodynamic numerical simulation was relatively mature. Compared with the hydrodynamic model, the process of water quality simulation was substantially more complex. Widely used numerical models include Water Quality Analysis Simulation Program (WASP) models, Environmental Fluid Dynamics Code (EFDC) models, and MIKE models [12-17]. The WASP model was flexible in application and can be widely used in various water quality simulations, but its hydrodynamic simulation was poor. The EFDC model had a wide range of applications, but was mainly used for simulation calculations in three-dimensional space. The MIKE21 model had good application inter-face and good simulation effects, and was used in our study. Widely used statistical models included the back-propagation (BP) neural network, the random forest algorithm, the bayesian model, and the cart decision tree algorithm [18-20]. The BP neural network model was widely used, had good self-adaptive ability and memory function, could learn and store a large number of input-output mode mapping relationships, and had clear advantages in predicting nonlinear problems [21, 22]. The BP neural network was applied in this study.

This study diminished the common difficulties of the early warning model through model coupling. When these numerical models are used alone, they must be based on the mechanisms of algae growth and reproduction. However, the mechanism of algae growth and reproduction was unclear. Our study used a statistical model to solve this problem and solved the problem

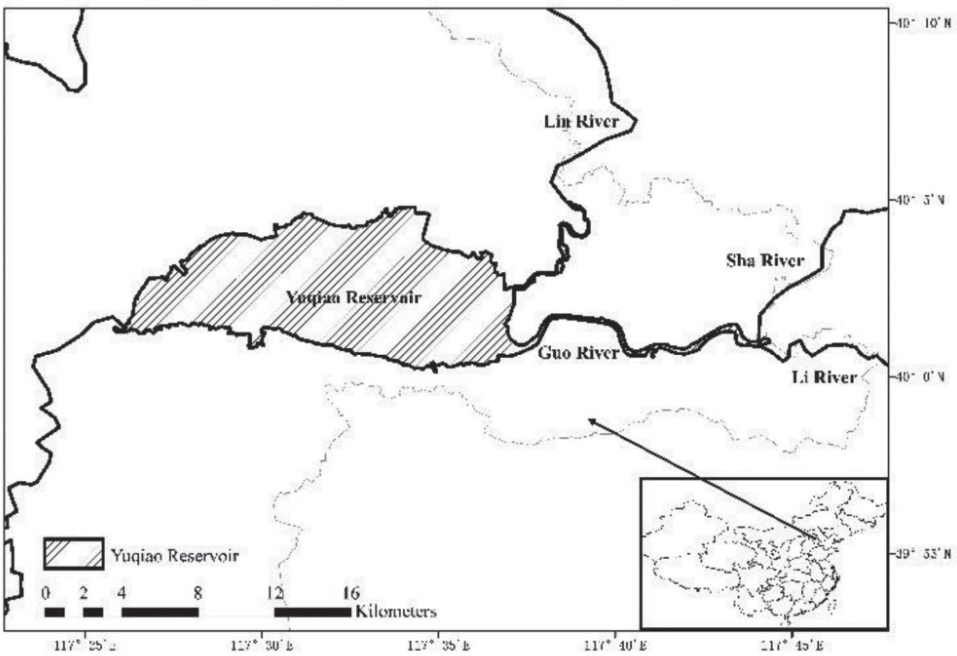

Fig. 1. Study area

Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy

of acquisition of key parameters to avoid studying the complex growth mechanism of algae and improve operability and accuracy.

The specific method was as follows: first, a large amount of historical monitoring data was used to train the BP neural network model to obtain the growth coefficient of cyanobacteria. Second, a MIKE21 flow model (FM) of Yuqiao Reservoir was established. This model was used to simulate the hydrodynamic force and water quality of the reservoir. Third, the cyanobacteria growth coefficient obtained after BP neural network training was used as the attenuation coefficient in the Mike21 water quality FM. In this way, model coupling was achieved.

Material and Methods

Research Area Profiles and Data Processing

Yuqiao Reservoir, which lies northeast of Tianjin, is a large, shallow reservoir in North China, as shown in Figure 1. The Yuqiao Reservoir Basin has an area of 2060 km^2. Its total designed capacity is 1.559 billion m^3, and the average water level elevation is 21.16 meters. The basin belongs to a temperate semihumid continental monsoon climate, with a multiyear average temperature of 12.5 °C and an average rainfall of 670 mm. Since 2014, the eutrophication level of Yuqiao Reservoir has increased. Large-scale algae blooms occurred in the reservoir area in the summer and autumn of 2015-2017, threatening the safety of the urban water supply in Tianjin [23]. There are three main rivers that flow into Yuqiao Reservoir: the Sha, Li, and Lin Rivers. The latter river dries up when there is no extreme rainfall. The water can outflow through an exit on the reservoir's western side.

Daily hydrometeorological data were used to construct a short-term early warning model for water blooms. The daily water level, water storage, inlet and outlet flow, and altitude data were obtained from the Yuqiao Reservoir Management Office. Daily meteorological data measured by meteorological monitoring stations was obtained from the China Meteorological Data Sharing Service System, which includes wind speed, humidity, temperature, precipitation, sun light hours, and atmospheric pressure. Daily water quality data, including water temperature, pH, conductivity, turbidity, permanganate index, ammonia nitrogen, total nitrogen, total phosphorus, chlorophyll A, and algal density, were obtained from the automatic water quality monitoring station of the Tianjin Ecological Environment Monitoring Center. On-site research was conducted once weekly from May to October 2021. A total of 21 locations were sampled according to grid monitoring points (Figure 2). All water samples were collected at approximately 0.5 m below the water surface, placed in polyethylene bottles, and immediately sent to the laboratory for treatment.

From May to October 2020, there was less rainfall in spring and autumn and more rainfall in summer, and the maximum daily rainfall was 139 mm. The characteristics of wind speed and direction were as follows: Southwest wind was the dominant wind direction applied, and the average wind speed was between 1.1 m/s and 2.9 m/s, as shown in Figure 3.

BP Neural Network Framework and Setup

The back-propagation algorithm was commonly referred to as the BP algorithm. A multilayer feed-forward neural network using the BP algorithm was termed a BP neural network.

The BP algorithm, also known as the negative gradient algorithm, adopted supervised delta learning. Its principle was to modify the connection weight between neurons in the network in the direction of gradient decline

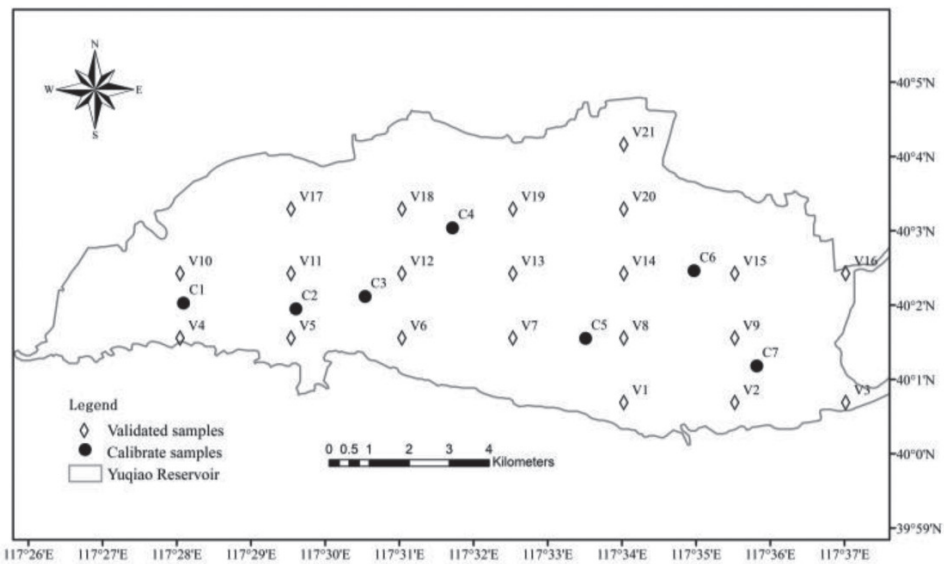

Fig. 2. Sample sites for model calibration and validation

Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy

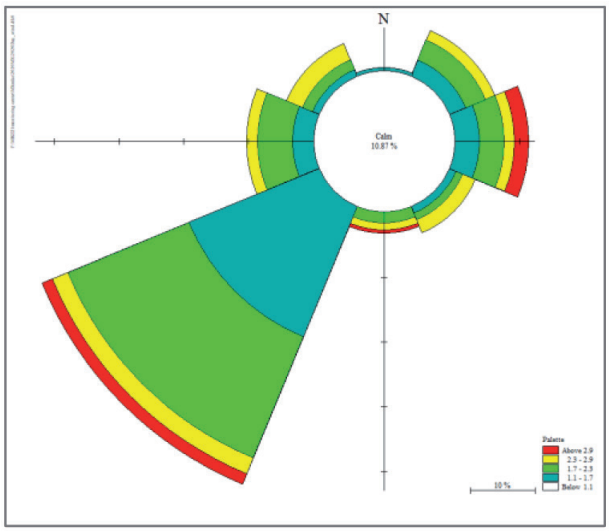

Fig. 3. Wind rose diagram

to minimize the sum of square errors of network output to achieve the expected learning results. When a BP algorithm was used, the activation function in a neural network often selected a sigmoid function that could be continuously differentiable. Compared with a single neural network, a BP neural network consists of simple nonlinear functions to realize highly nonlinear mapping of learning objectives. Therefore, BP neural networks were widely used in nonlinear modeling in the fields of pattern recognition and adaptive control.

Calculation and learning in a BP neural network can generally be divided into two steps with different directions: forward propagation of information and reverse adjustment of error. The first was the process in which information was propagated from the input layer to the hidden layer and finally output by the output layer. At this point, if the actual output was the same as the expected output, the learning and training were complete, and the results were presented. If the output is inconsistent with the expected value, an error signal will occur. By modifying the connection weight of the output layer and the hidden layer in turn, the error signal will decline along the negative gradient direction. That was done by adjusting the weight to reduce the error. Through this process, the final actual output of the network slowly approached the expected output.

In a three-layer feed-forward network [18, 22], the input vector is $X = (x_1, x_2, \dots x_n)^T$. The hidden layer output vector is $v = (v_{11}, v_{12}, \dots v_{1n})^T$. The expected output vector is $Y = (y_1, y_2, \dots y_j)^T$. The output of neurons in the hidden layer is as follows:

$$u_n = f_1\left(\sum_{m=1}^{M} \omega_{mn} x_m - \blacksquare_n\right)(n = 1,2 \dots, N) \quad (1)$$

In this formula, u_n is the output value for each node; ω_{mn} is the weight between node m and node n; $\blacksquare_n$ is the threshold value of node n; f_1 is the activation function.

The output of neurons in the output layer is as follows:

$$y_j = f_2\left(\sum_{n=1}^{M} w_{nj} v_n - \theta_j\right)(j = 1,2 \dots, J) \quad (2)$$

In this formula, y_j is the output value for each node; w_{nj} is the weight between node n and node j; θ_j is the threshold for node j; f_2 is the activation function. The sigmoid function is often used as an activation function:

$$f(x) = \frac{1}{1 + e^x} \quad (3)$$

A single-hidden layer BP neural network was established using MATLAB software. The input data of the neural network consisted of 22 hydrological, meteorological, and water quality indices, and the output result was algae density. To ensure the fast convergence of the network, the Levenberg-Marquardt algorithm was used. The activation function between neurons in the hidden layer and output layer was a hyperbolic tangent S-type function, Tansig.

Mike21 Model Framework and Setup

A 2D water quality model of Yuqiao Reservoir was established based on the MIKE21 FM developed by the Danish Hydraulic Institute and using MIKE21 software. Based on the numerical solution of the 2D shallow water equations and advection–dispersion equations, this model described the flow and water quality variations established by the hydrodynamic (HD) module and advection–dispersion (AD) module. The HD module was based on the solution of depth-integrated incompressible Reynolds averaged Navier-Stokes equations that described the evolution of the water level and two Cartesian velocity components, μ and v, for which solutions were numerically obtained from a finite difference form of the equations. The AD module simulated the physical transport processes of solutes at each grid point.

The HD module was used to simulate changes in water level and flow under the influence of various forces. The governing equations were two-dimensional diving equations. The horizontal momentum equation was integrated vertically in the Cartesian coordinate system. The continuity equation is expressed as follows [14, 24]:

$$\frac{\partial h}{\partial t} + \frac{\partial h}{\partial x} + \frac{\partial h\bar{u}}{\partial y} = 0 \quad (4)$$

The momentum equation in the X direction is as follows:

$$\frac{\partial h\bar{u}}{\partial t} + \frac{\partial h\bar{u}^2}{\partial x} + \frac{\partial h\overline{uv}}{\partial y} = f\bar{v}h - gh\frac{\partial \eta}{\partial x} - \frac{h}{\rho_0}\frac{\partial p_a}{\partial x} - \frac{gh^2}{2\rho_0}\frac{\partial \rho}{\partial x} + \frac{\tau_{sx}}{\rho_0} - \frac{\tau_{bx}}{\rho_0} - \frac{1}{\rho_0}\left(\frac{\partial s_{xx}}{\partial x} + \frac{\partial s_{xy}}{\partial y}\right) + \frac{\partial}{\partial x}(hT_{xx}) + \frac{\partial}{\partial y}(hT_{xy}) + hu_s S \quad (5)$$

The momentum equation in the Y direction is as follows:

Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy

$$\frac{\partial hv}{\partial t}+\frac{\partial huv}{\partial x}+\frac{\partial hv^2}{\partial y}=-f\bar{v}h-gh\frac{\partial \eta}{\partial y}-\frac{h}{\rho_0}\frac{\partial p_a}{\partial y}$$
$$-\frac{gh^2}{2\rho_0}\frac{\partial \rho}{\partial y}+\frac{\tau_{xy}}{\rho_0}-\frac{\tau_{by}}{\rho_0}-\frac{1}{\rho_0}\left(\frac{\partial s_{xy}}{\partial x}+\frac{\partial s_{yy}}{\partial y}\right) \quad (6)$$
$$+\frac{\partial}{\partial x}\left(hT_{xy}\right)+\frac{\partial}{\partial y}\left(hT_{yy}\right)+hv_sS$$

In this formula, x, y and z are Cartesian coordinates; t is time; h is the total water depth ($h = d + \eta$; in this formula d is still depth; η is the height of the water wave); ρ_0 is the density of fresh water, u、v、are the velocity components in the x、y、z directions, respectively. g is the acceleration of gravity; $f = 2\Omega \sin\varphi$ (Ω is the rotational angular velocity; φ is the geographic latitude); s_{xy} is the radiation stress tensor; S and (u_s, v_s) are the point source pollution emission and emission rate, respectively; $\bar{u}$, $\bar{v}$ are the average value at the vertical depth; $h\bar{u} = \int_{-d}^{\eta} u dz$ and $h\bar{v} = \int_{-d}^{\eta} v dz$ are the transverse stresses (including viscous friction, turbulent friction, and differential convective friction):

$T_{xx} = 2A\frac{\partial \bar{u}}{\partial x}$, $T_{xy} = A(\frac{\partial \bar{u}}{\partial y}+\frac{\partial \bar{v}}{\partial x})$ and $T_{yy} = 2A\frac{\partial \bar{u}}{\partial y}$, A is the eddy viscosity coefficient of horizontal flow.

The AD module describes the transport process for pollutants. The migration equation is as follows:

$$\frac{\partial}{\partial t}(hc)+\frac{\partial}{\partial x}(\bar{u}hc)+\frac{\partial}{\partial y}(\bar{v}hc)+K_dhc-S \quad (7)$$
$$=\frac{\partial}{\partial x}\left(h\lambda_x\frac{\partial c}{\partial x}\right)+\frac{\partial}{\partial y}\left(h\lambda_y\frac{\partial c}{\partial y}\right)$$

where is the concentration of pollutants, K_d is the linear attenuation coefficient, and λ_x and λ_y are the diffusion coefficients in the x and y directions, respectively.

The calculation area was discretized by an unstructured triangular mesh, and the water depth was interpolated. The entire model divided the calculation area into 7124 calculation units and 2519 calculation nodes. The topographic map construction and grid division of the reservoir area are shown in Figure 4.

The input data for hydrological models included flow, water quality, wind speed, and wind direction. The amount of water in the Lin River had been affected by the impoundment of the Longmenkou Reservoir in the upper reaches, and the flow had been cut off for a long time. This part of the flow was ignored in this simulation. The Sha River and Li River flowed into the Guo River and then entered the reservoir.

The hydrological parameters of the model mainly included the resistance coefficient (including the Manning number and the eddy viscosity coefficient) and the wind friction coefficient. The default value was adopted for the wind friction coefficient, and the trial-and-error method was adopted for the resistance coefficient. The water quality parameters of the model were mainly the algae attenuation coefficient, which was calculated according to the BP neural network prediction results, which will be introduced in the next section.

Model Coupling and Calibration Verification

The calibrated neural network model was used to predict the change trend of algae density over the following 7 days. Then, the change rate of algae density between every two days, that is, the attenuation coefficient, was calculated, which was made into a time series file and input into the MIKE21 model, so as to achieve model coupling.

Model verification was conducted at the temporal and spatial scales. On the temporal scale, the statistical index Nash coefficient was used to evaluate the model simulation effect:

$$NSE = 1-\frac{\sum_{i=1}^{n}(Q_{oi}-Q_{pi})^2}{\sum_{i=1}^{n}(Q_{oi}-\bar{Q}_o)^2} \quad (8)$$

In this formula, Q_{oi} is the measured value, $\bar{Q}_o$ is the average of the measured values, Q_{pi} is the simulated value, and n is the number of samples. The closer the r value and Nash coefficient are to 1, the higher the model's accuracy.

On the spatial scale, the spatial interpolation of measured values and the comparison of simulated values were used to evaluate the simulation effect of the model.

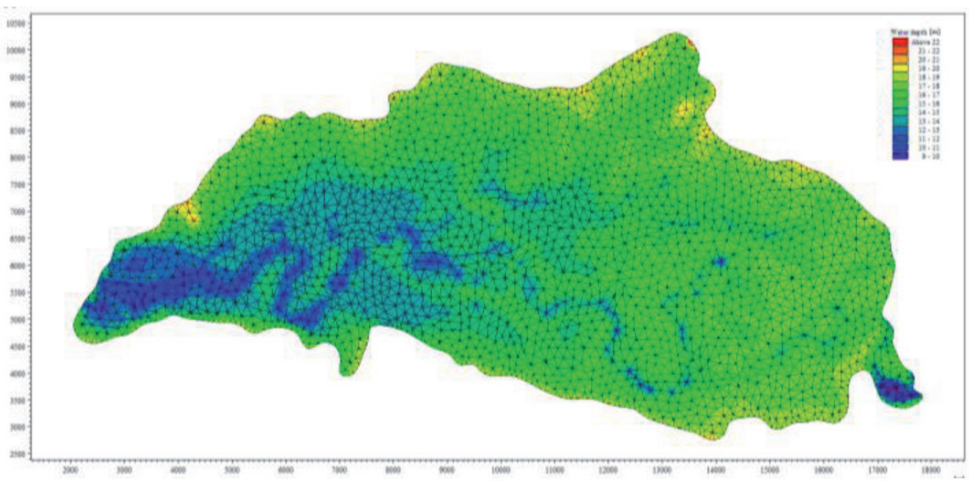

Fig. 4. Grid discrete graph

Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy

Results and Discussion

Results of Model Parameter Calibration

The Manning number was used to describe the bottom resistance of the reservoir in the hydrodynamic model. The Manning number was the main parameter for calculating the velocity of the open channel or the head loss along the channel. After debugging, the optimal value was 32 $m^{1/3}/s$. The eddy viscosity coefficient was used to describe turbulent flow. The optimal value of the eddy viscosity coefficient was 0.28 m^2/s. The wind friction coefficient changes with the wind speed. The higher the wind speed, the greater the wind friction coefficient. In this study, the model coefficients were set as a default range. When the wind speed was 0 m/s, the wind friction coefficient was 0.0016; when the wind speed was 9 m/s, the wind friction coefficient was 0.0026. The algae density change curve was created according to the prediction data in Figure 5. The linear slope between each of the two points was the attenuation coefficient. The attenuation coefficient of algae density over 7 days is shown in Table 1.

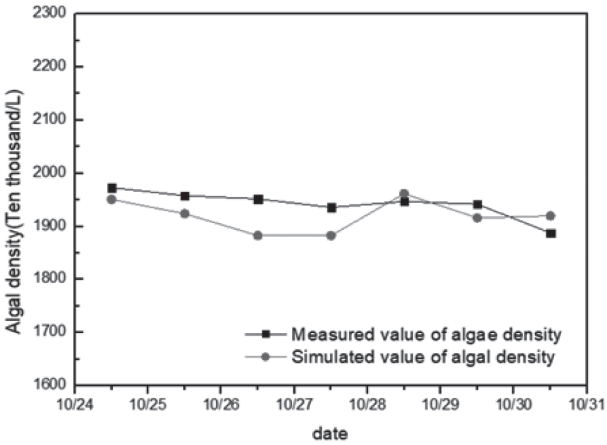

Fig. 5. Prediction of the algae growth curve using the neural network.

Table 1. Attenuation coefficient of the model

Date	Attenuation coefficient/s
10/25-10/26	3.125×10^{-7}
10/26-10/27	4.745×10^{-7}
10/27-10/28	0
10/28-10/29	-1.424×10^{-6}
10/29-10/30	5.324×10^{-7}
10/30-10/31	-4.630×10^{-8}

Results of the Model Efficiency Evaluation

As shown in Table 2, the Nash coefficient of the BP neural network model in the calibration period was 0.95. When the number of hidden layer neurons was 12, the simulation effect was better. The Nash coefficient in the validation period was 0.55. The fitting effect of the model was good during the simulation period. Its results could be applied to short-term prediction. The Nash coefficient of the hydrodynamic model was 0.94. The hydrodynamic operation results fit well with the measured values, indicating that the terrain construction was reliable. The fitting effects of points 2, 3, 4, 5, and 6 on the time scale were evaluated. The Nash coefficient ranged from 0.50-0.75.

After verification on the temporal scale, verification on the spatial scale was conducted. The 2D meshes were flexible, and this study compared the measured interpolated algae density spatial distribution map with the simulated spatial distribution results. As shown in Figure 6, the first row showed the interpolation of the measured values, and the second row showed the simulation results. The simulation range and distribution fitted well with the measured interpolation diagram from the spatial comparison, and the simulation results could be used to predict the future change trend of algae density.

Application of the Short-Term Early Warning Model

The short-term algae bloom early warning model based on the BP neural network and the MIKE21 FM had good applicability. It could be used for short-term

Table 2. Model prediction accuracy index

Model	Index	Nash coefficient (NSE)
BP neural network model	Algae density simulation period	0.95
	Algal density verification period	0.55
Hydrodynamic model	Water level	0.94
Coupling model	Point 1	0.68
	Point 2	0.70
	Point 3	0.63
	Point 4	0.75
	Point 5	0.54
	Point 6	0.50

Note: NSE was closer to 1.0, and the result was more accurate.

Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy

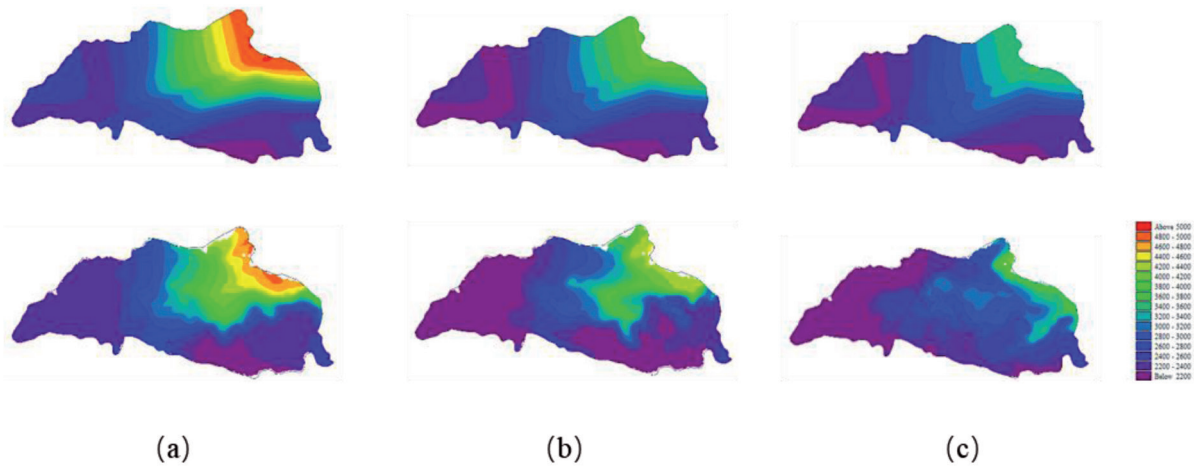

Fig. 6. Comparison of the spatial distribution between the interpolation diagram and simulation diagram of the coupled model automatic station data: (a) Date: 10/27; (b) Date: 10/29; (c) Date: 10/31.

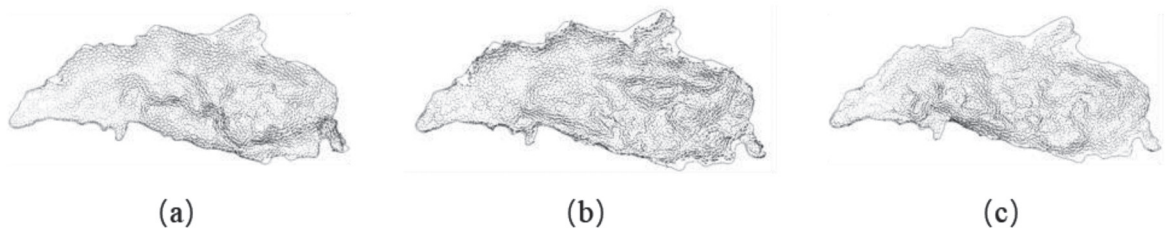

Fig. 7. Flow field in multiple cases: (a) Actual flow field; (b) No inflow and outflow flow field; (c) Static airflow field.

early warning of cyanobacterial algae blooms in Yuqiao Reservoir and could realize real-time spatial dynamic change prediction.

The model simulated the changes in the reservoir flow field under different hydrological and meteorological conditions and analyzed the factors of reservoir flow field change. The research outcomes indicated that inflow disturbance, wind speed, and wind direction were all factors affecting the reservoir flow field. Topography also affected the distribution of the flow field. Our study found that when there was water transfer, the reservoir had a regular flow field and active water flow. Because of the prominent channel topography at the bottom of Yuqiao Reservoir, the main track of water flow was clear when there was water transfer. The flow was not conducive to algae aggregation in the reservoir. Therefore, it was not easy for an algae bloom to occur (Figure 7(a)). When there was no water transfer, wind speed and direction were the main factors affecting the flow field. This study used the flow field of a single day as an example (Figure 7(b)). Based on the dynamic change, the change in the flow field was substantially affected by wind speed and direction. Because the wind speed and direction changed over time, the flow field was irregular and diverse. When the wind speed was strong, a stable circulation flow field did not form easily, and algae did not easily gather in this flow field environment. The risk of algal blooms occurring was low. When there was no water diversion, no wind, or stable low wind, small-scale circulation easily formed in the flow field. In this flow field environment, algae aggregation and algal blooms easily occurred (Figure 7(c)).

Discussion on Mechanisms and Model Applications

The formation of algal blooms can mainly be divided into two processes: algal growth and algal migration. In this study, the algae growth process was simulated by constructing a neural network model with time differences, whereby 22 meteorological, hydrological, and water quality indices were used as independent variables and algae density as the dependent variable. The algae migration process was simulated using a MIKE21 hydrodynamic model. The model fully considered the effects of wind, rainfall, water transfer, and topography on the migration and aggregation of algae within the flow field. Regarding the prediction effect, the constructed algae bloom short-term early warning model accurately reflected the entire process. It solved the problem associated with the statistical model only being used for single point prediction and smoothly addressed the problems posed by complex modeling principles and many parameters. The short-term early warning model of algal blooms superimposed temporal and spatial changes, which had advantages in terms of prediction accuracy and display effect.

This study referred to the previous research results of many scholars at home and abroad in the field of algae

Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy

bloom early warning. Several studies have analyzed the relationship between meteorological factors in Taihu and the algae bloom comprehensive index through satellite remote sensing interpretation and random forest machine learning algorithms [25]. The ElCOM-CAEDYM model was used to study the impact of meteorological factors on algal blooms in Chaohu Lake. The results show that continuous low winds are more likely to lead to algal blooms [26].

Considering meteorological factors and using their predictability in algae blooms, early warning could improve prediction accuracy to a certain extent. Based on a hydrodynamic model, this study investigated the migration pattern of algae in the flow field under the influence of different meteorological factors. Through the study of the flow field, it was found that when the wind speed is low and the wind direction is stable for a period, the risk of algal blooms in the shallow water area on the bank and the circulation area in the reservoir is high. Through the study of the flow field, it was found that when the wind speed was low and the wind direction was stable for a period, the risk of algal blooms in the shallow water area on the bank and the circulation area in the reservoir was high.

Many studies adopted models of some type, each with its own focus. However, there were limitations to such methods. The use of statistical models focuses on changes in time scale. In our case, the use of numerical models required in-depth knowledge of the growth mechanisms of algae. However, the complex mechanism of algae growth and reproduction was unclear. Fanxiang, and KONG applied the theory and technical system of algae bloom prevention, prediction, and early warning to predict the occurrence probability, location, and intensity of algae blooms in Tai Hu. The prediction accuracy of the model was 60%-80% [27]. To improve model accuracy, in lake and reservoir algae bloom early warning areas, combining models has become a research trend. However, the research time range in their study was short, and the results may have certain contingencies.

Conclusions

In this study, a BP neural network model was constructed based on meteorological, hydrological, water quality, and topography data in Yuqiao Reservoir to simulate the process of cyanobacteria bloom biomass. The hydrodynamic numerical model of MIKE21FM lake reservoir was constructed to simulate the migration and accumulation processes of cyanobacteria, and the two models were coupled for short-term algal bloom warning. From the perspective of prediction effect, the model ensured the accuracy of the whole process, solved the problem that only a single point can be predicted by using a statistical model alone, and solved the problem of lack of acquisition of key parameters such as attenuation coefficient in the water quality model. By superposing time and space changes, it provided effective technical support for the water environment management of Yuqiao Reservoir more intuitively.

Acknowledgements

We would like to thank Dr. Qi, Tianjin Research Institute for Water Transport Engineering, Ministry of Transport, for his technical support and guidance in the operation of the two-dimensional hydrodynamic water quality model (MIKE21 FM).

References

1 NETO, J.C. Environmental factors associated with toxic cyanobacterial blooms across 20 drinking water reservoirs in a semi-arid region of Brazil. Harmful Algae, **86**, 128, **2019**.
2 DONG H.Y., TAREQ A.M., RICHARDSON S.D. Transformation of algal Toxins during the oxidation/disinfection processes of Drinking Water: From Structure to Toxicity. Environmental Science and Technology, **57** (35), 12944, **2023**.
3 YE L., TAN L., WU H.X., CAI Q.H. Nonlinear causal analysis reveals an effective water level regulation approach for phytoplankton blooms controlling in reservoirs. Science of The Total Environment, **806**, 150948, **2022**.
4 LU W.K., YU L.X., OU X.K., LI F.R. Relationship between occurrence frequency of cyanobacteria bloom and meteorological factors in lake Dianchi. Journal of Lake Sciences, **29** (3), 534, **2017**.
5 KIM J.Y., ATIQUE U., MAMUN M., AN K.G. Long-term interannual and seasonal links between the nutrient regime, sestonic chlorophyll and dominant bluegreen algae under the varying intensity of monsoon precipitation in a drinking water reservoir. International Journal of Environmental Research and Public Health, **18** (6), 2871, **2021**.
6 LV X.Y., ZHU M.Y., MA Y.S., ZOU W., XU T.Y., ZHANG Z.B., ZHU G.W., Driving factors of algal blooms in drinking-water reservoirs in Lake Taihu Basin. Journal of Lake Sciences, **35** (5), 1516, **2023**.
7 SHA J., ZHAO Y., LI X., WANG Z.L. Assessing impacts of future climate change on hydrological processes in an urbanizing watershed with a multimodel approach. Journal of Water and Climate Change, **12** (4), 1023, **2020**.
8 MYERS D.T.L., RICHARD R.R., MCNAIR J.N., MATTHEW E. Watershed and streambank erosion modeling in a coldwater stream using the GWLF-E model: application and evaluation. Modeling Earth Systems and Environment, **7** (3), 1, **2021**.
9 CUI Y.J., YANG Y.D., ZHENG W.T., CHENG Z.Q., CHEN T.S., LIN X.F. Research on the succession characteristics of phytoplankton community structure and risk warnings in the mianhuatan Reservoir. China Rural Water and Hydropower, **3**, 127, **2022**.
10 CHOU J.S., HO C.C., HOANG H.S. Determining quality of water in reservoir using machine learning. Ecological informatics: an international journal on ecoinformatics and computational ecology, **44**, 57, **2018**.
11 LIANG Z. Simulate the forecast capacity of a complicated water quality model using the long short-term memory approach. Journal of Hydrology, **581**, 257, **2020**.
12 MBUH M.J., RICHARD M., COMFORT W. Water quality modeling and sensitivity analysis using Water Quality Analysis Simulation Program (WASP) in the Shenandoah River watershed. Physical Geography, **40** (2), 1, **2018**.

Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy • Author Copy

13 BOUCHARD D., KNIGHTES C., CHANG X.J., AVANT B. Simulating Multiwalled Carbon Nanotube Transport in Surface Water Systems Using the Water Quality Analysis Simulation Program (WASP). Environmental Science & Technology, **51** (19), 11174, **2017**.

14 XI B. , CHEN Q.Q., CHEN W., XI W., ZOU Y., HUANG H.T., CHEN Z.X., SHEN S.X., CHEN Y.X., KONG Q.H. River Water Quality Based on Mike21 Model Lifting Engineering Test Research. China Rural Water and Hydropower, **5**, 5, **2023**.

15 LI Q.F., GUO S., ZHANG R.Y., YAN X.M., ZHANG X.L. Application of MIKE11 Model in River Network Area of Pearl River Delta for Water Quality Improvement. Pearl River, **44** (2), 54, **2023**.

16 XU S., HE G.J., FANG H.W., BAI S., WU X.H. Parameter uncertainty and sensitivity analysis of the three Gorges Reservoir and Xiangxi River EFDC model. Journal of Hydrology, **610**, 127881, **2022**.

17 KIM J., LEE T., SEO D. Algal bloom prediction of the lower Han River, Korea using the EFDC hydrodynamic and water quality model. Ecological Modelling, **366**, 1730, **2017**.

18 QIN M., ZHANG M.F., HONG Y., SU Y.P., CHEN Y.F., LI H.L., CHEN Y.X. Study on the calculus model of cyanobacterial bloom based on BP artificial neural network in Dongzhang Reservoir. Journal of Fisheries Research, **41** (1), 18, **2019**.

19 SCHONLAU M., ZOU R.Y. The random forest algorithm for statistical learning: The Stata Journal, **20** (1), 3, **2020**.

20 LIANG Z., QIAN S.S., WU S., CHEN H., LIU Y., YU Y., YI X. Using Bayesian change point model to enhance understanding of the shifting nutrients-phytoplankton relationship. Ecological Modelling, **393**, 120, **2019**.

21 KUANG F., LONG Z., KUANG D., LIU X., GUO R. Application of back propagation neural network to the modeling of slump and compressive strength of composite geopolymers. Computational Materials Science, **206** (15), 111241, **2022**.

22 LAN W.Q., YANG X., GONG T.S., XIE J. Predicting the shelf life of Trachinotus Ovatus during frozen storage using a back propagation (BP) neural network model. Aquaculture and Fisheries, **8**(5), 544, **2023**.

23 LI Z.L., LUO N., SU D.Y., GAO K., ZHAO X.H., MEI P.Y., ZHANG Z. Load estimation and source variation analysis of nitrogen and phosphorus for the Yuqiao reservoir watershed based on the GWLF model. Journal of Agricultural Resources and Environment, **38** (1), 63, **2021**.

24 LI X., HUANG M., WANG R. Numerical Simulation of Donghu Lake Hydrodynamics and Water Quality Based on Remote Sensing and MIKE 21. International Journal of Geo-Information, **9** (2), 94, **2020**.

25 LUO X., HANG X., CAO Y., HANG R., LI Y.C. Dominant meteorological factors affecting cyanobacterial blooms under eutrophication in Lake Taihu. Journal of Lake Sciences, **35** (1), 1248, **2019**.

26 JIANG C.Y., TANG X.X., WANG C., YUAN Y., LI X.L., QIAN X. Influence of meteorological factors on outbreaks of cyanobacterial blooms in water resource region of Chaohu Lake. Jiangsu Agricultural Sciences, **47** (10), 281, **2019**.

27 KONG F.X., MA R.H., GAO J.F., WU X.D. The theory and practice of prevention, forecast and warning on cyanobacteria bloom in Lake Taihu. Journal of Lake Sciences, **21** (3), 314, **2009**.